Stratified Proportional Hazards Models. A GLIM oriented approach, with special reference to the problem of competing risks.

Camille Vanderhoeft

Vrije Universiteit Brussels

IPD-Working Paper 1985-5

Acknowledgements :

I would like to thank the Population and Family Study Centre (C.B.G.S.) for use of the NEGO IV data.
My greatest gratitude is expressed to Stanley Wijewickrema - who read the first drafts of this paper with extreme care and whose suggestions improved the text quite a lot - and to Paul Willems - with whom I had long and interesting discussions concerning the results of the present exercise.
Last but not least, I thank Etienne Vanden Balck for his careful typing of the text.

ORDERING INFORMATION

Extra copies of this working paper are available from :

Mr. E. Vanden Balck
Secretary/Librarian
Interuniversity Programme in Demography
c/o Centrum voor Sociologie
Vrije Universiteit Brussel
Pleinlaan 2
B-1050 Brussels
BELGIUM

at the cost of US $ 10.00 or BF 500 . Please attach check or international money order.

ORDER FORM

Name ....

Address ....

....

....

....

....

wishes to receive ...... copies of Working Paper Nr..........
and includes payment of US $ .... or BF........

Contents

0. INTRODUCTION

Three papers, containing ideas of great importance to statisticians and other investigators in a number of different scientific disciplines (biometrics, demography,...) were presented in the early seventies. In the first place, Cox (1972) introduced a proportional hazards model in which the base-line hazard (i.e. the nuisance function) was not given any specific form. Nelder and Wedderburn (1972) defined a class of generalized linear models and constructed a single algorithm for fitting any member of this class. Last but not least, Coale and McNeil (1972) presented a parametric model for the distribution of age at first marriage. Each of these three statistical issues has since been further developped in its own right, but over the years it has also been seen that they could be linked at different points.

Cox (1975) was responsible for the partial likelihood method related to the proportional hazards model found in his earlier paper. Several authors addressed their minds to the problem of making inferences through the partial likelihood method (see Kalbfleisch and Prentice (1980) or Lawless (1982) for an overview). Other authors suggested alternative and simpler methods designed to avoid the difficulties associated with Cox's 1972 ideas. Breslow (1972, 1974) for instance used the notion of constant hazards between two successive (non-censored) failure times; and Holford (1976, 1980) divided the period of follow-up into fixed intervals. This last approach is adopted in the present study.

The method developped by Cox (1972, 1975), and the alternative methods proposed by Breslow (1972, 1974) and Holford (1976, 1980), have been used by many investigators : e.g. Menken et al (1981), Trussell and Hammerslough (1983) and more recently Rodriguez (1984) in the world of demography.

The availability of computers and the development of special computer packages explains a good part of the attractiveness of the proportional hazards method. Laird and Olivier (1981) and Holford (1980) showed that iterative proportional fitting algorithms (used in connection with log-linear models for contingency tables) can be used here too. Baker and Nelder (1978) developped GLIM, a computer package for fitting the generalized linear models developped by Nelder and Wedderburn (1982). Since log-linear models are easily fitted through GLIM (Release 3) - as they are special generalized linear models - we preferred to use this computer package in the present study. Note also that the partial likelihood corresponding to the models in question can be maximized through GLIM, as it was pointed out by Whitehead (1980).

The Coale-McNeil (1972) model has also been in constant use among demographers. Maximum likelihood estimation of the parameters of the model is discussed by Rodriguez and Trussell (1980). Although the model was originally intended for the analysis of the distribution of age at first marriage, it has also been used in analysis of the age at first birth (Bloom, 1980). The fitting of the Coale-McNeil model has been simplified by the computer programme NUPTIAL (Rodriguez and Trussell, 1980). The application of the Coale-McNeil model is not considered in this paper, but similarities with the models presented in the present study are briefly noted (Section 4).

Theorètical issues related to proportional hazards models in the presence of competing risks have already been discussed by a number of authors (e.g. Cox (1972), Chiang (1968), Gail (1975), Holt (1978), Laird and Olivier (1981)). Extended applications are however, at least to our knowledge, rather rare - especially in demographic studies. The same remark can be made concerning the problem of stratifying the proportional hazards model (see Sections 2.3 and 3.3); though some connected theoretical aspects are considered by Holt (1978).

The present study tries to make some of these advanced statistical techniques more readily available to the user. With this end in view they are presented in one big framework. Reference to the mathematical foundations of the methods has not been avoided since we are convinced that a strong grasp of the formal mathematical background in question will help the investigator both to understand what he is doing and to see what may further be possible.

1. THE DATA

In a survey on fertility and family formation in Flanders (Belgium) organised by the Centrum voor Bevolkings- en Gezinsstudiën (Center for Population and Family Studies), 3101 women were interviewed during the period november 1982 - june 1983. These women were born between 1938 and 1961. Because of certain difficulties not all 3101 women were retained for analysis in the present paper. Two questionnaires had in fact been used in the survey : one for ever married, and one for never married women. Due to an error in the questionnaire for the never married women, 230 out of a total of 469 never married women had to be excluded from analysis. Various consistency checks led to the exclusion of a number of additional cases, so that the final sample contained only 2829 women.

For each of these women we calculated the age of entry into first union, i.e. either first cohabitation (meaning first cohabitation of never married women throughout this paper) or first marriage. The women who, at the time of interview, had never been married or had never cohabited, are said to be censored. The age at which these women are censored (i.e. their age at interview) forms part of the data. Formally, the data can be summarized in a set of vectors

$$(\underline{a}_i, \delta_i, z_i)_{i=1,\ldots N} \qquad (1.1)$$

where

N is the number of women in the sample,

$z_i=(z_{i1},\ldots,z_{im})$ is a vector coding m characteristics for woman i, e.g. her date of birth, the highest education level obtained, religion etc.,

δ_i is a status variable indicating that women i is censored (δ_i=0) or that her first union was a marriage (δ_i=1) or a cohabitation (δ_i=2),

$\underline{a}_i$ is the age in completed years at which woman i entered any of the states mentionned above (i.e. censoring, first marriage or first cohabitation).

Since first union occurs mainly after exact age 15, women entering into first union before age 15 were excluded. The 15th birthday was next reset as having a value equal to zero by the translation $\underline{t}_i=\underline{a}_i-15$. Note that $\underline{t}_i$ thus stands for the duration in completed years between the 15th birthday of a woman i and her entry into first union or the moment when she was censored. This duration will be referred to as *time* in the remainder of this text. The data set can now be symbolized as follows :

$$(\underline{t}_i,\delta_i,z_i)_{i=1,\ldots N} \tag{1.2}$$

This will hereafter be called the set of *observed individual data*. Note that $\underline{t}_i=0,1,2\ldots.$

If exact ages a_i of entry into first union or of censoring would have been recorded, then we could consider the data set

$$(t_i,\delta_i,z_i)_{i=1,\ldots N} \tag{1.3}$$

where $t_i=a_i-15$ (with $t_i \geqslant 0$). We will refer to (1.3) as the set of *(unknown) exact individual data*. Data set (1.3) is important for the construction of the models discussed later (see Sections 2 and 3) - the *likelihood function* (or simply the *likelihood*) for data (1.2) will be derived from the likelihood for data (1.3). Data set (1.3) makes reference to exact times t_i (as opposed to time $\underline{t}_i$ measured in completed years). Models in continuous time will first be constructed and they will next be transformed into discrete time models (through the use of certain assumptions, e.g. that of piecewise constant hazards - see Section 2.2). The data referred to in discrete time models take the form of the data set (1.2), whereas the data referred to in continuous time models take the form of the data set (1.3).

The characteristics - giving rise to three categorical variables (or covariates) in the present study - are :

- identification of the birth cohort, i.e. covariate COH taking the values
  1 for women born between 1948 and 1962 (48-62)
  2 for women born between 1938 and 1947 (38-47)

- highest educational level attained, i.e. covariate EDU taking the values
  - 1 for primary education (PRI)
  - 2 for secondary education (SEC)
  - 3 for higher education (HIGH)
- religious affiliation, i.e. covariate REL taking the values
  - 1 for Roman Catholics with regular Mass attendance (RC RMA)
  - 2 for Roman Catholics with irregular or no Mass attendance (RC IRMA)
  - 3 for women who claim no specific religious affiliation (NRA)
  - 4 for freethinkers (FREE)

In the rest of this paper we will in general speak about *covariates* z_i keeping in mind that z_i refers to a vector of covariate values $z_{i1},\dots z_{im}$ (here m=3) for woman i. All women with the same characteristics (or covariates) z will be referred to as belonging to *subgroup* z, and the set of distinct subgroups z will be symbolized by $\mathfrak{Z}$ (thus $z \varepsilon \mathfrak{Z}$). Note that time $\underline{t}_i$ refers in fact to the time interval $[\underline{t}_i, \underline{t}_i+1)$. Since $\underline{t}_i$ is measured in completed years, we can identify the interval $[\underline{t}_i, \underline{t}_i+1)$ for woman i through the index l_i, where $l_i=1$ if $\underline{t}_i=0$, $l_i=2$ if $\underline{t}_i=1$, etc. The index l_i varies between 1 and an upper value L which will be specified later.

For each subgroup z and for each time interval l we can then calculate

n_{lz} = the number of women in subgroup z who neither entered first union nor got censored before the starting point of the l-th interval (i.e. the number exposed to risk at the beginning of the l-th interval),

d_{1lz} = the number of women in subgroup z who entered the state of first marriage in the l-th interval,

d_{2lz} = the number of women in subgroup z who entered the state of first cohabitation in the l-th interval,

w_{lz} = the number of women in subgroup z who were censored in the l-th interval.

The observed individual data set (1.2) can thus be transformed into the *grouped data set* (1.4)

$$(l, n_{lz}, d_{1lz}, d_{2lz}, w_{lz}, z)_{z\varepsilon\mathfrak{Z};\, l=1,\dots L} \qquad (1.4)$$

Table 1.1 reproduces a few lines from our data set, organized as in (1.4).

Table 1.1 : Organization of grouped data set

					covariates z		
l	n_{lz}	d_{1lz}	d_{2lz}	w_{lz}	z_1=REL	z_2=EDU	z_3=COH
1	29	0	0	0	1	1	1
2	29	0	0	0	1	1	1
3	29	1	0	0	1	1	1
4	28	2	0	0	1	1	1
5	26	11	0	0	1	1	1
6	15	5	0	1	1	1	1
7	9	2	0	0	1	1	1
8	7	2	0	0	1	1	1
9	5	4	0	0	1	1	1
10	1	1	0	0	1	1	1
1	82	0	0	0	1	1	2
2	82	0	0	0	1	1	2
3	82	3	0	0	1	1	2
4	79	5	0	0	1	1	2
5	74	5	0	0	1	1	2
6	69	11	0	0	1	1	2
7	58	10	0	0	1	1	2
8	48	21	0	0	1	1	2
9	27	14	0	0	1	1	2

Thus, 29 women with covariates $z_1=1$, $z_2=1$, $z_3=1$ (i.e. Roman Catholic women with regular Mass attendance, whose highest educational level is primary education, and born between 1948 and 1962) are exposed to risk at the starting point of the first interval ($l=1$), i.e. at exact age 15.
No woman in this subgroup either entered the state of first union or was censored during the first interval ($l=1$), i.e. between exact ages 15 and 16, so that the 29 women were still exposed to risk at the beginning of the second time interval ($l=2$) (i.e. at exact age 16). Note that $n_{l+1z}= n_{lz}-d_{1lz}-d_{2lz}-w_{lz}$ in general.

Note further that some women enter the state of first union or are censored at ages beyond exact age 36. These women, however, constitute only a small fraction of the total sample. We have not excluded these women from analysis, but they are considered as being censored at exact age 36, i.e. at the end of the 21th interval. Consequently, the upper value L of the index l defined above is 21. The grouped data (1.4) is then said to be *time-censored.*(1)

If we were to omit the distinction between first marriage and first cohabitation and consider only first unions as such, we could define

$d_{lz}=d_{1lz}+d_{2lz}$ = the number of women in subgroup z experiencing first union in the lth time interval.

This would lead to the grouped data form

$$(l,n_{lz},d_{lz},w_{lz},z)_{z\epsilon\mathfrak{Z};l=1,\ldots L} \qquad (1.5)$$

The corresponding individual data are still of the form (1.2) or (1.3), but the status variable δ_i then takes only the values zero for censored women and 1 for women who enter either the state of first marriage or the state of first cohabitation.

For convenience, we will use the following terminology throughout the paper :

- if woman i enters the state of first marriage (δ_i=1), it will be said that she *enters the state of first union due to cause 1*,
- if woman i enters the state of first cohabitation (δ_i=2), it will be said that she *enters the state of first union due to cause 2.*

In general, a woman will be said to *enter the state of first union due to cause* j (j=1,2). The use of this terminology has various advantages : (1°)the discussion in Section 3 is substantially simplified; (2°)the terminology used is made similar to that used in cause-specific mortality studies; and (3°)extension from 2 to J causes (J > 2) is straightforward.

2. THE FIRST UNION MODEL

2.1. Mathematical formulation and likelihood construction

Let T be a continuous random variable, representing the *time* (i.e. *age since 15th birthday*) at which a woman enters the state of first union. Note that T cannot be observed for censored women.

The hazard function (or the instantaneous rate) of entering the state of first union at time t is, for women with characteristics z, defined as

$$\mu(t;z) = \lim_{\Delta t \downarrow 0} \frac{P(t \leqslant T < t + \Delta t \mid T \geqslant t, z)}{\Delta t}, \qquad t \geqslant 0 \tag{2.1}$$

This means that for an arbitrarily small interval $[t,t+\Delta t)$, the quantity $\mu(t;z).\Delta t$ can be interpreted as the probability that a woman with covariates z enters the state of first union in the time interval $[t,t+\Delta t)$, given that she has not done so before time t. Note that this is a conditional probability.

The probability, $S(t;z)$, that a woman with covariates z does not enter the state of first union in the time interval $[0,t)$ - i.e. the *survivor* function for women with covariates z - is defined as

$$S(t;z)=P(T \geqslant t \mid z).$$

It may be shown than that the hazard function and the survivor function for women with covariates z are related through the equation (see Appendix E1)

$$S(t;z)=\exp\left(-\int_0^t \mu(s;z)ds\right) \tag{2.2a}$$

or equivalently

$$\mu(t;z)=-\frac{d}{dt}\log S(t;z). \tag{2.2b}$$

Yet another function useful in the present discussion is the cumulative hazard function, $\Lambda(t;z)$ say. Defining it as

$$\Lambda(t;z) = \int_0^t \mu(s;z)ds,$$

we have

$$S(t;z) = \exp(-\Lambda(t;z)). \tag{2.3}$$

The (unconditional) probability that a woman with covariates z enters the state of first union in $[t,t+\Delta t)$ is approximated by the product $S(t;z).\mu(t;z)\Delta t$, which will be denoted by $f(t;z)\Delta t$, where $f(t;z)$ is the corresponding probability density function (p.d.f.). Thus

$$f(t;z) = S(t;z)\mu(t;z). \tag{2.4}$$

Note that $f(t;z) = -\frac{d}{dt} S(t;z)$. (Appendix E2).

The probability that a woman with covariates z experiences first union in $[0,t)$ is given by anyone of the following relations :

$$F(t;z) = 1 - S(t;z) \tag{2.5a}$$

$$= \int_0^t f(s;z)ds \tag{2.5b}$$

$$= \int_0^t S(s;z)\mu(s;z)ds. \tag{2.5c}$$

$F(t;z)$ is called the cumulative distribution function (c.d.f.). Note that $f(t;z) = \frac{d}{dt} F(t;z)$.

The ultimate proportion of women with covariates z who will ever experience first union, or the probability that a woman with covariates z ever enters the state of first union, is defined by

$$c(z) = F(\infty;z) = \lim_{t\to+\infty} F(t;z) \tag{2.6}$$

Clearly, $c(z)$ can be less than unity for some (or all) subgroups z. If $c(z)=1$, first union is said to be *universal* in subgroup z.

Finally, the conditional probability of entering first union in the interval $[t,t+h)$, given that first union has not occured before t and given covariates z, is

$$q(t,h;z) = \int_t^{t+h} \frac{S(s;z)\mu(s;z)}{S(t;z)}\, ds \qquad (2.7a)$$

$$= \int_t^{t+h} \frac{f(s;z)}{S(t;z)}\, ds \qquad (2.7b)$$

$$= \frac{F(t+h;z)-F(t;z)}{S(t;z)} \qquad (2.7c)$$

$$= \frac{S(t;z)-S(t+h;z)}{S(t;z)} \qquad (2.7d)$$

$$= 1 - \exp(-(\Lambda(t+h;z)-\Lambda(t;z))). \qquad (2.7e)$$

The likelihood for the (unknown) exact individual data (1.3) can now be constructed. In order to do this, note that the contribution to this likelihood of a woman i who experiences first union ($\delta_i=1$) at exact time t_i is $f(t_i;z_i)=S(t_i;z_i).\mu(t_i;z_i)$, and that the contribution of a woman i who is censored ($\delta_i=0$) at exact time t_i is $S(t_i;z_i)$. In general, the contribution of woman i is $S(t_i;z_i)(\mu(t_i;z_i))^{\delta_i}$. Under the assumption that the individual data are independent, and that the mechanisms of entering the state of first union and of censoring are independent, the likelihood for the (unknown) exact individual data (1.3) is proportional to

$$\mathcal{L} = \prod_{i=1}^{n} S(t_i;z_i)(\mu(t_i;z_i))^{\delta_i} . \qquad (2.8)$$

With $\mathcal{Z}$ defined as the set of distinct observed [(2)] covariates z as in Section 1, obvious reorganization of the factors involved yields

$$\mathcal{L} = \prod_{z\in\mathcal{Z}} \prod_{i\in\mathcal{R}(0,z)} S(t_i;z)\ (\mu(t_i;z))^{\delta_i} , \qquad (2.9)$$

where $\mathcal{R}(t,z)$ is the set of women i with covariates z who are at time t exposed to the risk of entering the state of first union, i.e. women with $t_i \geqslant t$.

Since it is clear from (2.9) that the likelihood ($\mathcal{L}$) is a product of likelihoods for different subgroups z, reference to covariates z may be dropped from the notation as long as the discussion does not concern them explicitly. The likelihood is then

$$\mathcal{L} = \prod_{i \in \mathcal{R}(0)} S(t_i)(\mu(t_i))^{\delta_i}. \tag{2.10}$$

Since it is convenient to work with the log-likelihood rather than with the likelihood as such, it is useful to keep the following formulae in mind.

$$\log \mathcal{L} = \sum_i \{ \delta_i \log \mu(t_i) + \log S(t_i)\} \tag{2.11a}$$

$$= \sum_i \{\delta_i \log \mu(t_i) - \Lambda(t_i)\} \tag{2.11b}$$

2.2. Piecewise exponential models

Consider a set of intervals $[a_o,a_1),[a_1,a_2),\ldots[a_{L-1},a_L)$, with $a_o=0$ and $a_{l-1}<a_l(l=1,\ldots L)$.(3) Assume that the hazard $\mu(t)$ (ignoring covariates z) is constant in each interval $[a_{l-1},a_l)$, say

$$\mu(t) = e^{\alpha_l} \qquad \text{for } a_{l-1} \leqslant t < a_l \ . \tag{2.12}$$

Note that the exponentiation ensures that the hazard is positive (without any further contraint on the parameters α_l) as it should be.

In particular, if the hazard is constant over the entire interval $[a_o,a_L)$, we get the model

$$\mu(t) = e^{\alpha} \qquad \text{for } a_o \leqslant t < a_L \ . \tag{2.13}$$

Since time is exponentially distributed in $[a_o,a_L)$ under model (2.13), whereas under model (2.12) its distribution is only separately exponential in each interval $[a_{l-1},a_l)$, the latter model is called a *piecewise exponential* model.

The lengths of the intervals in the above partition may vary and be different from unity. Clearly reality is better approximated by smaller intervals. However, since our unit of time measurement has been taken to be 1 year (Section 1), the length of the intervals we will deal with will not be less than 1 : we will continue to use intervals of unit length (i.e. $a_l=l$). The mathematical formulation developed for intervals of unit length can easily be adopted for a partition into intervals with different lengths, though this will not be treated here as it is not essential for our purpose.

Under model (2.12), the cumulative hazard function becomes

$$\Lambda(t) = \sum_{k<l} e^{\alpha_k} + e^{\alpha_l}.(t-a_{l-1}) \qquad \text{for } a_{l-1} \leqslant t < a_l \ . \tag{2.14}$$

The survivor function can be found from $S(t) = \exp(-\Lambda(t))$ using the value of $\Lambda(t)$ given in (2.14).(4)

The conditional probability of entering into first union in the l-th interval, given that first union has not been experienced before a_{l-1}, is (under model (2.12), using (2.14) and (2.7e)) found to be

$$q_l = q(a_{l-1},1) = 1 - e^{-e^{\alpha_l}} \qquad (2.15)$$

Substitution of (2.12) and (2.14) in (2.11b) gives us the log-likelihood for the piecewise exponential model, which, after some rearrangement of terms (Holford (1976, 1980); Laird and Olivier (1981)), can be expressed as follows

$$\log \mathcal{L} = \sum_{l=1}^{L} \{ d_l.\alpha_l - E_l.e^{\alpha_l} \}, \qquad (2.16)$$

where d_l is the number of women who experience first union in the l-th interval and where E_l stands for the exact exposure time (person-years lived outside the state of first union) in the l-th interval. The log-likelihood for the exponential model (2.13) is easily found from (2.16) using $\alpha_l=\alpha$:

$$\log \mathcal{L} = d.\alpha - E.e^{\alpha}, \qquad (2.17)$$

where $d = \Sigma\, d_l$ and $E = \Sigma\, E_l$ are respectively the total number of women who experience first union in $[a_o,a_L)$ and the total exposure time in $[a_o,a_L)$.

If the exact individual data (1.3) were known, the exposure times E_l could be calculated exactly. However, since we observe only the individual data (1.2), only an appropriate estimation $\tilde{E}_l$ for E_l can be arrived at. We will use the estimation

$$\tilde{E}_l = \sum_{i\in\mathcal{R}_l} \tilde{E}_{il} \qquad (2.18a)$$

with
$$\tilde{E}_{il} = 1 \text{ if } \underline{t}_i \geq a_l$$

$$= 1/2 \text{ if } a_{l-1} \leq \underline{t}_i < a_l \qquad (2.18b)$$

$$= 0 \text{ if } \underline{t}_i < a_{l-1},$$

where $\tilde{E}_{il}$ is an estimate of the exact exposure time E_{il} for woman i in the l-th (unit-) interval, and where $\mathcal{R}_l = \mathcal{R}(a_{l-1})$, the set of women who at the beginning of the l-th interval, have not yet experienced first union. The estimation (2.18b) assumes that all events (i.e. entering into first union or censoring) are uniformly distributed over each interval $[a_{l-1}, a_l)$, an assumption which seems to be appropriate for our data on first union (Section 1).(5)

Substitution of (2.18a-b) in (2.16) gives us the log-likelihood

$$\log \mathcal{L} = \sum_{l=1}^{L} \{ d_l . \alpha_l - \tilde{E}_l . e^{\alpha_l} \} \qquad (2.19)$$

which is in fact the log-likelihood for the individual data (1.2) - under the above assumption - and clearly also for the grouped data (1.5) (if covariates are ignored).

Reintroducing covariates z in the discussion, we have, by (2.9) and the comment which follows it :

$$\log \mathcal{L} = \sum_{z \in \mathcal{Z}} \sum_{l=1}^{L} \{ d_{lz} . \alpha_{lz} - \tilde{E}_{lz} . e^{\alpha_{lz}} \} \qquad (2.20)$$

which is the log-likelihood for the data (1.2) or (1.5) under the model

$$\mu(t;z) = e^{\alpha_{lz}} \qquad \text{for } a_{l-1} \leq t < a_l \qquad (2.21)$$

i.e. assuming a piecewise exponential model for each subgroup z. Note that $\tilde{E}_{lz} = n_{lz} - \frac{1}{2}(w_{lz} + d_{lz})$ where n_{lz}, w_{lz} and d_{lz} are defined in Section 1.

The solution of the system of *maximum likelihood equations* obtained by equating the partial derivatives of $\log \mathcal{L}$ in (2.20) with respect to each parameter α_{lz} to zero yields the following estimates for the hazards in (2.21).

$$\tilde{\mu}(t;z) = e^{\tilde{\alpha}_{lz}} = \frac{d_{lz}}{\tilde{E}_{lz}} \qquad \text{for } a_{l-1} \leq t < a_l \qquad (2.22)$$

Thus, under the piecewise exponential model the hazard in the l-th interval, and for subgroup z, is estimated by the *(observed) occurence-exposure rate* $d_{lz}/\tilde{E}_{lz}$.

2.3 Proportional hazards models

Model (2.21) does not specify any relation between subgroups. In fact, application of (2.21), with estimates as in (2.22), is equivalent to the application of standard life table techniques for each subgroup separately. The proportional hazards (PH) assumption provides a way of modelling the relation between subgroups.

The simplest form of the PH model is

$$\mu(t;\mathbf{z}) = e^{\alpha_l} . e^{\beta_{\mathbf{z}}} \qquad \text{for } a_{l-1} \leqslant t < a_l, \qquad (2.23)$$

where the parameter $\beta_{\mathbf{z}}$ depends merely on covariates $\mathbf{z}$ and the parameter α_l depends merely on time. If we suppose that $\beta_{\mathbf{z}_o} = 0$ for some $\mathbf{z}_o$ - $\mathbf{z}_o$ signifying a *reference subgroup* - then the series $\exp(\alpha_l)$ ($l=1,...L$) gives the hazard function for this reference subgroup. The parameter $\exp(\alpha_l)$ will in this case be referred to as the *base-line hazard* (for the l-th interval); and each parameter $\beta_{\mathbf{z}}$ becomes a measure of the difference between subgroup $\mathbf{z}$ and the reference subgroup $\mathbf{z}_o$. Alternatively, this difference is measured by the *relative risk* $\mu(t;\mathbf{z})/\mu(t;\mathbf{z}_o)$. Under the PH model (2.23) this relative risk is constant over time and equal to $\exp(\beta_{\mathbf{z}})$. It is said that covariates $\mathbf{z}$ act multiplicatively on the hazard. Note also that under the PH model there is no interaction between time t and covariates $\mathbf{z}$, whereas the general model (2.21) allows for such interactions.

A model which is less restrictive than the PH model (2.23) but more restrictive than the general (piecewise exponential) model (2.21) is the so called *stratified proportional hazards* (SPH) model. We now introduce this model through an example.

Consider the covariates REL, EDU and COH as defined in Section 1. Let z_1 = REL, z_2 = EDU and z_3 = COH, so that we have $\mathbf{z} = (z_1,z_2,z_3)$ = (REL,EDU,COH). As in Section 1 we shall, when necessary, use the expression *subgroup* $\mathbf{z}$ to refer in general to any one of the subgroups to which the sample population is partitioned by virtue of all covariates z_1, z_2 and z_3. In the present example we have 24 such subgroups : (1,1,1), (1,1,2), (1,2,1), (1,2,2),......(4,3,2). Note for instance that subgroup (1,2,1) is the set of Roman Catholic women with regular Mass attendance, having secondary education as the highest educational level attained and born between 1948 and 1962.

Under the PH model (2.23), the time dependence of experiencing first union in each subgroup z is measured by the same series of base-line hazards exp (α_l) $(l = 1,\ldots L)$. The PH model may however not provide a satisfactory fit for the data, and this could perhaps be attributed to the inappropriateness of using the same base-line hazards exp (α_l) for <u>all</u> subgroups. Suppose now that there is sufficient evidence that a better fit would emerge through the use of different base-line hazards for the different layers (or *strata*) into which the total sample population could be partitioned in relation to one or more covariates. Such a partition would arise in the example under consideration if, for instance, each group of women characterized by a specific religious affiliation were considered as a stratum. We would then have 4 strata to deal with : i.e.

REL = 1 : Roman Catholic women with regular Mass attendance,

REL = 2 : Roman Catholic women with irregular or no Mass attendance,

REL = 3 : women with no religious affiliation,

REL = 4 : freethinkers.

We will speak generally about a *stratum* z_1, in the same way that we speak about a *subgroup* z. Note that we can specify 6 subgroups referenced in relation to the remaining covariates z_2 = EDU and z_3 = COH in each of the 4 strata. For instance, in stratum z_1 = REL = 1 we have the subgroups (1,1,1), (1,1,2), (1,2,1), (1,2,2), (1,3,1) and (1,3,2). (See Figure C1.)

Suppose now that the time dependence of the experience of first union in stratum z_1 is measured by the series of hazards exp $(\alpha_{l z_1})$ $(l=1,\ldots L)$. There are 4 such series in the present example, and there is in general no simple relation between them. Suppose further that the hazards for an arbitrary subgroup z in stratum z_1 could be obtained simply by multiplying the series of hazards exp $(\alpha_{l z_1})$ $(l = 1,\ldots L)$ by a single factor exp (β_z). The series of hazards measuring the experience of first union of a subgroup z in stratum z_1 would then be given by exp $(\alpha_{l z_1} + \beta_z)$ $(l=1,\ldots L)$.

In other words, within each of the 4 strata z_1, we have considered a simple PH model. The resulting model valid for all the strata considered is called a SPH model.

In general, the SPH model can be symbolized as follows.

$$\mu(t;\mathbf{z}) = e^{\alpha_{l\mathbf{z}_1}} . e^{\beta_{\mathbf{z}}} \qquad \text{for } a_{l-1} \leqslant t < a_l \qquad (2.24)$$

where $\mathbf{z}$ is a subgroup in stratum $\mathbf{z}_1$. Note that this notation is quite general with respect to the stratum $\mathbf{z}_1$: i.e. the stratification may depend on more than one covariate – say covariates $z_1,\ldots.z_{m_1}$, in which case we define a $1 \times m_1$ vector $\mathbf{z}_1$ as being equal to $(z_1,\ldots.z_{m_1})$. If the remaining $m-m_1$ covariates $z_{m_1+1},\ldots.z_m$ are used to define a $1 \times (m-m_1)$ vector $\mathbf{z}_2 = (z_{m_1+1},\ldots,z_m)$, then the $1 \times m$ vector $\mathbf{z} = (\mathbf{z}_1,\mathbf{z}_2)$ denotes a general subgroup in stratum $\mathbf{z}_1$. Since the stratification effected above was done in reference to the covariate REL, $\mathbf{z}_1$ is the 1×1 vector $(z_1) = (\text{REL})$, and $\mathbf{z}_2$ becomes the 1×2 vector $(z_2,z_3) = (\text{EDU},\text{COH})$. In the applications given later (Section 2.5) we will have occasion to consider the case where $\mathbf{z}_1$ is taken as the 1×2 vector $(z_1,z_2) = (\text{REL}, \text{EDU})$. Note that in this case there are 2 stratifying variables REL and EDU.

Consider now a *reference subgroup* $\mathbf{z}_o$ in each stratum $\mathbf{z}_1$, and suppose that $\exp(\beta_{\mathbf{z}_o}) = 0$. The relative risk $\exp(\beta_{\mathbf{z}})$ would then measure the difference between the experiences of first union in subgroup $\mathbf{z} = (\mathbf{z}_1,\mathbf{z}_2)$ and the <u>corresponding</u> reference subgroup $\mathbf{z}_o$. Note that $\mathbf{z}_o$ is a $1 \times m$ vector $(\mathbf{z}_1,\mathbf{z}_{20})$ and that the first m_1 components of $\mathbf{z}$ and $\mathbf{z}_o$ are equal. The element $\mathbf{z}_{20}$ of vector $\mathbf{z}_o$ is merely a vector of $m-m_1$ covariates which, with the m_1 covariates of $\mathbf{z}_1$, serves to define the subgroup $\mathbf{z}_o$. (See Figure C1.)

The following relations should be noted (Appendix E3). Under the PH model (2.23) it can be shown that

$$S(t;\mathbf{z}) = (S(t;\mathbf{z}_o))^{e^{\beta_{\mathbf{z}}}} \qquad (2.25)$$

$$\Lambda(t;\mathbf{z}) = \Lambda(t;\mathbf{z}_o).e^{\beta_{\mathbf{z}}} \qquad (2.26)$$

$$q_l(\mathbf{z}) = 1 - (1-q_l(\mathbf{z}_o))^{e^{\beta_{\mathbf{z}}}} \qquad (2.27)$$

If z is an arbitrary subgroup in a specific stratum and if z_o is the corresponding reference subgroup in the same stratum - i.e. if $z = (z_1, z_2)$ and $z_0 = (z_1, z_{20})$ so that the two subgroups are situated in the same stratum z_1 - then formulae (2.25) to (2.27) are also valid under the SPH model (2.24).

Using the SPH model (2.24), the experience of first union is described by a set of series of base-line hazards exp (α_{lz_1}) $(l = 1, \ldots . L)$ - i.e. one series for each stratum z_1 - and a set of relative risks exp (β_z). There is in general no simple relationship between the base-line hazards exp $(\alpha_{lz'_1})$ of any one stratum z'_1 (say) and those, exp $(\alpha_{lz''_1})$, of another stratum z''_1 (say). Thus the two series of base-line hazards can be completely different, meaning that the time dependence of the experience of first union in the two strata can be quite different. Particular cases of the general SPH model may be obtained by assuming a specific relation between base-line hazards. In our case it was found useful to assume (1°) that the entrance into first union of women in stratum z'_1 starts b units ahead of that of women in stratum z''_1, and (2°) that once the process has begun in the two strata, the corresponding hazards are proportional. In other words : whereas the general SPH model leads to the loss of the assumption of proportional hazards across strata, our parametrization of the stratification will readopt this assumption after taking into account the different starting points. This can be symbolized by the equation

$$\mu(t;z'_o) = \mu(t+b;z''_o).e^{\omega} \qquad (2.28a)$$

where z'_o and z''_o are the reference subgroups in the two strata z'_1 and z''_1 respectively. If b is an integer (see Section 2.4), then we may rewrite (2.28a) as

$$e^{\alpha_{lz'_1}} = e^{\alpha_{l+bz''_1} + \omega} \qquad (2.28b)$$

The following relations can then be shown to hold (Appendix E4) :

$$S(t;z'_o) = (S(t+b;z''_o))^{e^{\omega}} \qquad (2.29)$$

$$\Lambda(t;z'_o) = \Lambda(t+b;z''_o).e^{\omega} \qquad (2.30)$$

$$q(a_{l-1},1;z'_o) = 1 - (1-q(a_{l-1}+b,1;z''_o))^{e^{\omega}}. \qquad (2.31a)$$

The latter equation may be written as

$$q_l(z'_o) = 1-(1 - q_{l+b}(z''_o))^{e^{\omega}} \qquad (2.31b)$$

if b is an integer.

Equations (2.29) to (2.31), which are all equivalent to equation (2.28), are useful for comparison of reference subgroups in different strata. For comparisons within a particular stratum z_1, one can use formulae (2.25) to (2.27). Comparison of an arbitrary subgroup in one stratum with an arbitrary subgroup in another stratum is now, in the SPH model with the additional assumption (2.38a), also possible through the use of just a few parameters. Before explaining this in detail we make the following remark concerning the concept of *reference subgroups* in SPH models.

It is convenient to define the *reference subgroups* z'_o and z''_o for any two strata z'_1 and z''_1 by the vectors (z'_1, z_{20}) and (z''_1, z_{20}) respectively. Note that z_{20}, a particular $1 \times (m-m_1)$ vector, is the same here for all strata. In our example, we could, for instance, take the reference subgroups (1,1,1), (2,1,1), (3,1,1) and (4,1,1). As a matter of fact, equation (2.28a) would then serve to compare any two reference subgroups of women with *different* religious affiliation - i.e. women in different strata - but with the *same* level of education attained and belonging to the *same* birth cohort. The shift parameters b and the relative risks exp (ω) then measure the effect of the covariate REL. In general, the shift parameters b and the relative risks exp (ω) measure the effect of the stratifying variables $z_1, \ldots z_{m_1}$ (adjusted for effects of the remaining covariates $z_{m_1+1}, \ldots z_m$).

From (2.28a) we see that it makes sense to introduce the concept of a *reference stratum*. If z_{10} is the $1 \times m_1$ vector denoting the reference stratum, if z_{oo} is the reference subgroup in the reference stratum z_{10}, and if z_o is the reference subgroup in an arbitrary stratum z_1, then equation (2.28a) may be rewritten as

$$\mu(t;z_o) = \mu(t + b_{z_1};\ z_{oo}).e^{\omega_{z_1}} \qquad (2.32)$$

where b_{z_1} and ω_{z_1} together, as mentionned in the previous paragraph, measure the difference between stratum z_1 and the (fixed) reference stratum z_{10}.(6) Further, if z is an arbitrary subgroup in stratum z_1, then we have

$$\mu(t;z) = \mu(t;z_o).e^{\beta_z},$$

and the combination of this relation with (2.32) gives :

$$\mu(t;z) = \mu(t+b_{z_1};z_{oo}).e^{\omega_{z_1}+\beta_z}.$$

If we define new parameters β'_z to be equal to $\omega_{z_1}+\beta_z$ (if z is a subgroup of stratum z_1), then we get the relation

$$\mu(t;z) = \mu(t+b_{z_1};z_{oo}).e^{\beta'_z}. \qquad (2.33)$$

Strictly speaking the model (2.33) is a SPH model. However, apart from the shift, the formula is that for ordinary PH models. This implies that the relative risks exp (β'_z) can be interpreted as the relative risks in ordinary PH models : i.e. exp (β'_z) measures the difference between subgroup z and the reference subgroup z_{oo} after adjusting for the difference in starting points in different strata.(7) Note that (2.33) refers only to one reference subgroup z_{oo}. The process of entry into first union is thus completely described by (1°) a base-line hazard $\mu(t;z_{oo})$ corresponding to the reference subgroup z_{oo}, (2°) the shift parameters b_{z_1} measuring the difference between starting points in stratum z_1 and the reference stratum z_{10} (to which z_{oo} belongs), and (3°) the (adjusted) relative risks exp (β'_z) which measure the remaining difference between the process in subgroup z (in stratum z_1) and in subgroup z_{oo}.

Consider now two *arbitrary subgroups* of women in different strata but whose characteristics measured by the covariates $z_{m_1+1},\dots z_m$ are the same. Such subgroups can in general be denoted by the 1 x m vectors $z' = (z'_1,z_2)$ and $z''= (z''_1,z_2)$. Since the SPH model implies a simple PH model within each stratum, the difference between subgroup z' and the corresponding reference subgroup $z'_o = (z'_1,z_{20})$ in stratum z'_1 is measured by the relative risk exp $(\beta_{Z'}) = \mu(t;Z')/\mu(t;Z'_o)$. Similarly, the relative risk exp $(\beta_{Z''}) = \mu(t;Z'')/\mu(t;Z''_o)$ measures the difference between subgroup Z'' and the corresponding reference subgroup $Z''_o = (Z''_1, Z_{20})$ in stratum Z''_1. In other words, exp $(\beta_{Z'})$ measures the effect of the covariates $z_{m_1+1},\ldots z_m$ in stratum Z'_1, and exp $(\beta_{Z''})$ measures the effect of the same covariates in stratum Z''_1. In general, estimates of exp $(\beta_{Z'})$ and exp $(\beta_{Z''})$ will be different, meaning that the effect of the covariates $z_{m_1+1},\ldots z_m$ is different across strata. A special SPH model would be obtained by assuming that the effect of covariates $z_{m_1+1},\ldots z_m$ does not depend on strata. Formally, exp $(\beta_{Z'})$ is then equal to exp $(\beta_{Z''})$, and these relative risks may then be denoted by exp (β_{Z_2}), where Z_2 is the $1 \times (m-m_1)$ vector of covariate values $z_{m_1+1},\ldots z_m$. In this case, the model formula (2.24) becomes

$$\mu(t;Z) = e^{\alpha_{l Z_1}} . e^{\beta_{Z_2}} \qquad \text{for } a_{l-1} \leqslant t < a_l,$$

where Z still remains (Z_1, Z_2). It is easy to see that there is then no interaction between the stratifying covariates $z_1,\ldots z_{m_1}$ and the remaining covariates $z_{m_1+1},\ldots z_m$.

Use of the assumption of no interaction between stratifying covariates $z_1,\ldots z_{m_1}$ and remaining covariates $z_{m_1+1},\ldots z_m$ under the SPH model (2.33) implies the relation $\beta'_Z = \omega_{Z_1} + \beta_{Z_2}$ (if $Z = (Z_1, Z_2)$) and hence yields the model formula

$$\mu(t;Z) = \mu(t+b_{Z_1};Z_{oo}).e^{\omega_{Z_1}}.e^{\beta_{Z_2}}, \qquad (2.34)$$

relating an arbitrary subgroup $Z = (Z_1, Z_2)$ in stratum Z_1 to the reference subgroup Z_{oo} in the reference stratum Z_{10}. Under this model the process of entry into first union can thus be described by (1°) a single base-line hazard $\mu(t;Z_{oo})$, (2°) shift parameters b_{Z_1} measuring the difference between starting points across strata, (3°) relative risks exp (ω_{Z_1}) measuring the remaining difference between strata, and (4°) relative risks exp (β_{Z_2}) measuring the difference between subgroups within each stratum.

Note that under the models (2.33) or (2.34) the formula relating survivor functions, cumulative hazard functions or conditional probabilities can be easily derived from formulae (2.29) to (2.31) by replacing z_o', z_o'', b and ω by z, z_{oo}, b_{z_1} and β_z' respectively.

Replacing the parameters α_{lz} in (2.20) by $\alpha_l+\beta_z$ yields the log-likelihood for the PH model (2.23) :

$$\log \mathcal{L} = \sum_{z\varepsilon\mathcal{Z}} \sum_{l=1}^{L} \{d_{lz}.(\alpha_l+\beta_z) - \tilde{E}_{lz}.e^{\alpha_l+\beta_z}\}. \tag{2.35}$$

Similarly, if we replace α_{lz} in (2.20) by $\alpha_{lz_1}+\beta_z$, then we get the log-likelihood for the SPH model (2.24) :

$$\log \mathcal{L} = \sum_{z\varepsilon\mathcal{Z}} \sum_{l=1}^{L} \{d_{lz}.(\alpha_{lz_1}+\beta_z) - \tilde{E}_{lz}.e^{\alpha_{lz_1}+\beta_z}\}. \tag{2.36}$$

Estimates of the parameters are found by solving the system of maximum likelihood equations, which are obtained by equating the partial derivatives of log $\mathcal{L}$ with respect to each parameter to zero. In general, the maximum likelihood equations must be solved by iteration, and a closed form expression (such as (2.22)) for the parameters does not exist. (Some details are given in Appendix E5.)

Under the special SPH model (2.33) the log-likelihood is

$$\log \mathcal{L} = \sum_{z\varepsilon\mathcal{Z}} \sum_{l=1}^{L} \{d_{lz}.(\alpha_{l+b_{z_1}} + \beta_z') - \tilde{E}_{lz}.e^{\alpha_{l+b_{z_1}}+\beta_z'}\} \tag{2.37}$$

where it is assumed again that the hazards are piecewise constant (over intervals of unit length) and that the parameters b_{z_1} are integers. We cannot obtain a system of maximum likelihood equations in the parameters β_z', b_{z_1} and α_l simultaneously. However, it is possible to estimate the parameters α_l and β_z' (simultaneously) <u>if the parameters b_{z_1} are fixed</u>. The construction of the appropriate system of maximum likelihood equations is tedious. However, the solution of the problem is quite simple if we use the idea that the time range for stratum z_1 should be translated over b_{z_1} units relative to the time range in the reference stratum. After such translations, the reduced problem is equivalent to the problem of estimating α_l and β_z under the PH model; i.e. maximization of log $\mathcal{L}$ in (2.35).

A method for maximization of the log-likelihoods (2.35) and (2.36) will be discussed in the following section.

2.4 Estimation through GLIM

It will be shown in this section that the (log-) likelihoods considered previously can be maximized through a method developped for a class of *generalized linear models* (GLM) by Nelder and Wedderburn (1972). The discussion which follows is GLIM3 oriented, since this computer package is extensively used in the applications.[8]

To begin with, note that the log-likelihoods (2.35), (2.36) and (2.37) are special cases of the log-likelihood (2.20), but that maximum likelihood estimation of the shift parameters in (2.37) is not possible. For a general discussion of the methods involved, we will therefore concentrate on the estimation of the parameters α_{lz} in the log-likelihood (2.20).

A GLM suitable for use with GLIM3 is defined by specifying the following three model components (Baker and Nelder, 1978) :

(1°) a set of dependent variables, which are statistically independent and whose distributions belong to the exponential family (The distribution of the dependent variables is called the *error structure*);

(2°) a set of independent variables and the way in which they are related to each other in producing their effects on the dependent variables (This is expressed in the *linear predictor* - a linear combination of the independent variables. Note that any categorical variables (covariates) in use and interactions between them should be transformed into *dummy variables* in order to allow a linear combination.);

(3°) the manner in which the independent variables act on the dependent variables (The function relating the linear predictor to the dependent variables is called the *link*).

It is now necessary to define a GLM - i.e. through specification of the three components discussed above - which would lead to the same parameter estimates as those obtained through the maximization of log $\mathcal{L}$ in (2.20). The required GLM is defined as follows :

(1°) the dependent variables - with realisations d_{lz} ($z\varepsilon$; $l=1,\dots L$) - are assumed to be statistically independent and Poisson distributed with means M_{lz}, say (The counts d_{lz} are then said to have the Poisson error structure);

(2°) the covariates (i.e. categorical independent variables) taken into count consist of a time covariate - with levels $l=1,\ldots L$ and other covariates denoted by the vector $\mathbf{z}$: the corresponding linear predictor will be denoted by $\alpha_{l\mathbf{z}}$;

(3°) the link between the linear predictor $\alpha_{l\mathbf{z}}$ and the mean $M_{l\mathbf{z}}$ of the dependent variable $d_{l\mathbf{z}}$ is presumed to be given by

$$\log M_{l\mathbf{z}} = \log \tilde{E}_{l\mathbf{z}} + \alpha_{l\mathbf{z}}. \tag{2.38}$$

where $\tilde{E}_{l\mathbf{z}}$ is the corresponding (approximate) exposure time.

Since the relevant Poisson probabilities take the form $M_{l\mathbf{z}}^{d_{l\mathbf{z}}}.e^{-M_{l\mathbf{z}}}/d_{l\mathbf{z}}!$, the likelihood for the GLM defined above will be proportional to

$$\bar{\mathcal{L}} = \prod_{\mathbf{z}} \prod_{l} M_{l\mathbf{z}}^{d_{l\mathbf{z}}}.e^{-M_{l\mathbf{z}}}. \tag{2.39}$$

Substitution of (2.38) in (2.39) gives :

$$\bar{\mathcal{L}} = \prod_{\mathbf{z}} \prod_{l} \tilde{E}_{l\mathbf{z}}^{d_{l\mathbf{z}}}.e^{d_{l\mathbf{z}}.\alpha_{l\mathbf{z}}}.e^{-\tilde{E}_{l\mathbf{z}}.e^{\alpha_{l\mathbf{z}}}}. \tag{2.40}$$

Taking logarithms we get

$$\log \bar{\mathcal{L}} = \sum_{\mathbf{z}} \sum_{l} (d_{l\mathbf{z}}.\log \tilde{E}_{l\mathbf{z}} + d_{l\mathbf{z}}.\alpha_{l\mathbf{z}} - \tilde{E}_{l\mathbf{z}}.e^{\alpha_{l\mathbf{z}}}). \tag{2.41}$$

Maximization of $\log \mathcal{L}$ in (2.20) and $\log \bar{\mathcal{L}}$ in (2.41) is equivalent, since the right hand sides of the two equations in question differ only by a constant term (i.e. $\sum_{\mathbf{z}} \sum_{l} d_{l\mathbf{z}}.\log \tilde{E}_{l\mathbf{z}}$).

A remark concerning equation (2.38) is in order at this point. This equation shows that the link function relates the means $M_{l\mathbf{z}}$ not merely to the covariates l (i.e. the time covariate) and $\mathbf{z}$, but also to the exposure times $\tilde{E}_{l\mathbf{z}}$. It follows that $\log \tilde{E}_{l\mathbf{z}}$ takes on the nature of a continuous independent variable. It would moreover have a coefficient equal to unity if it were used as a regressor in a regression model.

In the language of GLIM this would mean that if $\log \tilde{E}_{lz} + \alpha_{lz}$ is the linear predictor, part of it needs to be fixed. A fixed term forming part of the linear predictor is referred to in GLIM as an offset. The $OFFSET-statement effects this fixation.

Holford (1980) shows that the likelihood (2.40) may be derived in other ways : it can for instance be assumed that the <u>individual</u> exposure times in each interval are exponentially distributed, or it can be assumed that the vectors $(d_{lz})_{z\epsilon\mathcal{Z}}$ have a multinomial distribution. (Laird and Olivier (1981) on the other hand say that the d_{lz} have a Poisson distribution, conditional on $\tilde{E}_{lz}$, with mean $\tilde{E}_{lz}.e^{\alpha_{lz}}$.) It follows then that the maximum likelihood estimates of the parameters α_{lz} can be obtained using either techniques for log-linear models (in relation to rates)(9) or techniques for multinomial frequencies. Commonly used algorithms are often based either on iterative proportional fitting (IPF) or on Newton-Raphson methods. It has been shown above that the models presented in previous sections are equivalent to a particular GLM. Thus, the GLIM3 computer package can now be used. Note that the algorithms in question employ *iterative weighted least squares*, derived from the more general Newton-Raphson algorithm (Nelder and Wedderburn, 1972).

Since the time covariate l in the above GLM is a categorical variable, the length of the intervals need not necessarily be unity. In the latter eventuality, the (approximate) exposure times $\tilde{E}_{lz}$ should then be recalculated : i.e. multiply the individual exposure times in (2.18b) by the interval length $a_l - a_{l-1}$.

Having shown that the general piecewise exponential model (2.21) with log-likelihood (2.20) can be fitted to the data by using the GLIM3 package, it is now necessary to construct appropriate GLIM3 programmes which would suite the PH and SPH models - special cases of the general model (2.21) - described in Section 2.3. These GLIM3 programmes consist of three main parts : (1°) definition of appropriate vectors containing the data and/or other quantities used in the rest of the programme; (2°) construction of the model and its corresponding fitting device; (3°) further computations to get the results desired from the fit in an appropriate form, and the display of these results. We shall discuss these three parts in the following paragraphs. The discussion is based on the GLIM3-programme shown in Appendix B1.

The first part of the programme in Appendix B1 - called the *data definition part* hereafter - consists of the first seven lines. The $UNITS-statement defines the length of the vectors (to be defined later). In our case - and it is so in general - the number of *units* is equal to the length of the data file. The vectors which will contain the data are then defined ($DATA-statement) and the data are read in subsequently ($DINPUT-statement). The following vectors will be used in the present text.

BT = the lower boundary of an <u>age</u> interval (i.e. BT takes the values 15, 16,....35);

N = the number of women at risk at age BT (i.e. at the beginning of the age interval [BT,BT+1));

D1 = the number of women who enter first marriage in age interval [BT,BT+1);

D2 = the number of women who enter into first cohabitation in age interval [BT,BT+1);

W = the number of women withdrawn (or censored) in age interval [BT,BT+1);

REL = the religious category to which the women in question belong;

EDU = their education category;

COH = the birth-cohort to which they belong.

The age variable BT is transformed into the time variable T. In fact T stands then for the numbering l of the age intervals $[a_{l-1},a_l)$. Since all the covariates REL, EDU and COH are categorical, and since the time variable T as defined above is also categorical, the number of categories of each of these covariates (called *factors* in GLIM3) has to be defined through the $FACTOR-statement. The models described in Section 2 do not deal separately with the numbers of first marriage (D1) and first cohabitation (D2), but only with their sum, i.e. the number of first unions as such. These are stored in vector D. Finally, the log-exposure times are calculated and stored in vector LE.

It should be noted that vectors which will not be used after the data definition part may be deleted. In our example we deleted D1 and D2; we could also have deleted N, W and BT.

In the *model definition/fit part* of the GLIM3 programme, we specify the dependent variable ($YVAR-statement), the error structure ($ERROR-statement) and the offset ($OFFSET-statement). The linear predictor is defined and the model is fitted through the $FIT-statement. Note that the link function is implicitly defined in the $ERROR-statement - the log link function being the default setting for the Poisson error structure (Baker and Nelder, 1978).

In models used for the analysis of first union as such, D will always be the dependent variable, the error structure will always be Poisson and the offset will always be the log-exposure time LE. Hence, the $YVAR-, $ERROR- and the $OFFSET-statement may not be modified. Thus, only the $FIT-statement calls for modification so as to suite the model to be fitted. In other words, the GLIM3 expression defining the linear predictor α_{lz} has to be in line with the model to be fitted. Details about the construction of the linear predictor will be given later.

Whereas the data definition part and the model definition/fit part of a GLIM3 programme are strongly dependent on the format of the data and the nature of the model to be fitted, the *results part* is not bound by any of these links and is consequently very flexible : the investigator can thus compute and display any quantity of interest. One could, for instance, compute various test statistics, relative risks, (base-line) hazards, survival functions, standard errors of relative risks, etc. In most of our applications we opted for the calculation and/or display of the following results : the terms in the linear predictor (L option in the $DISPLAY-statement), the estimated values of the linear parameters and their standard errors (A option in the $DISPLAY-statement), and two test statistics with the corresponding number of degrees of freedom. These test statistics are defined and discussed in Appendix E7. Note that the log-likelihood statistic for the saturated model - i.e. the observed $-2.\log \mathcal{L}$ value - is computed already between the data definition part and the model definition/fit part.

In certain programmes (or applications) we inserted the computation of base-line hazards and relative risks, together with their (co)variances. An example is discussed in Appendix B4.

Attention should be focussed finally on the specification of the linear predictor α_{lz} in the GLIM3 programmes. Note that at least one term in the linear predictor depends on the time variable T. This implies that we are always dealing with the fit of a piecewise exponential model. The exponential model (i.e. constant hazard in the entire interval $[a_0, a_L)$) would be fitted if the linear predictor did not include the factor T (cfr. Section 2.2). If T is the only term in the linear predictor (i.e. if covariates REL, EDU and COH are not taken into account), the model will be referred to as the *null model*. The $FIT-statement for the null model is thus simply $FIT T. In order to obtain the GLIM3 expression for the linear predictor in PH models, it should be noted that (2.33) can be rewritten as

$$\log \mu(t;z) = \alpha_l + \beta_z$$

which shows that there is a term depending merely on the time covariate and a term depending only on the other covariates. PH models are therefore obtained through a $FIT-statement of the general form

$FIT T + Z

where Z stand for terms depending on covariates REL, EDU and/or COH. For instance, the least restrictive PH model is fitted through $FIT T + REL*EDU*COH, i.e. all main effects and interaction effects of the covariates REL, EDU and COH on the relative risks are taken into account. (Details about the * notation can be found in Baker and Nelder, 1978, Section 13). If the relative risks would only depend on REL, then the appropriate statement would be $FIT T + REL. If the relative risks depend on REL and EDU, but not on COH, then the appropriate statement would be either $FIT T + REL + EDU or $FIT T + REL*EDU.

The GLIM3 expression for the linear predictor in SPH models is easily obtained by rewriting (2.24) as

$$\log \mu(t;z) = \alpha_{lz_1} + \beta_z$$

Hence, one term includes both the time covariate T and some covariates corresponding to the stratifying covariate vector z_1, and the other term includes possibly all covariates REL, EDU and COH. The general form of the $FIT-statement for SPH models can thus be written as :

$FIT T*Z1 + Z

where Z is as before and Z1 stands for the stratifying covariates. For instance, in the case of a stratification according to religious affiliation, the $FIT-statement is $FIT REL*T + REL*EDU*COH if within-stratum relative risks depend on the stratum, or $FIT REL*T + EDU*COH if within-stratum relative risks do not depend on the stratum.

The $FIT-statement for parametrized models of the type (2.33) is formally that for ordinary PH models. However, as we noted in Section 2.3, the shift parameters cannot be estimated through maximum likelihood methods. Therefore, the (possible) shifts should be taken into account a priori by applying the correct transformation of the time variable T. An example is found in Appendix B2 : the shift is 2 years for HIGH-educated women relative to other ones; the time variable T is adjusted for the shifts in the data definition part of the programme.(10) The application of more complex shifts is shown in Appendix B3.

So far we have discussed how the models presented in Sections 2.2-3 can be fitted by using the GLIM3 computer package. In practice, of course, the investigator will always have to keep an eye on the following (closely related) problems :

(1°) Which model fits best?

(2°) Is the proportional hazards assumption (with respect to one or more covariates) appropriate?

(3°) How should the shift parameters be found?

All these problems have to do with the selection of a parsimoneous model, i.e. a model that fits the data adequately and, at the same time, allows for (relatively) easy interpretation of its parameters. To face up to all these problems, we suggest the following rough strategy :

(a) compare PH models mutually and with SPH models to detect the covariates which do not act multiplicatively on the hazard;

(b) select the "best" (general) SPH model;

(c) corresponding to the "best" SPH model, find the most appropriate set of shift parameters.

In steps (a) and (b), use should be made of the idea of *nested* models (e.g. Baker and Nelder, 1978) and the fact that they can easily be compared through the use of differences between their respective *scaled deviances* (and corresponding numbers of degrees of freedom). (See Appendix E7 for details.) In step (c), several sets of shift parameters should be compared to see which one gives the best fit. This is a trial and error procedure. An appropriate initial set, however, may for instance be obtained by careful examination of graphs of the cumulative distribution (or survivor) function as estimated by the best SPH model found in step (b). The initial set of shift parameters may then be improved step by step. The final parametrized SPH model should give estimated cumulative distribution functions (or hazard functions) which are close to the estimates obtained by fitting the (general) SPH model found in step (b).

2.5 Numerical application

This section (1°) deals with the selection of a parsimoneous model for the analysis of entry into first union, and (2°) contains a discussion of the results computed under this model.

A start was made by fitting the ordinary PH model T + REL + EDU + COH. The corresponding GLIM3-programme, together with some results, can be found in Appendix B1. Note that the formula for this PH model can be written as :

$$\mu(t;z) = \mu(t;z_o).e^{\beta_{z_1} + \beta_{z_2} + \beta_{z_3}}$$

Note also that the covariates REL (z_1), EDU (z_2) and COH (z_3) do not interact in their effects on the hazard. The goodness-of-fit statistics were found to be $\hat{\chi}^2_L = 659.93$ and $\hat{\chi}^2_P = 693.42$ with $\nu = 375$ degrees of freedom. The corresponding p-values are about zero, indicating a significant lack of fit. Hence the need to improve the simple PH model used.

As a first step in this direction one could add interactions between covariates to the PH model T + REL + EDU + COH. A number of such PH models - i.e. "between" the null model T and the least restrictive PH model T + REL*EDU*COH - could possibly be fitted with an accompanying analysis of deviance in each case (Appendix E7) aimed at finding the best among them. It may however be argued that none of the PH models was able to provide a good fit. This can be seen for instance by comparing the observed c.d.f. $\tilde{F}(t;z)$ and the c.d.f. $\hat{F}(t;z)$ as estimated from the PH model T + REL + EDU + COH. These c.d.f.'s are shown in Figure C2. Clearly the fit is inadequate in the case of the HIGH-education subgroups. The fact that the c.d.f. is overestimated at lower ages, and underestimated at higher ages indicates that none of the PH models would arrive at providing an adequate fit.

An alternative method of improving the fit is therefore indicated; and this is attempted via the use of a SPH model, i.e. by including interactions between time and covariates in the above PH model. Following the reasons given in the previous paragraph it is seen that the most obvious stratification is that related to the covariate EDU. The goodness-of-fit statistics for the SPH models EDU*(T + REL + COH) and EDU*T + REL + COH are shown in Table A1. The scaled deviance which measures the difference between the SPH model EDU*T + REL + COH and the PH model T + REL + EDU + COH is 284.85 with 40 degrees of freedom. This indicates a (highly) significant improvement. However, the scaled deviance between the SPH models EDU*(T + REL + COH) and EDU*T + REL + COH is 11.27 with 8 degrees of freedom, showing that the difference between these models is not significant (at the 10% level).

Each of these SPH models carries three strata (since EDU has three categories). The graphs of the estimated c.d.f.'s $\hat{F}(t;z)$ under the PH model T + REL + EDU + COH (Figure C2) however indicate that it would perhaps be sufficient to consider only two strata. A new stratifying covariate STR was therefore defined as follows :

STR = 1 if EDU = 1 or 2
      2   EDU = 3

Note that the first stratum is formed of PRI - and SEC - educated women, while the second stratum contains the HIGH - educated women. Two new stratified models were then fitted : STR*(T + REL + EDU + COH) and the more restrictive STR*T + REL + EDU + COH. Their respective goodness-of-fit statistics are shown in Table A1. The scaled deviance measuring the difference between these two SPH models is 5.32 with 4 degrees of freedom, showing that the more restrictive SPH model STR*T + REL + EDU + COH is not significantly worse (at the 25% level) than the SPH model STR*(T + REL + EDU + COH). The stratification related to the covariate EDU (3 strata) and that corresponding to the covariate STR (2 strata) can then be compared : the scaled deviance measuring the difference between the SPH models EDU*T + REL + COH and STR*T + REL + EDU + COH is 51.97 with 20 degrees of freedom. This is highly significant (even at the 1% level), indicating that the model with 3 education strata should be given preference. However, there is a significant lack of fit, even under the preferred model EDU*T + REL + COH, the p-values .0647 and .0096 (corresponding to the statistics $\hat{\chi}^2_L$ and $\hat{\chi}^2_p$ respectively) being still too low.

This lack of fit, which may be due to either the absence of interactions between the covariates EDU, REL and COH or the absence of stratifying covariates other than EDU (or eventually STR), seems to indicate that a more detailed form of stratification is required. (For instance, the SPH model REL*EDU*T + COH might be tried.) The number of time parameters (α_{lz_1}) would however increase fast and this in turn would lessen ease of interpretation. A prudent way out of these difficulties consisted of parametrizing the SPH models - a process which was outlined in Section 2.3.

On the basis of the above discussion and given the estimates under the PH model T + REL + EDU + COH as shown in Figure C2, it was decided to parametrize the SPH model STR*T + REL + EDU + COH. Figure C2 shows that HIGH-educated women (STR=2) tend to postpone entry into first union by about 2 years. Shifts of 1, 2 and 3 years were experimented with, and the corresponding models are denoted T' + REL + EDU + COH, T" + REL + EDU + COH and T"' + REL + EDU + COH respectively in the rest of the text. (The GLIM3-programme for fitting model T" + REL + EDU + COH and the results are shown in Appendix B2.) The corresponding goodness-of-fit statistics are shown in Table A1. It can be seen that the 2-years-shift gives the most satisfactory results.[(11)]

It will now be shown that, given the parametrization T", the log-additive model T" + REL + EDU + COH for the hazards is not significantly worse than any other model which incorporates interactions between the covariates REL, EDU and COH. In order to do this, all models of the type T" + [terms depending on covariates REL, EDU and/or COH] were fitted. The corresponding estimated statistics $\hat{\chi}^2_L$ and $\hat{\chi}^2_P$ with their degrees of freedom ν are listed in Table A2. Any of these models nested in the model T" + REL + EDU + COH can be compared with it by using the scaled deviance $\hat{\chi}^2_L - \hat{\chi}^2_{L,A}$ and the corresponding degrees of freedom $\nu - \nu_A$ (where $\hat{\chi}^2_{L,A}$ is the scaled deviance for the model T" + REL + EDU + COH, with ν_A degrees of freedom). Similar comparisons can also be made between the model T" + REL + EDU + COH and any model in which it is nested. The figures in the last two columns of Table A2 indicate that the main effects of the covariates REL, EDU and COH are all important, but that interactions can be ignored. A more detailed analysis of deviance is represented in Table A3. In sum, this analysis of deviance table shows that, given the parametrization T", the model T" + REL + EDU + COH is that which is most satisfactory.

The c.d.f.'s $\hat{F}(t;z)$, estimated under the model T" + REL + EDU + COH, are shown in Figure C2. Comparison with the PH model T + REL + EDU + COH shows a clear improvement of fit for HIGH-education subgroups, although the fits corresponding to some other subgroups suffer somewhat in this process (the worst affected being the two RC RMA - SEC subgroups). Examination of the observed and estimated schedules in these subgroups indicates that here too women tend to postpone entry into first union. A shift of 1 year seemed best suited to this case. Further, when the c.d.f.'s for RC RMA subgroups were compared with those for other subgroups, it was seen that all Roman Catholic women with regular Mass attendance tend to postpone entry into first union by about 1 year. The 2-years-shift for HIGH-educated women (relative to PRI- and SEC-educated women) and the 1-year-shift for RC RMA women (relative to other women) were consequently combined : the result is schematically represented in Figure C3 (panel b). Note that there are in fact four strata, and that the shift for one stratum relative to another is either 1, or 2, or 3 years. The notation T^* used hereafter will refer to this particular parametrized stratification.

It can be shown that model T^* + REL + EDU + COH is the most satisfactory of all models of the type T^* + [terms depending on covariates REL, EDU and/or COH]. In order to show that the fit obtained in this case of the model T^* + REL + EDU + COH is better than that corresponding to the model T" + REL + EDU + COH, the estimated c.d.f.'s $\hat{F}(t;z)$ were plotted in Figure C2.

No attempt was made to improve the model T^* + REL + EDU + COH by further stratification (or parametrization) in spite of the lack of fit in some subgroups. This attitude was adopted for the following reasons : (1°) the deficiency of the data used (as discussed in Section 1) did not warrant further expendition of effort; (2°) the possibility that inclusion of other covariates and/or redefinition of the covariates used might be more important had to be faced; and (3°) the fact that the present text is primarily intended as an introduction to a certain type of methodology. This section will therefore close with the presentation of some useful results, obtained under the model T^* + REL + EDU + COH.

The shift parameters b_z and the estimated relative risks $\exp(\hat{\beta}'_z)$ - adjusted for the shifts - are shown in Table A4. The interpretation of these parameters is discussed in Section 2.3. Such a table will come in useful when models of the type (2.33) are fitted. The model considered here however - i.e. T^* + REL + EDU + COH - has the property that the effect of any particular covariate on the hazard is not altered by changing the level (or category) of the other covariate(s) : i.e. the covariates REL, EDU and COH do not interact in their effects on the hazards. The effect of any particular covariate can therefore be represented as in Table A5. This table shows that :

(1°) the effect of REL consists mainly in bringing about a shift in the process. RC RMA women are seen merely to postpone entry into first union by about 1 year : for the rest the (adjusted) instantaneous risk (i.e. the hazard) of entering first union is almost the same across the four categories of REL;

(2°) the main effect of COH is to cause a spread of the process : younger women do not tend to postpone entry into first union, but the instantaneous risk of entering first union is much higher for them than for older women;

(3°) the effect of EDU is twofold : HIGH-educated women tend to postpone entry into first union by about 2 years, and the instantaneous risk of entering first union goes down if the level of education increases.

Table A5 does not give a complete picture of the effects of the three covariates on the process of entry into first union : the ultimate proportions $c(z)$ play an important role in this respect since they are the only measures for the (ultimate) level of entry into first union. Estimates can be found in Table A6 : i.e. the quantities $\hat{F}(a_{21};z)$ or the proportion of women with covariates z who have entered first union by the age of 36 years. However, because of the deficiency in the data discussed in Section 1, these estimates are not very reliable.(12)

Finally, estimates of the age Me(z) (resp. P10(z)) at which 50% (resp. 10%) of women with covariates z who will ever experience first union have already experienced this event, were computed and are shown in Table A7. The effect of any particular covariate, as seen through the use of the shift parameters and relative risks, can be found in this table too. For instance, REL has an effect on both P10 and Me : i.e. the whole process is simply shifted by about 1 year for RC RMA women. The effect of COH on P10 is small, but its effect on Me is seen to be more important. The same is true for covariate EDU if we considered only the two lower categories of this covariate, but both P10 and Me suddenly increase by more than 2 years for HIGH-educated women.

3. COMPETING RISKS : THE FIRST MARRIAGE/FIRST COHABITATION MODEL

3.1. Mathematical formulation and likelihood construction

Let T be a continuous random variable, representing the *time* (i.e. *age since 15th birthday*) at which a woman enters either the state of first marriage or the state of first cohabitation, depending on which state she enters first - i.e. T represents the time at which a woman enters the state of first union as in Section 2.1. Let C be a discrete random variable representing the specific state of union she enters first, i.e. either the state of first marriage (C = 1) or the state of first cohabitation (C = 2). For censored women, both T and C cannot be observed. With the definitions of Section 1, and omitting subscript i for woman i, we have T = t and C = δ for a woman who experiences first union (δ = 1 or 2) at time t, whereas we only know that T $>$ t for a woman who is censored (δ = 0) at time t.

The *cause-specific hazard functions*[13] $\mu_j(t;z)$, for women with characteristics z, are defined as

$$\mu_j(t;z) = \lim_{\Delta t \downarrow 0} \frac{P(t \leqslant T < t+\Delta t,\ C=j \mid T \geqslant t,z)}{\Delta t}. \qquad (3.1)$$

The quantity $\mu_j(t;z).\Delta t$, with arbitrarily small Δt, is then interpreted as the probability that a woman with characteristics z enters the state of first union due to cause j in the time interval $[t,t+\Delta t)$, given that she has not experienced first union before time t. A model which takes count of the simultaneous presence of many "causes" of the type under consideration is in general referred to as a *competing risks model*.

A distinction should be drawn between the probability of entering the state of first union due to cause j in $[t,t+\Delta t)$ (given that first union was not experienced before time t) <u>if only risk j is operative in</u> $[t,t+\Delta t)$, and the probability of entering the state of first union due to cause j in $[t,t+\Delta t)$ (given that first union was not experienced before time t) <u>if all risks are operative in</u> $[t,t+\Delta t)$.[14] Note however that these probabilities are assumed (following Makeham and subsequent actuarial practice) to be equal - i.e. equal to $\mu_j(t;z)\Delta t$; see Gail (1975).[15] It is only through this assumption that an estimation of the *pure* distribution of the time at which either entry into first marriage or entry into first cohabitation occurs, becomes possible. In other words, an estimate of what happens to women if they were exposed to only one of the risks is made possible. With the available data (Section 1), this assumption cannot be avoided, though its verification is not possible (Gail, 1975).

If reference is made globally to entry into first union and not to its constitutive "causes", then we have merely to deal with a hazard function defined as in formula (2.1). In that case the quantity $\mu(t;z).\Delta t$ retains the meaning it was given in Section 2, but $\mu(t;z)$ is now called the *total hazard function* (or the *total force of decrement* as in Pollard, 1973). If we assume that entry into first union due to several causes simultaneously is impossible - i.e. that the probability of entering first marriage and entering first cohabitation at the same time is zero - than, for an arbitrarily small interval $[t,t+\Delta t)$,

$$\mu(t;z).\Delta t = \mu_1(t;z).\Delta t + \mu_2(t;z).\Delta t,$$

whence

$$\mu(t;z) = \sum_j \mu_j(t;z). \qquad (3.2)$$

Note that the latter formula is valid for any number of causes J ($\geqslant 1$).

This last assumption is clearly valid in the present study which deals with the entry into first union due to one of the two causes marriage and cohabitation. In cause-specific mortality studies, however, the probability of dying from two or more causes simultaneously is not always zero. For instance, one may die from a car accident, because of a heart attack caused by the car accident. Problems arising from such situations are overcome by considering each simultaneous occurence of two or more causes as a new cause (e.g. Gail, 1975). Thus, formula (3.2) remains quite general.

When the discussion focusses on entry into first union as such and not on the causes of entry, the total survivor function $S(t;z)$, the total cumulative hazard function $\Lambda(t;z)$, the total p.d.f. $f(t;z)$, the total c.d.f. $F(t;z)$, the total ultimate proportion $c(z)$ and the total conditional probability $q(t,h;z)$ - just as explained above in the case of the total hazard function $\mu(t;z)$ - have the same meaning as in Section 2. Clearly, formulae (2.2) to (2.7) continue to be valid.

Besides the cause-specific hazard functions $\mu_j(t;z)$ as defined in (3.1) we also have need of the *cause-specific survivor function* $S_j(t;z)$ and the *cause-specific p.d.f.* $f_j(t;z)$, defined respectively as

$$S_j(t;z) = P(T \geq t,\ C=j|z) \tag{3.3}$$

and

$$f_j(t;z) = \frac{-dS_j(t;z)}{dt}. \tag{3.4}$$

Thus, $S_j(t;z)$ is the probability that a woman with covariates z does not enter into first union before time t and enters into first union due to cause j after t in the presence of other causes. For arbitrarily small Δt, the quantity $f_j(t;z).\Delta t$ is interpreted as the probability that a woman with covariates z enters into first union due to cause j in the interval $[t,t+\Delta t)$: i.e. $f_j(t;z).\Delta t$ is approximately $P(t \leq T < t + \Delta t, C=j|z)$. This probability can also be written as the product $S(t;z).\mu_j(t;z)\Delta t$ of the probability $S(t;z)$ to "survive" all causes of entry into first union till time t and the probability $\mu_j(t;z)\Delta t$ to enter into first union in the interval $[t,t+\Delta t)$. It follows then that

$$f_j(t;z) = S(t;z)\mu_j(t;z). \tag{3.5}$$

From (3.2) it follows immediately that

$$f(t;z) = \sum_j f_j(t;z). \tag{3.6}$$

If we denote the probability that a woman with covariates z enters into first union due to cause j in the interval $[0,t)$ by $Q_j(t;z)$, then

$$Q_j(t;z) = \int_0^t f_j(s;z)ds \tag{3.7}$$

and, by integration of (3.6) over $[0,t)$, we get

$$F(t;z) = \sum_j Q_j(t;z). \tag{3.8}$$

Following Chiang (1968) and Gail (1975), we call $Q_j(t;z)$ the *crude* probability of entering into first union due to cause j in $[0,t)$. The adjectif *crude* refers to the presence of all risks.(16)

Now let $c_j(z)$ stand for the crude probability that a woman with covariates z ever enters into first union due to cause j : i.e.

$$c_j(z) = P(C=j|z) = S_j(0;z). \tag{3.9}$$

Then (Appendix E8)

$$c(z) = \sum_j c_j(z) \tag{3.10}$$

and we have then subsequently

$$\begin{aligned} S(t;z) &= 1 - F(t;z) \\ &= 1 - c(z) + \sum_j c_j(z) - \sum_j Q_j(t;z) \\ &= 1 - c(z) + \sum_j (c_j(z) - Q_j(t;z)). \end{aligned}$$

Since (Appendix E8)

$$S_j(t;z) = c_j(z) - Q_j(t;z) \tag{3.11}$$

we have the relation

$$S(t;z) = 1 - c(z) + \sum_j S_j(t;z). \tag{3.12}$$

If $c(z) = 1$ (i.e. if the event of entering into first union is universal for women with covariates z; see Section 2.1), then equation (3.12) becomes

$$S(t;z) = \sum_j S_j(t;z)$$

which is similar to equations (3.2) and (3.6). It should be noted that

$$c_j(z) = \int_0^\infty f_j(t;z)\,dt = Q_j(\infty;z),$$

whence, using (3.11),

$$S_j(\infty;z) = 0.$$

Since the process corresponding to a cause j is not necessarily universal, the last result implies that the cause-specific survivor function $S_j(t;z)$

is only analogically true to its name. It cannot be interpreted univocally as <u>the</u> survivor function corresponding to cause j.

The crude probability $Q_j(t;z)$ as defined in (3.7) is an unconditional crude probability. The conditional crude probability of entering into first union due to cause j in an interval $[t,t+h)$, given first union has not been experienced before time t, is, for women with covariates z,

$$q_j(t,h;z) = \int_t^{t+h} \frac{S(s;z)\mu_j(s;z)}{S(t;z)} ds \tag{3.13a}$$

$$= \int_t^{t+h} \frac{f_j(s;z)}{S(t;z)} ds \tag{3.13b}$$

$$= \frac{Q_j(t+h;z) - Q_j(t;z)}{S(t;z)} \tag{3.13c}$$

$$= \frac{S_j(t;z) - S_j(t+h;z)}{S(t;z)}. \tag{3.13d}$$

Note that $Q_j(t;z) = q_j(0,t;z)$.

Further with (3.2) (or (3.6), (3.8) or (3.12)) it is easy to derive the following relation between the total conditional probability $q(t,h;z)$ and the crude conditional probabilities $q(t,h;z)$:

$$q(t,h;z) = \sum_j q_j(t,h;z). \tag{3.14}$$

Besides the *total* functions and the *cause-specific* functions defined above, use can also be made of *pseudo* functions. The *pseudo cumulative hazard function* is defined as

$$\Lambda_j(t;z) = \int_0^t \mu_j(s;z)ds \tag{3.15}$$

and the *pseudo survivor function* is defined as

$$G_j(t;z) = \exp\ (-\Lambda_j(t;z)). \tag{3.16}$$

Since, for small Δt, the quantity $\mu_j(t;z).\Delta t$ is approximately the conditional probability of entering into first union due to cause j <u>if only risk j is operative</u> in interval $[t,t+\Delta t)$ - i.e. due to the Makeham assumption - the pseudo functions describe the distribution of entry into first union due to cause j <u>after elimination of all other risks</u>. In other words, if women were only exposed to risk j. Thus the pseudo functions describe the distribution of the *pure* process of entry into first union due to a specified cause j. If $g_j(t;z)$ is the pseudo p.d.f., defined as $g_j(t;z) = -\frac{d}{dt} G_j(t;z)$, then

$$g_j(t;z) = G_j(t;z).\mu_j(t;z) \tag{3.17}$$

and $g_j(t;z).\Delta t$ is, for small Δt, approximately the unconditional probability of entry into first union due to cause j in $[t,t+\Delta t)$ for women with covariates z if they are only exposed to the risk j.

Using (3.2), we can easily derive the relations

$$\Lambda(t;z) = \sum_j \Lambda_j(t;z) \tag{3.18}$$

and

$$S(t;z) = \prod_j G_j(t;z). \tag{3.19}$$

There is no such simple relation between the total p.d.f. $f(t;z)$ and the pseudo p.d.f.'s $g_j(t;z)$.

Using cause-specific functions we have defined the crude probabilities $c_j(z)$, $Q_j(t;z)$ and $q_j(t,h;z)$. Similarly, using pseudo functions we can define *net* probabilities. The adjectif *net* refers to the presence of only one risk, j say.(17)

The net probability, $Q_{(j)}(t;z)$ say, that a woman with covariates z enters into first union due to cause j in the interval $[0,t)$ (in the absence of all other risks) is

$$Q_{(j)}(t;z) = \int_0^t g_j(s;z)ds. \tag{3.20}$$

The net probability, $\pi_j(z)$, that a woman with covariates z ever enters into first union due to cause j (in the absence of all other risks) is

$$\pi_j(z) = \int_0^\infty g_j(t;z)dt = Q_{(j)}(\infty;z). \tag{3.21}$$

The conditional net probability, $q_{(j)}(t,h;z)$, that a woman with covariates z enters into first union due to cause j in the interval $[t,t+h)$, given that she has not experienced first union due to cause j (in the absence of all other risks), is

$$q_{(j)}(t,h;z) = \int_t^{t+h} \frac{G_j(s;z)\mu_j(s;z)}{G_j(t;z)} ds \tag{3.22a}$$

$$= \int_t^{t+h} \frac{g_j(s;z)}{G_j(t;z)} ds \tag{3.22b}$$

$$= \frac{Q_{(j)}(t+h;z) - Q_{(j)}(t;z)}{G_j(t;z)} \tag{3.22c}$$

$$= \frac{G_j(t;z) - G_j(t+h;z)}{G_j(t;z)} \tag{3.22d}$$

$$= 1 - \exp\left(-(\Lambda_j(t+h;z)-\Lambda_j(t;z))\right) \tag{3.22e}$$

Using $g_j(t;z) = -\frac{dG_j(t;z)}{dt} = \frac{dQ_{(j)}(t;z)}{dt}$, $G_j(0;z) = 1$ and $Q_{(j)}(0;z) = 0$, it can be shown that

$$Q_{(j)}(t;z) = 1 - G_j(t;z). \tag{3.23}$$

Note that there is no simple relation between net probabilities and the corresponding total probabilities. Relations between crude and net probabilities will be derived later on, under special model assumptions (see (Sections 3.2 - 3).

The likelihood for the (unknown) exact individual data (1.3) (taking into account the cause of entering into first union) can now be constructed. With this end in view consider - as in Section 2.1 - the contribution to this likelihood for each woman i. The contribution of a woman i who enters into first union due to cause j ($\delta_i = j$) at time t_i is $S(t_i;z_i).\mu_j(t_i;z_i)$; the contribution of a woman i who is censored ($\delta_i = 0$) at time t_i is $S(t_i;z_i)$. In order to arrive at a general expression for the contribution of a woman i, we need to define indicator variables I_j (j = 1,...J) as follows :

$$I_j(\delta_i) = 1 \quad \text{if } \delta_i = j$$
$$= 0 \quad \text{if } \delta_i \neq j. \tag{3.24}$$

The contribution of woman i is then

$$S(t_i;z_i).\prod_j(\mu_j(t_i;z_i))^{I_j(\delta_i)}.$$

If it is assumed, as usual, that the individual experiences of women are independent, and if the mechanisms of entering into the state of first union and of censoring are independent (Lawless, 1982), then the likelihood for the (unknown) exact individual data (1.3) is proportional to

$$\mathcal{L} = \prod_{i=1}^{n} \{S(t_i;z_i).\ \prod_j(\mu_j(t_i;z_i))^{I_j(\delta_i)}\}. \tag{3.25}$$

As in Section (2.1), this likelihood can be rewritten as

$$\mathcal{L} = \prod_{z\in\mathcal{Z}}\ \prod_{i\in\mathcal{R}(0,z)} \{S(t_i;z)\ \prod_j(\mu_j(t_i;z_i))^{I_j(\delta_i)}\}. \tag{3.26}$$

Since it is clear that reference to covariates z may be dropped from the notation as long as the discussion does not concern them explicitly, we will (for the present) concentrate on the likelihood

$$\mathcal{L} = \prod_{i\in\mathcal{R}(o)} \{S(t_i).\ \prod_j(\mu_j(t_i))^{I_j(\delta_i)}\}, \tag{3.27}$$

or on the corresponding log-likelihood

$$\log \mathcal{L} = \sum_i \{ \sum_j [I_j(\delta_i).\log \mu_j(t_i)] + \log S(t_i)\} \tag{3.28a}$$

$$= \sum_i \{ \sum_j [I_j(\delta_i).\log \mu_j(t_i)] - \Lambda(t_i)\} \qquad (3.28b)$$

$$= \sum_j \{ \sum_i [I_j(\delta_i).\log \mu_j(t_i) - \Lambda_j(t_i)] \}. \qquad (3.28c)$$

3.2. <u>Piecewise exponential models and competing risks</u>

As in Section 2.2, consider a partition $[a_o,a_1)$, $[a_1,a_2)$,..... $[a_{L-1},a_L)$, with $a_o=0$ and $a_l-a_{l-1}=1$. Explanatory details concerning this partition as well as details about the piecewise exponential models discussed below are found in Section 2.2.

The formulation of a piecewise exponential model in the presence of competing risks follows the procedure outlined earlier when the simple piecewise exponential model (with only one risk in operation) was presented (see Section 2.2). Here too (i.e. in a competing risks model) it is assumed that each cause-specific hazard $\mu_j(t)$ is constant in each interval $[a_{l-1},a_l)$. Thus

$$\mu_j(t) = e^{\alpha_{jl}} \qquad \text{for } a_{l-1} \leqslant t < a_l. \qquad (3.29)$$

The model (3.29) has already been discussed by Chiang (1968, Ch. 11, Section 3) and is referred to by Laird and Olivier (1981).

Model (3.29) allows for different cause-specific hazards corresponding to different causes of entry into first union. A special case of interest is obtained from the general model by assuming that the cause-specific hazards corresponding to two causes j and k (say) are equal (i.e. that $\alpha_{jl} = \alpha_{kl}$ for all $l = 1,\dots L$). If this happens for each pair of causes, then we get the model

$$\mu_j(t) = e^{\alpha_l} \qquad \text{for } a_{l-1} \leqslant t < a_l \qquad (3.30)$$

where $\alpha_l = \alpha_{1l} = \dots = \alpha_{Jl}$

Under model (3.29), and using (3.2), the total hazard $\mu(t)$ can be written as

$$\mu(t) = \sum_j e^{\alpha_{jl}} \qquad \text{for } a_{l-1} \leqslant t < a_l. \qquad (3.31)$$

Thus, piecewise constant cause-specific hazards $\mu_j(t)$ imply a piecewise constant total hazard $\mu(t)$. Note that under the special model (3.30) the total hazard is

$$\mu(t) = J.e^{\alpha_l} \qquad \text{for } a_{l-1} \leqslant t < a_l. \tag{3.32}$$

Each pseudo cumulative hazard $\Lambda_j(t)$ can be written as

$$\Lambda_j(t) = \sum_{k<l} e^{\alpha_{jk}} + e^{\alpha_{jl}}.(t-a_{l-1}) \qquad \text{for } a_{l-1} \leqslant t < a_l \tag{3.33}$$

under the model (3.29). The total cumulative hazard $\Lambda(t)$ is then, using (3.18), equal to

$$\Lambda(t) = \sum_j \{ \sum_{k<l} e^{\alpha_{jk}} + e^{\alpha_{jl}}.(t-a_{l-1})\} \qquad \text{for } a_{l-1} \leqslant t < a_l. \tag{3.34}$$

Under the special model (3.30) we get

$$\Lambda(t) = J.\{ \sum_{k<l} e^{\alpha_k} + e^{\alpha_l}.(t-a_{l-1})\} \qquad \text{for } a_{l-1} \leqslant t < a_l. \tag{3.35}$$

Substitution of (3.29) and (3.33) (or (3.34)) in (3.28c) (or in (3.28b)) and reorganisation of the terms involved gives us the following expression for the log-likelihood :

$$\log \mathcal{L} = \sum_j \sum_l \{ d_{jl}.\alpha_{jl} - E_l.e^{\alpha_{jl}}\}, \tag{3.36}$$

where d_{jl} is the number of women entering into first union due to cause j in the l-th time interval $[a_{l-1}, a_l)$, and E_l is the total exact exposure time in the l-th interval. Note that the exposure time E_l is independent of the causes j, and that $d_{jl} = \sum_{i \in \mathcal{R}_l} I_j(\delta_i)$.

Under the special model (3.30), the log-likelihood becomes

$$\log \mathcal{L} = \sum_l \{d_l.\alpha_l - J.E_l.e^{\alpha_l}\}, \tag{3.37}$$

where d_l is the number of women entering first union (irrespective of the cause to which this is due) in the l-th interval, i.e. $d_l = \sum_j d_{jl}$. Note the difference between the log-likelihoods (2.16) and (3.37) which are alike in many respects. They are both log-likelihoods under models which ignore the cause of entry into first union. However, in order to test whether or not the cause of entering into first union can be ignored, a comparison has to be made between the maximized values of the log-likelihoods in (3.36) and (3.37) - e.g. through the log-likelihood ratio statistic $(\hat{\chi}_L^2)$. The maximized value of the log-likelihood in (2.16) cannot be used for that purpose.

The log-likelihood (3.36) is the log-likelihood for the exact individual data (1.3) under the piecewise exponential competing risks model (3.29). Since the exact individual data are usually unknown, the exact exposure times E_l cannot be computed. Therefore, proceeding as in Section 2.2, the exposure times E_l are approximated by some $\tilde{E}_l$ (see (2.18a-b)) and the log-likelihood

$$\log \mathcal{L} = \sum_j \sum_l \{d_{jl}.\alpha_{jl} - \tilde{E}_l.e^{\alpha_{jl}}\} \tag{3.38}$$

is then derived. This is thus the log-likelihood for the observed individual data (1.2) or for the grouped data (1.4) (ignoring covariates z).

Reintroducing covariates z, we have

$$\log \mathcal{L} = \sum_{z\in\mathcal{Z}} \sum_j \sum_l \{d_{jlz}.\alpha_{jlz} - \tilde{E}_{lz}.e^{\alpha_{jlz}}\} \tag{3.39}$$

where d_{jlz} is the number of women with covariates z entering the state of first union due to cause j in the l-th time interval, $\tilde{E}_{lz}$ is an approximation for the total exposure time E_{lz} in the l-th interval for women with covariates z, and the α_{jlz} are defined through

$$\mu_j(t;z) = e^{\alpha_{jlz}} \qquad \text{for } a_{l-1} \leqslant t < a_l. \tag{3.40}$$

Equation (3.40) is in fact the general piecewise exponential competing risks model for the data (1.2) or (1.4) involving covariates z.

From (3.39) we can see that *cause of entering first union* can be treated in the same way as the other covariates $z_1,\ldots.z_m$. With the intention of moving in that direction in mind, let z_{m+1} be a covariate taking on values j; let z^* be equal to $(z_1,\dots z_m, z_{m+1}) = (z, z_{m+1})$ and $\mathcal{Z}^*$ be equal to $\{z^* | z^* = (z, z_{m+1}),\ z\varepsilon\mathcal{Z},\ z_{m+1} = j\varepsilon\{1,\dots J\}\}$; let d_{lz^*} be equal to d_{jlz} if $z^* = (z,j)$ and $\tilde{E}_{lz^*}$ be equal to $\tilde{E}_{lz}$ if $z^* = (z, z_{m+1})$ (for any value of z_{m+1}). Then the log-likelihood (3.38) may be written as

$$\log \mathcal{L} = \sum_{z^*\varepsilon\mathcal{Z}^*} \sum_{l} \{d_{lz^*}.\alpha_{lz^*} - \tilde{E}_{lz^*}.e^{\alpha_{lz^*}}\}. \tag{3.41}$$

This has the same form as (2.20) which implies that the competing risks problem can, at least from a formal point of view, be treated as a problem involving only one (combined) cause of entering first union. This conclusion will be used in Section 3.4.

We now present some useful formulae for crude, net and total probabilities of entering first union due to a specified cause in a specified time interval, given that first union has not been experienced before the beginning of that interval. The formulae presented below do not make reference to covariates **z**. The derivation of formulae which incorporate **z** explicitly is left to the reader.

Under the piecewise exponential competing risks model (3.29), the crude probability q_{jl} of entering first union due to cause j in the l-th interval $[a_{l-1}, a_l)$, given first union has not been experienced before time a_{l-1} is, from (3.13),

$$q_{jl} = q_j(a_{l-1},1) = e^{\alpha_{jl}}.\ \frac{1 - e^{-\sum_j e^{\alpha_{jl}}}}{\sum_j e^{\alpha_{jl}}}\ ; \tag{3.42}$$

the net probability $q_{(j)l}$ is, from (3.22),

$$q_{(j)l} = q_{(j)}(a_{l-1},1) = 1 - e^{-e^{\alpha_{jl}}}\ ; \tag{3.43}$$

and the total probability q_l is, from (3.14) and (3.42),

$$q_l = q(a_{l-1},1) = 1 - e^{-\sum_j e^{\alpha_{jl}}} \tag{3.44}$$

(The proof of formulae (3.42) to (3.44) is given in Appendix E9.)

If we denote the j-th cause-specific hazard ($e^{\alpha_{jl}}$) for the l-th interval and the total hazard ($\sum_j e^{\alpha_{jl}}$) for the l-th interval respectively by μ_{jl} and μ_l, we can then rewrite (3.42-44) as

$$q_{jl} = \mu_{jl} \frac{1 - e^{-\mu_l}}{\mu_l} , \tag{3.42a}$$

$$q_{(j)l} = 1 - e^{-\mu_{jl}}, \tag{3.43a}$$

and $$q_l = 1 - e^{-\mu_l}. \tag{3.44a}$$

(Some approximate formulae are discussed in Appendix E10.)
From (3.42a-44a) we also get the relations

$$q_{jl} = \frac{\mu_{jl}}{\mu_l} \cdot q_l , \tag{3.45}$$

and $$q_{(j)l} = 1 - e^{-\frac{q_{jl}}{q_l} \cdot \mu_l} . \tag{3.46}$$

If p_l is the *total* probability to "survive" the l-th interval in the presence of all risks, given first union was not experienced before a_{l-1}, and if $p_{(j)l}$ is, similarly, the *net* (or *pure*) probability, then we have also the following formulae :

$$p_l = e^{-\mu_l} = 1 - q_l \tag{3.47}$$

$$p_{(j)l} = e^{-\mu_{jl}} = 1 - q_{(j)l} \tag{3.48}$$

$$q_{jl} = \frac{\mu_{jl}}{\mu_l} \cdot (1 - p_l) \tag{3.49}$$

$$q_{(j)l} = 1 - p_l^{(q_{jl}/q_l)} . \tag{3.50}$$

For the unconditional crude, net and total probabilities, defined in (3.7), (3.23) and (2.5) or (3.8), we can find the following commonly used formulae :

$$Q_{jl} = Q_j(a_l) = q_{j1} + p_1 . q_{j2} + \dots\dots + p_1 . \dots . p_{l-1} . q_{jl} \tag{3.51}$$

$$Q_{(j)l} = Q_{(j)}(a_l) = q_{(j)1} + p_{(j)1} . q_{(j)2} + \dots\dots + p_{(j)1} . \dots . p_{(j)l-1} . q_{(j)l} \tag{3.52}$$

$$Q_l = F(a_l) = q_1 + p_1 . q_2 + \dots\dots + p_1 . \dots . p_{l-1} . q_l . \tag{3.53}$$

The solution of the system of maximum likelihood equations, obtained by equating to zero the partial derivatives of log $\mathcal{L}$ in (3.39) with respect to the parameters $\alpha_{jl\mathbf{z}}$, yields the following estimate for the hazards in (3.40).

$$\tilde{\mu}_j(t;\mathbf{z}) = e^{\tilde{\alpha}_{jl\mathbf{z}}} = \frac{d_{jl\mathbf{z}}}{\tilde{E}_{l\mathbf{z}}} \quad \text{for } a_{l-1} \leqslant t < a_l . \tag{3.54}$$

Thus, under the piecewise exponential competing risks model, the j-th cause-specific hazard in the l-th interval, and for subgroup $\mathbf{z}$, is estimated by the *occurence-exposure rate* $d_{jl\mathbf{z}}/\tilde{E}_{l\mathbf{z}}$. It follows from (3.2) and (3.54) that the total hazard in the l-th interval, and for subgroup $\mathbf{z}$, is estimated by the occurence-exposure rate $\tilde{\mu}(t;\mathbf{z}) = \sum_j d_{jl\mathbf{z}}/\tilde{E}_{l\mathbf{z}} = d_{l\mathbf{z}}/\tilde{E}_{l\mathbf{z}}$, which is exactly the same occurence-exposure rate as obtained in Section 2.2 - formula (2.22).

Under the special model (3.30) we get - when covariates $\mathbf{z}$ are taken into consideration - the estimates

$$\hat{\mu}_j(t;\mathbf{z}) = \frac{d_{l\mathbf{z}}}{J.\tilde{E}_{l\mathbf{z}}} \qquad \text{for } a_{l-1} \leqslant t < a_l . \tag{3.55}$$

Note that the estimate of the total hazard $\mu(t;\mathbf{z})$ is the same under model (3.30) as under model (3.29), i.e. $\hat{\mu}(t;\mathbf{z}) = \tilde{\mu}(t;\mathbf{z})$.

We have not discussed the exponential model in the presence of competing risks in this section. In such a model the cause-specific hazards would be constant over the entire interval $[a_o, a_l)$. This model could be of special interest in a number of applications (e.g. if the entire interval $[a_o, a_l)$ is short). The appropriate formulae however can easily be obtained from the formulae given in this section.

3.3 Proportional hazards models in the presence of competing risks

In Section 2.3 we saw how the general (piecewise exponential) model for the analysis of age of entry into first union can substantially be simplified by assuming that the hazards are proportional (i.e. model (2.23)). Similar assumptions can be made under the competing risks model. In this section we will see how the ideas of proportional hazards and of stratification can be adopted in the competing risks model.

Under the competing risks model (Sections 3.1-2) we have to deal in general with the *total* hazards $\mu(t;z)$ and a number of *cause-specific* hazards $\mu_j(t;z)$. One or more of the following assumptions could be made in their regard.

(I) The ratio of the total hazards $\mu(t;z_1)$ and $\mu(t;z_2)$, for any two subgroups z_1 and z_2, is constant over the entire interval $[a_o,a_L)$: briefly, for any z_1 and z_2, $\mu(t;z_1)/\mu(t;z_2)$ does not depend on time t.

(II) The ratio of the cause-specific hazards $\mu_j(t;z_1)$ and $\mu_j(t;z_2)$, for any two subgroups z_1 and z_2, and for any specified cause j is constant over the entire interval $[a_o,a_L)$: briefly, for any z_1, z_2 and j, $\mu_j(t;z_1)/\mu_j(t;z_2)$ does not depend on time t.

(III) The ratio of the cause-specific hazards $\mu_{j_1}(t;z)$ and $\mu_{j_2}(t;z)$, for any two causes j_1 and j_2, and for each subgroup z, is constant over the entire interval $[a_o,a_L)$: briefly, for any j_1, j_2 and z, $\mu_{j_1}(t;z)/\mu_{j_2}(t;z)$ does not depend on time t.

As in Section 2.3 we can once again consider a *reference subgroup* z_o. In the same strain, we can also speak of a *reference cause* j_o. The assumptions I, II and III may then be formalized as follows.

(I) $\mu(t;z) = \mu(t;z_o).e^{\beta_z}$ for all z.

(II) $\mu_j(t;z) = \mu_j(t;z_o).e^{\tau_{jz}}$ for all z and j.

(III) $\mu_j(t;z) = \mu_{j_o}(t;z).e^{\gamma_{jz}}$ for all z and j.

Note that for the reference subgroup z_o the parameters β_{z_o} and τ_{jz_o} (j=1,...J) are zero, and that for the reference cause j_o the parameters γ_{j_oz} $(z\varepsilon\mathcal{Z})$ are zero.

Assumption I is clearly equivalent to the PH model (2.23) - the only difference is that the explicit reference to the piecewise constancy of hazards found in (2.23) is absent in the formulation of I.

Under assumption I an arbitrary subgroup z can be compared with the reference subgroup z_o using the single *relative risk* exp (β_z). Under assumption II an arbitrary subgroup z can be compared with the reference subgroup z_o using anyone of the relative risks exp (τ_{jz}), $j=1,\ldots J$: these relative risks exp (τ_{jz}) will be referred to as the *cause-specific relative risks for subgroup* z. Assumption III implies that the relative risk exp (γ_{jz}) can be used for the comparison - in subgroup z - of an arbitrary cause j with the reference cause j_o : this relative risk exp (γ_{jz}) will be referred to as the *subgroup-specific relative risk for cause* j.

The idea of a reference subgroup has been found to be convenient in PH models. The idea of a reference cause may be less convenient, and can, if necessary, be avoided as follows. We have

$$\mu(t;z) = \sum_{j'} \mu_{j'}(t;z) \qquad \text{(by (3.2))}$$

$$= (\sum_{j'} e^{\gamma_{j'z}}) . \ \mu_{j_o}(t;z) \qquad \text{(by III)},$$

whence

$$\mu_{j_o}(t;z) = \frac{1}{\sum_{j'} e^{\gamma_{j'z}}} \ \mu(t;z),$$

and

$$\mu_j(t;z) = \frac{e^{\gamma_{jz}}}{\sum_{j'} e^{\gamma_{j'z}}} \ \mu(t;z) \qquad \text{(by III)}.$$

If we define parameters θ_{jz} to be equal to $e^{\gamma_{jz}}/(\sum_{j'} e^{\gamma_{j'z}})$, then an alternative formalization of assumption III is :

$$\text{(III')} \quad \mu_j(t;z) = \mu(t;z).\theta_{jz} \qquad \text{for all } z \text{ and } j.$$

This implies that, in a specified subgroup z, an arbitrary cause j

can be compared with the totality of causes using the parameter θ_{jz}. The parameters θ_{jz} (which are in substance relative risks) satisfy the relation $\sum_j \theta_{jz}=1$: they will be referred to as the *subgroup-specific weights*.

There is, in general, no specific relation between any two of the assumptions I, II and III. For instance, proportionality of total hazards (i.e. I) does not necessarily imply proportionality of cause-specific hazards (i.e. II), and the converse is also not true in general. However, we can demonstrate the following interesting properties.

(A) If II and III are satisfied simultaneously, then I is satisfied and $\exp(\beta_z)$ is the weighted average $\sum_j \theta_{jz_o} . \exp(\tau_{jz})$ of the cause-specific relative risks $\exp(\tau_{jz})$ $(j=1,\ldots J)$.

(B) If II is satisfied, and if the cause-specific relative risks $\exp(\tau_{jz})$ for subgroup z do not depend on cause j - i.e. $\tau_{jz}=\tau_z$ for all $j=1,\ldots J$ - then I is satisfied and $\exp(\beta_z)$ is equal to $\exp(\tau_z)$.

Properties A and B are proved in Appendix E11. In Section 2 we discussed PH models and used them in connection with total hazards, i.e. we dealt with models for which assumption I holds.[18] We prefer to work here too - i.e. in the presence of competing risks - with models for which I holds. Properties A and B provide us with two conditions related to cause-specific hazards which give rise to proportionality of total hazards. It is therefore useful to discuss them in detail.

If II and III are satisfied simultaneously, we can write :

$$\mu_j(t;z) = \mu_{j_o}(t;z_o).e^{\tau_{jz}+\gamma_{jz_o}},$$

or

$$\mu_j(t;z) = \mu_{j_o}(t;z_o).e^{\gamma_{jz}+\tau_{j_o z}},$$

or

$$\mu_j(t;z) = \mu_{j_o}(t;z_o).e^{\lambda_{jz}} \qquad (3.56)$$

where $\lambda_{jz} = \tau_{jz} + \gamma_{jz_o} = \gamma_{jz} + \tau_{j_o z}$.

Under the assumption of piecewise constant hazards (Section 3.2), equation (3.56) can be written as

$$\mu_j(t;z) = e^{\alpha_l}.e^{\lambda_{jz}} \qquad \text{for } a_{l-1} \leqslant t < a_l. \tag{3.57}$$

where $\exp(\alpha_l)$ $(l=1,\ldots.L)$ is a series of hazards for cause j_o in subgroup z_o. This series may be called a base-line hazard. Equation (3.57) is a general formula for the model incorporating the piecewise exponential competing risks model and the assumptions II and III. Extending the terminology used in Section 2.3, we can say that there is (1°) no interaction between time t and cause j, (2°) no interaction between time t and covariates z, but that there is (3°) interaction between cause j and covariates z. The parameter $\exp(\lambda_{jz})$ in (3.57) is therefore a relative risk which can be used for the comparison of cause j in subgroup z with the reference cause j_o in the reference subgroup z_o. Such comparisons however do not seem to be of much use in practice : the investigator will not be particularly interested in the relative risks $\exp(\lambda_{jz})$ as such, but in the cause-specific relative risks $\exp(\tau_{jz})$ for subgroup z, and in the subgroup-specific relative risks $\exp(\gamma_{jz})$ for cause j. Fortunately, it is easy to compute both the cause-specific and the subgroup-specific relative risks from the relative risks $\exp(\lambda_{jz})$. Indeed, we have :

$$e^{\tau_{jz}} = e^{\lambda_{jz}}/e^{\lambda_{jz_o}} \tag{3.58a}$$

and

$$e^{\gamma_{jz}} = e^{\lambda_{jz}}/e^{\lambda_{j_oz}}. \tag{3.58b}$$

Further, the following expressions for θ_{jz} and $\exp(\beta_z)$ are also easily found :

$$\theta_{jz} = e^{\lambda_{jz}}/(\sum_{j'} e^{\lambda_{j'z}}), \tag{3.58c}$$

and

$$e^{\beta_z} = (\sum_j e^{\lambda_{jz}})/(\sum_j e^{\lambda_{jz_o}}). \tag{3.58d}$$

If II holds with $\tau_{jz} = \tau_z$, then we have the equation

$$\mu_j(t;z) = \mu_j(t;z_o).e^{\tau_z}. \tag{3.59}$$

Under the piecewise exponential model, this leads to the model formula

$$\mu_j(t;z) = e^{\alpha_{jl}}.e^{\tau_z} \qquad \text{for } a_{l-1} \leqslant t < a_l, \tag{3.60}$$

where exp (α_{jl}) $(l=1,\ldots.L)$ is a series of hazards for cause j in the reference subgroup z_o. The series exp (α_{jl}) $(l=1,\ldots L)$ will be called the *j-th cause-specific base-line hazard.* Here too we can say that there is (1°) no interaction between time t and covariates z, (2°) no interaction between cause j and covariates z, but that there is (3°) interaction between time t and cause j.

So far we have discussed two classes of models : i.e. those corresponding to the (general) model formulae (3.57) and (3.60). The intersection of these two classes is however not empty. Models belonging simultaneously to each of the two classes (3.57) and (3.60) are formalized through equation (3.61).

$$\mu_j(t;z) = e^{\alpha_l}.e^{\gamma_j}.e^{\tau_z} \qquad \text{for } a_{l-1} \leqslant t < a_l. \tag{3.61}$$

In other words, under model (3.61), assumptions II and III are satisfied, with $\tau_{jz} = \tau_z$.

Other classes of models are obtained by assuming that either II, or III, or III with $\gamma_{jz} = \gamma_j$, or any combination of these possibilities holds. Formally, we then get the following classes of models (using the assumptions of piecewise constant hazards throughout).

- If only III holds, we have

$$\mu_j(t;z) = e^{\alpha_{lz}}.e^{\gamma_{jz}} \qquad \text{for } a_{l-1} \leqslant t < a_l \tag{3.62}$$

  where exp (α_{lz}) $(l=1,\ldots L)$ is a series of hazards corresponding to the reference cause j_o in subgroup z - it will be called a *subgroup-specific base-line hazard.*

- If III holds with $\gamma_{jz} = \gamma_j$, we get the following class of models - i.e. a subclass of the class represented by (3.62) -

$$\mu_j(t;z) = e^{\alpha_{lz}}.e^{\gamma_j} \qquad \text{for } a_{l-1} \leqslant t < a_l. \tag{3.63}$$

  In this class of models the relative differences between causes is the same across subgroups, and they are measured in each subgroup by the same set of relative risks exp (γ_j) $(j=1,\dots J)$. Note that under model (3.63) the weights $\theta_{j\mathbf{z}}$ are also independent of covariates $\mathbf{z}$.

- If only assumption II holds, then we get the class of models represented by the formula

$$\mu_j(t;\mathbf{z}) = e^{\alpha_{jl}}.e^{\tau_{j\mathbf{z}}} \qquad \text{for } a_{l-1} \leqslant t < a_l \qquad (3.64)$$

where exp (α_{jl}) $(l=1,\dots.L)$ is a series of hazards corresponding to the j-th cause in the reference subgroup $\mathbf{z}_o$ - as in model (3.60) it will be called the *j-th cause-specific base-line hazard*. Note that (3.60) is a subclass of class (3.64).

The intersection of classes (3.57) and (3.63) is again class (3.61). It follows then that under assumptions II and III the additional assumptions $\tau_{j\mathbf{z}} = \tau_\mathbf{z}$ and $\gamma_{j\mathbf{z}} = \gamma_j$ (or $\theta_{j\mathbf{z}} = \theta_j$) are equivalent. These special assumptions are also equivalent to the statement $\lambda_{j\mathbf{z}} = \tau_\mathbf{z} + \gamma_j$ (see e.g. (3.58)) - if II and III hold.

It is useful to note that the different models discussed above are all piecewise exponential competing risks models, which are generally represented by the model formula (3.40). A schematic representation of the general class (3.40), its subclasses, and how the latter are obtained from the former is found in Figure C4.

The idea of stratification in relation to proportional hazards models (in the absence of competing risks) was introduced in Section 2.3. From the discussion in that section it follows that a stratified proportional hazards (SPH) model can be conceived as one which implies an ordinary proportional hazards (PH) model in each stratum. The same idea can be adopted here too to suite the presence of competing risks. The formal representation of SPH models in the presence of competing risks however tends to become very complicated since the expression "proportional hazards" could now stand for a number of different model assumptions - i.e. assumption I, or II, or III, either separately or in combination, could be taken as being satisfied. Further the concept of stratification presented earlier in this text - i.e. as accounting for the interaction between time and any specified subset of covariates - should not be used in conjunction with PH models in which only assumption III (i.e. proportionality with respect to the causes only) is satisfied. Hence, SPH models in the presence of competing risks will only be defined as modifications of the PH models (3.57), (3.60), (3.61) and (3.64) obtained by reintroducing interaction between time and specified covariates.[19]

As in Section 2.3 we will now split the covariate vector $\mathbf{z}$ into two parts $\mathbf{z}_1$ and $\mathbf{z}_2$ (i.e. $\mathbf{z} = (\mathbf{z}_1,\mathbf{z}_2)$) where $\mathbf{z}_1 = (z_1,\ldots z_{m_1})$ is the vector of m_1 *stratifying covariates* and $\mathbf{z}_2 = (z_{m_1+1},\ldots z_m)$ the vector of the $m-m_1$ *remaining covariates*. If the class of PH models defined by (3.57) is considered, both assumptions II and III (and hence also I) are seen to hold. The corresponding SPH models are therefore defined by the equation

$$\mu_j(t;\mathbf{z}) = e^{\alpha_{l\mathbf{z}_1}+\lambda_{j\mathbf{z}}} \qquad \text{for } a_{l-1} \leqslant t < a_l, \tag{3.65}$$

and it follows that assumptions II and III (and hence I) hold <u>in each stratum</u> $\mathbf{z}_1$. Note that (3.65) can be obtained from (3.57) by reintroducing interaction between time t and stratifying covariates $z_1,\ldots z_{m_1}$. Similarly, we obtain

$$\mu_j(t;\mathbf{z}) = e^{\alpha_{jl\mathbf{z}_1}+\tau_{j\mathbf{z}}} \qquad \text{for } a_{l-1} \leqslant t < a_l \tag{3.66}$$

from (3.64);

$$\mu_j(t;\mathbf{z}) = e^{\alpha_{jl\mathbf{z}_1}+\tau_{\mathbf{z}}} \qquad \text{for } a_{l-1} \leqslant t < a_l \tag{3.67}$$

from (3.60); and

$$\mu_j(t;\mathbf{z}) = e^{\alpha_{l\mathbf{z}_1}+\gamma_j+\tau_{\mathbf{z}}} \qquad \text{for } a_{l-1} \leqslant t < a_l \tag{3.68}$$

from (3.61). Note that there are a variety of different links, both between the SPH models (3.65), (3.66), (3.67) and (3.68) themselves, and between these SPH models and the PH models discussed above. For instance, (3.67) is a subclass of (3.66) obtained by assuming that $\tau_{j\mathbf{z}} = \tau_{\mathbf{z}} (j=1,\ldots J)$.

Similarly, (3.68) is a subclass of (3.65) resulting from the assumption that $\lambda_{jz} = \gamma_j + \tau_z$. Figure C5 - an extension of Figure C4 - summarizes these links and shows which assumption(s) is (are) needed to get one class from another.

The use of relative risks for the comparison of different subgroups in SPH models is equivalent to such comparisons in PH models, with the added proviso that the two subgroups to be compared should necessarily be in the same stratum. A complete overview of the possible relative risks for the classes of models presented in Figure C5 is given in Table A8.

The comparison of different strata in stratified models - e.g. through the comparison of the reference subgroups in different strata - is, in general, not straightforward, since the difference is not measured by a single (or just a few) parameter(s). However, we have already defined (in Section 2.3) a parametrized form of stratification as the combination of (1°) a difference between the starting points of the process in different strata, and (2°) a proportionality between the hazards corresponding to different strata once the process has started. The same ideas can be adopted in the presence of competing risks. Hence, proceeding as in Section 2.3, the SPH models (3.66), (3.67), (3.65) and (3.68) can be parametrized respectively as follows - ignoring the assumption of piecewise constant hazards (if as in Section 2.3 $\mathbf{z}$ is a subgroup of stratum $\mathbf{z}_1$ and $\mathbf{z}_{oo}$ is the reference subgroup in the reference stratum $\mathbf{z}_{10}$) :

$$\mu_j(t;\mathbf{z}) = \mu_j(t+b_{\mathbf{z}_1}; \mathbf{z}_{oo}).e^{\tau'_{j\mathbf{z}}} \tag{3.69}$$

with :

$$\tau'_{j\mathbf{z}} = \omega_{j\mathbf{z}_1} + \tau_{j\mathbf{z}} \tag{3.70a}$$

for a parametrization of (3.66),

or

$$\tau'_{j\mathbf{z}} = \omega_{j\mathbf{z}_1} + \tau_{\mathbf{z}} \tag{3.70b}$$

for a parametrization of (3.67);

$$\mu_j(t;z) = \mu_{j_o}(t+b_{z_1};z_{oo}).e^{\lambda'_{jz}} \qquad (3.71)$$

with :

$$\lambda'_{jz} = \omega_{j_o z_1} + \lambda_{jz} \qquad (3.72a)$$

for a parametrization of (3.65),

or,

$$\lambda'_{jz} = \omega_{j_o z_1} + \gamma_j + \tau_z \qquad (3.72b)$$

for a parametrization of (3.68).

If z_o is the reference subgroup in stratum z_1, then we have the following equations :

$$\mu_j(t;z_o) = \mu_j(t+b_{z_1};z_{oo}).e^{\omega_{jz_1}} \qquad (3.73)$$

under model (3.69-70a) or (3.69-70b), and

$$\mu_{j_o}(t;z_o) = \mu_{j_o}(t+b_{z_1};z_{oo}).e^{\omega_{j_o z_1}} \qquad (3.74)$$

under model (3.71-72a) or (3.71-72b). These formulae are useful for the interpretation of the parameters of the model. For instance, the parameters under model (3.71-3.72a) - i.e. the parametrization of (3.65) - are used as follows :

(1°) the shift parameter b_{z_1} indicates that the entrance into first union of women in stratum z_1 starts b_{z_1} units ahead of that of women in stratum z_{10};

(2°) exp $(\omega_{j_o z_1})$ measures the remaining difference - i.e. after the shift - between the entrance into first union due to cause j_o in the reference subgroups z_o (in stratum z_1) and z_{oo} (in stratum z_{10});

(3°) exp (λ_{jz}) measures the difference between the entrance into first union due to cause j in subgroup z (in stratum z_1) with the entrance into first union due to the reference cause j_o in the reference subgroup z_o (in stratum z_1).

As before relative risks exp (λ_{jz}) are taken to be of very limited use. In practice, these parameters may be replaced by the cause-specific relative risks exp (τ_{jz}) and the subgroup-specific relative risks exp (γ_{jz}), which are defined in (3.58a) and (3.58b) respectively. Note however that the relative risks exp $(\omega_{j_o z_1})$, exp (τ_{jz}) and exp (γ_{jz}) - in which the investegator will be interested - can all be derived from the relative risks exp (λ'_{jz}) by the formulae :

$$e^{\omega_{j_o z_1}} = e^{\lambda'_{j_o z_o}}, \tag{3.75a}$$

$$e^{\tau_{jz}} = e^{\lambda'_{jz}}/e^{\lambda'_{jz_o}}, \tag{3.75b}$$

$$e^{\gamma_{jz}} = e^{\lambda'_{jz}}/e^{\lambda'_{j_o z}}. \tag{3.75c}$$

We can also find the following expressions :

$$\theta_{jz} = e^{\lambda'_{jz}}/(\sum_{j'} e^{\lambda'_{j'z}}), \tag{3.57d}$$

$$e^{\beta_z} = (\sum_j e^{\lambda'_{jz}})/(\sum_j e^{\lambda'_{jz_o}}). \tag{3.75e}$$

Formulae (3.75a-e) are also valid under the model (3.71-72b), but here τ_{jz} does not depend on j, while γ_{jz} and θ_{jz} do not depend on z (Note also that γ_{jz} and θ_{jz} do not depend on the stratum z_1 of which z is a subgroup). Similar formulae can also be found under model (3.69-70a) or (3.69-70b). A complete overview of the shift parameters and relative risks (if they exist) under the four parametrized SPH models is given in Table A9.

The log-likelihoods under any of the models discussed in this section are obtained from the general formula (3.39) if the parameters α_{jlz} are replaced by the appropriate expression. As in Section 2.3, the following remarks concerning the estimation of the parameters of the models are in order. Firstly, the system of maximum likelihood equations has, in general, to be solved iteratively. Secondly, the shift parameters cannot be estimated by this method; the remaining parameters can be estimated only if the shift parameters are fixed.

Finally, we give some useful relations between total, cause-specific and pseudo-functions, and between total, crude and net probabilities. All these relations can easily be found from the appropriate relation for the hazard functions and from the formulae in Sections 3.1-2.

Under assumption III, we get the following relations between cause-specific and total functions (or between crude and total probabilities) :

$$f_j(t;z) = \theta_{jz}.f(t;z), \quad (3.76a)$$

$$Q_j(t;z) = \theta_{jz}.F(t;z), \quad (3.76b)$$

$$c_j(z) = \theta_{jz}.c(z), \quad (3.76c)$$

$$q_j(t,h;z) = \theta_{jz}.q(t,h;z); \quad (3.76d)$$

and between pseudo- and total functions (or between net and total probabilities) :

$$\Lambda_j(t;z) = \theta_{jz}.\Lambda(t;z), \quad (3.77a)$$

$$G_j(t;z) = (S(t;z))^{\theta_{jz}} \quad (3.77b)$$

$$q_{(j)}(t,h;z) = 1 - (1 - q(t,h;z))^{\theta_{jz}} \quad (3.77c)$$

Under assumption II, we get (as in ordinary (S)PH models - see Section 2.3) :

$$G_j(t;z) = (G_j(t;z_o))^{e^{\tau_j z}} \tag{3.78a}$$

$$\Lambda_j(t;z) = \Lambda_j(t;z_o).e^{\tau_j z}, \tag{3.78b}$$

$$q_{(j)}(t,h;z) = 1 - (1 - q_{(j)}(t,h;z_o))^{e^{\tau_j z}} \tag{3.78c}$$

If both II and III hold, then we get for instance :

$$G_j(t;z) = (S(t;z_o))^{\theta_{jz_o}.e^{\tau_j z}} \tag{3.79a}$$

$$\Lambda_j(t;z) = \theta_{jz_o}.e^{\tau_j z}.\Lambda(t;z_o), \tag{3.79b}$$

$$q_{(j)}(t,h;z) = 1 - (1 - q(t,h;z_o))^{\theta_{jz_o}.e^{\tau_j z}} \tag{3.79c}$$

The appropriate formulae for the parametrized SPH models can easily be obtained from the above formulae (cfr. Section 2.3).

3.4 Estimation of the parameters in competing risks models through GLIM

In this section we will discuss how the (log-) likelihoods under various competing risks models can be maximized through the GLIM3 computer package. As in Section 2.4 the following two remarks should be kept in mind : proper estimation of the shift parameters is not possible, and the log-likelihood of any competing risks model is a special case of the log-likelihood in equation (3.39). Our attention will therefore be focussed on the general likelihood (3.39).

As in Section 2.4, a GLM with a log-likelihood differing from $\log \mathcal{L}$ in (3.39) at most by a constant term needs to be constructed. The appropriate GLM is defined as follows :

(1°) the dependent variables (i.e. the counts d_{jlz} $(j=1,\ldots J; z\varepsilon\mathfrak{Z}; l=1,\ldots L)$) are assumed to be statistically independent and Poisson distributed with means M_{jlz} say (The counts d_{jlz} thus have a Poisson error structure);

(2°) the covariates consist of a time covariate - with levels $l=1,\ldots L$ -, the covariates $\mathbf{z}$, and a cause covariate - with levels $j=1,\ldots J$. The linear predictor is denoted by α_{jlz};

(3°) The link between the linear predictor α_{jlz} and the mean M_{jlz} of the dependent variable d_{jlz} is given by :

$$\log M_{jlz} = \log \tilde{E}_{lz} + \alpha_{jlz}, \tag{3.80}$$

where $\tilde{E}_{lz}$ is the corresponding (approximate) exposure time.

The likelihood under this GLM is proportional to

$$\overset{=}{\mathcal{L}} = \prod_j \prod_z \prod_l M_{jlz}^{d_{jlz}} . e^{-M_{jlz}} \tag{3.81}$$

$$= \prod_j \prod_z \prod_l \tilde{E}_{lz}^{d_{jlz}} . e^{d_{jlz}.\alpha_{jlz}} . e^{-\tilde{E}_{lz}.e^{\alpha_{jlz}}} . \tag{3.82}$$

Taking logarithms we get

$$\log \overline{\overline{\mathcal{L}}} = \sum_j \sum_z \sum_l (d_{jlz} \log \widetilde{E}_{lz} + d_{jlz}.\alpha_{jlz} - \widetilde{E}_{lz}.e^{\alpha_{jlz}}) \qquad (3.83)$$

which, apart from a constant term, is equal to log $\mathcal{L}$ in (3.39). Hence, estimates of the parameters α_{jlz} in competing risks models can be found by fitting the above GLM. In other words, we can still use GLIM3 in order to fit the competing risks models developped in Sections 3.2-3.

A GLIM3-programme for fitting competing risks models which specify cause of entry into first union has the same basic structure as a GLIM3-programme for fitting models in which no reference is made to cause of entry. However, there are some important modifications. We shall discuss these now, using the GLIM3-programme in Appendix B5.

To start with, note that the GLIM3-programme in Appendix B5 produces a fit for a parametrized SPH model of the form (3.71-72a). As mentionned before, however, shift parameters cannot be properly estimated. They are therefore fixed in advance, and the time variable (T) is adjusted for the shifts. Leaving aside this difficulty related to shifts, which can now be ignored in what follows, the model takes the form (3.57).

We have already argued in Section 3.2 that cause of entering first union (i.e. j) can be treated in the same way as the covariates $z_1,...z_m$: see the derivation of formula (3.41) from (3.39). There are 2 causes to be dealt with in the applications found in this text. They are taken count of through a dichotomous covariate TYPE which has been defined in the data definition part of the programme. Since each *unit* in GLIM is a particular combination of the covariates j, z and l, the total number of units is now doubled (i.e. from 402 to 804). The data are stored initially in vectors BT, N, D1, D2, W, REL, EDU and COH - which are all of length 402. However, since D1 and D2 correspond to different units (or j z l combinations), the initially formed vectors BT, N, D1, D2, W, REL EDU and COH are transformed into vectors T, LE, D, TYPE, ZR, ZE and ZC - which are all of length 804. The latter vectors can be described as follows :

T = the number corresponding to a time (or age) interval;

TYPE = the cause of entering first union;

ZR = the religious category;

ZE = the education category;

ZC = the birth-cohort;

D = the number of women with covariates ZR, ZE and ZC who enter first union due to cause TYPE in the T-th interval;

LE = the log-exposure time for the T-th interval and for the women with covariates ZR, ZE and ZC.

Note that LE does not depend on TYPE. This is as it should be since the exposure times $\tilde{E}_{lz}$ in (3.39) do not depend on cause j.

At the end of the data definition part, all vectors not used in the rest of the programme are deleted. This is of practical importance since the data space used should be reduced as often as possible because of computer memory space limitations.

The rest of the GLIM3-programme, except the $FIT-statement, is exactly the same as before (cf. Section 2.4). The specification of the linear predictor - through the $FIT-statement - is as follows. For competing risks models of the form (3.57),

$$\log \mu_j(t;z) = \alpha_l + \lambda_{jz} \qquad \text{for } a_{l-1} \leqslant t < a_l,$$

which shows incidentally that the $FIT-statement under models (3.57) has the general form

$FIT T + Z*TYPE

(Z standing for terms depending on covariates ZR, ZE and/or ZC).
The $FIT-statement for other competing risks models is found in the same way. The general forms of the $FIT-statement under the various classes of competing risks models discussed in Section 3.3 are listed in Table A10.

It is interesting to note that models which do not take the cause of entering first union into account can be fitted as special competing risks models if the covariate TYPE were omitted from the $FIT-statement. This allows for scaled deviance tests concerning the difference between the process of entry into first marriage and the process of entry into first cohabitation. For instance, the parametrized SPH model fitted through the GLIM3-programme in Appendix B3 can be effected through the GLIM3-programme in Appendix B5 if the $FIT-statement is changed into $FIT T + ZR + ZE + ZC. The difference between the scaled deviances corresponding to $FIT T + ZR + ZE + ZC and $FIT T + TYPE*(ZR + ZE + ZC) - both related to the programme in Appendix B5 - provides a chi-squared statistic which can be used to test whether the process of entering first marriage is (significantly) different from the process of entering first cohabitation. (See Appendix E7.) The scaled deviance obtained through the programme in Appendix B3 cannot be used for such a test.

Similarly, the scaled deviances can be used to test the PH assumption III (Section 3.3). For instance, the difference between the scaled deviances corresponding to the fits

$FIT T + TYPE*(ZR + ZE + ZC) and

$FIT T*TYPE + TYPE*(ZR + ZE + ZC)

provides a chi-squared statistic for testing the assumption III introduced earlier in Section 3.3.[20]

As in Section 2.4, it is very important to select a <u>parsimoneous</u> competing risks model. The same questions raised in Section 2.4 are relevant here too given that the cause covariate TYPE can be treated just like the other covariates. The strategy to be followed here too is similar to that followed in Section 2.4. However since the class of possible models is large it is best to follow relevant indications resulting from a prior cause-independent analysis of entry into first union. In other words, it may happen that stratification (with respect to covariates) and its parametrization is most conveniently effected as in the analysis of first union as such. Thus, the model used for the parsimoneous analysis of first union as such is at least a good starting point for the selection of a parsimoneous competing risks model. This procedure has been used in the operations described in the next section.

3.5 Numerical application

This section presents the results obtained from the application of a model of the type discussed in Sections 3.1-4. A detailed discussion of both the selection of the model and the results obtained from fitting it to the first marriage/first cohabitation data is not given here. The reader is referred to Willems and Vanderhoeft (1985) for more details. The main purpose of the following paragraphs is to indicate how the investigator should handle the basic results in order to transform these into a more usable form.

The model finally used in the analysis is specified by the GLIM3 expression for its linear predictor : T^* + (ZR + ZE + ZC)*TYPE. The notation T^* refers to the shifts of the time variable l for one stratum relative to an other : the (four) strata and the corresponding shifts are the same as those used in the application of the relevant models to data on first union as such (Section 2.5). In fact, the present model T^* + (ZR + ZE + ZC)*TYPE is closely related to the model T^* + REL + EDU + COH for analysis of first union as such : (1°) the strata and the shifts are exactly the same, and (2°) the action of the covariates REL, EDU and COH (here : ZR, ZE and ZC respectively) on the cause-specific hazards is analogously the same as there action on the total hazard of entry into first union - i.e. the covariates do not interact in their effects on the cause-specific hazards. The covariate TYPE - to be introduced in the competing risks model (Section 3.4) so as to distinguish the two causes of entry into first union - could be incorporated in the model in several ways. The resulting competing risks models are denoted as follows :

T^* + (ZR + ZE + ZC) + TYPE,

T^**TYPE + (ZR + ZE + ZC),

T^**TYPE + (ZR + ZE + ZC)*TYPE,

and the model mentionned above :

T^* + (ZR + ZE + ZC)*TYPE.

Each of these models can be regarded as being representative of one of the classes of SPH models shown (in boxes) in Figure C5 : i.e. the four models above belong to the classes (3.68), (3.67), (3.66) and (3.65) respectively. An analysis of deviance gives a first idea about the most suitable class of (parametric) SPH models. A model belonging to the class (3.65) was found to be the best in this regard. The final choice thus falls on the above mentionned model $T^* + (ZR + ZE + ZC)*TYPE$ (which can be described by the formulae (3.71-72a)). Note that the scaled deviance for this model is 680.47 with 747 degrees of freedom, indicating a fairly good fit.

The basic results - i.e. the linear parameters as they are estimated by application of a GLIM3 programme - are shown in Appendix B5. Using the formulae of Section 3.3 - i.e. the formula used in connection with the model (3.71-72a) - these parameter estimates can easily be transformed into other parameters more suitable for discussion.

To start with, the relative risks exp $(\hat{\lambda}'_{jz})$ were computed[21] and are presented in Table A11. Since these relative risks compare an arbitrary cause j in an arbitrary subgroup z with the reference cause j_o (here : j_o=1 ~ first marriage) in the reference subgroup z_o (here z_o=(1,1,1) ~ RC RMA - PRI - 48-62), after adjusting for the difference in the starting points of the corresponding processes, it is not easy to use them in a discussion of the differences found between first marriage and first cohabitation or found between subgroups. It has however been argued earlier that they can be transformed into subgroup-specific relative risks (for the comparison of causes) and cause-specific relative risks (for the comparison of subgroups) - cf. Section 3.3.

The subgroup-specific relative risks exp $(\hat{\gamma}_{jz})$ are shown in Table A12. They measure the difference between first cohabitation and first marriage in any specified subgroup z. How the covariates under consideration act on the relative difference between first marriage and first cohabitation can thus be seen. Alternatively, the subgroup-specific weights $\hat{\theta}_{jz}$ can be used. These quantities clearly show how the proportion of entry into first marriage (j=1) or first cohabitation (j=2) varies with covariates z.

The cause-specific relative risks exp $(\hat{\tau}_{jz})$ are shown in Table A13. Note that these quantities are in fact relative risks located within strata : they show how the process of entry into first marriage or the process of entry into first cohabitation taking place in a subgroup z located in a specified stratum differs from that occuring in the reference subgroup in the same stratum. (The four relevant strata are represented in Figure C3.) A table of within-stratum relative risks exp $(\hat{\beta}_z)$ (Table A14) can be constructed in similar fashion : they measure the difference between the process of entry into first union according to a subgroup z and the reference subgroup z_o in the same stratum.

The comparison of separate strata can be made on the basis of the parameters shown in Table A15, i.e. the shift parameters b_{z_1} and the relative risks exp $(\hat{\omega}_{j_o z_1})$. In fact, these parameters - particularly the relative risks exp $(\hat{\omega}_{j_o z_1})$ - enables one to compare the process of entry into first union due to marriage (i.e. the reference cause j_o) in different reference subgroups. (Note however that the shift parameters are valid for all subgroups.) In order to compare the process of entry into first union due to a specified cause, in two arbitrary subgroups in two separate strata, one should combine the relative risks exp $(\hat{\omega}_{j_o z_1})$ with the corresponding cause-specific and subgroup-specific relative risks.

The comparison of the process of entry into first union due to a specified cause in different subgroups - either in the same stratum or in separate strata - shows the effects of the covariates REL, EDU and COH. Since the model used here - i.e. T^* + (ZR + ZE + ZC)*TYPE - is additive in these covariates, their effects can be summarized as in Table A16. The relative risks in Table A16 measure the relative spread of the process of entry into first union due to either marriage or cohabitation, after adjustments for shifts have already been effected. Note that the relative risks corresponding to first marriage are different from those corresponding to first cohabitation, this being a consequence of the interaction between the cause covariate (TYPE) and the other covariates (ZR, ZE, ZC). A table representing the covariate effects on the total hazard, such as Table A5, could also be constructed under the model T^* + (ZR + ZE + ZC)*TYPE. The relative risks would be very close to the relative risks in Table A5. They are not exactly equal, since the model T^* + (ZR + ZE + ZC)*TYPE is more restrictive than the model T^* + ZR + ZE + ZC (or, in the notations of Section 2.5, T^* + REL + EDU + COH) : in the former model the assumptions I, II and III hold (in each stratum), whereas in the latter model only the assumption I holds (in each stratum).

We have also constructed a table of first deciles $P10_j(z)$ and medians $Me_j(z)$ (Table A17). These parameters can be defined as follows. The median $Me_j(z)$ (first decile $P10_j(z)$) is the age at which 50% (10%) of the women with covariates z who will ever experience first union due to cause j, if only cause j would exist, have already experienced this event. $Me_j(z)$, for instance, is the median age of entry into first union due to cause j if this phenomenon were to be observed in its pure form (Henry, 1959). Since the cause-specific hazard $\mu_j(t;z)$ is assumed to describe the phenomenon of entrance into first union due to cause j in its pure form, the estimation of $Me_j(z)$ and $P10_j(z)$ is only based on this hazard $\mu_j(t;z)$, and the mathematical procedure is equivalent to the problem outlined in Appendix E6. (In formula (E6.8) one has to replace $\hat{F}(a_L;z)$ by the pseudo c.d.f. $\hat{G}_j(a_L;z)$; $\hat{\Lambda}(a_{l-1};z)$ by the pseudo cumulative hazard $\hat{\Lambda}_j(a_{l-1};z)$ and $\hat{\mu}_{lz}$ by the cause-specific hazard $\hat{\mu}_{jlz}$.)

Finally, note that the description of the process of entry into first union due to either marriage or cohabitation is not as yet complete, since nothing has been said so far about the ultimate proportions of women who will ever make a cause-specified entry into first union. Estimates of these ultimate proportions were not given because of one of the data deficiencies already signalled in Section 1 - i.e. the exclusion of 469 never married women because of an error in the questionnaire. For the same reason, a table of conditional probabilities $\hat{q}_{jl}(z)$ (or $\hat{q}_{(j)l}(z)$) and of cumulative probabilities $\hat{Q}_j(a_l;z)$ (or $\hat{Q}_{(j)}(a_l;z)$) was also omitted. The estimated crude cumulative probabilities $\hat{Q}_j(a_l;z)$ have been plotted in Figure C6 together with the corresponding observed crude cumulative probabilities merely to show that the fit of the model $T^* + TYPE*(ZR + ZE + ZC)$ is fairly good.

4. CONCLUDING REMARKS

The PH model and other models closely related to it have been used in the work connected with the present text to study the entrance into first union - either due to marriage or to unmarried cohabitation - of Flemish women who were born between 1938 and 1962. It was found that ordinary PH models do not fit the data adequately. The reason for this unsatisfactory fit was seen to be linked to the postponement of entrance into first union in the case of women with certain characteristics.

The following socio-economic and demographic determinants were used in this study

(1°) religious affiliation (grouped into four categories, ranging from Roman Catholics with regular Mass attendance to freethinkers),

(2°) highest educational level attained (grouped into 3 categories : primary, secundary and higher educated women),

(3°) birth cohort (grouped into two broad categories : those born before or in 1947 and those born after 1947).

It was found that the starting point of the process of entry into first union is affected by both religious affiliation and educational level attained : Roman Catholics with regular Mass attendance tend to postpone entry into first union by about 1 year, and higher educated women tend to postpone it by about 2 years. If the duration variable - i.e. the time since the 15th birthday (see Section 1) is adjusted in view of these differences in starting points, then the effects of the above socio-economic and demographic determinants can further be accounted for through an ordinary PH model. In such a model, categories of women whose process of entry into first union can be seen as having different starting points are considered to form different strata. The model can then be described as a SPH model.

The ordinary PH model can be seen as a log-linear model for contingency tables - a result that is partially obtained by the piecewise exponential approach (Sections 2.2-3.2). Terms on the right hand side of the equation characterising such a model depend either on time or on the socio-economico-demographic determinants used, but not on both time and these determinants together. A SPH model is easily obtained by introducing terms depending on time and the other determinants simultaneously - i.e. by introducing interaction between time and these determinants. The general SPH model defined in this manner does not yield an estimation (through maximum likelihood techniques) of the amounts by which the process of entry into first union is shifted in moving from one stratum to another. However it does allow for the specification and inclusion of the optimum form of stratification suitable for the case in hand. Once an optimum stratification is found, the amounts by which the process has to be shifted in moving from one stratum to another can be estimated much more easily - for instance by visual inspection of observed and/or estimated schedules and comparison of them across strata.

The results of the analysis of first union as such have been put to further use in the analysis of cause specified first union. It was found that the particular stratification (and shifts) used in connection with first union as such was suitable for the cause-specific analysis too.

Technically, the cause of entering first union has been handled (by us) in the same way as the socio-economic and demographic determinants used. It was consequently easy to check whether or not the hazard of entry into first marriage and the hazard of entry into first (unmarried) cohabitation are proportional. It was found that they were in fact proportional. (If they were not, an examination of whether or not the process of first marriage starts later/earlier than the process of first (unmarried) cohabitation - depending perhaps on socio-economic and demographic characteristics - could be made.)

The methods used in this study are essentially related to multivariate multistate models. Extensions of the present methodology can readily be envisaged. Analyses of different kinds (e.g. analysis of entering into first union, analysis of termination of first union, analysis of remarriage, ...) which are usually done separately, could be done simultaneously : with consequent ease of comparison. Application of PH assumptions in such extended multistate analysis would advantageously allow for the comparison of different processes through the use of just a few parameters.

In more complex multistate analyses the importance of time dependent determinants will increase rapidly. The likelihood methods used in this study will then no longer be applicable as such. Suitable partial likelihood inference methods will consequently have to be dealt with. The GLIM3 oriented estimation techniques will then be unfortunately more difficult and less attractive for the user.

One of the important disadvantages of the methods presented in this paper consists of the fact that the range of shift suitable for any given process cannot be estimated by maximum likelihood techniques. This problem could be avoided by using parametric models instead of the semi-parametric models used in this text. Studies planned for the near future will examine how the semi-parametric methods used above could be replaced by others which are fully parametric. In this context note that such parametric PH methods would be very close to the Coale-McNeil nuptiality model (Coale and McNeil, 1972). Our own parametrized form of the SPH model is also, in point of fact, very close to the Coale-McNeil model : both these models carry a shift parameter (b in one case, a_o in the other), a speed parameter (relative risks as opposed to k) and the ultimate proportion (c in both models). Certain differences however need to be underlined. The Coale-McNeil model is (leaving aside the shift parameter a_o) an accelerated failure time model (Vanderhoeft, 1983). Hence, the interpretation of relative risks on the one hand and the speed parameter k on the other, is not the same : relative risks are multiplicative modifications of the hazard, whereas k is a multiplicative modification of the time variable itself.

Another disadvantage of the maximum likelihood methods used (in relation to general SPH models) is that the number of nuisance parameters (i.e. the base-line hazards needed) might become too large, entailing a consequent loss in efficiency (of the maximum likelihood estimates). This problem would perhaps be solved either by eliminating the nuisance parameters through partial likelihood techniques or by introducing suitable shift parameters. Note however that the strata to be used in both these cases should be known in advance.

Footnotes

(1) *Time-censoring* is also referred to as *Type I censoring*, see e.g. Lawless (1982). Because of this type of censoring, the observed individual data (1.2) - which is merely subject to *random censoring* (Lawless, 1982) - and the grouped data (1.4) are not equivalent, in the sense that (1.2) contains more information than (1.4). In other words, the likelihood for (1.2) and the likelihood for (1.4) with L=21 are not proportional. However, the difference is small since, as mentionned before, the number of women exposed to risk at exact age 36 is small compared to the total sample size. Any bias resulting from the time-censoring in question is therefore small and may be ignored in the steps leading from the likelihood for the individual data (1.2) to the likelihood for the time-censored grouped data (1.4) (with L=21) - see Section 2.1.

(2) The contribution to the likelihood of subgroups $\mathbf{z}$ for which no woman is observed is 1.

(3) We do not here specify the endpoint a_L of the L-th interval. This endpoint a_L can be either a time (or age) limit beyond which the event studied does not occur - as is customary in actuarial practice - or any other fixed (finite) timepoint. Moreover, the number of intervals L can be finite or infinite. Later on, we take both L and a_L, to be finite.

(4) To avoid technical problems in estimating the parameters α_l, we will work - in practice - with a finite number of intervals L (of unit length). (Otherwise, we would have an infinite number of parameters α_l (l=1,...+∞)). Similarly, if the last interval was $[a_{L-1},+\infty)$, then, with a constant hazard e^{α_L}, we would have $\Lambda(\infty)=\infty$ or $c=S(\infty)=1$, except if α_L is equal to $-\infty$. This too would cause computational problems. Note however that we don't need to specify L (or a_L) in theoretical discussions.

(5) For this and other assumptions see e.g. Holford (1976) and Menken et al (1981). Note that for the very fast interval (L=21) considered in the applications later on the censoring of the $w_{L\mathbf{z}}$ women is uniform

over the interval, but $n_{Lz}-w_{Lz}-d_{Lz}$ women are censored at exact time a_L; see Section 1 and footnote (1) for details.

(6) It may be shown that the shift parameter b_{z_1} does not depend on the choice of the reference subgroups. Whence the notation. The relative risks exp (ω_{z_1}), on the contrary, depend on the choice of the reference subgroups. However, we do not make reference to this in our notations.

(7) The parameters exp (β_z) will - whenever it is necessary - be referred to as *within-stratum relative risks*.

(8) GLIM3 is developed by the Numerical Algorithms Group (NAG). One should not read this text without having the GLIM3-manual (Baker and Nelder, 1978) at hand.

(9) From equation (2.38) we can derive the equation

$$\log \frac{M_{lz}}{\tilde{E}_{lz}} = \alpha_{lz}.$$

The expression "log-linear model in relation to rates" refers to the occurence-exposure rates $M_{lz}/\tilde{E}_{lz}$ and the corresponding linear predictors α_{lz} - i.e. the linear combinations of the covariates (or dummies representing these covariates) involved.

(10) Any shift parameter b_{z_1} used has in fact been measured as being equal to an integral number of years. In general the shifts could be measured in fractions of one year. Then, however, the length of the intervals $[a_{l-1}, a_l)$ would have to correspond to these fractions. Note that, in order to use GLIM3 and to fit parametrized SPH-models, (1°) all the intervals $[a_{l-1}, a_l)$ need to have the same length (taken as a unit and called a unit-length), and (2°) the shifts b_{z_1} should be measured in terms of this unit. This unit-length can thus be for instance 1 year, or 1 semester ($\frac{1}{2}$ year), or 1 month ($\frac{1}{12}$ year), etc.

(11) The models T' + REL + EDU + COH, T" + REL + EDU + COH and T"' + REL + EDU + COH are not *nested* (Baker and Nelder, 1978). They can therefore not be compared via the differences between their respective scaled deviances and corresponding degrees of freedom. Comparisons can however be made through their respective p-values, or through their respective *mean deviances* $\hat{\chi}^2_L/\nu$.

(12) Given that the ultimate proportion $c(z)$ is defined as follows

$$c(z) = \lim_{t\to\infty} F(t;z),$$

and since $F(t;z)$ increases with time t, the quantity $\hat{F}(a_L;z)$ with finite a_L - would normally underestimate the ultimate proportion $c(z)$. The data used however are not fully representative of the Flemish female population, since 469 never married women have in fact been excluded from the analysis. This causes $\hat{F}(a_L;z)$ to be an overestimate of the parameter $c(z)$: see the high values $\hat{F}(a_L;z)$ in Table A6.

(13) The terminology used by Lawless (1982) is adopted : thus terms like *cause-specific* hazard, p.d.f., ... and *pseudo* cumulative hazard,... will be used.

(14) *If only risk j is operative in* $[t,t+\Delta t)$ means that there is only a chance of entering the state of first union in $[t,t+\Delta t)$ due to cause j. Similarly, *if all risks are operative in* $[t,t+\Delta t)$ means that the chance of entering the state of first union in $[t,t+\Delta t)$ due to any specified cause could be influenced by the presence of other causes. Further, we use the terms *cause* and *risk* as in Chiang (1968, p. 243). I.e. the condition of first union is referred to as *cause* after the time of entering first union, but is called *risk* before the time of entering first union.

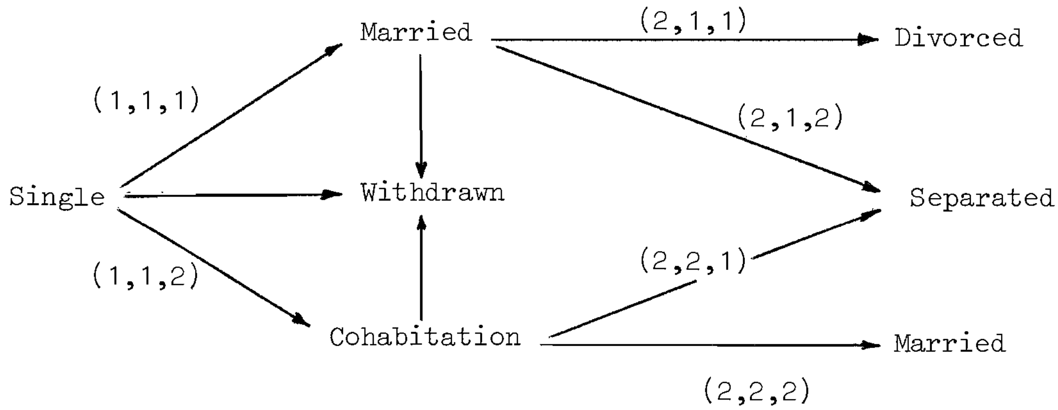
(15) This assumption is also adopted implicitly by Chiang (1968) where he defines a *net probability* $q_{i\delta}$ (p. 246), and by Pollard (1973) where he defines a *related single-decrement table* (p. 15, lines 4-5).

(16) In the terminology of Pollard (1973), $Q_j(t;z)$ is a *dependent* probability. Note that $f_j(t;z) = \frac{d}{dt} Q_j(t;z)$. Thus, we could call $Q_j(t;z)$ a *cause-specific c.d.f.*. However, we believe that the term *crude* (or *dependent*) is more suitable when reference is made to the presence of all risks.

(17) In the terminology of Pollard (1973), *net* probabilities would be called *independent* probabilities.

(18) The final model used in connection with data on age of entry into first union was a SPH model in which the stratification was parametrized through simple shifts. As explained earlier in the text, the general SPH model implies a PH model in each stratum (see Section 2.3). Moreover its parametrized version allows us to avoid a direct consideration of the shifts (or stratification) if the discussion were to focus on the relative risks. Therefore, we can discuss the proportionality of (total and/or cause-specific) hazards in the presence of competing risks in terms of the ordinary PH model.

(19) In more extended analyses stratification according to causes could be defined in a similar way. Consider for instance the following schema, representing the transitions to be dealt with in our analysis of the beginning and termination of first union.



The transitions could be coded by a three-dimensional vector $\underline{j} = (j_1, j_2, j_3)$ as indicated in the above schema. The "stratification" according to j_1 would then for example be equivalent to the assumption that the hazards corresponding to the transitions <u>into</u> first union are proportional and that likewise the hazards corresponding to the transitions out of first union are proportional. Further, the "stratification" according to j_1 and j_2 would be equivalent to the assumption that the hazards corresponding to the transitions into first union are proportional, that the hazards corresponding to the transitions out of first marriage are proportional, and that the hazards corresponding to the transitions out of first cohabitation are proportional. Note that the coding of the transitions depends on the stratifications to be examined. More complex schemas may be considered. In fact, any multidimensional schema may be treated by the same methods.

(20) This test is conditional on the stratification (according to covariates ZR and ZE) and its parametrized form, and on the additive structure used with the covariates ZR, ZE and ZC - i.e. ZR + ZE + ZC. An unconditional test concerning the effect of TYPE on the time parameters consists of the comparison of the scaled deviances corresponding to the fits

$FIT T*ZR*ZE*ZC + TYPE*ZR*ZE*ZC

and $FIT T*ZR*ZE*ZC*TYPE,

where T is not adjusted for a shift. Note that this is merely a goodness-of-fit test for the former model, since the latter denotes the saturated model.

(21) Each relative risk exp $(\hat{\lambda}_{jz})$ was treated separately in order to facilitate the construction of the GLIM3-macro PADE (Appendix D) which had to be used for the computation of the parameter estimate and its standard errors. Hence, no covariances were estimated for the relative risks exp $(\hat{\lambda}_{jz})$. The same procedure is followed for the other parameters discussed in the rest of Section 3.5.

REFERENCES

Baker, R.J., Nelder J.A. (1978) : The GLIM-system Release 3, Numerical Algorithms Group, Oxford.

Bloom D. (1980) : Age patterns of women at first birth, unpublished Ph.D.-thesis, Princeton University.

Breslow (1972) : Contribution to the discussion of the paper by D.R. Cox. J.R.S.S., Series B, 34, pp. 216-7.

Breslow, N. (1974) : Covariance analysis of censored survival data. Biometrics 30, pp. 89-100.

Burn, R. and Thompson, R. (1982) : A macro for calculating the covariance matrix of functions of parameter estimates; GLIM-Newsletter no. 5, pp. 26-29.

Chiang, C.L. (1968) : Introduction to stochastic processes in biostatistics, Wiley, New York.

Coale, A.J., McNeil D.R. (1972) : The distribution by age of the frequency of first marriage in a female cohort; JASA, Vol. 67, nr. 340, pp. 743-9.

Cox, D.R. (1972) : Regression models and life tables (with discussion); J.R.S.S., Series B, Vol. 34, pp. 187-220.

Cox, D.R. (1975) : Partial likelihood. Biometrika 62, pp. 269-276.

Gail, M. (1975) : A review and critique of some models used in competing risk analysis. Biometrics, 31, pp. 209-222.

Henry, L. (1959) : D'un problème fondamental de l'analyse démographique. Population, no. 1, pp. 9-32.

Holford, T.R. (1976) : Life tables with concomitant information. Biometrics 32, pp. 587-598.

Holford, T.R. (1980) : The analysis of rates and survivorship using log-linear models. Biometrics 36, pp. 299-306.

Holt, J.D. (1978) : Competing risks analyses with special reference to matched pair experiments. Biometrika, 65, 1, pp. 159-65.

Kalbfleisch, J.D., Prentice, R.L. (1980) : The statistical analysis of failure time data; Wiley, New York.

Laird, N., Olivier, D. (1981) : Covariance analysis of censored survival data using log-linear analysis techniques. JASA, Vol. 76, pp. 231-240.

Lawless, J.F. (1982) : Statistical models and methods for lifetime data. Wiley, New York.

Menken, J., Trussell, T.J., Stempel, D., Babakol, O. (1981) : Social and demographic influences on marriage dissolution in the U.S. : Applications of proportional hazards life table models. Demography, 18, nr. 2, pp. 181-200.

Nelder, J.A., Wedderburn, R.W. (1972) : Generalized linear models; J.R.S.S., Series A, Vol. 135, nr. 3, pp. 370-384.

O'Quigley J., Moreau T. (1984) : Testing the proportional hazards regression model against some general alternatives. Rev. Epidém. et Santé Publ., 32, pp. 185-98.

Pollard, J.H. (1973) : Mathematical models for growth of human populations, Cambridge University Press, Cambridge, England.

Rodríguez, G. (1984) : The analysis of birth intervals using proportional hazard models. WFS/TECH No. 2314, World Fertility Survey, London.

Rodríguez G., Trussell J. (1980) : Maximum likelihood estimation of the parameters of Coale's model nuptiality schedule from survey data. WFS/TECH No. 1261, World Fertility Survey, London.

Thompson, R. and Baker, R.J. (1981) : Composite link functions in generalized linear models; Appl. Statist., 30, no. 2, pp. 125-131.

Trussell, J., Hammerslough, C. (1983) : A hazards-model analysis of the covariates of infant and child mortality in Sri Lanka. Demography 20, pp. 1-26.

Vanderhoeft C. (1983) : A Unified approach to models for analysis of zero-one data with applications to intermediate fertility variables. IPD-Working Paper 83-5, Brussels.

Whitehead, J. (1980) : Fitting Cox's Regression model to survival data using GLIM. Appl. Statist., 29, 3, pp. 268-275.

Willems, P., Vanderhoeft, C. (1985) : Alleenwonen, samenwonen, huwen en scheiden. Forthcoming.

Table A1. Goodness-of-fit of a PH model and various SPH models for data on first union.

Model	$\hat{\chi}^2_L$	$\hat{\chi}^2_P$	ν
T + REL + EDUC + COH	659.93	693.42	375
EDU * T + REL + COH	375.08	398.64	335
EDU * (T + REL + COH)	363.81	378.92	327
STR * T + REL + EDU + COH	427.05	439.97	355
STR * (T + REL + EDU + COH)	421.73	427.15	351
T' + REL + EDU + COH	521.81	551.42	374
T" + REL + EDU + COH	459.82	476.60	373
T''' + REL + EDU + COH	483.02	503.87	372
T* + REL + EDU + COH	410.34	443.60	372

Legend : $\hat{\chi}^2_L$ = estimated likelihood ratio chi-squared statistic

$\hat{\chi}^2_P$ = estimated Pearson chi-squared statistic

ν = degrees of freedom

Table A2. Chi-squared statistics, with corresponding degrees of freedom, after fitting SPH models wherein the stratification is parametrized as a 2-years-shift for the HIGH-education group relative to the other ones.

Model	$\hat{\chi}^2_L$	$\hat{\chi}^2_P$	ν	$\lvert\hat{\chi}^2_L-\hat{\chi}^2_{L,A}\rvert$	$\lvert\nu-\nu_A\rvert$
T"	687.86	771.09	379	228.04	6
T" + REL	609.30	691.73	376	149.48	3
T" + EDU	628.33	658.79	377	168.51	4
T" + COH	609.94	689.00	378	150.12	5
T" + REL + EDU	551.35	579.50	374	91.53	1
T" + REL + COH	550.40	627.16	375	90.58	2
T" + EDU + COH	512.23	534.62	376	52.41	3
T" + REL + EDU + COH	$\hat{\chi}^2_{L,A}$=459.82	476.60	ν_A=373	–	–
T" + REL * EDU	539.39	561.01	368	~	~
T" + REL * COH	545.31	620.64	372	~	~
T" + EDU * COH	504.10	520.94	374	~	~
T" + REL * EDU + COH	448.14	459.47	367	11.68	6
T" + REL * COH + EDU	455.23	470.20	370	4.59	3
T" + EDU * COH + REL	452.77	463.84	371	7.05	2
T" + REL * EDU + REL * COH	444.67	455.78	364	15.15	9
T" + REL * EDU + EDU * COH	442.36	448.59	365	17.46	8
T" + REL * COH + EDU * COH	447.99	457.39	368	11.83	5
T" + REL * EDU + REL * COH + EDU * COH	438.70	444.68	362	21.12	11
T" + REL * EDU * COH	434.95	441.09	356	24.87	17

Legend : $\hat{\chi}^2_L$, $\hat{\chi}^2_P$ and ν : see Table A1.

$\lvert\hat{\chi}^2_L-\hat{\chi}^2_{L,A}\rvert$ = likelihood ratio chi-squared statistic corresponding to the difference between a model with likelihood ratio chi-squared statistic $\hat{\chi}^2_L$, and the model T" + REL + EDU + COH, if the two models are nested.

$\lvert\nu-\nu_A\rvert$ = the corresponding degrees of freedom.

~ : models are not nested.

Table A3. Analysis of deviance table for SPH models wherein the stratification is parametrized as a 2-years-shift for the HIGH-education group relative to the other ones.

Source of deviation	Effects controlled for	Deviance	Degrees of freedom	Mean Deviance (a)
REL	T"	78.56	3	26.19
EDU	T"	59.53	2	29.77
COH	T"	77.92	1	77.92
REL	T" + EDU + COH	52.41	3	17.47
EDU	T" + REL + COH	90.58	2	45.29
COH	T" + REL + EDU	91.53	1	91.53
REL	T" + EDU * COH	51.33	3	17.11
EDU	T" + REL * COH	90.08	2	45.04
COH	T" + REL * EDU	91.25	1	91.25
REL * EDU	main effects	11.68	6	1.95
REL * COH	main effects	4.59	3	1.53
EDU * COH	main effects	7.05	2	3.52
REL * EDU	main effects and other 2-way interactions	9.29	6	1.55
REL * COH	main effects and other 2-way interactions	1.01	3	.34
EDU * COH	main effects and other 2-way interactions	5.97	2	2.99
all 2-way interactions	main effects	21.12	11	1.92
3-way interaction	main effects and all 2-way interactions	3.75	6	.63
all interactions	main effects	24.87	17	1.46

(a) The mean deviance is the deviance per degree of freedom.

Table A4. Shift parameters b_z and estimated relative risks exp $(\hat{\beta}'_z)$, together with their standard errors, under the model T^* + REL + EDU + COH.

		Relative risk[(a)] (standard error) COH		Shift[(b)] (in years)
EDU	REL	48-62	38-47	
PRI	RC RMA	1.	.6637	0
			(.0284)	
	RC IRMA	.9743	.6466	-1
		(.0469)	(.0453)	
	NRA	1.1167	.7411	-1
		(.0862)	(.0680)	
	FREE	.9721	.6451	-1
		(.0864)	(.0657)	
SEC	RC RMA	.6192	.4109	0
		(.0325)	(.0306)	
	RC IRMA	.6033	.4004	-1
		(.0436)	(.0375)	
	NRA	.6915	.4589	-1
		(.0640)	(.0503)	
	FREE	.6019	.3995	-1
		(.0618)	(.0472)	
HIGH	RC RMA	.5738	.3808	+2
		(.0370)	(.0323)	
	RC IRMA	.5590	.3710	+1
		(.0475)	(.0389)	
	NRA	.6408	.4253	+1
		(.0658)	(.0507)	
	FREE	.5578	.3702	+1
		(.0590)	(.0452)	

(a) Relative to subgroup RC RMA-PRI-38-47; see footnote (18). Standard errors were calculated using the GLIM3-macros which are shown in Appendix B4 and D.

(b) Shift parameters b_z are relative to subgroup RC RMA-PRI. A negative (positive) shift parameter b_z indicates that women in subgroup z experience entry into first union earlier (later) than women in the reference subgroup.

Table A5. Covariate effects under the model T^* + REL + EDU + COH.

		Relative risk[(a)] (standard error)	Shift[(a)]
REL	RC RMA	1.	0
	RC IRMA	.9743 (.0469)	-1
	NRA	1.1167 (.0862)	-1
	FREE	.9721 (.0864)	-1
EDU	PRI	1.	0
	SEC	.6192 (.0325)	0
	HIGH	.5738 (.0370)	2
COH	48-62	1.	0
	38-47	.6637 (.0284)	0

(a) Relative to the first category of the covariate.

Table A6. Estimated (^) versus observed (~) conditional probabilities ($q_l(z)$) and cumulative probabilities ($F(a_l;z)$) of entry into first union, per subgroup.

SUBGROUP Z : RC RMA - PRI - 48-62

INDEX INTERVAL l	ENDPOINT AGE INTERVAL a_l+15	OBSERVED COND. PROB. ENTERING 1ST UNION $\tilde{q}_l(z)$	ESTIMATED COND. PROB. ENTERING 1ST UNION $\hat{q}_l(z)$	OBSERVED CUM. PROB. ENTERING 1ST UNION $\tilde{F}(a_l;z)$	ESTIMATED CUM. PROB. ENTERING 1ST UNION $\hat{F}(a_l;z)$
1	16.	0.0000	.0016	0.0000	.0016
3	18.	.0345	.0217	.0345	.0285
6	21.	.3448	.3129	.6611	.4937
11	26.	-	.5844	-	.9893
16	31.	-	.3824	-	.9990
21	36.	-	.0002	-	.9997

SUBGROUP Z : RC RMA - PRI - 38-47

1	16.	0.0000	.0011	0.0000	.0011
3	18.	.0366	.0144	.0366	.0190
6	21.	.1594	.2205	.2927	.3635
11	26.	.3750	.4416	.9390	.9508
16	31.	.3333	.2737	.9756	.9897
21	36.	0.0000	.0001	.9878	.9956

SUBGROUP Z : RC RMA - SEC - 48-62

1	16.	0.0000	.0010	0.0000	.0010
3	18.	.0097	.0135	.0144	.0178
6	21.	.2954	.2074	.4317	.3439
11	26.	.5556	.4194	.9434	.9398
16	31.	0.0000	.2580	.9830	.9860
21	36.	-	.0001	-	.9936

SUBGROUP Z : RC RMA - SEC - 38-47

1	16.	0.0000	.0007	0.0000	.0007
3	18.	0.0000	.0090	0.0000	.0118
6	21.	.1446	.1429	.2111	.2440
11	26.	.4524	.3029	.8722	.8451
16	31.	.2222	.1797	.9611	.9410
21	36.	0.0000	.0001	.9778	.9651

SUBGROUP Z : RC RMA - HIGH - 48-62

1	16.	0.0000	.0000	0.0000	.0000
3	18.	0.0000	.0009	0.0000	.0009
6	21.	.0600	.0451	.0662	.0609
11	26.	.3333	.3749	.8040	.8135
16	31.	.5000	.2099	.9571	.9710
21	36.	-	.2416	-	.9897

SUBGROUP Z : RC RMA - HIGH - 38-47

1	16.	0.0000	.0000	0.0000	.0000
3	18.	0.0000	.0006	0.0000	.0006
6	21.	.0141	.0302	.0141	.0408
11	26.	.2414	.2679	.6901	.6720
16	31.	.2500	.1448	.9155	.9045
21	36.	.5000	.1677	.9718	.9520

Table A6. Continued

SUBGROUP Z : RC IRMA - PRI - 48-62

INDEX INTERVAL l	ENDPOINT AGE INTERVAL a_l+15	OBSERVED COND. PROB. ENTERING 1ST UNION $\tilde{q}_l(z)$	ESTIMATED COND. PROB. ENTERING 1ST UNION $\hat{q}_l(z)$	OBSERVED CUM. PROB. ENTERING 1ST UNION $\tilde{F}(a_l;z)$	ESTIMATED CUM. PROB. ENTERING 1ST UNION $\hat{F}(a_l;z)$
1	16.	.0060	.0053	0.0000	.0053
3	18.	.1465	.0754	.1867	.0997
6	21.	.4714	.4429	.7711	.7125
11	26.	.2500	.5071	.9739	.9941
16	31.	-	.1989	-	.9990
21	36.	-	.0001	-	.9996

SUBGROUP Z : RC IRMA - PRI - 38-47

1	16.	.0170	.0035	0.0000	.0035
3	18.	.0819	.0507	.0909	.0673
6	21.	.2804	.3218	.5455	.5628
11	26.	.2727	.3747	.9375	.9667
16	31.	0.0000	.1369	.9602	.9899
21	36.	0.0000	.0001	.9716	.9949

SUBGROUP Z : RC IRMA - SEC - 48-62

1	16.	.0014	.0033	0.0000	.0033
3	18.	.0379	.0474	.0472	.0630
6	21.	.2979	.3039	.5121	.5379
11	26.	.3467	.3547	.9535	.9582
16	31.	.1667	.1283	.9811	.9863
21	36.	-	.0001	-	.9927

SUBGROUP Z : RC IRMA - SEC - 38-47

1	16.	0.0000	.0022	0.0000	.0022
3	18.	.0159	.0317	.0159	.0422
6	21.	.2556	.2137	.3694	.4009
11	26.	.3529	.2523	.9299	.8784
16	31.	.0769	.0871	.9618	.9421
21	36.	0.0000	.0001	.9809	.9620

SUBGROUP Z : RC IRMA - HIGH - 48-62

1	16.	0.0000	.0000	0.0000	.0000
3	18.	.0043	.0030	.0085	.0039
6	21.	.0673	.1038	.1068	.1569
11	26.	.2857	.3364	.8821	.8708
16	31.	.5000	.1252	.9806	.9722
21	36.	-	.1028	-	.9896

SUBGROUP Z : RC IRMA - HIGH - 38-47

1	16.	0.0000	.0000	0.0000	.0000
3	18.	.0156	.0020	.0156	.0026
6	21.	.0317	.0701	.0469	.1071
11	26.	.1905	.2382	.7344	.7428
16	31.	.3333	.0850	.9375	.9072
21	36.	0.0000	.0694	.9687	.9517

SUBGROUP z : NRA - PRI - 48-62

INDEX INTERVAL l	ENDPOINT AGE INTERVAL a_l+15	OBSERVED COND. PROB. ENTERING 1ST UNION $\tilde{q}_l(z)$	ESTIMATED COND. PROB. ENTERING 1ST UNION $\hat{q}_l(z)$	OBSERVED CUM. PROB. ENTERING 1ST UNION $\tilde{F}(a_l;z)$	ESTIMATED CUM. PROB. ENTERING 1ST UNION $\hat{F}(a_l;z)$
1	16.	0.0000	.0060	0.0000	.0060
3	18.	.1250	.0859	.1250	.1134
6	21.	.6250	.4886	.8750	.7604
11	26.	-	.5555	-	.9972
16	31.	-	.2245	-	.9996
21	36.	-	.0002	-	.9999

SUBGROUP z : NRA - PRI - 38-47

1	16.	0.0000	.0040	0.0000	.0040
3	18.	0.0000	.0579	0.0000	.0768
6	21.	.4000	.3592	.6000	.6126
11	26.	0.0000	.4162	.9333	.9798
16	31.	0.0000	.1552	.9333	.9949
21	36.	0.0000	.0001	.9333	.9976

SUBGROUP z : NRA - SEC - 48-62

1	16.	.0095	.0037	0.0000	.0037
3	18.	.0588	.0541	.0762	.0718
6	21.	.2951	.3398	.5810	.5872
11	26.	.6667	.3948	.9581	.9737
16	31.	-	.1456	-	.9927
21	36.	-	.0001	-	.9965

SUBGROUP z : NRA - SEC - 38-47

1	16.	0.0000	.0025	0.0000	.0025
3	18.	.0426	.0363	.0816	.0483
6	21.	.1282	.2409	.3061	.4441
11	26.	.2857	.2834	.8980	.9107
16	31.	0.0000	.0992	.9592	.9618
21	36.	0.0000	.0001	.9592	.9764

SUBGROUP z : NRA - HIGH - 48-62

1	16.	0.0000	.0000	0.0000	.0000
3	18.	0.0000	.0035	0.0000	.0045
6	21.	.1728	.1180	.2608	.1777
11	26.	.4000	.3750	.9190	.9042
16	31.	-	.1422	-	.9835
21	36.	-	.1169	-	.9947

SUBGROUP z : NRA - HIGH - 38-47

1	16.	0.0000	.0000	0.0000	.0000
3	18.	0.0000	.0023	0.0000	.0030
6	21.	.1000	.0799	.1000	.1218
11	26.	.3333	.2680	.8000	.7892
16	31.	-	.0968	-	.9344
21	36.	-	.0792	-	.9690

SUBGROUP z : FREE - PRI - 48-62

INDEX INTERVAL l	ENDPOINT AGE INTERVAL a_l+15	OBSERVED COND. PROB. ENTERING 1ST UNION $\tilde{q}_l(z)$	ESTIMATED COND. PROB. ENTERING 1ST UNION $\hat{q}_l(z)$	OBSERVED CUM. PROB. ENTERING 1ST UNION $\tilde{F}(a_l;z)$	ESTIMATED CUM. PROB. ENTERING 1ST UNION $\hat{F}(a_l;z)$
1	16.	0.0000	.0053	0.0000	.0053
3	18.	.2308	.0752	.2308	.0995
6	21.	.7500	.4422	.9231	.7117
11	26.	-	.5063	-	.9940
16	31.	-	.1985	-	.9990
21	36.	-	.0001	-	.9996

SUBGROUP z : FREE - PRI - 38-47

1	16.	0.0000	.0035	0.0000	.0035
3	18.	.0833	.0506	.0833	.0672
6	21.	.5714	.3212	.7500	.5620
11	26.	-	.3740	-	.9665
16	31.	-	.1366	-	.9898
21	36.	-	.0001	-	.9948

SUBGROUP z : FREE - SEC - 48-62

1	16.	0.0000	.0033	0.0000	.0033
3	18.	.0556	.0473	.0893	.0628
6	21.	.2933	.3033	.5205	.5371
11	26.	.5000	.3541	.9690	.9579
16	31.	0.0000	.1280	.9690	.9862
21	36.	-	.0001	-	.9927

SUBGROUP z : FREE - SEC - 38-47

1	16.	0.0000	.0022	0.0000	.0022
3	18.	0.0000	.0316	.0370	.0422
6	21.	.1905	.2133	.3704	.4002
11	26.	0.0000	.2518	.9630	.8778
16	31.	0.0000	.0869	.9630	.9417
21	36.	-	.0001	-	.9617

SUBGROUP z : FREE - HIGH - 48-62

1	16.	0.0000	.0000	0.0000	.0000
3	18.	0.0000	.0030	0.0000	.0039
6	21.	.1188	.1035	.1521	.1566
11	26.	.2308	.3358	.7664	.8702
16	31.	-	.1250	-	.9719
21	36.	-	.1026	-	.9895

SUBGROUP z : FREE - HIGH - 38-47

1	16.	0.0000	.0000	0.0000	.0000
3	18.	0.0000	.0020	0.0000	.0026
6	21.	.1000	.0700	.1000	.1069
11	26.	.5000	.2378	.8500	.7421
16	31.	0.0000	.0848	.9000	.9067
21	36.	0.0000	.0693	.9500	.9514

Table A7. Estimated first decile P10(z) and median Me(z) under the model T^{*}+ REL + EDU + COH.(a)

EDU	REL	COH 48-62 P10(z)	COH 48-62 Me(z)	COH 38-47 P10(z)	COH 38-47 Me(z)
PRI	RC RMA	18.9	21.0	19.2	21.6
	RC IRMA	18.0	20.1	18.3	20.6
	NRA	17.8	19.8	18.2	20.4
	FREE	18.0	20.1	18.3	20.6
SEC	RC RMA	19.3	21.7	19.7	22.4
	RC IRMA	18.3	20.8	18.7	21.4
	NRA	18.2	20.5	18.6	21.2
	FREE	18.3	20.8	18.7	21.4
HIGH	RC RMA	21.4	23.8	21.8	24.5
	RC IRMA	20.4	22.9	20.8	23.6
	NRA	20.3	22.7	20.7	23.3
	FREE	20.4	22.9	20.8	23.6

(a) Me(z) = t_1+15, where t_1 is the time point satisfying $\hat{F}(t_1;z)/\hat{F}(a_L;z) = .50$

P10(z) = t_2+15, where t_2 is the time point satisfying $\hat{F}(t_2;z)/\hat{F}(a_L;z) = .10$

(The procedure to calculate Me(z) and P10(z) is outlined in Appendix E6).

Table A8. Relative risks, used for comparison of subgroups or causes in competing risks models

Model	Cause-specific and subgroup-specific relative risk			Relative risk for total hazard	
	Assumptions satisfied[(a)]	Comparison of z with z_o, given j[(b)]	Comparison of j with j_o, given z[(c)]	Assumption satisfied[(a)]	Comparison of z with z_o[(d)]
(3.40)	-	-	-	-	-
(3.66) SPH	SII	$e^{\tau_{jz}}$	-	-	-
(3.62) PH	III	-	$e^{\gamma_{jz}}$	-	-
(3.64) PH	II(=SII+RII)	$e^{\tau_{jz}}$	-	-	-
(3.67) SPH	SII, $\tau_{jz}=\tau_z$	e^{τ_z}	-	SI	$e^{\beta_z} = e^{\tau_z}$
(3.65) SPH	SII,III	$e^{\lambda_{jz}}/e^{\lambda_{jz_o}}$	$e^{\lambda_{jz}}/e^{\lambda_{j_oz}}$	SI	$e^{\beta_z} = (\sum_j e^{\lambda_{jz}})/(\sum_j e^{\lambda_{jz_o}})$
(3.63) PH	III,$\gamma_{jz}=\gamma_j$	-	e^{γ_j}	-	-
(3.60) PH	II,$\tau_{jz}=\tau_z$	e^{τ_z}	-	I	$e^{\beta_z} = e^{\tau_z}$
(3.57) PH	II,III	$e^{\lambda_{jz}}/e^{\lambda_{jz_o}}$	$e^{\lambda_{jz}}/e^{\lambda_{j_oz}}$	I	$e^{\beta_z} = (\sum_j e^{\lambda_{jz}})/(\sum_j e^{\lambda_{jz_o}})$
(3.68) SPH	SII,III,$\lambda_{jz}=\gamma_j+\tau_z$	e^{τ_z}	e^{γ_j}	SI,$\lambda_{jz}=\gamma_j+\tau_z$	$e^{\beta_z} = e^{\tau_z}$
(3.61) PH	II,III,$\lambda_{jz}=\gamma_j+\tau_z$	e^{τ_z}	e^{γ_j}	I,$\lambda_{jz}=\gamma_j+\tau_z$	$e^{\beta_z} = e^{\tau_z}$

(a) SII and RII : footnote Fig. C5; SI : assumption I holds within each stratum.

(b) In PH models : $\tau_{jz_o}=\tau_{z_o}=0$; in SPH models : z and z_o are subgroups of the same stratum - i.e. $z = (z_1,z_2)$ and $z_o = (z_1,z_{20})$ - and $\tau_{jz_o} = \tau_{z_o} = 0$. Also : $\lambda_{j_oz_o} = 0$ but $\lambda_{jz_o} \neq 0$ in general.

(c) In general $\lambda_{j_oz} \neq 0$, but $\gamma_{j_oz} = \gamma_{j_o} = 0$.

(d) In PH and in SPH models : $\beta_{z_o} = 0$; see (b) for z and z_o in SPH models.

Table A9. Shift parameters and relative risks in parametrized SPH models

Parametrized Model	SPH Model	Comp. of origin in z_1 and z_{10} [(a)]	Comparison through cause-specific hazards: Comp. of z_1 and z_{10}, given j, after adjusting for origin [(b)]	Comparison through cause-specific hazards: Comp. of j with j_o, given z	Comparison through cause-specific hazards: Comp. of z with z_o, given j [(c)]	Comparison through total hazards: Comp. of z with z_o [(c)]
(3.69-70a)	(3.66)	b_{z_1}	$e^{\omega_{jz_1}} = e^{\tau'_{jz_o}}$	–	$e^{\tau_{jz}} = e^{\tau'_{jz}}/e^{\tau'_{j_o z}}$	–
(3.69-70b)	(3.67)	b_{z_1}	$e^{\omega_{jz_1}} = e^{\tau'_{jz_o}}$	–	$e^{\tau_z} = e^{\tau'_{jz}}/e^{\tau'_{jz_o}}$	$e^{\beta_z} = e^{\tau_z}$
(3.71-72a)	(3.65)	b_{z_1}	$e^{\omega_{j_o z_1}} = e^{\lambda'_{j_o z_o}}$	$e^{\gamma_{jz}} = e^{\lambda'_{jz}}/e^{\lambda'_{j_o z}}$	$e^{\tau_{jz}} = e^{\lambda'_{jz}}/e^{\lambda'_{jz_o}}$	$e^{\beta_z} = (\sum_j e^{\lambda'_{jz}})/(\sum_j e^{\lambda'_{jz_o}})$
(3.71-72b)	(3.68)	b_{z_1}	$e^{\omega_{j_o z_1}} = e^{\lambda'_{j_o z_o}}$	$e^{\gamma_j} = e^{\lambda'_{jz}}/e^{\lambda'_{j_o z}}$	$e^{\tau_z} = e^{\lambda'_{jz}}/e^{\lambda'_{jz_o}}$	$e^{\beta_z} = (\sum_j e^{\lambda'_{jz}})/(\sum_j e^{\lambda'_{jz_o}})$

(a) z_1 and z_{10} are different strata; z_{10} is the reference stratum : $b_{z_{10}}=0$.

(b) I.e. comparison of the reference subgroup z_o (in stratum z_1) with the reference subgroup z_{oo} (in stratum z_{10}); $\omega_{jz_{10}}=0$; only exp $(\omega_{j_o z_1})$ is given if there exists a reference cause (i.e. if III holds).

(c) z and z_o are subgroups in the same stratum (z_1).

Table A10. Specification of the linear predictor in GLIM3, under various competing risks models

Model	GLIM3 expression of the linear predictor[a]
(3.40)	T*TYPE*Z
(3.66)	T*TYPE*Z1+TYPE*Z
(3.62)	T*Z+TYPE*Z
(3.64)	T*TYPE+TYPE*Z
(3.67)	T*TYPE*Z1+Z
(3.65)	T*Z1+TYPE*Z
(3.63)	T*Z+TYPE
(3.60)	T*TYPE+Z
(3.57)	T+TYPE*Z
(3.68)	T*Z1+TYPE+Z
(3.61)	T+TYPE+Z

(a) Z stands for an expression depending on ZR (i.e. REL), ZE (i.e. EDU) and/or ZC (i.e. COH).

Z1 stands for an expression depending on the stratifying covariates only.

Table A11. Relative risks exp $(\hat{\lambda}'_{jz})$ under the model T^{*} + TYPE*(ZR + ZE + ZC)

		ZC			
		48-62		38-47	
ZE	ZR	$e^{\hat{\lambda}'_{1z}}$	$e^{\hat{\lambda}'_{2z}}$	$e^{\hat{\lambda}'_{1z}}$	$e^{\hat{\lambda}'_{2z}}$
PRI	RC RMA	1.	.0039	.6997	.0011
			(.0029)	(.0309)	(.0000)
	RC IRMA	.9194	.0713	.6433	.0200
		(.0448)	(.0161)	(.0461)	(.0051)
	NRA	.9164	.2403	.6412	.0673
		(.0763)	(.0625)	(.0627)	(.0191)
	FREE	.7087	.2615	.4959	.0733
		(.0724)	(.0709)	(.0567)	(.0215)
SEC	RC RMA	.6266	.0022	.4384	.0006
		(.0337)	(.0016)	(.0336)	(.0000)
	RC IRMA	.5761	.0399	.4031	.0112
		(.0425)	(.0057)	(.0387)	(.0024)
	NRA	.5742	.1345	.4017	.0377
		(.0565)	(.0245)	(.0464)	(.0092)
	FREE	.4441	.1464	.3107	.0410
		(.0511)	(.0287)	(.0404)	(.0104)
HIGH	RC RMA	.5396	.0041	.3776	.0012
		(.0365)	(.0030)	(.0334)	(.0000)
	RC IRMA	.4961	.0760	.3472	.0213
		(.0436)	(.0126)	(.0377)	(.0051)
	NRA	.4945	.2561	.3460	.0718
		(.0541)	(.0508)	(.0346)	(.0188)
	FREE	.3824	.2786	.2676	.0781
		(.0453)	(.0480)	(.0359)	(.0189)

Table A12. Subgroup-specific relative risks[(a)] exp ($\hat{\gamma}_{jz}$) and subgroup-specific weights $\hat{\theta}_{jz}$ under the model T^* + TYPE*(ZR + ZE + ZC).

		ZC					
		48-62			38-47		
ZE	ZR	$e^{\hat{\gamma}_{2z}}$	$\hat{\theta}_{1z}$	$\hat{\theta}_{2z}$	$e^{\hat{\gamma}_{2z}}$	$\hat{\theta}_{1z}$	$\hat{\theta}_{2z}$
PRI	RC RMA	.0039 (.0029)	.9961	.0039	.0016 (.0012)	.9984	.0016
	RC IRMA	.0776 (.0173)	.9280	.0720	.0311 (.0078)	.9699	.0301
	NRA	.2622 (.0699)	.7923	.2077	.1050 (.0303)	.9050	.0950
	FREE	.3689 (.1048)	.7305	.2695	.1478 (.0450)	.8712	.1288
SEC	RC RMA	.0035 (.0025)	.9965	.0035	.0014 (.0010)	.9986	.0014
	RC IRMA	.0693 (.0090)	.9351	.0649	.0278 (.0059)	.9730	.0270
	NRA	.2343 (.0435)	.8102	.1898	.0938 (.0232)	.9142	.0858
	FREE	.3297 (.0687)	.7520	.2480	.1320 (.0349)	.8833	.1167
HIGH	RC RMA	.0077 (.0055)	.9924	.0076	.0031 (.0022)	.9969	.0031
	RC IRMA	.1532 (.0246)	.8671	.1329	.0614 (.0146)	.9422	.0578
	NRA	.5178 (.1066)	.6589	.3411	.2074 (.0568)	.8283	.1717
	FREE	.7286 (.1363)	.5785	.4215	.2918 (.0743)	.7741	.2259

(a) The subgroup-specific relative risks exp ($\hat{\gamma}_{1z}$) are all equal to 1. since marriage is the reference cause.

Table A13. Cause-specific relative risks exp $(\hat{\tau}_{jz})$ under the model $T^* + TYPE*(ZR + ZE + ZC)$

		ZC			
		48-62		38-47	
ZE	ZR	$e^{\hat{\tau}_{1z}}$	$e^{\hat{\tau}_{2z}}$	$e^{\hat{\tau}_{1z}}$	$e^{\hat{\tau}_{2z}}$
PRI	RC RMA	1.	1.	.6997	.2802
				(.0309)	(.0553)
	RC IRMA	1.	1.	.6997	.2802
				(.0309)	(.0552)
	NRA	.9967	3.369	.6974	.9441
		(.0769)	(.6171)	(.0615)	(.2524)
	FREE	.7708	3.666	.5394	1.027
		(.0755)	(.6801)	(.0577)	(.2760)
SEC	RC RMA	.6266	.5599	.4384	.1569
		(.0337)	(.1279)	(.0336)	(.0507)
	RC IRMA	.6266	.5599	.4384	.1569
		(.0337)	(.1279)	(.0336)	(.0507)
	NRA	.6245	1.886	.4369	.5286
		(.0577)	(.5410)	(.0466)	(.1932)
	FREE	.4830	2.053	.3379	.5752
		(.0534)	(.5891)	(.0415)	(.2102)
HIGH	RC RMA	1.	1.	.6997	.2802
				(.0309)	(.0553)
	RC IRMA	1.	1.	.6997	.2802
				(.0309)	(.0553)
	NRA	.9967	3.369	.6974	.9441
		(.0769)	(.6171)	(.0615)	(.2524)
	FREE	.7708	3.666	.5394	1.027
		(.0755)	(.6801)	(.0577)	(.2760)

Table A14. Relative risks $e^{\hat{\beta}_Z}$ under the model $T^* + TYPE*(ZR + ZE + ZC)$

		ZC	
ZE	ZR	48-62	38-47
PRI	RC RMA	1.	.6997
			(.0308)
	RC IRMA	1.	.6997
			(.0302)
	NRA	.9967	.6974
		(.0718)	(.0584)
	FREE	.7708	.5394
		(.0673)	(.0536)
SEC	RC RMA	.6266	.4384
		(.0336)	(.0335)
	RC IRMA	.6266	.4384
		(.0329)	(.0329)
	NRA	.6245	.4369
		(.0578)	(.0448)
	FREE	.4830	.3379
		(.0529)	(.0394)
HIGH	RC RMA	1.	.6997
			(.0308)
	RC IRMA	1.	.6997
			(.0302)
	NRA	.9967	.6974
		(.0803)	(.0602)
	FREE	.7708	.5394
		(.0744)	(.0554)

Table A15. Shift parameters b_{z_1} and relative risks $e^{\hat{\omega}_{j_o z_1}}$, under the model

$T^* + TYPE * (ZR + ZE + ZC)$

Stratum z_1 (reference subgroup z_o)	b_{z_1}	$e^{\hat{\omega}_{j_o z_1}}$
PRI/SEC - RC RMA (PRI - RC RMA - 48-62)	0	1.
PRI/SEC - not RC RMA (PRI - RC IRMA - 48-62)	-1	.9194 (.0448)
HIGH - RC RMA (HIGH - RC RMA - 48-62)	+2	.5396 (.0365)
HIGH - not RC RMA (HIGH - RC IRMA - 48-62)	+1	.4961 (.0436)

Table A16. Covariate effects under the model T* + TYPE*(ZR + ZE + ZC).

		Relative risk[a] (standard error) Marriage	Cohabitation	Shift[a]
REL	RC RMA	1.	1.	0
	RC IRMA	.9194	18.37	-1
		(.0448)	(13.14)	
	NRA	.9164	61.87	-1
		(.0763)	(44.81)	
	FREE	.7087	67.33	-1
		(.0724)	(48.68)	
EDU	PRI	1.	1.	0
	SEC	.6266	.5599	0
		(.0337)	(.1279)	
	HIGH	.5396	1.066	2
		(.0365)	(.2575)	
COH	48-62	1.	1.	0
	38-47	.6997	.2802	0
		(.0309)	(.0553)	

(a) Relative to the first category of the covariate.

Table A17. Estimated first deciles $P10_j(z)$ and medians $Me_j(z)$ under the model $T^* + TYPE*(ZR + ZE + ZC)$

		First marriage (j=1)				First cohabitation (j=2)			
		ZC				ZC			
		48 - 62		38 - 47		48 - 62		38 - 47	
		$P10_1(z)$	$Me_1(z)$	$P10_1(z)$	$Me_1(z)$	$P10_2(z)$	$Me_2(z)$	$P10_2(z)$	$Me_2(z)$
PRI	RC RMA	19.0	21.1	19.2	21.6	21.2	25.5	21.2	25.5
	RC IRMA	18.0	20.2	18.3	20.7	19.9	23.8	20.1	24.3
	NRA	18.1	20.2	18.3	20.7	19.2	22.3	19.9	23.9
	FREE	18.2	20.5	18.5	21.1	19.1	22.2	19.9	23.8
SEC	RC RMA	19.3	21.7	19.7	22.3	21.2	25.5	21.2	25.5
	RC IRMA	18.4	20.9	18.8	21.5	20.1	24.1	20.2	24.4
	NRA	18.4	20.9	18.8	21.5	19.6	23.2	20.1	24.1
	FREE	18.7	21.3	19.0	21.9	19.5	23.0	20.0	24.1
HIGH	RC RMA	21.4	24.0	21.8	24.6	23.2	27.3	23.2	27.4
	RC IRMA	20.5	23.1	20.9	23.7	21.9	25.8	22.1	26.3
	NRA	20.5	23.1	20.9	23.7	21.2	24.2	21.9	25.8
	FREE	20.8	23.6	21.1	24.1	21.1	24.1	21.9	25.7

APPENDIX B. GLIM3 - PROGRAMMES AND -OUTPUTS

B1. GLIM3 - programme for fitting the PH model T + REL + EDU + COH.

```
$UNITS 402
$DATA BT N D1 D2 W REL EDU COH
$DINPUT 1
$CALC T=BT-14
$FACTOR T 21 REL 4 EDU 3 COH 2
$CALC D=D1+D2  $DEL D1 D2
$CALC LE=%LOG(N-(W+D)/2)
$CALC %S=0. : %S=%S-2*D*(%LOG(D)-LE)-D
$PRINT : : : "  SATURATED MODEL T*REL*EDU*COH "
         : : "  HAS -2*LOG(LIKELIHOOD) =       " *9 %S : : : :
$YVAR D
$ERROR P
$OFFSET LE
$FIT T+REL+EDU+COH
$DISPLAY L A
$CALC %L=%S+%DV
$PRINT : : : "  THE CURRENT MODEL HAS "
         : : "  -2*LOG(LIKELIHOOD) =              " *9 %L
         : : "  LIKELIHOOD RATIO CHI SQUARE =     " *9 %DV
         : : "  PEARSON CHI SQUARE =              " *9 %X2
         : : "  WITH DEGREES OF FREEDOM =         " *3 %DF
$STOP
```

B1. (continued) GLIM3 - output after fitting the PH model T + REL + EDU + COH

```
1 GLIM 3.11 (C)1977 ROYAL STATISTICAL SOCIETY, LONDON

  ----- $DATA LIST ABOLISHED
  ----- INVALID FUNCTION/OPERATOR ARGUMENT(S)


    SATURATED MODEL T*REL*EDU*COH

    HAS -2*LOG(LIKELIHOOD) =          3906.52813


             SCALED
   CYCLE   DEVIANCE       DF
     4      659.9        375

  LINEAR PREDICTOR
  %GM T REL EDU COH

            ESTIMATE         S.E.          PARAMETER
      1     -5.774          .4056          %GM
      0       ZERO        ALIASED          T(1)
      2      1.474          .4466          T(2)
      3      2.801          .4143          T(3)
      4      3.626          .4075          T(4)
      5      4.229          .4053          T(5)
      6      4.843          .4041          T(6)
      7      5.283          .4039          T(7)
      8      5.344          .4052          T(8)
      9      5.511          .4064          T(9)
     10      5.573          .4091          T(10)
     11      5.588          .4135          T(11)
     12      5.281          .4263          T(12)
     13      5.133          .4411          T(13)
     14      4.464          .4943          T(14)
     15      4.576          .5019          T(15)
     16      4.861          .4945          T(16)
     17      4.996          .5014          T(17)
     18      4.645          .5728          T(18)
     19      3.777          .7935          T(19)
     20      4.379          .6969          T(20)
     21      4.074          .7920          T(21)
      0       ZERO        ALIASED          REL(1)
     22      .3020          .4794E-01      REL(2)
     23      .4248          .7703E-01      REL(3)
     24      .3333          .8878E-01      REL(4)
      0       ZERO        ALIASED          EDU(1)
     25     -.4933          .5244E-01      EDU(2)
     26     -1.198          .6556E-01      EDU(3)
      0       ZERO        ALIASED          COH(1)
     27     -.4038          .4288E-01      COH(2)
     SCALE PARAMETER TAKEN AS              1.000


    THE CURRENT MODEL HAS

    -2*LOG(LIKELIHOOD) =                       4566.45922

    LIKELIHOOD RATIO CHI SQUARE =              659.931083

    PEARSON CHI SQUARE =                       693.422384

    WITH DEGREES OF FREEDOM =                  375.
```

B2. GLIM3-programme for fitting the SPH model T" + REL + EDU + COH

```
$UNITS 402
$DATA BT N D1 D2 W REL EDU COH
$DINPUT 1
$CALC T=BT-14 : T=%IF(%EQ(EDU,3),T,T+2)
$FACTOR T 23 REL 4 EDU 3 COH 2
$CALC D=D1+D2  $DEL D1 D2
$CALC LE=%LOG(N-(W+D)/2)
$CALC %S=0. : %S=%S-2*D*(%LOG(D)-LE)-D
$PRINT : : : "  SATURATED MODEL T*REL*EDU*COH "
         : : "  HAS -2*LOG(LIKELIHOOD) =       " *9 %S : : : :
$YVAR D
$ERROR P
$OFFSET LE
$FIT T+REL+EDU+COH
$DISPLAY L A
$CALC %L=%S+%DV
$PRINT : : : "  THE CURRENT MODEL HAS "
         : : "  -2*LOG(LIKELIHOOD) =                 " *9 %L
         : : "  LIKELIHOOD RATIO CHI SQUARE =        " *9 %DV
         : : "  PEARSON CHI SQUARE =                 " *9 %X2
         : : "  WITH DEGREES OF FREEDOM =            " *3 %DF
$STOP
```

B2. (continued) GLIM3 - Output after fitting the SPH model T" + REL + EDU + COH

```
1 GLIM 3.11 (C)1977 ROYAL STATISTICAL SOCIETY, LONDON

  ----- $DATA LIST ABOLISHED
  ----- INVALID FUNCTION/OPERATOR ARGUMENT(S)


    SATURATED MODEL T*REL*EDU*COH

    HAS -2*LOG(LIKELIHOOD) =          3906.52813


          SCALED
   CYCLE  DEVIANCE       DF
     8     459.8        373

  LINEAR PREDICTOR
  %GM T REL EDU COH

          ESTIMATE        S.E.          PARAMETER
      1   -12.22         13.25          %GM
      0     ZERO        ALIASED         T(1)
      2    6.169         13.29          T(2)
      3    6.633         13.26          T(3)
      4    7.963         13.25          T(4)
      5    9.237         13.25          T(5)
      6    10.15         13.25          T(6)
      7    10.83         13.25          T(7)
      8    11.40         13.25          T(8)
      9    11.74         13.25          T(9)
     10    11.72         13.25          T(10)
     11    11.79         13.25          T(11)
     12    11.73         13.25          T(12)
     13    11.89         13.25          T(13)
     14    11.25         13.25          T(14)
     15    11.17         13.25          T(15)
     16    11.06         13.25          T(16)
     17    10.95         13.26          T(17)
     18    10.62         13.26          T(18)
     19    10.94         13.26          T(19)
     20    11.31         13.26          T(20)
     21    10.91         13.26          T(21)
     22    3.335         20.90          T(22)
     23    3.148         22.45          T(23)
      0     ZERO        ALIASED         REL(1)
     24    .3009         .4800E-01      REL(2)
     25    .4395         .7711E-01      REL(3)
     26    .2996         .8878E-01      REL(4)
      0     ZERO        ALIASED         EDU(1)
     27   -.4826         .5242E-01      EDU(2)
     28   -.5642         .6454E-01      EDU(3)
      0     ZERO        ALIASED         COH(1)
     29   -.4030         .4263E-01      COH(2)
     SCALE PARAMETER TAKEN AS           1.000


    THE CURRENT MODEL HAS

    -2*LOG(LIKELIHOOD) =                    4366.34973

    LIKELIHOOD RATIO CHI SQUARE =           459.821591

    PEARSON CHI SQUARE =                    476.602581

    WITH DEGREES OF FREEDOM =               373.
```

B3. GLIM3-programme for fitting the SPH model T^{*} + REL + EDU + COH

```
$UNITS 402
$DATA BT N D1 D2 W REL EDU COH
$DINPUT 1
$CALC T=BT-14
     : T=%IF(%EQ(EDU,1)*%NE(REL,1),T+3,T)
     : T=%IF(%EQ(EDU,2)*%NE(REL,1),T+3,T)
     : T=%IF(%EQ(EDU,1)*%EQ(REL,1),T+2,T)
     : T=%IF(%EQ(EDU,2)*%EQ(REL,1),T+2,T)
     : T=%IF(%EQ(EDU,3)*%NE(REL,1),T+1,T)
$FACTOR T 24 REL 4 EDU 3 COH 2
$CALC D=D1+D2  $DEL D1 D2
$CALC LE=%LOG(N-(W+D)/2)
$CALC %S=0. : %S=%S-2*D*(%LOG(D)-LE)-D
$PRINT : : : "  SATURATED MODEL T*REL*EDU*COH "
         : : "  HAS -2*LOG(LIKELIHOOD) =        " *9 %S : : : :
$YVAR D
$ERROR P
$OFFSET LE
$FIT T+REL+EDU+COH
$DISPLAY L A
$CALC %L=%S+%DV
$PRINT : : : "  THE CURRENT MODEL HAS "
         : : "  -2*LOG(LIKELIHOOD) =                 " *9 %L
         : : "  LIKELIHOOD RATIO CHI SQUARE =        " *9 %DV
         : : "  PEARSON CHI SQUARE =                 " *9 %X2
         : : "  WITH DEGREES OF FREEDOM =            " *3 %DF
$STOP
```

B3. (continued) GLIM3 - output after fitting the SPH model T^* + REL + EDU + COH

```
1 GLIM 3.11 (C)1977 ROYAL STATISTICAL SOCIETY, LONDON

 ----- $DATA LIST ABOLISHED
 ----- INVALID FUNCTION/OPERATOR ARGUMENT(S)


   SATURATED MODEL T*REL*EDU*COH

   HAS -2*LOG(LIKELIHOOD) =          3906.52813



        SCALED
  CYCLE DEVIANCE        DF
    8    410.3          372

 LINEAR PREDICTOR
 %GM T REL EDU COH

        ESTIMATE        S.E.          PARAMETER
    1   -12.58         30.59          %GM
    0     ZERO       ALIASED          T(1)
    2    .5967         33.29          T(2)
    3    6.138         30.61          T(3)
    4    7.364         30.59          T(4)
    5    8.760         30.59          T(5)
    6    10.06         30.59          T(6)
    7    10.95         30.59          T(7)
    8    11.60         30.59          T(8)
    9    12.07         30.59          T(9)
   10    12.39         30.59          T(10)
   11    12.38         30.59          T(11)
   12    12.27         30.59          T(12)
   13    12.45         30.59          T(13)
   14    12.26         30.59          T(14)
   15    11.87         30.59          T(15)
   16    11.69         30.59          T(16)
   17    11.15         30.59          T(17)
   18    11.85         30.59          T(18)
   19    11.10         30.59          T(19)
   20    11.60         30.59          T(20)
   21    11.85         30.59          T(21)
   22    10.94         30.60          T(22)
   23    3.931         34.63          T(23)
   24    3.680         37.94          T(24)
    0     ZERO       ALIASED          REL(1)
   25   -.2609E-01    .4814E-01       REL(2)
   26    .1104        .7715E-01       REL(3)
   27   -.2834E-01    .8885E-01       REL(4)
    0     ZERO       ALIASED          EDU(1)
   28   -.4793        .5240E-01       EDU(2)
   29   -.5554        .6448E-01       EDU(3)
    0     ZERO       ALIASED          COH(1)
   30   -.4100        .4277E-01       COH(2)
   SCALE PARAMETER TAKEN AS          1.000


   THE CURRENT MODEL HAS

   -2*LOG(LIKELIHOOD) =                  4316.86769

   LIKELIHOOD RATIO CHI SQUARE =         410.339560

   PEARSON CHI SQUARE =                  443.599732

   WITH DEGREES OF FREEDOM =             372.
```

B4. Computation of (co)variances and standard errors of base-line hazards and relative risks in the model T^* + REL + EDU + COH.

Appendix B3 shows the GLIM3-programme for fitting the model T^* + REL + EDU + COH, and the estimates of the linear parameters with their standard errors. These linear parameters are - for the purpose of interpretation of the fitted results - transformed into base-line hazards and relative risks. It will be shown in this appendix how the (co) variances and standard errors of the new parameters can be obtained by using GLIM3.

The model formula can be written as

$$\mu(t;z) = \mu(t+b_{z_1};z_{oo}).e^{\beta'_z}$$

or $$\mu(t;z) = e^{\alpha_{l+b_{z_1}}}.e^{\beta'_z} \qquad \text{for } a_{l-1} \leqslant t < a_l,$$

where z is a subgroup in stratum z_1, and z_{oo} is the reference subgroup in the reference stratum z_{10}. Since l varies from 1 to 21, and since the shift b_{z_1} is at most 3 (years), the index $l+b_{z_1}$ of the time parameters α varies from 1 to 24. Hence, 24 base-line hazards exp (α_l), l=1,...24, are computed. Since there are 24 subgroups, 23 relative risks exp (β'_z) are also produced. In total, 47 new parameters, their standard errors and (co)variances are obtained. This is done by using the GLIM3-macros MUL5, MUL6, MUL7, MUL8 and COVA shown in Appendix D. Since the macro COVA uses a macro PADE - which computes the new parameters and the first order partial derivatives of these new parameters with respect to the linear parameters estimated by GLIM3 - the basic elements needed for the construction of this macro are discussed in the following paragraphs.

If the new parameters are denoted by ϕ_i (i=1,...47) - corresponding to the notations in Appendix D - and if the linear parameters are denoted by PE_j (j=1,...30) - corresponding to the interval names %PE(j) in GLIM3 -, then the ϕ_i are functions of the PE_j as follows :

$$\phi_1 = \exp\ (PE_1)$$

$$\phi_i = \exp\ (PE_1 + PE_i) \qquad \text{for } i=2,...24$$

$$\phi_i = \exp(PE_i) \qquad \text{for } i=25,\ldots 30$$

$$\phi_i = \exp(PE_{i-6} + PE_{28}) \qquad \text{for } i=31,\ldots 33$$

$$\phi_i = \exp(PE_{i-9} + PE_{29}) \qquad \text{for } i=34,\ldots 36$$

$$\phi_i = \exp(PE_{i-12} + PE_{30}) \qquad \text{for } i=37,\ldots 39$$

$$\phi_i = \exp(PE_{i-12} + PE_{30}) \qquad \text{for } i=40,41$$

$$\phi_i = \exp(PE_{i-17} + PE_{28} + PE_{30}) \qquad \text{for } i=42,\ldots 44$$

$$\phi_i = \exp(PE_{i-20} + PE_{29} + PE_{30}) \qquad \text{for } i=45,\ldots 47$$

Note that $\phi_1,\ldots\phi_{24}$ are the base-line hazards, and $\phi_{25},\ldots\phi_{47}$ are the relative risks $\exp(\beta_z')$. For instance, ϕ_{43} is the relative risks corresponding to the subgroup NRA - SEC - 38-47.

Since the ϕ_i are ordinary exponential functions of the PE_j, the first order partial derivatives $\frac{\partial \phi_i}{\partial PE_j}$ can easily be found to be either ϕ_i if ϕ_i is a function of PE_j or zero if ϕ_i is not a function of PE_j. A 47 x 30 matrix A is then defined as follows :

$$A = (a_{ij})_{\substack{i=1,\ldots 47 \\ j=1,\ldots 30}} \qquad \text{with } a_{ij}=1 \text{ if } \frac{\partial \phi_i}{\partial PE_j} = \phi_i$$

$$=0 \text{ if } \frac{\partial \phi_i}{\partial PE_j} = 0.$$

If FI is the 1 x 47 vector of new parameters ϕ_i, and F is the 47 x 47 matrix for which any row is equal to the vector FI, then the matrix D of first order partial derivatives can be written as the product of the matrix F with the matrix A :

$$D = F.A$$

Moreover, if PE is the 1 x 30 vector of linear parameters PE_j, then FI can be written as follows :

$$FI = \exp (PE.A^T)$$

where ^{T} denotes matrix transposition, and exp denotes matrix exponentiation. These ideas were used in order to construct the macro PADE.

The macro PADE uses another macro PHID. These two macros are as follows.

```
$M PHID ! %I-TH NEW PARAMETER & PARTIAL DERIVATIVES
$CAL %A=%A-1 : %I=%N-%A : FI(%I)=%CU(D(J+(%I-1)*%PL)*%PE(J)) !
: FI(%I)=%EXP(FI(%I)) !
: D(J+(%I-1)*%PL)=D(J+(%I-1)*%PL)*FI(%I) $$ENDM
$M PADE ! NEW PARAMETERS (FI) & PARTIAL DERIVATIVES (D)
$VAR %N FI : %M I1 I2 D : %PL J !
$CAL I1=%GL(%N,%PL) : I2=%GL(%PL,1) : D=%EQ(I1,I2) !
+%EQ(I2,1)*%GE(I1,2)*%LE(I1,24) !
+%EQ(I2,25)*(%EQ(I1,31)+%EQ(I1,34)+%EQ(I1,37)+%EQ(I1,42)+%EQ(I1,45)) !
+%EQ(I2,26)*(%EQ(I1,32)+%EQ(I1,35)+%EQ(I1,38)+%EQ(I1,43)+%EQ(I1,46)) !
+%EQ(I2,27)*(%EQ(I1,33)+%EQ(I1,36)+%EQ(I1,39)+%EQ(I1,44)+%EQ(I1,47)) !
: D=D !
+%EQ(I2,28)*(%GE(I1,31)*%LE(I1,33)+%EQ(I1,40)+%GE(I1,42)*%LE(I1,44)) !
+%EQ(I2,29)*(%GE(I1,34)*%LE(I1,36)+%EQ(I1,41)+%GE(I1,45)) !
+%EQ(I2,30)*%GE(I1,37) !
: J=%GL(%PL,1) : %A=%N !
$WHI %A PHID !
$DEL I1 I2 J !
$$ENDM
```

To start with, macro PADE computes the matrix A and stores its elements in the GLIM3-vector D. Then the macro PHID is used repeatedly to compute the i-th new parameter and its partial derivatives, and to store these derivatives in the corresponding components of the GLIM3-vector D.

Finally, the computation of the new parameters, their standard errors and their (co)variances is executed through the GLIM3-statements

```
$CALC %N=47 $USE COVA
```

since there are 47 new parameters. The estimated standard errors are as follows

PAR.	ESTIMATE	S. E.
1	.3441E-05	0.
2	.6250E-05	0.
3	.1594E-02	.1597E-02
4	.5430E-02	.1837E-02
5	.2192E-01	.3871E-02
6	.8076E-01	.8567E-02
7	.1956	.1645E-01
8	.3739	.2877E-01
9	.5991	.4539E-01
10	.8302	.6467E-01
11	.8172	.7143E-01
12	.7364	.7591E-01
13	.8822	.1001
14	.7228	.1054
15	.4919	.1001
16	.4115	.1035
17	.2382	.8575E-01
18	.4802	.1372
19	.2281	.1033
20	.3770	.1450
21	.4835	.1860
22	.1938	.1378
23	.1754E-03	.2847E-02
24	.1365E-03	.3062E-02
25	.9743	.4690E-01
26	1.117	.8615E-01
27	.9721	.8637E-01
28	.6192	.3245E-01
29	.5738	.3700E-01
30	.6637	.2838E-01
31	.6033	.4363E-01
32	.6915	.6403E-01
33	.6019	.6180E-01
34	.5590	.4746E-01
35	.6408	.6575E-01
36	.5578	.5899E-01
37	.6466	.4529E-01
38	.7411	.6801E-01
39	.6451	.6574E-01
40	.4109	.3056E-01
41	.3808	.3230E-01
42	.4004	.3750E-01
43	.4589	.5026E-01
44	.3995	.4724E-01
45	.3710	.3892E-01
46	.4253	.5068E-01
47	.3702	.4516E-01

The (co)variances of the new parameters are shown on the next five pages.

Parameter pair		Estimate
1.000	1.000	.1108E-07
2.000	1.000	.1008E-12
2.000	2.000	.6737E-08
3.000	1.000	.1915E-10
3.000	2.000	.4104E-10
3.000	3.000	.2550E-05
4.000	1.000	.5023E-10
4.000	2.000	.1301E-09
4.000	3.000	.2607E-07
4.000	4.000	.3376E-05
5.000	1.000	.2033E-09
5.000	2.000	.5259E-09
5.000	3.000	.1054E-06
5.000	4.000	.4037E-06
5.000	5.000	.1498E-04
6.000	1.000	.7565E-09
6.000	2.000	.1951E-08
6.000	3.000	.3915E-06
6.000	4.000	.1497E-05
6.000	5.000	.6050E-05
6.000	6.000	.7339E-04
7.000	1.000	.1865E-08
7.000	2.000	.4787E-08
7.000	3.000	.9630E-06
7.000	4.000	.3673E-05
7.000	5.000	.1485E-04
7.000	6.000	.5505E-04
7.000	7.000	.2708E-03
8.000	1.000	.3637E-08
8.000	2.000	.9308E-08
8.000	3.000	.1880E-05
8.000	4.000	.7168E-05
8.000	5.000	.2898E-04
8.000	6.000	.1074E-03
8.000	7.000	.2638E-03
8.000	8.000	.8278E-03
9.000	1.000	.5930E-08
9.000	2.000	.1513E-07
9.000	3.000	.3074E-05
9.000	4.000	.1172E-04
9.000	5.000	.4738E-04
9.000	6.000	.1757E-03
9.000	7.000	.4314E-03
9.000	8.000	.8424E-03
9.000	9.000	.2060E-02
10.00	1.000	.8360E-08
10.00	2.000	.2118E-07
10.00	3.000	.4328E-05
10.00	4.000	.1644E-04
10.00	5.000	.6648E-04
10.00	6.000	.2465E-03
10.00	7.000	.6055E-03
10.00	8.000	.1183E-02
10.00	9.000	.1937E-02
10.00	10.00	.4182E-02
11.00	1.000	.8399E-08
11.00	2.000	.2111E-07
11.00	3.000	.4331E-05
11.00	4.000	.1637E-04
11.00	5.000	.6620E-04
11.00	6.000	.2455E-03
11.00	7.000	.6030E-03
11.00	8.000	.1178E-02
11.00	9.000	.1930E-02
11.00	10.00	.2714E-02
11.00	11.00	.5102E-02
12.00	1.000	.7753E-08
12.00	2.000	.1937E-07
12.00	3.000	.3987E-05
12.00	4.000	.1502E-04
12.00	5.000	.6073E-04
12.00	6.000	.2252E-03
12.00	7.000	.5533E-03
12.00	8.000	.1081E-02
12.00	9.000	.1772E-02

Parameter pair		Estimate
12.00	10.00	.2493E-02
12.00	11.00	.2488E-02
12.00	12.00	.5763E-02
13.00	1.000	.9546E-08
13.00	2.000	.2363E-07
13.00	3.000	.4861E-05
13.00	4.000	.1818E-04
13.00	5.000	.7350E-04
13.00	6.000	.2726E-03
13.00	7.000	.6697E-03
13.00	8.000	.1309E-02
13.00	9.000	.2145E-02
13.00	10.00	.3019E-02
13.00	11.00	.3014E-02
13.00	12.00	.2771E-02
13.00	13.00	.1001E-01
14.00	1.000	.7876E-08
14.00	2.000	.1939E-07
14.00	3.000	.3987E-05
14.00	4.000	.1487E-04
14.00	5.000	.6011E-04
14.00	6.000	.2229E-03
14.00	7.000	.5478E-03
14.00	8.000	.1071E-02
14.00	9.000	.1755E-02
14.00	10.00	.2471E-02
14.00	11.00	.2468E-02
14.00	12.00	.2269E-02
14.00	13.00	.2752E-02
14.00	14.00	.1111E-01
15.00	1.000	.5368E-08
15.00	2.000	.1316E-07
15.00	3.000	.2715E-05
15.00	4.000	.1011E-04
15.00	5.000	.4087E-04
15.00	6.000	.1516E-03
15.00	7.000	.3726E-03
15.00	8.000	.7289E-03
15.00	9.000	.1196E-02
15.00	10.00	.1685E-02
15.00	11.00	.1684E-02
15.00	12.00	.1549E-02
15.00	13.00	.1879E-02
15.00	14.00	.1543E-02
15.00	15.00	.1002E-01
16.00	1.000	.4513E-08
16.00	2.000	.1112E-07
16.00	3.000	.2291E-05
16.00	4.000	.8550E-05
16.00	5.000	.3457E-04
16.00	6.000	.1282E-03
16.00	7.000	.3152E-03
16.00	8.000	.6166E-03
16.00	9.000	.1012E-02
16.00	10.00	.1426E-02
16.00	11.00	.1425E-02
16.00	12.00	.1311E-02
16.00	13.00	.1591E-02
16.00	14.00	.1307E-02
16.00	15.00	.8973E-03
16.00	16.00	.1072E-01
17.00	1.000	.2590E-08
17.00	2.000	.6385E-08
17.00	3.000	.1327E-05
17.00	4.000	.4970E-05
17.00	5.000	.2010E-04
17.00	6.000	.7453E-04
17.00	7.000	.1832E-03
17.00	8.000	.3586E-03
17.00	9.000	.5886E-03
17.00	10.00	.8299E-03
17.00	11.00	.8298E-03
17.00	12.00	.7637E-03
17.00	13.00	.9264E-03
17.00	14.00	.7609E-03

Parameter pair		Estimate
17.00	15.00	.5224E-03
17.00	16.00	.4433E-03
17.00	17.00	.7353E-02
18.00	1.000	.5199E-08
18.00	2.000	.1277E-07
18.00	3.000	.2690E-05
18.00	4.000	.1009E-04
18.00	5.000	.4078E-04
18.00	6.000	.1513E-03
18.00	7.000	.3719E-03
18.00	8.000	.7283E-03
18.00	9.000	.1197E-02
18.00	10.00	.1689E-02
18.00	11.00	.1691E-02
18.00	12.00	.1557E-02
18.00	13.00	.1890E-02
18.00	14.00	.1552E-02
18.00	15.00	.1067E-02
18.00	16.00	.9056E-03
18.00	17.00	.5285E-03
18.00	18.00	.1882E-01
19.00	1.000	.2520E-08
19.00	2.000	.6198E-08
19.00	3.000	.1310E-05
19.00	4.000	.4913E-05
19.00	5.000	.1986E-04
19.00	6.000	.7368E-04
19.00	7.000	.1812E-03
19.00	8.000	.3548E-03
19.00	9.000	.5834E-03
19.00	10.00	.8240E-03
19.00	11.00	.8253E-03
19.00	12.00	.7605E-03
19.00	13.00	.9229E-03
19.00	14.00	.7582E-03
19.00	15.00	.5215E-03
19.00	16.00	.4430E-03
19.00	17.00	.2586E-03
19.00	18.00	.5312E-03
19.00	19.00	.1066E-01
20.00	1.000	.4216E-08
20.00	2.000	.1043E-07
20.00	3.000	.2188E-05
20.00	4.000	.8210E-05
20.00	5.000	.3320E-04
20.00	6.000	.1231E-03
20.00	7.000	.3027E-03
20.00	8.000	.5927E-03
20.00	9.000	.9743E-03
20.00	10.00	.1376E-02
20.00	11.00	.1378E-02
20.00	12.00	.1270E-02
20.00	13.00	.1541E-02
20.00	14.00	.1267E-02
20.00	15.00	.8718E-03
20.00	16.00	.7409E-03
20.00	17.00	.4324E-03
20.00	18.00	.8879E-03
20.00	19.00	.4357E-03
20.00	20.00	.2103E-01
21.00	1.000	.5293E-08
21.00	2.000	.1323E-07
21.00	3.000	.2770E-05
21.00	4.000	.1046E-04
21.00	5.000	.4230E-04
21.00	6.000	.1569E-03
21.00	7.000	.3856E-03
21.00	8.000	.7550E-03
21.00	9.000	.1241E-02
21.00	10.00	.1753E-02
21.00	11.00	.1756E-02
21.00	12.00	.1618E-02
21.00	13.00	.1964E-02
21.00	14.00	.1616E-02

Parameter pair		Estimate
21.00	15.00	.1112E-02
21.00	16.00	.9457E-03
21.00	17.00	.5521E-03
21.00	18.00	.1134E-02
21.00	19.00	.5568E-03
21.00	20.00	.9328E-03
21.00	21.00	.3459E-01
22.00	1.000	.1984E-08
22.00	2.000	.5114E-08
22.00	3.000	.1086E-05
22.00	4.000	.4212E-05
22.00	5.000	.1703E-04
22.00	6.000	.6313E-04
22.00	7.000	.1551E-03
22.00	8.000	.3038E-03
22.00	9.000	.4999E-03
22.00	10.00	.7063E-03
22.00	11.00	.7074E-03
22.00	12.00	.6518E-03
22.00	13.00	.7911E-03
22.00	14.00	.6510E-03
22.00	15.00	.4480E-03
22.00	16.00	.3809E-03
22.00	17.00	.2228E-03
22.00	18.00	.4566E-03
22.00	19.00	.2254E-03
22.00	20.00	.3777E-03
22.00	21.00	.4850E-03
22.00	22.00	.1898E-01
23.00	1.000	.1675E-11
23.00	2.000	.4298E-11
23.00	3.000	.9379E-09
23.00	4.000	.3677E-08
23.00	5.000	.1487E-07
23.00	6.000	.5512E-07
23.00	7.000	.1355E-06
23.00	8.000	.2655E-06
23.00	9.000	.4376E-06
23.00	10.00	.6194E-06
23.00	11.00	.6209E-06
23.00	12.00	.5725E-06
23.00	13.00	.6955E-06
23.00	14.00	.5735E-06
23.00	15.00	.3950E-06
23.00	16.00	.3358E-06
23.00	17.00	.1969E-06
23.00	18.00	.4066E-06
23.00	19.00	.2001E-06
23.00	20.00	.3353E-06
23.00	21.00	.4317E-06
23.00	22.00	.1774E-06
23.00	23.00	.8103E-05
24.00	1.000	.1332E-11
24.00	2.000	.3752E-11
24.00	3.000	.7872E-09
24.00	4.000	.3227E-08
24.00	5.000	.1304E-07
24.00	6.000	.4831E-07
24.00	7.000	.1186E-06
24.00	8.000	.2320E-06
24.00	9.000	.3814E-06
24.00	10.00	.5384E-06
24.00	11.00	.5383E-06
24.00	12.00	.4954E-06
24.00	13.00	.6011E-06
24.00	14.00	.4957E-06
24.00	15.00	.3406E-06
24.00	16.00	.2899E-05
24.00	17.00	.1697E-06
24.00	18.00	.3492E-06
24.00	19.00	.1717E-06
24.00	20.00	.2888E-06
24.00	21.00	.3733E-06
24.00	22.00	.1541E-06
24.00	23.00	.1393E-09
24.00	24.00	.9376E-05

Parameter pair		Estimate
25.00	1.000	-.1526E-08
25.00	2.000	-.1116E-07
25.00	3.000	-.1710E-05
25.00	4.000	-.9938E-05
25.00	5.000	-.4011E-04
25.00	6.000	-.1479E-03
25.00	7.000	-.3603E-03
25.00	8.000	-.6984E-03
25.00	9.000	-.1131E-02
25.00	10.00	-.1558E-02
25.00	11.00	-.1522E-02
25.00	12.00	-.1376E-02
25.00	13.00	-.1641E-02
25.00	14.00	-.1345E-02
25.00	15.00	-.9256E-03
25.00	16.00	-.7929E-03
25.00	17.00	-.4532E-03
25.00	18.00	-.8906E-03
25.00	19.00	-.4299E-03
25.00	20.00	-.7337E-03
25.00	21.00	-.9529E-03
25.00	22.00	-.3917E-03
25.00	23.00	-.3298E-06
25.00	24.00	-.3385E-06
25.00	25.00	.2200E-02
26.00	1.000	-.1091E-08
26.00	2.000	-.1255E-07
26.00	3.000	-.1822E-05
26.00	4.000	-.1087E-04
26.00	5.000	-.4380E-04
26.00	6.000	-.1610E-03
26.00	7.000	-.3875E-03
26.00	8.000	-.7378E-03
26.00	9.000	-.1185E-02
26.00	10.00	-.1633E-02
26.00	11.00	-.1598E-02
26.00	12.00	-.1442E-02
26.00	13.00	-.1754E-02
26.00	14.00	-.1437E-02
26.00	15.00	-.8974E-03
26.00	16.00	-.7510E-03
26.00	17.00	-.4449E-03
26.00	18.00	-.8987E-03
26.00	19.00	-.4420E-03
26.00	20.00	-.7787E-03
26.00	21.00	-.1073E-02
26.00	22.00	-.4849E-03
26.00	23.00	-.4586E-06
26.00	24.00	-.4988E-06
26.00	25.00	.1810E-02
26.00	26.00	.7422E-02
27.00	1.000	.1193E-08
27.00	2.000	-.8328E-08
27.00	3.000	-.1236E-05
27.00	4.000	-.8336E-05
27.00	5.000	-.3365E-04
27.00	6.000	-.1237E-03
27.00	7.000	-.2996E-03
27.00	8.000	-.5789E-03
27.00	9.000	-.9381E-03
27.00	10.00	-.1308E-02
27.00	11.00	-.1308E-02
27.00	12.00	-.1198E-02
27.00	13.00	-.1396E-02
27.00	14.00	-.1058E-02
27.00	15.00	-.7003E-03
27.00	16.00	-.6172E-03
27.00	17.00	-.3631E-03
27.00	18.00	-.7677E-03
27.00	19.00	-.3938E-03
27.00	20.00	-.6614E-03
27.00	21.00	-.8305E-03
27.00	22.00	-.3413E-03
27.00	23.00	-.2423E-06
27.00	24.00	-.2519E-06

Parameter pair		Estimate
27.00	25.00	.1513E-02
27.00	26.00	.1738E-02
27.00	27.00	.7459E-02
28.00	1.000	-.4804E-08
28.00	2.000	-.8728E-08
28.00	3.000	-.2184E-05
28.00	4.000	-.7102E-05
28.00	5.000	-.2876E-04
28.00	6.000	-.1071E-03
28.00	7.000	-.2647E-03
28.00	8.000	-.5194E-03
28.00	9.000	-.8544E-03
28.00	10.00	-.1208E-02
28.00	11.00	-.1210E-02
28.00	12.00	-.1116E-02
28.00	13.00	-.1353E-02
28.00	14.00	-.1095E-02
28.00	15.00	-.7305E-03
28.00	16.00	-.6065E-03
28.00	17.00	-.3570E-03
28.00	18.00	-.7335E-03
28.00	19.00	-.3545E-03
28.00	20.00	-.5745E-03
28.00	21.00	-.7109E-03
28.00	22.00	-.2812E-03
28.00	23.00	-.2502E-06
28.00	24.00	-.1874E-06
28.00	25.00	.5035E-04
28.00	26.00	-.4336E-04
28.00	27.00	-.2936E-04
28.00	28.00	.1053E-02
29.00	1.000	-.8492E-08
29.00	2.000	-.1581E-07
29.00	3.000	-.2940E-05
29.00	4.000	-.7496E-05
29.00	5.000	-.3034E-04
29.00	6.000	-.1129E-03
29.00	7.000	-.2782E-03
29.00	8.000	-.5414E-03
29.00	9.000	-.8789E-03
29.00	10.00	-.1234E-02
29.00	11.00	-.1236E-02
29.00	12.00	-.1138E-02
29.00	13.00	-.1402E-02
29.00	14.00	-.1153E-02
29.00	15.00	-.7784E-03
29.00	16.00	-.6511E-03
29.00	17.00	-.3719E-03
29.00	18.00	-.7387E-03
29.00	19.00	-.3555E-03
29.00	20.00	-.5936E-03
29.00	21.00	-.7400E-03
29.00	22.00	-.2717E-03
29.00	23.00	-.2238E-06
29.00	24.00	-.1761E-06
29.00	25.00	.2042E-03
29.00	26.00	.1334E-03
29.00	27.00	-.2425E-03
29.00	28.00	.7609E-03
29.00	29.00	.1369E-02
30.00	1.000	-.2387E-08
30.00	2.000	-.4892E-08
30.00	3.000	-.1238E-05
30.00	4.000	-.4301E-05
30.00	5.000	-.1740E-04
30.00	6.000	-.6461E-04
30.00	7.000	-.1600E-03
30.00	8.000	-.3187E-03
30.00	9.000	-.5376E-03
30.00	10.00	-.7827E-03
30.00	11.00	-.8044E-03
30.00	12.00	-.7551E-03
30.00	13.00	-.9298E-03
30.00	14.00	-.7860E-03
30.00	15.00	-.5709E-03

Parameter pair		Estimate
30.00	16.00	-.4956E-03
30.00	17.00	-.2930E-03
30.00	18.00	-.6268E-03
30.00	19.00	-.3160E-03
30.00	20.00	-.5347E-03
30.00	21.00	-.6974E-03
30.00	22.00	-.2890E-03
30.00	23.00	-.2715E-06
30.00	24.00	-.2251E-06
30.00	25.00	.2454E-03
30.00	26.00	.2371E-03
30.00	27.00	.2129E-03
30.00	28.00	.1961E-03
30.00	29.00	.2299E-03
30.00	30.00	.8056E-03
31.00	1.000	-.5625E-08
31.00	2.000	-.1541E-07
31.00	3.000	-.3186E-05
31.00	4.000	-.1307E-04
31.00	5.000	-.5285E-04
31.00	6.000	-.1959E-03
31.00	7.000	-.4810E-03
31.00	8.000	-.9385E-03
31.00	9.000	-.1532E-02
31.00	10.00	-.2141E-02
31.00	11.00	-.2121E-02
31.00	12.00	-.1939E-02
31.00	13.00	-.2334E-02
31.00	14.00	-.1900E-02
31.00	15.00	-.1285E-02
31.00	16.00	-.1082E-02
31.00	17.00	-.6265E-03
31.00	18.00	-.1266E-02
31.00	19.00	-.6115E-03
31.00	20.00	-.1014E-02
31.00	21.00	-.1283E-02
31.00	22.00	-.5165E-03
31.00	23.00	-.4480E-06
31.00	24.00	-.3922E-06
31.00	25.00	.1411E-02
31.00	26.00	.1079E-02
31.00	27.00	.9084E-03
31.00	28.00	.1057E-02
31.00	29.00	.8678E-03
31.00	30.00	.3431E-03
31.00	31.00	.1904E-02
32.00	1.000	-.6040E-08
32.00	2.000	-.1752E-07
32.00	3.000	-.3567E-05
32.00	4.000	-.1466E-04
32.00	5.000	-.5923E-04
32.00	6.000	-.2193E-03
32.00	7.000	-.5356E-03
32.00	8.000	-.1037E-02
32.00	9.000	-.1688E-02
32.00	10.00	-.2359E-02
32.00	11.00	-.2340E-02
32.00	12.00	-.2139E-02
32.00	13.00	-.2597E-02
32.00	14.00	-.2113E-02
32.00	15.00	-.1371E-02
32.00	16.00	-.1142E-02
32.00	17.00	-.6742E-03
32.00	18.00	-.1376E-02
32.00	19.00	-.6696E-03
32.00	20.00	-.1124E-02
32.00	21.00	-.1459E-02
32.00	22.00	-.6144E-03
32.00	23.00	-.5634E-06
32.00	24.00	-.5182E-06
32.00	25.00	.1177E-02
32.00	26.00	.4548E-02
32.00	27.00	.1044E-02
32.00	28.00	.1149E-02
32.00	29.00	.9324E-03
32.00	30.00	.3659E-03
32.00	31.00	.1848E-02
32.00	32.00	.4099E-02
33.00	1.000	-.3931E-08
33.00	2.000	-.1364E-07
33.00	3.000	-.2888E-05
33.00	4.000	-.1207E-04
33.00	5.000	-.4879E-04
33.00	6.000	-.1807E-03
33.00	7.000	-.4428E-03
33.00	8.000	-.8634E-03
33.00	9.000	-.1411E-02
33.00	10.00	-.1984E-02
33.00	11.00	-.1986E-02
33.00	12.00	-.1826E-02
33.00	13.00	-.2180E-02
33.00	14.00	-.1719E-02
33.00	15.00	-.1144E-02
33.00	16.00	-.9718E-03
33.00	17.00	-.5719E-03
33.00	18.00	-.1188E-02
33.00	19.00	-.5884E-03
33.00	20.00	-.9680E-03
33.00	21.00	-.1205E-02
33.00	22.00	-.4847E-03
33.00	23.00	-.3932E-06
33.00	24.00	-.3382E-06
33.00	25.00	.9860E-03
33.00	26.00	.1034E-02
33.00	27.00	.4591E-02
33.00	28.00	.1005E-02
33.00	29.00	.5895E-03
33.00	30.00	.3225E-03
33.00	31.00	.1590E-02
33.00	32.00	.1763E-02
33.00	33.00	.3820E-02
34.00	1.000	-.9149E-08
34.00	2.000	-.2181E-07
34.00	3.000	-.3845E-05
34.00	4.000	-.1301E-04
34.00	5.000	-.5258E-04
34.00	6.000	-.1949E-03
34.00	7.000	-.4778E-03
34.00	8.000	-.9282E-03
34.00	9.000	-.1505E-02
34.00	10.00	-.2096E-02
34.00	11.00	-.2078E-02
34.00	12.00	-.1899E-02
34.00	13.00	-.2307E-02
34.00	14.00	-.1895E-02
34.00	15.00	-.1289E-02
34.00	16.00	-.1089E-02
34.00	17.00	-.6224E-03
34.00	18.00	-.1231E-02
34.00	19.00	-.5931E-03
34.00	20.00	-.9993E-03
34.00	21.00	-.1268E-02
34.00	22.00	-.4895E-03
34.00	23.00	-.4073E-06
34.00	24.00	-.3658E-06
34.00	25.00	.1461E-02
34.00	26.00	.1169E-02
34.00	27.00	.6321E-03
34.00	28.00	.7702E-03
34.00	29.00	.1451E-02
34.00	30.00	.3649E-03
34.00	31.00	.1655E-02
34.00	32.00	.1584E-02
34.00	33.00	.1140E-02
34.00	34.00	.2252E-02
35.00	1.000	-.1011E-07
35.00	2.000	-.2486E-07
35.00	3.000	-.4328E-05
35.00	4.000	-.1461E-04
35.00	5.000	-.5901E-04
35.00	6.000	-.2185E-03
35.00	7.000	-.5331E-03
35.00	8.000	-.1028E-02
35.00	9.000	-.1662E-02
35.00	10.00	-.2315E-02
35.00	11.00	-.2297E-02
35.00	12.00	-.2099E-02
35.00	13.00	-.2572E-02
35.00	14.00	-.2112E-02
35.00	15.00	-.1384E-02
35.00	16.00	-.1158E-02
35.00	17.00	-.6706E-03
35.00	18.00	-.1341E-02
35.00	19.00	-.6507E-03
35.00	20.00	-.1110E-02
35.00	21.00	-.1442E-02
35.00	22.00	-.5817E-03
35.00	23.00	-.5131E-06
35.00	24.00	-.4829E-06
35.00	25.00	.1267E-02
35.00	26.00	.4408E-02
35.00	27.00	.7268E-03
35.00	28.00	.8249E-03
35.00	29.00	.1605E-02
35.00	30.00	.3928E-03
35.00	31.00	.1588E-02
35.00	32.00	.3651E-02
35.00	33.00	.1252E-02
35.00	34.00	.2291E-02
35.00	35.00	.4322E-02
36.00	1.000	-.7570E-08
36.00	2.000	-.2015E-07
36.00	3.000	-.3567E-05
36.00	4.000	-.1207E-04
36.00	5.000	-.4880E-04
36.00	6.000	-.1807E-03
36.00	7.000	-.4424E-03
36.00	8.000	-.8584E-03
36.00	9.000	-.1393E-02
36.00	10.00	-.1950E-02
36.00	11.00	-.1952E-02
36.00	12.00	-.1794E-02
36.00	13.00	-.2164E-02
36.00	14.00	-.1727E-02
36.00	15.00	-.1159E-02
36.00	16.00	-.9871E-03
36.00	17.00	-.5699E-03
36.00	18.00	-.1159E-02
36.00	19.00	-.5716E-03
36.00	20.00	-.9565E-03
36.00	21.00	-.1196E-02
36.00	22.00	-.4599E-03
36.00	23.00	-.3566E-06
36.00	24.00	-.3157E-06
36.00	25.00	.1067E-02
36.00	26.00	.1127E-02
36.00	27.00	.4045E-02
36.00	28.00	.7228E-03
36.00	29.00	.1192E-02
36.00	30.00	.3457E-03
36.00	31.00	.1365E-02
36.00	32.00	.1505E-02
36.00	33.00	.3207E-02
36.00	34.00	.1773E-02
36.00	35.00	.1978E-02
36.00	36.00	.3479E-02
37.00	1.000	-.3319E-08
37.00	2.000	-.1217E-07
37.00	3.000	-.2341E-05
37.00	4.000	-.1079E-04
37.00	5.000	-.4357E-04
37.00	6.000	-.1611E-03
37.00	7.000	-.3950E-03
37.00	8.000	-.7740E-03
37.00	9.000	-.1274E-02

B4. (Continued)

Parameter pair		Estimate
37.00	10.00	-.1797E-02
37.00	11.00	-.1794E-02
37.00	12.00	-.1649E-02
37.00	13.00	-.1995E-02
37.00	14.00	-.1659E-02
37.00	15.00	-.1170E-02
37.00	16.00	-.1009E-02
37.00	17.00	-.5862E-03
37.00	18.00	-.1202E-02
37.00	19.00	-.5932E-03
37.00	20.00	-.1008E-02
37.00	21.00	-.1312E-02
37.00	22.00	-.5415E-03
37.00	23.00	-.4834E-06
37.00	24.00	-.4440E-06
37.00	25.00	.1699E-02
37.00	26.00	.1432E-02
37.00	27.00	.1212E-02
37.00	28.00	.2245E-03
37.00	29.00	.3596E-03
37.00	30.00	.9477E-03
37.00	31.00	.1271E-02
37.00	32.00	.1138E-02
37.00	33.00	.9685E-03
37.00	34.00	.1325E-02
37.00	35.00	.1224E-02
37.00	36.00	.1045E-02
37.00	37.00	.2051E-02
38.00	1.000	-.3368E-08
38.00	2.000	-.1379E-07
38.00	3.000	-.2592E-05
38.00	4.000	-.1201E-04
38.00	5.000	-.4849E-04
38.00	6.000	-.1790E-03
38.00	7.000	-.4359E-03
38.00	8.000	-.8455E-03
38.00	9.000	-.1387E-02
38.00	10.00	-.1958E-02
38.00	11.00	-.1959E-02
38.00	12.00	-.1800E-02
38.00	13.00	-.2202E-02
38.00	14.00	-.1831E-02
38.00	15.00	-.1233E-02
38.00	16.00	-.1052E-02
38.00	17.00	-.6225E-03
38.00	18.00	-.1296E-02
38.00	19.00	-.6463E-03
38.00	20.00	-.1114E-02
38.00	21.00	-.1491E-02
38.00	22.00	-.6445E-03
38.00	23.00	-.6076E-06
38.00	24.00	-.5825E-06
38.00	25.00	.1476E-02
38.00	26.00	.5191E-02
38.00	27.00	.1392E-02
38.00	28.00	.1903E-03
38.00	29.00	.3453E-03
38.00	30.00	.1057E-02
38.00	31.00	.1099E-02
38.00	32.00	.3427E-02
38.00	33.00	.1047E-02
38.00	34.00	.1183E-02
38.00	35.00	.3364E-02
38.00	36.00	.1134E-02
38.00	37.00	.2009E-02
38.00	38.00	.4625E-02
39.00	1.000	-.1509E-08
39.00	2.000	-.1028E-07
39.00	3.000	-.2024E-05
39.00	4.000	-.9713E-05
39.00	5.000	-.3924E-04
39.00	6.000	-.1449E-03
39.00	7.000	-.3544E-03
39.00	8.000	-.6940E-03
39.00	9.000	-.1145E-02

Parameter pair		Estimate
39.00	10.00	-.1629E-02
39.00	11.00	-.1650E-02
39.00	12.00	-.1529E-02
39.00	13.00	-.1830E-02
39.00	14.00	-.1466E-02
39.00	15.00	-.1020E-02
39.00	16.00	-.8913E-03
39.00	17.00	-.5258E-03
39.00	18.00	-.1119E-02
39.00	19.00	-.5686E-03
39.00	20.00	-.9587E-03
39.00	21.00	-.1229E-02
39.00	22.00	-.5074E-03
39.00	23.00	-.4247E-06
39.00	24.00	-.3860E-06
39.00	25.00	.1243E-02
39.00	26.00	.1384E-02
39.00	27.00	.5157E-02
39.00	28.00	.1712E-03
39.00	29.00	.6257E-04
39.00	30.00	.9243E-03
39.00	31.00	.9364E-03
39.00	32.00	.1048E-02
39.00	33.00	.3360E-02
39.00	34.00	.7741E-03
39.00	35.00	.8642E-03
39.00	36.00	.3020E-02
39.00	37.00	.1725E-02
39.00	38.00	.1951E-02
39.00	39.00	.4321E-02
40.00	1.000	-.4654E-08
40.00	2.000	-.8822E-08
40.00	3.000	-.2216E-05
40.00	4.000	-.7376E-05
40.00	5.000	-.2986E-04
40.00	6.000	-.1111E-03
40.00	7.000	-.2747E-03
40.00	8.000	-.5421E-03
40.00	9.000	-.8999E-03
40.00	10.00	-.1286E-02
40.00	11.00	-.1301E-02
40.00	12.00	-.1208E-02
40.00	13.00	-.1474E-02
40.00	14.00	-.1213E-02
40.00	15.00	-.8383E-03
40.00	16.00	-.7094E-03
40.00	17.00	-.4184E-03
40.00	18.00	-.8749E-03
40.00	19.00	-.4309E-03
40.00	20.00	-.7123E-03
40.00	21.00	-.9037E-03
40.00	22.00	-.3656E-03
40.00	23.00	-.3342E-06
40.00	24.00	-.2638E-06
40.00	25.00	.1854E-03
40.00	26.00	.1180E-03
40.00	27.00	.1123E-03
40.00	28.00	.8203E-03
40.00	29.00	.6474E-03
40.00	30.00	.6290E-03
40.00	31.00	.9139E-03
40.00	32.00	.9891E-03
40.00	33.00	.8669E-03
40.00	34.00	.7371E-03
40.00	35.00	.7907E-03
40.00	36.00	.6938E-03
40.00	37.00	.7358E-03
40.00	38.00	.7808E-03
40.00	39.00	.6860E-03
40.00	40.00	.9339E-03
41.00	1.000	-.6994E-08
41.00	2.000	-.1330E-07
41.00	3.000	-.2661E-05
41.00	4.000	-.7443E-05
41.00	5.000	-.3012E-04

Parameter pair		Estimate
41.00	6.000	-.1120E-03
41.00	7.000	-.2765E-03
41.00	8.000	-.5422E-03
41.00	9.000	-.8918E-03
41.00	10.00	-.1268E-02
41.00	11.00	-.1282E-02
41.00	12.00	-.1189E-02
41.00	13.00	-.1464E-02
41.00	14.00	-.1216E-02
41.00	15.00	-.8442E-03
41.00	16.00	-.7165E-03
41.00	17.00	-.4150E-03
41.00	18.00	-.8499E-03
41.00	19.00	-.4173E-03
41.00	20.00	-.7007E-03
41.00	21.00	-.8913E-03
41.00	22.00	-.3461E-03
41.00	23.00	-.3043E-06
41.00	24.00	-.2460E-06
41.00	25.00	.2764E-03
41.00	26.00	.2246E-03
41.00	27.00	-.3877E-04
41.00	28.00	.6176E-03
41.00	29.00	.1041E-02
41.00	30.00	.6148E-03
41.00	31.00	.7728E-03
41.00	32.00	.8287E-03
41.00	33.00	.5763E-03
41.00	34.00	.1172E-02
41.00	35.00	.1291E-02
41.00	36.00	.9892E-03
41.00	37.00	.7824E-03
41.00	38.00	.8357E-03
41.00	39.00	.5719E-03
41.00	40.00	.7906E-03
41.00	41.00	.1043E-02
42.00	1.000	-.5161E-08
42.00	2.000	-.1318E-07
42.00	3.000	-.2861E-05
42.00	4.000	-.1127E-04
42.00	5.000	-.4557E-04
42.00	6.000	-.1690E-03
42.00	7.000	-.4157E-03
42.00	8.000	-.8151E-03
42.00	9.000	-.1341E-02
42.00	10.00	-.1893E-02
42.00	11.00	-.1893E-02
42.00	12.00	-.1743E-02
42.00	13.00	-.2110E-02
42.00	14.00	-.1735E-02
42.00	15.00	-.1197E-02
42.00	16.00	-.1017E-02
42.00	17.00	-.5939E-03
42.00	18.00	-.1218E-02
42.00	19.00	-.5965E-03
42.00	20.00	-.9955E-03
42.00	21.00	-.1272E-02
42.00	22.00	-.5171E-03
42.00	23.00	-.4611E-06
42.00	24.00	-.3961E-06
42.00	25.00	.1085E-02
42.00	26.00	.8590E-03
42.00	27.00	.7313E-03
42.00	28.00	.8198E-03
42.00	29.00	.7146E-03
42.00	30.00	.7137E-03
42.00	31.00	.1470E-02
42.00	32.00	.1447E-02
42.00	33.00	.1250E-02
42.00	34.00	.1319E-02
42.00	35.00	.1291E-02
42.00	36.00	.1114E-02
42.00	37.00	.1415E-02
42.00	38.00	.1367E-02
42.00	39.00	.1179E-02

Parameter pair		Estimate
42.00	40.00	.9860E-03
42.00	41.00	.8838E-03
42.00	42.00	.1406E-02
43.00	1.000	-.5646E-08
43.00	2.000	-.1501E-07
43.00	3.000	-.3223E-05
43.00	4.000	-.1270E-04
43.00	5.000	-.5134E-04
43.00	6.000	-.1902E-03
43.00	7.000	-.4661E-03
43.00	8.000	-.9085E-03
43.00	9.000	-.1492E-02
43.00	10.00	-.2107E-02
43.00	11.00	-.2109E-02
43.00	12.00	-.1942E-02
43.00	13.00	-.2367E-02
43.00	14.00	-.1946E-02
43.00	15.00	-.1305E-02
43.00	16.00	-.1101E-02
43.00	17.00	-.6501E-03
43.00	18.00	-.1346E-02
43.00	19.00	-.6629E-03
43.00	20.00	-.1116E-02
43.00	21.00	-.1450E-02
43.00	22.00	-.6076E-03
43.00	23.00	-.5617E-06
43.00	24.00	-.4996E-06
43.00	25.00	.9510E-03
43.00	26.00	.3182E-02
43.00	27.00	.8399E-03
43.00	28.00	.8982E-03
43.00	29.00	.7778E-03
43.00	30.00	.7999E-03
43.00	31.00	.1464E-02
43.00	32.00	.2973E-02
43.00	33.00	.1393E-02
43.00	34.00	.1303E-02
43.00	35.00	.2695E-02
43.00	36.00	.1238E-02
43.00	37.00	.1410E-02
43.00	38.00	.3005E-02
43.00	39.00	.1335E-02
43.00	40.00	.1091E-02
43.00	41.00	.9752E-03
43.00	42.00	.1454E-02
43.00	43.00	.2526E-02
44.00	1.000	-.4034E-08
44.00	2.000	-.1200E-07
44.00	3.000	-.2662E-05
44.00	4.000	-.1060E-04
44.00	5.000	-.4285E-04
44.00	6.000	-.1588E-03
44.00	7.000	-.3902E-03
44.00	8.000	-.7648E-03
44.00	9.000	-.1260E-02
44.00	10.00	-.1788E-02
44.00	11.00	-.1802E-02
44.00	12.00	-.1667E-02
44.00	13.00	-.2006E-02
44.00	14.00	-.1614E-02
44.00	15.00	-.1103E-02
44.00	16.00	-.9432E-03
44.00	17.00	-.5559E-03
44.00	18.00	-.1166E-02
44.00	19.00	-.5807E-03
44.00	20.00	-.9643E-03
44.00	21.00	-.1220E-02
44.00	22.00	-.4956E-03
44.00	23.00	-.4244E-06
44.00	24.00	-.3599E-06
44.00	25.00	.8021E-03
44.00	26.00	.8292E-03
44.00	27.00	.3175E-02
44.00	28.00	.7853E-03
44.00	29.00	.5296E-03
44.00	30.00	.6989E-03
44.00	31.00	.1262E-02
44.00	32.00	.1390E-02
44.00	33.00	.2729E-02
44.00	34.00	.9763E-03
44.00	35.00	.1067E-02
44.00	36.00	.2337E-02
44.00	37.00	.1213E-02
44.00	38.00	.1331E-02
44.00	39.00	.2786E-02
44.00	40.00	.9539E-03
44.00	41.00	.7525E-03
44.00	42.00	.1259E-02
44.00	43.00	.1406E-02
44.00	44.00	.2232E-02
45.00	1.000	-.7395E-08
45.00	2.000	-.1721E-07
45.00	3.000	-.3244E-05
45.00	4.000	-.1104E-04
45.00	5.000	-.4462E-04
45.00	6.000	-.1654E-03
45.00	7.000	-.4066E-03
45.00	8.000	-.7942E-03
45.00	9.000	-.1299E-02
45.00	10.00	-.1829E-02
45.00	11.00	-.1828E-02
45.00	12.00	-.1682E-02
45.00	13.00	-.2051E-02
45.00	14.00	-.1697E-02
45.00	15.00	-.1175E-02
45.00	16.00	-.1000E-02
45.00	17.00	-.5768E-03
45.00	18.00	-.1167E-02
45.00	19.00	-.5703E-03
45.00	20.00	-.9621E-03
45.00	21.00	-.1231E-02
45.00	22.00	-.4864E-03
45.00	23.00	-.4221E-06
45.00	24.00	-.3686E-06
45.00	25.00	.1107E-02
45.00	26.00	.9082E-03
45.00	27.00	.5385E-03
45.00	28.00	.6208E-03
45.00	29.00	.1092E-02
45.00	30.00	.6925E-03
45.00	31.00	.1290E-02
45.00	32.00	.1256E-02
45.00	33.00	.9369E-03
45.00	34.00	.1699E-02
45.00	35.00	.1740E-02
45.00	36.00	.1370E-02
45.00	37.00	.1409E-02
45.00	38.00	.1376E-02
45.00	39.00	.1031E-02
45.00	40.00	.8408E-03
45.00	41.00	.1122E-02
45.00	42.00	.1274E-02
45.00	43.00	.1312E-02
45.00	44.00	.1039E-02
45.00	45.00	.1514E-02
46.00	1.000	-.8226E-08
46.00	2.000	-.1963E-07
46.00	3.000	-.3666E-05
46.00	4.000	-.1245E-04
46.00	5.000	-.5031E-04
46.00	6.000	-.1864E-03
46.00	7.000	-.4563E-03
46.00	8.000	-.8864E-03
46.00	9.000	-.1447E-02
46.00	10.00	-.2038E-02
46.00	11.00	-.2040E-02
46.00	12.00	-.1877E-02
46.00	13.00	-.2303E-02
46.00	14.00	-.1905E-02
46.00	15.00	-.1284E-02
46.00	16.00	-.1086E-02
46.00	17.00	-.6328E-03
46.00	18.00	-.1291E-02
46.00	19.00	-.6344E-03
46.00	20.00	-.1079E-02
46.00	21.00	-.1404E-02
46.00	22.00	-.5712E-03
46.00	23.00	-.5145E-06
46.00	24.00	-.4647E-06
46.00	25.00	.9981E-03
46.00	26.00	.3077E-02
46.00	27.00	.6188E-03
46.00	28.00	.6731E-03
46.00	29.00	.1213E-02
46.00	30.00	.7769E-03
46.00	31.00	.1274E-02
46.00	32.00	.2657E-02
46.00	33.00	.1037E-02
46.00	34.00	.1754E-02
46.00	35.00	.3120E-02
46.00	36.00	.1534E-02
46.00	37.00	.1419E-02
46.00	38.00	.2910E-02
46.00	39.00	.1166E-02
46.00	40.00	.9278E-03
46.00	41.00	.1251E-02
46.00	42.00	.1314E-02
46.00	43.00	.2301E-02
46.00	44.00	.1156E-02
46.00	45.00	.1599E-02
46.00	46.00	.2569E-02
47.00	1.000	-.6344E-08
47.00	2.000	-.1610E-07
47.00	3.000	-.3058E-05
47.00	4.000	-.1041E-04
47.00	5.000	-.4209E-04
47.00	6.000	-.1560E-03
47.00	7.000	-.3828E-03
47.00	8.000	-.7475E-03
47.00	9.000	-.1224E-02
47.00	10.00	-.1731E-02
47.00	11.00	-.1744E-02
47.00	12.00	-.1612E-02
47.00	13.00	-.1955E-02
47.00	14.00	-.1585E-02
47.00	15.00	-.1087E-02
47.00	16.00	-.9315E-03
47.00	17.00	-.5416E-03
47.00	18.00	-.1119E-02
47.00	19.00	-.5556E-03
47.00	20.00	-.9331E-03
47.00	21.00	-.1183E-02
47.00	22.00	-.4664E-03
47.00	23.00	-.3881E-06
47.00	24.00	-.3351E-06
47.00	25.00	.8449E-03
47.00	26.00	.8804E-03
47.00	27.00	.2803E-02
47.00	28.00	.5891E-03
47.00	29.00	.9191E-03
47.00	30.00	.6787E-03
47.00	31.00	.1097E-02
47.00	32.00	.1203E-02
47.00	33.00	.2308E-02
47.00	34.00	.1380E-02
47.00	35.00	.1532E-02
47.00	36.00	.2502E-02
47.00	37.00	.1222E-02
47.00	38.00	.1342E-02
47.00	39.00	.2520E-02
47.00	40.00	.8113E-03
47.00	41.00	.9994E-03
47.00	42.00	.1138E-02
47.00	43.00	.1268E-02
47.00	44.00	.1941E-02
47.00	45.00	.1295E-02
47.00	46.00	.1451E-02
47.00	47.00	.2039E-02

B5. GLIM3-programme for fitting the model T^* + TYPE*(ZR + ZE + ZC)

```
$UNITS 804
$DATA 402 BT N D1 D2 W REL EDU COH
$DINPUT 1
$CALC BT=BT-14
     : BT=%IF(%EQ(EDU,1)*%NE(REL,1),BT+3,BT)
     : BT=%IF(%EQ(EDU,2)*%NE(REL,1),BT+3,BT)
     : BT=%IF(%EQ(EDU,1)*%EQ(REL,1),BT+2,BT)
     : BT=%IF(%EQ(EDU,2)*%EQ(REL,1),BT+2,BT)
     : BT=%IF(%EQ(EDU,3)*%NE(REL,1),BT+1,BT)
$FACTOR T 24 ZR 4 ZE 3 ZC 2 TYPE 2
$CALC TYPE=%GL(2,402) : I2=%GL(402,1) : I1=%GL(804,1)
: LE=N(I2)-(W(I2)+D1(I2)+D2(I2))/2 : LE=%LOG(LE)
: D=0 : D(I1)=%LE(I1,402)*D1(I2)+%GT(I1,402)*D2(I2)
: T(I1)=BT(I2) : ZR(I1)=REL(I2)
: ZE(I1)=EDU(I2) : ZC(I1)=COH(I2)
$DEL BT N D1 D2 W REL EDU COH I1 I2
$CALC %S=0. : %S=%S-2*D*(%LOG(D)-LE)-D
$PRINT : : : "  SATURATED MODEL T*REL*EDU*COH "
         : : "  HAS -2*LOG(LIKELIHOOD) =       " *9 %S : : : :
$YVAR D
$ERROR P
$OFFSET LE
$FIT T+TYPE*(ZR+ZE+ZC)
$DISPLAY L A
$CALC %L=%S+%DV
$PRINT : : : "  THE CURRENT MODEL HAS "
         : : "  -2*LOG(LIKELIHOOD) =               " *9 %L
         : : "  LIKELIHOOD RATIO CHI SQUARE =      " *9 %DV
         : : "  PEARSON CHI SQUARE =               " *9 %X2
         : : "  WITH DEGREES OF FREEDOM =          " *3 %DF
$STOP
```

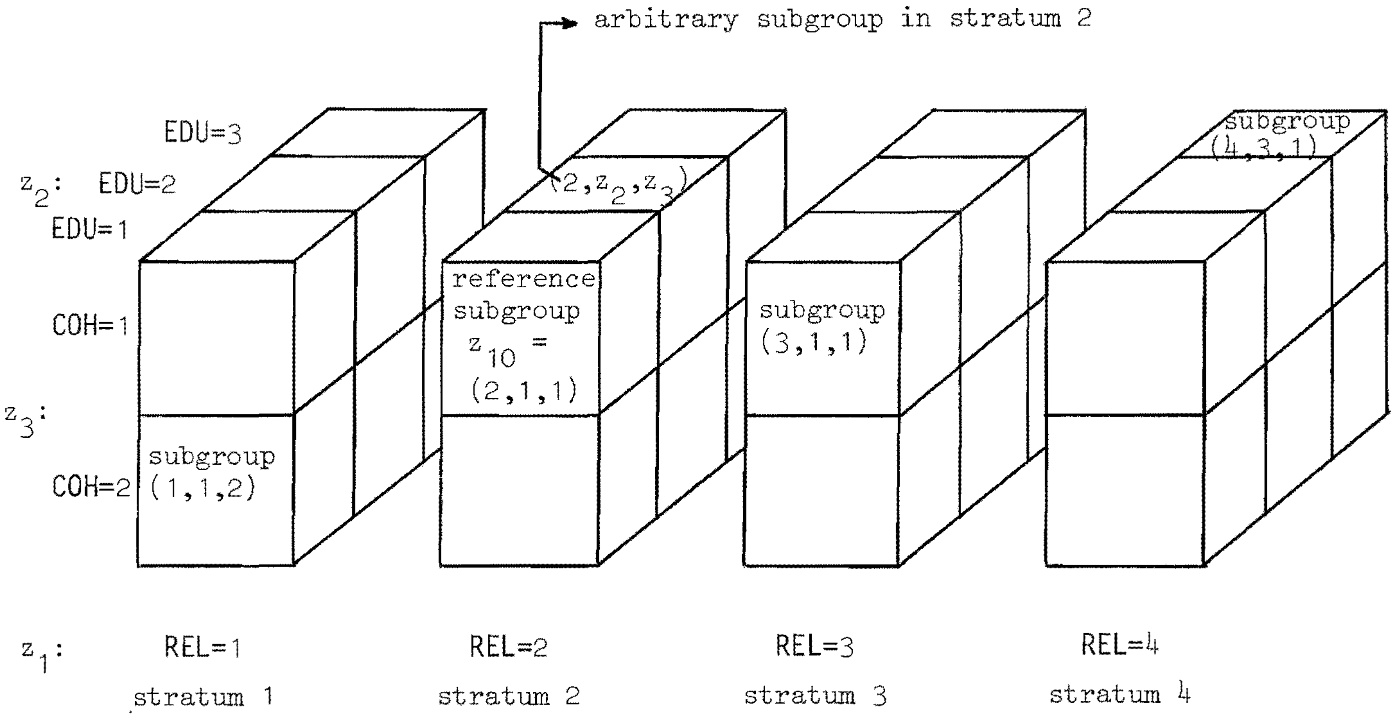
B5. (continued) GLIM3-output after fitting the model T* + TYPE*(ZR + ZE + ZC)

```
GLIM 3.11 (C)1977 ROYAL STATISTICAL SOCIETY, LONDON
----- $DATA LIST ABOLISHED
----- INVALID FUNCTION/OPERATOR ARGUMENT(S)


  SATURATED MODEL T*REL*EDU*COH
  HAS -2*LOG(LIKELIHOOD) =          4764.94422



         SCALED
 CYCLE  DEVIANCE      DF
   8     680.5       767

LINEAR PREDICTOR
%GM T TYPE ZR ZE ZC TYPE.ZR TYPE.ZE TYPE.ZC

        ESTIMATE        S.E.        PARAMETER
    1   -11.26         16.11        %GM
    0     ZERO       ALIASED        T(1)
    2    .1649         18.13        T(2)
    3    4.785         16.14        T(3)
    4    6.018         16.12        T(4)
    5    7.414         16.11        T(5)
    6    8.718         16.11        T(6)
    7    9.602         16.11        T(7)
    8    10.25         16.11        T(8)
    9    10.72         16.11        T(9)
   10    11.05         16.11        T(10)
   11    11.03         16.11        T(11)
   12    10.92         16.11        T(12)
   13    11.10         16.11        T(13)
   14    10.91         16.11        T(14)
   15    10.53         16.11        T(15)
   16    10.35         16.11        T(16)
   17    9.803         16.12        T(17)
   18    10.50         16.11        T(18)
   19    9.761         16.12        T(19)
   20    10.27         16.12        T(20)
   21    10.52         16.12        T(21)
   22    9.598         16.13        T(22)
   23    3.847         18.27        T(23)
   24    3.456         20.63        T(24)
    0     ZERO       ALIASED        TYPE(1)
   25   -5.551         .7411        TYPE(2)
    0     ZERO       ALIASED        ZR(1)
   26   -.8399E-01    .4871E-01     ZR(2)
   27   -.8734E-01    .8327E-01     ZR(3)
   28   -.3443         .1022        ZR(4)
    0     ZERO       ALIASED        ZE(1)
   29   -.4675        .5381E-01     ZE(2)
   30   -.6169        .6760E-01     ZE(3)
    0     ZERO       ALIASED        ZC(1)
   31   -.3571        .4409E-01     ZC(2)
    0     ZERO       ALIASED        TYPE(1).ZR(1)
    0     ZERO       ALIASED        TYPE(1).ZR(2)
    0     ZERO       ALIASED        TYPE(1).ZR(3)
    0     ZERO       ALIASED        TYPE(1).ZR(4)
    0     ZERO       ALIASED        TYPE(2).ZR(1)
   32    2.995         .7170        TYPE(2).ZR(2)
   33    4.212         .7289        TYPE(2).ZR(3)
   34    4.554         .7302        TYPE(2).ZR(4)
    0     ZERO       ALIASED        TYPE(1).ZE(1)
    0     ZERO       ALIASED        TYPE(1).ZE(2)
    0     ZERO       ALIASED        TYPE(1).ZE(3)
    0     ZERO       ALIASED        TYPE(2).ZE(1)
   35   -.1125         .2346        TYPE(2).ZE(2)
   36    .6805         .2508        TYPE(2).ZE(3)
    0     ZERO       ALIASED        TYPE(1).ZC(1)
    0     ZERO       ALIASED        TYPE(1).ZC(2)
    0     ZERO       ALIASED        TYPE(2).ZC(1)
   37   -.9150         .2018        TYPE(2).ZC(2)
   SCALE PARAMETER TAKEN AS         1.000
```

B5. (Continued).

```
THE CURRENT MODEL HAS
-2*LOG(LIKELIHOOD) =                5445.41574
LIKELIHOOD RATIO CHI SQUARE =       680.471521
PEARSON CHI SQUARE =                1753.73717
WITH DEGREES OF FREEDOM =           767.
```

Figure C1. Stratification according to covariate z_1 = REL



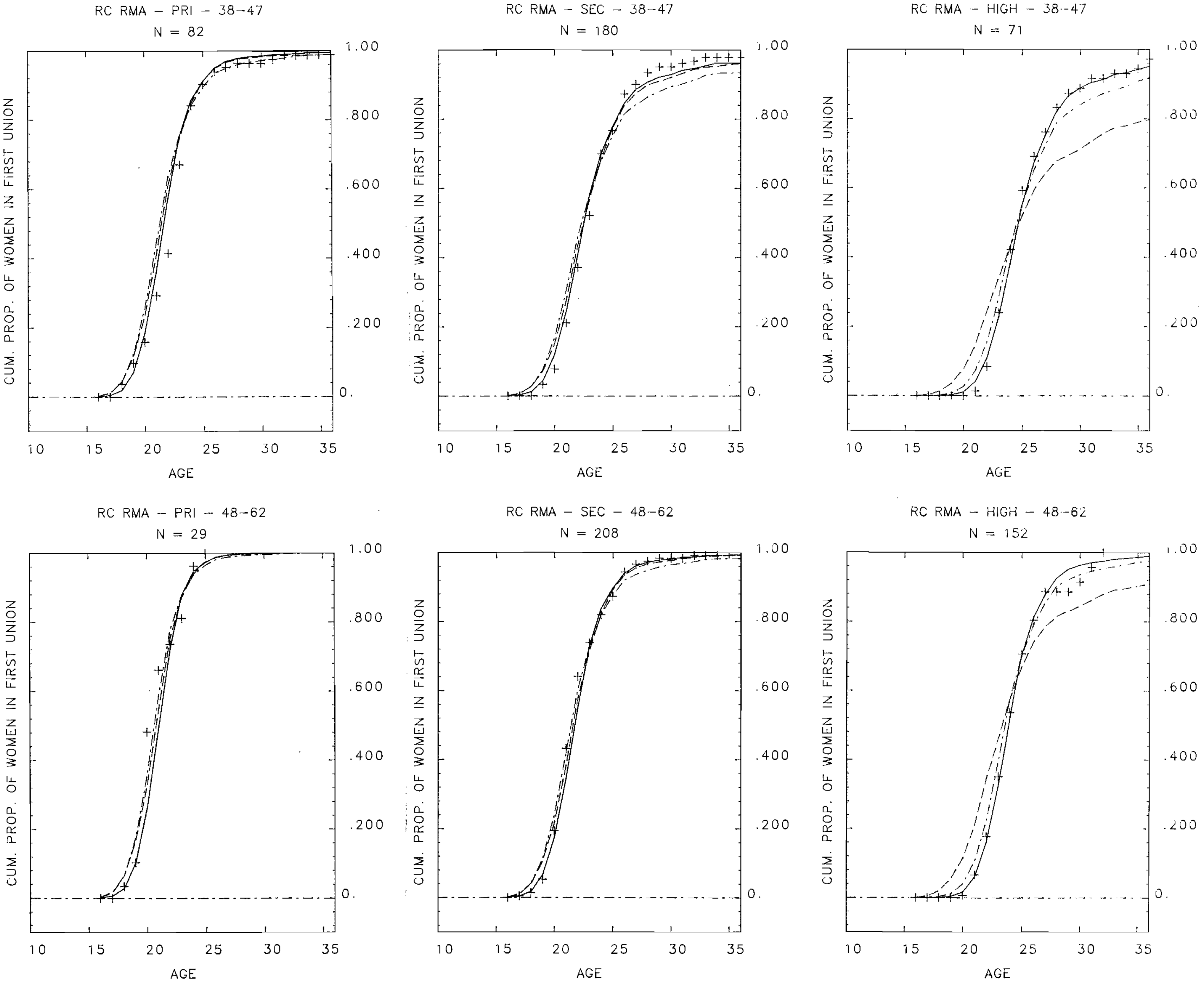


Figure C2. Observed versus estimated cumulative probabilities of entry into first union, per subgroup.

Legend : observed : + + +
estimated : T + REL + EDU + COH : – – – ; T" + REL + EDU + COH : —·—·—
T* + REL + EDU + COH : ———

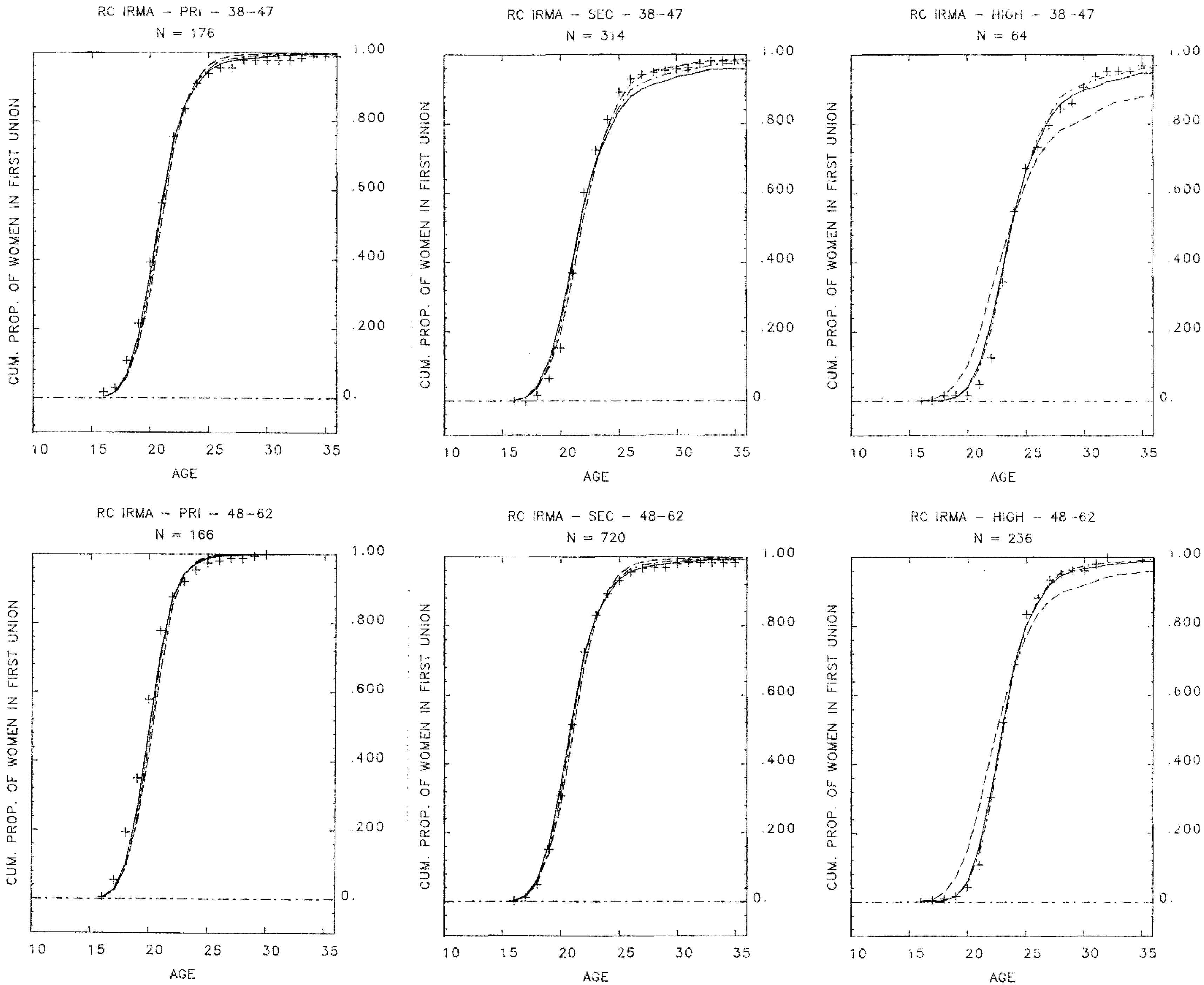


Figure C2. Continued

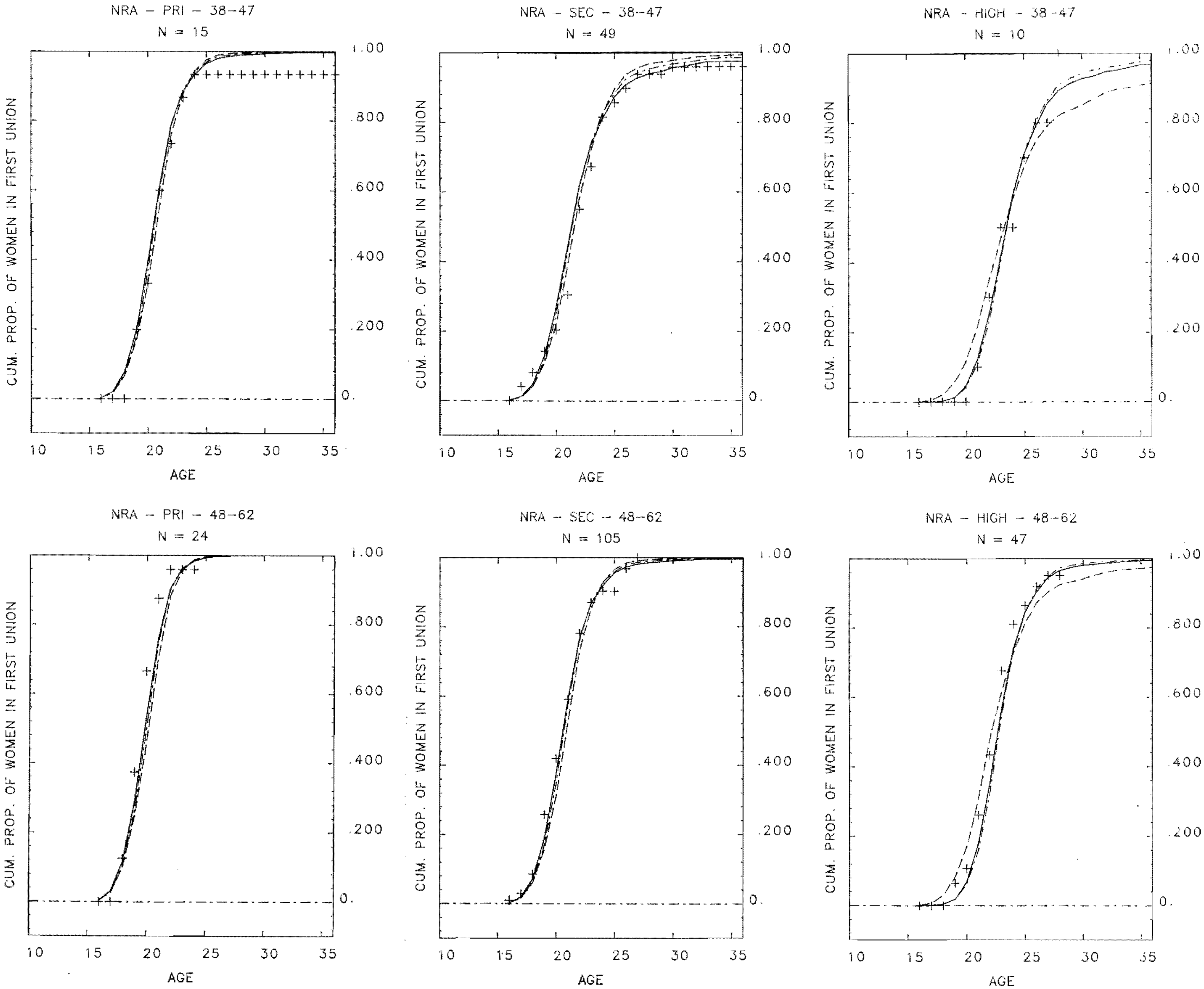


Figure C2. Continued

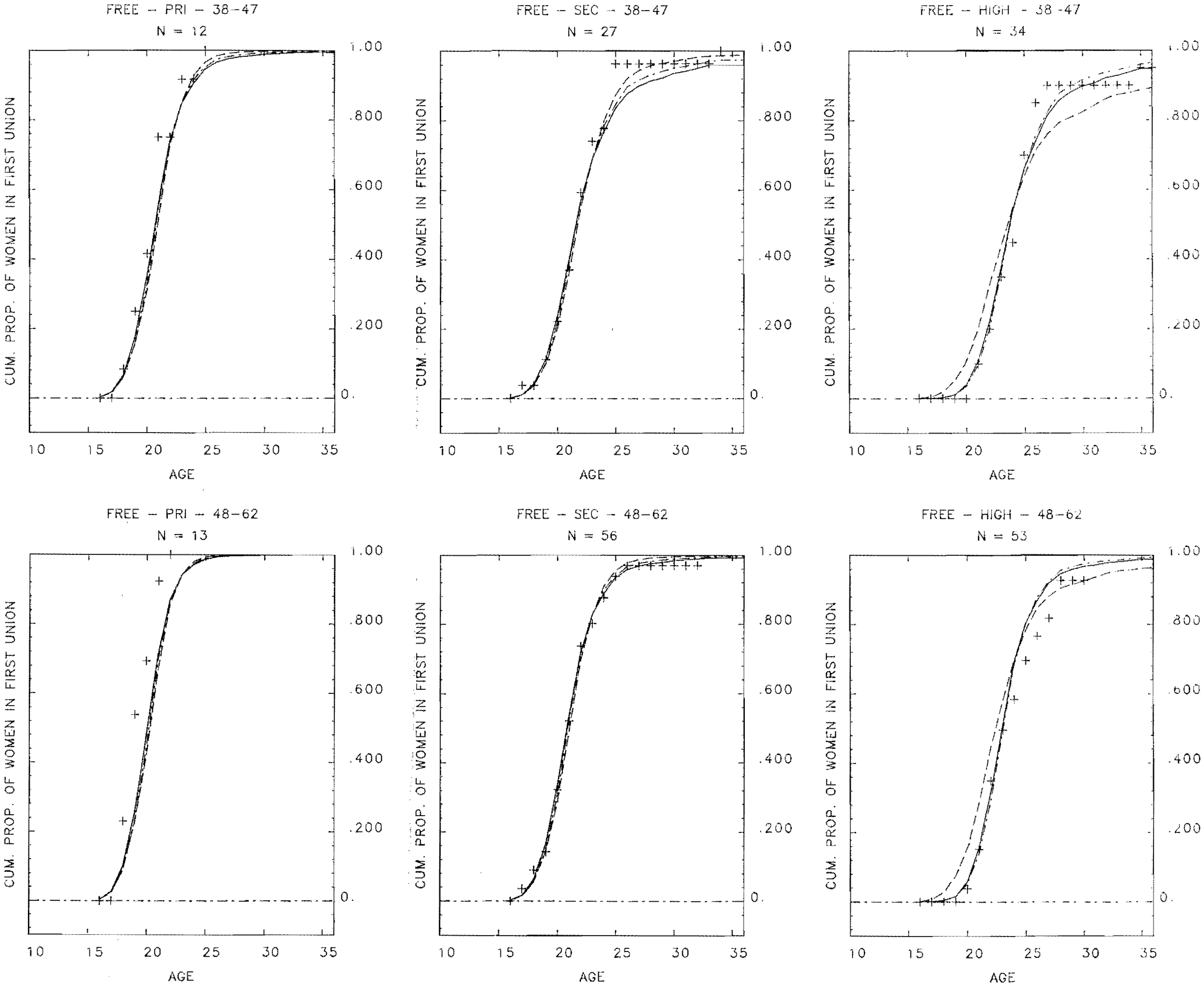


Figure C2. Continued

Figure C3. Schematic presentation of the shifts (or stratifications) denoted by T" and T*

a. T", i.e. 2 strata and a shift over 2 years

PRI; SEC subgroups :

HIGH subgroups :

b. T*, i.e. 4 strata and shifts over 1, 2 or 3 years

PRI - not RC RMA; SEC - not RC RMA subgroups :

PRI - RC RMA; SEC - RC RMA subgroups :

HIGH - not RC RMA subgroups :

HIGH - RC RMA subgroups :

Figure C4. Schematic presentation of various classes of PH models in the presence of competing risks and under the assumption of piecewise constant hazards

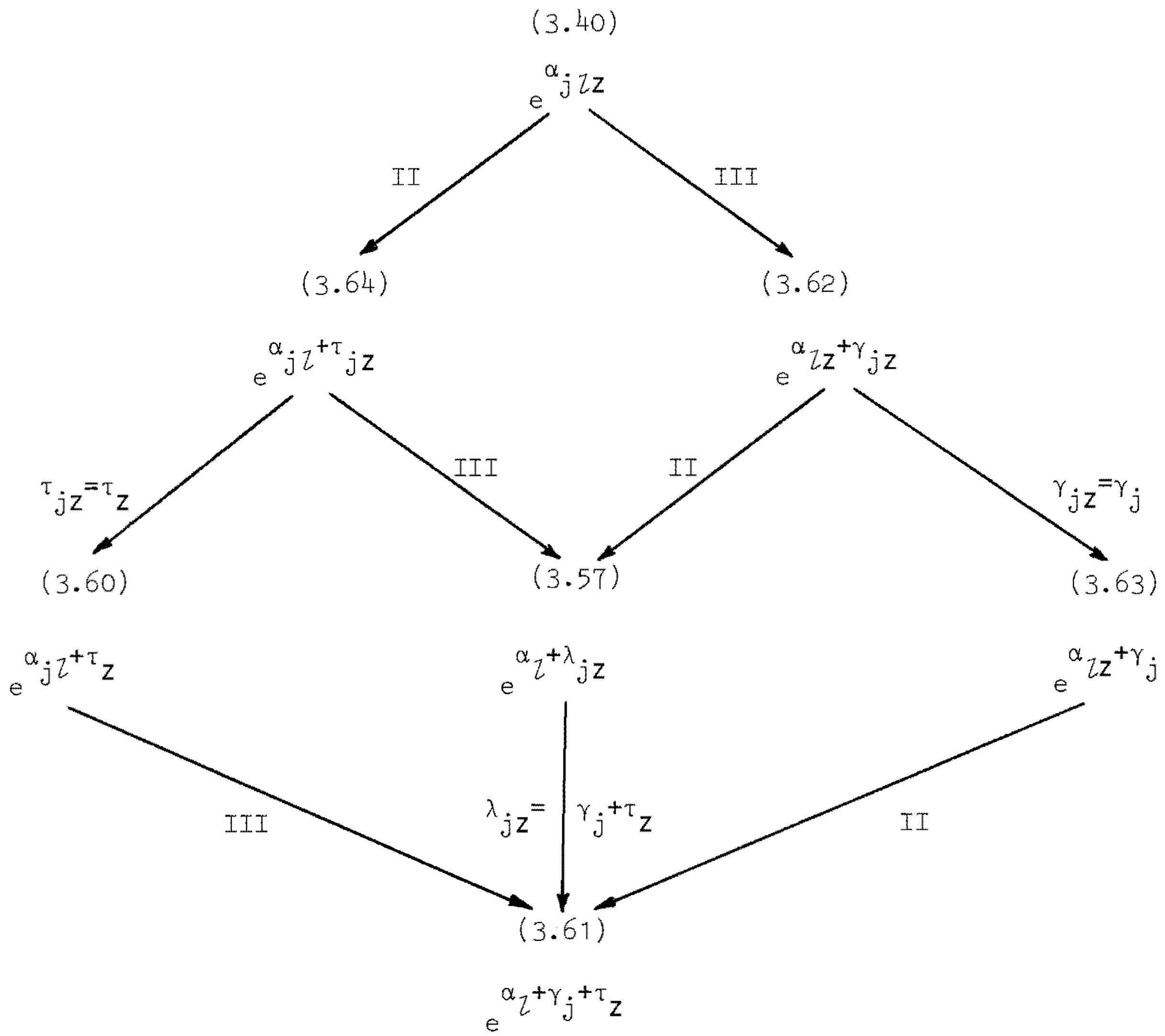
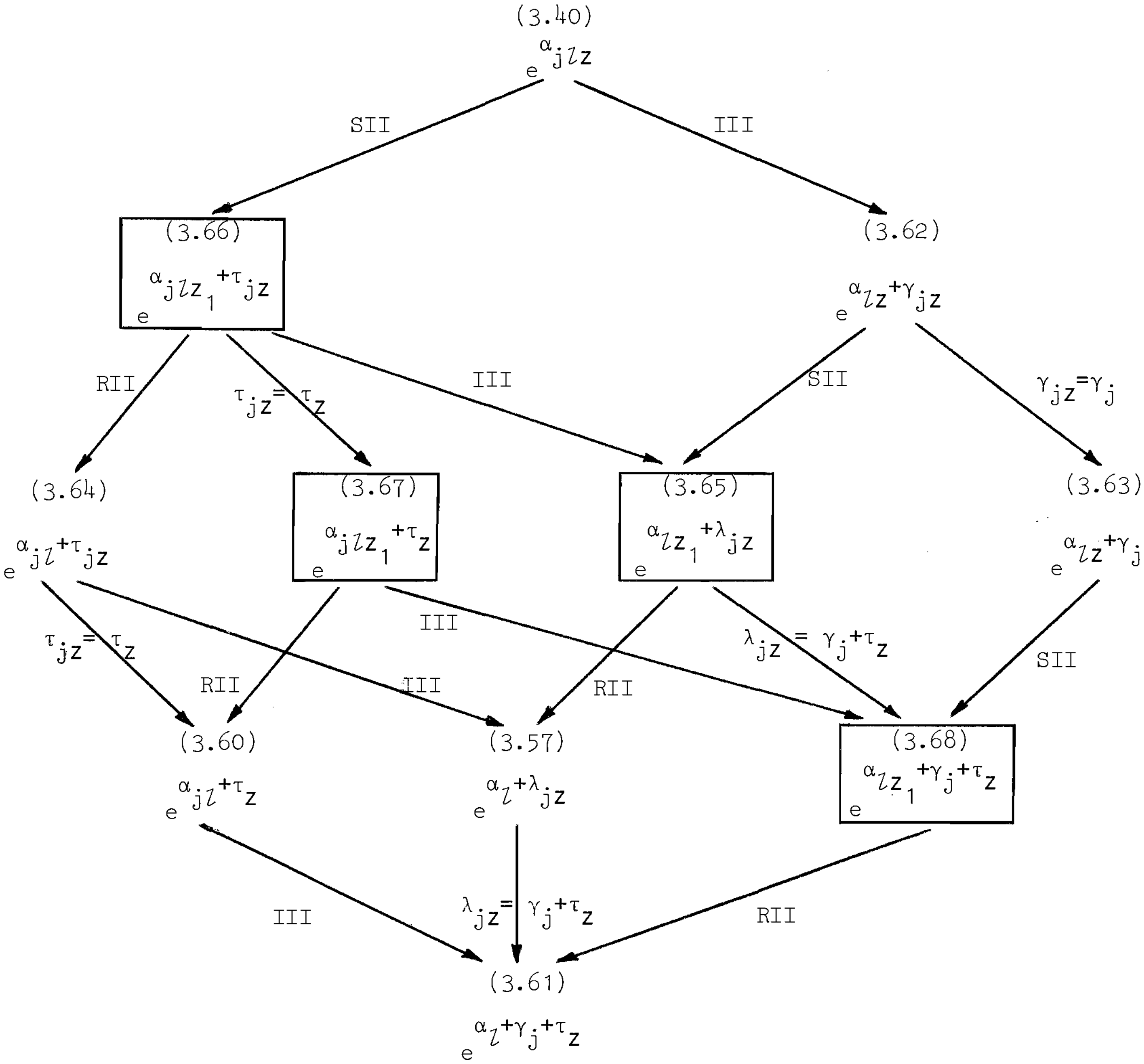


Figure C5. Schematic presentation of various classes of PH and SPH models (in boxes) in the presence of competing risks and under the assumption of piecewise constant hazards[a]



(a) SII : assumption II holds within each stratum

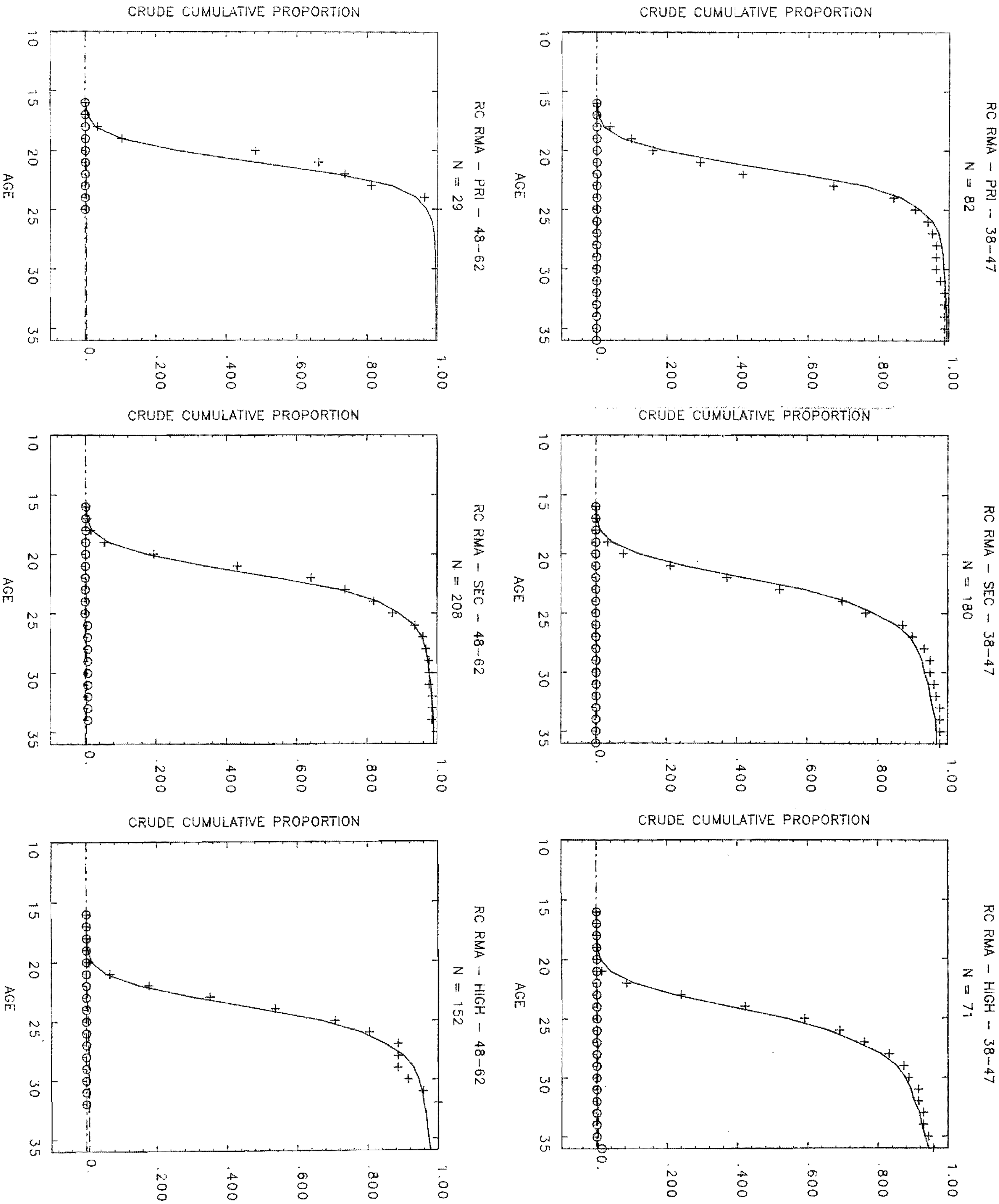
RII : assumption II holds also across strata

(RII can be assumed if SII is already assumed)

Figure C6. Observed versus estimated crude cumulative probabilities of entry into first marriage and first cohabitation, per subgroup. Model T* + TYPE * (ZR + ZE + ZC).

Legend : observed : marriage +++; cohabitation 000

estimated : marriage ——; cohabitation ---



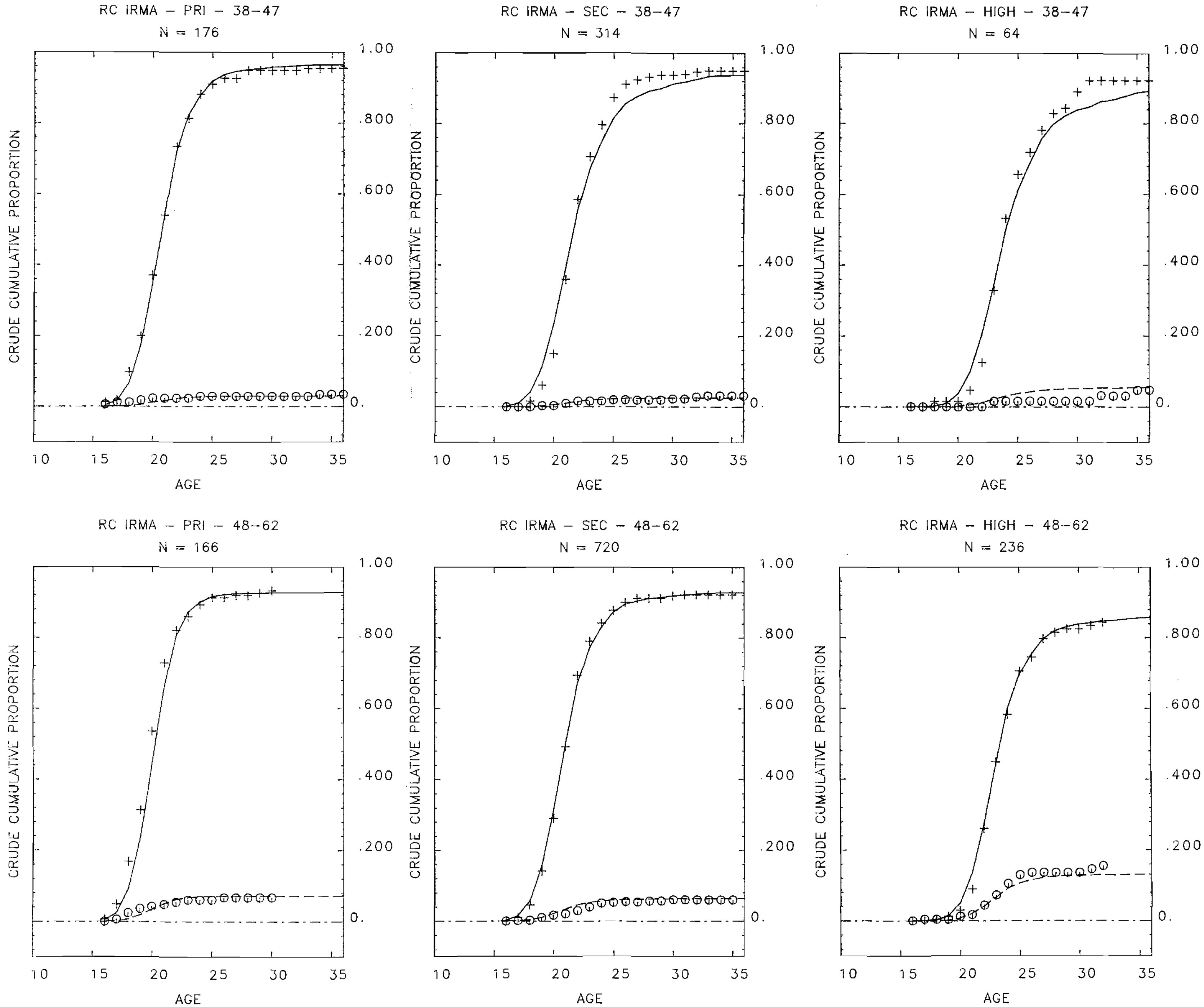


Figure C6. Continued.

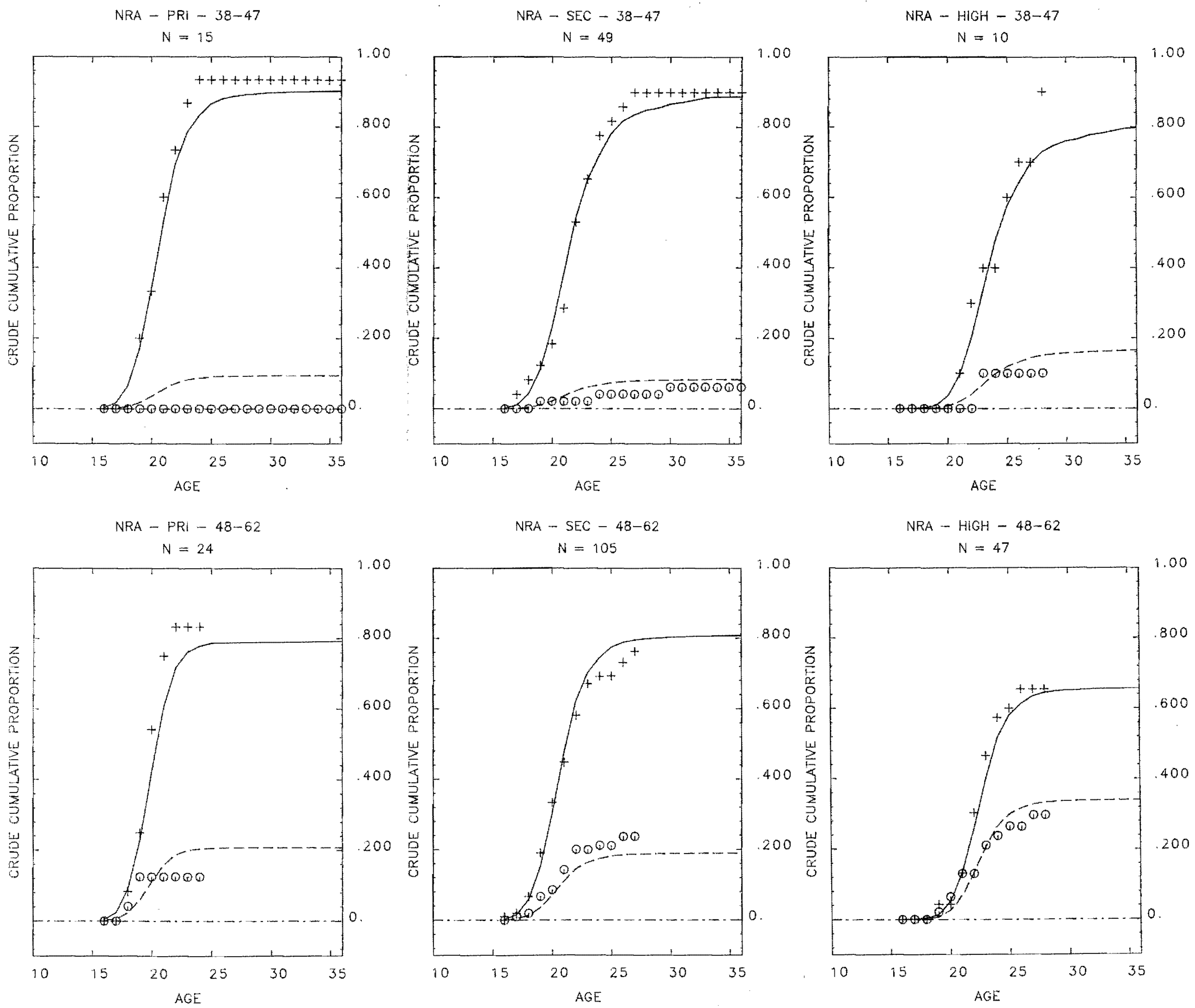


Figure C6. Continued.

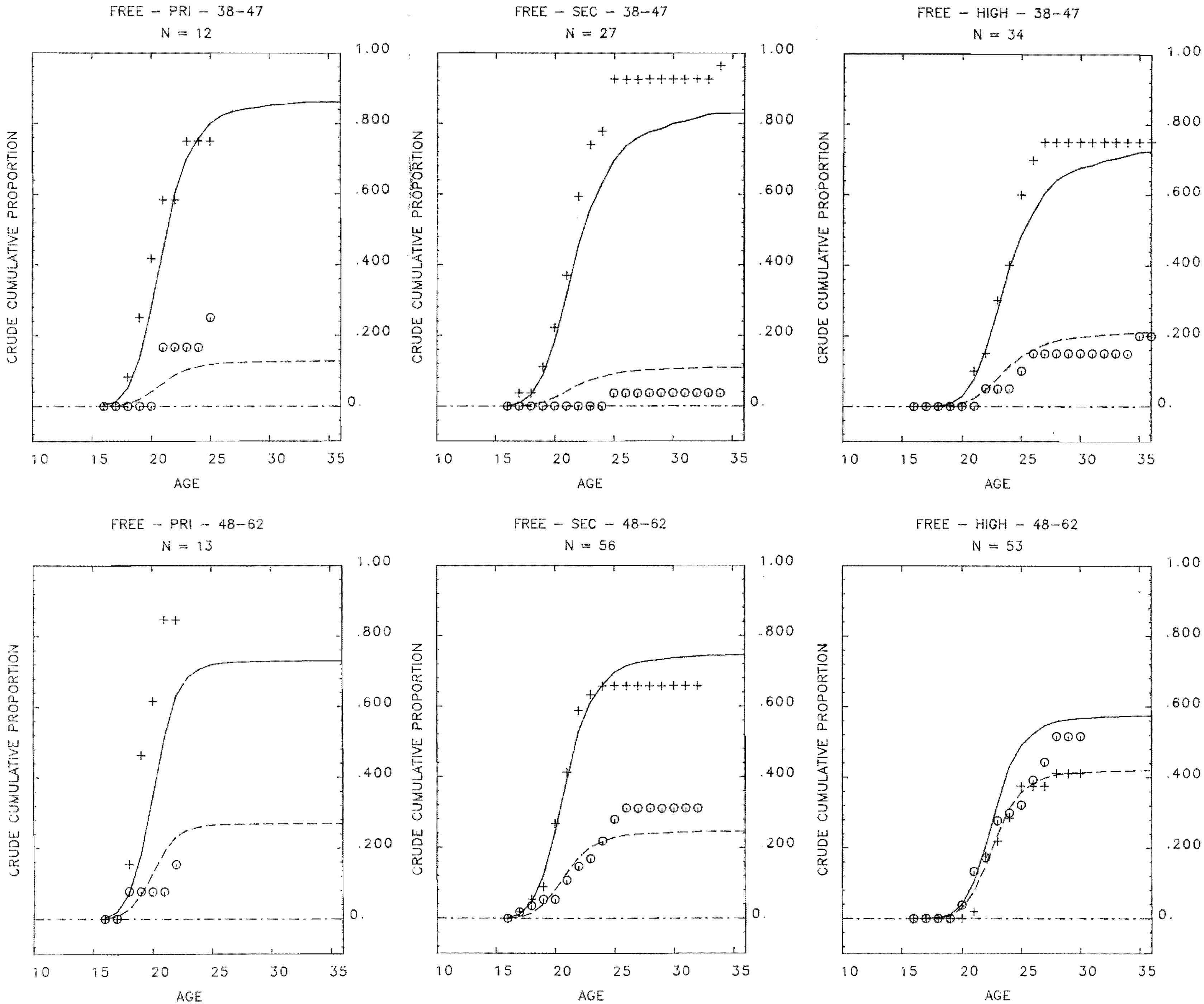


Figure C6. Continued.

Appendix D. GLIM3-macros for calculating the covariance matrix of functions of parameter estimates.

Let $\underline{\theta} = (\theta_1,\ldots\ldots,\theta_k)^T$ be a vector of k parameters, let $\underline{\phi} = (\phi_1,\ldots\ldots,\phi_l)^T$ be a vector of l new parameters defined as functions of the parameters $\theta_1,\ldots\ldots\theta_k$ by

$$\phi_i = f_i(\underline{\theta}) \qquad i = 1,\ldots\ldots l$$

where the f_i are known functions.

If the f_i are sufficiently well-behaved, so that, among other things, the partial derivatives $\frac{\partial f_i}{\partial \theta_j}$ exist, then it can be shown, by considering Taylor expansions of the functions f_i about $\underline{\theta} = \hat{\underline{\theta}}$ (i.e. the maximum likelihood estimates of the parameters $\theta_1,\ldots\ldots\theta_k$), that the following approximate formula holds.

$$C \cong D.V.D^T$$

where V is the kxk covariance matrix of $\underline{\theta}$, C is the lxl covariance matrix of $\underline{\phi}$, and D is the lxk matrix of partial derivatives (evaluated in the m.l.e. $\hat{\underline{\theta}}$).

In GLIM-Newsletter no. 5 R. Burn and R. Thompson presented macros for calculating the matrix C, but they only considered the case $k=l$, i.e. k parameters $\theta_1,\ldots\ldots\theta_k$ are transformed in k new parameters $\phi_1,\ldots\ldots\phi_k$. Consequently, their macros MUL1, MUL2, MUL3, MUL4, as well as the macro COVA can only be used in the special case $k=l$. This note presents macros MUL5, MUL6, MUL7 and MUL8 - for multiplication of D with V and D.V with D^T - which can be used in the general case considered. The main differences between the macros MUL1, MUL2, MUL3, MUL4 and the macros MUL5, MUL6, MUL7, MUL8 are caused by

1° the use of a new scalar %N, which stands for the number l of new parameters;

2° the use of a new vector VCM, which, after the calculations, contains the (co)variances of the new parameters $\phi_1,\ldots\ldots\phi_l$.

Note that VCM needs to be defined since the system vector %VC (used by Burn and Thompson) does not have, in general, the correct length $l(l+1)/2$. Note also that, after the calculations, the vector %VC still contains the (co)variances of the parameters $\theta_1,\ldots\ldots\theta_k$.

The macros MUL5, MUL6, MUL7 and MUL8 are as follows.

```
$M MUL5 $CAL %A=%A-1 : %K=%PL-%A !
: V1=%LE(J,%K)*(%K*(%K-1)/2+J)+%GT(J,%K)*(J*(J-1)/2+%K) !
: V1=%VC(V1) : %B=%N $WHI %B MUL6 $$ENDM
$M MUL6 $CAL %B=%B-1 : %I=%N-%B : J1=%EQ(I1,%I) !
: D1(J1*%CU(J1))=D : DV(%I+%N*(%K-1))=%CU(D1*V1) $$ENDM
$M MUL7 $CAL %A=%A-1 : %I=%N-%A : J1=%EQ(I2,%I) !
: V1(J1*%CU(J1))=DV : %B=%I $WHI %B MUL8 $$ENDM
$M MUL8 $CAL %B=%B-1 : %K=%I-%B : J1=%EQ(I1,%K) !
: D1(J1*%CU(J1))=D : VCM(%K+%I*(%I-1)/2)=%CU(V1*D1) $$ENDM
```

The macro COVA constructed by Burn and Thompson depends heavily on the analysis to be done. However, part of it is quite general - i.e. the definition (through $VAR) of the vectors needed for the calculations, the use of the macros for the matrix multiplications, and the displaying of the results. Therefore, the following general macro COVA is presented.

```
$M COVA !
$CAL %M=%N*%PL : %T=%N*(%N+1)/2 : %S=%N*%N !
$EXT %PE !
$USE PADE !
$VAR %PL J V1 D1 : %M J1 I1 I2 DV : %T VCM !
$CAL I1=%GL(%N,%PL) : J=%GL(%PL,1) : I2=%GL(%N,1) !
$EXT %VC !
$CAL %A=%PL $WHI %A MUL5 $CAL %A=%N $WHI %A MUL7 !
$DEL D DV J V1 D1 I1 I2 J1 !
$PRI : : " PAR.  ESTIMATE       S.E." : !
$VAR %N I $CAL I=%GL(%N,1) !
$VAR %T J $CAL J=0 : J(I*(I+1)/2)=I !
$VAR %N V1 $CAL V1(J)=VCM : V1=%SQRT(V1) !
$LOO FI V1 !
$DEL FI I J V1 !
$VAR %S I1 I2 J1 J !
$CAL I1=%GL(%N,%N) : I2=%GL(%N,1) : J1=%GE(I1,I2) : J=J1*%CU(J1) !
$DEL J1 !
$VAR %T K1 $CAL K1(J)=I1 $DEL I1 !
$VAR %T K2 $CAL K2(J)=I2 $DEL I2 J !
$PRI : : "         PARAMETER PAIR        (CO)VARIANCE" : !
$LOO K1 K2 VCM !
$DEL K1 K2 VCM !
$$ENDM
```

Macro COVA uses a macro PADE. This macro is intended for calculation of the new parameters $\phi_1.....\phi_l$ - which should be stored in a vector FI - and the partial derivatives $\frac{\partial f_i}{\partial \theta_j}(\hat{\underline{\theta}})$ $(i = 1,.....l;$ $j = 1,.....k)$ - which should be stored in a vector D. It must be noted that all vectors used in the macro COVA may also be used in the macro PADE, but they should be deleted as soon as FI and D have got their values. Note also that the scalars %M, %T and %S may not be redefined in PADE. The macro COVA can be used after the scalar %N has been assigned the value l, i.e. the number of new parameters.

EXAMPLE : ABO problem.

Thompson and Baker (1981) presented a loglinear model with composite link function. The GLIM3-programme for the ABO problem was presented in their article. Burn and Thompson (1982) calculated the (co)variances of new parameters p, r and q, i.e. the gene frequencies. The calculations were repeated using the macros MUL5, MUL6, MUL7, MUL8 and COVA shown above. The macro PADE for this problem is as follows.

```
$M PADE ! NEW PARAMETERS (FI) & PARTIAL DERIVATIVES (D)
$VAR %N FI !
$CAL FI=%EXP(%PE) : %Q=%CU(FI) : FI=FI/%Q !
$VAR %M I1 I2 J1 D !
$CAL I1=%GL(%N,%PL) : I2=%GL(%PL,1) : J1=%EQ(I1,I2) !
: D=FI(J1*%CU(J1))-FI(I1)*FI(I2) !
$DEL I1 I2 J1 !
$$ENDM
```

Since the number of new parameters equals the number of original parameters in this problem, the macros are used through

```
$CALC %N=%PL $USE COVA
```

The results correspond exactly to those in Burn and Thompson (1982).

APPENDIX E. MATHEMATICAL ADDENDUM

E1. From (2.1) we have - for small Δt -

$$\mu(t;z).\Delta t \cong P(t \leqslant T < t + \Delta t | T \geqslant t, z).$$

Hence

$$\mu(t;z).\Delta t \cong \frac{P(T \geqslant t|z) - P(T \geqslant t + \Delta t|z)}{P(T \geqslant t|z)}$$

$$\cong \frac{S(t;z) - S(t + \Delta t;z)}{S(t;z)}$$

But $\lim\limits_{\Delta t \downarrow 0} \dfrac{S(t;z) - S(t + \Delta t;z)}{\Delta t} = - \dfrac{dS(t;z)}{dt}$

from the definition of the derivative of a function.

Hence

$$\mu(t;z) = - \frac{1}{S(t;z)} \frac{dS(t;z)}{dt}$$

$$= - \frac{d}{dt} \log S\,(t;z)$$

Integration of the latter formula gives

$$\int_0^t \mu(s;z)\,ds = - [\log S(t;z) - \log S(0;z)]$$

$$= - \log S(t;z)$$

since $S(0;z) = 1$ and $\log 1 = 0$.
Formula (2.2a) then follows immediately.

E2. By definition, a p.d.f. $f(t;z)$ is related to its c.d.f. $F(t;z)$ as follows

$$f(t;z) = \frac{dF(t;z)}{dt}.$$

The c.d.f. $F(t;z)$ is defined by

$$F(t;z) = P(T < t|z).$$

Since

$$S(t;z) = P(T \geqslant t|z) = 1 - P(T < t|z)$$

$$= 1 - F(t;z),$$

we have

$$\frac{dS(t;z)}{dt} = -\frac{dF(t;z)}{dt}.$$

Hence

$$f(t;z) = -\frac{dS(t;z)}{dt}.$$

From the results in E1 we get then :

$$\mu(t;z) = f(t;z)/S(t;z),$$

and (2.4) follows then immediately.

E3. From (2.23) and the definition of the reference subgroup (i.e. $\beta_{z_o}=0$), we have

$$\mu(t;z) = \mu(t;z_o).e^{\beta_z}$$

Integration gives (2.26), and substitution of (2.26) in

$$S(t;z) = \exp(-\Lambda(t;z))$$

gives

$$S(t;z) = \exp(-\Lambda(t;z_o).e^{\beta_z})$$

$$= (e^{-\Lambda(t;z_o)})^{e^{\beta_z}}$$

$$= (S(t;z_o))^{e^{\beta_z}}$$

which is (2.25).

Further

$$1 - q_l(z) = e^{-e^{\alpha_l}.e^{\beta z}}$$

by a generalization of (2.15).

Hence

$$1 - q_l(z) = e^{-\mu(t;z_o).e^{\beta z}} \qquad \text{for } a_{l-1} \leqslant t < a_l$$

$$= (e^{-\mu(t;z_o)})^{e^{\beta z}}$$

$$= (1 - q_l(z_o))^{e^{\beta z}},$$

which gives (2.27).

E4. Since $\mu(t;z_o') = 0$ if $t < 0$, it follows from (2.28a) that $\mu(t+b;z_o'') = 0$ if $t < 0$, or $\mu(s;z_o'') = 0$ if $s < b$.

Hence $\Lambda(s;z_o'') = 0$ if $s \leqslant b$ and $S(s;z_o'') = 1$ if $s \leqslant b$.

We get then from (2.28a) :

$$\Lambda(t;z_o') = \int_o^t \mu(u;z_o')\, du$$

$$= (\int_o^t \mu(u+b;z_o'')\, du).e^{\omega}$$

$$= (\int_b^{t+b} \mu(s;z_o'')\, ds).e^{\omega}$$

$$= (\Lambda(t+b;z_o'') - \Lambda(b;z_o'')).e^{\omega}$$

$$= \Lambda(t+b;z_o'').\, e^{\omega}$$

Further

$$S(t;z_o') = \exp(-\Lambda(t;z_o'))$$

$$= \exp(-\Lambda(t+b;z_o'').e^{\omega})$$

$$= (S(t+b;z_o''))^{e^{\omega}}.$$

Generalization of (2.15) gives

$$1 - q(a_{l-1},1;z_o') = e^{-\mu(a_{l-1};z_o')}$$

$$= e^{-\mu(a_{l-1}+b;z_o'').e^{\omega}}$$

$$= (1 - q(a_{l-1}+b,1;z_o''))^{e^{\omega}}.$$

This proves formulae (2.29) to (2.31a).

Note also the following simple relations if $\omega = 0$:

$$\mu(t;z_o') = \mu(t+b;z_o''),$$

$$\Lambda(t;z_o') = \Lambda(t+b;z_o''),$$

$$S(t;z_o') = S(t+b;z_o''),$$

$$q(a_{l-1},1;z_o') = q(a_{l-1}+b,1;z_o'').$$

E5. The system of maximum likelihood equations obtained from (2.35) is :

$$\frac{\partial \log \mathcal{L}}{\partial \alpha_l} = \sum_z d_{lz} - \sum_z \tilde{E}_{lz}.e^{\alpha_l+\beta_z} = 0 \qquad (l=1,\ldots L)$$

$$\frac{\partial \log \mathcal{L}}{\partial \beta_z} = \sum_l d_{lz} - \sum_l \tilde{E}_{lz}.e^{\alpha_l+\beta_z} = 0 \qquad (z\varepsilon\ )$$

Equivalently :

$$d_{l.} - (\sum_z \tilde{E}_{lz}.e^{\beta_z}).e^{\alpha_l} = 0 \qquad (l=1,\ldots L)$$

$$d_{.z} - (\sum_l \tilde{E}_{lz}.e^{\alpha_l}).e^{\beta_z} = 0 \qquad (z\varepsilon\ )$$

where $d_{l.} = \sum_z d_{lz}$ and $d_{.z} = \sum_l d_{lz}$.

This system of equations should, in general, be solved iteratively, and computer packages should therefore be used. Note also that if e^{α_l} is solved from the first set of equations, i.e.

$$e^{\alpha_l} = d_{l.}/(\sum_z \tilde{E}_{lz}.e^{\beta_z}), \qquad (l=1,\ldots L)$$

and if these expressions are substituted in the second set of equations, then we obtain a set of equations in the parameters β_z alone, i.e.

$$d_{.z} - e^{\beta_z}.(\sum_l \frac{\tilde{E}_{lz}.d_{l.}}{\sum_z \tilde{E}_{lz}.e^{\beta_z}}) = 0. \qquad (z\varepsilon\ )$$

This system is then solved iteratively, yielding estimates $\hat{\beta}_z$, and these estimates are substituted in the above expressions for the e^{α_l} $(l=1,\ldots L)$, yielding estimates $e^{\hat{\alpha}_l}$ $(l=1,\ldots L)$.

E6. By definition, Me(z) is the age at which 50% of women in subgroup z ever experiencing first union have already entered into first union. Similarly, P10(z) is the age at which 10% of women in subgroup z ever experiencing first union have already entered into first union. In general, let t_z be the age at which 100p% of women in subgroup z ever experiencing first union have already entered into first union. Then, the following relation is satisfied :

$$\frac{\hat{F}(t_z;z)}{\hat{c}(z)} = p \tag{E6.1}$$

An estimate of t_z can now be obtained by interpolation (of some function) between the endpoints a_{l-1} and a_l which should satisfy the inequalities :

$$\frac{\hat{F}(a_{l-1};z)}{\hat{c}(z)} \leqslant p < \frac{\hat{F}(a_l;z)}{\hat{c}(z)} \tag{E6.2}$$

or the inequalities :

$$\hat{F}(a_{l-1};z) \leqslant p.\hat{c}(z) < \hat{F}(a_l;z). \tag{E6.3}$$

Several interpolation methods can be used : e.g. linear or quadratic interpolation of the cumulative distribution function $\hat{F}(t;z)$,....
A linear interpolation of the cumulative hazard function $\hat{\Lambda}(t;z)$ is preferred since this function is in each interval $[a_{l-1},a_l)$ linear in t. The linear interpolation formula is obtained as follows.
From (E6.3) we get subsequently :

$$\hat{S}(a_{l-1};z) \geqslant 1-p.\hat{c}(z) > \hat{S}(a_l;z), \tag{E6.4}$$

or

$$\hat{\Lambda}(a_{l-1};z) \leqslant -\log(1-p.\hat{c}(z)) < \hat{\Lambda}(a_l;z). \tag{E6.5}$$

Note that $-\log(1-p.\hat{c}(z)) = \hat{\Lambda}(t_z;z)$. Hence, the linear interpolation formula is :

$$t_z = a_{l-1} + \frac{a_l - a_{l-1}}{\hat{\Lambda}(a_l;z) - \hat{\Lambda}(a_{l-1};z)}.(-\log(1-p.\hat{c}(z))-\hat{\Lambda}(a_{l-1};z)) \tag{E6.6}$$

This formula may be simplified to :

$$t_z = a_{l-1} + \frac{-\log(1-p.\hat{\partial}(z))-\hat{\Lambda}(a_{l-1};z)}{\hat{\mu}_{lz}} \qquad \text{(E6.7)}$$

where $\hat{\mu}_{lz}$ is the estimated (constant) hazard in the l-th interval. Further, we used $\hat{F}(a_L;z)$ as an estimate for $\hat{\partial}(z)$. The interpolation formula used is thus finally :

$$t_z = a_{l-1} + \frac{-\log(1-p.\hat{F}(a_L;z))-\hat{\Lambda}(a_{l-1};z)}{\hat{\mu}_{lz}} \qquad \text{(E6.8)}$$

E7. Use is made of the generalized Pearson chi-square statistic ($\hat{\chi}_P^2$) and the scaled deviance or log-likelihood ratio chi-square statistic ($\hat{\chi}_L^2$). Those statistics are defined as follows. Let $\hat{\mathcal{L}}_C$ be the estimated likelihood under some model which is called the 'current model', and let $\hat{\mathcal{L}}_S$ be the estimated likelihood under the 'saturated model' (i.e. the model with estimates as in (2.22)). Let $\hat{M}_{lz}$ be the estimated mean of the dependent variable d_{lz} under the current model, and note that the estimate of M_{lz} under the saturated model is d_{lz}. Then, the (estimated) generalized Pearson chi-square statistic is defined as

$$\hat{\chi}_P^2 = \sum_z \sum_l \frac{(d_{lz}-\hat{M}_{lz})^2}{\hat{M}_{lz}}$$

and the scaled deviance is defined as

$$\hat{\chi}_L^2 = -2 \log (\hat{\mathcal{L}}_C/\hat{\mathcal{L}}_S).$$

The latter statistic can also be written as

$$\hat{\chi}_L^2 = 2 \sum_z \sum_l d_{lz} \log (d_{lz}/\hat{M}_{lz}).$$

It is known that the statistics $\hat{\chi}_P^2$ and $\hat{\chi}_L^2$ are asymptotically chi-square distributed with, for example, ν degrees of freedom. The number of degrees of freedom ν is, in GLIM-terminology, the number of 'units' minus the number of independent parameters in the current model. Note that the number of units is in fact the number of combinations of intervals (l) and covariates (z) with non-zero number of observations n_{lz}. The statistics $\hat{\chi}_P^2$ and $\hat{\chi}_L^2$ can be used to test the goodness-of-fit of the current model. The scaled deviance is useful for comparison of two nested models, say model 1 and model 2. Let the estimated scaled deviance, its degrees of freedom, the estimated likelihood and means under model i be $\hat{\chi}_{Li}^2$, ν_i, $\hat{\mathcal{L}}_i$ and $\hat{M}_{lz}^{(i)}$ respectively. If model 1 is nested in model 2, then :

$$\hat{\chi}_{L1}^2 - \hat{\chi}_{L2}^2 = -2 \log (\hat{\mathcal{L}}_1/\hat{\mathcal{L}}_2).$$

Thus, the difference between the scaled deviances of model 1 and model 2 is again a log-likelihood ratio statistic and, as such, it is asymptotically chi-square distributed with $\nu_1 - \nu_2$ degrees of freedom. Note that $\nu_1 - \nu_2$ is also equal to the number of independent parameters added to model 1 to get model 2. 'Analysis of deviance' tables (Baker and Nelder, 1978) are constructed with scaled deviances and differences between scaled deviances for nested models. It is not a problem to extend the above ideas to the competing risks problem. In the notation used to derive formula (3.41) from formula (3.39) - i.e. $z^* = (z,j)$ - the scaled deviance can be written as

$$\hat{\chi}_L^2 = 2 \sum_{z^*} \sum_{l} d_{lz^*} \log (d_{lz^*}/\hat{M}_{lz^*}),$$

and the generalized Pearson chi-squared statistic is

$$\hat{\chi}_P^2 = \sum_{z^*} \sum_{l} \frac{(d_{lz^*} - \hat{M}_{lz^*})^2}{\hat{M}_{lz^*}} .$$

Note that the formula $\hat{\chi}_L^2 = -2 \log (\hat{\mathcal{L}}_C/\hat{\mathcal{L}}_S)$ is still useful. From those formulae, it is clear that the cause covariate can be treated in the same way as the other covariates, as has been argued above in the text.

E8. $c_j(z)$ is the probability that a woman with covariates z ever enters into first union due to cause j, and c(z) is the probability that a woman with covariates z ever enters into first union, irrespective of cause. We assumed that entry into first union is never due to more than one cause simultaneously (i.e. the assumption leading to (3.2)). Equation (3.10) follows immediately from this assumption and from the fundamental probability theory. Similarly, a woman with covariates z who ever experiences first union due to cause j (with probability $c_j(z)$), experiences it either before time t (with probability $Q_j(t;z)$) or after time t (with probability $S_j(t;z)$) Hence, the fundamental probability theory gives

$$c_j(z) = Q_j(t;z) + S_j(t;z),$$

which is equivalent to (3.11).

E9. By ignoring covariates z and putting t equal to a_{l-1} and h equal to 1 in (3.13a), the following can be derived :

$$q_{jl} = q_j(a_{l-1},1)$$

$$= \int_{a_{l-1}}^{a_l} \frac{S(s).\mu_j(s)}{S(a_{l-1})} \, ds$$

$$= \int_{a_{l-1}}^{a_l} e^{-(\Lambda(s)-\Lambda(a_{l-1}))} .\mu_j(s).ds$$

$$= e^{\alpha_{jl}} . \int_{a_{l-1}}^{a_l} e^{-(\sum_j e^{\alpha_{jl}}).(s-a_{l-1})} ds$$

(since $\mu_j(s) = e^{\alpha_{jl}}$ for all s in the l-th interval, and since $\sum_j e^{\alpha_{jl}}$ is the constant total hazard in the l-th interval)

$$= e^{\alpha_{jl}} . \int_0^1 e^{-(\sum_j e^{\alpha_{jl}}).u} du$$

$$= e^{\alpha_{jl}} . \left[\frac{e^{-\sum_j e^{\alpha_{jl}}.u}}{-\sum_j e^{\alpha_{jl}}} \right]_{u=o}^{u=1}$$

$$= e^{\alpha_{jl}} . \frac{e^{-\sum_j e^{\alpha_{jl}}} - 1}{-\sum_j e^{\alpha_{jl}}}$$

which gives (3.42).

Formula (3.43) and (3.44) are easily obtained from (3.22e) and (2.7e) respectively : again t is replaced by a_{l-1}, h by 1 and covariates Z are ignored.

E10. The following approximations are based on the Taylor series expansion for the exponential function, i.e.

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + ...$$

Hence, if $|x|$ is sufficiently small, we get the first order approximation :

$$e^x \cong 1 + x$$

Thus, if the total hazard μ_l for the l-th interval is sufficiently small, we have from (3.42a) :

$$q_{jl} \cong \mu_{jl}$$

and from (3.44a) :

$$q_l \cong \mu_l .$$

Similarly, if μ_{jl} is sufficiently small (note that this is so if μ_l is sufficiently small), we get from (3.43a) :

$$q_{(j)l} \cong \mu_{jl}.$$

Instead of using a first order approximation we can use the second order approximation

$$e^x \cong 1 + x + \frac{x^2}{2}.$$

We get then the following approximate formulae :

$$q_{jl} \cong \mu_{jl}.(1 - \frac{\mu_l}{2})$$

and $$q_l \cong \mu_l.(1 - \frac{\mu_l}{2})$$

if μ_l is sufficiently small,

and
$$q_{(j)l} \cong \mu_{jl}.(1 - \frac{\mu_{jl}}{2})$$

if μ_{jl} is sufficiently small.

Possibly we can use a first order approximation for $e^{-\mu_{jl}}$ and a second order approximation for $e^{-\mu_l}$. Then we can derive the approximate formula :

$$q_{jl} \cong q_{(j)l}.(1 - \frac{\mu_l}{2}).$$

This gives an approximate relationship between the net probability $q_{(j)l}$ and the crude probability q_{jl}; the correction factor uses the total hazard μ_l. It is interesting to note here that Pollard (1973) presents a similar approximate relationship between q_{jl} and $q_{(j)l}$ (which he calls respectively dependent and independent probabilities). The correction factor in his formula, however, does not depend on the total hazard but on the net probability corresponding to the alternative cause(s).

E11. Property A can be demonstrated as follows :

$$\mu(t;z) = \sum_j \mu_j(t;z) \qquad \text{(by (3.2))}$$

$$= \sum_j \mu_j(t;z_o)\ e^{\tau_j z} \qquad \text{(by II)}$$

$$= \sum_j \mu(t;z_o).\theta_{jz_o}.e^{\tau_j z} \qquad \text{(by III')}$$

$$= \mu(t;z_o).(\sum_j \theta_{jz_o}.e^{\tau_j z})$$

$$= \mu(t;z_o).e^{\beta_z}$$

if e^{β_z} is defined as the weighted average $\sum_j \theta_{jz_o}.e^{\tau_j z}$.

Property B can be demonstrated as follows.

$$\mu(t;z) = \sum_j \mu_j(t;z) \qquad \text{(by (3.2))}$$

$$= \sum_j \mu_j(t;z_o)\ e^{\tau_j z} \qquad \text{(by II)}$$

$$= (\sum_j \mu_j(t;z_o)).e^{\tau_z} \qquad \text{(since } \tau_{jz}=\tau_z\text{)}$$

$$= \mu(t;z_o).e^{\tau_z} \qquad \text{(by (3.2))}$$

$$= \mu(t;z_o).e^{\beta_z}$$

if e^{β_z} is defined as being equal to e^{τ_z}.

IPD-WORKING PAPERS 1983

1983-1. R. Lesthaeghe : "A Century of Demographic and Cultural Change in Western Europe : An Exploration of Underlying Dimensions".

X 1983-2. R. Lesthaeghe, C. Vanderhoeft, S. Becker, M. Kibet : "Individual and Contextual Effects of Education on Proximate Determinants and on Life Time Fertility in Kenya".

1983-3. F. Eelens : "Impact of Breast-feeding on Infant and Child Mortality with varying Incidence of Malaria - Evidence from the Kenya Fertility Survey 1977-78".

1983-4. S. Becker, A. Chowdhury, S. Huffman : "Determinants of Natural Fertility in Matlab, Bangladesh".

1983-5. C. Vanderhoeft : "A Unified Approach to Models for Analysis of Zero-One Data with Applications to Intermediate Fertility Variables".

1983-6. S. Wijewickrema, R. Bulté : "Migration Impact on Population Growth in Belgium : A Multiregional Investigation with Detailed Projection Results for the Period 1976-2001".

1983-7. S. Wijewickrema et Alii : "Marital Status Trends in Belgium, 1961-77 : Application of Multi-State Analysis".

1983-8. H.J. Page : "Fertility and Family : New Currents and emerging Emphases in Research and Policy" (Overview of the UN Expert Group Meeting on Fertility and Family (New Delhi, 1983) preparatory to the International Conference on Population).

1983-9. R. Lesthaeghe, C. Vanderhoeft, S. Becker, M. Kibet : "Individual and Contextual Effects of Female Education on the Kenyan Marital Fertility Transition" (abbreviated non-technical version of Working Paper 1983-2).

X = OUT OF PRINT

IPD-WORKING PAPERS 1984

1984-1. F. Rajulton : "An Age Dependent Semi-Markov Model of Marital Status in Belgium : An Application of Littman's Algorithm to Period Data, 1970".

1984-2. R. Lesthaeghe : "Fertility and its Proximate Determinants in sub-Saharan Africa : The Record of the 1960's & 70's".

1984-3. C. Vanderhoeft : "Life Tables and Relational Models : A Unified Approach, and Applications in Demography".

1984-4. R. Lesthaeghe : "Demographic Renewal and Problems of Social Organization".

1984-5. S. Becker : "A Response Bias in the reporting of Month of Birth in Pregnancy History Surveys".

1984-6. R. Lesthaeghe : "On the Adaptation of sub-Saharan Systems of Reproduction".

IPD-WORKING PAPERS 1985

1985-1. R. Lesthaeghe and F. Eelens : "Social Organization and Reproductive Regimes : Lessons from sub-Saharan Africa and historical Western Europe".

1985-2. P. Willems and S. Wijewickrema : "The Evolution of Nuptiality in Belgium from 1954 to 1981".

1985-3. F. Eelens and L. Donné : "The Proximate Determinants of Fertility in Sub-Saharan Africa: A Factbook based on the Results of the World Fertility Survey".

1985-4. Ronald C. Schoenmaeckers : "Current Fertility Behaviour in Africa. Results from a Birth Interval Analysis of WFS Data".

1985-5. Camille Vanderhoeft : Stratified Proportional Hazards Models. A GLIM oriented approach, with special reference to the problem of competing risks.

1985-6. Stanley Wijewickrema : "Childlessness in Belgium and Flanders".

1985-7. R. Lesthaeghe : "Value Orientations, Economic Growth and Demographic Trends - Toward a Confrontation?

1985-8. Fernando Rajulton and Stan Wijewickrema : "An Enquiry into the Two Basic Assumptions of Multistate Increment-Decrement Life Tables".

IPD-WORKING PAPERS 1986

1986-1. F. Rajulton : "Marriageability and divorceability : a simulation of the unobservables through the conditional gaussian diffusion process.