HIDDEN HETEROGENEITY IN COMPOSITE LINK MODELS : FURTHER DEVELOPMENTS

C. Vanderhoeft

Vrije Universiteit, Brussels

IPD-Working Paper 1987-1

Acknowledgements :
The author is grateful to the National Bank of Belgium for financial support.

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 $ 6.50 or BF 250. 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........

HIDDEN HETEROGENEITY IN COMPOSITE LINK MODELS : FURTHER DEVELOPMENTS

Camille Vanderhoeft *

1. INTRODUCTION

Recently, Vanderhoeft (1986) considered the problem of taking unobserved heterogeneity into account when the basic tool for analysis is a composite link model. In the present notes we try to give a more complete overview of such models and methods. Attention is focused mainly on models for binomial and Poisson data.

In Section 2 we discuss an extension of Thompson and Baker's (1981) composite link model. In Section 3 we introduce in this model multidimensional hidden heterogeneity (for binomial and Poisson data). In Section 4 and Section 5 we discuss approximate models which allow for estimation of the parameters of the composite link model with hidden heterogeneity. Section 6 applies the methods in a reanalysis of binomial data used by Williams (1982) and Poisson data used by Breslow (1984). Finally, Section 7 introduces a simulation study.

The discussion is far from complete. Mathematical details are not included. The reader is strongly advised to have a look at the preceding paper by Vanderhoeft (1986).

* Camille Vanderhoeft, Interuniversity Programme in Demography, Vrije Universiteit, Pleinlaan 2, B-1050 Brussels, Belgium.

2. A COMPOSITE LINK MODEL

Consider observations y_i (i=1,...N), with expected values μ_i. The μ_i depend on linear parameters $\beta_1,\dots \beta_p$ as follows :

$$\mu_i = c_i(\boldsymbol{\eta}_i), \tag{2.1}$$

where the $c_i(.)$ are known functions from R^k into R, and $\boldsymbol{\eta}_i$ is a vector $(\eta_{i1},\dots \eta_{ik})^T$ of k linear predictors which are defined as follows. Consider the block-diagonal matrix

$$\mathbf{X}_i = \begin{bmatrix} \mathbf{X}_{i1} & \mathbf{0} & & \\ \mathbf{0} & \mathbf{X}_{i2} & & \\ & & \cdot & \\ & & & \mathbf{X}_{ik} \end{bmatrix}, \tag{2.2}$$

the supervector

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \cdot \\ \cdot \\ \boldsymbol{\beta}_k \end{bmatrix}, \tag{2.3}$$

and their product

$$\mathbf{X}_i . \boldsymbol{\beta} = \begin{bmatrix} \mathbf{X}_{i1} . \boldsymbol{\beta}_1 \\ \cdot \\ \cdot \\ \mathbf{X}_{ik} . \boldsymbol{\beta}_k \end{bmatrix} = \begin{bmatrix} \eta_{i1} \\ \cdot \\ \cdot \\ \eta_{ik} \end{bmatrix} = \boldsymbol{\eta}_i . \tag{2.4}$$

Note that $\mathbf{X}_{ij}$ is a vector of covariates for the i-th observation and corresponding to the j-th linear predictor; $\boldsymbol{\beta}_j$ is the vector of linear parameters corresponding to the j-th linear predictor. It is not difficult to see that the model of Section 1 can be treated as a composite link model as defined here above; note that the linear predictors may indeed depend on different sets of covariates.

If $\boldsymbol{\beta}_j$ and $\mathbf{X}_{ij}$ have p_j components (so that $p=\Sigma_j p_j$), denoted respectively as β_{js} and X_{ijs} ($s=1,\ldots p_j$), and if $c'_{ij}(\boldsymbol{\eta}_i)$ is the j-th partial derivative of $c_i(\boldsymbol{\eta}_i)$ then we can prove the following theorem.

Theorem 1

If the y_i are independent with distribution belonging to Nelder and Wedderburn's (1972) exponential family

$$p(y_i,\theta_i,\phi)=\exp[\alpha_i(\phi)\{y_i.\theta_i-g(\theta_i)+h_0(y_i)\}+c(\phi,y_i)], \tag{2.5}$$

then the maximum likelihood estimates of the linear parameters of the above composite link model can be found by iterative reweighted least squares (IRLS) with

working dependent variables $z_i=\Sigma_{j,s}X^*_{ijs}.\beta_{js}+(y_i-\mu_i)$,

working independent variables $X^*_{i(js)}=c'_{ij}(\boldsymbol{\eta}_i).X_{ijs}$ ($j=1,\ldots k$; $s=1,\ldots p_j$), and

iterative weights $w_i=1/\tau_i^2$ (where $\tau_i^2=d\mu_i/d\theta_i$).

Corollary 1

If $p_1=\ldots=p_k=p°$, $\mathbf{X}_{ij}=\mathbf{X}_j=(X_{j1},\ldots X_{jp°})$for all observations i and $\boldsymbol{\beta}_j=\boldsymbol{\beta}=(\beta_1,\ldots\beta_{p°})^T$ for all $j=1,\ldots k$, then $\boldsymbol{\eta}_i=\boldsymbol{\eta}$ for all i and we can apply Theorem 1 with

working dependent variables $z_i=\Sigma_j\, c'_{ij}(\boldsymbol{\eta}).\eta_j+(y_i-\mu_i)$,

working independent variables $X^*_{is}=\Sigma_j\, c'_{ij}(\boldsymbol{\eta}).X_{js}$ ($s=1,\ldots p°$), and

iterative weights $w_i=1/\tau_i^2$ (where $\tau_i^2=d\mu_i/d\theta_i$).

We can write $c'_{ij}(\boldsymbol{\eta})=\partial\mu_i/\partial\eta_j$, hence Corollary 1 is a result discussed by

Thompson and Baker (1981:Section 5).

Corollary 2

If k=1 in Theorem 1, then the ML estimates of the linear parameters can be found by application of the IRLS procedure with

working dependent variables $z_i=\Sigma_s X_{is}.\beta_s+(y_i-\mu_i)/c'_i(\eta_i)$,

unchanged working independent variables X_{is} $(s=1,...p_1)$, and

iterative weights $w_i=c'_i(\eta_i)^2/\tau_i^2$ (where $\tau_i^2=d\mu_i/d\theta_i$).

This is Nelder and Wedderburn's (1972) original result for a single linear predictor.

Remarks

i) $E(y_i) = \mu_i = g'(\theta_i)$

ii) $var(y_i) = \tau_i^2/\alpha_i(\phi) = g''(\theta_i)/\alpha_i(\phi) = (d\mu_i/d\theta_i)/\alpha_i(\phi)$

iii) It is assumed that $\alpha_i(\phi)=\alpha_i/\phi$. The α_i are usually called *prior weights*.
For one-parameter exponential families $\phi=1$.

3. MULTIDIMENSIONAL HIDDEN HETEROGENEITY (HH)

In this text we will discuss the following models for, respectively, binomial and Poisson data :

$$y_i | \mathbf{X}_i, \mathbf{U}_i \approx \mathcal{B}(n_i, \mu_i) \qquad (3.1a)$$

$$\mu_i = c_i(\boldsymbol{\eta}_i + \mathbf{U}_i) = n_i \mathcal{F}(\boldsymbol{\eta}_i + \mathbf{U}_i) \qquad (3.1b)$$

$$\mathbf{U}_i \approx (\mathbf{0}, \boldsymbol{\Sigma}) \qquad (3.1c)$$

and

$$y_i | \mathbf{X}_i, \mathbf{U}_i \approx \mathcal{P}(\mu_i) \qquad (3.2a)$$

$$\mu_i = c_i(\boldsymbol{\eta}_i + \mathbf{U}_i) = n_i \mathcal{F}(\boldsymbol{\eta}_i + \mathbf{Z}_i) \qquad (3.2b)$$

$$\mathbf{U}_i \approx (\mathbf{0}, \boldsymbol{\Sigma}) \qquad (3.2c)$$

where (3.1c) means that the random vector $\mathbf{U}_i$, which represents the hidden or unobserved information for the i-th observation, has some (multidimensional) distribution with mean vector $\mathbf{0}$ and dispersion matrix $\boldsymbol{\Sigma}$. Note that μ_i is the conditional expectation of y_i, given $\mathbf{X}_i$ and $\mathbf{U}_i$. More details can be found in Vanderhoeft (1986).

In the next sections we will discuss how the parameters of the models (3.1) and (3.2) can be estimated.

4. APPROXIMATE MODELS - METHOD 1

4.1. The approximate models.

From (3.1) and (3.2) one can derive the following approximate model (cf. Vanderhoeft, 1986) :

$$E(y_i) \cong \mu_{io} \tag{4.1a}$$

$$\mathrm{var}(y_i) \cong \tau_i^2/\alpha_i \tag{4.1b}$$

where for binomial data : $\mu_{io} = c_i(\boldsymbol{\eta}_i) = n_i.\mathcal{F}(\boldsymbol{\eta}_i)$ (4.2a)

$$\tau_i^2 = \mu_{io}.(1-\mu_{io}/n_i) \tag{4.2b}$$

$$\alpha_i^{-1} = 1 + n_i(n_i-1)\mathbf{J}_i^T.\boldsymbol{\Sigma}.\mathbf{J}_i/\tau_i^2 \tag{4.2c}$$

and for Poisson data : $\mu_{io} = c_i(\boldsymbol{\eta}_i) = n_i.\mathcal{F}(\boldsymbol{\eta}_i)$ (4.3a)

$$\tau_i^2 = \mu_{io} \tag{4.3b}$$

$$\alpha_i^{-1} = 1 + n_i^2.\mathbf{J}_i^T.\boldsymbol{\Sigma}.\mathbf{J}_i/\tau_i^2 \tag{4.3c}$$

with

$$\mathbf{J}_i = \begin{bmatrix} \mathcal{F}_1(\boldsymbol{\eta}_i) \\ \cdot \\ \cdot \\ \mathcal{F}_k(\boldsymbol{\eta}_i) \end{bmatrix}$$

where $\mathcal{F}_j(\boldsymbol{\eta}_i)$ is the j-th partial derivative of $\mathcal{F}(\boldsymbol{\eta}_i)$.

4.2. Estimation of the parameters.

A general algorithm is as follows.

Step 1 : Take some initial estimates for the linear parameters and for $\boldsymbol{\Sigma}$, and compute initial estimates for the prior weights α_i.

Step 2 : Reestimate the linear parameters.

Step 3 : Reestimate $\boldsymbol{\Sigma}$, and recalculate the prior weights α_i. Check wether some convergence criterion is satisfied; if not go back to Step 2.

From Theorem 1 it follows that the new estimates of the linear parameters $\boldsymbol{\beta}$ (Step 2), given some estimates of the prior weights α_i, are found by iterative minimization of the quadratic form

$$Q(\boldsymbol{\beta}) = \Sigma_i\, \alpha_i . w_i .(z_i - \Sigma_{(js)}\, X^*_{i\,(js)} . \beta_{js})^2$$

$$= (\mathbf{Z}-\boldsymbol{\mathcal{X}}^*.\boldsymbol{\beta})^T.\, \mathbf{A}.\mathbf{W}.(\mathbf{Z}-\boldsymbol{\mathcal{X}}^*.\boldsymbol{\beta}) \qquad (4.4)$$

where

$\mathbf{Z} = (z_1, \ldots z_N)^T$

$\boldsymbol{\beta}$ is the supervector of linear parameters as defined in Section 2

$\boldsymbol{\mathcal{X}}^* = (X^*_{i\,(js)})$ (Nxp matrix)

$\mathbf{W} = \mathrm{diag}(w_i)$

$\mathbf{A} = \mathrm{diag}(\alpha_i)$.

(A proper justification of this statement should be based on quasi-likelihood estimation techniques.) New estimates of $\boldsymbol{\beta}$ in Step 2 are thus found by solving iteratively the equations

$$(\boldsymbol{\mathcal{X}}^{*T}.\mathbf{A}.\mathbf{W}.\,\boldsymbol{\mathcal{X}}^*).\ \boldsymbol{\beta} = \boldsymbol{\mathcal{X}}^{*T}.\mathbf{A}.\mathbf{W}.\ \mathbf{Z}. \qquad (4.5)$$

If the solution is $\boldsymbol{\beta}_{min}$ and if $Q_0 = Q(\boldsymbol{\beta}_{min})$, then it can be shown that

$$Q_0 = (\mathbf{Z}-\boldsymbol{\mathcal{X}}^*.\boldsymbol{\beta})^T.\mathbf{G}.(\mathbf{Z}-\boldsymbol{\mathcal{X}}^*.\boldsymbol{\beta}), \qquad (4.6)$$

with

$$\mathbf{G} = (\mathbf{I}-\mathbf{A}.\mathbf{W}.\boldsymbol{\mathcal{X}}^*(\boldsymbol{\mathcal{X}}^{*T}.\mathbf{A}.\mathbf{W}.\boldsymbol{\mathcal{X}}^*)^-\boldsymbol{\mathcal{X}}^{*T})\mathbf{A}.\mathbf{W} \qquad (4.7)$$

whence (using Rao(1973) : formula 4a.1.7) :

$$E(Q_0) = \mathrm{Tr}(\mathbf{G}.D(\mathbf{Z}-\boldsymbol{\mathcal{X}}^*.\boldsymbol{\beta})) + E(\mathbf{Z}-\boldsymbol{\mathcal{X}}^*.\boldsymbol{\beta})^T.\mathbf{G}.E(\mathbf{Z}-\boldsymbol{\mathcal{X}}^*.\boldsymbol{\beta}), \qquad (4.8)$$

where D(**H**) denotes the dispersion matrix of a vector **H**. But, since $E(Z-\mathbf{X}^*.\boldsymbol{\beta}) \cong \mathbf{0}$ and $D(Z-\mathbf{X}^*.\boldsymbol{\beta}) \cong \mathbf{W}^{-1}.\mathbf{B}$, with $\mathbf{B} = \text{diag}(\alpha_i^{-1}) = \text{diag}(b_i)$, it follows that

$$E(Q_0) \cong Tr((\mathbf{I}-\mathbf{A}.\mathbf{W}.\mathbf{X}^*(\mathbf{X}^{*T}.\mathbf{A}.\mathbf{W}.\mathbf{X}^*)^-\mathbf{X}^{*T})\mathbf{A}.\mathbf{B}). \qquad (4.9)$$

This formula can be used in Step 3 in order to find a new estimate of $\boldsymbol{\Sigma}$. Note that Williams (1982) suggests the Pearson chi-squared statistic obtained in Step 2 as an estimate of $E(Q_0)$ and that the unknowns only appear in the matrix **B**. The latter formula can be written more explicitly as

$$E(Q_0) \cong \Sigma_i\, \alpha_i.(1-\alpha_i.w_i.q_i).b_i. \qquad (4.10)$$

This formula is useful in GLIM, since the prior weights α_i are found in the system vector %PW, the variances q_i of the linear predictors in the system vector %VL and the iterative weights w_i in the system vector %WT. The only problem is to find an appropriate computational formula for the b_i (i.e. by rewriting the r.h.s. of (4.2c) or (4.3c)), but in some applications this too may be rather easy. For an example, see Vanderhoeft (1986 : Section 9).

Unfortunately, equation (4.10) generally involves k^2 unknown parameters σ_{jg} (i.e. the entries of the matrix $\boldsymbol{\Sigma}$), in which case there is no unique solution for (4.10). Thus only situations in which there is a single unknown can be solved with the methods described before. Such a particular situation may be, for instance, $\boldsymbol{\Sigma} = \sigma^2.\boldsymbol{\Sigma}_0$ where σ^2 is unknown but $\boldsymbol{\Sigma}_0$ is known.

5. APPROXIMATE MODELS - METHOD 2

5.1. The approximate model.

An alternative method for estimation of the parameters of models (3.1) and (3.3) is based on the *emperical transformation*. I.e. we consider for each observation y_i a function $\mathbf{g}_i(\,.\,)$ from R into R^k such that

$$\mu_i = c_i(\mathbf{g}_i(\mu_i)). \tag{5.1}$$

The emperical transformation (for observation y_i) is than defined as

$$\mathbf{e}_i = \mathbf{g}_i(y_i). \tag{5.2}$$

Note that

$$\mathbf{g}_i(\,.\,) = \begin{bmatrix} g_{i1}(\,.\,) \\ . \\ . \\ g_{ik}(\,.\,) \end{bmatrix}. \tag{5.3}$$

It can than be shown that

$$E(\mathbf{e}_i) \cong \boldsymbol{\eta}_i \tag{5.4a}$$
$$D(\mathbf{e}_i) \cong \boldsymbol{\Sigma} + \tau_i^2.(F_i.\, F_i^T) \tag{5.4b}$$

where τ_i^2 is defined by (4.2) or (4.3) and

$$\mathbf{F}_i = \begin{bmatrix} (n_i.\, \mathbf{f}_1(\boldsymbol{\eta}_i))^{-1} \\ . \\ . \\ (n_i.\, \mathbf{f}_k(\boldsymbol{\eta}_i))^{-1} \end{bmatrix} \tag{5.5}$$

where $\mathbf{f}_j(\boldsymbol{\eta}_i)$ is the j-th partial derivative of $\mathbf{f}(\boldsymbol{\eta}_i)$.

5.2. Estimation of the parameters - Single linear predictor.

If k =1, then (5.4) reduces to

$$E(e_i) \cong \eta_i \qquad (5.5a)$$

$$var(e_i) \cong \alpha_i^{-1} \qquad (5.5b)$$

where for binomial data : $\alpha_i^{-1} = \sigma^2 + n_i.\mathcal{F}(\eta_i)(1-\mathcal{F}(\eta_i))/(n_i.\mathcal{F}'(\eta_i))^2$ (5.6)

and for Poisson data : $\alpha_i^{-1} = \sigma^2 + n_i.\mathcal{F}(\eta_i)/(n_i.\mathcal{F}'(\eta_i))^2$. (5.7)

A suitable algorithm for estimation of the parameters is :

Step 1 : Take some initial estimates for the linear parameters and for σ^2, and compute initial estimates for the prior weights α_i and for the dependent variables e_i.

Step 2 : Reestimate the linear parameters.

Step 3 : Reestimate σ^2, and recalculate α_i and e_i. Check wether some convergence criterion is satisfied; if not go back to Step 2.

As in Section 4.2, the new estimates of the linear parameters $\boldsymbol{\beta}$ (=$\boldsymbol{\beta}$ since k =1), given some estimates of α_i and e_i, are found by iterative minimization of the quadratic form

$$Q(\boldsymbol{\beta}) = \Sigma_i\, \alpha_i.(e_i - \Sigma_s\, X_{is}.\beta_s)^2$$

$$= (\mathbf{E}-\boldsymbol{\mathcal{X}}.\boldsymbol{\beta})^T.\ \mathbf{A}.(\mathbf{E}-\boldsymbol{\mathcal{X}}.\boldsymbol{\beta}) \qquad (5.8)$$

where

$\mathbf{E} = (e_1,... e_N)^T$

$\boldsymbol{\beta} = (\beta_1,... \beta_p)^T$ is the vector of linear parameters (Section 2)

$\boldsymbol{\mathcal{X}} = (X_{is})$ (Nxp matrix)

$\mathbf{A} = diag(\alpha_i)$.

As in Section 4.2 one can show that

$$E(Q_0) \cong Tr((\mathbf{I}-\mathbf{A}.\boldsymbol{\mathcal{X}}(\boldsymbol{\mathcal{X}}^T.\mathbf{A}.\boldsymbol{\mathcal{X}})^-\boldsymbol{\mathcal{X}}^T)\mathbf{A}.\mathbf{B}), \qquad (5.9)$$

or more explicitely :

$$E(Q_0) \cong \Sigma_i\, \alpha_i.(1-\alpha_i.q_i).b_i. \qquad (5.10)$$

Substitution of the r.h.s. of (5.6) or (5.7) for the b_i then yields a simple formula from which σ^2 can easily be reestimated (Step 3).

5.3. Estimation of the parameters - General case.

We consider first the linear transformation

$$\mathbf{e}_i^* = \boldsymbol{\Omega}_i.\mathbf{e}_i \qquad (5.11a)$$

$$\boldsymbol{\Omega}_i = [\boldsymbol{\Sigma} + \tau_i^2.(F_i.\, F_i^T)]^{-1/2}. \qquad (5.11b)$$

Then

$$E(\mathbf{e}_i^*) \cong \boldsymbol{\eta}_i^* = \boldsymbol{\mathcal{X}}_i^*.\boldsymbol{\beta} \qquad (5.12a)$$

$$D(\mathbf{e}_i^*) \cong \mathbf{I}, \qquad (5.12b)$$

where $\boldsymbol{\mathcal{X}}_i^* = \boldsymbol{\Omega}_i.\boldsymbol{\mathcal{X}}_i$. Thus the independent variables in the model (5.12) also depend on the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$. (Note the difference between models (5.5) and (5.12). Of course, the case k =1 can also be treated as a special case of the general model (5.12).) However, an appropriate algorithm based on ordinary linear models is easily constructed :

- Step 1 : Take some initial estimates for the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ (or for $\boldsymbol{\Omega}_i$), and compute initial estimates for the independent variables $\boldsymbol{\mathcal{X}}_i^*$ and the dependent variables $\mathbf{e}_i^*$.
- Step 2 : Reestimate the linear parameters.
- Step 3 : Reestimate $\boldsymbol{\Sigma}$, and recalculate $\boldsymbol{\Omega}_i$, $\boldsymbol{\mathcal{X}}_i^*$ and $\mathbf{e}_i^*$. Check wether some convergence criterion is satisfied; if not go back to Step 2.

Reestimation of $\boldsymbol{\beta}$ is done through minimization of the quadratic form

$$Q(\beta) = \Sigma_i (\mathbf{e}_i^* - \mathbf{X}_i^* . \beta)^T . (\mathbf{e}_i^* - \mathbf{X}_i^* . \beta)$$
$$= (E^* - \mathbf{X}^* . \beta)^T . (E^* - \mathbf{X}^* . \beta), \qquad (5.13)$$

where

$$\mathbf{E}^* = (\mathbf{e}_1^{*T}, \ldots \mathbf{e}_N^{*T})^T \quad \text{(Nkx1 vector)}$$
$$\mathbf{X}^* = (\mathbf{X}_1^{*T}, \ldots \mathbf{X}_N^{*T})^T \quad \text{(Nkxp matrix).}$$

And for reestimation of $\mathbf{\Sigma}$, one can use the formula

$$E(Q_0) \cong Tr(\mathbf{I} - \mathbf{X}^*(\mathbf{X}^{*T} . \mathbf{X}^*)^- \mathbf{X}^{*T}). \qquad (5.9)$$

6. NUMERICAL EXAMPLES - MODELS WITH SIMPLE LINK

6.1. Application to binomial data.

We apply the methods discussed in Sections 4 and 5 to the binomial seed-data used by Crowder (1978) and by Williams (1982) and Vanderhoeft (1986). See the latter paper for more details. In order to have comparible results from the two methods, we add 1/2 to the number of germinated seeds (DX→DX+0.5) and 1 to the batch sizes (NX→NX+1), which corresponds to Williams' (1982) definition of the emperical logits.

We apply three models - with simple link (i.e. k=1) - specified as follows:

$\mathcal{F}(\eta_i) = 1/(1+\exp(-\eta_i))$	for the logit model,
$= 1-\exp(-\exp(\eta_i))$	for the cloglog model,and
$= \int_{-\infty}^{\eta_i}(2\pi)^{-1}\exp(-x^2/2)dx$	for the probit model.

The Method 1 formula (4.2c) for the prior weights has the simple form

$$\alpha_i^{-1} = 1 + \sigma^2(1-1/n_i)w_i$$

where the w_i are the iterative weights (Section 2). This and the corresponding formula for reestimation of σ^2 are easy to handle in GLIM3 (Baker and Nelder, 1978). The GLIM3 program is shown in Appendix A.1.; details can be found in Vanderhoeft (1986). This program has been used in order to produce Table 1.

The Method 2 formula (5.6) for the prior weights becomes

$\alpha_i^{-1} = \sigma^2 + (n_i\mathcal{F}(\eta_i).(1-\mathcal{F}(\eta_i))^{-1}$	for the logit model,
$= \sigma^2 + \mathcal{F}(\eta_i)/[n_i(1-\mathcal{F}(\eta_i))(\log(1-\mathcal{F}(\eta_i)))^2]$	for the cloglog model,and
$= \sigma^2 + \mathcal{F}(\eta_i)(1-\mathcal{F}(\eta_i))\exp[\eta_i^2 + \log(2\pi) - \log(n_i)]$	for the probit model.

The GLIM3 program can be constructed as before (although it is slightly more complicated now); it is shown in Appendix A.2. The results from application of this program to the binomial seed-data are reproduced in Table 2.

The $\hat{\chi}^2$-values and the ANOVA tables from Methods 1 and 2 (Tables 1 and 2) are comparible, which gives equivalent significance testing. On the contrary, the analysis shows that estimates of σ^2 can be quite different between the two Methods.

Table 1. Binomial data.
Method 1; DX→DX+0.5, NX→NX+1.0

GLIM program options and results. Binomial data.

Contents of MOD (1)	LINK option (2)	Initial σ^2 (%P) (3)	Fix/est. σ^2 (%F) (4)	Nbr. of iter.(%R) (5)	$\hat{\sigma}^2$ (6)	Heter. $\hat{\chi}^2$ (7)	No heter. $\hat{\chi}^2$ (8)
X1*X2	G	.0	1	5	.0824	17.00	29.69
X1*X2	C	.0	1	5	.0430	17.00	29.69
X1*X2	P	.0	1	5	.0314	17.00	29.69
X1+X2	G	.0824	0	5	-	20.98	36.39
X1+X2	C	.0430	0	5	-	20.53	35.43
X1+X2	P	.0314	0	5	-	21.01	36.39
X1	G	.0824	0	5	-	43.50	88.50
X1	C	.0430	0	5	-	42.81	88.50
X1	P	.0314	0	5	-	44.08	88.50
X2	G	.0824	0	5	-	23.34	39.23
X2	C	.0430	0	5	-	22.99	39.24
X2	P	.0314	0	5	-	23.38	39.23
-	G	.0824	0	5	-	45.44	90.72
-	C	.0430	0	5	-	45.35	90.72
-	P	.0314	0	5	-	46.06	90.72
X1+X2	G	.0	1	5	.1196	18.00	36.39
X1+X2	C	.0	1	5	.0597	18.00	35.43
X1+X2	P	.0	1	5	.0457	18.00	36.39

Table 1. Cont'd

ANOVA tables.

Panel A : Heterogeneity between replicates.

Source	d.f.	Link : Logit	Cloglog	Probit
Interaction	1	3.98	3.53	4.01
X1	1	2.36	2.46	2.37
X2	1	22.52	22.28	23.07
Main effects	2	24.46	24.82	25.05

Panel B : No heterogeneity between replicates.

Source	d.f.	Link : Logit	Cloglog	Probit
Interaction	1	6.70	5.74	6.70
X1	1	2.84	3.81	2.84
X2	1	52.11	53.07	52.11
Main effects	2	54.33	55.29	54.33

Table 2. Binomial data.
Method 2; DX→DX+0.5, NX→NX+1.0

GLIM program options and results.

Contents of MOD (1)	Link (%A) (2)	Initial σ^2 (%P) (3)	Fix/est. σ^2 (%F) (4)	Nbr. of iter.(%R) (5)	$\hat{\sigma}^2$ (6)	Heter. $\hat{\chi}^2$ (7)	No heter. $\hat{\chi}^2$ (8)
X1*X2	1	.0	1	5	.1092	17.00	28.42
X1*X2	2	.0	1	5	.0655	17.00	27.92
X1*X2	3	.0	1	5	.0385	17.00	29.31
X1+X2	1	.1092	0	5	-	20.21	34.55
X1+X2	2	.0655	0	5	-	19.40	33.51
X1+X2	3	.0385	0	5	-	20.46	35.68
X1	1	.1092	0	5	-	44.50	81.80
X1	2	.0655	0	5	-	41.87	81.17
X1	3	.0385	0	5	-	45.11	87.01
X2	1	.1092	0	5	-	23.13	36.86
X2	2	.0655	0	5	-	23.06	36.18
X2	3	.0385	0	5	-	23.25	38.27
-	1	.1092	0	5	-	47.42	83.44
-	2	.0655	0	5	-	44.34	85.09
-	3	.0385	0	5	-	47.82	89.04
X1+X2	1	.0	1	5	.1416	18.00	34.55
X1+X2	2	.0	1	5	.0777	18.00	33.51
X1+X2	3	.0	1	5	.0516	18.00	35.68

Table 2. Cont'd

ANOVA tables.

Panel A : Heterogeneity between replicates.

Source	d.f.	Link : Logit	Cloglog	Probit
Interaction	1	3.21	2.40	3.46
X1	1	2.92	3.66	2.79
X2	1	24.39	22.47	24.65
Main effects	2	27.21	24.94	27.36

Panel B : No heterogeneity between replicates.

Source	d.f.	Link · Logit	Cloglog	Probit
Interaction	1	6.13	5.59	6.37
X1	1	2.31	2.67	2.59
X2	1	47.25	47.66	51.33
Main effects	2	48.89	51.58	53.36

6.2. Application to Poisson data.

Next, we apply Method 1 and Method 2 to the Poisson data analysed previously by Margolin *et al.* (1981) and Breslow (1984). See the latter paper for more details.

The model used is log-linear - with a simple link - :

$$f(\eta_i) = \exp(\eta_i),$$

with $n_i=1$.

The Method 1 formula (4.3c) for the prior weights now becomes

$$\alpha_i^{-1} = 1 + \sigma^2 w_i,$$

where the w_i are the iterative weights (Section 2).It is thus again not difficult to write an appropriate GLIM3 program. This is shown in Appendix A.3. Table 3 shows the results from application of this analysis. They are comparible with Breslow's (1984) results. The figures between parenthesis are the estimates for σ^2 obtained by using the GLIM3-macros given in Breslow (1984 : Appendix). Note that Breslow's method is only computationally different from our Method 1.

The Method 2 formula (5.7) for the prior weights is

$$\alpha_i^{-1} = \sigma^2 + \exp(-\eta_i).$$

The appropriate GLIM3-program is shown in Appendix A.4. and the results are presented in Table 4. Again, inference making from Method 1 and Method 2 would be equivalent, particularly if the model includes hidden heterogeneity. But, as for the binomial data, the estimates of σ^2 can be quite different.

Table 3. Poisson data.
Method 1.

GLIM program options and results.

Contents of MOD (1)	LINK option (2)	Initial σ^2 (%P) (3)	Fix/est. σ^2 (%F) (4)	Nbr. of iter.(%R) (5)	$\hat{\sigma}^2$ (6)	Heter. $\hat{\chi}^2$ (7)	No heter. $\hat{\chi}^2$ (8)
X1+X2	L	.0	1	5	.0717 (.0718)	15.00	46.23
X1	L	.0717	0	5	-	20.84	65.48
X2	L	.0717	0	5	-	26.57	81.40
-	L	.0717	0	5	-	26.77	82.85
X1	L	.0	1	5	.1037 (.1039)	16.00	65.48

ANOVA tables.

Panel A : Heterogeneity between replicates.

Source	d.f.	Link : Log
X1	1	11.57
X2	1	5.84
Main effects	2	11.77

Panel B : No heterogeneity between replicates.

Source	d.f.	Link : Log
X1	1	35.17
X2	1	19.25
Main effects	2	36.62

Table 4. Poisson data.
Method 2.

GLIM program options and results

Contents of MOD (1)	LINK option (2)	Initial σ^2 (%P) (3)	Fix/est. σ^2 (%F) (4)	Nbr. of iter.(%R) (5)	$\hat{\sigma}^2$ (6)	Heter. $\hat{\chi}^2$ (7)	No heter. $\hat{\chi}^2$ (8)
X1+X2	L	.0	1	5	.0596 (.0611)	15.00	44.58
X1	L	.0596	0	5	-	20.70	61.95
X2	L	.0596	0	5	-	26.03	76.15
-	L	.0596	0	5	-	27.53	77.60
X1	L	.0	1	5	.0875 (.0873)	16.00	65.48

ANOVA tables.

Panel A : Heterogeneity between replicates.

Source	d.f.	Link : Log
X1	1	11.03
X2	1	5.70
Main effects	2	12.53

Panel B : No heterogeneity between replicates.

Source	d.f.	Link : Log
X1	1	31.57
X2	1	17.37
Main effects	2	33.02

7. SIMULATION - MODELS WITH COMPOSITE LINK

In this section we present some results from the analysis of simulated binomial data through a composite link model (i.e. k=2) following Method 1 (Section 4). The composite link is defined as :

$$F(\eta_i) = \Phi((x_i+\eta_{i2})/\exp(-\eta_{i1})),$$

where $\Phi(.)$ is the cumulative standard normal distribution function. Note that the partial derivatives are :

$$F_1(\eta_i) = \varphi((x_i+\eta_{i2})/\exp(-\eta_{i1})).((x_i+\eta_{i2})/\exp(-\eta_{i1})),$$
$$F_2(\eta_i) = \varphi((x_i+\eta_{i2})/\exp(-\eta_{i1}))/\exp(-\eta_{i1}),$$

where $\varphi(.)$ is the standard normal probability density function. The data used are given in the form (y_i, n_i, x_i, $A1_i$, $A2_i$) for i=1,... N, with dichotomous independent variables or covariates A1 and A2 (values 0 and 1), and y_i "items" out of n_i - all characterized by covariates equal to $A1_i$ and $A2_i$ - which have "failed" by "time" x_i. The linear predictors are

$$\eta_{ij} = X_{ij1}.\beta_{j1} + X_{ij2}.\beta_{j2} + X_{ij3}.\beta_{j3} + X_{ij4}.\beta_{j4},$$

where X_{ij1} =1 for all i and j, $X_{ijs} = A(s-1)_i$ if the model assumes that covariate A(s-1) has an effect on the j-th linear predictor and X_{ijs} =0 otherwise (s=2,3), $X_{ij4} = A1_i.A2_i$ if the model assumes that A1 and A2 have an interaction effect on the j-th linear predictor and X_{ij4} =0 otherwise. (Note that x, which is in fact a third independent variable, could be included in the linear predictors. The corresponding β's would have known values 0 and 1, respectively.)

For simulation of the data we set n_i=200, β_{11}=-1.386, β_{12}=0, β_{13}=-0.5, β_{14}=0, β_{21}=-25, β_{22}=2, β_{23}=0 and β_{24}=0. The reference population A1=A2=0 has then a mean failure time of 25.0 with a variance of 16.0, given $F(\eta_i)$ as above. The simulation procedure is as follows. Firstly, we generate N independent random vectors $\mathbf{U}_i=(U_{i1},U_{i2})^T$, assuming a bivariate normal distribution with mean vector **0** and some covariance matrix $\Sigma=(\sigma_{jl})$. (We will use the notation $\sigma_j^2 = \sigma_{jj}$, j=1,2. hereafter.) Note that the bivariate normal distribution may be replaced by any other bivariate distribution. Secondly, we compute the vector of linear predictors $\boldsymbol{\eta}_i$ and add the hidden heterogeneity vector $\mathbf{U}_i$. Thirdly, we generate for each "batch" i n_i independent random failure times and count the number y_i of failure times smaller than or equal to x_i. In the following paragraphs we present some results using simulated data given that σ_2^2=.01 and $\sigma_1^2 = \sigma_{12} = \sigma_{21}$=0 (Data I) or given that σ_1^2=.01 and $\sigma_2^2 = \sigma_{12} = \sigma_{21}$=0 (Data II).

Appendix A.5. shows a GLIM3 program to fit the composite link model as described here before. The general structure of this program is the same as for the programs in Appendices A.1. to A.4. Details are therefore not given. Note only that we use the assumption $\boldsymbol{\Sigma} = \sigma^2.\ \boldsymbol{\Sigma}_0$ where $\boldsymbol{\Sigma}_0$ is a "normalised" form of the covariance matrix used in generating the data. The program in Appendix A.5. fits the *starting model* (see Tables 5 and 6), including the estimation of σ_2^2 (or σ^2), and all *submodels* given the estimates for σ_2^2 and for the prior weights. The estimate for σ_2^2 is .0723 and the estimated Pearson chi-squared statistics can be found in Table 5 (col. 2a). (Of course, for the starting model we find all estimated Pearson chi-squared statistics on the same output.)

Thus, Table 5 and Table 6 present results from fitting various composite link models - ignoring or including hidden heterogeneity - to Data I and Data II respectively. The differences between estimated Pearson chi-squared statistics show once again that model selection generally depends on the method used : effects of covariates can be significant when hidden heterogeneity is ignored, while they are not significant when hidden heterogeneity is taken into account.

The ideas presented in this section are merely indicative for further research. A lot of problems have to be solved. We intend to do this via more detailed and refined simulations. Such simulations are certainly useful in learning how composite link models - which are extended to incorporate hidden heterogeneity - can be applied for analysis of real data sets.

Table 5. Pearson chi-squared statistics: (1) HH ignored; (2) HH taken into account (σ^2=.0723 fixed): (2a) prior weights fixed,(2b) prior weights reestimated.

Model terms	D.F.	(1)	(2a)	(2b)
Starting model				
X11,X12, X13, X21,X22, X23	78	91.7	78.0	78.0
Submodels :				
X11,X12, X21,X22, X23	79	506.2	417.9	431.3
X11, X21,X22, X23	80	507.2	422.8	432.5
X11, X13, X21,X22, X23	79	92.2	78.7	79.3
X11,X12, X13, X21,X22	79	91.6	77.9	77.9
X11,X12, X21,X22	80	509.8	421.0	434.5
X11, X21,X22	81	511.4	426.0	435.7
X11, X13, X21,X22	80	92.2	78.6	79.2
X11,X12, X13, X21	80	347.9	279.3	278.4
X11,X12, X21	81	800.1	660.9	672.2
X11, X21	82	795.7	664.7	672.5
X11, X13, X21	81	348.2	279.0	278.7
X11,X12, X13, X21, X23	79	348.1	279.5	278.5
X11,X12, X21, X23	80	798.2	659.2	671.2
X11, X21, X23	81	793.9	662.9	671.0
X11, X13, X21, X23	80	348.4	279.2	278.8
Effect* of A1 on slope k	1	0.56	0.73	1.27
Effect* of A2 on slope k	1	414.54	339.90	353.30
Effect* of A1 on shift b	1	256.44	201.50	200.50
Effect* of A2 on shift b	1	-0.08	-0.09	-0.06

* Controlled for all other main effects

Table 6. Pearson chi-squared statistics: (1) HH ignored; (2) HH taken into account (σ^2=.0656 fixed): (2a) prior weights fixed,(2b) prior weights reestimated.

Model terms	D.F.	(1)	(2a)	(2b)
Starting model				
X11,X12, X13, X21,X22, X23	78	308.4	78.0	78.0
Submodels :				
X11,X12, X21,X22, X23	79	669.2	146.5	153.0
X11, X21,X22, X23	80	669.5	147.3	153.0
X11, X13, X21,X22, X23	79	307.9	78.8	78.6
X11,X12, X13, X21,X22	79	351.3	79.0	79.1
X11,X12, X21,X22	80	687.1	146.1	153.7
X11, X21,X22	81	688.3	147.0	153.6
X11, X13, X21,X22	80	351.5	79.8	79.6
X11,X12, X13, X21	80	633.9	162.4	169.7
X11,X12, X21	81	1045.0	231.3	243.0
X11, X21	82	1040.0	231.3	243.3
X11, X13, X21	81	636.7	160.7	169.5
X11,X12, X13, X21, X23	79	565.9	161.0	167.6
X11,X12, X21, X23	80	1023.0	231.5	241.6
X11, X21, X23	81	1019.0	231.6	241.9
X11, X13, X21, X23	80	590.4	160.4	167.4
Effect* of A1 on slope k	1	-0.50	0.84	0.56
Effect* of A2 on slope k	1	360.80	68.50	75.00
Effect* of A1 on shift b	1	257.50	83.00	89.60
Effect* of A2 on shift b	1	42.90	1.02	1.08

* Controlled for all other main effects

REFERENCES

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

Breslow N.E. (1984) : Extra-Poisson Variation in Log-linear Models. Applied Statistics, 33, No. 1, pp. 38-44.

Crowder M.J. (1978) : Beta-binomial Anova for Proportions. Applied Statistics, 27, No. 1, pp. 34-37.

Margolin B.H., Kaplan N. and Zeiger E. (1981) : Statistical analysis of the Ames *Salmonella* /microsome test. Proc. Natl. Acad. Sci. USA, 76, pp. 3779-3783.

Nelder J.A. and R.W.M. Wedderburn (1972) : Generalized Linear Models. J. R. Statist. Soc. A, 135, No. 3, pp. 370-384.

Thompson R and R.J. Baker (1981) : Composite Link Functions in Generalized Linear Models. Applied Statistics, 30, No.2, pp.125-131.

Vanderhoeft C. (1986) : Heterogeneity in Composite Link Models. Working Paper 86-60. IIASA, Laxenburg, Austria.

Williams D.A. (1982) : Extra-binomial Variation in Logistic Linear Models. Applied Statistics, 31, No. 2, pp. 144-148.

APPENDICES : GLIM3 programs

A.1. Binomial data, Method 1 - simple links.

```
$C MACROS

$M MOD X1*X2 $ENDM

$M ESTP !
 $EXT %VL $CA WVQ=%PW*(1-%PW*%WT*%VL) !
 : %P=(%X2-%CU(WVQ))/%CU(WVQ*A) !
 $PR : : " NEW SIGMA2 = " %P : $$ENDM

$M HET !
 $CA %R=%R-1 : A=(1-1/%BD)*%WT !
 $SW %F ESTP !
 $CA W=1/(1+%P*A) $W W !
 $F #MOD $PR : : " NEW CHI2    = " %X2 : $$ENDM

$C 'MAIN' PROGRAM

$UNITS 21
$DATA DX NX X1 X2 $READ
 10 39 1 1
 23 62 1 1
 23 81 1 1
 26 51 1 1
 17 39 1 1
  5  6 1 2
 53 74 1 2
 55 72 1 2
 32 51 1 2
 46 79 1 2
 10 13 1 2
  8 16 2 1
 10 30 2 1
  8 28 2 1
 23 45 2 1
  0  4 2 1
  3 12 2 2
 22 41 2 2
 15 30 2 2
 32 51 2 2
  3  7 2 2
$FACTOR X1 2 X2 2
$CALC DX=DX+0.5 : NX=NX+1.0
$YVAR DX $ERR B NX $LINK G

$C CHOOSE (INITIAL) VALUE FOR SIGMA2 : %P
   CHOOSE PROGRAM CONSTANTS :  %F=0 IF SIGMA2 IS FIXED
                                  1 IF SIGMA2 IS REESTIMATED
                               %R=MAX NBR. OF ITERATIONS
                                  0 IF NO HETEROGENEITY IS TAKEN INTO
                                    ACCOUNT (I.E. %P=%F=0)
$CALC %P=.0000 : %F=1 : %R=5
$PR : : " (INITIAL) SIGMA2 = " %P :
$CALC W=1 $WEIGHT W
$FIT #MOD $DISP M A
$PR : : " CHI2 (NO HETER.) = " %X2 :
$WHILE %R HET $DISP M A
$STOP
```

A.2. Binomial data, Method 2 - simple links.

```
$C MACROS

$M MOD X1*X2 $ENDM

$M ESTP !
 $EXT %VL $CA WVQ=%PW*(1-%PW*%VL) !
 : %P=(%X2-%CU(WVQ*A))/%CU(WVQ) !
 $PR :: " NEW SIGMA2 = " %P : $$ENDM

$M HET !
 $CA %R=%R-1 $SW %A ALOG ACLL APRO !
 $SW %F ESTP !
 $CA W=1/(%P+A) $W W !
 $F #MOD $PR :: " NEW CHI2    = " %X2 : $$ENDM

$M ALOG !
 $CA FP=1/(1+%EXP(-%LP)) !
 $CA A=1/(NX*FP*(1-FP)) $$ENDM
$M ACLL !
 $CA FP=1-%EXP(-%EXP(%LP)) !
 $CA A=FP/(NX*(1-FP)*%LOG(1-FP)**2) $$ENDM
$M APRO !
 $CA FP=%NP(%LP) !
 $CA A=FP*(1-FP)*%EXP(%FV*%FV+%LOG(2*%PI)-%LOG(NX)) $$ENDM

$M WLOG !
 $PR :: " LOGIT MODEL " : !
 $CA W=NX*OP*(1-OP) $$ENDM
$M WCLL !
 $PR :: " C-LOGLOG MODEL " : !
 $CA W=%LOG(1-OP)**2*NX*(1-OP)/OP $$ENDM
$M WPRO !
 $PR :: " PROBIT MODEL " : !
 $CA W=%EXP(-ET*ET-%LOG(2*%PI)+%LOG(NX))/(OP*(1-OP)) $$ENDM

$M ELOG $CA ET=%LOG(OP/(1-OP)) $$ENDM
$M ECLL $CA ET=%LOG(-%LOG(1-OP)) $$ENDM
$M EPRO $CA ET=%ND(OP) $$ENDM

$C 'MAIN' PROGRAM

$UNITS 21
$DATA DX NX X1 X2 $READ
 10 39 1 1
 23 62 1 1
 23 81 1 1
 26 51 1 1
 17 39 1 1
  5  6 1 2
 53 74 1 2
 55 72 1 2
 32 51 1 2
 46 79 1 2
 10 13 1 2
  8 16 2 1
 10 30 2 1
  8 28 2 1
 23 45 2 1
  0  4 2 1
  3 12 2 2
 22 41 2 2
 15 30 2 2
 32 51 2 2
  3  7 2 2
$FACTOR X1 2 X2 2
$CALC DX=DX+0.5 : NX=NX+1.0
 : OP=DX/NX
```

```
$C CHOOSE (INITIAL) VALUE FOR SIGMA2 : %P
   CHOOSE PROGRAM CONSTANTS : %F=0 IF SIGMA2 IS FIXED
                                 1 IF SIGMA2 IS REESTIMATED
                              %R=MAX NBR. OF ITERATIONS
                                 0 IF NO HETEROGENEITY IS TAKEN INTO
                                   ACCOUNT (I.E. %P=%F=0)
                              %A=1 FOR LOGIT MODEL
                                 2 FOR C-LOGLOG MODEL
                                 3 FOR PROBIT MODEL
$CALC %P=.0000 : %F=1 : %R=5 : %A=1

$C INITIALIZE

$SWITCH %A ELOG ECLL EPRO
$YVAR ET $SCALE 1
$PR :: " (INITIAL) SIGMA2 = " %P :
$SWITCH %A WLOG WCLL WPRO $WEIGHT W

$C FIT

$FIT #MOD $DISP A
$PR :: " CHI2 (NO HETER.) = " %X2 :
$WHILE %R HET $DISP A
$STOP
```

A.3. Poisson data, Method 1 - simple link.

```
$C MACROS

$M MOD X1+X2 $ENDM

$M ESTP !
 $EXT %VL $CA WVQ=%PW*(1-%PW*%WT*%VL) !
 : %P=(%X2-%CU(WVQ))/%CU(WVQ*A) !
 $PR :: " NEW SIGMA2 = " %P : $$ENDM

$M HET !
 $CA %R=%R-1 : A=%WT !
 $SW %F ESTP !
 $CA W=1/(1+%P*A) $W W !
 $F #MOD $PR :: " NEW CHI2    = " %X2 : $$ENDM

$C 'MAIN' PROGRAM

$UNITS 18
$DATA DX X $READ
 15 0    21 0    29 0
 16 10   18 10   21 10
 16 33   26 33   33 33
 27 100 41 100 60 100
 33 333 38 333 41 333
 20 1000 27 1000 42 1000
$CALC X1=%LOG(X+10) : X2=-X
$YVAR DX $ERR P $LINK L

$C CHOOSE (INITIAL) VALUE FOR SIGMA2 : %P
   CHOOSE PROGRAM CONSTANTS : %F=0 IF SIGMA2 IS FIXED
                                 1 IF SIGMA2 IS REESTIMATED
                              %R=MAX NBR. OF ITERATIONS
                                 0 IF NO HETEROGENEITY IS TAKEN INTO
                                   ACCOUNT (I.E. %P=%F=0)
$CALC %P=.0000 : %F=1 : %R=5
$PR :: " (INITIAL) SIGMA2 = " %P :
$CALC W=1 $WEIGHT W
$FIT #MOD $DISP M A R
$PR :: " CHI2 (NO HETER.) = " %X2 :
$WHILE %R HET $DISP M A R
$STOP
```

A.4. Poisson data, Method 2 - simple link.

```
$C MACROS

$M MOD X1+X2 $ENDM

$M ESTP !
 $EXT %VL $CA WVQ=%PW*(1-%PW*%VL) !
 : %P=(%X2-%CU(WVQ*A))/%CU(WVQ) !
 $PR :: " NEW SIGMA2 = " %P : $$ENDM

$M HET !
 $CA %R=%R-1 : A=1/%EXP(%FV) !
 $SW %F ESTP !
 $CA W=1/(%P+A) $W W !
 $F #MOD $PR :: " NEW CHI2    = " %X2 : $$ENDM

$M OUTP !
 $CA R=%EXP(%FV) $LO X R $$ENDM

$C 'MAIN' PROGRAM

$UNITS 18
$DATA DX X $READ
 15 0    21 0    29 0
 16 10   18 10   21 10
 16 33   26 33   33 33
 27 100  41 100  60 100
 33 333  38 333  41 333
 20 1000 27 1000 42 1000
$CALC X1=%LOG(X+10) : X2=-X
$CALC ET=%LOG(DX)
$YVAR ET $SCALE 1

$C CHOOSE (INITIAL) VALUE FOR SIGMA2 : %P
   CHOOSE PROGRAM CONSTANTS : %F=0 IF SIGMA2 IS FIXED
                                 1 IF SIGMA2 IS REESTIMATED
                              %R=MAX NBR. OF ITERATIONS
                                 0 IF NO HETEROGENEITY IS TAKEN INTO
                                   ACCOUNT (I. E. %P=%F=0)
$CALC %P=.0000 : %F=1 : %R=5
$PR :: " (INITIAL) SIGMA2 = " %P :
$CALC W=DX $WEIGHT W
$FIT #MOD $DISP M A $USE OUTP
$PR :: " CHI2 (NO HETER.) = " %X2 :
$WHILE %R HET $DISP M A $USE OUTP
$STOP
```

A.5. Binomial data, Method 1 -composite link.

```
$C MACROS

$M FSB $F . $D A !
 $PR : " CHI2 = " %X2 : $$ENDM
$M SUB1 !
 $U FSB !
 $CA A12=0 $U FSB !
 $CA A11=0 $U FSB !
 $CA A12=A2-1 $U FSB !
 $CA A11=A1-1 $$ENDM
$M SUB0 !
 $U SUB1 !
 $CA A22=0 $U SUB1 !
 $CA A21=0 $U SUB1 !
 $CA A22=A2-1 $U SUB1 $$ENDM

$M MEXT $EXT %PE $$ENDM

$M FV ! CALCULATE FITTED VALUES
 $CA %FV=N*%NP((X+LP2)*%EXP(LP1)) $$ENDM

$M DENS ! CALCULATE DENSITY FUNCTION FOR '%NP'
 $CA F=(X+LP2)*%EXP(LP1) !
 : F=%EXP(-(%LOG(2*%PI)+F*F)/2) $$ENDM

$M PD ! CALCULATE PARTIAL DERIVATIVES
 $U DENS $CA C2=N*F*%EXP(LP1) !
 : C1=C2*(X+LP2) $$ENDM

$M M1 !
 $CA %A=%NE(%PL,0) $SW %A MEXT !
! CALCULATE LINEAR PREDICTORS
 $CA LP1=%PE(1)+%PE(2)*A11+%PE(3)*A12+%PE(4)*A13 !
 : LP2=%PE(5)+%PE(6)*A21+%PE(7)*A22+%PE(8)*A23 !
! CALCULATE FITTED VALUES AND PARTIAL DERIVATIVES
 $U FV $U PD !
! CALCULATE WORKING INDEPENDENT VARIABLES
 $CA X11=C1 : X12=C1*A11 : X13=C1*A12 : X14=C1*A13 !
 : X21=C2 : X22=C2*A21 : X23=C2*A22 : X24=C2*A23 !
! CALCULATE LINEAR PREDICTOR
 : %LP=%PE(1)*X11+%PE(2)*X12+%PE(3)*X13+%PE(4)*X14 !
 +%PE(5)*X21+%PE(6)*X22+%PE(7)*X23+%PE(8)*X24 !
$$ENDM

$M M2 $CA %DR=1 $$ENDM

$M M3 !
 $CA %VA=%FV*(1-%FV/N) : %VA=%IF(%LE(%VA,0),.001,%VA) $$ENDM

$M M4 $CA %DI=2*(%YV*%LOG(%YV/%FV)+(N-%YV)*%LOG((N-%YV)/(N-%FV))) $$ENDM

$M HET !
 $CA %R=%R-1 !
 : B=(1-1/N)*(%X*C1*C1+%Y*C2*C2+2*%Z*C1*C2)/%VA !
 $SW %F ESTP !
 $CA W=1/(1+%P*B) $W W !
 $F . $PR :: " NEW CHI2 = " %X2 : $$ENDM

$M ESTP !
 $EXT %VL $CA WVQ=%PW*(1-%PW*%VL/%VA) !
 : %P=(%X2-%CU(WVQ))/%CU(WVQ*B) !
 $PR :: " NEW SIGMA2 = " %P : $$ENDM
```

```
$C PROGRAM CONSTANTS AND PARAMETERS

$DATA 1 NU $DINPUT 3
$DATA 8 PAR $DINPUT 3
$DATA 3 S $DINPUT 3
$CALC %U=NU : %F=1 : %R=5
 : %P=%IF(%GE(S(1),S(2)),S(1),S(2)) : S=S/%P
 : %X=S(1) : %Y=S(2) : %Z=S(3)
$PR :: " INITIAL SIGMA2 = " %P ::
$DEL NU S

$C READ DATA

$UNITS %U
$DATA D N X A1 A2
$FORMAT
(5X,3F5.0,2F5.0)
$DINPUT 3

$C DEFINITION OF MODEL

$CALC A11=(A1-1) : A12=(A2-1) : A13=0
    : A21=(A1-1) : A22=(A2-1) : A23=0
$CALC B=0 : W=1
$YVAR D $WEIGHT W
$OWN M1 M2 M3 M4 $SCALE 1

$C FITTING THE MODEL

$VAR 8 %PE $CALC %PE=PAR $DEL PAR
$CALC %LP=X11=X12=X13=X14=X21=X22=X23=X24=0
$FIT X11+X12+X13+X14+X21+X22+X23+X24-%GM $DISP L A
$PR :: " CHI2 (NO HETER.) = " %X2 :
$WHILE %R HET $DISP L A

$C RESTRICTED MODELS

$U SUBO

$STOP
```

IPD-WORKING PAPERS 1983

PRICE PER COPY	
BF 230 US$ 6.00	1983-1. R. Lesthaeghe : "A Century of Demographic and Cultural Change in Western Europe : An Exploration of Underlying Dimensions".
BF 300 US$ 8.00	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".
BF 230 US$ 6.00	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".
BF 230 US$ 6.00	1983-4. S. Becker, A. Chowdhury, S. Huffman : "Determinants of Natural Fertility in Matlab, Bangladesh".
BF 350 US$ 9.00	1983-5. C. Vanderhoeft : "A Unified Approach to Models for Analysis of Zero-One Data with Applications to Intermediate Fertility Variables".
BF 300 US $ 8.00	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".
BF 230 US$ 6.00	1983-7. S. Wijewickrema et Alii : "Marital Status Trends in Belgium, 1961-77 : Application of Multi-State Analysis".
BF 230 US$ 6.00	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).
BF 230 US$ 6.00	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).

IPD-WORKING PAPERS 1984

PRICE PER COPY	
BF 230 US$ 6.00	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".
BF 400 US$10.00	1984-2. R. Lesthaeghe : "Fertility and its Proximate Determinants in sub-Saharan Africa : The Record of the 1960's & 70's".
BF 300 US$ 8.00	1984-3. C. Vanderhoeft : "Life Tables and Relational Models : A Unified Approach, and Applications in Demography".
BF 230 US$ 6.00	1984-4. R. Lesthaeghe : "Demographic Renewal and Problems of Social Organization".
BF 230 US$ 6.00	1984-5. S. Becker : "A Response Bias in the reporting of Month of Birth in Pregnancy History Surveys".
BF 230 US$ 6.00	1984-6. R. Lesthaeghe : "On the Adaptation of sub-Saharan Systems of Reproduction".

IPD-WORKING PAPERS 1985

PRICE PER COPY	
BF 330 US$ 8.50	1985-1. R. Lesthaeghe and F. Eelens : "Social Organization and Reproductive Regimes : Lessons from sub-Saharan Africa and historical Western Europe".
BF 300 US$ 8.00	1985-2. P. Willems and S. Wijewickrema : "The Evolution of Nuptiality in Belgium from 1954 to 1981".
BF 400 US$10.00	1985-3. F. Eelens and L. Donné : "The Proximate Determinants of Fertility in Sub-Saharan Africa: A Factbook based on the Results of the World Fertility Survey".
BF 130 US$ 3.50	1985-4. Ronald C. Schoenmaeckers : "Current Fertility Behaviour in Africa. Results from a Birth Interval Analysis of WFS Data".
BF 400 US$10.00	1985-5. Camille Vanderhoeft : Stratified Proportional Hazards Models. A GLIM oriented approach, with special reference to the problem of competing risks.
BF 250 US$ 6.50	1985-6. Stanley Wijewickrema : "Childlessness in Belgium and Flanders".
BF 250 US$ 6.50	1985-7. R. Lesthaeghe : "Value Orientations, Economic Growth and Demographic Trends - Toward a Confrontation?
BF 250 US$ 6.50	1985-8. Fernando Rajulton and Stan Wijewickrema : "An Enquiry into the Two Basic Assumptions of Multistate Increment-Decrement Life Tables".

IPD-WORKING PAPERS 1986

PRICE PER COPY	
BF 250 US$ 6.50	1986-1. F. Rajulton : "Marriageability and divorceability : a simulation of the unobservables through the conditional gaussian diffusion process.
BF 180 US$ 4.50	1986-2. H.J. Page : "Child-bearing versus Child-rearing : Co-residence of Mothers and Children in Sub-Saharan Africa".
BF 450 US$11.50	1986-3. R. Lesthaeghe, G. Kaufmann, D. Meekers : "The Nuptiality Regimes in Sub-Saharan Africa".
BF 230 US$ 6.00	1986-4. R. Lesthaeghe, G. Kaufmann, D. Meekers : "The Nuptiality Regimes in Sub-Saharan Africa : Data Files NUPFILE 1 and NUPFILE 2".
BF 200 US$ 5.00	1986-5. Hwa Young LEE and Fernando RAJULTON : "A Semi-Markovian Multistate Analysis of Sexual Union Formation in Flanders, Belgium".
BF 450 US$11.50	1986-6. Zeng Yi : "Fertility and Household Dynamics in China".
BF 300 US$ 8.00	1986-7. R. Lesthaeghe & D. Meekers : "Value Changes and the Dimensions of Familism in the EEC".