Berger, Fahrmeir, Klasen:

Dynamic Modelling of Child Mortality in Developing Countries: Application for Zambia

Sonderforschungsbereich 386, Paper 299 (2002)

Online unter: http://epub.ub.uni-muenchen.de/

Projektpartner

Dynamic Modelling of Child Mortality in Developing Countries: Application for Zambia

Ursula Berger

Medical Research Council, Social and Public Health Sciences Unit

University of Glasgow

Ludwig Fahrmeir

Institut für Statistik

Ludwig-Maximilians Universität München

Stephan Klasen

Seminar für Empirische Wirtschaftsforschung

Ludwig-Maximilians Universität München

Abstract

In this paper, we analyse the causes of under five mortality in Zambia, with a particular emphasis on assessing possible time-variations in the effects of covariates, i.e. whether the effects of certain covariates vary with the age of the child. The analysis is based on micro data from the 1992 Demographic and health Survey. Employing a Bayesian dynamic logit model for discrete time survival data and Markov-Chain Monte Carlo methods, we find that there are several variables, including the age of the mother and the breastfeeding duration whose effects exhibit distinct age-dependencies. In the case of breastfeeding, this age dependency is intimately linked with the reasons for stopping breastfeeding. Incorporating such age dependencies greatly improves the explanatory power of the model and yields new insights on the differential role of covariates on child survival.

Key words: Child Mortality, Dynamic Survival Models, Bayesian Model Comparison, Discrete Failure Time.

1 Introduction

Mortality and its converse indicator, longevity or life expectancy, are among the most important measures of well-being and development in poor countries (Sen, 1998). Since child mortality has an overwhelming influence on life expectancy, it is particularly important to analyse the determinants of child mortality in poor countries. The general medical definition distinguishes mortality of a child with respect to the child's age: Death within the first week of life is included with perinatal mortality (which also includes late foetal mortality) and death within the first month is referred to as neonatal mortality. Since peri- and neonatal mortality is heavily influenced by prematurity, fatal genetic conditions of the foetus, and problems associated with delivery, analysis of mortality often separate between the determinants of this type of mortality and mortality after the first month, which is mostly related to socio-economic and health conditions of the household and the child which will be the focus of the analysis here (e.g. Waldron, 1998). It is possible to analyse the determinants of child mortality at various levels of causality (Mosley and Chen, 1984). The biomedical and epidemiological literature typically focuses on the immediate determinants of child mortality, in particular the impact of various diseases and weakened resistance. In contrast, socio-economic analyses of child mortality are usually focused on underlying determinants of child mortality that make children more vulnerable to the attack of various diseases. These determinants usually involve the education of the parents, the income or wealth situation of the household, access to water and sanitation services and access to health services. Sometimes, analysts also examine intermediate causes of child mortality that lie between the underlying socio-economic determinants and the immediate factors. Such variables typically include maternal factors such as birth spacing or maternal age at birth, environmental contamination, nutritional deficiency, and care practices (Mosley and Chen, 1984). The precise separation between underlying and intermediate factors is, however, not always very precise as some variables might be proxies of either (Schultz, 1984; Klasen, 1999).

In this study we will investigate the underlying and proximate determinants of child mortality in Zambia in 1992 excluding neonatal mortality. We will jointly consider socio-economic factors as well as maternal, environmental, and nutrition-related issues. The analysis is based on the 1992 Demographic and Health Surveys

in which for a representative sample of women from Zambia data was collected on the monthly survival times of their children.

Apart from this question of the nature of the causal chain, there are question concerning possible time-variations in the effects of the socio-economic determinants of child mortality. In particular, there are good reasons to believe that some determinants of child mortality should have a larger impact in the early phase of life. Breastfeeding is likely to have a different impact on mortality in different life periods. In particular, while there is strong biomedical evidence that exclusive breastfeeding in the first months of life lowers child mortality (WHO,1995), the importance of breastfeeding is less clear in later months, particularly after the child has reached the age of one year. Such time-variations are typically modelled using piecewise fits based on interaction terms with time (e.g. Guilkey and Riphahn, 1997). But this procedure has many disadvantages. First it is discontinuous and depends on a pre-set partitioning of the time axis where finer partitioning increases the flexibility of the fit but at the same time uses up a great number of degrees of freedom. Moreover, this method can be quite unreliable if there are only a few observations or events. It is therefore preferable to model these dynamic effects smoothly without being too data-intensive.

In this paper, we use discrete-time survival models (see Fahrmeir and Tutz, 2001, ch.9), with months as time units, for a refined analysis of the causes of child mortality. Such an approach can make considerably more efficient use of the information contained in the data. In particular, we can model and estimate in a flexible way non-linear and time-varying patterns of the baseline hazard and of covariate effects, in addition to usual fixed effects of covariates. Inference for such dynamic discrete-time survival models is fully Bayesian and uses MCMC techniques as in Fahrmeir and Knorr-Held (1997) and Fahrmeir and Lang (2001). For model selection and validation Bayesian model criteria are employed, including the Bayes factors, the posterior deviance, and the Deviance Information criterion (Spiegelhalter et al., 2002).

The paper is organised as follows. Section 2 provides a outline of the study and a

descriptive summary of the data. Section 3 briefly outlines dynamic discrete-time survival models and inferential methods used in the analysis. Results are discussed in Section 4. Finally in Section 5 we draft conclusions and point out some directions for future research.

2 The Study

2.1 Zambia

Zambia is a low income developing country in Southern Africa, belonging to the ten poorest countries in the world. While it was one of the richer African countries in the 1960s, largely related to its great mineral wealth (especially copper), the economy stagnated and regressed throughout most of the 1970s and 1980s, largely as a result of falling copper prices, economic mismanagement, and the impact of subsequent austerity packages by the IMF and the World Bank. As a result, poverty in Zambia has increased considerably so that by 1996, 72.6% of the population was living on less than $1 per capita per day (World Bank, 2000). Child mortality rose throughout the 1980s and early 1990s. This deterioration in mortality was reinforced by the onset of the AIDS epidemic in the country which increased mortality in all age groups, particularly since the mid-1990s.

2.2 The Data

The analysis of child mortality is based on data from the 1992 Demographic and Health Survey (DHS) for Zambia. The DHS program is funded by United States Agency for International Development to collect population, health, and nutrition data from developing countries and is implemented by Macro International Inc. in conjunction with a local institutional partner. For the survey a nationally representative sample of women in reproductive ages (15-49) is interviewed using a household questionnaire and a women's questionnaire. The questionnaires consist

of different sections including respondent basic data such as age and educational achievements, type of the district and condition of residence, data on reproduction and birth history with child's individual characteristics, health history of the child, information on maternity and feeding of the child, and more. (For more details on the data see http://www.measuredhs.com/.) In order to undertake survival modelling a data set was constructed from a child individual recode data file, where each record represents a child and consists of survival information and the values of the explanatory covariates.

Our analysis focuses on child mortality until age five excluding peri- and neonatal mortality. Individual data records were available for 5965 children that were born within the last five years before interview and had survived their first month of life. 633 children had died. The Kaplan-Meier estimate for the survival of these children is given in Figure 1. It shows a 5-years mortality rate of nearly 14%, where most of the deaths occur within the first two years of life. (Note that here, as well as in the following KM-plots, the y-axis is cut to the interval $[0.7, 1]$.)

2.3 Coding of the Covariates

From the basic variables of the DHS data a set of covariates is constructed and their values are grouped to form categorical and binary coded factors. Some of them contain missing values. Table 1 gives a short description of these constructed covariates, their coding, and their distribution. Child and birth specific information is given by the covariates gender, birth order and the preceding birth interval, where birth order is reduced to a binary factor, indicating whether the child is first born. The preceding birth interval is calculated as months between the birth of the child and the previous birth the mother had and is naturally only defined for children not first born. A 'short' birth interval is defined to be no longer than two years. Maternal information includes the age of the mother at birth, calculated from the DHS data on the date of birth of the mother and the date of birth of the child and coded in three categories. For the highest educational attainment of the mother we distinguish no education or incomplete/complete primary school versus higher attainment than primary school. Also the size of the household is

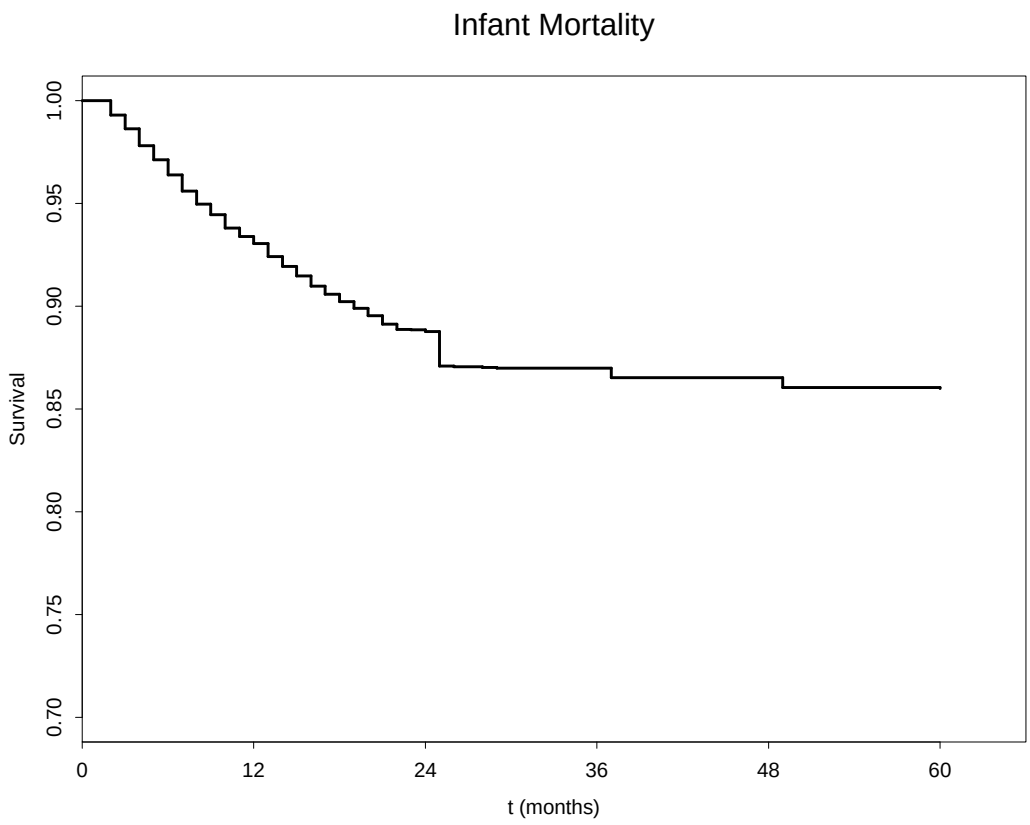

Figure 1: Overall survival of children younger than five in Zambia (excluding peri- and neonatal mortality).

Table 1: Factors analysed in the child mortality study

factor	frequency	coding	interpretation
SEX	3003 (50.3%)	0: female	gender
	2962 (49.7%)	1: male	
BFIRST	4669 (78.3%)	0: no	whether the child
	1296 (21.7%)	1: yes	is first born
BIRTHIN	1035 (22.2%)	0: $\leq$ 24 months	preceding
	3628 (77.8%)	1: $>$ 24 months	birth interval
MAGE	1780 (28.8%)	0: $\leq$ 21	age of mother
	3486 (58.4%)	1: 22-35	at birth
	699 (11.7%)	2: $>$ 35	
MEDUCATION	4822 (80.9%)	0: $\leq$ primary	educational
	1141 (19.1%)	1: $>$ primary	attainment
HHSIZE	1177 (19.7%)	0: 1 - 4	total number
	1522 (25.5%)	1: 5 - 6	of household members
	3266 (54.8%)	2: 7 -	
URBAN	3393 (56.9%)	0: rural	type of district
	2572 (43.1%)	1: urban	of residence
WATER	2835 (47.7%)	0: residence/tap	source of
	3108 (52.3%)	1: else	water
HOUSE	3104 (52.3%)	0: wood/sand	material
	2836 (47.7%)	1: higher quality	of the floor
ELECTRICITY	4776 (80.3%)	0: no	electricity
	1170 (19.7%)	1: yes	
BREASTF		0: no	currently breastfeeding
		1: yes	(time-dependent)

taken into account, which is recorded for the time of the interview and is split into three categories. Information on the type and condition of the residence comprises the type of district by distinguishing rural and urban, whether the household has electricity and the main material of the floor. The major source of drinking water is divided with respect to its quality, where water access in the residence or from public tab is assumed have controlled quality, while water from a public well, springs, rivers or streams, ponds or lakes or rainwater is not controlled. Water from tanker trucks is also added to the later category since, even if it is controlled, it is usually rather costly and scarce.

Some covariates naturally have special characteristics which must be taken into account when fitting the data. This applies for the duration of breastfeeding, which is given in month. It is an internal covariate that is observed only as long as a child is alive. It is also time-dependent, in contrast to most of the other variables which are fixed in time. In consequence, the covariate 'duration of breastfeeding' carries survival information of the corresponding child as it can never exceed its survival time. Therefore this covariate cannot be included into the model as a fixed covariate, but its time dependent and survival-dependent nature must be recognised. It is described in a binary covariate process (BREASTF), which has the value one during the months the child was breastfed and zero when the child was fed differently.[1]

The Kaplan-Meier estimate of Figure 2 gives an impression of the distribution of duration of breastfeeding within the collective. Here, an observation is censored when duration of breastfeeding ended due to death or censoring. It shows that the majority of women would breastfeed their children about one to two years.

[1]To give an example, think of a child that survived only 5 months but was breastfed during all this time. Duration of breastfeeding is then equal to 5. The corresponding covariate process is equal 1 for all five month and not defined later. In contrary, if a mother stopped breastfeeding after 5 months for some other reasons, e.g. illness, duration of breastfeeding is equal to 5 as well, while the covariate process is equal 1 for five month and equal zero for the following month until end of observation.

Figure 2: Duration of breastfeeding (corrected for break-off by death or censoring of the child).

3 Dynamic Discrete-Time Survival Models

Many studies of child mortality simply use a static logit framework, thus modelling the probability of being alive at a certain age (e.g. Prevost, 1996; Mohamed et al. 1998). While this is partly driven by the lack of time-varying covariates, it is nevertheless problematic as it fails to model the precise timing of death and thus wastes information available in most cross-sectional micro data sets. Moreover, it does not allow to include internal covariates such as breastfeeding. Many data sets, including the DHS used here, collect a large number of retrospective information for each child so that it is possible to model a dynamic survival model with time-varying covariates.

For the data set at hand, survival time is not recorded continuously but is only known to lie within a month or, generally, a time interval. Data of this kind are known as interval censored. Since many ties occur, these data cause problems when continuous-time models are used. Instead of continuous time one observes the discrete time T with values $t = 1, \ldots, k$, where $T = t$ denotes death in month or interval t, and k is the last observation interval. In addition to survival time T, a sequence of possibly time-varying covariate vectors $x_t = (x_{t1}, \ldots, x_{tk})$ is observed. Let $x(t) = (x_1, \ldots, x_t)$ denote the history of covariates up to month t. The discrete hazard function is defined as the conditional probability

$$\lambda(t|x(t)) = P(T = t|T \geq t, x(t))$$

for death in month t, given survival up to this month and the history of covariates. Since our data set comprises children that have died within the preceding five years, as well as children still alive, we recorded for each child survival information by (T_i, δ_i), $i = 1, \ldots, 5965$, where $T_i \in \{1, \ldots, 60\}$ denotes its observation time (measured in months), and δ_i is the indicator of survival, which takes the value 1 if child i had died, and zero if it is still alive (i.e. censored). For formal reasons we assume that censoring is non-informative and occurs at the end of the interval, which implies that the set of children under risk at age t also includes children who are censored at t. Thus, for $\delta_i = 1$, T_i is the age of the child at death, and for $\delta_i = 0$, T_i is the current age of the child at interview.

Discrete-time survival models can be cast into the framework of binary regression models by defining binary event indicators $y_i(t)$, $t = 1, \ldots, T_i$, with

$$y_i(t) = \begin{cases} 0, & t < T_i \\ \delta_i, & t = T_i. \end{cases}$$

The hazard function for child i can then be written as a binary response model

$$P(y_i(t) = 1 | x_i(t)) = h(\eta_i(t)) \tag{1}$$

where $x_i(t)$ are the covariate processes for child i, h is an appropriate response or link function, and the predictor $\eta_i(t)$ is a function of the covariates.
Note that breastfeeding is an internal covariate, i.e. its path may carry information on survival because it can be observed only as long as a child is alive. Fahrmeir und Wagenpfeil (1996) and Fahrmeir and Tutz (2001, ch.9) give formal conditions, which guarantee that internal covariates are non-informative for the discrete-time survival process. Informally, these conditions can be described as follows: For each t, the joint distribution of the risk set R_t of children alive at month t and the covariates $x_i(t), i \in R_t$, given past risk sets, covariates and observed deaths, does not depend on the parameters contained in the predictor $\eta_i(t-1)$. This seems to be a plausible assumption in our application.

Common choices for discrete survival models of the form (1) are the grouped Cox model and logit or probit models. In this paper, we use a logit form. The conventional model is then

$$P(y_i(t) = 1 | x_i(t)) = \frac{\exp(\eta_i(t))}{1 + \exp(\eta_i(t))}$$

with partially linear predictor

$$\eta_i(t) = \beta_0(t) + x_{it}'\beta\,. \tag{2}$$

The baseline hazard effect $\beta_0(t)$, $t = 1, 2, \ldots$ is an unknown, usually non-linear function of t to be estimated from the data. Treating the effect $\beta_0(t), t = 1, 2, \ldots$ as separate parameters usually gives either very unstable estimates or may even lead to divergence of the estimation procedure. In a purely parametric framework the baseline hazard is therefore often modelled by a few dummy variables dividing

the time-axis t into a number of segments or by some low order polynomial. In general it is difficult to correctly specify such parametric functional forms for the baseline effects in advance. Non-parametric modelling based on some qualitative smoothness restriction offer a more flexible solution to explore unknown dynamic patterns in β_0. Then (2) can be regarded as the basic form of a semi-parametric predictor where the effects β of covariates are fixed and time-constant.
In many applications, the restriction to constant covariate effects is not realistic. Rather, the effect of some covariates may vary over time. Therefore, the predictor is generalised to

$$\eta_i(t) = \beta_0(t) + z'_{it}\beta(t) + w'_{it}\gamma$$

where the effect $\beta(t)$ of the covariates in z_{it} are time-varying, while w_{it} comprises covariates with an effect γ that remains constant over time. Also the time-dependent effect functions $\beta(t)$ will be modelled non-parametrically.

To estimate smooth effect functions and model parameters, we use a fully Bayesian approach, as developed in Fahrmeir and Knorr-Held (1997) and Fahrmeir and Lang (2001). For fixed effect parameters γ we assume diffuse priors. For the dynamic effect functions, more exactly for the sequence $\beta(t)$, $t = 1, \ldots, k$ of function values, we assign smoothness priors in form of a second order random walk model

$$\beta(t) = 2\beta(t-1) - \beta(t-2) + u(t),$$

with i.i.d. Gaussian errors $u(t) \sim N(0, \tau^2)$. The random walk model locally penalises deviations from straight lines or, equivalently, deviations of second differences from zero. The variance τ^2 controls the amount of smoothness: the penalty increases or decreases as the variance becomes smaller or larger, respectively. Thus, the variance acts as a smoothness parameters. Data driven estimation, jointly with effect functions and fixed effects is possible by assigning highly dispersed inverse Gamma priors.
Fully Bayesian inference is based on the posterior distribution of the model parameters, which is not of known form, however. Therefore, MCMC sampling from full conditionals for dynamic effect functions, fixed effects and smoothing parameters is used for posterior analysis. Details are given for example in Fahrmeir and Knorr-Held (1997) and Fahrmeir and Lang (2001).

An essential task of the model building process is the comparison of a set of plausible models, for example to rate the impact of covariates and to assess if their effects are time-varying or not. The comparison of models intends to select the model that takes all relevant structure into account while remaining parsimonious. A model criterion should therefore trade off between goodness of fit and model complexity.

Consider that comparison is about a set of non-nested hierarchical models $\mathcal{M}_1(\theta), \ldots, \mathcal{M}_m(\theta)$, such as described above, where θ_j comprises the entire set of parameters of the model $\mathcal{M}_j$. A classical approach to Bayesian model selection relies on Bayes factors (Jeffreys, 1961). For comparing two models, $\mathcal{M}_0$ and $\mathcal{M}_1$ say, it is defined by

$$\mathrm{BF}(\mathcal{M}) = \frac{\int L(Y|\theta_0, \mathcal{M}_0)\pi(\theta_0|\mathcal{M}_0)}{\int L(Y|\theta_1, \mathcal{M}_1)\pi(\theta_1|\mathcal{M}_1)} \tag{3}$$

where $\pi(\theta_j|\mathcal{M}_j)$ denotes the prior specifications made under model $\mathcal{M}_j$ and $L(Y|\theta_j, \mathcal{M}_j)$ is the likelihood, i.e. the conditional probability of the data given the fully specified model. Thus the Bayes factor is the ratio of the marginal likelihoods of models M_0 and M_1 and expresses the weight of evidence for model $\mathcal{M}_0$ over model $\mathcal{M}_1$. A serious challenge is its computation for complex, hierarchical models as the integral, i.e. the marginal likelihood, are again intractable, see Han and Carlin (2001) for a comparative review. We use an approximation of the Bayes factor via importance sampling from the posterior distribution which is equivalent to the posterior harmonic mean of the likelihood (Kass and Raftery, 1995) and is computational feasible also for large data sets as in our application.

Additionally, we routinely employ the Bayesian model deviance, which is defined as minus twice the log-likelihood (Dempster, 1997), i.e.

$$D(\mathcal{M}) = -2 \log L(Y|\theta, \mathcal{M}) \,.$$

For a given set of data the deviance $D(\mathcal{M})$ is a function of the model parameters θ. Characteristics of its posterior distribution can therefore directly be derived from the posterior distribution of θ, which is provided by the MCMC output. A central characteristic is, of course, the posterior mean of the deviance $\overline{D(\mathcal{M})}$, which is frequently used to summarise the goodness of fit of a model.

Based on the Bayesian model deviance Spiegelhalter et al. (2001) suggested the

Deviance Information Criterion DIC

$$\mathrm{DIC}(\mathcal{M}) \;=\; \overline{D(\mathcal{M})} + df_{\mathcal{M}}\,. \tag{4}$$

Here the posterior mean of the deviance $\overline{D(\mathcal{M})}$ is penalised by the effective number of model parameters $df_{\mathcal{M}}$ measuring the complexity of the model, where $df_{\mathcal{M}} = \overline{D(\mathcal{M})} - D(\bar{\mathcal{M}})$, and $D(\bar{\mathcal{M}}) := D(\bar{\theta}, \mathcal{M})$ is the deviance evaluated at the posterior mean of the model parameters. Like the posterior deviance also the DIC is easy to compute in a MCMC analysis.

Note that $\overline{D(\mathcal{M})}$ is the logarithm of the posterior geometric mean of the likelihood whereas the integrals of the Bayes factor are approximated by the posterior harmonic means of the likelihood, and so, despite their different origin, the criteria we determine for model comparison are closely related. A systematic comparison of the criteria together with a simulation study is given in Berger (2002).

4 Results of the survival analysis

To get an insight to the data structure, check for possible interactions between the covariates and explore time-variation in the effect structures we studied different Kaplan-Meier curves and univariate models in a preliminary analysis. This allows us to reduce the complexity of the multivariate model building process.
At first sight it seems that gender has no impact on child's survival. If additionally interaction with the order of birth is considered the Kaplan-Meier estimates reveal that generally boys have a higher mortality risk than girls, but for first-borns, where girls have a higher mortality risk (Figure 3). Overall first borns have higher mortality than their younger siblings. No relevant interactions between other covariates have been found.

To explore which of the covariate effects might vary with the age of the child,

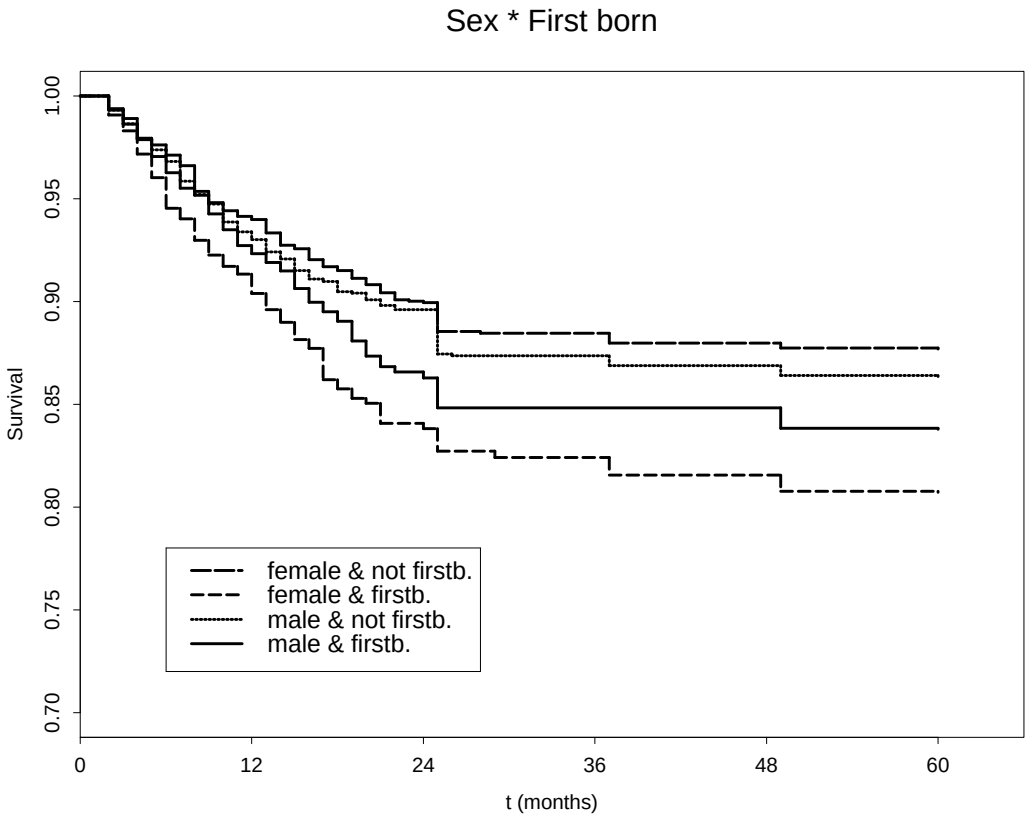

Figure 3: Child mortality dependent on gender and birth order.

we fitted univariate dynamic logit models with time-varying covariate effects and compare them with the corresponding constant estimates. Model comparison is based on the criteria described above, i.e. the Bayes factor, the posterior mean of the deviance and the DIC, where smaller values indicate a better fit. We found the most distinct time-variation for breastfeeding. Also the age of the mother at birth shows time-variation which should be considered. The dynamic structures of the other factors are rather weak and their improvement of the fit does not justify the increased complexity of the model.
Using this information we fitted a multivariate model (M^*) jointly evaluating all covariates, the interaction of gender and birth order (SEX*BFIRST) and time-varying effects for BREASTF and MAGE.

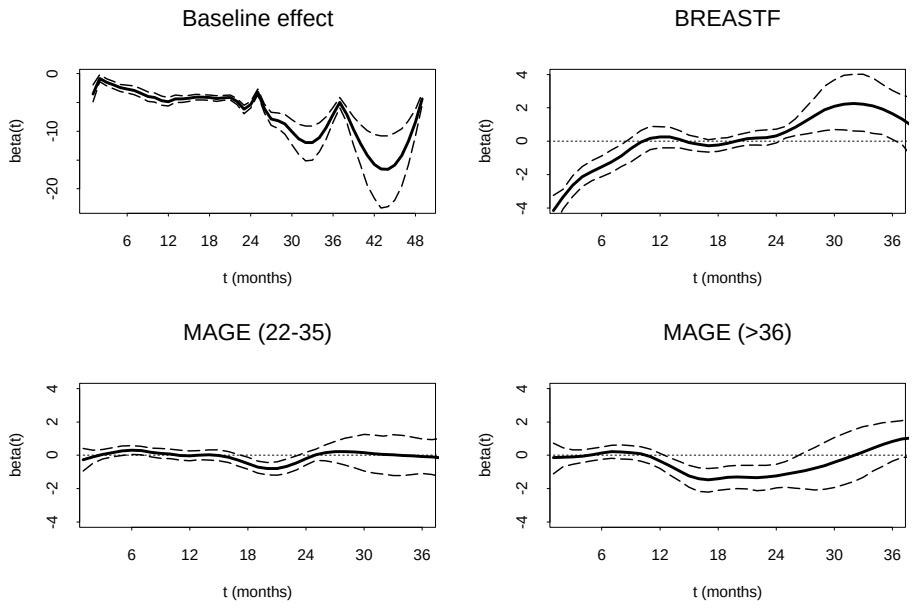

Figure 4: Dynamic effect for the multivariate model.

The results for the constant (fixed) effects are summarised in Table 2 where the

posterior mean of the parameters expresses the impact of the covariate. Additionally the posterior standard deviation is given, together with the 10% and 90% credibility regions enclosing 80% of the posterior sample of the effects. It confirms the findings from the Kaplan-Meier analysis showing a strong interaction between birth order and gender. A larger preceding birth interval has a clear negative impact on the survival of children while a higher education of the mother improve the life expectation of a child. Also children living in a larger household benefit. The effects of the factors describing the condition of residence and the type of district of residence seem to carry partially same information and thus have a smaller impact. The 10%-90% credibility bands of the effects for the variables URBAN, HOUSE and WATER even include the zero. In general children living in better equipped houses such as houses with electricity, have a lower mortality rate.

Table 2: Results for the constant effects from the selected multivariate logit model

factor		post. mean	Post. StdDev	10% CI	90% CI
BFIRST		-0.028	0.149	-0.216	0.161
BIRTHIN		-0.290	0.099	-0.419	-0.165
SEX		0.104	0.092	-0.015	0.219
SEX*BFIRST		-0.313	0.180	-0.539	-0.082
DMEDUCATION		-0.167	0.122	-0.325	-0.014
HHSIZE	D1	-0.662	0.104	-0.799	-0.535
HHSIZE	D2	-0.963	0.096	-1.086	-0.842
URBAN		0.151	0.126	-0.007	0.319
WATER		0.093	0.136	-0.077	0.267
HOUSE		0.001	0.125	-0.158	0.158
ELECTRICITY		-0.289	0.139	-0.472	-0.113

The development of the effects over the age of the child are displayed in Figure 4 together with the 80% credibility region. The estimated baseline effect declined over time with a steep descent within the first year. The peaks jutting at month 24, 36 and 48 are caused by heaping of children reported to have died at this age, a typical misrepresentation in these kind of surveys. Controlling for this heaping through a flexible baseline is a further advantage of our approach as it is likely to reduce biases of the coefficients on the fixed and time-varying effects. For the

covariates we only display the effect for the first 36 months since later the credibility intervals are too wide to provide any reliable information on the dynamic structure. The figure shows the distinct time-variation for breastfeeding causing a high risk of mortality when not done in the first year but having no impact when stopped within the second year. This goes along with the general breastfeeding conduct of mothers in Zambia, which was illustrated in Figure 2. As described before, break-off by death or censoring of the child does not influence this result. The effect of age of the mother indicates that the benefits of an older mother are rather relevant in the second year, probably after breastfeeding is stopped. For comparison we also fit a model where all covariate effects are fixed constant ($M0$) and only the baseline effect is allowed to smoothly vary over time, as well as a simple parametric model ($M1$) which captures possible time-variations of the baseline effect by a piecewise fit, that is by including five interaction terms with time. Figure 5 shows the values for the Bayesian model criteria for these models, i.e. minus twice the logarithm of the Bayes Factor respectively the marginal likelihood, the posterior mean of the deviance and the DIC. It becomes obvious that the data contains dynamic patterns which cannot be adequately captured by a piecewise baseline effect.

To validate the dynamic effects of the covariates we further fitted models, where in turn the effects of BREASTFT and MAGE are fixed constant (models M2 and M3). Their fits are also shown in Figure 5. Clearly all criteria support the dynamic effects for breastfeeding. Also the dynamic effect for the age of the mother improves the goodness of fit, although the difference in $DIC(M^*)$ and $DIC(M3)$ is very small.

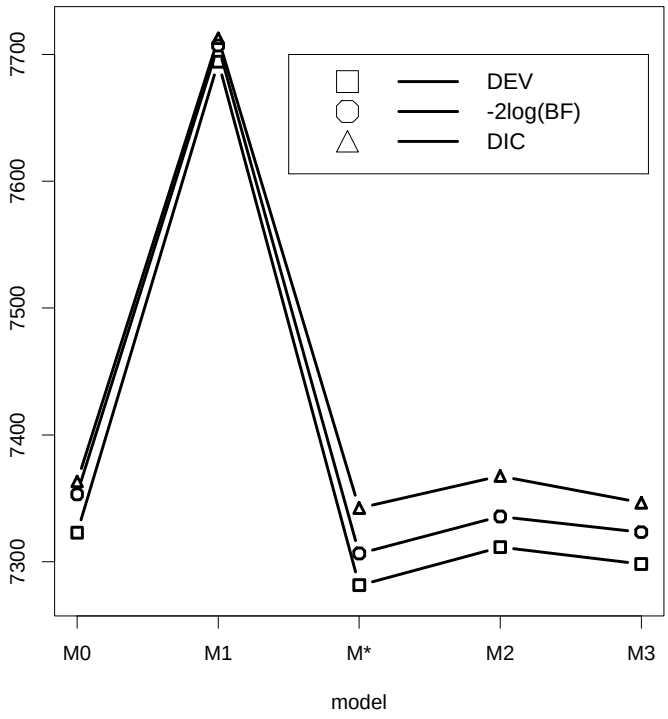

Figure 5: Comparison of different multivariate models. $\mathcal{M}0 : \beta_0(t) + \beta' w + \beta\, BREASTF + \beta\, MAGE$. $\mathcal{M}1 : \beta_0(Dummy) + \beta' w + \beta\, BREASTF + \beta\, MAGE$. $\mathcal{M}^* : \beta' w + \beta(t) BREASTF + \beta(t) MAGE$. $\mathcal{M}2 : \beta' w + \beta\, BREASTF + \beta(t) MAGE$. $\mathcal{M}3 : \beta' w + \beta(t) BREASTF + \beta\, MAGE$.

4.1 Reasons for stopping breastfeeding

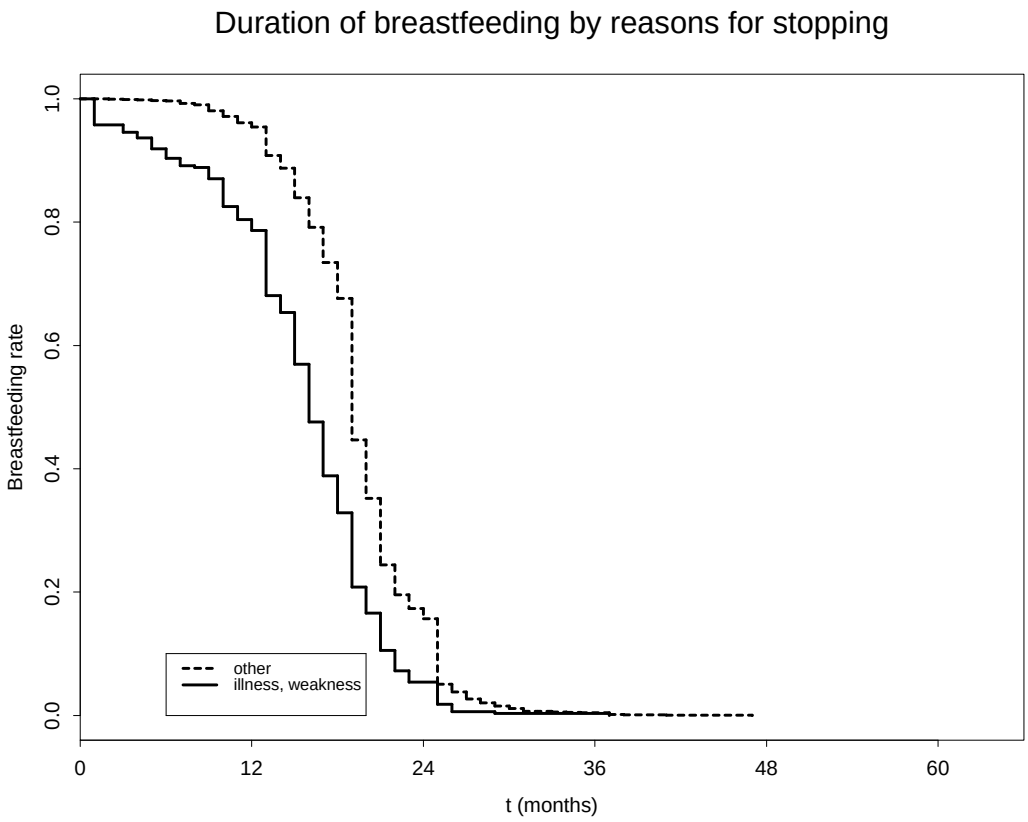

Figure 6: Duration of breastfeeding dependent on reasons for stopping breastfeeding.

The strongest effect on survival shows breastfeeding, which induces a very high risk of mortality when stopped already within the first year of life. One reason might be the low quality of substituting nutrition. However, we would then expect that the quality of water shows the converse time-structure, which is not the case here. Another explanation becomes apparent when studying the reasons for which breastfeeding was stopped within the first year. Figure 6 illustrates the dependence of duration of breastfeeding on the reasons for stopping, distinguishing illness and weakness of the mother or the child from innocuous reasons such as weaning age, pregnancy, working ect. . It shows, that if not due to illness or weakness, most women in Zambia tend to breastfeed longer than one year. We would expect that life expectancy of the child is reduced when breastfeeding was stopped due to

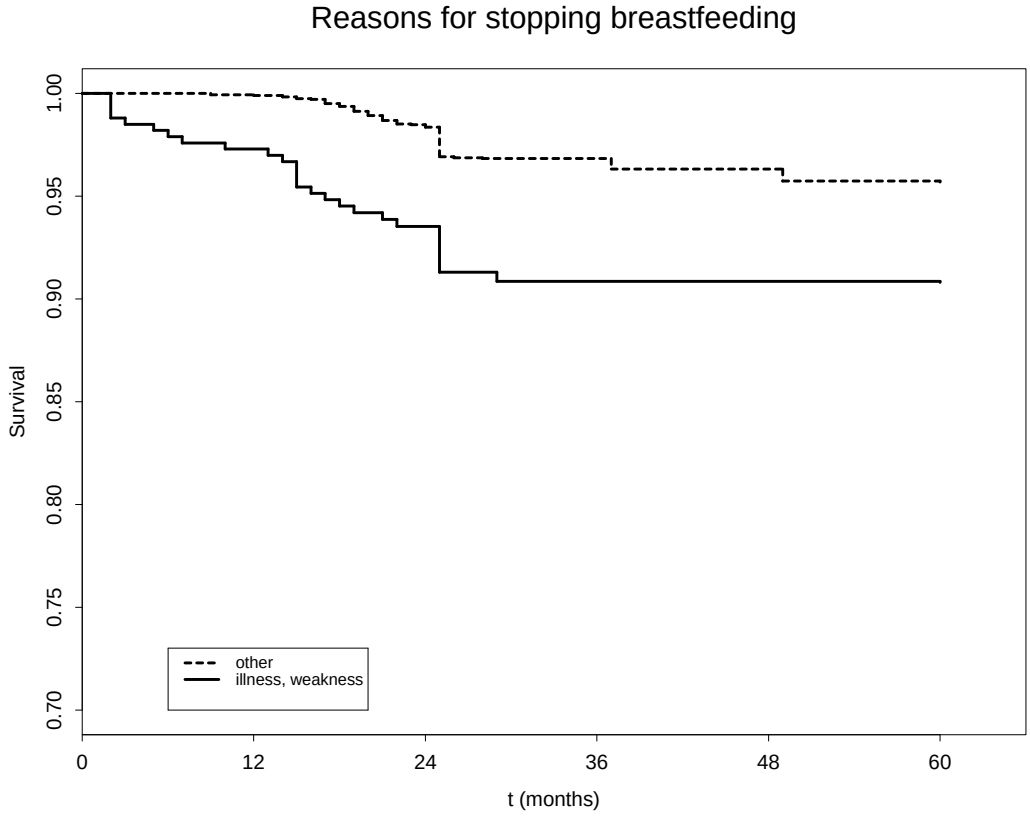

Figure 7: Child mortality dependent on reasons for stopping breastfeeding.

illness or weakness compared to other stopping reasons (see Figure 7).

As a consequence we included the stopping reasons in the selected model $\mathcal{M}^*$ from above. Similar to BREASTFT we constructed a time-dependent covariate STOPR that is zero as long as the child is breastfed and changes to one in the month before breastfeeding was stopped due to illness or weakness. For all other children it is zero throughout observation time. Hence STOPR is not identical to an interaction between a fixed covariate representing stopping reason and the time-dependent covariate BREASTFT, since in the last moth of breastfeeding both covariates have value one when illness or weakness occurred (see Figure 8).

Figure 9 shows the resulting dynamic effects of the covariates of this extended model where the effect of the stopping reasons is allowed to vary over time (model $M4$). The effect of breastfeeding becomes much flatter revealing that the high risk of stopping breastfeeding within the first months can mostly be explained by the

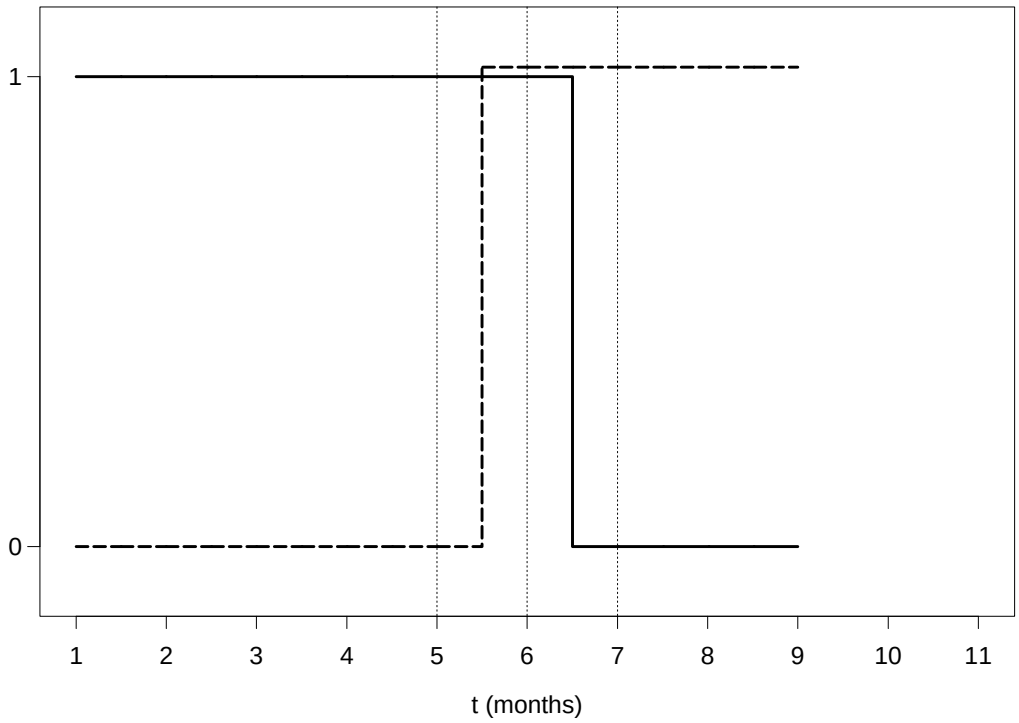

Figure 8: Construction of the time-dependent covariates BREASTF (——) and STOPR (- - - -), when assuming that at month 6 breastfeeding was stopped due to illness or weakness and death occurred at month 9.

stopping reasons. This is of course not achieved when STOPR is included with a fixed effect (model $M5$, effects not shown). Figure 10 shows the model criteria for those models and model M^*. All criteria are in agreement that the reasons for stopping breastfeeding carry important information and should be included with a time-varying effect. However, STOPR obviously cannot explain all time-variation of the effect of breastfeeding, since the criteria shown in Figure 10 still reject a fixed effect for BREASTFT (model $M6$). The constant effects of the remaining parameters are unaffected by the inclusion of STOPR, and therefore not shown.

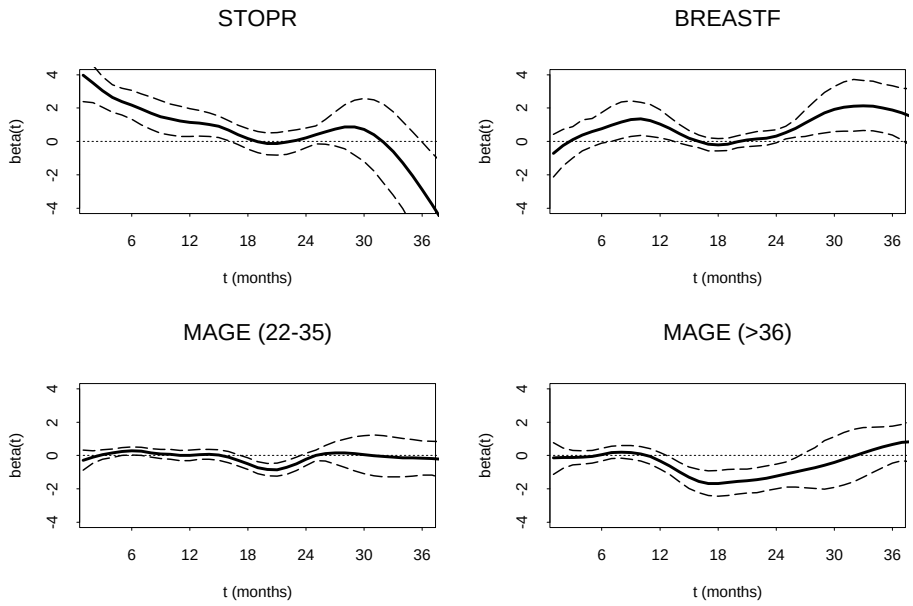

Figure 9: Dynamic effect for the multivariate model including STOPR.

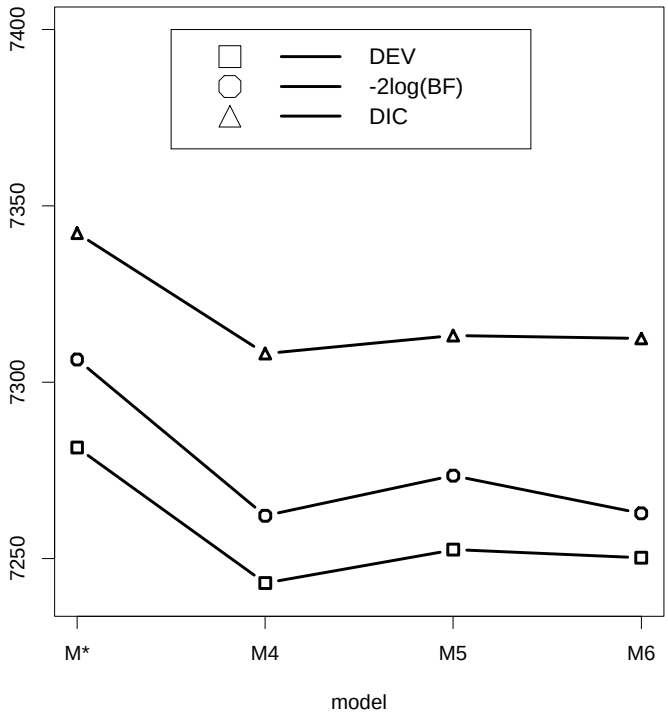

Figure 10: Model comparison for reasons for stopping breastfeeding. $\mathcal{M}^* : \beta' w + \beta(t) BREASTF + \beta(t) MAGE$. $\mathcal{M}4 : \mathcal{M}^* + \beta(t)\, STOPR$. $\mathcal{M}5 : \mathcal{M}^* + \beta\, STOPR$. $\mathcal{M}6 : \beta' w + \beta(t) MAGE + \beta(t) STOPR + \beta\, BREASTF$.

5 Discussion

The analysis examines the fixed and time-varying impacts of socio-economic, maternal, environmental and nutrition-related factors on child mortality in Zambia using a dynamic logit model. Our modelling approach yields a much better fit than traditional modelling approaches with fixed effects and time-variation being modelled using dummy variables. The comparison of this quite large number of univariate and multivariate models of such a huge data set was only possible as all fit criteria can be derived as a 'by-product' from the MCMC sampling of the posterior distribution of the parameters and hence do not demand much additional computation. Thus it is computationally possible to evaluate a large class of non-nested models using very large data sets.

Apart from illustrating the opportunities of such a modelling approach, the application also generates some new insights which will be discussed presently. The substantive results support some known findings, for example about the higher mortality of first-borns, about the importance of the socioeconomic conditions of the household (proxied by urbanization, housing quality, and electricity access), the importance of mother's age and education, and the harmful effects of short birth-intervals (e.g. Mosley and Chen, 1984; Guilkey and Riphahn, 1998, Schultz, 1984).

We report two unusual findings, one related to household size and the other related to gender differences in mortality. In larger households, mortality is significantly lower. Given that we consider household size (rather than family size) and given that households in Zambia often include relatives beyond the nuclear family or even non-relatives, this effect may be a reflection of well-endowed households attracting additional members. Also, there may be the intertemporal selection effect that small households may be small due to past mortality.
Second, while boys suffer overall from higher mortality, girls do badly among the first-borns. The first finding is consistent with boys higher susceptibility to mortality, particularly within the first year (see e.g. Waldron, 1998) combined with little evidence of gender bias in any direction in Sub Saharan African countries (Klasen,

1996). The particularly poor state of first-born girls is a new finding and merits further investigation. It may not show up in other studies due to the fact that boys have much higher neonatal mortality which would lead to higher mortality of all birth orders (e.g. Arnold, 1997). Since we exclude neonatal mortality, the relative disadvantage of first-born females only becomes apparent. The causes for it remain to be explored further in future work.

Regarding the time-varying effects of covariates, breastfeeding and maternal age play an important role. The benefit of greater maternal age at birth are particularly apparent in the second year of life. This is consistent with expectations as care and nutrition decisions of the mother play an increasing role in determining mortality when breast-feeding is ceased in the second year. Here the experience and maturity that comes with age appears to have a significant impact.

The most important substantive finding of the paper is the dynamic effect of breastfeeding. When just looking at breastfeeding in our dynamic setting, we reproduce the known result that it reduces mortality, particularly in the first nine months of life (e.g. WHO, 1995). We also find that breastfeeding beyond one year of age has little impact on mortality. Very long breastfeeding (beyond 30 months) appears to be associated with even slightly higher mortality, which may be due to the fact that only very poor mothers with little alternative nutrition on offer would breastfeed this long (e.g. Klasen, 1999).

More importantly, we show that care must be exercised when interpreting such a finding. When we consider the reasons for stopping to breastfeed, we find that the effect of breastfeeding is intimately related to the reasons for stopping it. Many mothers stop breastfeeding due to illness and weakness of them or their child and they are the ones who then suffer from elevated childhood mortality rates.

Three important message emanate from this. First, an important policy implication of this finding is that health (and possibly nutrition) interventions should particularly be targeted towards women who are suffering from illness and weakness to allow them to continue breastfeeding. Second, the large effects of the benefits in breastfeeding (e.g. Guilkey and Riphahn, 1998) are hard to interpret without considering the reasons for stopping breastfeeding. Third, once we have controlled for the reasons for stopping to breastfeed, it is not possible to make a

firm judgement about the benefits of breastfeeding in the first year as virtually all women continue to breastfeed beyond that time, unless forced by illness to stop. We thus show that our modelling approach not only generates a superior fit of the data but also generates a number of important and policy-relevant findings on the determinants of child mortality in a poor developing country.

References

[1] Arnold, F. (1997). Gender Preferences for Children. DHS Comparative Studies No. 23. Washington DC: DHS.

[2] Berger, U. (2002). Exploring Time-Variation in Survival Models. *Doctoral thesis*, Ludwig-Maximilians Universität, München.

[3] Demographic and Health Surveys (DHS). *http://www.measuredhs.com/*.

[4] Dempster, A. (1997). Comment on 'The calibration of P-values, posterior Bayes factors and the AIC from the posterior distribution of the likelihood'. *Statistics and Computing*, 7, pp. 265-269.

[5] Fahrmeir, L., Knorr-Held, L. (1997). Dynamic discrete time duration models: Estimation via Markov Chain Monte Carlo. *Sociological Methodology*, pp. 417-452.

[6] Fahrmeir, L., Lang, S. (2001). Bayesian inference for generalized additive mixed models based on Markov random field priors. *Applied Statistics*, 50, 2, pp. 201-220.

[7] Fahrmeir, L., Tutz, G. (2001). *Multivariate statistical modelling based on generalized linear models* (2nd edn.). Springer-Verlag, New York.

[8] Fahrmeir, L., Wagenpfeil, S. (1996). Smoothing Hazard Functions and Time-Varying Effects in Discrete Duration and Competing Risks Models. *Journal of the American Statistical Association*, 91, pp. 1584-1594

[9] Guilkey, D., Riphahn, R.,(1997). The determinants of child mortality in the Philippines: Estimation of a structural model. *Journal of Development Economics*, 56, 2, pp. 281-305.

[10] Han, C., Carlin, B.P. (2000). MCMC methods for computing Bayes factors: A comparative review. Research Report 2000-001, Division of Biostatistics, University of Minnesota.

[11] Jeffreys, H. (1961). *Theory of probability* (3rd edn.). Oxford University Press, London. (reprint of 1939)

[12] Kass, R., Raftery, A. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, pp. 773-795.

[13] Klasen, S. (1999). Malnourished and Surviving in South Asia, better Nourished and Dying Young in Africa? What can explain this puzzle. Discussion paper 214, SFB 386, Ludwig-Maximilians Universität, München.

[14] Klasen, S. (1996). Nutrition, Health, and Mortality in Sub-Saharan Africa: Is there a Gender Bias? *Journal of Development Studies*, 32, pp. 913-932.

[15] Mohamed, W.N. , Diamond, I.D., Smith, P. (1998). The determinants of infant mortality in Malaysia: a graphical chain modelling approach. *Journal of the Royal Stat. Society*, A, 161, 3, pp. 349-366.

[16] Mosley, W., Chen, L. (1984). An Analytical Framework for the Study of Child Survival in Developing Countries. In: *Child Survival: Research and Policy*(ed.: Mosley, W.H.). Population Council, New York.

[17] Prevost, A.T. (1996). Gibbs sampling for multilevel logistic models of Tunisian infant mortality. *Multilevel Modelling Newsletter*, 8(1), p. 9.

[18] Schultz, T.P. (1984). Studying the Impact of Household Economic and Community Variables on Child Mortality. In: *Child Survival: Research and Policy*(ed.: Mosley, W.H.). Population Council, New York.

[19] Sen, A. (1998) Mortality as an Indicator of Economic Success and Failure. *Economic Journal*, 108, pp. 1-25.

[20] Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, B, (to appear).

[21] Waldron, I. (1998). Sex Differences in Infant and Early Childhood Mortality. In: *Too Young to Die: Genes or Gender*(ed.: UN). United Nations, New York.

[22] WHO (1995). Physical Status: The use and interpretation of anthropometry. *WHO Technical Report Series*, 854, Geneva.

[23] World Bank (2000). *World Development Report 2000/2001: Attacking Poverty.* Oxford University Press, New York. (http://www.worldbank.org/poverty/wdrpoverty/report/index.htm)