9

LOG-MULTIPLICATIVE MODELS FOR DISCRETE-TIME, DISCRETE-COVARIATE EVENT-HISTORY DATA

*Yu Xie**

In this paper I develop a new class of discrete-time, discrete-covariate models for modeling nonproportionality in event-history data within the log-multiplicative framework. The models specify nonproportionality in hazards to be a log-multiplicative product of two components: a nonproportionality pattern *over time and a nonproportionality* level *per group. Illustrated with data from the U.S. National Longitudinal Mortality Study (Rogot et al. 1988) and from the 1980 June Current Population Survey (Wu and Tuma 1990), the log-multiplicative models are shown to be natural generalizations of proportional hazards models and should be applicable to a wide range of research areas.*

An earlier version of this paper was presented at the annual meeting of the American Sociological Association in Pittsburgh, Pennsylvania (August 1992). The author is grateful to Mark Becker, Chris Bettinger, Bruce Christenson, Clifford Clogg, Robert Mare, Peter Marsden, Adrian Raftery, Lawrence Wu, Kazuo Yamaguchi, and anonymous reviewers for advice and comments on earlier versions of this paper; and to Lawrence Wu for parts of the data analyzed. This research was supported by a research grant and a Young Investigator Award from the National Science Foundation and a Spencer Fellowship from the National Academy of Education. The opinions expressed herein are those of the author. GLIM macros used to estimate the models presented in this paper are available from the author upon request.

*University of Michigan

1. EVENT-HISTORY DATA IN TABULAR FORM

When time and covariates are measured discretely, event-history data can generally be summarized in a tabular form. Tables 1 and 2 are two such examples. The first example presents mortality differentials by educational attainment among U.S. males, drawn from the U.S. National Longitudinal Mortality Study (Rogot et al. 1988, Table 6). The upper panel of Table 1 gives the exposure (O), the number of persons at risk of dying by age and years of schooling; and the lower panel gives the number of deaths (d) for the same classification. Table 2 illustrates an example from Wu and Tuma's (1990) local hazard analysis of age patterns of first marriages by educational attainment and race/ethnicity. As in Table 1, exposures (O) and events (d) are given in two separate panels. The data sets are explained in more detail in Appendix A.

TABLE 1
Exposure and Deaths by Age and Educational Attainment: U.S. Males

	Years of Schooling				
Age	0–8	9–11	12	13–15	16+
Exposure (O)					
15–24	9096.0	27871.5	20498.5	10308.0	2679.0
25–34	3111.0	6231.0	24775.5	15800.0	16522.0
35–44	4306.5	5875.0	17523.0	7646.5	11627.5
45–54	7174.5	6627.5	14802.5	5453.0	8241.5
55–64	9558.5	6461.0	13103.0	4581.5	5563.0
65–74	8445.5	3645.0	5207.0	1882.5	2365.0
75–84	4788.5	1236.5	1390.0	638.0	803.0
85+	1113.0	149.5	210.5	104.5	144.0
Deaths (d)					
15–24	32	71	41	18	0
25–34	8	36	67	32	16
35–44	43	54	80	35	31
45–54	121	99	185	60	61
55–64	425	274	390	125	120
65–74	629	234	310	125	120
75–84	655	159	144	72	92
85+	260	43	55	19	30

Note: Data are extracted from the U.S. National Longitudinal Mortality Study (Rogot et al. 1988, Table 6). For each age × education cell, a death is assumed to contribute 0.5 person-period of exposure.

Taking the ratio between a cell in the lower panel (d) and the corresponding cell in the upper panel (O) in Tables 1 and 2 gives the observed hazard rate at time t given covariate vector $\underset{\sim}{z}$, denoted here by $h(t|\underset{\sim}{z})$. For example, the hazard rate of first marriage for white women at age 18 with 12 years of schooling is observed to be 0.20 (318/1622.35). The hazard rate is interpretable as the "instantaneous rate" in general and as the "force of mortality" in the mortality literature (Namboodiri and Suchindran 1987, p. 32). As in the life-table approach in demography, observed hazard rates can easily be calculated and reported for discrete-time, discrete-covariate event-history data. There is, however, an important distinction between the usual life-table approach and event-history models for discrete-time, discrete-covariate event-history data: Whereas the demographic literature commonly treats observed data as populations and observed hazard rates as exact rates, the tabular presentation of event-history data in the form of Tables 1 and 2 implicitly treats observed data as samples and observed hazard rates as being inexact. Thus it is important to preserve original information on exposures and events and to not reduce them into observed hazard rates, so that statistical techniques for smoothing sampled data and assessing sampling errors can be applied later.

One limitation of the tabular presentation of event-history data and the statistical treatment of such tabular data is the requirement that time and covariates be discrete. When time and covariates are measured in continuous units, they need to be discretized into intervals, as in life-tables. However, this limitation is offset by the advantage that observed tabular data can be easily reported in their entirety and interpreted substantively before any statistical models are applied. In practice, time and covariates are often measured in discrete terms. When they are not, social scientists often categorize them. As argued by Manton et al. (1992), "discrete coding of continuous variables may not lose significant information, may better reflect the information present, and because the information used in estimation is restricted to what can be reliably reported, may be more robust" (p. 324).

In this paper, I consider statistical models for event-history data with discrete time and discrete covariates in the form of Tables 1 and 2. In general, $\underset{\sim}{z}$ is a vector of dummy variables identifying categories of covariates and their combinations, and t takes on a fixed number of

TABLE 2

Exposure and First Marriages among U.S. Women by Age and Educational Attainment

	Years of Schooling (0–11)			Years of Schooling (12)			Years of Schooling (13–15)			Years of Schooling (16+)		
Age	White	Black	Mexican	White	Black	Mexican	White	Black	Mexican	White	Black	Mexican
Exposure (*O*)												
<15	2153.48	4116.75	1480.50	3080.10	2781.75	668.25	1212.68	1115.25	185.25	962.63	651.00	70.50
15	1279.90	2448.00	886.50	2040.25	1833.50	439.50	806.00	740.50	122.50	640.60	433.50	47.00
16	1076.95	2102.50	762.50	2002.55	1795.00	430.50	800.00	729.50	120.50	639.00	430.00	46.00
17	842.15	1725.50	633.50	1891.45	1702.00	408.00	785.85	707.50	117.00	634.90	424.00	45.00
18	635.20	1409.00	510.00	1622.35	1517.50	359.50	745.60	663.00	107.50	625.05	415.50	44.00
19	490.20	1189.00	414.50	1266.05	1278.00	291.50	653.70	592.00	92.50	605.20	400.00	42.50
20	392.75	1020.00	346.50	964.45	1043.00	225.00	522.00	499.00	75.50	569.60	376.00	40.50
21	319.25	875.50	284.50	727.65	847.50	176.00	386.10	392.50	58.50	508.70	340.50	38.50
22	260.90	762.00	236.50	550.80	704.00	139.00	278.25	305.50	42.50	420.60	295.00	32.50
23	215.50	673.50	197.00	422.90	582.00	108.00	207.30	248.50	29.00	327.65	243.50	24.00
24–25	336.30	1129.00	305.00	606.60	886.00	145.00	289.70	370.00	37.00	460.50	370.00	36.00
26–27	247.60	900.00	217.00	404.80	632.00	85.00	186.00	234.00	22.00	294.90	252.00	22.00
28–30	277.80	1060.50	228.00	414.15	636.00	82.50	180.15	211.50	16.50	286.35	223.50	12.00
31–39	558.45	2070.00	369.00	751.50	972.00	112.50	299.25	310.50	9.00	467.55	310.50	4.50
>40	230.50	785.00	100.00	282.00	300.00	20.00	105.50	85.00	0.00	158.50	90.00	0.00

First Marriages (d)												
<15	29	94	42	9	19	3	1	1	1	1	0	0
15	57	123	47	17	22	9	4	5	1	1	1	0
16	112	175	75	57	52	9	8	17	3	2	6	2
17	139	195	71	144	99	28	19	22	4	6	6	0
18	138	163	96	318	170	49	51	44	13	13	11	2
19	103	133	59	286	158	57	91	49	10	26	20	1
20	77	115	62	234	160	46	99	64	15	44	27	3
21	61	102	48	184	111	28	95	67	7	73	35	1
22	49	78	36	127	85	26	64	32	11	88	36	11
23	37	62	31	96	75	21	44	28	5	69	35	4
24–25	52	93	44	107	80	37	50	42	6	81	36	6
26–27	31	62	29	59	45	12	27	26	3	39	32	4
28–30	27	55	23	44	57	8	21	19	2	29	27	3
31–39	27	80	29	38	25	7	16	14	0	20	12	0
>40	9	39	7	10	14	0	3	4	0	4	1	0

Note: From an extract of the 1980 June Current Population Survey kindly supplied by Lawrence Wu (see Wu and Tuma 1990). A marriage or a censored observation is assumed to contribute 0.5 person-duration of exposure. Duration is set to 1.5 years for the first age interval and 10 years for the last age interval.

intervals (i.e., $t = 1, 2, \ldots T$).[1] In a research setting, the researcher is interested in comparing the hazard function over time and across different groups or strata. That is, interest centers on how $h(t|\underset{\sim}{z})$ varies as a function of t as well as of $\underset{\sim}{z}$. In order to separate the effects of $\underset{\sim}{z}$ from those of t, however, further constraints on $h(t|\underset{\sim}{z})$ are necessary.

1.1 *Loglinear Models*

Cox (1972) popularized the following specification:

$$h(t|\underset{\sim}{z}) = \exp(\underset{\sim}{z}'\underset{\sim}{\beta})\, h_0(t), \tag{1}$$

where $h_0(t)$ is the baseline time-dependency function. The model is called the *proportional hazards model* because the ratio of hazards between any two groups, represented here by $\underset{\sim}{z}_2$ and $\underset{\sim}{z}_1$, is a constant, invariant with respect to t:

$$h(t|\underset{\sim}{z}_2)/h(t|\underset{\sim}{z}_1) = \exp[(\underset{\sim}{z}_2 - \underset{\sim}{z}_1)'\underset{\sim}{\beta}]. \tag{2}$$

Thus a covariate affects the hazard rate as a multiplier of the baseline time-dependency function. It either lowers or raises it by the same amount regardless of t. That is to say, the effects of $\underset{\sim}{z}$ do not depend on t, and likewise the effects of t do not depend on $\underset{\sim}{z}$.

When t is truly continuous, there should be no ties—that is, there should be no two events that occur exactly at the same time. In this case, estimation of $h_0(t)$ is impossible without a parametric assumption, because there are as many unknown parameters for $h_0(t)$ as the number of observations in a sample. Cox's (1972) most significant contribution was to show how partial maximum likelihood estimates of the β parameters in equation (1) could be obtained without a parametric assumption about $h_0(t)$, thus leaving $h_0(t)$ to be free but not estimated.

With discrete time and discrete covariates, estimation of equation (1) is rather simple and dates back at least to Glasser (1967, p. 562). The basic idea is to treat $h_0(t)$ as a set of unknown parameters to be estimated along with the β's in equation (2). As Breslow (1974), Holford (1976, 1980), and Laird and Oliver (1981) later observed, this approach is equivalent to Cox's (1972) model, but it

[1]Time intervals may vary in length, depending on the density of observations within each interval (for example, see Laird and Oliver 1981).

provides an alternative estimation method for grouped data. In fact, Holford (1980) and Laird and Oliver (1981) showed that algorithms for contingency table analysis assuming the Poisson distribution can be directly adapted to estimate equation (1).

To see this, first let $d(t|\underset{\sim}{z})$ denote the number of events that occur at time t given $\underset{\sim}{z}$. In expectation,

$$d(t|\underset{\sim}{z}) = O(t|\underset{\sim}{z})\, h(t|\underset{\sim}{z}), \tag{3}$$

where $O(t|\underset{\sim}{z})$ is the "exposure," or total time at risk of experiencing the event, at time t conditional on $\underset{\sim}{z}$. Substituting equation (1) into (3) and taking the natural logarithm on both sides of (3) leads to the following log-linear model:[2]

$$\begin{aligned} \log[d(t|\underset{\sim}{z})] &= \log[O(t|\underset{\sim}{z})] + \log[h_0(t)] + \underset{\sim}{z}'\underset{\sim}{\beta} \\ &= \log[O(t|\underset{\sim}{z})] + \lambda_t + \underset{\sim}{z}'\underset{\sim}{\beta}. \end{aligned} \tag{4}$$

In the language of log-linear analysis of contingency tables (e.g., Bishop, Fienberg, and Holland 1975; Agresti 1990), $d(t|\underset{\sim}{z})$ is the frequency variable to be explained, $\log[O(t|\underset{\sim}{z})]$ is included as a control variable with a constrained coefficient of unity,[3] $\lambda_t = \log[h_0(t)]$ is the main effect of time, and $\underset{\sim}{z}'\underset{\sim}{\beta}$ as a general expression contains the main effects of covariates and their interactions. For the first example in Table 1, the model can be written in notation for a two-way contingency table as

$$\log(d_{ti}) = \log(O_{ti}) + \lambda_t + \alpha_i, \tag{5}$$

where λ_t ($t = 1, \ldots T$) is the log-additive effect of time t, and α_i ($i = 1, \ldots I$) is the log-additive effect of the ith category of educational attainment, subject to some normalization to ensure that only $T + I - 1$ nonredundant parameters are estimated. In this paper, I conveniently set $\alpha_1 = 0$. It is easy to see that equation (5) implies the same proportionality constraint as does equation (1):

$$\begin{aligned} h_{ti}/h_{ti'} &= (d_{ti}/O_{ti})/(d_{ti'}/O_{ti'}) \\ &= [\exp(\lambda_t)\exp(\alpha_i)] \,/\, [\exp(\lambda_t)\exp(\alpha_{i'})] \\ &= \exp(\alpha_i - \alpha_{i'}), \end{aligned} \tag{6}$$

[2]There is no so-called "intercept" because T instead of $T - 1$ dummies are used for $h_0(t)$. As will be made clear, normalization is achieved by placing constraints on $\underset{\sim}{z}'\underset{\sim}{\beta}$.

[3]In GLIM, $\log[O(t|\underset{\sim}{z})]$ is treated as the "offset." For its implementation in other computer programs, see Clogg and Eliason (1988, p. 241).

where i and i' are any two categories of the covariate variable. Note that λ_t and α_i should be interpreted differently from comparable parameters in log-linear models for contingency tables. In equation (5) the λ_t and α_i parameters represent the *effects* of time and the covariate on the logged hazard rate; in log-linear models for contingency tables, the parameters represent the marginal distributions of the variables. The difference in interpretation occurs because exposure is controlled, making this a model for a rate rather than for a cell frequency.

For the second example (Table 2), the log-linear model for proportional hazards can take several forms. For example, we can specify that the hazard rate of first marriage is affected by education alone, race alone, both education and race but no interaction between them, or both education and race and their interactions. Two specific examples are given as follows:

$$\log(d_{tij}) = \log(O_{tij}) + \lambda_t + \alpha_i; \tag{7a}$$

$$\log(d_{tij}) = \log(O_{tij}) + \lambda_t + \alpha_i + \beta_j, \tag{7b}$$

where λ_t ($t = 1, \ldots T$), α_i ($i = 1, \ldots I$), and β_j ($j = 1, \ldots J$) represent the log-additive effects of age, educational attainment, and race/ethnicity. As a normalization, I set $\alpha_1 = \beta_1 = 0$.

It should be pointed out that the log-linear model for hazard rates in the form of equations (5) and (7) is sometimes called the "log-rate model." Its specification differs from that of the logit model for discrete-time data (e.g., Allison 1984), although the difference diminishes as time intervals narrow. For a comparison of the two alternative specifications, see Yamaguchi (1991). Equation (4) is estimable via maximum likelihood (ML) under the assumption that d follows an independent Poisson distribution within each covariate-by-time classification (i.e., within each cell in the lower panels of Tables 1 and 2). Computer programs for contingency table analysis can be readily adapted for estimation. Different *a priori* assumptions about the β parameters determine different model specifications. For example, we can assume that a particular covariate affects, or does not affect, the hazard rate. The log-linear model for hazard rates of equation (4) implicitly fits the joint distribution of all covariates, analogous to the logit model for binary outcomes (Fienberg 1980, pp. 95–116). The inclusion of the interaction of two covariates in the log-rate model means that the *effects* of one covariate depend on the other covariate.

2. NONPROPORTIONALITY IN EVENT-HISTORY DATA

2.1 *The Issue*

In a research setting, proportional hazards models in the parsimonious form of equation (1) for continuous time or equation (4) for discrete time are often inadequate, and nonproportional hazards models are called for. Sometimes the main theoretical interest is to test whether there are interaction effects between time and covariates. For example, House et al. (1990) proposed that social inequality in health declines with age in later years of life due to the increasing dominance of biological forces over social factors, thus suggesting that socioeconomic differentials in morbidity and functional limitations should converge with age. Relatedly, Manton and his associates (Manton, Poss, and Wing 1979; Manton, Patrick, and Johnson 1987) have considered various explanations for the black/white mortality crossover in the United States. The death rates of blacks are higher than those of whites in earlier ages and then converge with and fall below those of whites in advanced ages. Clearly, such theories of convergence and crossover cannot be tested using proportional hazards models.

One likely source of nonproportionality is population heterogeneity in unmeasured characteristics. Vaupel and Yashin (1985) demonstrate, for example, that mortality crossovers can be attributed to artifacts of population heterogeneity, since "the frailer members of the disadvantaged population die off relatively quickly, leaving a surviving population that largely consists of the robust subcohort" (p. 179). I do not review the literature on latent heterogeneity here. Interested readers should consult other sources (e.g., Heckman and Singer 1982; Vaupel and Yashin 1985; Mare 1992) for methods aimed at modeling unobserved heterogeneity. Instead, the objective of this paper is to model *observed* patterns of nonproportionality with observed variables. Cautions should always be exercised in interpreting nonproportionality parameters as to whether they reveal true nonproportionality or merely reflect selectivity biases due to population heterogeneity.

As an illustration, Figure 1 presents observed hazard rates (in

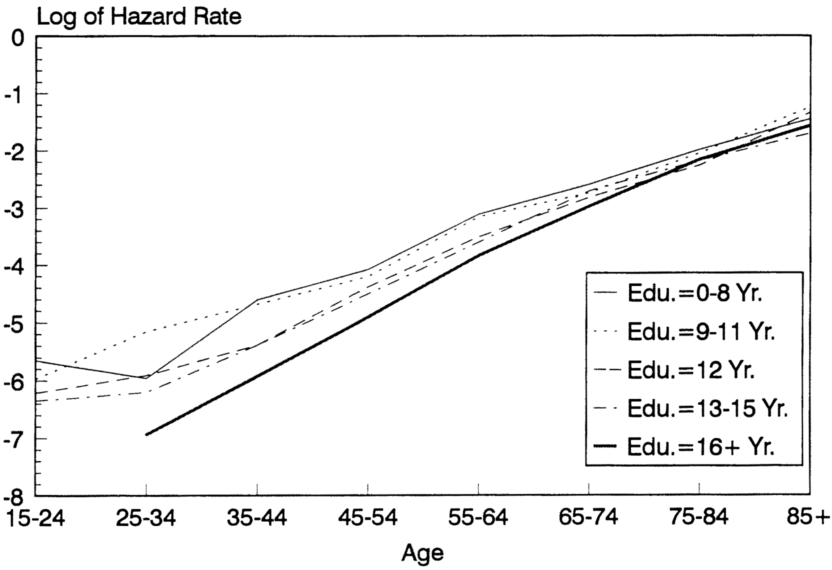

FIGURE 1. Hazard rate of death among U.S. males by age and educational attainment. *Source:* U.S. National Longitudinal Mortality Study (Rogot et al. 1988).

the natural logarithmic scale) from the first example (Table 1). It is evident that the curves for the five educational groups are not parallel, which means that the hazard rates are not proportional. The educational differentials in mortality reach a maximum in the second age interval (24–35) and gradually decline with age. Proportional hazards models are inappropriate for the problem, as they would necessarily force the five lines to differ by a constant.[4] Even though in this case a proportional hazards model would give the correct rank order of the educational groups in terms of social inequality in the average hazard of death, such a model necessarily prevents the researcher from testing the theoretically interesting hypothesis that (observed) social inequality in mortality diminishes with age.

In many other situations, model misspecification due to nonproportionality may lead to incorrect conclusions. Based on the data shown in Table 2, the curves in Figures 2(a) and 2(b) also violate proportionality. That is, the differentials in the hazard rate of first

[4]This is so because Figure 1 plots the logarithm of the hazard rate on the vertical axis.

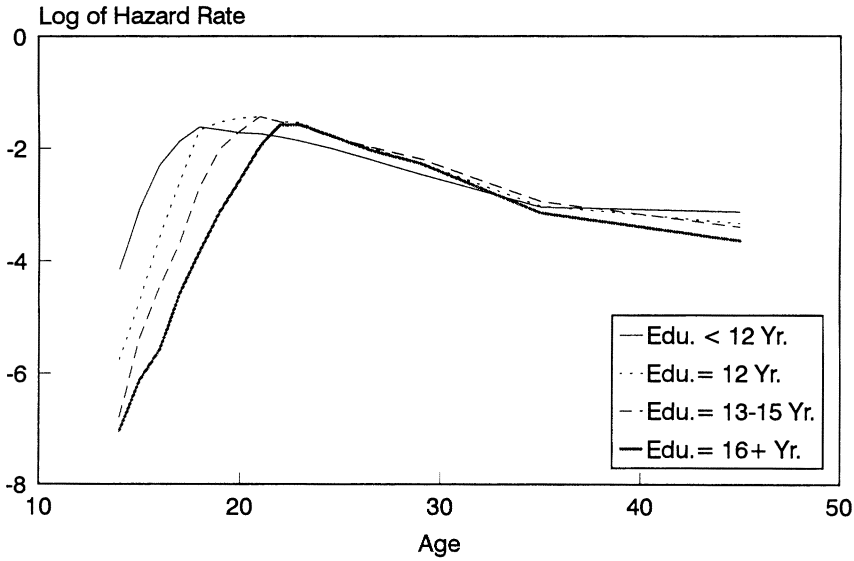

FIGURE 2(a). Hazard rate of first marriage by age and educational attainment. *Source:* 1980 June Current Population Survey (Wu and Tuma 1990).

marriage by educational attainment and race/ethnicity are so age-dependent that any discussions of group differences should be qualified by age. Commenting on the differences between Cox's model and their local hazard models, Wu and Tuma (1990) concluded that "[t]hese results suggest a different behavioral process than the results of the Cox model" (p. 172).

2.2 *Prior Approaches*

There exist numerous approaches to handling nonproportionality—i.e., interaction between time and covariates in event-history analysis. For Cox's model, one standard recommendation (e.g., Lawless 1982, p. 365; Allison 1984, p. 39) is to stratify on covariates for which proportionality does not hold. Stratification means that the researcher specifies multiple baseline hazards functions, one for each stratum, within which proportionality holds true with respect to other covariates. That is, $h_0(t)$ of equation (1) is changed to $h_{0s}(t)$, where the second subscript s refers to the sth stratum. In partial

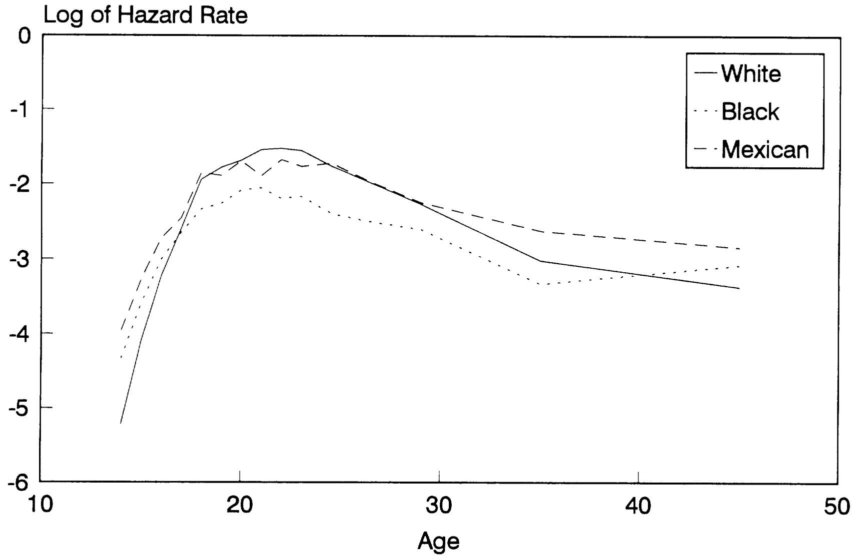

FIGURE 2(b). Hazard rate of first marriage by age and race/ethnicity. *Source:* 1980 June Current Population Survey (Wu and Tuma 1990).

maximum likelihood estimation with continuous measures of time, the stratum-specific $h_{0s}(t)$ function is specified but not estimated.

Nonproportionality can also be modeled more explicitly in the $\exp(\underline{z}'\underline{\beta})$ part of equation (1). This is achieved by including the interaction between a time-varying covariate and other covariates (e.g., Allison 1984, p. 38). The simplest example is to add the product of t and a covariate variable, testing whether the effect of the covariate departs from proportionality as a linear function of time. This solution presumes that the nonproportionality structure is known *a priori,* and that empirical measures for the structure are available. In practice, however, these two conditions are not always met. For example, we know that mortality differentials by educational attainment depend on differences in work and living conditions, lifestyles, knowledge, access to medical care, and a host of other factors. However, it is very difficult to measure all these factors over the whole life span and disentangle their independent and joint contributions.

For discrete-time, discrete-covariate event-history data, stratification is tantamount to the *full* interaction between t and covariates. That is, different hazards curves can be flexibly specified and empirically estimated (e.g., Laird and Oliver 1981), if we change the log-linear model of equation (4) into

$$\log[d(t|\underline{z})] = \log[O(t|\underline{z})] + \lambda_{tz} + \underline{z}'\underline{\beta}, \tag{8}$$

subject to the requirement that λ_{tz} and $\underline{z}'\underline{\beta}$ are normalized. However, the unstructured modeling strategy of equation (8) is likely to lead to the loss of information and imprecision of estimation due to over-parameterization. For example, when all covariates (including their interactions) are nonproportional, equation (8) is equivalent to rewriting observed hazards in the *saturated* model. As a result, we lose the usual advantages of fitting a statistical model to sampled data.[5]

However, parameters from an over-parameterized model can be further analyzed. For example, Yamaguchi (1993) proposed to model nonproportional effects using parameters estimated from the saturated model. In essence, Yamaguchi's method consists of two steps. In the first step, parameters for the saturated model are estimated via ML along with their variance-covariance matrix. In the second step, these parameter estimates and their variance-covariance matrix are then used in testing more constrained models. Starting with a very general and conservative model (i.e., the saturated model), Yamaguchi's method should appeal to researchers who are either at the stage of exploratory data analysis or find other models too restrictive. In contrast, the log-multiplicative approach introduced in this paper specifies and tests a family of restricted models in a single step.

Nonproportionality can also be easily handled in such parametric models as the commonly used exponential, Gompertz, and Weibull models. This is typically done by including the interaction between time-dependency parameters and covariates so that individuals of different covariates follow different time-dependency parametric functions (albeit within the same family). One problem with this approach, however, is that interactions between covariates and time-dependency parameters are usually specified globally—i.e., for the entire range of time. A better approach is to allow for such interactions within local intervals, resulting in piecewise parametric models (e.g., Wu 1991) and local parametric hazard models (Wu and Tuma 1990).

[5]Namely: (1) smoothing scattered observed data, (2) reducing large amounts of observed data, and (3) testing theoretically interesting hypotheses.

Wu and Tuma's paper on local hazard models is worth special mention, for it represents a serious attempt to model the kind of nonproportionality shown in Figures 2(a) and 2(b) within a flexible framework. In essence, a local hazard model is a hybrid between nonparametric and parametric models—globally nonparametric because it breaks down the entire regression curve into separate pieces and locally parametric because it assumes a parametric form (i.e., exponential or Gompertz) in a small neighborhood. In fact, we may view the local hazard approach as stratification, with a parametric hazard function being stratified across time.

3. LOG-MULTIPLICATIVE MODELS FOR EVENT-HISTORY DATA

In this paper, I present a class of models that generalize proportional hazards models by permitting nonproportional effects, while avoiding some of the drawbacks of stratified, interactive, or local hazard models. The solution is flexible in form and applicable to a variety of situations when the proportionality assumption does not hold. As will be shown later, it is a restricted form of the stratified model and reducible to the proportional hazards model when proportionality holds true.

3.1 *Log-Multiplicative Specification*

The interaction model as expressed by equation (8) generally puts no constraints on the interaction effects between time and covariates, as λ_{tz} is allowed to vary freely. Though the model has the advantage of describing nonproportionality at its fullest, it suffers from certain problems. First, the model lacks structure, and its estimated patterns of nonproportionality are the same as the observed patterns. Second, the model is not parsimonious, as the number of parameters for nonproportionality is the product of the number of covariate categories (I or J) involved in nonproportionality and the number of time categories (T). Third, the resulting parameters describing the interaction between time and the covariates are difficult to interpret, not only because there are too many of them but also because they do not follow any prescribed regularity.

I propose to remedy the above shortcomings by borrowing a structural constraint first introduced by Goodman (1979) and later named "log-multiplicative" by Clogg (1982b). The common use of the log-multiplicative specification is to model two-way associations in two-way contingency tables (e.g., Goodman 1979) or variations in two-way log-multiplicative associations across other dimensions in multiway tables (Clogg 1982*a*; Becker 1989; Becker and Clogg 1989). More recently, Goodman (1986, pp. 262–66) proposed to model cross-layer variations in unspecified two-way associations of primary interest in a multiplicative specification for three-way tables. This proposal has been implemented by Xie (1991) and Xie and Pimentel (1992) to study the age patterns of marital fertility rates across different populations and by Xie (1992) to compare social mobility across nations. Here, I extend the log-multiplicative approach to model discrete-time, discrete-covariate event-history data.

For the first example in Table 1, the log-multiplicative model is

$$\log(d_{ti}) = \log(O_{ti}) + \lambda_t + \alpha_i + \tau_t\,\xi_i. \tag{9}$$

τ_t's ($t = 1, \ldots T$) are parameters representing the time pattern of nonproportionality, and ξ_i's ($i = 1, \ldots I$) are parameters representing the nonproportionality levels for different educational groups. For the second example in Table 2, the model can take several forms, three of which are:

$$\log(d_{tij}) = \log(O_{tij}) + \lambda_t + \alpha_i + \beta_j + \tau_t\,\xi_i; \tag{10a}$$
$$\log(d_{tij}) = \log(O_{tij}) + \lambda_t + \alpha_i + \beta_j + \tau_t\,(\xi_i + \psi_j); \tag{10b}$$
$$\log(d_{tij}) = \log(O_{tij}) + \lambda_t + \alpha_i + \beta_j + \tau_t\,\xi_i + \varpi_t\,\psi_j. \tag{10c}$$

As in equation (9), the τ_t's and ϖ_t's ($t = 1, \ldots T$) represent the time pattern of nonproportionality, and ξ_i's ($i = 1, \ldots I$) and ψ_j's ($j = 1, \ldots J$) respectively represent the nonproportionality levels for educational attainment and race/ethnicity. In the special case when the log-multiplicative term (or terms) disappears, the log-multiplicative model of equations (9) and (10) reduces to the proportional hazards model of equations (5) and (7b). Since the log-multiplicative terms in equation (10) are all special cases of λ_{tz} in equation (8), the log-multiplicative model is more restricted than the stratified model, which allows the full interaction between time and covariates.

To illustrate the implications of the log-multiplicative model, let us further explore an example using equation (10b). It follows that

$$h_{tij} = d_{tij}/O_{tij} = \exp(\lambda_t) \exp(\alpha_i) \exp(\beta_j) \exp[\tau_t (\xi_i + \psi_j)]; \qquad \text{or}$$
$$\log(h_{tij}) = \lambda_t + \alpha_i + \beta_j + \tau_t (\xi_i + \psi_j).$$

At the ith level of the first covariate (educational attainment), we have the following hazards model as a function of time and the second covariate (race/ethnicity):

$$\log(h_{tj}|i) = (\lambda_t + \tau_t \xi_i) + \alpha_i + \beta_j + \tau_t \psi_j. \tag{11}$$

If $\tau_t \psi_j$ is absent, as in equation (10a), equation (11) is essentially a proportional hazards model conditional on the first covariate in the form of equation (7b), with the time-dependency function depending on the first covariate $(\lambda_t + \tau_t \xi_i)$. When $\tau_t \xi_i$ and $\tau_t \psi_j$ are both present, equation (11) means that the time pattern of hazards is nonproportional with respect to either covariate.

It is important that α_i be interpreted along with ξ_i, β_j with ψ_j, and λ_t with τ_t, for there can be no true proportional effects in the presence of nonproportionality. In accordance with earlier terminology for the log-linear model, however, I continue to refer to λ_t, α_i, and β_j as "log-additive" effects—i.e., additive effects on logged hazards if nonproportionality were held constant. As will be shown later, λ_t, α_i, and β_j are indeterminate until τ_t, ξ_i, and ψ_j are normalized.

3.2 *Normalization and Estimation*

For the first example in Table 1, the contingency table under consideration is three-way: time by covariate (education) by data type (events versus exposure). The last dimension is used to construct the observed hazard rate (events per unit exposure) and is embedded into log-linear/log-multiplicative models discussed in this paper through equation (4). Similarly, the second example with two covariates involves a four-way table. As mentioned earlier, λ_t, α_i, and β_j represent the effects of time and covariates. Likewise, $\tau_t (\xi_i + \psi_j)$ of equation (10b) represents the interaction effects of time and the covariates—i.e., of nonproportionality.

As latent scores, τ_t, ξ_i, and ψ_j in equation (10b) are indeterminate without normalization (see Xie 1991). Four normalization constraints are necessary, as the locations of τ_t, ξ_i, and ψ_j are indistinguishable from the scales of λ_t, α_i, β_j, and the scale of τ_t is indistinguishable from that of $(\xi_i + \psi_j)$. Whether to normalize by weighting the scores with marginal probabilities or to use unweighted normalization has been a matter of dispute in recent years (e.g., Becker and Clogg 1989; Goodman 1991; Clogg and Rao 1991). The unweighted solution is usually preferred, especially in a comparative context, because unweighted scores are invariant to changes in marginal distributions.[6] Thus I normalize the locations of τ_t, ξ_i, and ψ_j, and then the scale of τ_t, without resorting to weights, imposing the constraints $\Sigma\ \tau_t = \Sigma\ \xi_i = \Sigma\ \psi_j = 0$, and $\Sigma\ \tau_t^2 = 1$. Hence, of the $T + I + J$ parameters in $\tau_t\,(\xi_i + \psi_j)$, only $T + I + J - 4$ are nonredundant. As compared to the model of equation (10b), $T - 2$ additional parameters are needed to identify $\bar{\omega}_t$'s for the model of equation (10c).

When only one covariate is included in the log-multiplicative specification, as in equations (9) and (10a), three normalization constraints are required. For consistency, I set $\Sigma\ \tau_t = \Sigma\ \xi_i = 0$, and $\Sigma\ \tau_t^2 = 1$. If the only covariate involved in the log-multiplicative specification is dichotomous, the log-multiplicative model is equivalent to the interaction model. That is, there is no information reduction, or structural constraint, in equations (9) and (10a), if $I = 2$. The same is true if $T = 2$, although this latter situation is unlikely to occur because time is usually coded in many intervals in order to preserve information.

Because a log-multiplicative term involves the product of two unknown parameters [e.g., $\tau_t\,(\xi_i + \psi_j)$ in equation (10b)], estimation of log-multiplicative models is more complicated than the usual loglinear models. Yet, the complication is revolvable with a proper modification of any standard computer package with a procedure for ML estimation of log-linear models. I here briefly describe an iterative procedure that I used to estimate the models reported in this paper.[7] The first step is to give starting values to log-multiplicative effects (i.e., τ_t, ξ_i, and ψ_j). The second step is to estimate the log-multiplicative effects of the covariates (ξ_i's and ψ_j's) via ML while temporarily treating the starting values for τ_t's as known. With ξ_i's and ψ_j's thus updated and temporarily treated as known, the next

[6]The same invariance property is obtained if the same set of weights is applied to all populations being considered. Since I do not have standard weights *a priori,* I choose to use uniform weights (i.e., not to weight).

[7]All empirical results reported in this paper were estimated using GLIM macros, which are available from the author upon request. I would like to thank Mark Becker for providing some useful GLIM macros, from which my macros were developed.

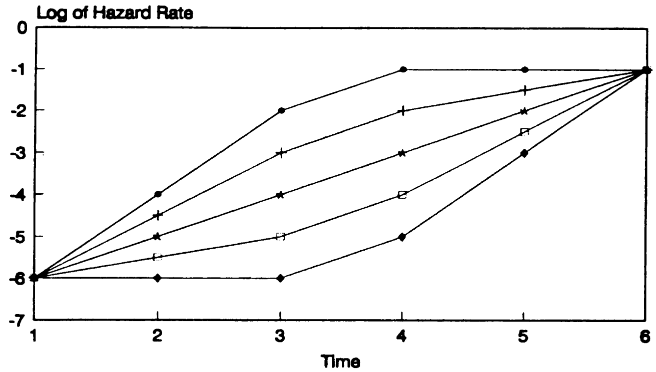

FIGURE 3(a). A stylized example of convergent curves described by a log-multiplicative model.

step is to estimate and update τ_t's. This process continues until all estimated parameters stabilize.

3.3 *Interpretation*

Using the case of equation (9) as an example, we further explore the interpretation of the model parameters. The τ_t parameters can be interpreted as representing the typical *pattern* of deviation from proportionality;[8] and ξ_i can be interpreted as representing the *levels* of deviation from proportionality. The log-multiplicative model specification is parsimonious because the number of parameters required to model nonproportionality is in the order of $(T + I - 3)$, instead of $(T - 1)(I - 1)$ as in the interaction model. For our example in Table 1, the log-multiplicative model uses 10 degrees of freedom for interactions between time and the covariate, whereas the full interactive model uses 28 degrees of freedom. Nonetheless, the log-multiplicative model is flexible, capable of capturing various kinds of deviations from proportionality, such as convergence and crossover of different curves. Three different numerical realizations of equation (9) are given as stylized examples in Figures 3(a)

[8] I use the term "pattern" to describe the interaction effects between the hazard rate and time. Similar uses of the term "pattern" are found in the mobility literature (e.g., Featherman, Jones, and Hauser 1975; Xie 1992) and the demographic literature (e.g., Coale 1971; Xie and Pimentel 1992).

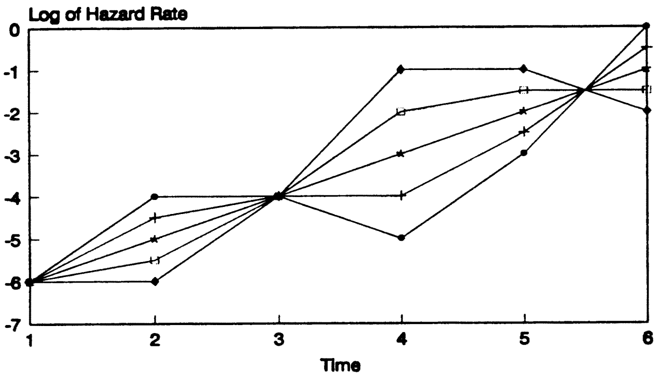

FIGURE 3(b). A stylized example of cross-over curves described by a log-multiplicative model.

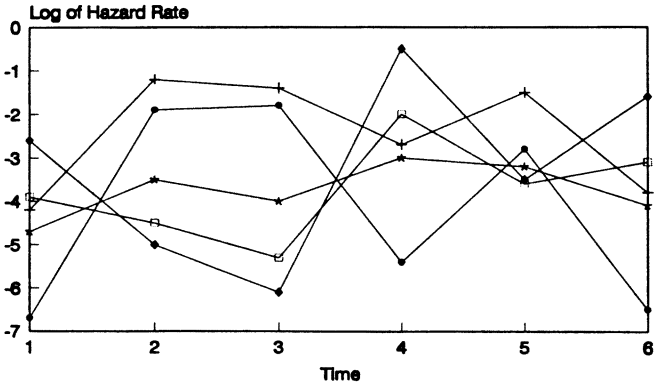

FIGURE 3(c). A stylized example of seemingly irregular curves described by a log-multiplicative model.

through 3(c). A more detailed account of the three examples is given in Appendix B.

The log-multiplicative specification is flexible because the deviation pattern parameters (τ_t's) as well as the baseline time-dependency parameters (λ_t's) are not specified *a priori*. What is specified is the *structure* for nonproportionality in a log-multiplicative form. What is common among Figures 3(a) through 3(c) is the constraint that the distance between any two lines follows the same pattern along the time dimension but is raised or lowered by a multiplier for different contrasts. This can be seen more clearly if we take the ratio of hazards between two groups (say, i and i') at time t:

$$h_{ti}/h_{t,i'} = \exp(\alpha_i - \alpha_{i'}) \exp[\tau_t(\xi_i - \xi_{i'})], \quad (12)$$

a quantity that varies with t only because of τ_t. The τ_t parameters thus determine the nonproportionality pattern in the sense that a large absolute value of (normalized) τ_t indicates a high level of departure from proportionality along the time dimension, and vice versa. Note that τ_t's do not vary with covariates. This means that the basic pattern (or shape) of nonproportionality is the same, but the level of nonproportionality varies across different groups, as shown in equation (12). An analogy with usual log-multiplicative models can be readily made if we turn equation (12) from the relative risk into the log-odds-ratio for any pair of time periods (t and t') and two categories of the covariate (i and i'):

$$\log(h_{ti}\, h_{t',i'}/h_{t',i}\, h_{t,i'}) = (\tau_{t'} - \tau_t)\, (\xi_{i'} - \xi_i). \quad (13)$$

That is, the relative likelihood for transition is determined by the product of the distance between the temporal pattern scores and the distance between the levels of the covariate.

For the log-multiplicative model to be successful, the researcher needs to assume that the pattern of deviation from proportionality is the same for all categories of all the covariates involved in a specification. In many research settings, this assumption may be very sensible. In others, it may not be so sensible, but at least it is *always testable.* Hence, researchers are urged to consider the log-multiplicative specification before adopting a fully interactive model when proportionality fails to hold. Let me now discuss different cases in turn.

Case 1: One covariate present (equation 9). Very often, the covariate explaining differences in the hazard rate is an ordinal variable with respect to its effects on the outcome variable (hazard rate). In the first example, this is explicit. However, this can be implicit when variables such as social origin, cohort, or region are used as the covariate. The order of the categories of the covariate is inconsequential, as the estimation of equation (9) reveals the correct order. For the sake of demonstration, let us assume that the categories are correctly ordered 1, 2, . . . I. It is reasonable to assume that the same mechanism that differentiates categories 1 and 2 is likely to differentiate categories 2 and 3, etc., so that the same time-specific pattern of deviation from proportionality is shared by all categories. For instance, in the first example (Table 1) we may assume that the differentials between those with 0–8 years of schooling and those with 9–11 years of schooling follow the same *pattern,* albeit with a different level, as the differentials between those with 13–15 years of schooling and those with 16+ years of schooling. This reasoning is consistent with the hypothesis that over time biological factors increase their importance relative to social factors. As mortality differentials by education converge with age (Figure 1), the advantages or disadvantages of social groups diminish, regardless of their relative positions in the social hierarchy. This hypothesis predicts that the pattern parameters (τ_t) monotonically decreases over time.

Case 2: Two covariates present, one covariate involved in the log-multiplicative specification (equation 10a). The interpretation is the same as in Case 1, as long as the second covariate has only log-additive effects. Case 2 is a hybrid between the log-multiplicative model (with respect to the first covariate) and the log-linear model (with respect to the second covariate). A variation of this, combining the interaction model and the log-multiplicative model, occurs when the first covariate is included in the log-multiplicative specification, and analysis is stratified such that the second covariate has an unrestricted interaction with time. Interpretation of this hybrid model is similar to that of equation (10a) in that the two covariates are distinct.

Case 3: Two covariates present, and both used in a single log-multiplicative term (equation 10b). In this case, the two covariates can be thought of as sharing a common time-specific pattern of deviation from proportionality. In the second example (Table 2), for instance, both educational attainment and race/ethnicity may cause nonproportionality in similar patterns by age. Formally, the constraint is τ_t $(\xi_i + \psi_j)$ of equation (10b). Affecting the hazard rate nonproportionally through a common vector of pattern parameters (τ_t's), ξ_i and ψ_j parameter estimates enable the researcher to compare relative degrees of nonproportionality of the two covariates.

Case 4: Two covariates present, and involved in two separate log-multiplicative terms (equation 10c). The key feature of this model is to partition the log-multiplicative term into two dimensions. Multidimensional log-multiplicative models have received extensive treatments in various contexts (Goodman 1986; Becker 1989; Becker and Clogg 1989; Xie 1992). Interpretation of this case is similar to Case 3. The main difference is that in Case 4 two sets of time-dependent deviation parameters (τ_t and ϖ_t) are separately normalized and estimated, as the two covariates no longer share the same pattern of deviation from nonproportionality.

Case 5: With two covariates, three-way interactions present involving the two covariates and time. There could be many different specifications for two-way and three-way interactions. A special log-multiplicative specification for three-way interactions takes the form $\tau_t \; \xi_i \; \psi_j$. In this case, the researcher conditions the log-multiplicative specification of one covariate's nonproportionality on another covariate (see Goodman 1986, p. 263; Xie 1992, p. 386). This model is more difficult to interpret and thus less desirable than those in Cases 2 through 4 involving only two-way interactions.

Case 6: More than two covariates present. Even though I do not consider this case in the present paper, the models discussed in Cases 1 through 5 are readily generalizable to situations when there are more than two covariates. Interpretations of such models remain the same.

As is true with any model-based approach to analyzing empirical data, the different specifications of log-linear and log-multiplicative models should be initiated from a theoretical point of view and tested against observed data. For this purpose, I recommend the use of such standard tools as the log-likelihood ratio chi-squared statistic, the Pearson chi-squared statistic, and the BIC statistic.

4. EMPIRICAL EXAMPLES

The first three rows of Table 3 present three log-linear models for hazard rates applied to the first example. The goodness-of-fit of the models is assessed first by the log-likelihood ratio chi-squared statistic (L^2) and the Pearson chi-squared statistic (X^2) along with their degrees of freedom (DF). However, it is well known that, with large samples, the log-likelihood ratio chi-squared test and the Pearson chi-squared test are likely to reject a good model. For this reason, I also use Schwarz's (1978) Bayesian criterion (BIC) as adapted by Raftery (1986) for contingency table settings:

TABLE 3
Goodness-of-Fit Statistics of Log-linear and Log-Multiplicative Models Applied to Data in Table 1

Model	Description	L^2	X^2	DF	Δ	BIC
E	Constant	9778.90	16387.11	39	55.71%	9443.94
T	Age	265.56	265.20	32	8.07	−9.28
A	Age + Edu	94.19	94.55	28	4.29	−146.29
X	Age + Edu + Age × Edu	27.71	24.57	18	2.03	−126.89
X^*	Age + Edu + Age* × Edu*	34.99	31.07	22	2.12	−153.97

Note: L^2 is the log-likelihood ratio chi-squared statistic, and X^2 is the Pearson chi-squared statistic, both with the degrees of freedom reported in column DF. Δ is the Index of Dissimilarity. BIC = $L^2 - (DF)\log(D)$, where D is the total number of deaths (5,371). Edu is an abbreviation for educational attainment. Age* is a constrained function of age, and Edu* is a constrained function of education (see Table 4 and text). The multiplicative sign (×) denotes log-multiplicative specification. Two cells of zero exposure are blocked out from all the models.

$$\text{BIC} = L^2 - (DF)\log D, \tag{14}$$

where D is the total number of events. The rule is to select the model with the lowest BIC value. When BIC is negative, the null hypothesis is preferred relative to the saturated model, whose L^2 and BIC are by definition zero. The saturated model in this framework means full stratification for all possible combinations of covariates. As a purely descriptive measure, I also use the Index of Dissimilarity (Shryock and Siegel 1976, p. 131), denoted as Δ. The Index of Dissimilarity here can be interpreted as the proportion of events that would have to be reclassified in order for the statistical model to achieve perfect prediction of the observed events.

As a point of reference, the first model in Table 3 is the exponential (E) model, which naively assumes that the hazard does not vary either by time or by education. Clearly, the model does not fit the data. The second model is the time-dependency (T) model. With seven additional parameters for age, model T greatly improves the fit. The third model is of the form described by equation (5) with one covariate, and it is referred to as the "log-additive (A) model" because its parameters have log-additive effects on hazards. From Table 3, it is easy to observe that model A fits the data better than does model T: L^2 is reduced by 171.37 to 94.19 using only four degrees of freedom; Δ and BIC also decline respectively from 8.07

TABLE 4
Estimated Parameters of Models in Table 3

Variables	Model T	Model A	Model X	Model X^*
Log-Additive Effects of Age (λ_t)				
15–24	−6.075	−5.930	−5.936	−5.867
25–34	−6.035	−5.729	−5.594	−5.545
35–44	−5.264	−4.985	−4.894	−4.862
45–54	−4.387	−4.156	−4.067	−4.037
55–64	−3.382	−3.188	−3.113	−3.082
65–74	−2.721	−2.578	−2.437	−2.399
75–84	−2.066	−1.961	−1.795	−1.770
85+	−1.442	−1.354	−1.160	−1.122
Log-Additive Effects of Edu (α_i, excluded = 0–8)				
9–11		−0.025	−0.062	−0.072
12		−0.274	−0.383	−0.419
13–15		−0.314	−0.444	−0.536
16+		−0.567	−0.782	−0.795
Log-Multiplicative Effects of Age or Age* (τ_t)				
15–24			0.220	0.103
25–34			0.580	0.549
35–44			0.392	0.413
45–54			−0.019	−0.056
55–64			−0.007	
65–74			−0.328	−0.446
75–84			−0.381	
85+			−0.457	−0.562
Log-Multiplicative Effects of Edu or Edu* (ξ_i)				
0–8			0.528	0.501
9–11			0.593	
12			0.000	−0.014
13–15			−0.242	−0.487
16+			−0.881	

Note: Edu (abbreviation for educational attainment) is measured in years of schooling completed. Log-multiplicative effects of age are normalized with a mean of zero and a sum of squares equal to one, and those of Edu are normalized with a mean of zero.

percent and −9.28 to 4.29 percent and −146.29. In other words, the comparison of models A and T reveals that there are significant differences in hazards across the five educational groups.

Parameter estimates for models T and A are displayed in the first two columns in Table 4. The log-additive effects of educational attainment are normalized as contrasts to the excluded category of

0–8 years of schooling. As expected, both models T and A show that the logged hazard of death (i.e., force of mortality) increases with age. The log-additive effects of education in Model A agree with the expectation from the stratification literature that more education is associated with a lower likelihood of death. The differentials occur most drastically at two points: (1) with or without a high school education (12 years of schooling), and (2) with or without a college education (16 years of schooling). For example, compared to those with 0–8 years of schooling, a high school education reduces the hazard of death by about 24 percent; compared to those with just a high school education, a college education further reduces the hazard of death by 25 percent.[9]

The last two rows of Table 3 report goodness-of-fit statistics for two log-multiplicative models. The difference between models X and X^* is that X^* is a restricted version of X with certain equality constraints imposed on the log-multiplicative effects of age and education. The log-multiplicative model X and the loglinear model A are nested, and according to the difference in the L^2 statistic, model X is better than model A (a chi-squared reduction of 66.48 for 10 degrees or freedom). The Index of Dissimilarity drops significantly from 4.29 percent to 2.03 percent. By the BIC criterion, however, model X is not as good as model A, due to the fact that model X uses up too many degrees of freedom. The problem of over-parameterization in model X is solved by constraining some of the log-multiplicative effects in model X^*. (The constraints were derived after inspection of parameter estimates for model X.) As shown in Table 4, I constrain the log-multiplicative effects of age to be equal for age categories 45–54 and 56–64 and for age categories 65–74 and 75–84, and the log-multiplicative effects of education to be equal before 12 years of schooling and after 12 years of schooling. The constraints saved 4 degrees of freedom, making model X^* ($L^2 = 34.99$ with 22 degrees of freedom, BIC $= -153.97$) preferable to model X.

The estimated parameters of models X and X^* are displayed

[9]Note the following formulas: $1 - \exp(-0.274) = 24$ percent; $1 - \exp(-0.567 + 0.274) = 25$ percent. The total benefit of having a college education through high school is $1 - (1 - 24 \text{ percent})(1 - 25 \text{ percent}) = 0.43$, which is the same as what is calculated directly from the effect of a college education compared to 0–8 years of schooling: $1 - \exp(-0.567)$.

in Table 4. Given the currently implemented normalization $\Sigma\ \tau_t = \Sigma\ \xi_i = 0$ and $\Sigma\ \tau_t^2 = 1$, the log-additive effects can be interpreted as "unweighted" mean effects on the log of the hazard. Hence, the parameter estimates of models X and X^* are consistent with those of model A, indicating that on average the hazard rate of death monotonically increases with age and declines with educational attainment. The normalized log-multiplicative effect of age in model X first rises from 0.220 for age 15–24 to 0.580 for age 25–34 and then drops continuously with the slight aberration for age 55–64. The exception for age 55–64 can be attributed to sampling error, as it is smoothed in model X^*. Also after smoothing in model X^*, the normalized log-multiplicative effects of education decrease monotonically. The two sets of log-multiplicative effects should be interpreted jointly. For example, multiplying the positive effect of 0–11 years of schooling by the generally decreasing trend of age effects means that the age pattern of nonproportionality favors those with higher educational attainment at early ages (with lower relative risks of death) and penalizes them at later ages (with higher relative risks of death). To compare the educational groups on an absolute scale, however, it is necessary to add the log-additive effects of education. In other words, it cannot be directly inferred from the log-multiplicative effects how hazard rates for groups differ in the time range actually observed. Rather, the log-multiplicative effects can tell us only how the group differences vary over time—i.e., the interaction between time and the covariates. For the first example, the estimated ξ_i parameters indicate that mortality differentials by educational attainment decrease with age. Figure 4 presents the hazard rates predicted from these parameter estimates for the educational groups. It is no surprise that the five curves in Figure 4 closely match those in Figure 1, but the curves in Figure 4 are smoother.

While the first example has only one covariate (educational attainment), the second example has twó covariates (educational attainment and race/ethnicity). Table 5 reports the goodness-of-fit statistics for models that were fitted to the data displayed in Table 2. As before, models E, T, and A denote the exponential model, the time-dependency model, and the log-additive model, respectively. Models A_1 and A_2 were previously written respectively in equations (7a) and (7b). In model I_1, the full interaction between age and education is fitted; and model I_2 adds the full interaction between age

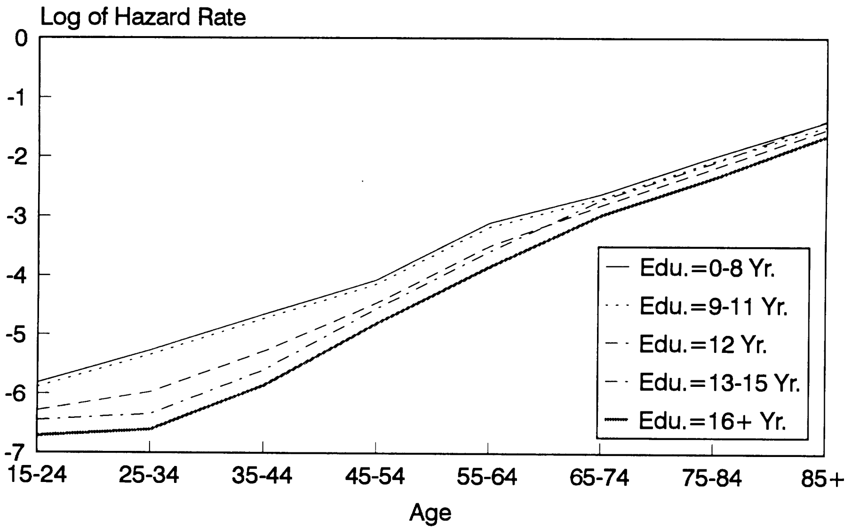

FIGURE 4 Predicted hazard rate of death among U.S. males by age and educational attainment.

and race/ethnicity.[10] In estimating the models via ML, I blocked out two cells with zero observed exposure by including two dummy variables for them in all statistical models. The effect is to ignore the cells for which data are missing while leaving other cells intact (see Goodman 1968).

The goodness-of-fit statistics reported in the first four lines in Table 5 confirm the expected importance of age, education, and race/ethnicity in explaining the hazard of first marriage, as the goodness-of-fit improves steadily from model E to model A_2. However, model A_2 still does not fit the observed data well. The lack of fit is mostly due to nonproportionality in the effects of the covariates, as the introduction of the interactions between age and the covariates in models I_1 and I_2 greatly improves the goodness-of-fit. From model A_2 to I_1, for example, L^2 is reduced by 1092.58 to 219.88 using 42 degrees of freedom. Δ and BIC are also drastically reduced from 15.02 percent and 114.92 to 5.76 percent and −828.07, respectively. The addition of the interaction between age and race/ethnicity in

[10]The interaction model was not listed for the first example because it is a saturated model—exhausting all degrees of freedom and reproducing the observed data exactly.

TABLE 5

Goodness-of-Fit Statistics of Log-linear and Log-Multiplicative Models Applied to Data in Table 2

Model	Description	L^2	X^2	DF	Δ	BIC
E	Constant	6294.07	6032.30	177	35.48%	4695.03
T	Age	1911.34	1827.93	163	18.86	438.79
A_1	Age + Edu	1642.73	1634.72	160	18.32	197.27
A_2	Age + Edu + Race	1312.46	1278.02	158	15.02	−114.92
I_1	Age + Edu + Race + Age•Edu	219.88	220.32	116	5.76	−828.07
I_2	Age + Edu + Race + Age•Edu + Age•Race	98.01	98.78	88	3.24	−696.99
X_1	Age + Edu + Race + Age × Edu	327.54	380.96	142	6.92	−955.30
X_2	Age + Edu + Race + Age × (Edu + Race)	264.20	322.48	140	5.97	−1000.58
Y_2	Age + Edu + Race + Age × Edu + Age•Race	226.74	269.97	127	5.47	−920.59

Note: L^2 is the log-likelihood ratio chi-squared statistic, and X^2 is the Pearson chi-squared statistic, both with the degrees of freedom reported in column DF. Δ is the Index of Dissimilarity. BIC = $L^2 - (DF)\log(D)$, where D is the total number of first marriages (42305). Edu is an abbreviation for educational attainment. Dot (•) denotes full interaction between two categorical variables; and the multiplicative sign (×) denotes log-multiplicative specification. Two cells of zero exposure are blocked out from all the models.

model I_2 results in similar but less impressive gains in goodness-of-fit: Model I_2 is better than model I_1 by the chi-squared test (a difference of 121.87 in L^2 for 28 degrees of freedom), but model I_1 is better than model I_2 according to the BIC criterion. This seems to suggest that the effects of race/ethnicity are proportional after the nonproportionality in the effect of education is taken into account. Before rushing to such a conclusion, we should realize that many degrees of freedom (30 for model I_1 and 58 for model I_2) are used in describing nonproportionality. A more sensitive test of nonproportionality with fewer degrees of freedom may yield different results, as will be shown with log-multiplicative models.

Estimated parameters of models T and A_2 are displayed in the first two columns of Table 6. The parameters representing age show the same pattern as in Figure 2(a) and 2(b): The hazard of first marriage rises rapidly from ages less than 15 to the early 20s and then gradually declines. The log-additive effects of education indicate that the hazard of first marriage is on average inversely affected by educational attainment. Likewise, the log-additive effects of race/ethnicity can be interpreted to mean that overall white women have higher hazards of first marriage than women of Mexican origin, who in turn have higher hazards of first marriage than black women. These results are consistent with those from Wu and Tuma's (1990, Table 2) application of Cox's continuous-time proportional hazards model.

Three log-multiplicative models are listed in the last three rows of Table 5, as X_1 (equation 10a), X_2 (equation 10b), and Y_2 (equation 10c).[11] Model X_1 is an intermediate model between A_2 and I_1, and model X_2 between A_2 and I_2. The advantages of the log-multiplicative model specification emerge most clearly when compared to the log-linear model on the one hand and the interaction model on the other. Let us first stay with the reduction in L^2 as the criterion for assessing relative goodness-of-fit of a model. Relative to model A_2 (L^2 = 1312.46 with 158 degrees of freedom), X_1 reduces L^2 to 327.54 with 142 degrees of freedom. This compares to model I_1's reduction in L^2 to 219.88 with 116 degrees of freedom. That is, for merely 16 degrees of freedom (from A_2 to X_1), model X_1 explains

[11]Further simplified models (models X_1^*, X_2^*, and Y_2^*) with constrained functions of age were fitted and are found to be slightly better. The results of models X_1^*, X_2^*, and Y_2^* are not reported here but are available from the author upon request.

TABLE 6
Estimated Parameters of Selected Models in Table 5

Variables	Model T	Model A_2	Model X_2	Model Y_2	
Log-Additive Effects of Age (λ_t)					
<15	−4.526	−4.189	−5.685	−5.599	
15	−3.709	−3.367	−4.518	−4.532	
16	−3.050	−2.700	−3.353	−3.411	
17	−2.605	−2.244	−2.567	−2.674	
18	−2.092	−1.717	−1.697	−1.843	
19	−1.997	−1.606	−1.455	−1.592	
20	−1.860	−1.456	−1.243	−1.347	
21	−1.809	−1.397	−1.127	−1.239	
22	−1.835	−1.418	−1.110	−1.228	
23	−1.867	−1.448	−1.124	−1.234	
24–25	−2.059	−1.641	−1.303	−1.417	
26–27	−2.249	−1.834	−1.465	−1.568	
28–30	−2.444	−2.042	−1.641	−1.731	
31–39	−3.147	−2.763	−2.454	−2.539	
>40	−3.165	−2.796	−2.752	−2.744	
Log-Additive Effects of Edu (α_i, excluded = 0–11)					
12		−0.141	−0.174	−0.147	
13–15		−0.387	−0.581	−0.529	
16+		−0.651	−1.423	−1.327	
Multiplicative Effects of Race (β_j, excluded = White)					
Black		−0.423	−0.378	−0.277	
Mexican		−0.083	−0.060	0.015	
Log-Multiplicative Effects of Age (τ_t)				(For Edu)	(For Race)
<15			0.573	0.537	0.625
15			0.474	0.479	0.362
16			0.313	0.328	0.177
17			0.194	0.222	−0.010
18			0.030	0.065	−0.194
19			−0.054	−0.028	−0.234
20			−0.104	−0.096	−0.135
21			−0.164	−0.158	−0.198
22			−0.222	−0.215	−0.247
23			−0.227	−0.226	−0.202
24–25			−0.215	−0.209	−0.213
26–27			−0.227	−0.233	−0.117
28–30			−0.230	−0.255	0.039
31–39			−0.137	−0.147	−0.020
>40			−0.002	−0.064	0.367

(continued on next page)

TABLE 6 (continued)

Variables	Model T	Model A_2	Model X_2	Model Y_2	
Log-Multiplicative Effects of Edu (ξ_i)					
0–11			3.447	3.298	
12			1.128	1.087	
13–15			−0.647	−0.646	
16+			−3.928	−3.739	
Log-Multiplicative Effects of Race (ψ_j)					
White			−0.417		−0.736
Black			0.517		0.595
Mexican			−0.100		0.142

Note: Edu (abbreviation for educational attainment) is measured in years of schooling completed. Log-multiplicative effects of age are normalized with a mean of zero and sum of squares equal to one, and those of covariates are normalized with a mean of zero.

about 90 percent of the nonproportionality explained by the I_1 model, which uses 42 degrees of freedom. Similarly, for merely 19 degrees of freedom, model X_2 explains about 86 percent of the nonproportionality explained by the I_2 model, which uses 70 degrees of freedom. In addition, model X_2 fits the data well in absolute terms of all available criteria. The comparison of models X_1 and X_2 reveals that X_2 is a better model by either the L^2 or the BIC criterion. Moving from X_1 to X_2 reduces L^2 by 63.34 for two degrees of freedom, which is highly significant; BIC for X_2 is more negative than that of X_1. This result confirms the nonproportional effects of race/ethnicity earlier observed by Wu and Tuma (1990). If I had relied on interaction models to detect nonproportionality, I might have concluded that race/ethnicity has only proportional effects, for I_2 is better than I_1 according to the BIC criterion.

Model Y_2 relaxes the assumption that the two covariates do not share the same age pattern of deviation from nonproportionality. Model X_2 can be seen as a restricted version of model Y_2, which uses 13 (15 minus 2 for normalization) additional degrees of freedom. However, the null assumption of a common deviation pattern cannot be rejected, as model Y_2 does not improve the goodness-of-fit over model X_2. Thus model X_2 is preferred over model Y_2.

The parameter estimates of models X_2 and Y_2 are given in Table 6. The interpretation of them is similar to those of models X

and X^* in Table 4, except that the effects of education and race/ethnicity should now be interpreted in a multivariate context, that is, net effects after the statistical control of the other covariate. It is evident that the two sets (one for education and the other for race/ethnicity) of log-multiplicative effects of age in model Y_2 are *estimated* to be similar. This explains why the restriction imposing a common age deviation pattern in model X_2 does not result in much deterioration of the fit. Whether this result can be duplicated in other research situations with other covariates is an empirical question subject to empirical investigation. However, the restricted version of the log-multiplicative model (equation 10b) is preferred to the unrestricted version of the log-multiplicative model (equation 10c), because the former is much more parsimonious and easier to interpret than the latter.

The log-multiplicative effects of age for model X_2 can be interpreted as the typical age-specific distance between any two groups of an education and race/ethnicity combination. The normalized estimates reveal that the typical distance, net of differing levels, starts very large and then rapidly narrows down before it grows a little toward the last age interval. This pattern is a reverse image of the log-additive effects of age, which represent the (unweighted) average age-variation in the hazard rate of first marriage. The normalized log-multiplicative effects of educational attainment parallel the log-additive effects of educational attainment in that the effects are the largest for those with less than 12 years of schooling and decline monotonically. We observe that for lower educational groups the log-multiplicative term moves in the direction compensating the general age pattern, and for higher educational groups the log-multiplicative term moves in the direction exacerbating the general age pattern. Thus lower educational attainment is associated with an overall higher hazard rate of first marriage but a flatter age pattern, and vice versa. This is shown graphically in Figure 5(a). Note that the difference between Figures 5(a) and 2(a) is partly due to smoothing and partly due to the control of race/ethnicity.

For the covariate of race/ethnicity, the rank order of the log-additive effects is the opposite of the log-multiplicative effects. The log-additive effects of race/ethnicity tell us that on average whites have higher hazards than those of Mexican origin, who in turn have higher hazards than blacks. The log-multiplicative effects reveal that

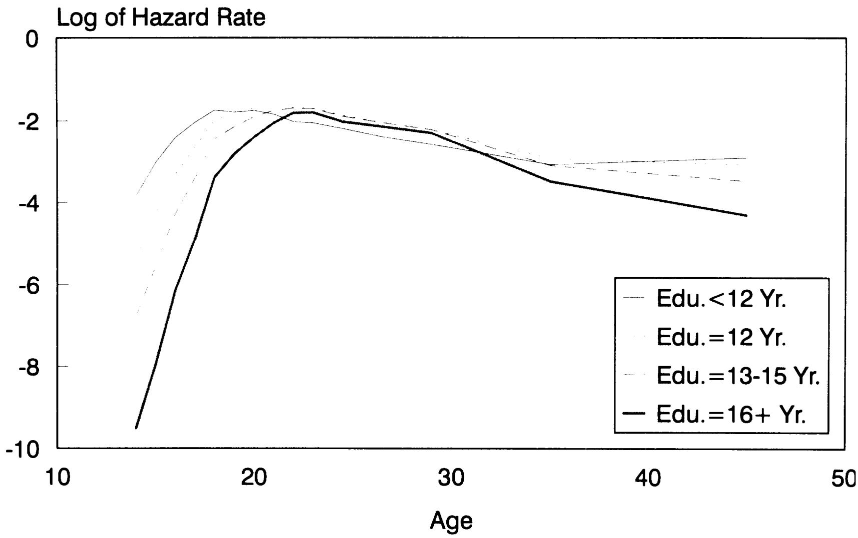

FIGURE 5(a). Predicted hazard rate of first marriage by age and educational attainment, controlling for race/ethnicity.

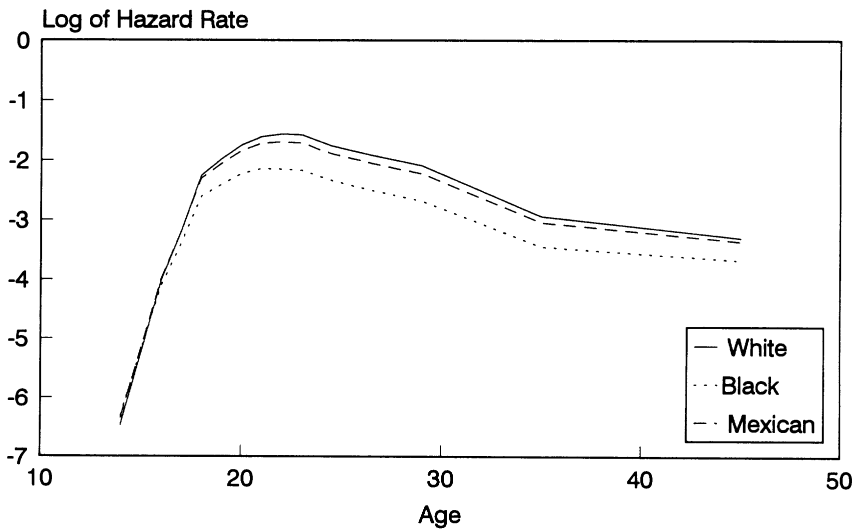

FIGURE 5(b). Predicted hazard rate of first marriage by age and race/ethnicity, controlling for education.

the race/ethnicity differences start small in early ages, grow to maxima in the middle range (age 20s) and narrow slightly down toward the end. Figure 5(b) plots the logged hazards predicted by age and race/ethnicity. Due to the statistical control of educational attainment, Figure 5(b) differs noticeably from Figure 2(b). In Figure 2(b) the observed hazard of whites appears to be lower than that of blacks and those of Mexican origin in early and later ages because of the indirect effect of whites' overall higher educational attainment.

5. CONCLUSION

This paper proposes a new class of models, log-multiplicative models, for the analysis of nonproportionality in hazard rates. The new models are variants of the log-linear (or log-rate) model for discrete-time, discrete-covariate data, with the number of events as the dependent variable and exposure as a control. Both event-history data and synthetic cohort data can be used. The new approach builds upon Goodman's (1979) innovative log-multiplicative specification for two-way tables and later development of the log-multiplicative specification for multiway tables (notably Clogg 1982*a*; Goodman 1986; Becker and Clogg 1989; Xie 1991, 1992).

The new models fall between the simple case of log-additive models for proportionality and the complicated case of unconstrained interaction models for nonproportionality. In a variety of applications, log-multiplicative models may fit observed data almost as well as the comparable unconstrained interaction models. But log-multiplicative models in general are much simpler and easier to interpret. The parsimony of log-multiplicative models is achieved by separating nonproportionality into two components: the nonproportionality *pattern* over time and the nonproportionality *level* across groups. This assumption may be reasonable for many research situations in social science and is always testable.

APPENDIX A: THE DATA

Two data sets were used in this paper. The first (given in Table 1) was drawn from the U.S. National Longitudinal Mortality Study (Rogot et al. 1988, Table 6). Exposure (upper panel) was measured not in person-years, but in person-periods, with a period being the length of the study. The study design was such that the length of the follow-up period differed for eight cohorts of subjects. Ideally, the researcher should distinguish the eight cohorts in an analysis of the data. For the illustrative purposes of this paper, I took the simplistic approach of combining all subjects in order to have a large enough sample, as was done in Rogot et al. (1988). For each age × education combination, a death was assumed to contribute 0.5 person-period of exposure. Strictly speaking, the data were not event-history data, because the study did not follow the same individuals over time. Rather, the data in Table 1 provide us a snapshot of individuals of different ages at one point in time. Thus there are no logical constraints on amounts of exposure between two adjacent age intervals.

In contrast, the second data set (given in Table 2), extracted from the 1980 June Current Population Survey by Wu and Tuma (1990), is more typical of event-history data. Educational attainment was measured using four categories, and race/ethnicity using three. For cells in the upper right-hand corner, educational attainment was likely to follow rather than precede the event of first marriage, thus violating the direction of causality implicit in regression models treating education as a covariate. Still, I chose this example because nonproportionality was well researched in this case by Wu and Tuma (1990), with the caution that the effects of education should be interpreted as results of a host of factors, including population heterogeneity and true causality of education. Since substantive interest lies in comparisons of the hazard rates of first marriage by educational attainment and race/ethnicity, I weighted the data for whites by 10 percent so that the three race/ethnicity groups would have comparable numbers of observations. Otherwise, whites would dominate the age patterns and the educational effects. This treatment is tantamount to a stratified sample with whites undersampled and blacks and women of Mexican origin oversampled.

For the second data set, I divided the age variable originally measured in months into 15 intervals, with small intervals at early ages and wide intervals at later ages. As in the case of smoothing within a neighborhood for local hazard models (Wu and Tuma 1990, pp. 155–56), aggregation for discrete-time models imposes the assumption of within-interval homogeneity. The choice of interval widths should balance the need for smoothing, on the one hand, against that for preserving information on time-dependency, on the other. Intervals that are too wide, for example, may seriously conceal variation in hazards over time. In practice, the researcher should try to ensure that there are enough observed events (d) in each cell after covariates are introduced. I experimented with several schemes for categorizing time before choosing the 15-category scheme presented in Table 2. Similar results were found with tables obtained using other schemes.

In calculating exposure for the second data set, I used the following formulas (adapted from Laird and Oliver 1981, Table 8):

$$O(t|\underset{\sim}{z}) = \{n(t|\underset{\sim}{z}) - [d(t|\underset{\sim}{z}) + w(t|\underset{\sim}{z})]/2\}\delta(t), \qquad \text{(A1)}$$

where n is the number of never married respondents at the beginning of an age interval, d and w are respectively the number of respondents who were married and the number of respondents who were censored during an age interval, and δ is the interval duration in years. Duration for the first age interval was set to 1.5 years, and for the last age interval to 10 years. Arbitrary assignment was necessary but would not greatly affect the substantive results, because all models presented in this paper included the main effects of discrete time. The same approach was adopted by Laird and Oliver (1981). Also note that exposure in Table 2 is constrained across age intervals within a classification of covariates. That is, the number of subjects at risk at any age equals the number of subjects at risk at the earlier age minus the number of subjects who either experienced the event or withdrew from the study at the earlier age:

$$n(t|\underset{\sim}{z}) = n(t-1|\underset{\sim}{z}) - [d(t-1|\underset{\sim}{z}) + w(t-1|\underset{\sim}{z})]. \qquad \text{(A2)}$$

Because $\log(O)$ is included as a control variable with a constrained coefficient of unity (equation 4), the implicit relationship of equation (A2) in Table 2 has no bearing on the statistical estimation of the models presented. Thus this distinction between the two data sets is inconsequential.

APPENDIX B: HYPOTHETICAL EXAMPLES

Table B1 describes the three hypothetical examples graphed in Figures 3(a) through 3(c). They were all constructed using equation (9), but with different combinations of model parameters. For each exam-

TABLE B1
Three Hypothetical Examples

Example A, as in Figure 3(a)

Log of Hazard (Main Entries)

				Covariate Dimension				
			i	1	2	3	4	5
			α_i	0.0	0.0	0.0	0.0	0.0
Time Dimension			ξ_i	2.0	1.0	0.0	−1.0	−2.0
t	λ_t	τ_t						
1	−6.0	0.0		−6.0	−6.0	−6.0	−6.0	−6.0
2	−5.0	0.5		−4.0	−4.5	−5.0	−5.5	−6.0
3	−4.0	1.0		−2.0	−3.0	−4.0	−5.0	−6.0
4	−3.0	1.0		−1.0	−2.0	−3.0	−4.0	−5.0
5	−2.0	0.5		−1.0	−1.5	−2.0	−2.5	−3.0
6	−1.0	0.0		−1.0	−1.0	−1.0	−1.0	−1.0

Example B, as in Figure 3(b)

Log of Hazard (Main Entries)

				Covariate Dimension				
			i	1	2	3	4	5
			α_i	−2.0	−1.0	0.0	1.0	2.0
Time Dimension			ξ_i	2.0	1.0	0.0	−1.0	−2.0
t	λ_t	τ_t						
1	−6.0	1.0		−6.0	−6.0	−6.0	−6.0	−6.0
2	−5.0	1.5		−4.0	−4.5	−5.0	−5.5	−6.0
3	−4.0	1.0		−4.0	−4.0	−4.0	−4.0	−4.0
4	−3.0	0.0		−5.0	−4.0	−3.0	−2.0	−1.0
5	−2.0	0.5		−3.0	−2.5	−2.0	−1.5	−1.0
6	−1.0	1.5		0.0	−0.5	−1.0	−1.5	−2.0

Example C, as in Figure 3(c)

Log of Hazard (Main Entries)

				Covariate Dimension				
			i	1	2	3	4	5
			α_i	−1.4	0.8	0.0	0.5	1.5
Time Dimension			ξ_i	2.0	1.0	0.0	−1.0	−2.0
t	λ_t	τ_t						
1	−4.7	−0.3		−6.7	−4.2	−4.7	−3.9	−2.6
2	−3.5	1.5		−1.9	−1.2	−3.5	−4.5	−5.0
3	−4.0	1.8		−1.8	−1.4	−4.0	−5.3	−6.1
4	−3.0	−0.5		−5.4	−2.7	−3.0	−2.0	−0.5
5	−3.2	0.9		−2.8	−1.5	−3.2	−3.6	−3.5
6	−4.1	−0.5		−6.5	−3.8	−4.1	−3.1	−1.6

ple, parameters along the time dimension are given in the first two columns, and parameters along the covariate dimension in the first two rows. Note that the parameters in Table B1 are not normalized as are those in Tables 4 and 6. The main entries are the natural logarithm of the hazard rate (i.e., $\log(h_{ti})$'s).

REFERENCES

Agresti, Alan. 1990. *Categorical Data Analysis.* New York: Wiley.

Allison, Paul D. 1984. *Event History Analysis: Regression for Longitudinal Event Data.* Beverly Hills: Sage Publications.

Becker, Mark P. 1989. "Models for the Analysis of Association in Multivariate Contingency Tables." *Journal of American Statistical Association* 84:1014–19.

Becker, Mark P., and Clifford C. Clogg. 1989. "Analysis of Sets of Two-Way Contingency Tables Using Association Models." *Journal of the American Statistical Association* 84:142–51.

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul W. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice.* Cambridge: MIT Press.

Breslow, N. 1974. "Covariance Analysis of Censored Survival Data." *Biometrics* 30:89–99.

Clogg, Clifford C. 1982*a*. "Some Models for the Analysis of Association in Multiway Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 77:803–15.

———. 1982*b*. "Using Association Models in Sociological Research: Some Examples." *American Journal of Sociology* 88:114–34.

Clogg, Clifford C., and Scott R. Eliason. 1988. "Some Common Problems in Log-Linear Analysis." In *Common Problems/Proper Solutions: Avoiding Error in Quantitative Research,* edited by J. Scott Long, 226–57. Beverly Hills: Sage Publications.

Clogg, Clifford C., and C. R. Rao. 1991. "Comment on 'Measures, Models, and Graphical Displays in the Analysis of Cross-Classified Data,' by L. A. Goodman." *Journal of American Statistical Association* 86:1118–20.

Coale, Ansley. J. 1971. "Age Patterns of Marriage." *Population Studies* 25:193–214.

Cox, D. R. 1972. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society B* 34:187–220.

Featherman, David L., F. Lancaster Jones, and Robert M. Hauser. 1975. "Assumptions of Social Mobility Research in the U.S.: The Case of Occupational Status." *Social Science Research* 4:329–60.

Fienberg, Stephen E. 1980. *The Analysis of Cross-Classified Categorical Data.* 2d ed. Cambridge: MIT Press.

Glasser, Marvin. 1967. "Exponential Survival with Covariance." *Journal of the American Statistical Association* 62:561–68.

Goodman, Leo A. 1968. "The Analysis of Cross-Classified Data: Independence, Quasi-Independence, and Interactions in Contingency Tables with or without Missing Entries." *Journal of the American Statistical Association* 63:1091–131.

———. 1979. "Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 74:537–52.

———. 1986. "Some Useful Extensions of the Usual Correspondence Analysis Approach and the Usual Log-Linear Models Approach in the Analysis of Contingency Tables." *International Statistical Review* 54:243–309.

———. 1991. "Measures, Models, and Graphical Displays in the Analysis of Cross-Classified Data." *Journal of the American Statistical Association* 86:1085–138.

Heckman, James J., and Burton Singer. 1982. "Population Heterogeneity in Demographic Models." In *Multidimensional Mathematical Demography,* edited by Kenneth C. Land and Andrei Rogers, 567–99. New York: Academic Press.

Holford, Theodore R. 1976. "Life Tables with Concomitant Information." *Biometrics* 32:587–97.

———. 1980. "The Analysis of Rates and of Survivorship Using Log-Linear Models." *Biometrics* 36:299–305.

House, James S., Ronald C. Kessler, A. Regula Herzog, Richard P. Mero, Ann M. Kinney, and Martha J. Breslow. 1990. "Age, Socioeconomic Status, and Health." *Milbank Quarterly* 68:383–411.

Laird, Nan, and Donald Oliver. 1981. "Covariance Analysis of Censored Survival Data Using Log-Linear Analysis Techniques." *Journal of the American Statistical Association* 76:231–40.

Lawless, J. F. 1982. *Statistical Models and Methods for Lifetime Data.* New York: Wiley.

Manton, Kenneth G., Clifford H. Patrick, and Katrina W. Johnson. 1987. "Health Differentials between Blacks and Whites: Recent Trends in Mortality and Morbidity." *Milbank Quarterly* 65:129–98.

Manton, Kenneth G., Sharon S. Poss, and Steven Wing. 1979. "The Black/White Mortality Crossover: Investigation from the Perspective of the Components of Aging." *The Gerontologist* 19:291–300.

Manton, Kenneth G., Max A. Woodbury, Eric Stallard, and Larry S. Corder. 1992. "The Use of Grade-of-Membership Techniques to Estimate Regression Relationships." In *Sociological Methodology 1992,* edited by Peter V. Marsden, 321–81. Washington: American Sociological Association.

Mare, Robert D. 1992. "Discrete-Time Bivariate Hazards with Unobserved Heterogeneity: An Indirect Contingency Table Approach." Center for Demography and Ecology Working Paper 92–7. University of Wisconsin-Madison.

Namboodiri, Krishnan, and C. M. Suchindran. 1987. *Life Table Techniques and Their Applications.* Orlando: Academic Press.

Raftery, Adrian E. 1986. "Choosing Models for Cross-Classifications (Comment on Grusky and Hauser)." *American Sociological Review* 51:145–46.

Rogot Eugene, Paul D. Sorlie, Norman J. Johnson, Claudia S. Glover, and David W. Treasure. 1988. *A Mortality Study of One Million Persons by Demographic, Social, and Economic Factors: 1979–1981 Follow-Up*. Washington: U.S. Department of Health and Human Services.

Schwarz, G. 1978. "Estimating the Dimension of a Model." *The Annals of Statistics* 6:461–64.

Shryock, Henry S., and Jacob S. Siegel. 1976. *The Methods and Materials of Demography*. Orlando: Academic Press.

Vaupel, James W., and Anatoli I. Yashin. 1985. "Heterogeneity Ruses: Some Surprising Effects of Selection on Population Dynamics." *American Statistician* 39:176–85.

Wu, Lawrence L. 1991. "Modeling the Shape of the First Marriage Rate: Pooling, Statistical Power, and Nonproportionality." Center for Demography and Ecology Working Paper 91-17. University of Wisconsin-Madison.

Wu, Lawrence L., and Nancy Brandon Tuma. 1990. "Local Hazard Models." In *Sociological Methodology 1990*, edited by Clifford C. Clogg, 141–80. Washington: American Sociological Association.

Xie, Yu. 1991. "Model Fertility Schedules Revisited: The Log-Multiplicative Model Approach." *Social Science Research* 20:355–68.

———. 1992. "The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables." *American Sociological Review* 57:380–95.

Xie, Yu, and Ellen E. Pimentel. 1992. "Age Patterns of Marital Fertility: Revising the Coale-Trussell Method." *Journal of the American Statistical Association* 87:977–84.

Yamaguchi, Kazuo. 1991. *Event History Analysis*. Newbury Park, Calif.: Sage Publications.

———. 1993. "Modeling Time-Varying Effects of Covariates in Event-History Analysis Using Statistics from the Saturated Hazard Rate Model." In *Sociological Methodology 1993*, edited by Peter V. Marsden, 279–317. Washington: American Sociological Association.