Title
Accelerated Failure Time Regression Model with a Regression Model of Surviving Fraction: An Application to the Analysis of RPermanent EmploymentS in Japan

Author
Kazuo Yamaguchi

Publication Date
2011-10-24

Accelerated Failure-Time Regression Models With a Regression Model of Surviving Fraction: An Application to the Analysis of "Permanent Employment" in Japan

KAZUO YAMAGUCHI*

Accelerated failure-time regression models with an additional regression model for the surviving fraction are proposed for the analysis of events that may never occur, regardless of censoring, for some people in the population risk set. The models attempt to estimate simultaneously the effects of covariates on the acceleration/deceleration of the timing of a given event and the surviving fraction; that is, the proportion of the population for which the event never occurs. The extended family of the generalized Gamma distribution is used for the accelerated failure-time regression model; the logistic function is used for the regression model of the surviving fraction. The models are applied to the data of interfirm job mobility in Japan to assess variability in "permanent employment" among white collar and blue collar employees in firms of different sizes, independent from their variability in the timing of interfirm job separations.

KEY WORDS: Accelerated failure time; Generalized gamma model; Job mobility; Permanent employment; Surviving fraction.

Hazard rate models can reflect two distinct factors associated with high/low hazard rates: (1) acceleration/deceleration in the timing of the event and (2) high/low *limiting survival probability.* The limiting survival probability implies the limiting value of the survivor function when time goes to infinity. A nonzero limiting survival probability often is called a *surviving fraction* (Miller 1981). Although we never can observe subjects up to time infinity, the assumption of a nonzero limiting survival probability seems consistent with the characteristics of empirical data for many life events. Some specific parametric functions with a nonzero limiting survival probability have been shown to fit the data of marriage, divorce, or occupational mobility better than representative parametric functions with a zero limiting survival probability (Diekmann and Mitter 1983, 1984; Diekmann 1989; Hernes 1972; Wu 1990).

Most parametric hazard rate models implicitly assume a zero limiting survival probability. On the other hand, proportional hazards models based on Cox's partial likelihood method are compatible with nonzero limiting survival probabilities, because the unspecified baseline hazard function can take a value of 0 for any time t at which the event does not occur in the sample. In analyses based on the proportional hazards and related hazard rate models, however, the interpretations of high/low hazard rates often are ambiguous or arbitrary regarding the distinction between timing and limiting survival probability of the event's occurrence.

For example, for an event experienced by most people in the population risk set, such as marriage, a higher hazard rate generally is interpreted as an advancement or acceleration of timing with respect to age; a lower hazard rate is interpreted as a postponement or deceleration. On the other hand, for events not experienced by a fairly large proportion of people in the population risk set, such as divorce or the birth of a second child, a higher hazard rate usually is interpreted as a higher lifetime probability for the occurrence of the event rather than an acceleration in the timing of the event.

However, alternative interpretations of high/low hazard rates can be given for both situations. For example, blacks have lower rates of marriage than whites not because they marry later, but rather because they are less likely to marry (Bennett et al., 1989; Schoen and Kluegel 1988; Spanier 1983). Similarly, an early age of first childbirth predicts not only a higher lifetime probability of having a second child (Millman and Hendershot 1980), but also a shorter interval between the first and second births (Marini and Hodsdon 1981).

Proportional hazards models do not permit an useful *parametric* distinction between the effects of covariates on the timing of the event and the effects of covariates on the limiting survival probability of the event. Unlike proportional hazards models, accelerated failure-time models assume that high/low hazard rates result solely from acceleration/deceleration in the timing of the event. They also typically assume a zero limiting survival probability. Although the assumption of a zero limiting survival probability is unrealistic for many life events, the fact that their parameters only govern acceleration/deceleration and do not affect the limiting survival probability makes it possible to modify and extend these models to include another set of parameters governing the limiting survival probability. The modification is made by (1) introducing a surviving fraction into the accelerated failure-time models and (2) applying the logistic regression model for the surviving fraction.

An important point here is that although both the accelerated failure-time model and the proportional hazard model can include a surviving fraction, the surviving fraction becomes independent of other parameters only for the former model. For the accelerated failure time model that satisfies

* Kazuo Yamaguchi is Professor, Department of Sociology, University of Chicago, Chicago, IL 60637. This article was written when the author was affiliated with the Department of Sociology, UCLA. The research is partially supported by National Science Foundation Grant SES-9008163 and a UCLA Academic Senate Grant. The author thanks James S. Coleman, an associate editor, and all referees for their comments on earlier drafts of the article. In addition, he thanks Linda Ferguson for her research and editorial assistance.

$S(t, \theta) = S_0(\theta t)$, where $S(t, \theta)$ is the survivor function, $S(\infty, \theta)$ is independent of parameter θ regardless whether $S_0(\infty) = 0$ or not. However, for the proportional hazard model that satisfies $S(t, \theta) = S_0(t)^\theta$, $S(\infty, t)$ is not independent of θ unless $S_0(\infty) = 0$.

The regression model for the surviving fraction was described by Farewell (1977) and Miller (1981) for the exponential model without covariates. In this paper, the regression model is extended to a general class of accelerated failure-time models with covariates. Thus, the modified model incorporates variability in both timing and limiting survival probability and, through a pair of regression equations, permits separate predictors to be identified simultaneously for both timing and limiting survival probability.

One sociological topic for which the distinction between timing and limiting survival probability is important is the assessment of "permanent employment" or "lifetime employment" in Japan. Two distinct conceptualizations of permanent employment exist: One refers to the management practices of firms and the other to the tenure of employees. In the first concept, permanent employment may imply that some Japanese firms employ only people who have just graduated from school and that they never fire or lay off employees (Abegglen 1958; Dore 1973). The second concept, may imply that employees tend to work for their first employer until retirement (Cole 1972; Tominaga 1964), which is similar to the concept of employees' "life-time commitment" to their firms (Marsh and Mannari 1971). This article deals with the latter concept.

Permanent employment can be studied as a theoretical concern (i.e., regarding why it exists) and also as an empirical concern (i.e., regarding the extent to which it exists); this article is concerned primarily with the latter concern. Previous studies that assessed the extent of permanent employment in Japan analyzed the rate of interfirm job mobility (Cheng 1988; Cole 1979; Cole and Siegel 1980; Taira 1962; Tominaga 1964; Tsuzuki 1989). A major limitation in these studies was a lack of separation between the two "components" responsible for the low rates of interfirm job mobility; late timing of job separations and a high surviving fraction. Generally, the interpretation given to a low rate of interfirm job separations is that employees in a given category of firms have a high tendency to remain in their firms. A low rate, however, may reflect a long duration in the firm before leaving. This article eliminates the ambiguity of interpretation by using new models that simultaneously identify the predictors of high-versus-low limiting probability of remaining in the same firm and the predictors of early-versus-late timing of interfirm job separations. An interfirm job separation is defined as a voluntary or involuntary separation from the employer, excluding retirement.

1. MODELS

Let T and $Y = \log(T)$ be random variables for failure time and the logarithm of failure time. Let $f(y)$ be the conditional pdf of Y, given that the event occurs, and let $g(y)$ be the unconditional pdf of Y. We assume a positive surviving fraction p. It follows that

$$\begin{aligned} g(y) &= (1 - p)f(y) && \text{when } y < \infty \\ g(y) &= p && \text{when } y = \infty. \end{aligned} \tag{1}$$

Then the survivor function corresponding to $g(y)$, $S_g(y)$, can be expressed using the survivor function corresponding to $f(y)$, $S_f(t)$, as follows:

$$S_g(y) = (1 - p)S_f(y) + p. \tag{2}$$

For the conditional pdf of $T = \exp(Y)$, we assume a general class of accelerated failure time models, namely, the extended family of generalized Gamma models. [Alternately, for $Y = \log(T)$, we assume the extended family of generalized log-Gamma models.] The generalized Gamma model was introduced by Stacy (1962) and extended by Prentice (1974). The extended family of generalized Gamma models, described in detail by Lawless (1982) and Kalbfleisch and Prentice (1980), includes exponential, Weibull, reciprocal Weibull, log-normal, and Gamma as its special cases. Applying the extended family to failure-time data is now a standard SAS procedure—procedure LIFEREG in the SAS statistical package (SAS Institute 1985).

The extended family of generalized Gamma models has considerable flexibility in capturing the characteristics in the distribution of T. The shape of the survivor function becomes even more flexible by introducing the surviving fraction p in Formula (2). Further, using the extended family rather than just the generalized Gamma model has certain advantages. For instance, the generalized Gamma model reflects only negatively skewed pdf for $\log(T)$, but the extended family can reflect both positively skewed and negatively skewed pdf for $\log(T)$. Empirical distributions of the pdf for the logarithm of employment duration usually reveal positive skewness.

The error term for the generalized Gamma regression model for the conditional pdf of $y = \log(T)$ follows the generalized log-Gamma distribution such that

$$y = \mathbf{X}'\boldsymbol{\beta} + \sigma z, \tag{3}$$

where $\mathbf{X}$ is the covariate vector, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)'$ is the parameter vector, σ is the scale parameter, and z has the standard log-Gamma distribution with shape parameter k such that:

$$\begin{aligned} f(z; k) &= \frac{k^{k-1/2}}{\Gamma(k)} \exp(\sqrt{k}z - ke^{z/\sqrt{k}}) && \text{when } 0 \le k < \infty \\ &= \frac{1}{(2\pi)^{1/2}} \exp(-z^2/2) && \text{when } k = \infty. \end{aligned} \tag{4}$$

The extended family is obtained by replacing parameter k by parameter $\lambda = k^{-1/2}$ and allowing both positive and negative values of λ. Its pdf then becomes

$$\begin{aligned} f(z; \lambda) &= \frac{|\lambda|}{\Gamma(\lambda^{-2})} (\lambda^{-2})^{\lambda-2} \exp[\lambda^{-2}(\lambda z - e^{\lambda z})] && \text{when } \lambda \neq 0 \\ &= \frac{1}{(2\pi)^{1/2}} \exp(-z^2/2) && \text{when } \lambda = 0. \end{aligned} \tag{5}$$

We also need to specify the functional form for the dependence of the surviving fraction p on covariates $\mathbf{X}$. To do

so, the logit model suggested by Farewell (1977) and Miller (1981) is employed:

$$p = \exp(\mathbf{X}'\alpha)/[1 + \exp(\mathbf{X}'\alpha)], \quad (6)$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)'$ is the parameter vector.

In this article, we apply two types of models. The first type, models with a *constant surviving fraction,* assumes a single constant value for p with $p = 0$ as a special case. The second type, models with *heterogeneous surviving fractions,* assumes in Formula (6) the same set of covariates used for the model of $\log(T)$ in Formula (3) with a separate set of parameters—that is, α ($\neq\beta$)—as their coefficients.

The log-likelihood function for the ith observation of $Y = \log(T)$ for the generalized log-Gamma model modified by the introduction of surviving fractions becomes:

$$\begin{aligned}\log L_i(\alpha, \beta, \sigma, k) &= \delta_i[\log(1 - p(\alpha)) + \log f(z_i; k, \beta, \sigma) - \log \sigma] \\ &\quad + (1 - \delta_i)\log[(1 - p(\alpha))Q(k, ke^{z_i\sqrt{k}}) + p(\alpha)] \quad (7)\end{aligned}$$

for $0 < k < \infty$. Here $z_i = [\log(T_i) - \mathbf{X}_i'\beta]/\sigma$; $\delta_i = 1$ if T_i is an observation of failure time, and $\delta_i = 0$ if T_i is an observation of censoring time; $f(z; k, \beta, \sigma)$ is given by Formula (4), and $Q(k, a)$ is the incomplete gamma integral

$$Q(k, a) = \int_a^\infty \frac{x^{k-1}}{\Gamma(k)} e^{-x}\, dx. \quad (8)$$

For $k = \infty$ or $\lambda = 0$, we simply need to replace $Q(k, ke^{z_i\sqrt{k}})$ by the normal integral from z_i to infinity. For the extended family with a *positive* λ, we need to replace k by λ^{-2} in Formula (7). For a *negative* λ, we obtain the corresponding log-likelihood function by replacing (a) $f(z_i; k, \beta, \sigma)$ by $f(-z_i; \lambda^{-2}, \beta, \sigma)$ and (b) $Q(k, ke^{z_i\sqrt{k}})$ by $[1 - Q(\lambda^{-2}, \lambda^{-2}e^{-z_i|\lambda|})]$ in Formula (7).

Generally, a good estimate of the surviving fraction is obtained when ample data for the end of the normal risk period are available. It follows that a mechanical application of the model could pose a problem. In the analysis of first marriages based on cross-sectional data, for example, the sample should include a sufficient number of people whose ages are over the normal period of risk, such as over age 40. Applying the model to a sample of young people (such as those age 30 or younger) would present a problem, because the data do not provide any direct observation of survival probabilities around the end of the normal risk period. Hence, the estimated values of p will depend heavily on (a) fitting a particular parametric function—the extended family of the generalized Gamma—to the observed survivor function for young ages and (b) the capacity of this parametric distribution to extrapolate the survivor function for older ages. Such an extrapolation will not be accurate in many cases, however.

The model introduced here implicitly assumes two latent subpopulations: one with zero risk of having the event and the other with a nonzero risk subject to the generalized Gamma model. In the model, the proportion of the latent subpopulation with zero risk in the total population is fixed and equal to the surviving fraction p. Although the assumption of a time-invariant proportion of the latent population with zero risk may not hold true empirically, the estimate of p always reflects the "empirical surviving fraction" expected from the shape of the empirical survivor function around the end of normal period of risk.

2. METHODS

Parameter estimation for the extended family of the generalized Gamma model was described in detail by Lawless (1982). This procedure is modified here to incorporate the simultaneous estimation of parameters for the surviving fraction. First, the maximum likelihood estimates for parameters α, β, and $\log(\sigma)$ are obtained for each fixed value of λ, using the Newton–Raphson algorithm. Then, the search for the value of λ that maximizes the log-likelihood function is made. The calculation of the incomplete Gamma integral is based on the formula of a series development presented by Abramowitz and Stegun (1972, p. 262).

One problem exists in comparing parameters for models that introduce surviving fractions as a logistic function of covariates. Because the estimate of surviving fraction p may approach 0 for some categories of nominal covariates, the estimates of certain parameters thereby approach minus infinity, and, further, their standard errors approach infinity more rapidly than these parameter estimates approach minus infinity. Although we still can attain the convergence of estimates for other parameters and the log-likelihood in such situations, the significance of parameter estimates that approach minus infinity is lost asymptotically. Hence the difference between a parameter with an estimate of minus infinity and another parameter is tested with a one-degree-of-freedom likelihood ratio test, instead of the test based on the variance-covariance matrix of parameter estimates. The likelihood ratio test compares the model that imposes the identity of two parameters with the model that does not make this imposition.

Model selection is based primarily on the Bayesian information criterion (BIC) introduced by Raftery (1986). This statistic for comparing models among exponential families was initially advocated by Schwarz (1978). Raftery (1986) showed its general usefulness in selecting among models for contingency tables, and Heckman and Walker (1987) recommended its use for selecting among models for duration data. Although Raftery's version uses the saturated model for comparison, here BIC is calculated using the constant rate model for comparison:

$$\text{BIC} = L^2 - (np)\log(N), \quad (9)$$

where L^2 is the likelihood ratio chi-squared statistic (with respect to the likelihood function of $Y = \log(T)$ for the significance test of parameters in the model against the constant rate model, np is the number of parameters that the tested model adds to the constant rate model, and N is the number of observations. The model that maximizes BIC is the best model among those compared.

3. DATA AND HYPOTHESES

The data for the analyses presented in this article come from the 1975 Social Stratification and Mobility Survey in

Japan, whose major results were edited by Tominaga (1979). Cheng (1988), Grusky (1983), and Yamaguchi (1987) also used this data set to analyze job shift patterns, status attainment, and intergenerational occupational mobility. The survey collected work history data from a national representative sample in Japan. Because the timing of job separations was coded by the age of occurrence rather than by the year and month of occurrence, the failure-time data are somewhat crude. Nevertheless, this survey provides the only publicly available data of complete work histories in Japan.

The present analysis is based on a sample of nonfarm male employees age 20–64 in 1975. The dependent variable is the log-duration of the first full-time employment after the completion of schooling. The following variables are employed as predictors for the timing of interfirm job separations and the surviving fraction: (a) a six-category scheme for firm size (private firms with 1–4 employees, 5–29 employees, 30–299 employees, 300–999 employees, and 1,000 or more employees, and government) and (b) the distinction between white collar and blue collar occupations. These two variables are selected because of their centrality in previous research (Cheng 1988; Cole 1979; Koike 1983a,b; Tominaga 1964; Tsuzuki 1989).

Firm size has been used in previous research as the major correlate of interfirm job mobility and age-wage profiles in Japan (Cheng 1988; Cole 1979; Hashimoto 1990; Koike 1983a,b; Tan 1980; Tominaga 1964; Tsuzuki 1989). Theoretically, permanent employment with the seniority-based wage system found in large Japanese firms is considered as a functional alternative to the internal labor market (Cole 1973; Sumitani 1974a,b) found in some American firms to retain workers with firm specific skills (Doeringer and Piore 1971). To the extent that internal labor markets are associated with large firms, we expect firm size to have a positive effect on the surviving fraction of interfirm job separations. Hence we expect:

Hypothesis 1. As firm size becomes larger, the surviving fraction will be larger.

Employees in small firms, who have lower salaries or wages, security, and fringe benefits (Koike 1983b; Steven 1983), are expected to have a higher probability of finding better jobs in other firms. This is especially true when employees are young, because labor markets are less rigid regarding their job mobility. Hence, employees in small firms have a significantly higher relative frequency of interfirm job separations than do employees in larger firms, especially when the duration of their first employment is short. Thus we expect:

Hypothesis 2. As firm size becomes smaller, the timing of interfirm job separation will be more accelerated, given that job separation occurs.

Although government and large private firms are known for the wide practice of permanent employment (Cole 1979; Yamaguchi 1983), the government is known to have a larger salary return for tenure. This salary schedule is characterized by a smaller salary at the beginning of employment and a larger salary (including retirement allowance) at the end of employment (Yamaguchi 1983). Compared with employees of large private firms, government workers who leave their jobs have less to lose in terms of salary in the beginning and more to lose later. Hence we expect:

Hypothesis 3. Compared to employees of large private firms, government employees will have an accelerated timing of interfirm job separation, given that job separation occurs.

However, the extent of permanent employment as employees' tendency to work for their employer until retirement may not be different between the two groups of firms, because both groups typically have permanent employment policies. In other words, the surviving fraction can be the same in government and large private firms.

The importance of the distinction between white collar and blue collar employees has been documented in previous studies of job mobility (e.g., Cheng 1988; Tominaga 1964) and age–wage profiles (Koike 1983a,b). Studies of large manufacturing firms (Abegglen 1958; Dore 1973) found permanent employment policies applied equally to white collar and blue collar workers. However, studies of job mobility found differences between white collar and blue collar employees regarding the rate of interfirm job separations (Cheng 1988; Tominaga 1964). We shall examine whether differences between white collar and blue collar workers tend to reflect differences in the tendency to remain in the same firm "permanently" or differences in the timing of job separations.

The raw data used for the analysis are given in Table 1. All the results from the application of models presented in subsequent tables can be reproduced from this table. The duration data represent the difference between the age of the first full-time employment and either the age of the first interfirm job separation, for those who had a separation, or the age at the time of the survey for censored cases. Age differences greater than 30 are treated as censored at the 31st year, because interfirm job separations after more than 30 years of employment tend to involve retirement and the data do not distinguish retirement from other voluntary job separations. By setting the 31st year of employment as an additional censoring time, we effectively can eliminate the confounding of retirement with occurrences of the event. This is necessary because many large private firms had early retirement policies during the time period covered in this study, such as mandatory retirement at age 55 (Cole 1979; Koike 1983a). This strategy may not seem workable for subjects entering full-time employment at an older age (age 26 or over), who may retire before serving 30 years in their firms. Only 1.3% of subjects in the risk set, however, had a late first employment and left their firms before the survey date; further, the oldest age of departure among them was 50, still too early for a retirement.

Some cases have a duration value of 0, occurring when subjects left their employers at the same age as they were hired. Because duration values of 0 generate a problem, two alternative treatments are tested. The first treatment is to add 1 to all duration values. The second treatment is to assign the expected value of duration to cases with 0 duration according to (a) certain assumptions about the distribution of entries into the risk set and (b) the assumption of a uniform

Table 1. Number of Events and Censored Observations by Firm Size, Occupation, and the Difference Between the Age at Entry Into the First Job and Age at the Occurrence of First Interfirm Job Separation/Censoring: Employed Med Age 20–64 in 1975

Age difference	Number of events: Firm size[a] 1	2	3	4	5	6	Number of censored observations: Firm size[a] 1	2	3	4	5	6	Number of events: Firm size[a] 1	2	3	4	5	6	Number of censored observations: Firm size[a] 1	2	3	4	5	6
	I. White collar workers												II. Blue collar workers											
0	2	2	3	3	2	1	0	0	1	1	1	6	5	4	7	3	2	0	0	1	0	0	1	0
1	7	11	11	4	8	18	0	0	1	2	4	3	8	33	18	13	13	2	0	0	1	0	0	0
2	6	13	15	9	8	14	0	1	3	2	4	5	5	37	31	9	19	2	0	2	0	1	2	0
3	5	16	14	7	9	16	1	1	4	1	9	2	13	26	19	15	19	3	1	2	4	1	4	0
4	4	12	12	3	7	6	0	2	3	2	6	5	16	24	15	4	13	9	0	3	2	2	2	0
5	6	6	12	1	12	8	2	1	2	1	6	1	15	23	8	9	12	4	1	1	3	2	4	0
6	5	6	7	2	5	3	0	2	3	1	2	1	8	11	9	5	6	1	0	1	4	1	3	0
7	2	6	8	3	1	3	1	1	7	4	6	3	3	11	8	5	3	2	0	5	6	4	1	0
8	3	2	3	4	1	8	0	4	4	6	4	4	4	9	3	0	2	1	0	5	4	0	5	0
9	0	4	2	0	2	3	0	0	5	3	6	3	5	6	4	1	1	1	1	5	2	0	3	0
10	0	4	6	0	5	2	0	2	3	2	8	8	3	7	5	2	8	1	0	4	2	0	3	0
11	0	1	1	0	3	1	1	0	3	1	4	5	0	3	1	2	3	0	0	0	1	0	6	0
12	0	2	1	2	4	0	0	1	2	3	5	1	2	3	2	1	4	1	1	1	3	1	0	0
13	0	1	1	0	0	0	0	1	4	2	5	3	1	1	2	0	3	0	0	2	1	1	2	0
14	0	0	1	0	1	1	1	3	2	0	3	5	0	1	1	1	1	1	1	0	2	0	5	0
15	0	3	2	0	2	0	1	1	2	2	4	6	1	2	1	0	0	0	0	3	2	0	2	0
16	0	2	1	0	2	1	0	0	3	2	3	2	0	1	0	0	0	0	0	0	3	0	4	0
17	0	3	0	0	0	1	0	0	3	2	3	3	0	2	1	0	2	1	0	2	1	2	1	0
18	0	0	1	0	1	3	0	0	5	1	3	3	0	1	0	1	1	0	0	1	1	0	3	0
19	0	1	1	0	1	0	0	0	3	3	3	2	1	2	2	3	1	0	1	2	4	0	3	0
20	0	1	1	0	1	0	0	2	2	0	1	0	2	0	1	1	0	0	0	1	2	0	4	0
21	0	0	0	0	1	0	0	1	1	0	3	0	0	0	1	0	1	1	0	4	0	0	0	0
22	0	1	0	0	0	1	1	2	1	1	1	4	0	1	0	0	1	0	1	0	1	0	4	0
23	0	0	0	0	0	0	0	2	0	1	4	1	0	0	0	0	0	0	1	2	0	0	0	0
24	0	0	0	0	1	1	0	1	2	0	2	6	0	0	0	0	2	0	0	0	1	2	1	0
25	0	0	0	1	0	0	0	0	0	1	1	3	1	1	0	0	0	0	0	1	0	0	4	0
26	0	0	0	0	0	0	0	0	1	1	2	6	0	0	1	0	0	0	0	0	1	0	2	0
27	0	0	0	0	0	1	0	0	0	0	1	4	0	0	0	0	0	0	1	0	1	0	2	0
28	0	0	0	1	1	0	0	0	1	0	1	2	0	0	0	0	0	0	0	0	0	1	2	0
29	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1
30	0	0	0	1	1	0	0	0	1	0	3	0	0	0	0	0	0	0	0	0	0	0	1	0
31							[0	1	2	3	10	10][b]							[1	8	5	2	10	0][b]

NOTE: $N = 1731$.

[a] Number of employees: 1: 1–4; 2: 5–29; 3: 30–299; 4: 300–999; 5: 1,000 and over; 6: Government.

[b] The number of cases censored because they have reached the upper limit of the 30-year duration period.

distribution of events during the age of employment. The first method generated a substantially larger log-likelihood than did the second method for each tested model, indicating a greater probability of predicting the observed data according to BIC. For this reason, the following analyses use $Y = \log[(\text{age difference}) + 1]$ for the observation of $\log(T)$.

4. ANALYSES

First, 12 models are tested. Then an additional model (which turned out to be the best model) is also tested. The first 12 models represent combinations of (a) two parametric types (log-normal and generalized Gamma), (b) three types of surviving fractions (zero surviving fraction, a positive constant surviving fraction, and heterogeneous surviving fractions that depend on covariates), and (c) two types of covariate effects (main effects of firm size and occupation, and both main and interaction effects of firm size and occupation). The results for the log-normal model among special cases of the extended family of the generalized Gamma model are presented, because this model is equivalent to the linear regression model of censored log-duration data with a normally distributed error term. In the following, the extended family of the generalized Gamma model is referred to simply as the generalized Gamma model. Table 2 presents the likelihood ratio chi-squared statistic, degrees of freedom, and BIC for the 13 models.

The results show that, according to BIC, the generalized Gamma model that hypothesizes the main effects of covariates on both the timing and the surviving fraction (i.e., Model I-3G) is the best model among the first 12 models). It is worthwhile to note that this new model—which introduces heterogeneous surviving fractions—greatly improves the fit of the corresponding accelerated failure-time model (Model I-3G compared with Model I-1G) in terms of both likelihood ratio test and BIC.

The interaction effects between firm size and occupation are significant according to the likelihood ratio test at the .01 level (Model II-3G versus Model I-3G), but make the BIC value smaller. Model II-3G, which introduces full interaction effects of occupation and firm size, makes the BIC statistic smaller in part because it uses 10 parameters to characterize the interaction effects. Results from Model

Table 2. Comparison of Models

	L^2 [a]	np [b]	BIC
I. Models with only main effects of firm size and occupation			
1. Models with a zero surviving fraction			
L. Log-normal	428.92	7	376.72
G. Generalized Gamma	515.00	8	455.35
2. Models with a constant nonzero surviving fraction			
L. Log-normal	494.04	8	434.38
G. Generalized Gamma	518.18	9	451.07
3. Models with heterogeneous surviving fractions			
L. Log-normal	595.74	14	491.35
G. Generalized Gamma	620.94	15	509.09
II. Models with full interaction effects of firm size and occupation			
1. Models with zero surviving fraction			
L. Log-normal	445.60	12	356.12
G. Generalized Gamma	521.58	13	424.64
2. Models with a constant nonzero surviving fraction			
L. Log-normal	506.80	13	409.87
G. Generalized Gamma	526.10	14	421.71
3. Models with heterogeneous surviving fraction			
L. Log-normal	626.14	24	447.19
G. Generalized Gamma	645.86	25	459.45
III. Additional Models			
Model I-3G plus an interaction effect, [blue collar work] × [firms with 30–299 and 300–999 employees and government], on the surviving fraction	628.88	16	509.58

NOTE: N = 1,731.

[a] Likelihood-ratio chi-squared compared with the constant rate model regarding the likelihood function of $Y = \log(T)$. The log-likelihood of the constant rate model is −2,313.68.

[b] The number of parameters added to the constant rate model.

II-3G (not reported), however, suggest a specific pattern of interaction effects in the regression on surviving fraction. This pattern is tested by introducing one parameter to Model I-3G. As shown in Table 2, the 13th model thus formulated (Model III) improves the fit of Model I-3G in terms of both likelihood ratio test and BIC. It becomes the best of the 13 models according to both comparison criteria. The added single parameter explains about one-third of the chi-squared difference between Models I-3G and II-3G.

The absolute fit of Model III with the data cannot be easily tested. But a fair extent of agreement exists between the survivor function estimated from Model III and the survivor function estimated nonparametrically by the Kaplan–Meier product limit method. The former estimate of the survivor function lies consistently within the 95% confidence interval of the Kaplan–Meier estimate for most covariates' values, although the Kaplan–Meier estimates for the survivor function are less smooth. Figure 1 presents a typical example of this situation for the category of white collar workers in firms with 1,000 or more employees. Figure 1 presents the Kaplan–Meier estimate of the survivor function, the upper and lower 95% confidence limits, and survivor functions estimated by Models III and I-1G. The graph shows that the estimate from the generalized Gamma model without the surviving fraction (i.e., Model I-1G) falls outside the confidence interval of the Kaplan–Meier estimate for a significant range of employment duration, but the estimate from Model III falls consistently within the interval.

Table 3 presents parameter estimates from Models I-1G, I-3G, and III and, for comparison, the results from the proportional hazards models based on Cox's partial likelihood method.

The results from Model III (the best-fitting model) and Model I-3G show that both firm size and occupation affect the surviving fraction much more strongly than they affect acceleration/deceleration in the timing of job separations. In fact, the occupation effect on acceleration/deceleration of the timing of job separations is statistically insignificant. On the other hand, the results from these two models indicate that blue collar workers have a significantly larger lifetime probability of an interfirm job separation than do white collar workers. The interaction effect of Model III further indicates that this tendency is especially strong for workers employed in government, in private firms with 30–299 employees, and in private firms with 300–999 employees.

In contrast, the results from both the generalized Gamma model without the surviving fraction (Model I-1G) and the Cox model are less informative, even though they both consistently reveal that smaller firms and blue collar occupations are associated with shorter duration of employment. Parameter estimates in the Cox model change signs, because positive effects on hazard rates imply shorter duration. The Cox model does not separate the effects of covariates on the surviving fraction from those on acceleration/deceleration. Model I-1G assumes a zero limiting survival probability and therefore erroneously attributes all the effects of covariates to differences in the acceleration/deceleration of the timing of job separations.

Table 4 presents results from Model III regarding differences in the timing and the surviving fraction among firms of different sizes. The significance test for differences between firm-size effects is based on the variance–covariance matrix of parameter estimates, except for differences that include minus infinity as a parameter estimate. For the latter cases, a one-degree-of-freedom likelihood ratio test is used by comparing Model III against the model that imposes each par-

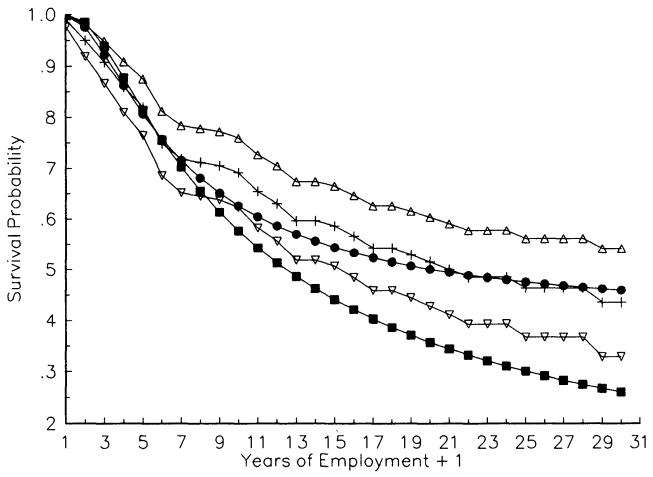

Figure 1. Estimated Survivor Functions. White Collar, Firms with 1,000+ Employees. + = Kaplan–Meier; ▽ = Lower 95% C.L.; △ = Upper 95% C.L.; ● = Model III; ■ = Model I-1G.

Table 3. Parameter Estimates From Selected Models

	Model I-1G		Model I-3G		Model III		Cox's model	
	b	*b/s.e.*	*b*	*b/s.e.*	*b*	*b/s.e.*	*b*	*b/s.e.*
Shape parameter (λ)[a]	−1.063		−.625		−.631		(N.A.)	
Log (scale parameter)	.023	(1.0)	−.172	(−5.7)	−.167	(−5.3)	(N.A.)	
1. Parameters for covariates on log (duration)								
A. Intercept	2.122	(34.4)	1.815	(24.5)	1.808	(24.4)	(N.A.)	
B. Firm size (versus 1,000+)								
1–4	−.548	(−5.6)	−.240	(−2.6)	−.236	(−2.5)	.868	(7.6)
5–29	−.409	(−5.5)	−.149	(−1.9)	−.150	(−1.9)	.687	(7.4)
30–299	−.347	(−4.6)	−.174	(−2.1)	−.143	(−1.7)	.486	(5.5)
300–999	−.381	(−4.1)	−.257	(−2.6)	−.251	(−2.5)	.479	(4.1)
government	−.210	(−2.4)	−.176	(−1.7)	−.177	(−1.8)	.281	(2.4)
C. Blue collar (versus white collar)	−.220	(−4.2)	−.080	(−1.5)	−.077	(−1.5)	.351	(5.4)
2. Parameters for covariates on surviving fractions								
A. Intercept	—		−.154	(−.8)	−.379	(−1.8)	(N.A.)	
B. Firm size (versus 1,000+)								
1–4	—		[−∞][b]	(−.0[c])	[−∞][b]	(−.0[c])	(N.A.)	
5–29	—		−2.466	(−3.6)	−2.362	(−3.8)	(N.A.)	
30–299	—		−1.015	(−3.5)	−.753	(−2.2)	(N.A.)	
300–399	—		−.667	(−2.1)	−.183	(−.5)	(N.A.)	
government	—		−.418	(−1.7)	−.111	(−.4)	(N.A.)	
C. Blue collar (versus white collar)	—		−1.120	(−4.8)	−.597	(−2.1)	(N.A.)	
D. Interaction: [blue collar]*[30–999 + government]	—		—		−1.425	(−2.2)	(N.A.)	

NOTE: Coeff./s.e. in parentheses.
[a] The standard error of the shape parameter is not estimated. See the method section in the text.
[b] The coefficient asymptotically goes to −∞.
[c] The ratio, coeff./s.e., asymptotically goes to 0.

ticular hypothesized equality. For white collar workers, the hypotheses tested by the likelihood ratio test include hypotheses $\alpha_1 = \alpha_i$ for $i = 2$, $i = 3$, $i = 4$ and $i = 5$ and hypothesis $\alpha_1 = 0$. Some hypotheses for blue collar workers are identical to those for white collar workers, and others include hypotheses $\alpha_1 = \alpha_i + \alpha_7$ for $i = 3$, $i = 4$, and $i = 5$, where α_7 represents the parameter for the interaction effect described in 2D of Table 3.

The results in Table 4 show that, with respect to timing, significant differences exist only between the largest private firms with 1,000 or more employees and all other firms; the latter has an accelerated timing of interfirm job separations compared to the former. This finding is consistent with hypothesis H3, which posits an accelerated timing of job separation for government workers compared to that for employees of large private firms. It partially supports hypothesis H2, which posits an accelerated timing of job separations for smaller firms.

As for differences in the surviving fraction, the results in Table 4 indicate that among white collar workers, firms of different sizes are clustered into three groups: (1) the group with the highest surviving fraction, which includes government and large private firms with 300–999 or 1,000 or more employees; (2) the group with a middle-level surviving fraction, which includes medium-sized firms with 30–299 employees; and (3) the group with the lowest surviving fraction, which includes small firms with 1–4 or 5–29 employees. This finding supports hypothesis H1, which posits a positive effect of firm size on the surviving fraction.

Among blue collar workers, significant differences in surviving fraction exist only between the largest private firms and other firms, with the former having a significantly larger surviving fraction. Hence the results for blue collar workers only partially support hypothesis H1. Table 5 presents the expected values of surviving fractions based on Model III, which clearly indicate three levels of surviving fractions for white collar workers and two levels for blue collar workers. For each group of workers, the lowest level of the surviving fraction is not significantly different from 0, as shown in Table 4. Descriptions of higher levels of surviving fractions are included in the summary of findings presented in the following section.

5. SUMMARY AND DISCUSSION

Using the new models yields the following findings for the analysis of interfirm job mobility in Japan:

1. Among white collar workers, the surviving fraction is about 36–41% for government and large private firms with 300 or more employees and about 24% for medium-sized firms. It follows that for white collar workers, a tendency for permanent employment exists in a wide range of relatively large firms.

Table 4. Pairwise Comparisons of Parameter Estimates Between Employees of Different Firm Sizes Based on Model III

	Comparisons of β parameters	Comparisons of α parameters	
	All employees	White collar employees	Blue collar employees
1–4 vs. 5–29	−.086	$[-\infty]^{a}$	$[-\infty]^{a}$
1–4 vs. 30–299	−.093	$[-\infty]^{***,a}$	$[-\infty]^{a}$
1–4 vs. 300–999	.015	$[-\infty]^{***,a}$	$[-\infty]^{a}$
1–4 vs. govern.	−.059	$[-\infty]^{***,a}$	$[-\infty]^{a}$
1–4 vs. 1,000+	−.236*	$[-\infty]^{***,a}$	$[-\infty]^{***,a}$
5–29 vs. 30–299	−.007	−1.589*	−.184
5–29 vs. 300–999	.101	−2.179**	−.754
5–29 vs. govern.	.026	−2.251***	−.826
5–29 vs. 1,000+	−.150†	−2.362***	−2.362***
30–299 vs. 300–999	.108	−.570	−.570
30–299 vs. govern.	.034	−.642†	−.642†
30–299 vs. 1000+	−.143†	−.753*	−2.178***
300–999 vs. govern.	.034	−.072	−.072
300–999 vs. 1,000+	−.251*	−.183	−1.608**
govern. vs. 1,000+	−.177†	−.111	−1.536*

NOTE: *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$

[a] The significance level is based on the one-degree-of-freedom likelihood ratio test that compares Model III with the model that imposes the parametric equality of relevant two effects. (See text for details.)

2. Among blue collar workers, the surviving fraction is about 27% for private firms with 1,000 or more employees and is not significantly larger than 0 for other firms. It follows that for blue collar workers, a tendency for permanent employment exists only for employees in the largest private firms and that this tendency is weaker than that for white collar workers in government and large private firms.

3. Controlling for differences in the surviving fraction, the effects of firm size on the acceleration/deceleration of timing are weak. Significant differences are found only between employees of the largest private firms and those of all other firms, with the latter group having a relatively accelerated timing of job separation. This may be due to a higher starting salary for employees in the largest private firms as compared to that of employees in other firms.

4. Controlling for firm size, white collar workers have a significantly larger surviving fraction than do blue collar workers, especially for government and private firms with 30–999 employees.

5. Controlling for differences in surviving fraction and firm size, there is no difference in the timing of job separations between white collar and blue collar workers.

Extending the accelerated failure time regression model by introducing a regression model of surviving fractions has advantages and disadvantages compared to proportional hazards and related hazard rate regression models. The major advantage is that we can distinguish the effects of covariates on the timing of occurrence and on the lifetime probability of occurrence. As demonstrated in the summary findings described in the previous paragraphs, the new model provides substantive insights into the duration data that cannot be gained—at least in as definite and parsimonious a way as is done here—by using other models.

The application presented in this article also shows that parameter estimates governing the surviving fraction are largely independent of those governing acceleration/deceleration. Table 6 presents correlations between the two sets of parameter estimates derived from the variance–covariance matrix of parameter estimates from Model I-3G, which uses the same set of covariates for α and β parameters. The negative correlation between the estimates of α_i and β_i in the diagonal cells of Table 6 shows that a larger (smaller) estimate for α_i will be compensated for to some extent by making the estimate of β_i smaller (larger). Even for the pair with the largest absolute value of correlation, however, one can explain less than 12% [$.118 = .343^2$] of the variability of the other.

Table 5. Estimated Surviving Fraction From Model III by Firm Size and Occupation

Firm size	White collar	Blue collar
1–4	.000	.000
5–29	.061	.034
30–299	.244	.041
300–999	.363	.070
1,000+	.406	.274
Government	.380	.075

The new model has several major disadvantages as compared to the proportional hazards model, however. First, like accelerated failure-time models, the modified models introduced in this article cannot use time-dependent covariates. Second, even though the generalized Gamma model is relatively flexible in capturing the shape of hazard and surviving functions, with the addition of the surviving fraction making it even more flexible, the modeling of time dependence is still much more flexible in the proportional hazards and related hazard rate regression models. Therefore, the chance of achieving a good fit of models with data is higher for the latter models. Finally, the models introduced in this article require a richer data set in terms of information about the survivor function near the end of the normal period of risk.

Given these disadvantages of the new models compared to the proportional hazards and related hazard rate regression models, the former models will not become a substitute for the latter models in most empirical situations. But when the key issues involve the distinction between timing and lifetime probability of occurrence and the effects of time-independent covariates on these outcomes, the models introduced in this article can be very useful.

Table 6. Correlations Between Two Sets of Parameter Estimates Based on Model I-3G[a]

	α_1	α_2	α_3	α_4	α_5	α_6
β_1	.000[b]	.015	−.103	−.130	−.145	.103
β_2	.000[b]	−.245	−.148	−.154	−.179	.055
β_3	.000[b]	−.049	−.343	−.143	−.173	.026
β_4	.000[b]	−.019	−.109	−.325	−.129	.100
β_5	.000[b]	−.006	−.101	−.108	−.292	.052
β_6	.000[b]	.043	.021	.063	−.040	−.245

[a] Subscripts indicate for both α and β: 1: 1–4 employees (versus 1,000+); 2: 5–29 employees; 3: 30–299 employees; 4: 300–999 employees; 5: government; 6: blue collar (versus white collar).

[b] The estimate asymptotically goes to 0.

On the other hand, the models introduced in this article can be considered improvements of accelerated failure-time models. Although representative accelerated failure-time models can be applied using SAS-LIFEREG, they carry a built-in assumption of a zero limiting survival probability. But because this assumption may not be realistic in the analysis of most life events, except for death, these models likely will have a poor fit with data (Wu 1990). On the other hand, for any given accelerated failure-time model, we always can test whether the additional regression model for the surviving fraction improves the fit. If it does, then we not only obtain a better fitting model but also gain deeper and more accurate insights into the data.

[*Received December 1989. Revised October 1991.*]

REFERENCES

Abegglen, J. (1958), *The Japanese Factory,* New York: Free Press.

Abramowitz, M., and Stegun, I. A. (1972), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables;* Washington DC: U.S. Government Printing Office.

Bennett, N. G., Bloom, D. E., and Craig, P. H. (1989), "The Divergence of Black and White Marriage Patterns," *American Journal of Sociology,* 95, 692–722.

Cheng, M. T. (1988), "Job Shift Patterns of the Japanese Male Labor Force," unpublished Ph.D. dissertation, University of California, Los Angeles, Department of Sociology.

Cole, R. E. (1972), "Permanent Employment in Japan: Facts and Fantasies," *Industrial and Labor Relations Review,* 26, 615–630.

——— (1973), "Functional Alternatives and Economic Development: An Empirical Example of Permanent Employment in Japan," *American Sociological Review,* 38, 424–438.

——— (1979), *Work, Mobility, and Participation: A Comparative Study of American and Japanese Industry.* Berkeley: University of California Press.

Cole, R. E., and Siegel, P. (1980), "Alternative Job-Changing Patterns: A Detroit-Yokohama Comparison," Paper presented at the Japan-United States Conference on Social Stratification and Mobility, Honolulu, Hawaii.

Diekmann, A. (1989), "Diffusion and Survival Models for the Process of Entry Into Marriage," *Journal of Mathematical Sociology,* 14, 31–44.

Diekmann, A., and Mitter, P. (1983), "The 'Sickle-Hypothesis': A Time-Dependent Poisson Model With Applications to Deviant Behavior and Occupational Mobility," *Journal of Mathematical Sociology,* 9, 85–101.

——— (1984), "A Comparison of the 'Sickle Function' With Alternative Stochastic Models of Divorce Rates," in *Stochastic Modeling of Social Processes,* eds. A. Diekmann and P. Mitter, New York: Academic Press, pp. 123–153.

Doeringer, P. B., and Piore, M. J. (1971), *Internal Labor Market and Manpower Analysis.* Lexington, MA: Lexington Books.

Dore, R. P. (1973), *British Factory—Japanese Factory: Origins of National Diversity in Industrial Relations,* Berkeley, CA: University of California Press.

Farewell, V. T. (1977), "A Model for a Binary Variable with Time-Censored Observations," *Biometrika,* 64, 43–46.

Grusky, D. B. (1983), "Industrialization and the Status Attainment Process: The Thesis of Industrialization Reconsidered," *American Sociological Review,* 48, 494–506.

Hashimoto, M. (1990). *The Japanese Labor Market in a Comparative Perspective with the United States.* Kalamazoo, MI: W. E. Upjohn Institute.

Heckman, J. J., and Walker, J. R. (1987), "Using Goodness of Fit and Other Criteria to Choose Among Competing Duration Models: A Case Study of Hutterite Data," in *Sociological Methodology 1987,* ed. C. C. Clogg, Washington DC: American Sociological Association, pp. 205–246.

Hernes, G. (1972), "The Process of Entry into Marriage," *American Sociological Review,* 37, 173–182.

Kalbfleisch, J. D., and Prentice, R. L. (1980), *The Statistical Analysis of Failure Time Data,* New York: John Wiley.

Koike, K. (1983a), "Internal Labor Markets: Workers in Large Firms," in *Contemporary Industrial Relations in Japan,* ed. T. Shirai, Madison, WI: University of Wisconsin Press, pp. 29–62.

——— (1983b), "Workers in Small Firms and Women in Industry," in *Contemporary Industrial Relations in Japan,* ed. T. Shirai, Madison, WI: University of Wisconsin Press, pp. 89–116.

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

Marini, M. M., and Hodsdon, P. (1981), "Effects of the Timing of Marriage and the First Birth on the Spacing of Subsequent Births," *Demography,* 18, 529–548.

Marsh, R., and Mannari, H. (1971), "Lifetime Commitment in Japan: Roles, Norms, and Values," *American Journal of Sociology,* 76, 795–812.

Miller, R. G. (1981), *Survival Analysis,* New York: John Wiley.

Millman, S. R., and Hendershot, G. E. (1980), "Early Fertility and Lifetime Fertility," *Family Planning Perspectives,* 12, 139–140, 145–149.

Prentice, R. L. (1974), "A Log-Gamma Model and Its Maximum Likelihood Estimation," *Biometrika,* 61, 539–544.

Raftery, A. E. (1986), "A Note on Bayes Factors for Log-Linear Contingency Table Models With Vague Prior Information," *Journal of the Royal Statistical Society,* Ser. B, 48, 249–250.

SAS Institute (1985), *SAS User's Guide: Statistics* (Version 5) Cary, NC: Author.

Schoen, R., and Kluegel, J. R. (1988), "Changing Pattern of First Marriage in the United States," *Demography,* 22, 265–279.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics,* 6, 461–464.

Spanier, G. B. (1983), "Married and Unmarried Cohabitation in the United States," *Journal of Marriage and the Family,* 45, 277–288.

Stacy, E. W. (1962), "A Generalization of the Gamma Distribution," *The Annals of Mathematical Statistics,* 33, 1187–1192.

Steven, R. (1983), *Classes in Contemporary Japan,* Cambridge, U.K.: Cambridge University Press.

Sumitani, M. (1974a), "Nihonteki Roshi Kankei no Saikento, Jo [Japanese Labor-Management Relations Revisited—Part I]," *Nihon Rodo Kyokai Zasshi [The Journal of the Japanese Labor Association],* 185, 2–16.

——— (1974b), "Nihonteki Roshi Kankei no Saikento, Ge [Japanese Labor-Management Relations Revisited—Part II]," *Nihon Rodo Kyokai Zasshi [The Journal of the Japanese Labor Association],* 187, 2–11.

Taira, K. (1962), "Characteristics of Japanese Labor Markets," *Economic Development and Cultural Change,* 10, 150–168.

Tan, H. W. (1980), "Human Capital and Technical Change: A Study of Wage Differentials in Japanese Manufacturing," unpublished Ph.D. dissertation, Yale University, Department of Economics.

Tominaga, K. (1964), "Hinon Shakai to Rodo Ido [Japanese Society and Labor Mobility]," in *Gijutsu Kakushin to Ningen no Mondai [Technological Development and Human Problems],* ed. K. Odaka, Tokyo, Japan: Diamond Press, pp. 277–306.

Tsuzuki, K. (1989), "Shokureki Kenkyu ni Okeru Ido to Jikan [Career Mobility as a Dynamic Process]," *Sociological Theory and Methods,* 4, 25–40.

Wu, L. L. (1990), "Simple Graphical Goodness-of-Fit Test for Hazard Rate Models," in *Event History Analysis in Life Course Research,* eds. K. U. Mayer and N. B. Tuma, Madison, WI: University of Wisconsin Press, pp. 109–117.

Yamaguchi, K. (1987), "Models for Comparing Mobility Tables: Toward Parsimony and Substance," *American Sociological Review,* 52, 482–494.

Yamaguchi, K. (1983), "The Public Sector: Civil Servants," in *Contemporary Industrial Relations in Japan,* ed. T. Shirai, Madison, WI: University of Wisconsin Press, pp. 295–312.