

DISCRETE DURATION MODEL HAVING AUTOREGRESSIVE RANDOM EFFECTS WITH APPLICATION TO JAPANESE DIFFUSION INDEX

Yasuhiro Omori*

This article describes a semiparametric estimation method for a discrete duration model with autoregressive random effects using Markov chain Monte Carlo techniques, and analyzes the duration of ten monthly economic times series which are components of the Japanese leading diffusion index. By introducing common time-dependent random effects to the individual duration, we capture a co-movement among durations that represents external macroeconomic factors. A dynamic modelling approach is employed assuming smoothness conditions on the baseline hazard function for long duration times with sparse observations.

Key words and phrases: Autoregressive random effects, discrete proportional hazards model, Markov chain Monte Carlo method, sequential probit model, state space model.

1. Introduction

Duration models have been widely applied to various economic and financial data recently. Among them are durations of unemployment (e.g. Kennan, 1985; Lancaster, 1990; McCall, 1996), bankruptcies (Lane *et al.*, 1986; Bandopadhyaya, 1994; Li, 1999), mortgage prepayments (Green and Shoven, 1986; Schwartz and Torous, 1989; Deng *et al.*, 2000), business cycles (Diebold and Rudebusch, 1990, 1992; Durland and McCurdy, 1994; Ghysels, 1997) and financial transaction data (Engle and Russell, 1998).

While most duration models are based on continuous time models that are designed for biomedical data, economic data are often observed at discrete time points such as weeks, months and quarters. For data observed at discrete time points, Prentice and Gloeckler (1978) first described a simple discrete semiparametric proportional hazards model and applied it to survival data on a large set of breast cancer data (11,442 observations) where no assumption was made for the baseline hazard function. However, as it is often the case with economic duration data, we may not have enough observations to conduct the nonparametric estimation of parameters, especially when there is a large number of time intervals. Therefore, we need to introduce some smoothness conditions on the baseline hazard function to construct a semiparametric discrete duration model for economic data (e.g. Gamerman, 1992; Fahrmeir, 1994; Omori and Johnson, 1998; Biller, 2000).

On the other hand, economic data are subject to some external macroe-

Received November 20, 2002. Revised February 3, 2003. Accepted February 21, 2003.

*Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

conomic factors represented by common (possibly unobserved) time-dependent variables. In multiple economic times series analysis, for example, a dynamic factor model is often considered to explain such a co-movement among time series (e.g. Engle and Watson, 1981; Stock and Watson, 1991; Forni and Reichlin, 1998). In the analysis of a panel of time series observations, Hjellvik and Tjøstheim (1999) considered individual effect as well as common effects over time to explain the intercorrelation across the panel and, further, showed ignoring the intercorrelation would lead to the inconsistency of estimators. This possibly unobserved common dynamic component can be modelled as a time-dependent random effect and needs to be considered in discrete duration models for economic data as well.

It has come to attention that ignoring such unobserved heterogeneity or random effects would mislead the conclusion as to the duration data. For continuous time models with time independent random effects, it is known that the hazard rates are underestimated and that survival times are stochastically larger than expected when the existence of random effects is ignored (Lee and Klein, 1988; Lancaster, 1990; Omori and Johnson, 1993, 1999). Similar results are shown to hold for the discrete time model with time independent random effects (Omori and Johnson, 1998).

Accordingly, estimation procedures have been developed to account for heterogeneity. For continuous time models with time independent random effects, Heckman and Singer (1984) first derived a consistent nonparametric maximum likelihood estimator for a proportional hazards model where a distribution of random effects is not specified. Honoré (1990) further proved asymptotic normality of a class of estimators based on order statistics for Weibull mixture model, and Ishwaran (1996) also proposed an estimator that converges at uniform rate for a finite dimension parameter space. For semiparametric models, Nielsen *et al.* (1992) assumed a gamma distribution for the random effect and proposed EM algorithm for maximum likelihood estimation (see also Andersen *et al.* (1993)). Murphy (1995) gave the asymptotic distribution of the estimator of the cumulative hazard rates and variance of the gamma random effect. Horowitz (1999) further established the asymptotic normality of the nonparametric estimator for the baseline hazard and the distribution for random effects assuming some smoothness conditions. Also, Sahu *et al.* (1997) gave an estimation method using Markov chain Monte Carlo method for parametric and semiparametric models. For discrete semiparametric models with time independent random effects, Han and Hausman (1990) and Sueyoshi (1992) extended Prentice and Gloeckler (1978) using an alternative formulation with competing risks. Omori and Johnson (1998) imposed smoothness conditions on the hazards rate function using a random walk process and gave an estimation method using Markov chain Monte Carlo, while Albert and Chib (2001) used a polynomial equation for the smoothness. We extend these models to consider an individual basic duration structure as well as unobserved common dynamics.

In Section 2, we introduce a dynamic modelling of discrete duration time

using the setup given in Albert and Chib (1993, 2001). Then we propose an estimation method using Markov chain Monte Carlo techniques in Section 3. In Section 4, we analyze the duration of ten monthly economic time series that are components of the Japanese leading diffusion index. Section 5 concludes the paper.

2. Model

We consider a discrete duration time Y_{it} for the i -th subject at calendar time t ($i = 1, \dots, n$, $t = 1, \dots, T$). The duration has a discrete distribution at duration time $j \in \{1, 2, \dots, J + 1\}$ given the parameter θ . The discrete duration time points correspond to certain intervals A_1, A_2, \dots, A_{J+1} where $A_j = (a_{j-1}, a_j]$, $j = 1, 2, \dots, J + 1$ with $a_0 = 0 < a_1 < \dots < a_{J+1} = \infty$. We denote the failure at duration time j by $Y_{it} = j$ for the i -th subject at calendar time t . Since the duration of the i -th subject may start at any calendar time, it is a function of both i and t (i.e. $j = j(i, t)$). We consider the discrete survival model with a hazard rate function

$$(2.1) \quad Pr(Y_{it} = j | Y_{it} \geq j, \alpha, \beta, \gamma_t^*) = F(\theta_{it}), \quad \theta_{it} = \alpha_j - x'_{it}\beta - z'_{it}\gamma_t,$$

where F is a cumulative distribution function, and $\alpha = (\alpha_1, \dots, \alpha_J)'$ represents a baseline hazards function (or cutpoints), and x_{it} , z_{it} are vectors of (calendar) time-varying covariates, and β is a regression parameter vector. The coefficient γ_t is a vector of (calendar) time-varying parameters, and $\gamma_t^* = \{\gamma_1, \dots, \gamma_t\}$.

Incorporating latent variables can derive this setup of the model. Suppose that the duration of the i -th subject starts at calendar time $t - j + 1$ and $Y_{it} \geq j$. Then define latent variables W_{it} and w_{it} such that

$$(2.2) \quad W_{it} = x'_{it}\beta + z'_{it}\gamma_t + e_{it}, \quad w_{it} = W_{it} - \alpha_j,$$

where $e_{it} \sim$ i.i.d. F . Conditional on $Y_{it} \geq j$, we observe $Y_{it} = j$ if $W_{it} \leq \alpha_j$ ($w_{it} \leq 0$). Thus we obtain

$$(2.3) \quad \begin{aligned} Pr(Y_{it} = j | Y_{it} \geq j, \alpha, \beta, \gamma_t^*) &= Pr(W_{it} \leq \alpha_j | \alpha, \beta, \gamma_t^*) \\ &= F(\alpha_j - x'_{it}\beta - z'_{it}\gamma_t). \end{aligned}$$

If we take a cumulative distribution function of an extreme value distribution for $F(\theta)$, we obtain the proportional hazards model. If we use that of a standard normal distribution, the sequential probit model is obtained. The γ_t 's introduce autoregressive random effect components to account for unobserved common dynamics subject to the macroeconomic condition. Also they represent missing covariates or a set of miscellaneous covariates that cannot be included without increasing the number of regression coefficients. We need to model them properly as we need to do for serially correlated disturbances in the regression analysis of time series data. We will assume that the γ_t 's are assumed to follow a stationary process.

To find a probability mass function for Y_{it} , define the latent variables

$$W_{i,t-j+k} = x'_{i,t-j+k}\beta + z'_{i,t-j+k}\gamma_{t-j+k} + e_{i,t-j+k}, \quad w_{i,t-j+k} = W_{i,t-j+k} - \alpha_k$$

for $k = 1, \dots, j$ where $e_{i,t-j+k} \sim$ i.i.d. F . Then we observe $Y_{it} = j$ when $W_{i,t-j+1} > \alpha_1, \dots, W_{i,t-1} > \alpha_{j-1}$ and $W_{it} \leq \alpha_j$ (equivalently, $w_{i,t-j+1} > 0, \dots, w_{i,t-1} > 0$ and $w_{it} \leq 0$). Thus the probability mass function of Y_{it} is given by

$$Pr(Y_{it} = j | \alpha, \beta, \gamma_t^*) = F(\theta_{it}) \prod_{s=1}^{j-1} \{1 - F(\theta_{i,t-s})\}, \quad 1 \leq j \leq J,$$

$$Pr(Y_{it} = J + 1 | \alpha, \beta, \gamma_t^*) = \prod_{s=1}^J \{1 - F(\theta_{i,t-s})\}.$$

When the observation Y_{it} is right censored, we have

$$Pr(Y_{it} = j, d_i = 0 | \alpha, \beta, \gamma_t^*) = \prod_{s=1}^{j-1} \{1 - F(\theta_{i,t-s})\},$$

where $d_i = 0$ if Y_{it} is right censored and $d_i = 1$ otherwise.

2.1. Modelling baseline hazard rates through the α_j 's

It is often that we do not have sufficient observations to estimate hazard rates nonparametrically for larger duration time. Thus, we need to impose smoothness conditions on the baseline hazard function by some prior process to obtain stable estimators. We assume that α_j follows a simple random walk as in Gamerman (1992), Fahrmeir (1994), Omori and Johnson (1998) and Biller (2000).

$$(2.4) \quad \alpha_{j+1} = \alpha_j + u_j, \quad u_j \sim \text{i.i.d. } N(0, \sigma_u^2), \quad j = 1, \dots, J-1,$$

and $\alpha_1 \sim N(0, \sigma_{u0}^2)$. For a second order random walk process for α_j , we have the state space representation of the process

$$\begin{pmatrix} \alpha_{j+1} \\ \alpha_j \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_j \\ \alpha_{j-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u_j, \quad u_j \sim \text{i.i.d. } N(0, \sigma_u^2),$$

for $j = 1, \dots, J-1$. In general, we may consider the process

$$(2.5) \quad \alpha_j = G'\delta_j, \quad \delta_{j+1} = H\delta_j + Gu_j, \quad u_j \sim N(0, \sigma_u^2),$$

where $G = (1, 0, \dots, 0)'$, $\delta_1 \sim N(0, \sigma_{u0}^2 I)$ and H is a known matrix. Instead of the above random walk prior process, we may assume that the α_j 's satisfy the polynomial model (e.g. Albert and Chib (2001)) such as

$$(2.6) \quad \alpha_j = \psi_0 + \psi_1(j-1) + \psi_2(j-1)^2 + u_j, \quad u_j \sim \text{i.i.d. } N(0, \sigma_u^2), \\ j = 1, \dots, J.$$

Both models are considered and compared using estimated marginal likelihood in Section 4.

2.2. Modelling the unobserved process through γ_t

We assume that the γ_t 's follow a stationary process. For example, we consider the following first order autoregressive process for the univariate γ_t ,

$$(2.7) \quad \begin{aligned} \gamma_{t+1} &= \phi\gamma_t + \epsilon_t, & \epsilon_t &\sim \text{i.i.d. } N(0, \sigma_\epsilon^2), \\ \gamma_1 &\sim N(0, \sigma_\epsilon^2/(1 - \phi^2)), & |\phi| &< 1 \end{aligned}$$

for $t = 1, \dots, T - 1$ where the ϕ and σ_ϵ^2 are unknown parameters. Higher order stationary process may be constructed similarly through the state space representation,

$$\gamma_{t+1} = \Phi\gamma_t + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d. } N(0, \Sigma),$$

for $t = 1, \dots, T - 1$ where Φ and γ_1 satisfy a stationary condition.

3. Markov chain Monte Carlo implementation

This section proposes an estimation method using a Markov chain Monte Carlo method.

3.1. Sequential probit model

In the sequential probit model, F is a cumulative distribution function of a standard normal random variable. We generate latent variables w_{it} 's so that we obtain samples from conditional posterior distributions using normal random variables.

Generation of latent variables, w_{it} 's. First we consider the conditional posterior distribution of w_{it} 's. Suppose that we observe $Y_{it} = j$. This implies $w_{i,t-j+k} > 0$ for $k = 1, \dots, j - 1$ and $w_{it} \leq 0$. Using the conditional prior $w_{i,t-j+k} | \alpha, \beta, \gamma \sim N(-\alpha_k + x'_{i,t-j+k}\beta + z'_{i,t-j+k}\gamma_{t-j+k}, 1)$ for $k = 1, \dots, j$, we obtain the conditional posterior distribution such that

$$(3.1) \quad \begin{aligned} &w_{i,t-j+k} | (y_{it} = j), \alpha, \beta, \gamma \\ &\sim \begin{cases} TN_{(0,\infty)}(-\alpha_k + x'_{i,t-j+k}\beta + z'_{i,t-j+k}\gamma_{t-j+k}, 1), & 1 \leq k \leq j - 1, \\ TN_{(-\infty,0]}(-\alpha_j + x'_{it}\beta + z'_{it}\gamma_t, 1), & k = j, \end{cases} \end{aligned}$$

where $TN_{(a,b)}(\mu, \sigma^2)$ denotes the normal distribution (μ, σ^2) truncated to the interval (a, b) .

Generation of a regression parameter β . We assume a multivariate normal prior for β , $\beta \sim N(b_0, B_0)$, which is conjugate. Then the conditional posterior distribution is also multivariate normal with mean $\hat{\beta}$ and covariance matrix B_1 such that

$$(3.2) \quad \hat{\beta} = B_1 \left\{ B_0^{-1}b_0 + \sum_{t=1}^T \sum_{i \in R_t} (W_{it} - z'_{it}\gamma_t)x_{it} \right\},$$

$$(3.3) \quad B_1 = \left(B_0^{-1} + \sum_{t=1}^T \sum_{i \in R_t} x_{it}x'_{it} \right)^{-1},$$

where R_t denotes a set of labels attached to the subjects at risk at calendar time t .

Generation of the γ_t 's. Let \hat{W}_t denote the $n_t \times 1$ vector of W_{it} 's for observations at risk at calendar time t where n_t is the number of subjects at risk at calendar time t . Also let $\hat{x}_t, \hat{z}_t, \hat{e}_t$ denote corresponding vectors of x_{it}, z_{it}, e_{it} . Then we consider a linear Gaussian system

$$(3.4) \quad \hat{W}_t = \hat{x}_t' \beta + \hat{z}_t' \gamma_t + \hat{e}_t, \quad \hat{e}_t \sim N(0, I_{n_t}),$$

$$(3.5) \quad \gamma_{t+1} = \Phi \gamma_t + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma),$$

for $t = 1, \dots, T$ where $\gamma_1 \sim N(0, \Sigma_0)$. By using the simulation smoother given by de Jong and Shephard (1995), we sample disturbances term and obtain samples of the γ_t 's.

Generation of Φ . We set our prior to ensure the stationarity of γ_t and sample from the conditional posterior density using accept/rejection algorithm. In the univariate case with $\Phi = \phi$, we may assume a beta prior distribution for $(\phi+1)/2$. For example, if we take a uniform prior on $(-1, 1)$ (Beta(1,1) prior for $(\phi+1)/2$) and assume a stationarity distribution for the initial value, $\gamma_1 \sim N(0, \sigma_\epsilon^2/(1-\phi^2))$, then the conditional posterior distribution is given by

$$(3.6) \quad \phi | \gamma, \sigma_\epsilon^2 \propto TN_{(-1,1)} \left(\frac{\sum_{t=2}^T \gamma_{t-1} \gamma_t}{\sum_{t=2}^{T-1} \gamma_t^2}, \frac{\sigma_\epsilon^2}{\sum_{t=2}^{T-1} \gamma_t^2} \right).$$

Generation of Σ . As a prior distribution for Σ^{-1} , we take Wishart distribution $W(\rho_0, R_0)$ given by

$$(3.7) \quad g(\Sigma^{-1}) \propto |\Sigma^{-1}|^{(\rho_0 - r_\alpha - 1)/2} \exp \left\{ -\frac{1}{2} \text{tr} R_0^{-1} \Sigma^{-1} \right\},$$

where r_α is a rank of Σ . Then the resulting conditional posterior is $W(\rho_0 + 1, \{R_0^{-1} + \sum_{t=1}^{T-1} (\gamma_{t+1} - \Phi \gamma_t)(\gamma_{t+1} - \Phi \gamma_t)'\}^{-1})$. When the γ_t 's are scalar, $\Sigma = \sigma_\epsilon^2$ and $\gamma_1 \sim N(0, \sigma_\epsilon^2/(1-\phi^2))$, we take Gamma(φ_0, φ_1) for the prior distribution of σ_ϵ^{-2} . The conditional posterior distribution is Gamma($\varphi_0 + 0.5 * (T-1), \{\varphi_1^{-1} + 0.5(1-\phi^2)\gamma_1^2 + 0.5 \sum_{t=1}^{T-1} (\gamma_{t+1} - \phi \gamma_t)^2\}^{-1}$).

Generation of the α_j 's. Similarly to the case of the γ_t 's, we generate α conditional on β, γ, w . Define an artificial variable \hat{w}_j such that

$$\hat{w}_j = -\frac{1}{m_j} \sum_{t=1}^T \sum_{i \in R_{jt}} (w_{it} - x'_{it} \beta - z'_{it} \gamma_t),$$

where $R_{jt} = \{i : Y_{it} \geq j\}$ and m_j is the number of such observations. Then we consider the following linear Gaussian system with $\alpha_j = G' \delta_j$,

$$(3.8) \quad \hat{w}_j = G' \delta_j + v_j, \quad v_j \sim N(0, m_j^{-1}),$$

$$(3.9) \quad \delta_{j+1} = H \delta_j + G u_j, \quad u_j \sim N(0, \sigma_u^2),$$

for $j = 1, \dots, J$ where $\delta_1 \sim N(0, \sigma_{u0}^2 I)$. By using the simulation smoother, we obtain the samples of α_j 's.

Generation of σ_u^2 . We take $\text{Gamma}(\xi_0, \xi_1)$ for the prior distribution of σ_u^{-2} and the conditional posterior distribution is $\text{Gamma}(\xi_0 + 0.5 * (J - 1), \{\xi_0^{-1} + 0.5 \sum_{j=1}^{J-1} \{G'(\delta_{j+1} - H\delta_j)\}^2\}^{-1})$. For a first order random walk process with $\alpha_j = \delta_j$, we have $\text{Gamma}(\xi_0 + 0.5 * (J - 1), \{\xi_1^{-1} + 0.5 \sum_{j=1}^{J-1} (\alpha_{j+1} - \alpha_j)^2\}^{-1})$.

Acceleration of the Gibbs sampler. We may add the following step using generalized Gibbs sampler by Liu and Sabatti (2000) to accelerate the convergence of the α_j 's and γ_t 's. For example, when the γ_t 's are assumed to be scalar and $z_{it} = 1$ for all i and t , generate $\delta \sim N(\mu_\delta, \sigma_\delta^2)$ where

$$(3.10) \quad \mu_\delta = -\sigma_\delta^2 \left\{ \frac{\alpha_1}{\text{var}(\alpha_1)} + \frac{(1 - \phi^2)\gamma_1 + (1 - \phi) \sum_{t=2}^T (\gamma_t - \phi\gamma_{t-1})}{\sigma_\gamma^2} \right\},$$

$$(3.11) \quad \sigma_\delta^2 = \left\{ \text{var}(\alpha_1)^{-1} + \frac{(1 - \phi)^2(T - 1) + (1 - \phi^2)}{\sigma_\epsilon^2} \right\}^{-1},$$

and set $\alpha_j + \delta \rightarrow \alpha_j$ and $\gamma_t + \delta \rightarrow \gamma_t$ for all j, t .

3.2. Extension to the duration model with non-normal F

This section focuses on the survival model where F is not a cumulative distribution function of a standard normal distribution. Generations of Φ, Σ, σ_u^2 can be done as in Section 3.1. For α, β, γ , we need the Metropolis-Hastings algorithm with a pseudo-dominating rejection algorithm to sample from non-Gaussian conditional posterior distribution as follows (see e.g. Tierney, 1994; Chib and Greenberg, 1995).

To sample from a continuous multivariate density, for example, $f(x_1, x_2, \dots, x_n)$ which is assumed to be everywhere positive, we simulate x using Gibbs sampling algorithm. Suppose that it is difficult to sample directly from $f(z|x_{\setminus i}^{(j)})$ where $x_i^{(j)}$ denote the j -th sample of x_i and $x_{\setminus i}^{(j)} = (x_1^{(j)}, \dots, x_{i-1}^{(j)}, x_{i+1}^{(j)}, \dots, x_n^{(j-1)})$. We apply the adaptive/rejection sampling method by using a density $h(z)$ which is easy to sample from and satisfies $f \leq ch(z)$ for all z with some constant c as follows: (Step 1) generate a candidate z from $h(z)$ and a uniform random variable u on $(0, 1)$ (Step 2) if $u \leq f(z|x_{\setminus i}^{(j)})/ch(z)$ return z , otherwise return to Step 1. In our case, it is hard to find a density $h(z)$ that dominates the true density $f(z|x_{\setminus i}^{(j)})$, and we apply an independence chain Metropolis-Hastings algorithm with pseudo-dominating rejection algorithm (Tierney 1994): First we sample z from a density proportional to $q(z, x_{\setminus i}^{(j)}) = \min\{f(z|x_{\setminus i}^{(j)}), ch(z)\}$ by applying the acceptance/rejection sampling above. Then we accept z with probability $\min\{f(z|x_{\setminus i}^{(j)})q(x_i^{(j-1)}, x_{\setminus i}^{(j)})/f(x_i^{(j-1)}|x_{\setminus i}^{(j)})q(z, x_{\setminus i}^{(j)}), 1\}$.

Generation of β . Let $h(\beta)$ denote a logarithm of a prior density of β and $l(\theta_{it}) = I(y_{it} = j) \log F(\theta_{it}) + I(y_{it} > j) \log(1 - F(\theta_{it}))$ given $y_{it} \geq j$. Since

the conditional posterior distribution of β is complicated, we approximate it by a normal distribution using a Taylor expansion of the conditional posterior density around the mode, $\hat{\beta}$. Since the log conditional posterior density for β given y and other parameters is (excluding constant terms)

$$\begin{aligned} h(\beta) + \sum_{i,t} l(\theta_{it}) \\ \approx (\beta - \hat{\beta})' \left\{ h_{\beta}(\hat{\beta}) + \sum_{i,t} x'_{it} l'(\hat{\theta}_{it}) \right\} \\ + \frac{1}{2} (\beta - \hat{\beta})' \left\{ h_{\beta\beta}(\hat{\beta}) + \sum_{i,t} x'_{it} x_{it} l''(\hat{\theta}_{it}) \right\} (\beta - \hat{\beta}) \end{aligned}$$

where $h_{\beta}(\beta) = \partial h(\beta) / \partial \beta$ and $h_{\beta\beta}(\beta) = \partial^2 h(\beta) / \partial \beta \partial \beta'$, $l'(\theta_{it}) = \partial l(\theta_{it}) / \partial \theta_{it}$ and $l''(\theta_{it}) = \partial^2 l(\theta_{it}) / \partial \theta_{it}^2$. For the discrete proportional hazards model, we have $F(\theta) = 1 - \exp\{-\exp(\theta)\}$ and

$$\begin{aligned} l'(\theta_{it}) &= \frac{I(y_{it} = j) \exp\{\theta_{it} - \exp(\theta_{it})\}}{1 - \exp\{-\exp(\theta_{it})\}} - I(y_{it} > j) \exp(\theta_{it}), \\ l''(\theta_{it}) &= \frac{I(y_{it} = j) \exp\{\theta_{it} - \exp(\theta_{it})\} [1 - \exp(\theta_{it}) - \exp\{-\exp(\theta_{it})\}]}{[1 - \exp\{-\exp(\theta_{it})\}]^2} \\ &\quad - I(y_{it} > j) \exp(\theta_{it}). \end{aligned}$$

Our proposal density is $N(\mu^*, \Sigma^*)$ with

$$(3.12) \quad \mu^* = \hat{\beta} + \Sigma^* \left\{ h_{\beta}(\hat{\beta}) + \sum_{i,t} x'_{it} l'(\hat{\theta}_{it}) \right\},$$

$$(3.13) \quad \Sigma^* = - \left\{ h_{\beta\beta}(\hat{\beta}) + \sum_{i,t} x'_{it} x_{it} l''(\hat{\theta}_{it}) \right\}^{-1},$$

and apply the Metropolis-Hastings pseudo-dominating rejection algorithm.

Generation of γ_t . To reduce the high correlation among samples, we employ a multi-move sampler which generates a block of γ_t 's at the same time as in Shephard and Pitt (1997) and Watanabe and Omori (2001). To perform this multi-move sampler, first we approximate a true conditional posterior density of $(\gamma_{s+1}, \dots, \gamma_{s+k})$ given other variables. Since the log $f(\epsilon_s, \dots, \epsilon_{s+k-1} | \gamma_s, \gamma_{s+k+1}, y)$ excluding constant terms is given by

$$\begin{aligned} & - \frac{1}{2} \sum_{t=s}^{s+k} \epsilon'_t \Sigma^{-1} \epsilon_t + \sum_{t=s+1}^{s+k} \sum_{i \in R_t} l(\theta_{it}) \\ & \approx - \frac{1}{2} \sum_{t=s}^{s+k-1} \epsilon'_t \Sigma^{-1} \epsilon_t + \sum_{t=s+1}^{s+k} \sum_{i \in R_t} l(\hat{\theta}_{it}) \end{aligned}$$

$$\begin{aligned}
& + \sum_{t=s+1}^{s+k} \sum_{i \in R_t} z'_t(\gamma_t - \hat{\gamma}_t) l'(\hat{\theta}_{it}) + \frac{\{z'_t(\gamma_t - \hat{\gamma}_t)\}^2}{2} l''(\hat{\theta}_{it}) \\
& - \frac{1}{2}(\gamma_{s+k+1} - \Phi \hat{\gamma}_{s+k})' \Sigma^{-1} (\gamma_{s+k+1} - \Phi \hat{\gamma}_{s+k}) \\
& + (\gamma_{s+k+1} - \Phi \hat{\gamma}_{s+k})' \Sigma^{-1} \Phi (\gamma_{s+k} - \hat{\gamma}_{s+k}) \\
& - \frac{1}{2}(\gamma_{s+k} - \hat{\gamma}_{s+k})' \Phi' \Sigma^{-1} \Phi (\gamma_{s+k} - \hat{\gamma}_{s+k}),
\end{aligned}$$

we define artificial variables, $\hat{y}_t = z'_t \hat{\gamma}_t - \sum_{i \in R_t} l'(\hat{\theta}_{it}) / \sum_{i \in R_t} l''(\hat{\theta}_{it})$ for $t = s + 1, \dots, s + k - 1$, and

$$\begin{aligned}
(3.14) \quad \hat{y}_{s+k} & = \left(\Phi' \Sigma^{-1} \Phi - z_{s+k} z'_{s+k} \sum_{i \in R_{s+k}} l''(\hat{\theta}_{s+k}) \right)^{-1} \\
& \times \left\{ z_{s+k} \sum_{i \in R_{s+k}} l'(\hat{\theta}_{i,s+k}) - z_{s+k} z'_{s+k} \hat{\gamma}_{s+k} \right. \\
& \quad \left. \times \sum_{i \in R_{i,s+k}} l''(\hat{\theta}_{i,s+k}) + \Phi' \Sigma^{-1} \gamma_{s+k+1} \right\}.
\end{aligned}$$

Then consider the following approximating linear Gaussian model to sample from the true conditional posterior density.

$$(3.15) \quad \hat{y}_t = z'_t \gamma_t + e_t, \quad e_t \sim N(0, \tau_t^2), \quad \gamma_{t+1} = \Phi \gamma_t + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma),$$

where $\tau_t^2 = - \left[\sum_{i \in R_t} l''(\hat{\theta}_{it}) \right]^{-1}$ for $t = s + 1, \dots, s + k - 1$ and

$$\begin{aligned}
(3.16) \quad \hat{y}_{s+k} & = \gamma_{s+k} + \epsilon_{s+k}, \\
\epsilon_{s+k} & \sim N \left(0, \left\{ \Phi' \Sigma^{-1} \Phi - z_{s+k} z'_{s+k} \sum_{i \in R_{s+k}} l''(\hat{\theta}_{s+k}) \right\}^{-1} \right).
\end{aligned}$$

We use a Kalman filter and the simulation smoother to sample the errors $(\epsilon_s, \dots, \epsilon_{s+k-1})$ and the γ_s 's, and then apply the Metropolis-Hastings algorithm using this approximate density.

Acceleration of the Gibbs sampler. As in Section 3.1, we add the step to accelerate the convergence. When the γ_t 's are assumed to be scalar and $z_{it} = 1$ for all i and t , generate $\delta \sim N(\mu_\delta, \sigma_\delta^2)$ and set $\alpha_j + \delta \rightarrow \alpha_j$ and $\gamma_t + \delta \rightarrow \gamma_t$ for all j, t .

4. Duration dependence structure of Japanese Diffusion Index

In this section, we apply our semiparametric discrete duration model to Japanese leading diffusion index. We analyze eleven economic time series adopted

Table 1. Ten monthly series in Japanese leading diffusion index.

(1)	Index of producer's final demand goods and inventory shipment ratio (average of values in 1995 = 100)*.
(2)	Index of raw materials inventory to consumption ratio (manufacturing, 1995=100)*.
(3)	New job offers except for new graduates (persons).
(4)	New orders for machinery (million yen).
(5)	Total floor area of building construction started ($1000m^2$).
(6)	Total floor area of new housing construction started ($1000m^2$).
(7)	Number of new passenger car registrations and reports.
(8)	Nikkei commodity price index (17items, percent compared with the corresponding month of the previous year).
(9)	Money supply (M2+CD, percent compared with the corresponding month of the previous year).
(10)	Index of investment climate (manufacturing, percent).

* The rise and fall of the series are inverted.

Table 2. Summary statistics of durations (rise/fall).

	1973–1979	1980–1989	1990–1999
Counts	109/111	180/180	184/184
Mean	3.76/3.60	3.66/3.01	3.14/3.46
Median	3/2	3/2	2/3
Max	20/25	26/15	19/29
Min	1/1	1/1	1/1

in Japanese leading diffusion index, consisting of one quarterly and ten monthly series (however, they were revised in 2001). The series are considered to precede business cycle turning points and the index is the proportion of these series which increased over the past three months. The index shows overall direction of movements across various sectors of the economy rather than the amplitude of business cycle fluctuations. The longer duration of rise (fall) in the time series is considered to suggest an economic expansion (contraction) of the corresponding sector. If the index continues to stay above (below) 50 percent, the business cycle is considered to be turning from contraction (expansion) to expansion (contraction). It is used to judge the current business cycle phase and forecast its turning points. Therefore we investigate the duration dependence structure of the times series comprised in the diffusion index.

Since monthly data are available earlier than quarterly data, we focus on ten monthly series adopted in the leading diffusion index below and analyze their durations of rises and falls. Table 1 shows these ten time series. Three periods are considered: (i) April 1973–December 1979 (ii) January 1980–December 1989 (iii) January 1990–December 1999, corresponding to the seventies, eighties and nineties since the currently adopted times series are applied to calculate the

indexes from April 1973. Further, for simplicity, we assume that the duration (of rise or fall) still continues when there is no change in times series due to rounding error (though it is counted as 0.5 in the calculation of Japanese diffusion index). Summary statistics of durations are given in Table 2.

We assume the dynamic model with

$$(4.1) \quad \theta_{it} = \alpha_j + \beta_1 X_i + \sum_{j=1}^9 \beta_{2j} D_{ij} + \gamma_t,$$

$$(4.2) \quad \gamma_{t+1} = \phi \gamma_t + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d. } N(0, \sigma_\epsilon^2), \quad t \geq 1,$$

$\gamma_1 \sim N(0, \sigma_\epsilon^2 / (1 - \phi^2))$ for the i -th observation at time t where

$$X_i = \begin{cases} 1 & \text{if } i\text{-th observation is rising,} \\ 0 & \text{otherwise,} \end{cases}$$

$$D_{ij} = \begin{cases} 1 & \text{if } i\text{-th observation is the } j\text{-th series of diffusion index,} \\ 0 & \text{otherwise.} \end{cases}$$

We use dummy variables to show the difference among ten series and take the index of investment climate as the reference series for convenience sake. The variable X_i is introduced to explain the asymmetric durations of rises and falls.

A stationary distribution is assumed for γ_1 . We take a normal prior $N(0, 1000)$ for $\beta_1, \beta_{21}, \dots, \beta_{29}$ and Gamma(0.01, 100) prior for σ_ϵ^{-2} (which implies a unit mean and a large variance, 100). A uniform prior on $(-1, 1)$ is taken for ϕ . For the α_j 's, the following three models are considered for comparison.

M_1 : a first order random walk process. $\alpha_j = \alpha_{j-1} + u_{j-1}$, $u_j \sim N(0, \sigma_u^2)$, for $j = 2, \dots, J$, with $\alpha_1 \sim N(0, 100)$ and $\sigma_u^{-2} \sim \text{Gamma}(0.01, 100)$.

M_2 : a quadratic function. $\alpha_j = \psi_0 + \psi_1(j-1) + \psi_2(j-1)^2$ for $j = 1, \dots, J$ with $(\psi_0, \psi_1, \psi_2)' \sim N(0, 1000I_3)$.

M_3 : a quadratic function with disturbances. $\alpha_j = \psi_0 + \psi_1(j-1) + \psi_2(j-1)^2 + u_j$, $u_j \sim N(0, \sigma_u^2)$ for $j = 1, \dots, J$ with $(\psi_0, \psi_1, \psi_2)' \sim N(0, 1000I_3)$.

Three models are estimated for both the sequential probit model and proportional hazards model. We also estimate these models without time-dependent random effects γ_t 's for comparison. The results were generated using Ox version 2.20 (Doornik, 1999).

The initial iterations (1000–4500) are discarded and the following 10000 iterations are recorded. The number of blocks for sampling γ_t is set equal to three or four so that each block contains approximately 30 parameters. The sample paths look stable and autocorrelations decay quickly. Their convergences to the target distributions are tested via a simple test statistic for comparing two means (see Geweke (1992)) where the estimated variances for the sample means are obtained by using a Parzen window with bandwidth 300.

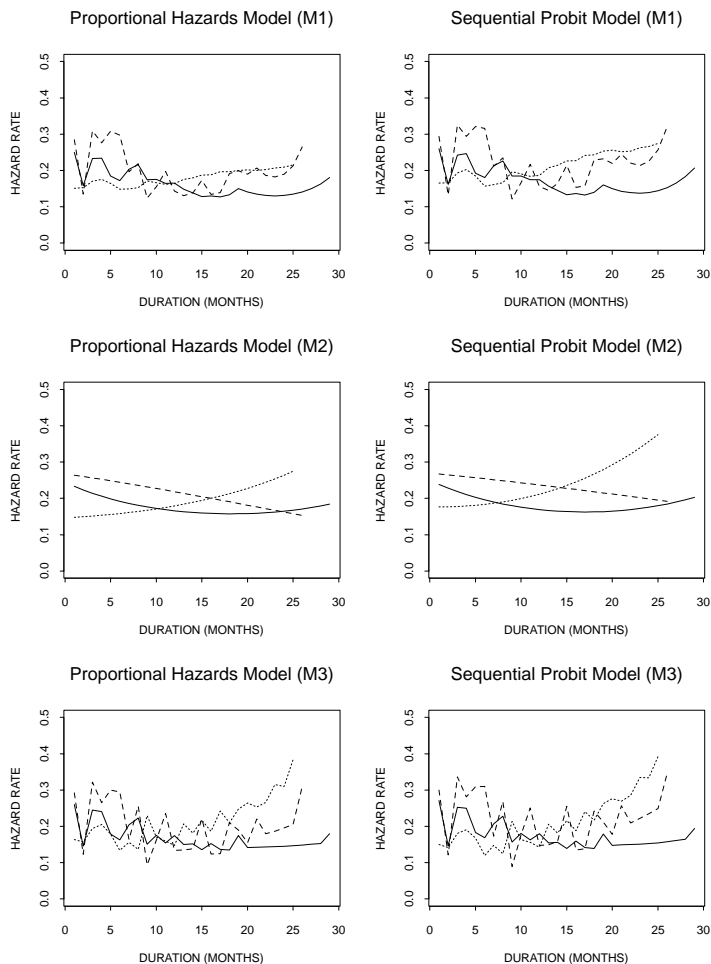


Figure 1. Estimated hazard rates. Dotted line: 1973–1979. Dashed line: 1980–1989. Solid line: 1990–1999.

Figure 1 shows the baseline hazard rate functions for three models with random effects. They look linearly increasing and stable in the seventies. However, as we shall see later, durations of ten times series are strongly influenced by the autoregressive random effects γ_t 's for the first two years (April 73–March 74). This suggests that there existed some external macroeconomic factors to explain a co-movement of durations in addition to the individual baseline hazard rate function. In the eighties, the hazard rates become higher (30%) for the first seven months (except the second month), and flat with low values for longer durations. The long economic boom in the late 80's resulted in the low hazard rates for longer durations. On the other hand, in the nineties, they become flat with low values (15%–20%), which correspond to the long economic contractions after the collapse of the bubble economy in Japan.

As shown in Figure 1, similar hazard rate functions are obtained for both sequential probit models and proportional hazards models, but the proportional

Table 3. Marginal log likelihood followed by standard error.

1973–1979	Sequential Probit Model		Proportional Hazards Model	
	With γ_t	Without γ_t	With γ_t	Without γ_t
M_1	-515.24 (0.25)	-522.33 (0.14)	-515.22 (0.20)	-521.52 (0.10)
M_2	-511.33 (0.05)	-518.64 (0.06)	-507.40 (0.06)*	-513.65 (0.05)
M_3	-527.79 (0.42)	-537.10 (0.66)	-522.73 (0.33)	-529.32 (0.30)

1980–1989	Sequential Probit Model		Proportional Hazards Model	
	With γ_t	Without γ_t	With γ_t	Without γ_t
M_1	-761.99 (0.10)	-764.04 (0.12)	-757.78 (0.07)*	-759.98 (0.06)
M_2	-782.21 (0.05)	-785.33 (0.05)	-777.49 (0.04)	-779.97 (0.03)
M_3	-776.41 (0.38)	-779.03 (0.37)	-769.15 (0.13)	-771.36 (0.13)

1990–1999	Sequential Probit Model		Proportional Hazards Model	
	With γ_t	Without γ_t	With γ_t	Without γ_t
M_1	-792.96 (0.21)	-792.95 (0.16)	-788.80 (0.10)	-787.50 (0.10)*
M_2	-797.29 (0.05)	-797.16 (0.04)	-792.62 (0.04)	-792.14 (0.04)
M_3	-811.66 (0.84)	-809.37 (0.54)	-802.92 (0.48)	-800.96 (0.49)

1973–1999 (TOTAL)	Sequential Probit Model		Proportional Hazards Model	
	With γ_t	Without γ_t	With γ_t	Without γ_t
M_1	-2070.20	-2079.32	-2061.80*	-2068.99
M_2	-2090.83	-2101.13	-2077.50	-2085.76
M_3	-2115.86	-2125.51	-2094.80	-2101.64

* Selected model.

hazards models tend to give smoother hazard rate functions than sequential probit models for M_1 and M_3 . This is probably because the standard normal distribution has tight tails and, correspondingly, sequential probit models are more sensitive to the magnitude of changes in the parameter values. A first order random walk prior for the baseline hazard does not smooth hazard rates very much in the eighties since there are large bumps in the short durations. A second order random walk prior may be preferred to smooth such bumps.

To compare six models ($M_1 - M_3$ with and without random effects), we calculated their marginal likelihoods. Chib (1995) and Chib and Jeliazkov (2001) are used to calculate the posterior ordinate from outputs of Gibbs sampling and M-H algorithm respectively, and, further, the accelerated Gaussian sampler by Danielsson and Richard (1993) is used to calculate the log likelihood for the model with time-dependent random effects. The calculated marginal likelihoods (in natural logarithm) are shown in Table 3. In the seventies, the proportional hazards model M_2 with random effects have the largest marginal likelihood and the model is selected. In fact, as shown in Figure 1, the hazard rates functions look linear or slightly quadratic in M_1 and M_3 . Also all models with random effects have much larger values of marginal likelihoods than those without random

Table 4. Estimated posterior means followed by standard deviations, and 95% credible intervals for $\beta_1, \phi, \sigma_\epsilon^2$. Proportional hazards model (M_1).

	1973–1979	1980–1989	1990–1999
β_1	0.159 (0.156) (−0.150, 0.458)	0.253 (0.114) (0.025, 0.476)	−0.044 (0.107) (−0.253, 0.165)
ϕ	0.807 (0.205) (0.240, 0.998)	0.538 (0.278) (−0.189, 0.919)	0.166 (0.385) (−0.607, 0.797)
σ_ϵ^2	0.097 (0.076) (0.014, 0.296)	0.062 (0.040) (0.012, 0.162)	0.037 (0.025) (0.006, 0.101)

Table 5. Estimated posterior means followed by standard deviations, and 95% credible intervals for $\beta_1, \phi, \sigma_\epsilon^2$. Sequential Probit Model (M_1).

	1973–1979	1980–1989	1990–1999
β_1	0.120 (0.106) (−0.089, 0.325)	0.172 (0.081) (0.015, 0.330)	−0.034 (0.078) (−0.185, 0.119)
ϕ	0.854 (0.122) (0.535, 0.997)	0.605 (0.276) (−0.149, 0.956)	0.207 (0.430) (−0.765, 0.823)
σ_ϵ^2	0.048 (0.031) (0.011, 0.127)	0.033 (0.020) (0.007, 0.080)	0.025 (0.017) (0.004, 0.066)

effects, suggesting with strong evidence that there existed some external macroeconomic factors represented by common unobserved time-dependent variables in the seventies. In the eighties, the proportional hazards model M_1 with random effects is selected. The baseline hazard rate functions have a large bump and no longer look like a function of polynomials as in Figure 1. The quadratic function models with disturbances (M_3) are not flexible enough to explain the irregular behaviour of the hazard rate function. The marginal likelihood values indicate that random effects are still significant, but that they are not strongly significant any more. In the nineties, the proportional hazards model M_1 without random effects has the largest marginal likelihood. The random effects are no longer significant, and the durations are explained by baseline hazard rate functions.

Throughout three periods, we found that the values of marginal likelihoods are very similar for both sequential probit models and proportional hazards models. Thus, estimated hazard rate functions in Figure 1 look similar for both models, correspondingly. However, the proportional hazards models have always larger marginal likelihoods than those of sequential probit models. Since the proportional hazards models M_1 with random effects have the largest sum of marginal likelihoods over three periods, we further discuss their estimation results in detail as below. Also, the sequential probit models M_1 with random effects are discussed for comparison.

The posterior means and their standard deviations of the parameters for $(\beta_1, \phi, \sigma_\epsilon^2)$ are given in Tables 4 and 5 for the proportional hazards models and

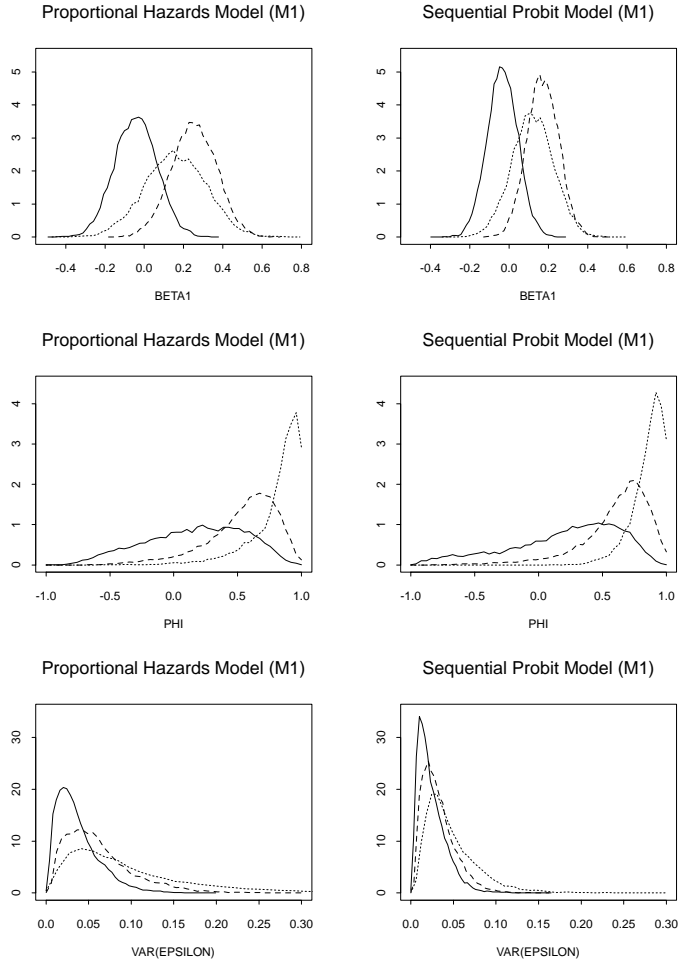


Figure 2. Estimated posterior densities for $\beta_1, \phi, \sigma_\epsilon^2$. Dotted line: 1973–1979. Dashed line: 1980–1989. Solid line: 1990–1999.

the sequential probit models respectively. Note that the coefficient β_1 shows the effect of rise on the duration. It is often pointed out that the rises of the series tend to continue longer than the falls and hence that β_1 is supposed to be positive. The estimates are expectedly positive in the seventies and eighties, but turned negative in the nineties (though 95% credible intervals include 0 in the seventies and nineties). This reflects the fact that Japan experienced the long recession in the nineties and the falls tend to continue longer than they used to be. The posterior densities of β_1 in Figure 2 also show this significant negative bias in the nineties. The densities for the sequential probit models are steeper than those for the proportional hazards models, but similar patterns are obtained.

The estimate of σ_ϵ^2 shows the significance of time-dependent random effects. The estimated posterior mean becomes smaller gradually from the seventies to the nineties. As we discussed above using marginal likelihoods, it indicates that the random effects are significant in the seventies and become less and less signif-

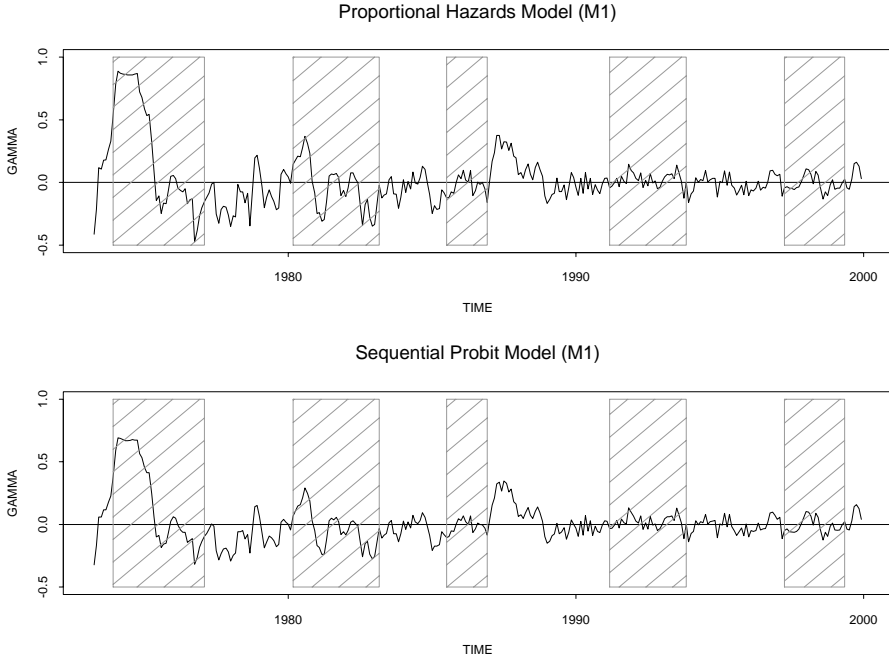


Figure 3. Estimated γ_t 's. Shaded areas correspond to contractions of the business cycle.

icant in the eighties and nineties. Similar results are obtained for the coefficient ϕ of the autoregressive process, which shows the first order autocorrelation for the time-dependent random effects. The correlation is high (close to one) and significantly positive in the seventies, but become smaller and less significant in the eighties and nineties. The posterior densities of $(\phi, \sigma_\epsilon^2)$ in Figure 2 show these gradual changes. Further, the estimated γ_t 's are shown in Figure 3. The random effects, or the unobserved common dynamic factor seems to have a great impact on the diffusion index in the seventies and early eighties. However, the factor is less influential in the nineties as its magnitude becomes smaller.

The posterior means and their standard deviations of the parameters for the β_{2j} 's are given in Tables 6 and 7. Their posterior densities are also shown in Figures 4 and 5. Throughout three periods, estimated posterior means are negative for the index of producer's final demand goods and inventory shipment ratio (β_{21}), Index of raw materials inventory to consumption ratio (β_{22}), new job offers (β_{23}), the new orders for machinery (β_{24}), total floor area of building construction started (β_{25}), total floor area of new housing construction started (β_{26}) and the number of new passenger car registrations and reports (β_{27}). This implies these seven series have shorter durations than the index of investment climate (the reference series). Among the seven series, the new orders for machinery has the highest hazard rates throughout three time periods. The $\beta_{21}, \beta_{25}, \beta_{26}, \beta_{27}$ are also significantly negative in at least one of the three periods. On the other hand, the percent change of M2+CD has the longest durations among all series in both the seventies and the nineties. Similar results are obtained for both the

Table 6. Estimated posterior means followed by standard deviations, and 95% credible intervals for β_{2j} 's. Proportional hazards model (M_1).

	1973–1979	1980–1989	1990–1999
β_{21}	−0.273 (0.365) (−0.986, 0.437)	−0.621 (0.237)* (−1.089, −0.161)	−0.403 (0.260) (−0.938, 0.100)
β_{22}	−0.318 (0.370) (−1.028, 0.415)	−0.086 (0.263) (−0.591, 0.435)	−0.389 (0.263) (−0.929, 0.120)
β_{23}	−0.236 (0.376) (−0.960, 0.510)	−0.163 (0.266) (−0.690, 0.344)	−0.305 (0.266) (−0.843, 0.205)
β_{24}	−1.497 (0.330)* (−2.147, −0.869)	−0.916 (0.232)* (−1.380, −0.477)	−0.715 (0.248)* (−1.228, −0.234)
β_{25}	−1.065 (0.334)* (−1.715, −0.421)	−0.256 (0.252) (−0.759, 0.237)	−0.714 (0.251)* (−1.233, −0.231)
β_{26}	−0.927 (0.335)* (−1.589, −0.290)	−0.270 (0.256) (−0.780, 0.226)	−0.374 (0.259) (−0.908, 0.114)
β_{27}	−0.825 (0.346)* (−1.506, −0.172)	−0.402 (0.244) (−0.897, 0.066)	−0.587 (0.252)* (−1.103, −0.113)
β_{28}	−0.493 (0.357) (−1.194, 0.193)	0.313 (0.289) (−0.244, 0.889)	−0.018 (0.280) (−0.585, 0.530)
β_{29}	0.441 (0.437) (−0.398, 1.331)	−0.214 (0.256) (−0.728, 0.292)	0.234 (0.301) (−0.363, 0.819)

* 95% credible interval does not include 0.

Table 7. Estimated posterior means followed by standard deviations, and 95% credible intervals for β_{2j} 's. Sequential Probit model (M_1).

	1973–1979	1980–1989	1990–1999
β_{21}	−0.118 (0.224) (−0.554, 0.323)	−0.451 (0.156)* (−0.755, −0.145)	−0.265 (0.162) (−0.577, 0.053)
β_{22}	−0.139 (0.225) (−0.574, 0.298)	−0.046 (0.165) (−0.369, 0.274)	−0.245 (0.162) (−0.556, 0.077)
β_{23}	−0.080 (0.226) (−0.523, 0.365)	−0.069 (0.166) (−0.389, 0.259)	−0.172 (0.166) (−0.493, 0.156)
β_{24}	−1.047 (0.211)* (−1.464, −0.634)	−0.653 (0.161)* (−0.965, −0.333)	−0.489 (0.159)* (−0.801, −0.174)
β_{25}	−0.696 (0.213)* (−1.117, −0.285)	−0.147 (0.163) (−0.459, 0.177)	−0.500 (0.161)* (−0.812, −0.180)
β_{26}	−0.597 (0.212)* (−1.009, −0.181)	−0.165 (0.160) (−0.480, 0.150)	−0.247 (0.162) (−0.558, 0.082)
β_{27}	−0.503 (0.216)* (−0.928, −0.078)	−0.279 (0.159) (−0.582, 0.037)	−0.387 (0.159)* (−0.693, −0.074)
β_{28}	−0.277 (0.220) (−0.711, 0.153)	0.196 (0.171) (−0.137, 0.533)	−0.008 (0.168) (−0.334, 0.332)
β_{29}	0.317 (0.240) (−0.145, 0.792)	−0.133 (0.163) (−0.451, 0.185)	0.160 (0.173) (−0.171, 0.509)

* 95% credible interval does not include 0.

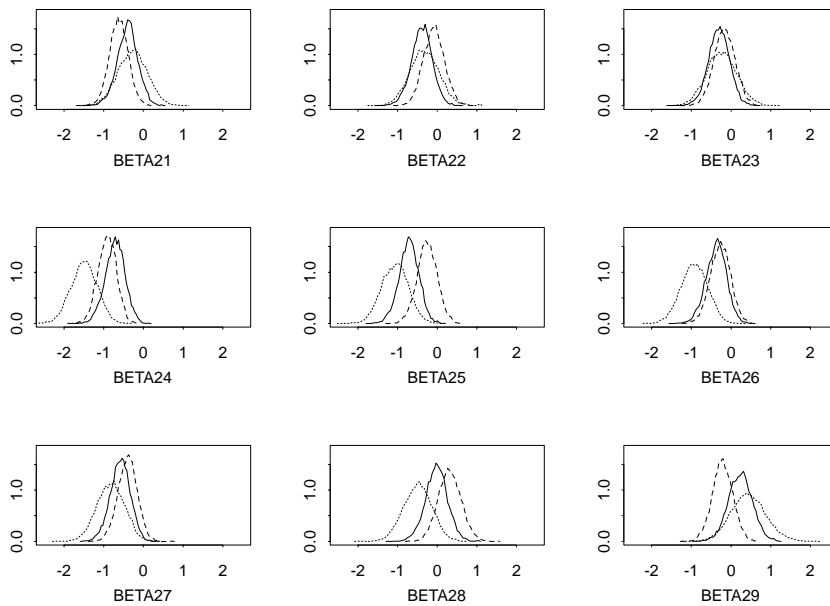


Figure 4. Estimated posterior densities for β_{2j} 's for proportional hazards model (M_1). Dotted line: 1973–1979. Dashed line: 1980–1989. Solid line: 1990–1999.

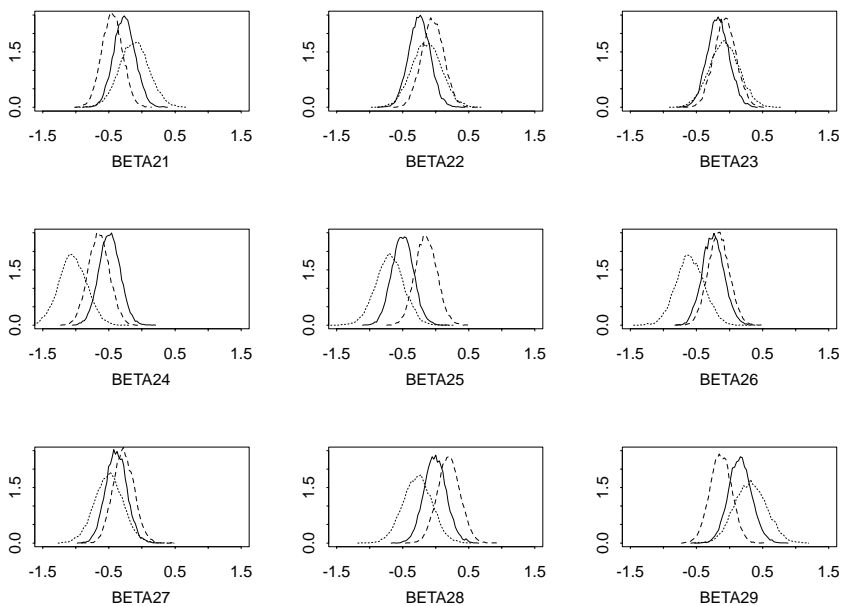


Figure 5. Estimated posterior densities for β_{2j} 's for sequential probit model (M_1). Dotted line: 1973–1979. Dashed line: 1980–1989. Solid line: 1990–1999.

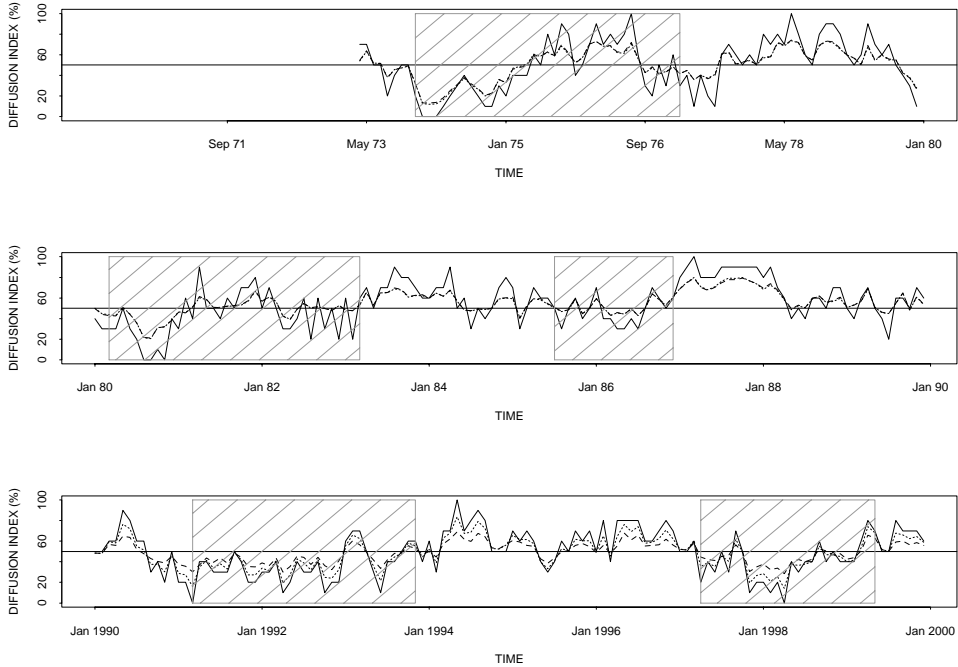


Figure 6. Diffusion index. Solid line: observed diffusion index. Shaded areas correspond to contractions of the business cycle. Dotted line: estimated posterior means of expected values (sequential probit model, M_1). Dashed line: estimated posterior means of expected values (proportional hazards model, M_1).

proportional hazards model and sequential probit models.

Finally, Figure 6 shows posterior means for the expected values of the diffusion index. Since the different series shows different movement, the observed diffusion index often fluctuates from month to month and it is difficult to judge the current phase of business cycle. To see expected direction of movements, we may use the posterior mean for the expected value of the index. It is an average of probabilities that each time series would rise in that month. The sequence of expected values is stable and smoother than the observed indexes, which would make it easier to find the turning point of business cycle. Similar estimates are obtained for both the proportional hazards model and sequential probit models in the seventies and eighties. In the nineties, however, the estimates given by the proportional hazards model are much smoother than those by the sequential probit model.

5. Conclusion

Using Markov chain Monte Carlo techniques, this article proposed a semi-parametric estimation method for a discrete duration model with autoregressive random effects, and the durations of rises and falls are analysed for ten monthly series of Japanese leading diffusion index. Using marginal likelihoods, we found the best model is the proportional hazards model with autoregressive random effects overall. However, we also note that the obtained results are very similar

for both sequential probit models and proportional hazards models. The baseline hazard function looks quadratic in the seventies but shows irregular patterns in the eighties and nineties. The assumptions of the quadratic function models (even with disturbances) are not flexible enough to explain a bump in the hazard rates, and the simple first order random walk process is found to be preferred.

As expected, the rises of the series tend to continue longer than the falls in the seventies and eighties. On the contrary, in the nineties, the falls tend to continue longer than the rises since Japan experienced the long recession during the period. The marginal likelihoods also indicate that common unobserved macroeconomic factors (or random effects) are significant in the seventies, but that they become less and less significant in the eighties and nineties. Thus the durations are explained by the individual baseline hazard rate functions without a co-movement recently. Among the time ten series, the new orders for machinery has the highest hazard rates throughout three time periods, while the percent change of M2+CD has the longest durations among all series in both the seventies and nineties.

Since economic data are often observed at discrete time points, we may not have enough observations for a nonparametric estimation of a baseline hazards function. Further, a co-movement among durations need to be considered to explain external macroeconomic factors. As is shown in this article, our dynamic modelling approach is useful and feasible when we assume smoothness conditions on the baseline hazard function and consider such factors by introducing common autoregressive random effects to the individual duration.

Acknowledgements

The author would like to thank anonymous referees for their helpful comments, which resulted in an improved manuscript.

REFERENCES

- Albert, J. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data, *J. Amer. Statist. Assoc.*, **88**, 669–679.
- Albert, J. and Chib, S. (2001). Sequential ordinal modeling with applications to survival data, *Biometrics*, **57**, 829–836.
- Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*, Springer-Verlag, New York.
- Bandopadhyaya, A. (1994). An estimation of the hazard rate of firms under chapter 11 protection, *Rev. Econ. Statist.*, **76**, 346–350.
- Billier, C. (2000). Discrete duration models combining dynamic and random effects, *Lifetime Data Analysis*, **6**, 375–390.
- Chib, S. (1995). Marginal likelihood from the Gibbs output, *J. Amer. Statist. Assoc.*, **90**, 1313–1321.
- Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm, *Amer. Statist.*, **49**, 327–335.
- Chib, S. and Jeliazkov, I. (2001). Marginal likelihood from the Metropolis-Hastings output, *J. Amer. Statist. Assoc.*, **96**, 270–281.
- Danielsson, J. and Richard, J.-F. (1993). Accelerated Gaussian importance sampler with application to dynamic latent variable models, *J. Appl. Econometrics*, **8**, S153–S173.

- de Jong, P. and Shephard, N. (1995). The simulation smoother for time series models, *Biometrika*, **82**, 339–350.
- Deng, Y., Quigley, J. M. and Van Order, R. (2000). Mortgage terminations, heterogeneity and the exercise of mortgage options, *Econometrica*, **68**, 275–307.
- Diebold, F. X. and Rudebusch, G. D. (1990). A nonparametric investigation of duration dependence in the American business cycle, *J. Pol. Econ.*, **98**, 596–616.
- Diebold, F. X. and Rudebusch, G. D. (1992). Have postwar economic fluctuations been stabilized?, *Amer. Econ. Rev.*, **82**, 993–1005.
- Doornik, J. A. (1999). *Object-oriented Matrix Programming Using Ox*, 3rd ed., Timberlake Consultants Press and Oxford, London.
- Durland, J. M. and McCurdy, T. H. (1994). Duration-dependent transitions in a markov model of U.S. GNP growth, *J. Bus. Econ. Statist.*, **12**, 279–288.
- Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data, *Econometrica*, **66**, 1127–1162.
- Engle, R. F. and Watson, M. W. (1981). A one-factor multivariate time series of metropolitan wage rates, *J. Amer. Statist. Assoc.*, **76**, 774–781.
- Fahrmeir, L. (1994). Dynamic modeling and penalized likelihood estimation for discrete time survival data, *Biometrika*, **81**, 317–330.
- Forni, M. and Reichlin, L. (1998). Let's get real: a factor analytical approach to disaggregated business cycle, *Rev. Econ. Stud.*, **65**, 453–473.
- Gamerman, D. (1992). A dynamic approach to statistical analysis of point processes, *Biometrika*, **79**, 39–50.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, *Bayesian Statistics*, (eds. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith), **4**, pp. 169–188, Oxford University Press, New York.
- Ghysels, E. (1997). On seasonality and business cycle durations: a nonparametric investigation, *J. Econometrics*, **79**, 269–290.
- Green, J. Y. and Shoven, J. B. (1986). The effects of interest rates on mortgage prepayments, *J. Money, Credit and Banking*, **18**, 41–59.
- Han, A. and Hausman, J. A. (1990). Specification and semiparametric estimation of duration models, *J. Appl. Econometrics*, **5**, 1–28.
- Heckman, J. and Singer, B. (1984). A method of minimizing the impact of distributional assumptions in econometric models of duration data, *Econometrica*, **52**, 271–320.
- Hjellvik, V. and Tjøstheim, D. (1999). Modeling panels of intercorrelated autoregressive time series, *Biometrika*, **86**, 573–590.
- Honoré, B. E. (1990). Simple estimation of a duration model with unobserved heterogeneity, *Econometrica*, **58**, 453–473.
- Horowitz, J. L. (1999). Semiparametric estimation of a proportional hazard model with unobserved heterogeneity, *Econometrica*, **67**, 1001–1028.
- Ishwaran, H. (1996). Uniform rates of estimation in the semiparametric Weibull mixture model, *Ann. of Statist.*, **24**, 1572–1585.
- Kennan, J. (1985). The duration of contract strikes in U.S. manufacturing, *J. Econometrics*, **28**, 5–28.
- Lancaster, T. (1990). *The Econometric Analysis of Transition Data*, Cambridge University Press, Cambridge.
- Lane, W. R., Looney, S. W. and Wansley, J. W. (1986). An Application of the Cox Proportional Hazards Model to Bank Failure, *J. Bank. Fin.*, **10**, 511–531.
- Lee, S. and Klein, J. P. (1988). Bivariate models with a random environmental factor, *IAPQR Trans.*, **12**(2), 1–18.
- Li, K. (1999). Bayesian analysis of duration models: an application to chapter 11 bankruptcy, *Econ. Lett.*, **63**, 305–312.
- Liu, J. S. and Sabatti, C. (2000). Generalised Gibbs sampler and multigrid Monte Carlo for Bayesian computation, *Biometrika*, **87**, 353–370.
- McCall, B. P. (1996). Unemployment insurance rules, joblessness, and part-time work, *Econometrica*, **64**, 647–682.

- Murphy, S. A. (1995). Asymptotic theory for the frailty model, *Ann. of Statist.*, **23**, 182–198.
- Nielsen, G. G., Gill, R. D., Andersen, P. K. and Sørensen, T. I. A. (1992). A counting process approach to maximum likelihood estimation in frailty models, *Scand. J. Statist.*, **19**, 25–43.
- Omori, Y. and Johnson, R. A. (1993). The influence of random effects on the unconditional hazard rate and survival functions, *Biometrika*, **80**, 910–914.
- Omori, Y. and Johnson, R. A. (1998). Multivariate discrete proportional hazards model with random effects, Research Paper Series, 8, Faculty of Economics, Tokyo Metropolitan University.
- Omori, Y. and Johnson, R. A. (1999). Some consequences of random effects in multivariate survival models, *Multivariate Analysis, Experimental Design and Survey Sampling*, (ed. S. Ghosh), pp. 301–347, Marcel Dekker, New York.
- Prentice, R. L. and Gloeckler, L. A. (1978). Regression analysis of grouped survival data with application to breast cancer data, *Biometrics*, **34**, 57–67.
- Sahu, S. K., Dey, D. K., Aslanidou, H. and Sinha, D. (1997). A Weibull regression model with gamma frailties for multivariate survival data, *Lifetime Data Analysis*, **3**, 123–137.
- Schwartz, E. S. and Torous, W. N. (1989). Prepayment and the valuation of mortgage-backed securities, *J. Fin.*, **44**, 375–392.
- Shephard, N. and Pitt, M. K. (1997). Likelihood analysis of non-Gaussian measurement time series, *Biometrika*, **84**, 653–667.
- Stock, J. H. and Watson, M. W. (1991). A probability model of the coincident and leading economic indicators, *Leading Economic Indicators: New Approaches and Forecasting Records*, (eds. K. Lahiri and G. H. Moore), pp. 63–89, Cambridge University Press, Cambridge.
- Sueyoshi, G. T. (1992). Semiparametric proportional hazards estimation of competing risk models with time-varying covariates, *J. Econometrics*, **51**, 25–58.
- Tierney, L. (1994). Markov chains for exploring posterior distributions, *Ann. of Statist.*, **22**, 1701–1728.
- Watanabe, T. and Omori, Y. (2001). Multi-move sampler for estimating non-Gaussian times series models: Comments on Shephard and Pitt (1997), Research Paper Series, Faculty of Economics, Tokyo Metropolitan University, No. 25.