Fixed Effects Maximum Likelihood Estimation of a Flexibly Parametric Proportional Hazard Model with an Application to Job Exits

Published in *Economic Letters* 2012

Audrey Light*
Department of Economics
Ohio State University
410 Arps Hall
1945 N. High Street
Columbus, OH 43210 USA
light.20@osu.edu

Yoshiaki Omori
Faculty of Economics
Yokohama National University
79-3 Tokiwadai, Hodogaya Ward
Yokohama 240-8501 JAPAN
omori@ynu.ac.jp

Abstract: We extend the fixed effects maximum likelihood estimator to a proportional hazard model with a flexibly parametric baseline hazard. We use the method to estimate a job duration model for young men, and show that failure to account for unobserved fixed effect causes negative schooling and union effects to be downward biased.

Key words: Proportional hazard model, fixed effects

JEL code: C41

*Corresponding author. +1(614)292-0493 (phone), +1(614)292-3906 (fax)
1. Introduction

We extend Yamaguchi’s (1986) fixed effects maximum likelihood (FEML) estimator to the proportional hazard model (PHM) with a flexibly parametric baseline hazard similar to the one proposed by Han and Hausman (1990) and Meyer (1990). After describing the estimation method—which requires that the baseline hazard and covariates are constant within each duration interval—we use Monte Carlo experiments to demonstrate that FEML performs well when these key assumptions are met, even with as few as five spells per individual. We then use the model to identify effects of union membership, schooling attainment, and other potentially endogenous variables on young men’s job durations.

This extension of FEML provides a tractable method for contending with endogenous covariates in a variety of hazard model applications. A more general solution to endogeneous covariates is to estimate multiple duration models simultaneously (van den Berg 2000). However, this approach cannot incorporate multiple endogenous regressors without a substantial increase in computational burden, and it requires that endogenous variables be modeled as duration variables. Simultaneous estimation of a PHM and discrete choice or linear models for endogenous regressors skirts the need for multiple duration models, but generally requires distributional assumptions for identification.

Fixed effects duration models overcome these problems as long as the endogenous covariates are only correlated with time-invariant, individual-specific unobservables. Three approaches—fixed effects conditional likelihood (FECL), fixed effects partial likelihood (FEPL) and fixed effects marginal likelihood (FEMGL)—use stratification on individuals to remove individual-specific parameters from the likelihood function, thereby avoiding inconsistency due to the presence of these “nuisance” parameters (Chamberlain 1985; Cox and Lewis 1966; Ridder and Tunali 1999; Yamaguchi 1986). As long as the observation period is long enough to minimize this inconsistency, however, the FEML approach dominates FECL, FEPL and FEMGL by (i) requiring less stringent assumptions about the underlying stochastic process; (ii) allowing all forms of time-varying covariates to be included; and (iii) identifying the baseline hazard, which allows conditional survival probabilities to be predicted. We note that a fourth approach to the PHM with fixed effects that entails nonparametric estimation of the baseline hazard (Horowitz and Lee 2004) might dominate FEML in cases with dependent censoring or a baseline hazard that even a flexible parameterization fails to fit adequately.
2. Fixed Effects Maximum Likelihood Estimator

Given data on \(j = 1, \ldots, J\) job durations for \(i = 1, \ldots, N\) individuals, we model the hazard rate for individual \(i\) on job \(j\) at tenure \(t\) as a standard PHM:

\[
\lambda(t|X_{ij}(t), v_i) = v_i \lambda_0(t) \exp[\beta X_{ij}(t)]
\]  

(2.1)

where \(v_i\) represents time-constant, individual unobservables, \(\lambda_0(t)\) represents the baseline hazard, \(X_{ij}(t)\) is a vector of time-varying covariates, and \(\beta\) is a vector of coefficients to be estimated. In contrast to the random effects PHM, we do not have to make a distributional assumption for \(v_i\) or assume \(v_i\) is uncorrelated with the covariates. Multiple spells for each individual and time-varying covariates aid identification.

Yamaguchi (1986) proposes a FEML estimator for the case where (i) precise failure times are assumed known; and (ii) the baseline hazard is assumed to be a constant, exponential function. He shows that the first order condition for the maximization of the resulting log-likelihood function can be solved analytically for \(v_i\) as a function of \(\beta\) and the baseline parameter, given the data; this function is then substituted into the log-likelihood function to obtain FEML estimators for the parameters.

We extend Yamaguchi’s FEML approach as follows. Following Han and Hausman (1990), Kiefer (1988) and Prentice and Gloeckler (1978), we divide the time axis into unit intervals \((0, 1], \ldots, (t-1, t]\), although any uniform spell length can be used in estimation. We assume all covariates are constant within each interval and, in contrast to Yamaguchi (1986), we assume that the baseline hazard is constant within each interval rather than constant for all \(t\).

The log-likelihood function is

\[
\ln L = \sum_{i=1}^{N} \left\{ \sum_{j=1}^{J_i} \left[ D_{ij} \ln[v_i \exp(\beta X_{ij}(T_{ij}') + \delta(T_{ij}'))] - v_i \exp[\beta X_{ij}(T_{ij}')] + \ln[T_{ij} - (T_{ij}' - 1)] 
+ \delta(T_{ij}') + \sum_{t=1}^{T_{ij}'-1} [-v_i \exp(\beta X_{ij}(t) + \delta(t))]] \right\}
\]  

(2.2)

where \(\delta(t) = \ln \int_{t-1}^{t} \lambda_0(u)du\) is the log of the baseline hazard integrated over the interval, \(D_{ij}\) is a dummy variable indicating a failure, \(T_{ij}\) is the “exact” duration, and \(T_{ij}'\) is the grouped duration such that \(T_{ij}' - 1 < T_{ij} < T_{ij}'\).

Differentiating (2.2) with respect to \(v_i\) and setting the result to zero can be solved individually.
for each $i$ due to the interval-constant baseline hazard assumption:

$$
\hat{v}_i = \frac{\sum_{j=1}^{I_i} D_{ij}}{\sum_{j=1}^{I_i} \exp[\beta X_{ij}(T'_{ij}) + \ln[T_{ij} - (T'_{ij} - 1)] + \delta(T'_{ij})] + \sum_{t=1}^{T'_{ij}-1} \exp[\beta X_{ij}(t) + \delta(t)]}.
$$

(2.3)

We could treat $\delta(t)$ as free parameters for each interval, but to reduce the number of parameters we express it as a flexible polynomial $\sum_{p=0}^{5} \alpha_p t^p$. FEML estimates for $\beta$ and $\alpha_p$ are obtained by replacing $\nu_i$ in (2.2) with $\hat{\nu}_i$ and maximizing the resulting function with respect to the parameters.

As discussed in Yamaguchi (1986), the FEML estimator has two limitations. First, it does not identify coefficients for covariates that are time-constant within and across jobs. Second, the likelihood function depends on $\hat{\nu}_i$, which are incidental parameters, so estimates are inconsistent. As demonstrated below, neither shortcoming appears to be significant in our application.

### 3. Monte Carlo Results

We generate data using model (2.1) with one endogenous covariate $X_{1ij}(t) = \mu_i + \omega_{ij}(t)$ and one exogenous covariate $X_{2ij}(t)$. We assume $X_{2ij}(t)$ and $\omega_{ij}(t)$ are independently and identically distributed as standard normal, and that $\ln\nu_i$ and $\mu_i$ are bivariate normal with variances equal to one and covariance equal to -0.5.\(^1\) We set $\beta_1 = 1$ and $\beta_2 = -1$, and make three alternative assumptions about the log of the baseline hazard integrated over the interval: (1) exponential ($\delta(t) = \gamma t$) with $\gamma = 1.4$; (2) Weibull ($\delta(t) = \gamma t^\alpha$) with $\alpha = 0.5, \gamma = 1.4$; and (3) Weibull with $\alpha = 1.5, \gamma = 0.86$. The exponential function is consistent with the assumption (section 2) that the baseline hazard is constant over each interval; this assumption is violated when we use the Weibull function, especially as the degree of duration dependence increases.

For each parameterization we generate 1,000 samples consisting of 100 individuals with, alternately, 5, 10 and 100 spells per person. To generate data for each spell for each individual, we first draw a pair of heterogeneity terms ($\ln\nu_i, \mu_i$) from the bivariate normal distribution. We then set $t=0$, draw $X_{2ij}$ and $\omega_{ij}$ from the standard normal distribution, and compute the conditional probability that the spell survives the first interval. We draw a random number from

---

\(^1\) In our application, innate ability, determination, and other unobserved factors that explain endogenous covariates such as union status and schooling attainment are likely to be negatively correlated with stick-to-itness and other unobservables that increase the job exit hazard.
a uniform distribution and judge the spell to fail if this number exceeds the computed conditional probability; otherwise, we increment $t$ and repeat the process, censoring all spells at $t=32$.

Finally, after judging a spell to fail within a given interval, we compute its completed duration in continuous time by applying the “inverse transform technique” on the cumulative distribution function for duration, conditional on completion in the given interval.

We use the FEML approach described in section 2 to estimate the coefficients $\beta_1$ and $\beta_2$ plus the six parameters in the function $\delta(t) = \sum_{p=0}^{5} \alpha_p t^p$ for each alternative model. In addition to estimating the FEML with alternative parameterizations of the baseline function, we also produce maximum likelihood (ML) estimates of a model that ignores unobserved heterogeneity ($\nu_i=1$). In table 1, we report estimates for $\hat{\beta}_1$ and $\hat{\beta}_2$ that represent the means and standard errors across samples of estimated coefficients.

Table 1 reveals that FEML performs well when we use the exponential baseline hazard. With 100 spells per person, we obtain estimates of 1.002 and -1.001, which are close to the true values of 1 and -1. The estimates remain reasonably close to their true values (while becoming less precise) when the number of spells per person falls to 10 and even to 5. In contrast, ML estimates are severely biased ($\hat{\beta}_1 = .515$ and $\hat{\beta}_2 = -.713$) when we ignore unobserved heterogeneity and use 10 spells per person. When we assume a Weibull hazard, the estimated coefficients are biased (in opposite directions) by 42% when $\alpha=.5$ and $\gamma=1.4$, and by 12% when $\alpha=1.5$ and $\gamma=.86$. The Weibull with $\alpha=.5$ and $\gamma=1.4$ exhibits extreme, negative duration over the first interval, so this model’s poor performance demonstrates that violating the condition that the baseline hazard be constant within each interval becomes more of a problem as duration dependence increases. In further experiments not reported in table 1, we find that the bias falls by only 10% when, using $\alpha=.5$ and $\gamma=1.4$, we decrease the interval length for $\delta(t)$ by half while holding constant the frequency of changes in time-varying covariates.

4. **Application to Job Duration Data**

We use the PHM with fixed effects—and, for comparison, a PHM with no unobserved heterogeneity and a PHM with $\ln \nu_i$ treated as a random draw from a standard normal distribution—to estimate a job exit hazard with 1979-2008 data from the National Longitudinal Survey of Youth (NLSY79). Among the 5,579 men in the NLSY79 who were not in the military oversample, we select a sample of 3,357 men who were born in 1960-64 and still in school when
the survey began tracking employment in January 1978 (thus ensuring non-left censored work histories), and who are observed working for at least two unique employers between initial school exit and the 2008 interview. Using 13-week intervals, we obtain a sample with 232,211 interval-specific observations for 32,630 jobs held by these men. The number of spells per person ranges from two to 40 with a mean of 9.7 and standard deviation of 6.2. Based on our Monte Carlo results we believe that bias due to the incidental parameters problem will be minimal.

In Table 2, we report means and standard deviations for select covariates included in the hazard model. Years of schooling and union status are key covariates insofar as they are time-varying, and are likely to be correlated with time-constant unobservables represented by \( v_i \) (innate ability, patience, etc.).\(^2\) Other time-varying covariates include years of work experience and its square, and the average hourly wage (in 1996 dollars). For the non-FE models, we include time-constant indicators of race/ethnicity and an age-adjusted, 1980 score from the Armed Forces Qualifications Test. Other covariates not shown in table 2 include industry and occupation dummies and controls for the local unemployment rate, rural location, and average unemployment insurance benefits in the state of residence.

Table 2 reveals that the estimated coefficient for union status is -0.270 when unobserved heterogeneity is ignored (column A), -0.283 when unobservables are assumed to be normally distributed random variables that are uncorrelated with covariates (column B), and -0.247 when the FEML approach is used to account for endogeneity. Similarly, the estimated schooling coefficient is -0.044 in column A, -0.047 in column B, and -0.037 in column C. For these and other coefficients, we do not always reject the null hypothesis of equality across specifications. Focusing on point estimates, however, we conclude that the failure to account for unobservables causes the estimated effects of union membership and schooling on job exits to be downward biased by 9% and 19%, respectively. The biases become even more severe—15% for union status and 27% for schooling—when we use the random effects model and assume all covariates are exogenous.

In contrast to the FECL, FEPL and FEMGL models, the FEML model identifies the baseline

\[^2\]Years of schooling is often thought of as time-invariant once the career begins, but 35% of our sample members increment their highest grade completed after their initial school exit, which we define as the start date of the first nonenrollment spell lasting 13 months or longer.
hazard and thus allows us to assess the magnitudes of the effects of endogenous covariates by computing conditional survival probabilities. Focusing on the schooling effect, we predict that a man with 16 years of schooling, three years of job tenure, and mean or modal values for all other covariates has a 62% chance of remaining with his current employer for another 13-week interval. This predicted survival probability falls to 50% if we use the random effects model (with exogenous schooling) and 44% if we ignore unobserved heterogeneity.
Acknowledgements: This research was funded by a grant to Light from the National Science Foundation and a grant to Omori from the Japan Society for the Promotion of Science (Grant-in-Aid for Scientific Research (C)20530193). We thank Shiyiing Zhang for excellent research assistance.

References


Table 1: Monte Carlo Estimates

<table>
<thead>
<tr>
<th>True Model</th>
<th>Spells per person</th>
<th>β₁</th>
<th>β₂</th>
<th>β₁</th>
<th>β₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential baseline, γ=1.4</td>
<td>5</td>
<td>1.042</td>
<td>-1.042</td>
<td>(0.066)</td>
<td>(0.063)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1.024</td>
<td>-1.021</td>
<td>0.515</td>
<td>-0.713</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>1.002</td>
<td>-1.001</td>
<td>(0.029)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Weibull baseline, α=.5, γ=1.4</td>
<td>10</td>
<td>1.427</td>
<td>-1.428</td>
<td>0.541</td>
<td>-0.752</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.881</td>
<td>-0.881</td>
<td>0.541</td>
<td>-0.752</td>
</tr>
<tr>
<td>True values</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

Note: Each simulation is based on 1,000 samples containing 100 individuals. Standard errors are in parentheses. The estimated model in the right-most columns (maximum likelihood with ν=1) ignores unobserved heterogeneity.
Table 2: Maximum Likelihood Estimates for Three Alternative Job Duration Hazard Models

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Mean (S.D.)</th>
<th>Coeff. (S.E.)</th>
<th>Coeff. (S.E.)</th>
<th>Coeff. (S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of school</td>
<td>12.73 (.230)</td>
<td>-.044 (.004)</td>
<td>-.047 (.009)</td>
<td>-.037 (.010)</td>
</tr>
<tr>
<td>1 if union member</td>
<td>.15 (.022)</td>
<td>-.270 (.027)</td>
<td>-.283 (.025)</td>
<td>-.247 (</td>
</tr>
<tr>
<td>Years of experience</td>
<td>9.93 (7.42)</td>
<td>-.075 (.003)</td>
<td>-.094 (.004)</td>
<td>-.069 (</td>
</tr>
<tr>
<td>Average hourly wage</td>
<td>13.63 (22.44)</td>
<td>-.014 (.001)</td>
<td>-.014 (.001)</td>
<td>-.013 (</td>
</tr>
<tr>
<td>1 if black</td>
<td>.28 (.018)</td>
<td>.025 (.051)</td>
<td>.072 (</td>
<td></td>
</tr>
<tr>
<td>1 if Hispanic</td>
<td>.18 (.020)</td>
<td>.010 (.051)</td>
<td>-.018 (</td>
<td></td>
</tr>
<tr>
<td>AFQT score</td>
<td>1.35 (27.53)</td>
<td>-.001 (.000)</td>
<td>-.000 (</td>
<td></td>
</tr>
</tbody>
</table>

Note: Model A ignores unobserved heterogeneity; B assumes unobserved factors are random effects that are uncorrelated with all covariates; C assumes fixed effects. All three models are estimated with 232,211 observations for 3,357 men using NLSY79 data. See text for other covariates included in each model.