

Econometric Estimation of Proportional Hazard Models

By: Gary Koop and Christopher J. Ruhm

Koop, Gary and [Christopher J. Ruhm](#). "Econometric Estimation of Proportional Hazard Models." *Journal of Economics & Business*, Vol. 45, No. 5, October 1993, 421-30.

Made available courtesy of ELSEVIER:

http://www.elsevier.com/wps/find/journaldescription.cws_home/505734/description#description

*****Note: Figures may be missing from this format of the document**

Abstract:

When analyzing duration data, covariates are typically assumed to modify hazard rates through the use of the proportional hazard model, in which the baseline hazard is multiplied by some function of the regressors and associated parameters, $g(\bullet)$, to determine individual hazard rates. The specification for $g(\bullet)$ which is in virtually universal use, $g(\bullet) = \exp(\bullet)$, may be overly restrictive. This paper considers one alternative—setting $g(\bullet)$ equal to the cumulative density function of the standard normal distribution. The relative performance of this functional form for $g(\bullet)$, compared to the exponential transformation, is shown to depend on the parametrization of the baseline hazard rate.

Article:

1. Introduction

Proportional hazard models have been increasingly used in the social and biological sciences to analyze duration data. In their basic form, such models may suffer from a number of overly restrictive assumptions. For instance, the failure to adequately control for heterogeneity can inaccurately indicate the occurrence of duration dependence, estimates may depend critically on the assumed parametric form of the baseline hazard, and the effects of covariates on hazard rates may vary over time. Previous researchers (e.g., Cox 1972, Heckman and Singer 1984, Katz and Meyer 1990) have devoted considerable effort to overcoming these problems.

Less attention has been paid to how individual hazard rates are modified by covariates. In the proportional hazard specification, the baseline hazard, h_0 , is multiplied by a function of the covariates $g(X\beta)$, to determine individual hazard rates. In principle, $g(\bullet)$ can take any form, subject to the restriction that it is always nonnegative. In practice, the specification in virtually universal use is $g(X\beta) \exp(X\beta)$, which implies that the natural logarithm of the hazard rate is linear in the covariates.¹

This paper evaluates the relative performance of a possible substitute form of $g(\bullet)$ —the cumulative density function of the standard normal distribution—using data on nonemployment durations from the January 1988 *Current Population Survey*. The relative performance of this functional form for $g(\bullet)$ depends on the parameterization of the baseline hazard. When h_0 is Weibull, the cumulative normal specification slightly outperforms the exponential transformation

¹ Less frequently estimated families of hazard models include log-linear and piecewise-constant hazards (see Lancaster (1990, chapter 3) for discussion).

typically estimated. Conversely, the latter functional form is marginally preferred when using the Cox semi-parametric hazard model.

I. An Alternative Specification of the Hazard Model

The hazard rate is the probability of leaving a specified state during the interval $t + \Delta t$, conditional on remaining in the state through time t :

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t}. \quad (1)$$

Individual covariates are typically assumed to influence the hazard function as specified by the proportional hazard model $h_i(t) = h_0(t) \cdot g(\mathbf{X}_i \beta)$, where $h_i(t)$ is the hazard rate for person i at time t , h_0 is the baseline hazard function, \mathbf{X}_i and β are vectors of covariates and regression parameters, and $g(\cdot)$ is an arbitrary nonnegative function. The exponential specification of $g(\cdot)$ generally used, $g(\mathbf{X}_i \beta) = \exp(\mathbf{X}_i \beta)$, yields the estimation model

$$h_i(t) = h_0(t) \cdot \exp(\mathbf{X}_i \beta). \quad (2)$$

Equation (2) implies that the log hazard rate is linear in the covariates, and that the marginal effects of changes in the latter are $\partial(\ln h_i) / \partial \mathbf{X} = \beta$.²

There is neither strong theoretical nor empirical justification for setting $g(\cdot) = \exp(\cdot)$. One alternative is to set $g(\cdot) = \Gamma(\cdot)$, where $\Gamma(\cdot)$ is the cumulative density function of some distribution evaluated at $\mathbf{X}_i \beta$. In this usage, $\Gamma(\cdot)$ has no probabilistic interpretation; it simply provides a monotonically increasing mapping of $\mathbf{X}_i \beta$ into the interval $[0, 1]$.³ The specification we focus upon in this paper is $\Gamma(\cdot) = \Phi(\cdot)$, for Φ the c.d.f. of the standard normal distribution. Thus, we estimate

$$h_i(t) = h_0(t) \cdot \Phi(\mathbf{X}_i \beta). \quad (3)$$

The cumulative normal hazard model differs from the exponential specification in several ways. 1) It is symmetric around zero, whereas $g(\cdot) = \exp(\cdot)$ is not, and exhibits less convexity; as a result, predicted hazard rates are generally lower (higher) at low and high (intermediate) values of $\mathbf{X}_i \beta$. Figure 1 illustrates these relationships. 2) Cumulative distribution functions restrict predicted hazard rates to values less than 1 if the baseline hazard is similarly limited.⁵ 3) The marginal

² See Kalbfleish and Prentice (1980) or Kiefer (1988) for further discussion.

³ We experimented with setting $g(\cdot) = \exp[\lambda(Z_i)]$, where $\lambda(\cdot)$ is the Edgeworth transformation: $\lambda(Z_i) = \alpha_1 Z_i + \alpha_2 Z_i^2 + \dots + \alpha_n Z_i^n$, for $Z_i = \mathbf{X}_i \beta$. The maximized log-likelihood was virtually unaffected by the inclusion of higher order terms.

⁴ Similar results were obtained using the cumulative logistic distribution.

⁵ For instance, with time-invariant baseline hazards, the range of predicted hazard rates is the closed interval $[0, 1]$ when $g(\cdot) = \Phi(\cdot)$.

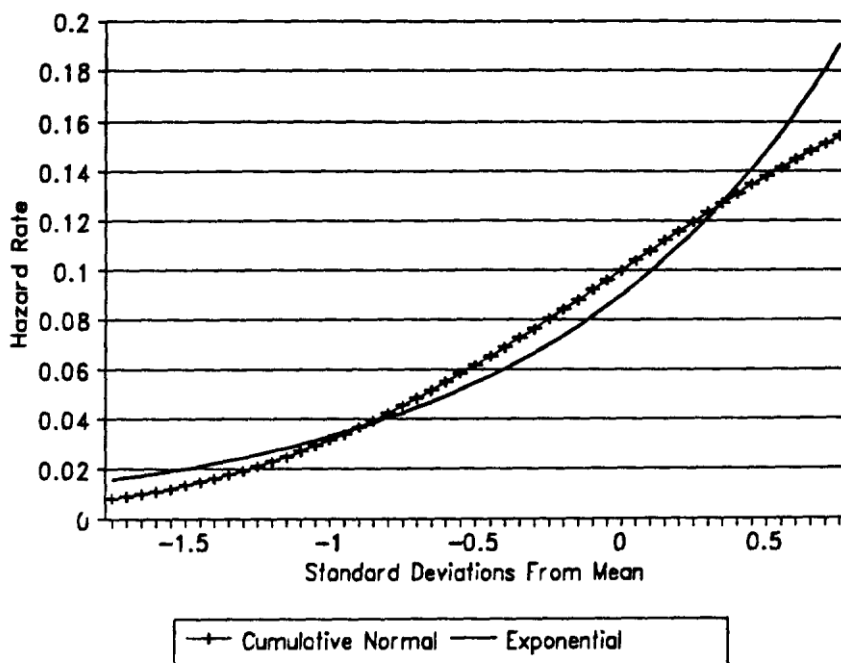


Figure 1. Hazard rates for exponential and cumulative normal specification of $g(\cdot)$.

effects of particular covariates depend on values of the other regressors; when $g(\cdot) = \Phi(\cdot)$, $\partial(\ln h)/\partial X = [\phi(\cdot)/\Phi(\cdot)]\beta$, for ϕ the p.d.f. of the standard normal distribution.

II. Data

Nonemployment spell durations following permanent job loss are analyzed below, using data from the January 1988 *Current Population Survey* (CPS)⁶ The sample includes 25-60 year old workers who lost jobs between 1983 and 1985 as a result of a business failure, plant closure or relocation, or a layoff resulting from slack work or position or shift abolished. The dependent variable is the number of weeks of nonemployment in the first spell following loss of the job. Due to top-coding, nonemployment durations are right censored at 98 weeks. This censoring is taken into account in the maximum likelihood estimates reported below. Covariates are incorporated to control for sex, race (white versus nonwhite), household status (head versus nonhead), reason for job loss (plant closing versus partial layoff), residence in right-to-work states, work experience, education, job seniority, number of children, the state or SMSA unemployment rate, industry growth rate, and an earnings residual remaining after controlling for observable characteristics.⁷

III. Results

⁶ The January 1988 *CPS* includes a special supplement containing information on involuntary job terminations occurring during the previous 5 years. Similar data, also available in the January 1984 and 1986 *CPS*, have been used by other researchers (Addison and Portugal 1987, Podgursky and Swaim 1987, Ehrenberg and Jakubson 1988) to analyze postdisplacement joblessness.

⁷ Further information on the data and sample is contained in Ruhm (1992).

The exponential and cumulative normal proportional hazard specifications are estimated using two functional forms for h_0 . The first is the Weibull baseline hazard, $h_0(t) = h_0 \alpha t^{\alpha-1}$, for which the log-likelihood function is

$$L(\theta) = \sum_{i=1}^n d_i \ln f(t_i, \theta) + \sum_{i=1}^n (1 - d_i) \ln [1 - F(t_i, \theta)] \quad (4)$$

where f and $1 - F$ indicate density and survivor probabilities, d_i equals 0 (1) if the i th spell is censored (uncensored), and θ includes the parameters (h_0, α, β) .

Second, we use the Cox (1972) partial likelihood model, which leaves the parametric form of the baseline hazard unspecified and relies on information on the rank ordering of survival times.⁸ The contribution of the j th shortest duration to the partial likelihood is the ratio of the corresponding hazard rate divided by the sum of hazards for spells still in progress through the immediately preceding period. Absent censoring or ties, the log-likelihood function is

$$L(\theta) = \sum_{i=1}^n \left(\ln g(\mathbf{X}_i, \beta) - \ln \left[\sum_{j=i}^n g(\mathbf{X}_j, \beta) \right] \right). \quad (5)$$

In addition to estimating the parameters, β , we calculate survival probabilities at various durations of joblessness. Since the integrated hazard is unspecified in the Cox model, we use Kiefer's (1988) nonparametric estimator:

$$\hat{H}_0(t_k) = \sum_{j=1}^k d(t_j) / D(t_j) \quad (6)$$

where $d(t_j)$ is the number of spells ending at t_j and $D(t_j)$ is the denominator of the likelihood contribution of the k th ordered (uncensored) observation.

The maximized log-likelihood (L) is used to measure relative performance of the alternative specifications of $g(\bullet)$. With Weibull baseline hazards, L is higher for $g(\bullet) = \Phi(\bullet)$ than for the exponential transformation (— 7113.08 in the former case and — 7114.54 in the latter). Conversely, for the Cox model, L is marginally higher for $g(\bullet) = \exp(\bullet)$ than for the cumulative normal specification (— 11425.38 versus — 11425.97). Thus, the desired form for $g(\bullet)$ is sensitive to the parametric assumptions on the baseline hazard.

To distinguish, in a statistical sense, which specification for $g(\bullet)$ is closest to the true data generating process, we constructed the likelihood ratio test for model selection among nonnested hypotheses proposed by Vuong (1989):

⁸ More general parametric specifications of the baseline hazard are also available. For instance, $h_0(t) = h_0 \alpha \exp[(\alpha - 1)\tau]$ uses the Box—Cox transformation $\tau = (t^\lambda - 1)/\lambda$, with $\lambda \geq 0$, and reduces to the Weibull hazard when $\lambda = 0$ and to the Gompertz form when $\lambda = 1$. We do not examine such extensions in this paper.

$$V = \frac{LR}{n^{1/2}\bar{\omega}} \quad (7)$$

where LR is the difference in the maximized log-likelihoods between specifications, n is the sample size, and $\bar{\omega}$ is a normalizing constant. Defining $f_1(t_i, \theta)$ and $f_2(t_i, \theta)$, respectively, as the estimated densities for the i th observation using the exponential and cumulative normal forms of $g(\bullet)$,

$$\bar{\omega}^2 = n^{-1} \sum_{t=1}^n \left[\log \frac{f_1(t_i, \theta)}{f_2(t_i, \theta)} \right]^2. \quad (8)$$

Since the limiting distribution of V is the standard normal, the relative performance of the alternative specifications for $g(\bullet)$ can be evaluated by choosing a critical test value C, corresponding to the desired level of statistical significance, and comparing V to C. If $V > C$ ($V < -C$), the exponential (cumulative normal) functional form for $g(\bullet)$ is superior; if $-C < V < C$, we are unable to conclude that either specification dominates. Calculated values for V are -0.64 and 0.14 , respectively, for the Weibull and Cox models. Thus, although neither form of $g(\bullet)$ is clearly superior ($\Pr(V \leq 0.64) = 0.26$ and $\Pr(V \geq 0.14) = 0.44$), the cumulative normal specification performs at least as well as the exponential form.

We next calculated predicted marginal effects of changes in the covariates for an "average" individual, and then compared survivor probabilities for three hypothetical persons. The form of $g(\bullet)$ is most critical when analyzing survivor probabilities, since predicted marginal effects differ substantially for outliers. The first two columns of Tables 1 and 2 provide maximum likelihood estimates for $g(\bullet) = \exp(\bullet)$, with marginal effects of the covariates calculated as $e^\beta - 1$. The third and fourth columns show findings for $g(\bullet) = \Phi(\bullet)$, with marginal effects $(e^{[\Phi(\bullet)/\Phi(\cdot)]\beta} - 1)$ normalized such that the experience estimate is the same as for the exponential specification. Qualitative findings are not sensitive to this particular normalization. The experience, education, seniority, area unemployment rate, industry growth rate, and wage residual variables are divided by their standard deviations so that a one unit change corresponds to a difference of one standard deviation.

The covariates generally have large and statistically significant effects which are entirely consistent with the findings of previous research (e.g., Podgursky and Swaim 1987, Ruhm 1992). Marginal effects are fairly similar across specifications of $g(\bullet)$. With Weibull baseline hazards, the differences never exceed 2.1 percentage points or 15% (Table 1). For the Cox model (Table 2), the disparities are somewhat larger, surpassing 3 percentage points in 3 cases and 15% in 5. The parameters are less precisely estimated for the cumulative normal specification of $g(\bullet)$, especially when considering reason for displacement, industry growth rate, or predisplacement wage residual.

Table 1. Regression Results for Weibull Hazard Model

Regressor	$g(\cdot) = \exp[\mathbf{X}\beta]$		$g(\cdot) = \Phi[\mathbf{X}\beta]$	
	Coefficient	Marginal Effect	Coefficient	Marginal Effect
Intercept	-3.242 (0.116)		-1.333 (0.084)	
Experience	-0.226 (0.039)	-20.2%	-0.093 (0.021)	-20.2%
Education	0.086 (0.030)	9.0%	0.040 (0.019)	10.2%
Seniority	-0.159 (0.042)	-14.7%	-0.065 (0.021)	-14.6%
Household Head	0.426 (0.088)	53.1%	0.176 (0.057)	53.3%
Female	-0.477 (0.069)	-37.9%	-0.207 (0.043)	-39.5%
Nonwhite	-0.751 (0.109)	-52.8%	-0.291 (0.069)	-50.7%
Closing	0.147 (0.056)	15.8%	0.059 (0.032)	15.4%
Area Unemployment Rate	-0.235 (0.040)	-20.9%	-0.094 (0.016)	-20.4%
Number of Children	-0.085 (0.028)	-8.1%	-0.040 (0.010)	-9.3%
Right-to-Work State	0.171 (0.076)	18.6%	0.076 (0.036)	20.3%
Industry Growth Rate	0.097 (0.042)	10.2%	0.035 (0.020)	8.9%
Wage Residual	0.078 (0.035)	8.1%	0.032 (0.018)	8.1%
α	0.742 (0.023)		0.742 (0.019)	
$-L$		7114.54		7113.08

Note: Standard errors in parentheses. Marginal effects in the cumulative normal specification for $g(\cdot)$ are normalized such that the marginal impact of experience is the same as in the exponential functional form.

To demonstrate the sensitivity of predicted survival probabilities to the specification of $g(\cdot)$, for persons with characteristics associated with rapid or slow reemployment, survival probabilities were calculated for three hypothetical individuals. The first worker is "average" in that all regressors were set equal to the sample mean values. The second possesses characteristics associated with rapid reemployment.⁹ The third individual has contrasting characteristics associated with slow reemployment!¹⁰

⁹ He is a white male childless household head, displaced by a plant closing, and living in a right-to-work state. His years of education, earnings residual, and predisplacement industry growth rate are one standard deviation above the sample mean, while his previous job seniority, labor market experience, and the unemployment rate (in his state or SMSA) are one standard deviation below it.

¹⁰ She is a black female nonhead with three children, involved in a partial layoff in a state without right-to-work laws. Her education, earnings residual, and industry employment growth are one standard deviation lower than average, while her experience, seniority, and local unemployment rate are one standard deviation above it.

Table 2. Regression Results for Cox Hazard Model

Regressor	$g(\cdot) = \exp[X\beta]$		$g(\cdot) = \Phi[X\beta]$	
	Coefficient	Marginal Effect	Coefficient	Marginal Effect
Experience	-0.146 (0.025)	-13.6%	-0.162 (0.021)	-13.6%
Education	0.063 (0.029)	6.5%	0.080 (0.034)	7.5%
Seniority	-0.105 (0.032)	-10.0%	-0.112 (0.034)	-9.6%
Household Head	0.289 (0.059)	33.5%	0.274 (0.071)	28.0%
Female	-0.308 (0.057)	-26.5%	-0.389 (0.060)	-29.6%
Nonwhite	-0.478 (0.087)	-38.0%	-0.465 (0.089)	-34.2%
Closing	0.110 (0.057)	11.6%	0.097 (0.060)	9.1%
Area Unemployment Rate	-0.145 (0.020)	-13.5%	-0.170 (0.030)	-14.2%
Number of Children	-0.055 (0.020)	-5.4%	-0.084 (0.030)	-7.3%
Right-to-Work State	0.119 (0.046)	12.6%	0.123 (0.051)	11.7%
Industry Growth Rate	0.063 (0.029)	6.5%	0.059 (0.032)	5.5%
Wage Residual	0.047 (0.022)	4.8%	0.054 (0.035)	5.0%
-L		11425.38		11425.97

Note: Standard errors in parentheses. Marginal effects in the cumulative normal specification for $g(\cdot)$ are normalized such that the marginal impact of experience is the same as in the exponential functional form.

The cumulative normal form of $g(\cdot)$ yields lower survival probabilities for "average" individuals, but higher rates of survival for persons with attributes associated with rapid or delayed reemployment. To show this, Figures 2 and 3 indicate ratios of predicted survival probabilities, with the numerator calculated for $g(\cdot) = \Phi(\cdot)$ and the denominator for $g(\cdot) = \exp(\cdot)$. Ratios less (greater) than one therefore indicate that predicted survival probabilities are lower (higher) in the cumulative normal than in exponential specification for $g(\cdot)$. With Weibull baseline hazards, the individual with characteristics associated with rapid reemployment is predicted to remain jobless at least 2 weeks 5.5% more often (0.613 versus 0.581) in the former than in the latter case. By 4 weeks, the gap rises to 8.4% (0.490 versus 0.452), at 13 weeks to 18.7% (0.216 versus 0.182), and by 52 weeks to 60.0% (0.016 versus 0.010). The differences are even more dramatic in the Cox model. At 2 weeks, the gap is 10.6% (0.733 versus 0.663), at 4 weeks 18.5% (0.597 versus 0.504), by 13 weeks 52.2% (0.280 versus 0.184), and at 52 weeks 156% (0.062 versus 0.025). The disparity in relative survival probabilities also grows over time for the "low" hazard rate and "average" individuals, although less significantly.

While relative differences in survival rates are greatest for individuals expected to obtain rapid reemployment, the absolute disparities, beyond 5 months, are

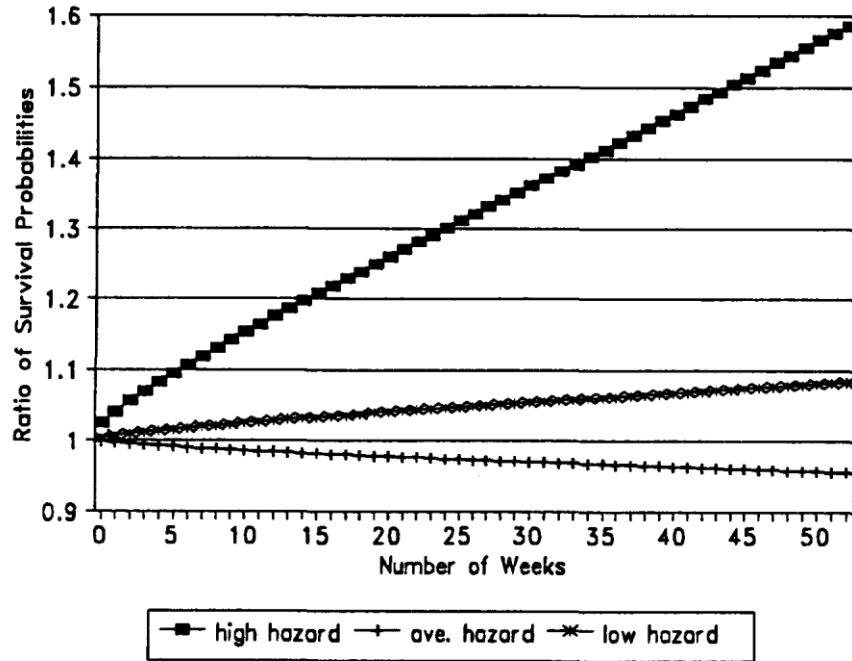


Figure 2. Ratio of survival probabilities for Weibull model.

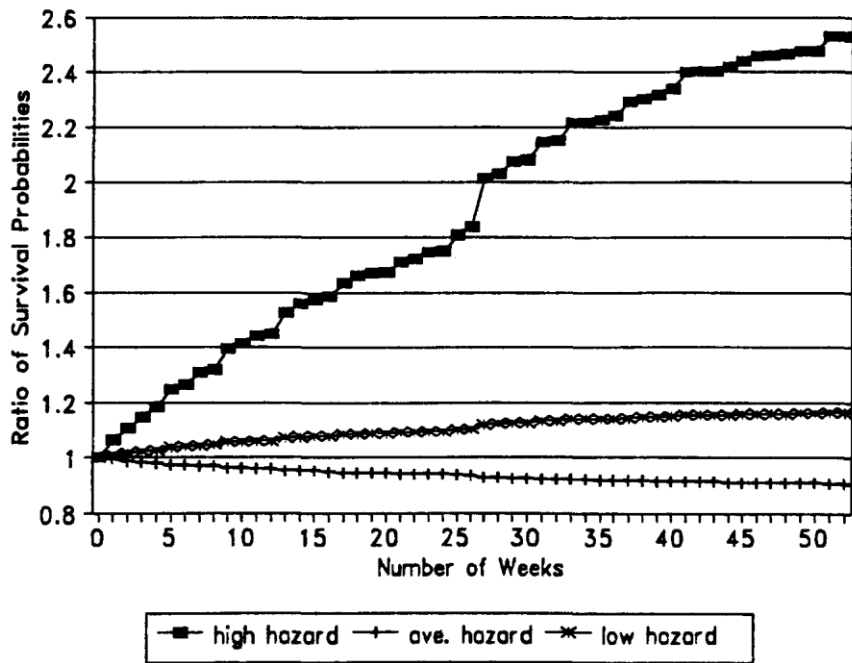


Figure 3. Ratio of survival probabilities for Cox model.

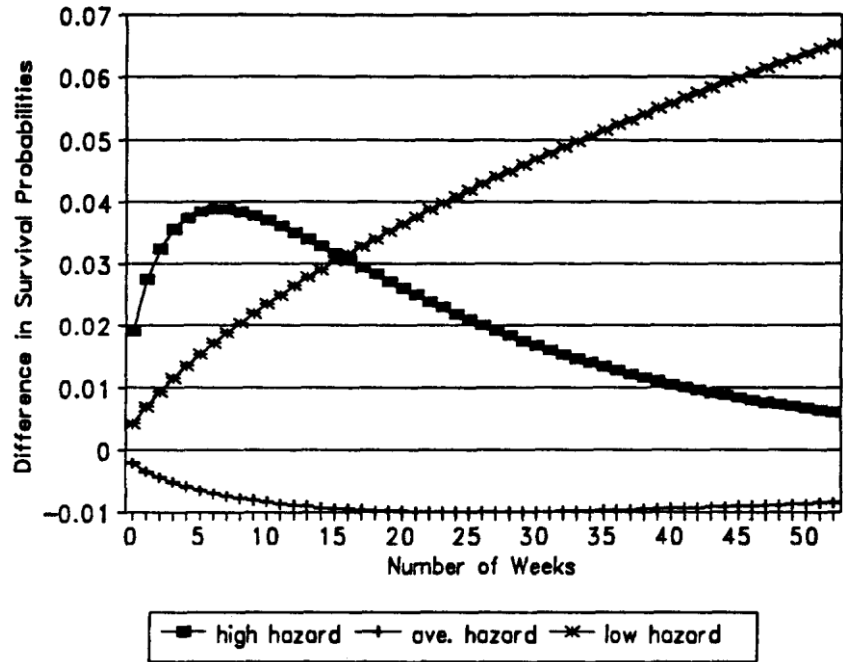


Figure 4. Differences in survival probabilities for Weibull model.

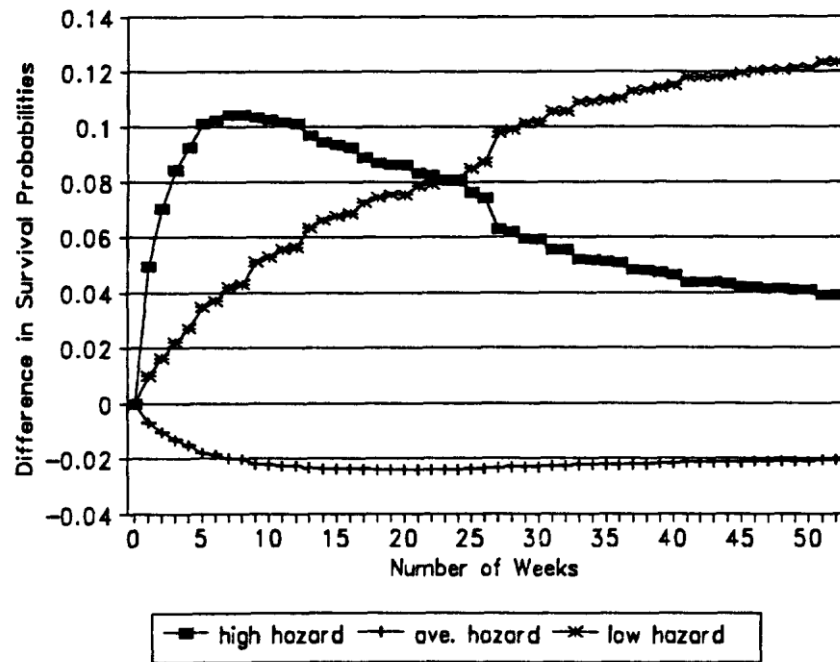


Figure 5. Differences in survival probabilities for Cox model.

largest for persons with characteristics predicting slow reemployment (see Figures 4 and 5). This occurs because survival probabilities for the two other groups quickly become quite low. For

instance, in the Weibull model, the "high" hazard individual is predicted to remain out of work through 52 weeks 60% more frequently using the cumulative normal rather than the exponential specification for $g(\bullet)$. Since the estimated survival probability is less than 2% in either case, however, the absolute difference is only 0.6 percentage points. Conversely, at 52 weeks, the corresponding gap is only 8.4% for the "low" hazard individual, but the estimated survival probability exceeds 77% (using either specification for O_i), and so the absolute disparity is a 10 times larger 6.5 percentage points.

IV. Summary

This paper proposes an alternative to the proportional hazard model conventionally estimated. Rather than multiplying the baseline hazard by $\exp(\mathbf{X}_i \beta)$, we suggest using $\Phi(\mathbf{X}_i \beta)$, for $\Phi(\cdot)$ the c.d.f. of the standard normal distribution. This alternative is compared to the conventional specification using data on nonemployment durations for displaced workers surveyed in January 1988. Using Weibull baseline hazards, the cumulative normal form of $g(\bullet)$ outperforms the exponential specification. Conversely, for the Cox semi-parametric model, $g(\bullet) = \Phi(\bullet)$ does slightly worse. Although neither functional form for $g(\bullet)$ statistically dominates, the findings provide weak evidence favoring the cumulative normal specification. Marginal effects of the covariates are typically insensitive to the specification of $g(\bullet)$, suggesting that parameter estimates may be fairly robust across functional forms. Predicted survival probabilities vary substantially, however, for persons with characteristics associated with rapid or slow reemployment.

References

- Addison, J., and Portugal, P. 1987. On the distributional shape of unemployment duration. *Review of Economics and Statistics* 69:520-526.
- Cox, D. 1972. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society* 34(B):187-220.
- Ehrenberg, R., and Jakubson, G. 1988. *Advance Notice Provisions in Plant Closing Legislation*. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.
- Heckman, J., and Singer, B. 1984. A method of minimizing the impact of distributional assumptions in economic models for duration data. *Econometrica* 52(2):271-320.
- Kalbfleisch, J., and Prentice, R. 1980. *The Statistical Analysis of Failure Time Data*. New York: Wiley.
- Katz, L., and Meyer, B. 1990. The impact of the potential duration of unemployment benefits on the duration of unemployment. *Journal of Public Economics* 41(1):45-72.
- Kiefer, N. 1988. Economic duration data and hazard functions. *Journal of Economic Literature* 26(2):646-679.
- Lancaster, T. 1990. *The Econometric Analysis of Transition Data*. New York: Cambridge Univ. Press.
- Podgursky, M., and Swaim, P. 1987. Duration of joblessness following job displacement. *Industrial Relations* 6:213-226.
- Ruhm, C. 1992. Advance notice and postdisplacement joblessness. *Journal of Labor Economics* 10(1):1-32.
- Vuong, Q. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2):307-333.