VATT-KESKUSTELUALOITTEITA
VATT-DISCUSSION PAPERS

107

COMPETING RISKS FOR VACANCY DURATIONS AND ENDOGENOUS COMMON DEPENDENCE

Ilpo Suoniemi

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Helsinki 1995

ACKNOWLEDGEMENT

A preliminary version of the paper has been presented at the Seventh EALE conference in Lyon, 7-10.9.1995. In preparing the study I have been assisted by Juha Rantala. I am grateful to him for useful discussions on the subject and comments on an earlier version of the paper. All errors are naturally my own making. Furthermore, I wish to thank the newspaper Helsingin Sanomat for providing me with data on the total volume of vacancy advertisement.

SAATTEKSI

Tämän työn alustava versio on esitetty kokouksessa, Seventh EALE conference in Lyon, 7-10.9.1995. Tutkimusta valmistellessani olen saanut apua Juha Rantalalta. Kiitän häntä hyödyllisistä keskusteluista ja aiempaa versiota koskevista huomautuksista. Virheet ovat tietenkin omaa käsialaani. Lisäksi kiitän Helsingin Sanomia, että sain käyttöön työpaikkailmoitusten kokonaismäärätietoja.

ISBN 951-561-149-0

ISSN 0788-5016

Valtion taloudellinen tutkimuskeskus

Government Institute for Economic Research

Hämeentie 3, 00530 Helsinki, Finland

Painatuskeskus Pikapaino Opastinsilta

Helsinki, October 1995

Suoniemi, Ilpo: Competing risks for vacancy durations and endogenous common dependence. Helsinki, VATT, Government Institute for Economic Research. 1995. (C ISSN 0788-5016; No 107) ISBN 951-561-149-0.

ABSTRACT: The paper presents a discussion how firms rationally adjust the level of their recruitment effort in relation to the ease of filling the vacancy through public unemployment offices. This induces dependence between the durations in the two channels of recruitment. Multivariate models with random proportional hazards generated by mixtures, the frailty distributions are used in the paper to discuss and estimate a competing risks model with mutually dependent recruitment channels using Finnish vacancy duration data. The channels of recruitment are found to be (positively) associated. Vacancy durations vary with respect to region, industry, occupational status and local labour market conditions. The explanatory variables have a more moderate effect in the private recruitment channel possibly reflecting a rational adjustment in the search effort by the employer.

KEY WORDS: Vacancy duration, competing risks model, multivariate distributions, mixtures, frailties.

Suoniemi, Ilpo: Avoimet työpaikat, kilpailevat kanavat ja endogeeninen riippuvuus. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus. 1995. (C ISSN 0788-5016; No 107) ISBN 951-561-149-0.

TIIVISTELMÄ: Tutkimuksessa tarkastellaan, miten yritykset sopeuttavat työntekijän etsintäponnistelujaan suhteessa siihen, miten helposti avoin työpaikka täyttyy työnvälityksen kautta. Tämä saa aikaan riippuvuutta näiden kahden täyttökanavan välille. Työssä käytetään moniulotteisia malleja, joissa generoidaan suhteellisen satunnaishasardin malleja sekoitejakauman (frailty) avulla. Uudenmaan avointen työpaikkojen aineistoa käyttäen estimoidaan kilpailevien riskien malli, jossa täyttökanavat ovat positiivisessa yhteydessä toisiinsa. Kestot eri kanavissa riippuvat aluetta, toimialaa ja ammattiasemaa kuvaavista muuttujista sekä paikallisesta työmarkkinatilanteesta. Vaikutukset ovat suhteellisesti suurempia työnvälityskanavan kautta tapahtuneeseen täyttöön. Tämä voi viitata siihen, että työnantaja rationaalisesti sopeuttaa etsintäponnistelujaan.

ASIASANAT: Avoimet työpaikat, kilpailevien riskien malli, moniulotteiset jakaumat, sekoitusjakaumat

Contents

1 Introduction

Vacancy durations are in many empirical studies found to depend on various characteristics of both the employer and jobs offered, on regional labour market conditions and on the recruitment channel (van Ours 1989, and van Ours and Ridder 1991). Measuring the effects of those determinants may provide information on the matching function which is a key concept in the flow approach to labour markets (Blanchard and Diamond 1989). Van Ours and Ridder (1991) consider the role of hiring standards in search for new employees. Employers may lower the job requirements if the waiting cost due to an unfilled vacancy is high and the wage associated with a vacancy is given. They find that job requirements are not lowered over the duration of the vacancy which suggests that the reservation productivity level does not decline over time. In a companion paper Van Ours and Ridder (1992) present some evidence that employer search is nonsequential. The choice of recruitment channels and methods are rarely revised as the search continues. These observations are used here to present the recruitment decision in a simplified, 'reduced form' framework.

In the present paper the focus is on the recruitment channels. It is argued that the recruiment channels are endogenously affected by firms' decisions how much resources to spend in using private recruitment channels in contrast to using the public employment offices that are free. In an environment where there are costs of obtaining information the employers buy advertisement and make other efforts to lower the corresponding information costs of jobseekers. There is a natural trade-off between the effort costs and the waiting costs of an unfilled vacancy.

Various market frictions may affect that the duration of vacancies is too long. This may hold even under the conditions of substantial unemployment and may inhibit firms from fully exploiting their growth potential. In the case of an upswing these effects may slow down the growth in employment. The examination in the paper gives some indication that the recruitment costs may play a role here. Improving labour market conditions for the employers make them reduce vacancy advertisement, and therefore the duration of vacancies is longer than it would be without this rational adjustment in effort level. The gain to the firms accrues both in terms of less advertisement costs and partly in the form of reduced waiting cost. On the other hand, the improved performance of public employment offices in their job-worker matching effort accrues to the society partly in the form of reduced advertisement costs to the employers but is shown in a dampened manner in the mean vacancy duration and unemployment rate.

Multivariate proportional hazards models with random coefficients of proportionality generated by mixtures, the frailty distributions (Vaupel et. al. 1979 and Hougaard, 1984) are used in the paper to consider and estimate a competing risks model with mutually dependent recruitment channels using Finnish vacancy duration data. A related dependency measure, the cross-ratio function (Oakes, 1989) is used to characterize association between the durations. The present models are based on parametric assumptions such as exponential, Weibull or Burr distribution which are widely used to analyze count processes in applied labour economics. The parameters are estimated by solving numerically the nonlinear estimating equations given by the principle of maximum likelihood.

The remainder of the paper is organized as follows. Section 2 presents a simplified model of job recruitment to discuss the choice of the optimal effort level by the employer. Section 3 introduces a class of multivariate distributions that are generated by a mixture model having a random proportional hazards component, called a frailty. Section 4 develops the analysis in a competing risks framework and reports the findings from the estimations. Section 5 contains the concluding remarks.

2 A simplified model of job recruitment

Consider the following situation: a firm is about to fill a vacancy and incures losses L(t) in present value if the vacancy is unfilled up to the time t.[1] The expected waiting cost is

$$L = \int_0^\infty L(t)dF_t = \int_0^\infty S(t)dL_t, \tag{1}$$

by partial integration, $L(0) = 0$. Above F is the distribution function of vacancy durations, T, and S is the corresponding survivor function, S = 1-F, S(t) = Pr(T > t). In the case of a constant current cost, L_0, $dL_t = e^{-\delta t} L_0$, where δ is the discount rate.

Introduce next competing channels for job recruitment. The first channel is through unemployment offices, the use of which is free to the firm. In contrast, the second channel can be affected by advertisement bought by the firm. It is assumed that the firm has no preference over applicants from either channel.[2] The wage associated with a vacancy is considered fixed. Lastly one allows for a withdrawal of an unfilled vacancy. In the following one considers the model

$$S = \exp\{w \log B_1 + wa \log B_2 + \log B_3\}, \tag{2}$$

where B_i is the baseline survivor function of channel i, i= 1,2. Symbol $a, a \geq 0$ indicates the effort level with an attached cost, $c(a)$. We have assumed that the above survivor functions B_1 and B_2 are dependent on firm and vacancy specific individual factors, w which are known to the firm in making the decision how much resources c to spend to get a desired effect by advertising.

In (2) the customary proportional hazard assumption in the two channels w.r.t individual effects w and advertisement decision a is invoked. The common dependence through w is written in a form which will later prove useful in the econometric formulations of the model, on the assumption that conditional on w, called a frailty (Vaupel

[1]Since the primary interest lies in econometric model building the decision problem is not presented in a dynamic optimization framework with a sequential advertisement decision. At least in a time-homogenous set-up no significant new insight was found. In contrast, in what follows the advertisement decision is an once and for all decision made at the beginning of the vacancy duration. Empirical studies by van Ours and Ridder (1992) indicate that employer search is nonsequential. The choice of recruitment channel and methods are rarely revised as the search continues.

[2]In cases with adverse selection there may be a separating equilibria with self-selection of the applicants and with employers specializing in the exclusive use of one of the search channels, see Barron and Mellow (1989).

et.al. 1979), $w > 0$, the channels are independent. Note that the effect of advertisement in the second channel is through the term $h = wa$ which means that advertisement is more effective for those vacancies that are easier to fill by the first channel, i.e. have a high w, than for those that are generally hard to fill, i.e. have a low w. This does not sound too unrealistic.[3] The final survivor function B_3 captures the effect of vacancy withdrawal. In addition the present value dL can be easily absorbed in the last term.[4]

To make things easier, assume that the baseline survivor functions satisfy $B_1 = B_2 = B$.[5] The firm minimizes total costs (TC) which are due to the direct cost of advertisement (c), and expected waiting costs (L),

$$TC = c(a) + L(w,a) = c(a) + \int_0^\infty \exp\{w(1+a)\log B(t) + \log B_3(t)\}\, dt, \tag{3}$$

w.r.t advertisement effort a.

The first order condition for an optimum is

$$c'(a) = -L_a \equiv -\frac{\partial L}{\partial a} = -w\int_0^\infty \log B(t) \exp\{w(1+a)\log B(t) + \log B_3(t)\}\, dt. \tag{4}$$

For fixed parameter values the R.H.S of (4) is decreasing in effort level, a.[6] Assume increasing marginal costs, $c'' \geq 0$, and that marginal reduction in waiting costs $-L_a$ is decreasing in w. It is easy to show that the optimal effort level is a decreasing function of w.[7] A simple comparative statics exercise reveals that at the optimum

$$\eta_{a,w} \equiv \frac{da/dw}{a/w} = \frac{c' - (1+a)f}{a(c''+f)}, \tag{5}$$

where the positive function f is defined by

$$f(a,w) = w^2 \int_0^\infty (\log B(t))^2 \exp\{w(1+a)\log B(t) + \log B_3(t)\}\, dt. \tag{6}$$

It has been observed that under plausible assumptions the effort level is a decreasing function of w, where the latter variable measures the effectiveness of employment offices. However, below we are primarly interested in the observable effect of advertisement on the hazard of job-fulfillment through the second channel, i.e. the variable $h = wa(w)$ and its dependence on w. Since $a = h(w)/w$ is decreasing in w, -h is a star-shaped

[3] Later we will relax the assumption on the specific functional form $h = wa$.

[4] Consider waiting costs which stay (ex ante) constant in current value through time, L_0 and a fixed discount rate δ. In this case $L(t) = L_0 \int \exp\{-\delta u\}du$, and $dL_t = L_0 \exp\{-\delta t\}$. The term $\exp\{-\delta t\}$ can be included in B_3 as an exponential stopping rule, and the advertisement costs can be measured directly in units of L_0.

[5] Here some sort of normalisation of the distribution should be used in order to separate the effects due to w and the effort level a from those of the baseline survivor function.

[6] In this paper the terms (strictly) increasing are used in lieu of (increasing) nondecreasing.

[7] Using these observations it is immediately obvious that advertisement effort increases if marginal costs diminish, or equivalently, the fixed current loss L_0, see footnote 4, increases or the discount rate δ goes down, or the probability of withdrawal decreases.

function. Therefore h is a subadditive function suggesting that unimodality may prevail in commonly encountered cases. Below examples and simple conditions are given that guarantee unimodality.

Consider briefly the case: $c'' > 0$. Now

$$\eta_{h,w} = 1 + \frac{da/dw}{a/w} = \frac{c''a + c' - f}{a(c'' + f)}. \tag{7}$$

The denominator in (7) is positive by the second order condition for an optimum. The numerator can be nonpositive only if $c''a - f < -c'$. To get some intuition on this condition recall that $c'(a) = -L_a$, and $L_{aa} = f$. Variable L can be interpreted as the expected value of a waiting time, T where the duration has the survivor function $B(t)^{w+h(w)}B_3(t)$. At the optimal effort level, a^* the above condition is therefore equivalent with $ac''/c' + L_{aa}/L_a < -1$. The L.H.S is the sum of two terms. The first one is the proportional change in marginal direct costs if advertisement is increased and the second one which has a negative value is the corresponding proportional saving in marginal waiting costs if the effort level is adjusted on the envelope.

In addition one can see that with a high value of the parameter w, i.e. with a very short expected value L, the above condition is more likely to hold. This may result in that an initially increasing dependence of h on w is eventually a decreasing one. In this case the particular vacancy is so much sought after that it does not pay to use advertisement. Let us examine some simple examples, to get further feel of the problem.

Example 1. Consider the simplest possible case where B corresponds to a time-homogenous Poisson arrival process with a unit time intensity parameter, and the withdrawal process (B_3) is a similar process with the parameter δ. The last parameter captures the combined effect, i.e. it is the sum, of the discount factor and the actual cancelling intensity. In addition, the proportional hazard of the second channel, $h = \nu wa$ where ν captures the effectiveness of advertising effort relative to that of employment offices.

In the present case one can write (4) in the form

$$\frac{c'(h/(\nu w))}{\nu w} = -\frac{\partial \mathcal{E}T}{\partial h}, \tag{8}$$

where $h = \nu wa$. Duration T has an exponential distribution with parameter $(w + h + \delta)$, and has an expected value $(w + h + \delta)^{-1}$. This gives

$$c'(h/(\nu w)) = \frac{\nu w}{(h + w + \delta)^2}. \tag{9}$$

Let first $c' = \gamma$. Then the observed hazard of the channel 2 is

$$\nu wa \equiv h = -w - \delta + \sqrt{\nu w/\gamma}, \tag{10}$$

if the R.H.S. is positive, and zero otherwise.

The upper left panel of Figure 1 presents the marginal cost of advertisement and the corresponging marginal reduction in waiting costs with some values of w. The upper right panel provides a typical shape of the observed hazard related to channel two, $h(w) = \nu w a(w)$, as a function of w. One can observe that if employment offices are very effective i.e. w is high enough it is not worthwhile to buy any advertisement. In this case the marginal reduction in waiting cost is everywhere below the line in the upper left panel which refers to the constant marginal cost of advertisement. This corresponds to the region in the right corner of the upper right panel.[8] Note that no advertisement is bought in the region near the origin. In this case the non-zero discount rate (or probability of vacancy withdrawal) makes it not wortwhile to advertise those vacancies that are almost impossible to fill.

Assume next linear marginal costs, $c' = \gamma a$. In this case the observed hazard of the second channel is a root of a third degree polynomial in h. It can be shown that the trinomial has only one real root which in addition is always positive. In the panels in the centre of Figure 1 we have given the corresponding graphics for this case. The same general apperance as in the fixed marginal cost is maintained execpt that with increasing marginal costs the hazard is increased less steeply and gets later to those values of w where the monotone relationship is lost. Furthermore, assuming zero marginal costs for a zero level of advertisement removes the region in the right where advertisement is never worthwhile in the fixed marginal cost case.

Example 2. Consider the case where B corresponds to a Weibull distribution with a duration shape parameter $\alpha < 1$. In the previous example both unemployment offices and advertisement have a constant hazard, i.e. they had a time-homogenous, 'ever-lasting' effect on the vacancy duration. In the present case it is assumed that the effects of their efforts may decay over time. The waiting cost is now

$$L = \int_0^\infty \exp\{-(w+h)t^\alpha - \delta t\}\, dt = \frac{1}{\delta}\left[1 + \int_0^\infty \exp\{-\delta t\}\, d\exp\{-(w+h)t^\alpha\}\right], \quad (11)$$

where $h = \nu w a$. Change of variables $x = (w+h)^{1/\alpha} t$ gives

$$L = \frac{1}{\delta}\left[1 - (w+h)^{1/\alpha}\mathcal{L}_x[\delta(w+h)^{1/\alpha}]\right], \quad (12)$$

where $\mathcal{L}_x$ is the Laplace transform of a standardized Weibull variable with the parameter α, and $\mathcal{L}_x[u] = \mathcal{E}e^{-ux}$.

In the present case one cannot give any closed form expression on the dependence between the two hazards. Instead we have given some numerical calculations on the above relation in the two bottom panels of Figure 1. The marginal cost of advertisement is here a linear function. The bottom panels are constructed similarly as those corresponding to Example 1. Note that the improved performance of public employment offices

[8]The observation is intimately related to the fact that marginal cost is fixed. In a more realistic case, see below, one would allow a very small amount of advertisement to be bought at a very low marginal cost to ensure an interior solution for the optimal advertisement decision.

in their job-worker matching effort moves the marginal savings curve (in waiting costs) inwards from $-L_a(w_0)$ to $-L_a(w_1)$. The benefit to the society is partly in the form of reduced advertisement costs to the employers which is shown in the bottom left panel as an area under the marginal cost curve and partly as an increase in the hazard of the second channel (bottom right panel in Figure 1). If the labour market conditions for the employers improve further they will reduce effort level more and there is practically no observable effect in the recruitment hazard of the private channel. The total effect on the labour market is shown in a dampened manner in the mean vacancy duration and unemployment rate.

The following characterization of (1) is useful to derive sharper results in comparative statics since it covers most commonly encountered classes of survival distributions. Start with

Lemma 1 *Assume that with a monotone change of variables*[9] $t \mapsto y$ *the expected waiting cost may be written using a member in a location family of distributions:*

$$L(w,h) = \int_B \exp\{-g(y - \mu(w+h))\} \exp\{\xi(y)\}\ dy, \tag{13}$$

where g *is an increasing function with the location* μ *depending on variables affecting the first channel,* w *and the second channel,* h*, respectively, with* $\mu' < 0$*. Let the implicit density* $g'e^{-g}$ *be strongly unimodal, i.e. the logarithmic density is a concave function of* μ*.*[10] *The function* ξ *which subsumes the withdrawal, discount process and the Jacobian of the transformation* $y \mapsto t$ *defines a* $\sigma-$*finite measure on* $B, B \subset R$*. In this case the derivative*

$$L_\mu(\mu) = \int_B g' \exp\{-g(y - \mu(w+h))\}\ \exp\{\xi(y)\}\ dy, \tag{14}$$

and $\log L_\mu(\mu)$ *is a concave function of* μ*.*

Proof: By application of Artin's theorem, see Marshall and Olkin (1979), p. 452.

The first order condition for an optimum corresponding to (4) is given by

$$\log c_h(w,h) = \log(-\mu') + \log L_\mu(\mu), \tag{15}$$

and the dependence of h on w can be written as

$$\begin{aligned} h'(w) &= \frac{-c_{hw}/c_h + \mu''/\mu' + \mu' L_{\mu\mu}/L_\mu}{c_{hh}/c_h - \mu''/\mu' - \mu' L_{\mu\mu}/L_\mu} \\ &= \frac{-c_{hw}/(\mu' c_h) + \mu''/(\mu')^2 + L_{\mu\mu}/L_\mu}{c_{hh}/(\mu' c_h) - \mu''/(\mu')^2 - L_{\mu\mu}/L_\mu}. \end{aligned} \tag{16}$$

[9] In most cases one considers the transformation $y(t) = \log t$.

[10] This faciliates relatively straightforward estimation of the location function μ, using for example ML- or L-estimators.

Since (16) is obtained through division by $\mu' < 0$ the second order condition implies that the denominator is negative in the present formula. The map $w \mapsto w + h(w)$ is increasing everywhere if $c_{hh} > -c_{hw}$. Furthermore, the distributional assumptions of the lemma guarantee that $L_{\mu\mu}/L_\mu$ is a decreasing function of μ. This implies that $-L_{\mu\mu}/L_\mu$ is a decreasing function of w, since $w \mapsto w + h(w)$ is an increasing map. One obtains

Proposition 1 *If $c_{hh} > -c_{hw}$, and $\mu''/(\mu')^2$ is decreasing in μ (increasing in w) and in addition $c_{hw}/(\mu' c_h)$ increases while $h'(w) < 0$ then the dependence $w \mapsto h(w)$ is at most unimodal.*

As a sequel to Example 2 consider Weibull duration with the functions $y : t \mapsto \log t$, $g : x \mapsto \exp\{\alpha x\}$, $\mu : x \mapsto -\frac{1}{\alpha}\log x$.[11] Assume that the marginal costs are represented by a power function in w and h, $\log c(w, h) = c_0 + (1+\gamma)\log h - \beta \log w$, with $\gamma > \beta$. Now

$$c_{hw}/(\mu' c_h) = -\alpha\beta(1 + h/w). \tag{17}$$

If $\beta > 0$, this goes to $-\infty$ as $w \downarrow 0$, if h has a nonzero limit on the right at zero. Consider the smallest value, w_m with $h'(w_m) = 0$. If $w > w_m$, $h' < 0$, and the function (17) increases having the limit value $-\alpha\beta$ as $w \uparrow \infty$. Therefore h is at most a unimodal function of w.

In the empirical part of the paper a model with lognormal durations is considered. Similarly, as above the model can be subsumed under a setup guaranteing unimodality using the above results.

3 Bivariate survival models induced by frailties

This section introduces two variants of the base line model (3) where the individual effects w affecting vacancy duration are partly unobserved and follow a given distribution. Since these individual effects are known to the firms in making the advertisement decision but not to the econometrician this induces dependence between the channels of recruitment. After optimisation on the part of the firm one can write the joint survivor function for the channels 1 and 2 conditional on w as

$$B(t_1, t_2) = \exp\{w \log B_1(t_1) + h(w) \log B_2(t_2)\}, \tag{18}$$

where B_j is the baseline survivor function of channel j, j= 1,2.

Below the econometric model is based on a class of distributions which have proven fruitful in survival analysis. These are generated by a mixture model having a random proportional hazards component, w, called a 'frailty' (see Vaupel et. al. 1979, and Hougaard, 1984) on the assumption that conditional on w, the channels are independent. Introduce a frailty distribution, G_w. Furthermore, assume to simplify things $h(w) = \nu w$, see below.

[11] In this example $-\mu''/(\mu')^2 = -1$, a decreasing function of w, and the assumptions of the lemma clearly hold.

Now the unconditional marginal survivor functions are

$$S_1(t_1) = \int \exp\{w \log B_1(t_1)\} dG_w = \mathcal{L}_w[-\log B_1(t_1)], \tag{19}$$

$$S_2(t_2) = \int \exp\{w\nu \log B_2(t_2)\} dG_w = \mathcal{L}_w[-\nu \log B_2(t_2)], \tag{20}$$

where $\mathcal{L}_w$ is the Laplace transform of w, $\mathcal{L}_w[u] = \mathcal{E}_w e^{-uw}$. The mixture with G_w has the following property, for any pair of G_w and S_1 there exist a B_1 such that (19) holds, $B_1(t_1) = \exp\{-\mathcal{L}_w^{-1}[S_1(t_1)]\}$.

Since the durations T_1, T_2 are conditionally independent the previous representation extends to the bivariate survivor function $S(t) = S(t_1, t_2) = Pr(T_1 > t_1, T_2 > t_2)$,

$$\begin{aligned} S(t) &= \int \exp\{w[\log B_1(t_1) + \nu \log B_2(t_2)]\} dG_w \\ &= \mathcal{L}_w[-\log B_1(t_1) - \nu \log B_2(t_1)]. \end{aligned} \tag{21}$$

Recall that all bivariate distributions having the joint survivor function $S(t_1, t_2)$ satisfy the following inequality $S_l \leq S \leq S_u$, where $S_u = \min\{S_1(t_1), S_2(t_2)\}$ is the Fréchet upper bound and $S_l = \max\{S_1(t_1) + S_2(t_2) - 1, 0\}$. In the more general case where h is an increasing function of w it can be shown that S(t) is a totally positive function in t_1, and t_2.[12] Therefore, the random variables T_1, and T_2 are *associated* and in particular there is no possibility of negative correlation.[13]

Next we introduce some concepts which have been used by Oakes (1989) to develop the analysis of frailty models. First, these models are a subclass of the archimedean distributions studied by Genest and MacKay (1986) which have the marginal distributions as parameters.[14] These have a general form

$$S(t_1, t_2) = p\left(q\{S_1(t_1)\} + q\{S_2(t_2)\}\right), \tag{22}$$

where p is a nonnegative decreasing function with p(0) = 1 and nonnegative second derivative, and q is its inverse function, and $S_1(t_1) = S(t_1, 0)$, and S_2 are the marginal survivor functions. Here p satisfies a stronger condition, it is a Laplace transform.

The cross-ratio function which forms the basis of Oakes (1989) work is defined by

$$\theta^*(t) = \frac{S(t) D_{12} S(t)}{D_1 S(t) D_2 S(t)}, \tag{23}$$

where D_j denotes the operator $-\partial/\partial t_j$. The cross-ratio function introduced by Clayton (1978) may be interpreted as the ratio of the hazard rate of the conditional distribution of

[12]The result follows since w and h and monotone in the same direction, $x^y, x > 0$ is a totally positive function, and by Theorem 18.A.4.a, p. 488 in Marshall and Olkin (1979).

[13]Random variables (T_1, T_2) are associated if $Cov(u(T_1, T_2), v(T_1, T_2)) \geq 0$ for all increasing functions u, v such that the covariance exists, see Marshall and Olkin (1989).

[14]Marshall and Olkin (1989) consider more general families of multivariate distributions having marginals as parameters which are generated by mixtures. These allow also for initially dependent survivals and they consider their relation to the various concepts of dependence introduced by Lehmann (1966). These are appropriate to survival analysis but space limitations refrain the discussion here.

T_1 given $\{T_2 = t_2\}$ to that of T_1 given $\{T_2 > t_2\}$. The former is equal to $f(t_1|t_2)/S(t_1|t_2) = D_{12}S(t)/D_2S(t)$ and the latter is $f(t_1|T_2 > t_2)/S(t_1|T_2 > t_2) = D_1S(t)/S(t)$. Note that the cross-ratio function is symmetric in (T_1, T_2). In addition the function is dependent only on the rank of the observations, i.e. the model is invariant under a monotone transformation of the time-scales. This property enables one to consider, say the logarithmic durations with no change in the definition of θ^*.

Oakes has shown that for archimedean distributions the cross-product function θ^* depends on $t = (t_1, t_2)$ only through some function $\theta(v)$ of v = S(t). The function θ is given by the formula

$$\theta(v) = \frac{-vq''(v)}{q'(v)}. \tag{24}$$

In addition he shows that $\theta(v)$ can be interpreted as a measure of local dependence which is closely associated with Kendall's (1938) coefficient of concordance, $\tau = \mathcal{E}sign\{(T_1^{(1)} - T_1^{(2)})(T_2^{(1)} - T_2^{(2)})\}$, where $T^{(i)}$, i = 1,2 are independent copies of $T = (T_1, T_2)$. A pair $(T^{(1)} - T^{(2)})$ is called concordant if $(T_1^{(1)} - T_1^{(2)})(T_2^{(1)} - T_2^{(2)}) > 0$, and otherwise discordant. In fact the ratio $(\theta(v) - 1)/(\theta(v) + 1)$ is a conditional version of Kendall's τ in frailty models.

Model 1. Consider the case where individual effects have a positive stricly stable distribution. These have Laplace transforms $\mathcal{L}(u) = exp(-u^{\delta})$, with $0 < \delta < 1$.[15] For the bivariate model (21), $\theta(v) = 1 + (1 - \delta)/(-\delta \log v)$ which decreases from infinity to unity as the survivor function v decreases from 1 to 0. The case $\delta \to 1$ corresponds to independence between T_1 and T_2, and as $\delta \to 0$ we obtain the bound $S(t) = \min\{S_1(t_1), S_2(t_2)\}$. The corresponding joint survivor function is

$$S(t) = \exp\{-\left[(-\log S_1(t_1))^{1/\delta} + (-\log S_2(t_2))^{1/\delta}\right]^{\delta}\}, \tag{25}$$

and $(-T_1, -T_2)$ have Gumbel's (1960) bivariate Type B distribution of extreme values. The marginal survivor functions have the form $S_1(t_1) = \exp\{\delta \log B_1(t_1)\}$ Note that if B has a proportional hazard property only the coefficient of proportionality is changed in formula (25).

The foundation for Model 1 is that in view of the central limit theorem it has a close analogy with normal random effects model. On the other hand, the gamma form is both flexible in shape and tractable for the purpose of describing the common random component.

Model 2. In this case one considers the class of frailty distributions with a constant cross-ratio function $\theta(v) = \theta$. Solving for the differential equation corresponding to (24) gives Clayton's (1978) original model with $q(v) = (1/v)^{\theta-1} - 1$, with $\theta > 1$, giving the Laplace transform of w, $\mathcal{L}_w(u) = p(u) = (1 + u)^{-1/(\theta-1)}$. This corresponds to a gamma$(1/(\theta - 1))$ random variable, a class of heterogeneity distributions widely used in econometric (univariate) duration literature.

[15]The following derivations are presented by Oakes in Example 5 of his 1989 paper.

For the bivariate model (22) the case $\theta \to 1$ corresponds to independence between T_1 and T_2, and as $\theta \to \infty$, the Fréchet upper bound $S(t) = \min\{S_1(t_1), S_2(t_2)\}$ is obtained. If $\theta > 1$ the corresponding joint survivor function is

$$S(t) = \left[(1/S_1(t_1))^{\theta-1} + (1/S_2(t_2))^{\theta-1} - 1\right]^{-1/(\theta-1)}, \tag{26}$$

where the marginal survivor functions have the form $S_1(t_1) = (1 - \log B_1(t_1))^{-1/(\theta-1)}$.

If $\theta < 1$ the corresponding joint survivor function exhibits negative association between T_1 and T_2[16]

$$S(t) = \left[\max\{(S_1(t_1))^{1-\theta} + (S_2(t_2))^{1-\theta} - 1, 0\}\right]^{1/(1-\theta)}, \tag{27}$$

and the support depends on θ. As $\theta \to 0$, one obtains the Fréchet lower bound $S(t) = \max\{S_1(t_1) + S_2(t_2) - 1, 0\}$, a singular distribution concentrated on the curve, $S_1(t_1) + S_2(t_2) = 1$.

The above models are particularly attractive to examine cases where several separate durations, say spells of unemployment are observable on identifiable individuals. In the present study the model has to confined to a less informative case of a competing risks model where the duration variables have a latent variable interpretation and only the minimum duration is actually observed.

4 Competing risks with dependent duration channels

Below the previous model of job recruitment is assessed in light of estimation exercises utilizing a competing risks framework. A common theme in the subsequent analysis is the association among observable variables induced by an unobservable latent variable. Here the underlying dependence is interpreted in terms of an endogenous advertisement decision by the firms affecting the recruitment channel. An additional feature of duration data is the possibility of censoring. The firm whose market environment affecting price, sales or other characteristic change may adjust its labour demand in such a way that a previously announced vacancy is withdrawn.

The data concern vacancies reported to employment offices in the province of Uusimaa, Finland in 1989.[17] Vacancies are for a homogenous category of employees, with upper secondary or lower level of high education in technology. In the data (2531 observations) two exit channels are considered. The first is recruitment through employment offices (30 per cent of the cases) and the second, recruitment through other channels

[16] However, this survivor function is not generated by a frailty distribution since no proper distribution is defined by the corresponding Laplace transform, cf. footnote 13.

[17] The province is the most populated part in Finland, and the capital, Helsinki, is situated there. The data have been provided by the Ministry of Employment, and it has been previously used by J. Rantala. I thank them for giving me access to the data.

(45 per cent), and the mean duration in the data is 46 days. The exits through the latter category of channels are here taken to reflect the endogenous recruitment efforts by the firms.[18] The remaining vacancies (25 per cent) were withdrawn by the employers from the employment office registers. Observations in the last category are considered as censored in the estimations. In most cases these consist of observations where there exists no information on the recruitment channel.

The duration may end in, say m alternative ways which are called exit channels. Both the length of duration and the label of the corresponding exit channel are observed. A competing risks model for duration is obtained by defining m, possibly mutually dependent random variates y_j^*, $j = 1, \ldots, m$, and setting the observed duration y,

$$y \equiv \min\{y_1^*, \ldots, y_m^*, c\}. \tag{28}$$

The censoring variable c is independent of all y^*'s, and has density function ψ and survivor function Ψ.

Oakes considers in his work bivariate frailty models when both y_1^*, and y_2^* are observable. Here we have a competing risks framework and only $\min\{y_1^*, \ldots, y_m^*, c\}$, and the channel indicators $I(j), j = 1, \ldots, m$ are observed. The loglikelihood of an individual observation, $(y_k, I_k(j)|X_k(j); j = 1, \ldots, m)$, $k = 1, \ldots, n$, where $I_k(j)$ is an indicator function for a completed, uncensored duration through exit channel j, and $X_k(j)$ is a vector of explanatory variables, is given by

$$\begin{aligned} \ell_k &\equiv \ell(\theta|I_k(j), y_k, X_k(j)) \\ &= \sum_{j=1}^{m} I_k(j) \left[\log D_j S((y_k, \ldots, y_k); X_k, \theta) + \log \Psi_k(y_k)\right] \\ &+ \left(1 - \sum_{j=1}^{m} I_k(j)\right) \left[\log S((y_k, \ldots, y_k); X_k, \theta) + \log \psi_k(y_k)\right]. \end{aligned} \tag{29}$$

where D_j denotes the operator $-\partial/\partial y_j$ applied to the function $S(y_1, \ldots, y_m)$. The partial hazard functions are defined by $\lambda_j = D_j \log S$ for each individual exit channel.

In the above formulation the censoring mechanism may vary between the observations. The coefficients $\Psi_k(y_k)$, and $\psi_k(y_k)$ may be regarded as nuisance parameters which do not depend on the direct objects of interest, θ. Therefore it suffices to maximize that part of the loglikelihood which involves the parameters of interest.

$$\ell_k(\theta) = \sum_{j=1}^{m} I_k(j) \log\left[\lambda_j((y_k, \ldots, y_k); X_k, \theta)\right] + \log S((y_k, \ldots, y_k); X_k, \theta). \tag{30}$$

[18] Some contamination is naturally present in the data. For example the second channel includes recruitment by direct contact which is only in part affected by employer's effort. However, more elaborate identification of the various search channels was not possible in the data. The stigma effect resulting from a signalling equilibria of the applicants, see Barron and Mellow (1989) is not relevant to the present data since the occupations do not have high educational requirements and the public employment offices are reasonably successful as a recruitment channel.

4.1 Models with frailties

The first two of econometric models are based on frailty Models 1 and 2, with Weibull baseline survivor functions. The baseline duration t, $t \geq 0$, has the survivor function

$$B(t; X^T\beta) = \exp\{-t^\alpha e^{-X^T\beta}\}. \tag{31}$$

The expression $\exp\{-X^T\beta\}$ defines the proportional hazard component of the model. Here X is a column vector of k explanatory variables[19] and β is a vector of parameters to be estimated together with a shape parameter α. For the purposes of this study it is convenient to work with the variable $y = \log t$.[20] It is well known that y follows the type I extreme value distribution, $y \sim EV(X^T\beta, \alpha)$, with the survivor function

$$B(y; X^T\beta) = \exp\{-e^{\alpha y - X^T\beta}\}, \tag{32}$$

The variable, $u = \alpha y - X^T\beta$ is a standardized disturbance term in the sense that $\mathcal{E}u = \psi(0)$, and $Var(u) = \psi'(0)$.[21]

Introduce extreme value forms (32) for survivor functions in (25), after some simplifications one can write the loglikelihood of Model 1, the observation index k is dropped for convenience,

$$\begin{aligned} \ell(\theta) &= \sum_{j=1}^{2} I_j \left[\log\delta + \log\alpha_j + \alpha_j y - X_j^T\beta_j + (\delta - 1)\log\sum_{j=1}^{2}\exp\{\alpha_j y - X_j^T\beta_j\}\right] \\ &- \left[\sum_{j=1}^{2}\exp\{\alpha_j y - X_j^T\beta_j\}\right]^{\delta}. \end{aligned} \tag{33}$$

[19]The following explanatory variables are used, dummy variables accounting for occupation (10 categories), industry (5 categories), type of work (permanent/temporary), worktime (regular/shift work), and region (3 categories). In addition U/V -ratios, i.e. the number of unemployed divided by the number of vacancies, are calculated separately for each sub-region. This continuous variable is used in the analysis to capture the effects of local labour market conditions. To control for the aggregate level of advertising effort we use the monthly vacancy advertisement volume in Helsingin Sanomat, by far the dominant newspaper in the area.

[20]To study the dependence between the observable hazards of the two channels a competing risks Weibull duration model was estimated subject to the restriction that the shape parameters of the two exit channels are equal to get a close resemblance with the theoretical model. The estimated logarithms of proportional hazards, m_j, j = 1, 2 are calculated for each observation. If the true hazards μ_j, j = 1,2, are dependent on each other, one would expect to see some correlation between the m_j's where the latter variables are interpreted as loglinear approximations of the true form of the hazard. The dependence was found to be a near linear one. In the examination allowance is made for the fact that estimates are used instead of the true parameters (details available on request).

[21]Disturbance term has a moment generating function, $\varphi(s) = \Gamma(1+s)$, and ψ refers to the psi function, $\psi(x) = d\log\Gamma(1+x)/dx$. The higher order cumulants of u are obtained through the derivatives of the psi function at zero, $\psi(0) = -\gamma_0$ (Euler's constant), and $Var(u) = \psi'(0) = \pi^2/6$, Abramovitz and Stegun (1970). The distribution is an asymptotic distribution of extreme order statistics. Therefore an argument could be made that the distribution of durations really should be Weibull in the present circumstances.

In fact the model can be reparametrized in a more convenient way for estimation (Appendix A). The maximum likelihood estimates of Model 1 converge to a value $\delta = 0$ in the data. This represents an extreme case of positive dependence between the channels with the joint survivor function obtaining the upper bound $S(y_1, y_2) = \min\{S_1(y_1), S_2(y_2)\}$. However, as $\delta \to 0$, the loglikelihood of the above model tends to the loglikelihood of a model where duration is governed by a common Weibull distribution, and the choice between the channels is determined by a logit probability model. This limit case may be useful in some applications to describe markets with adverse selection and separating equilibria in types of agents. However, in our data the highest value for the loglikelihood is obtained for a degenerate case where the upper bound is obtained with the condition, $u_1 = u_2$, where $u_j = \alpha_j y - X\beta_j$, $j = 1, 2$, holding for all observations (Appendix A). In this case the identification of separate channels is lost. In Model 2

$$\begin{aligned}\ell(\theta) &= \sum_{j=1}^{2} I_j \left[\log\delta + \log\alpha_j + \alpha_j y - X_j^T\beta_j - \log\left(1 + \sum_{j=1}^{2}\exp\{\alpha_j y - X_j^T\beta_j\}\right)\right] \\ &- \delta\log\left[1 + \sum_{j=1}^{2}\exp\{\alpha_j y - X_j^T\beta_j\}\right], \qquad (34)\end{aligned}$$

where $\delta = 1/(\theta^* - 1) > 0$, see Appendix B.

Model 2 represents a bivariate Burr distribution. In the univariate case it is a more familiar one to economists and leads to the extended logistic (or Burr type XII) duration model. The Burr distribution can be succesfully fitted to almost any set of unimodal data and it has been widely used in the literature to control for unobservable heterogeneity, see Tadikamalla (1980). In this paper the frailties are utilized to extend the setup to include the case of separate and correlated duration channels. A nice interior solution is obtained for the maximum likelihood estimation. If the likelihoods are compared Model 2 seems to perform better than the reparametrized form of Model 1 which maintains the separate identification of the channels.

Table I presents the estimation results for the individual parameters of Model 2.[22] There is significant positive association between the durations with the estimated value for cross-ratio equal to 2.3. The structure of the frailty models allows only for associated random variables. The implied parameter value for our gamma mixing variable is 0.78 and its density is a decreasing one and tends to infinity at zero. Returning to the earlier 'theory discussion' this may imply that traces of the (possibly) decreasing part of the function h are lost in the smoothing with the frailty distribution. However, the discussion in the start of the paper broadly suggests that the coefficients relating to channel two should be closer to zero in absolute value. At first glance this seems to hold. One can find cases where a coefficient referring to an easily filled vacancy in channel one is associated with a relative reduction in size in channel two, but the pattern is far from a consistent one. In particular, the coefficients referring to the occupational status and especially

[22]The results of Model 1 are not reported here. The asymptotic normality of estimated parameters should hold reasonably well since one utilizes maximum likelihood estimation within a location family of distributions with a strongly unimodal density.

those of industry dummies are more difficult to interpret since here the current waiting costs may differ due to differences in worker productivity.

If the parameters are compared between the channels the coefficients of the regional variables suggest that channel two represents more integrated labour markets with less regional variance. Similarly, the U/V -ratio has less influence on the exit through channel two. Total advertisement level shortens the vacancy durations in both channels implying that the variable which is constructed at the monthly level captures the seasonal component in employer search. The effect is a more moderate one in channel two. These observations may reflect that advertisement is adjusted on individual conditions with an off-setting effect.

Figure 2 shows the forms of the partial hazard functions with the curves referring to channel two lying highest. The baseline proportional hazard components have been calculated at the mean values of the explanatory variables in channels one and two, respectively. The dotted curves refer to calculations where the positive association between the channels has not been accounted for.

4.2 The model with dependent lognormal variables

The third econometric model is based on a natural extension of a bivariate lognormal model to a duration framework. In this case the proportional hazard property in baseline survivor function is lost. However, the model is well-known and useful for comparison purposes. Furthermore, it has the random effects interpretation if the the durations share a common unobservable variable affecting the mean duration. In addition, the model allows for negative dependence in contrast to the models that are based on frailties.

Here the underlying latent durations have lognormal distributions, where the logarithmic durations y_j, j = 1,2, have the marginal distribution functions,

$$f(y_j; X_j^T b_j) = \frac{1}{\sigma_j}\phi\left(\frac{y_j - X_j^T b_j}{\sigma_j}\right), \tag{35}$$

j = 1,2. The variables y are correlated, with the correlation coefficient ρ, $-1 < \rho < 1$.

To obtain close resemblance with the models used earlier in the paper reparametrisize the model by setting $\alpha_j = 1/\sigma_j$, $\beta_j = (1/\sigma_j)b_j$, and finally for the convenience of estimation $r = \rho/(1-\rho^2)^{\frac{1}{2}}$ with $-\infty < r < \infty$, and $\rho = r/(1+r^2)^{\frac{1}{2}}$. Now one can write the bivariate survivor function $S(t) = S(t_1, t_2) = S(y_1, y_2)$,

$$\begin{aligned} S(y_1, y_2) &= \alpha_1\alpha_2\sqrt{1+r^2}\int_{y_1}^{\infty}\int_{y_2}^{\infty}\phi\left(\alpha_1 v_1 - X_1^T\beta_1\right) \\ &\times \phi\left(\sqrt{1+r^2}(\alpha_2 v_2 - X_2^T\beta_2) - r(\alpha_1 v_1 - X_1^T\beta_1)\right) dv_1 dv_2 \end{aligned} \tag{36}$$

which can be easily computed.

To obtain the cross-ratio function (23) calculate $D_j S(y, y)$, where D_j denotes the operator $-\partial/\partial y_j$ applied to the function $S(y_1, y_2)$. For example,

$$D_1 S(y, y) = \alpha_1 \phi(u_1)\left[1 - \Phi\{u_2\sqrt{1+r^2} - r u_1\}\right], \tag{37}$$

where $u_j = \alpha_j y - X_j^T \beta_j$, j = 1,2, define the two standardized disturbance terms of the model.[23]

The estimation results are given in Table II and they are qualitatively similar to those given by Model 2 (Table I). In addition the loglikelihoods are on closely similar levels. The estimate of the parameter r implies a positive value for the correlation coefficient equal to 0.40. The variance for the latent logarithmic duration through channel two is less than half of that of channel one suggesting that variance reducing adjustment of effort level may be operating here. The random effects interpretation implies that the common unobservable variable accounts for about 26 and 62 per cent of total variation in channels one and two, respectively.

5 Conclusion

The discussion about the effects of endogenous adjustment of recruitment effort is in the paper used as a starting point to analyze the data. Unfortunately, we have not been able to present specific parametric restrictions and tests to support the qualitative implications of the theoretical discussion. Instead, the data have confined the analysis to interpretation of individual coefficients with possibly introducing post-sample bias. However, it is felt that the basic idea merits a more careful analysis and further evalution with possibly alternative data to get a feel of the quantitative significance of the qualitative implications of theory.

Three models in the competing risks framework have been estimated with dependent latent durations. The statistical models generated by mixtures with stochastic proportional hazard components should find more extensive use in labour econometrics. In the paper the parameters are estimated by solving numerically the nonlinear estimating equations given by the principle of maximum likelihood. On the other hand, since frailty models are constructed with the marginals as parameters they are amendable to semiparametric analysis. Furthermore, the methods are immediately applicable to situations where for example several spells of unemployment are observable. In this case the conditions for indentifying the models are far less restrictive than in the case of competing risks, and they may offer additional tools to obtain insight to the functioning of complicated real life labour markets.

[23] In the present case the formula for the cross-ratio function θ^* is far less transparent than in Models 1 or 2 (for details, see Appendix C). Furthermore, since Model 3 is not a member of the class of models induced by frailties, the cross-ratio function is dependent on both y_1 and y_2. In Figure 3 the cross-ratio is represented as a function of u_1 in the case of selected values of u_2.

References

Abramowitz, M. and Stegun, I.A. (1970), *"Handbook of Mathematical Functions"*, *Dover Publications*, New York.

Barron, J.M. and Mellow, W. (1989), "Labor contract formation, search requirements, and use of a public employment service," *Economic Inquiry*, 20, 381-387.

Blanchard, O.J. and Diamond P.A. (1989), "The Beveridge curve," *Brookings Papers on Economic Activity*, (1), 1-60.

Clayton, D.G. (1978), "A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease tendency," *Biometrika*, 65, 141-151.

Genest, C. and MacKay, R.J. (1986), "The joy of copulas: bivariate distributions with given marginals," *The American Statistician*, 40, 280-283.

Gumbel, E.J. (1960), "Bivariate exponential distributions," *Journal of the American Statistical Association*, 55, 698-707.

Hougaard, P. (1984), "Life table methods for heterogenous populations: distributions describing the heterogeneity," *Biometrika*, 71, 75-83.

Kendall, M.G. (1938), "A new measure of rank correlation," *Biometrika*, 30, 81-93.

Lehmann, E.L. (1966), "Some concepts of dependence," *Annals of Mathematical Statistics*, 37, 1137-1153.

Marshall, A.W., and Olkin I. (1979), *Inequalities: Theory of Majorization and Applications*, Academic Press, New York.

Marshall, A.W., and Olkin I. (1988), "Families of multivariate distributions," *Journal of the American Statistical Association*, 83, 834-841.

Rudin, W. (1973), *"Functional Analysis,"* *McGraw-Hill, Inc.*, New York.

Oakes, D. (1989), "Bivariate survival models induced by frailties," *Journal of the American Statistical Association*, 84, 487-493.

Tadikamalla, P.R. (1980), "A look at the Burr and related distributions," *International Statistical Review*, 48, 337-344.

van Ours, J. (1989), "Durations of Dutch job vacancies," *De Economist*, 137, 309-327.

van Ours, J. and Ridder, G. (1991), "Job requirements and the recruitment of new employees," *Economics Letters*, 36, 213-218.

van Ours, J. and Ridder, G. (1992), "Vacancies and the recruitment of of new employees," *Journal of Labor Economics*, 10, 138-155.

Vaupel, J.W., Manton, K.G. and Stallard, E. (1979), "The impact of heterogeneity in individual frailty and the dynamics of mortality," *Demography*, 16, 439-454.

A Appendices

The joint survivor function of Model 1 is

$$-\log S(t) = \left[(-\log S_1(t_1))^{1/\delta} + (-\log S_2(t_2))^{1/\delta}\right]^{\delta}, \tag{38}$$

where $-\log S_j(t_j) = \exp\{\alpha_j y_j - X_j^T \beta_j\}$, $y_j = \log t_j$, j = 1,2, and $0 < \delta < 1$. Let $u_j = \alpha_j y_j - X^T \beta_j$, j = 1,2. For the derivations below it is convenient to write[24]

$$-\log S(y) = e^{\max\{u_1,u_2\}}\left[1 + e^{-\frac{1}{\delta}||u_2-u_1||}\right]^{\delta}, \tag{39}$$

Next differentiate, to get

$$\begin{aligned}
\frac{-\partial \log S(y)}{\partial y_1} &= \alpha_1 e^{\max\{u_1,u_2\}}\left[1 + e^{-\frac{1}{\delta}||u_2-u_1||}\right]^{\delta-1} e^{-\frac{1}{\delta}I||u_2-u_1||}, \\
\frac{-\partial \log S(y)}{\partial y_2} &= \alpha_2 e^{\max\{u_1,u_2\}}\left[1 + e^{-\frac{1}{\delta}||u_2-u_1||}\right]^{\delta-1} e^{-\frac{1}{\delta}(1-I)||u_2-u_1||}.
\end{aligned} \tag{40}$$

The loglikelihood of an individual observation, $(y, I_j|X; j = 1,2)$, where I_j is an indicator function for a completed duration through channel j, is given by

$$\begin{aligned}
\ell_\delta(\theta) &= I_1\left[\log\alpha_1 + \max\{u_1,u_2\} - \frac{1}{\delta}I||u_2-u_1|| + (\delta-1)\log\left[1 + e^{-\frac{1}{\delta}||u_2-u_1||}\right]\right] \\
&+ I_2\left[\log\alpha_2 + \max\{u_1,u_2\} - \frac{1}{\delta}(1-I)||u_2-u_1|| + (\delta-1)\log\left[1 + e^{-\frac{1}{\delta}||u_2-u_1||}\right]\right] \\
&- e^{\max\{u_1,u_2\}}\left[1 + e^{-\frac{1}{\delta}||u_2-u_1||}\right]^{\delta}.
\end{aligned} \tag{41}$$

Above both u_1 and u_2 are taken at the point $y_1 = y_2 = y$. Reparametrize the model by setting $a = (1/\delta)(\alpha_2 - \alpha_1)$, and $\gamma = (1/\delta)(\beta_2 - \beta_1)$. With this notation $u_2 - u_1 = \delta(ay - X^T\gamma)$, and $I = I\{ay - X^T\gamma > 0\}$. Let $v_2 = ay - X^T\gamma$. Now one can write

$$\begin{aligned}
\ell_\delta &= I_1\left[\log\alpha_1 + u_1 + \delta I v_2 - I||v_2|| + (\delta-1)\log\left[1 + e^{-||v_2||}\right]\right] \\
&+ I_2\left[\log(\alpha_1 + \delta a) + u_1 + \delta I v_2 - (1-I)||v_2|| + (\delta-1)\log\left[1 + e^{-||v_2||}\right]\right] \\
&- e^{u_1}\left[1 + e^{v_2}\right]^{\delta}.
\end{aligned} \tag{42}$$

For a fixed $\delta, 0 < \delta < 1$, the function (42) is concave in the estimable parameters.[25]

[24]To keep the R.H.S of (39) bounded, as $\delta \to 0$. Note that $\max\{u_1,u_2\} = u_1 + I(u_2 - u_1)$, and $||u_2 - u_1|| = I(u_2 - u_1) + (1 - I)(u_1 - u_2)$, where $I = I\{u_2 > u_1\}$.

[25]To see this establish first that the function is concave w.r.t. u_1 and v_2. The result follows by observing that these two 'residuals' are linear in the parameters and since $\log\alpha_1$, and $\log(\alpha_1 + \delta a)$ are concave in the parameters.

The above formula shows that as $\delta \longrightarrow 0$, the loglikelihood tends to the limit

$$\begin{aligned}
\ell_0 &= I_1 \left[\log \alpha_1 + u_1 + \log\left(\frac{e^{-I\|v_2\|}}{1+e^{-\|v_2\|}}\right)\right] \\
&+ I_2 \left[\log \alpha_1 + u_1 + \log\left(\frac{e^{-(1-I)\|v_2\|}}{1+e^{-\|v_2\|}}\right)\right] - e^{u_1} \\
&= I_1 \left[\log \alpha_1 + u_1 - \log\left(1+e^{v_2}\right)\right] \\
&+ I_2 \left[\log \alpha_1 + u_1 - \log\left(1+e^{-v_2}\right)\right] - e^{u_1}. \qquad (43)
\end{aligned}$$

An interesting observation is that as $\delta \longrightarrow 0$, the loglikelihood of the model tends to the loglikelihood of the model where duration is governed by a common Weibull model, and the choice between the channels is determined by a logit probability model with log-odds [26]

$$\log \frac{Pr\{I_2 = 1\}}{Pr\{I_1 = 1\}} = v_2. \qquad (44)$$

On the other hand, the limit model for survivor function is $S(y) = \min\{S(y_1), S(y_2)\}$. The direct derivation of the limit model gives the loglikelihood[27]

$$\begin{aligned}
\ell_c &= I_1 \left[\log \alpha_1 + \max\{u_1, u_2\} + \log\left[1 - I + \delta_I(u_2 - u_1)\right]\right] \\
&+ I_2 \left[\log \alpha_2 + \max\{u_1, u_2\} + \log\left[I + \delta_I(u_2 - u_1)\right]\right] - e^{\max\{u_1,u_2\}}. \qquad (45)
\end{aligned}$$

In the model (45) the loglikelihood is bounded from below only if $\{I(j) = 1\} \subseteq \{u_j = \max\{u_1, u_2\}\}$, for j = 1,2. Therefore, maximization entails that the above condition holds and this is equivalent with $\{I(j) = 1\} \Rightarrow S(y_j) = \min\{S(y_1), S(y_2)\}$. A special case is that $u_1 = u_2$ holds over the whole data set.[28] Note that (43) shows that the above limit condition can be obtained through our reparametrization. However, the reparametrization produces an additional log-odds component (44) always having a lower value of the loglikelihood than the previous degenerate special case where separate identification of the channels is lost.

In maximum likelihood estimation of the reparametrized model (42) one may use

$$\begin{aligned}
\frac{\partial \ell_\delta}{\partial \alpha_1} &= \frac{I_1}{\alpha_1} + \frac{I_2}{\alpha_1 + \delta a} + \left[I_1 + I_2 - e^{u_1}\left[1 + e^{v_2}\right]^\delta\right] y, \qquad (46) \\
\frac{\partial \ell_\delta}{\partial \beta_1} &= -\left[I_1 + I_2 - e^{u_1}\left[1 + e^{v_2}\right]^\delta\right] X, \qquad (47)
\end{aligned}$$

[26]This property is somewhat related to McFadden's (1971) result, where he shows that independent latent variables, u_1 and u_2 that follow type I extreme-value distribution generate logit choice probabilities, $P_j = Pr\{u_j = \max\{u_1, u_2\}\}$, j = 1,2.

[27]Here we extend differential calculus by considering distributions, i.e. integrable functionals that satisfy a functional equation, see for example Rudin (1973). In this case the 'derivative' of the Heaviside function, $I\{x \geq a\}$ is equal to the delta function δ_a, $\delta_a(x) = 1$, if $x = a$, and zero otherwise.

[28]In the data no parameter estimates could be found to satisfy the above condition with some u_j's unequal.

$$\frac{\partial \ell_\delta}{\partial a} = \left(\frac{\delta}{\alpha_1+\delta a}+y\right) I_2 + (\delta-1)\left[I_1+I_2\right]\left[I+(1+e^{\|v_2\|})^{-1}(1-2I)\right] y$$
$$- \quad \delta e^{u_1}\left[1+e^{v_2}\right]^\delta \left[1+e^{-v_2}\right] y, \tag{48}$$
$$\frac{\partial \ell_\delta}{\partial \gamma} = -\left[I_2+(\delta-1)\left[I_1+I_2\right]\right]\left[I+(1+e^{\|v_2\|})^{-1}(1-2I)\right] X$$
$$+ \quad \delta e^{u_1}\left[1+e^{v_2}\right]^\delta \left[1+e^{-v_2}\right] X, \tag{49}$$
$$\frac{\partial \ell_\delta}{\partial \delta} = \frac{aI_2}{\alpha_1+\delta a} + \left[I_1+I_2\right]\left[Iv_2+\log(1+e^{-\|v_2\|})\right] - e^{u_1}\left[1+e^{v_2}\right]^\delta \log\left[1+e^{v_2}\right]. \tag{50}$$

B

The joint survivor function of Model 2 is

$$-\log S(t) = \frac{1}{\theta-1}\log\left[1-\log B_1(t_1)-\log B_2(t_2)\right], \tag{51}$$

where $-\log B_j(t_j) = \exp\{\alpha_j y_j - X^T\beta_j\}$, and $y_j = \log t_j$, j = 1,2. Let $u_j = \alpha_j y_j - X^T\beta_j$, j = 1,2, and $\delta = (\theta-1)^{-1} > 0$. Differentiate, to get

$$\frac{-\partial \log S(y)}{\partial y_j} = \alpha_j \delta e^{u_j}\left[1+e^{u_1}+e^{u_2}\right]^{-1}, \tag{52}$$

for j = 1,2.

The loglikelihood for an individual observation, $(y, I_j|X; j = 1,2)$, where $y = \log(\min\{T_1,T_2\})$, and I_j is an indicator function for a completed, uncensored duration through exit channel j, is given by

$$\ell_\delta = I_1\left[\log\delta+\log\alpha_1+u_1-\log\left[1+e^{u_1}+e^{u_2}\right]\right]$$
$$+ \quad I_2\left[\log\delta+\log\alpha_2+u_2-\log\left[1+e^{u_1}+e^{u_2}\right]\right]$$
$$- \quad \delta\log\left[1+e^{u_1}+e^{u_2}\right]. \tag{53}$$

Similarly as above one can show that (53) is a concave function in the estimable parameters.

In maximum likelihood estimation of the parameters in (53) one may use the gradient formulae

$$\frac{\partial \ell_\delta}{\partial \alpha_j} = I_j\left[\frac{1}{\alpha_j}+\left(1-\frac{e^{u_j}}{1+e^{u_1}+e^{u_2}}\right)y\right] - \frac{\delta e^{u_j} y}{1+e^{u_1}+e^{u_2}}, \tag{54}$$
$$\frac{\partial \ell_\delta}{\partial \beta_j} = -\left[I_j\left(1-\frac{e^{u_j}}{1+e^{u_1}+e^{u_2}}\right) - \frac{\delta e^{u_j}}{1+e^{u_1}+e^{u_2}}\right] X, \tag{55}$$
$$\frac{\partial \ell_\delta}{\partial \delta} = \frac{1}{\delta}\left[I_1+I_2\right] - \log\left[1+e^{u_1}+e^{u_2}\right]. \tag{56}$$

for j = 1,2.

C

The logarithmic durations y_j, j = 1,2, are normally distributed and have the marginal distribution functions,

$$f(y_j; X_j^T b_j) = \frac{1}{\sigma_j}\phi\left(\frac{y_j - X_j^T b_j}{\sigma_j}\right), \tag{57}$$

j = 1,2. The variables y_j are correlated, with the correlation coefficient ρ, $-1 < \rho < 1$. Reparametrisize the model by setting $\alpha_j = 1/\sigma_j$, $\beta_j = (1/\sigma_j)b_j$, and finally for the convenience of estimation $r = \rho/(1-\rho^2)^{\frac{1}{2}}$. Now we can write the bivariate survivor function $S(t) = S(t_1, t_2) = S(y_1, y_2)$,

$$\begin{aligned} S(y_1, y_2) &= \alpha_1\alpha_2\sqrt{1+r^2}\int_{y_1}^{\infty}\int_{y_2}^{\infty}\phi\left(\alpha_1 v_1 - X_1^T\beta_1\right) \\ &\times \phi\left(\sqrt{1+r^2}(\alpha_2 v_2 - X_2^T\beta_2) - r(\alpha_1 v_1 - X_1^T\beta_1)\right) dv_1 dv_2. \end{aligned} \tag{58}$$

Differentiate to get

$$\frac{-\partial S(y,y)}{\partial y_j} = \alpha_j\phi(u_j)\left[1 - \Phi\left(u_{\bar{j}}\sqrt{1+r^2} - ru_j\right)\right], \tag{59}$$

where $u_j = \alpha_j y - X_j^T\beta_j$, j = 1,2, define the two standardized disturbance terms of the model. With the new notation the survivor function can be written

$$\begin{aligned} S &= \sqrt{1+r^2}\int_{-\infty}^{-u_1}\int_{-\infty}^{-u_2}\phi(\tau_1)\,\phi\left(\tau_2\sqrt{1+r^2} - r\tau_1\right) d\tau_1 d\tau_2 \\ &= \int_{-\infty}^{-u_1}\phi(\tau_1)\,\Phi\left(\sqrt{1+r^2}(-u_2) - r\tau_1\right) d\tau_1. \end{aligned} \tag{60}$$

The loglikelihood of an individual observation (y, I_j), j=1,2 is

$$\begin{aligned} \ell_r &= \sum_{j=1}^{2} I_j\left[\log\alpha_j + \log\phi(u_j) + \log\Phi\left(\sqrt{1+r^2}(-u_{\bar{j}}) - ru_j\right)\right] \\ &= \left(1 - \sum_{j=1}^{2} I_j\right)\log\int_{-\infty}^{-u_1}\phi(\tau_1)\,\Phi\left(\sqrt{1+r^2}(-u_2) - r\tau_1\right) d\tau_1, \end{aligned} \tag{61}$$

where $\bar{j} \neq j$, j = 1,2.

In maximum likelihood estimation of Model 3 (59) one may use the gradient formulas

$$\begin{aligned} \frac{\partial\ell_r}{\partial\alpha_j} &= I_j\left[\frac{1}{\alpha_j} + \left(-u_1 + \frac{r\phi}{1-\Phi}\right)y\right] - I_{\bar{j}}\frac{y\phi\sqrt{1+r^2}}{1-\Phi} \\ &+ \left(1 - \sum_{j=1}^{2} I_j\right)\frac{y\phi(u_j)\Phi\left(\sqrt{1+r^2}(-u_{\bar{j}}) + ru_j\right)}{S(y_1, y_2)}, \\ \frac{\partial\ell_r}{\partial\beta_j} &= -\left[I_j\left(-u_1 + \frac{r\phi}{1-\Phi}\right) - I_{\bar{j}}\frac{\phi\sqrt{1+r^2}}{1-\Phi}\right] X \end{aligned} \tag{62}$$

$$
\begin{aligned}
& - \left(1-\sum_{j=1}^{2} I_j\right) \frac{\phi(u_j)\Phi\left(\sqrt{1+r^2}(-u_{\bar{j}})+ru_j\right)}{S(y_1,y_2)} X, & (63)\\
\frac{\partial \ell_r}{\partial r} &= -\sum_{j=1}^{2} I_j \left[\frac{ru_{\bar{j}}}{\sqrt{1+r^2}} - u_j\right] \frac{\phi}{1-\Phi} \\
& + \left(1-\sum_{j=1}^{2} I_j\right) \frac{e^{-\frac{1}{2}u_2^2}\phi\left(\sqrt{1+r^2}(-u_1)+ru_2\right)}{(1+r^2)S(y_1,y_2)}. & (64)
\end{aligned}
$$

In the above formulae the expression $\phi/(1-\Phi)$ is taken at the point $(1+r^2)^{\frac{1}{2}}u_{\bar{j}} - ru_j$, if it is associated with I_j, j =1,2.

The formula for cross-ratio function, θ^*,

$$
\theta^*(t) = \frac{S(t)D_{12}S(t)}{D_1S(t)D_2S(t)},
$$

gives

$$
\theta^*(t) = \frac{\sqrt{1+r^2}\phi\left(u_2\sqrt{1+r^2}-ru_1\right)\int_{-\infty}^{-u_1}\phi(\tau_1)\,\Phi\left(\sqrt{1+r^2}(-u_2)-r\tau_1\right)d\tau_1}{\phi(u_2)\Phi\left(\sqrt{1+r^2}(-u_2)+ru_1\right)\Phi\left(\sqrt{1+r^2}(-u_1)+ru_2\right)}. \quad (65)
$$

In the present case the formula for θ^* is far less transparent than in the case of Models 1 or 2. Further, since Model 3 is not a member of the class of models induced by frailties, the cross-ratio function is dependent on both y_1 and y_2. In Figure 3 the cross-ratio function is represented as a function of u_1 in the case of selected values of u_2 using the estimate of the correlation coefficient.

D Tables

TABLE I

Results with Weibull model and Gamma frailty

	Channel 1		Channel 2	
Parameters	Estimates	Est./s.e.	Estimates	Est./s.e.
CONSTANT	0.710828	20.744	0.912039	24.899
Region:				
METROPOL	0.951115	8.073	0.347573	3.119
REST	1.056205	8.466	0.782241	6.402
Industry:				
FOODBEV	-0.084784	-0.218	-1.160932	-3.633
PAPER	-1.139636	-3.810	-1.164215	-3.459
CONSTRUC	-0.619560	-4.648	-0.262983	-1.961
SERVICES	-0.166525	-0.489	-0.338069	-1.084
Occupation in the technical field:				
SUPERVISOR	0.818951	3.587	0.189975	0.869
BUILDER	-0.842481	-4.256	-1.222486	-5.865
FOODBEV	1.300364	2.937	1.096900	3.075
METALWARE	0.637627	2.633	0.103596	0.540
WELDER	-0.481188	-2.236	-0.079745	-0.363
FITTER	-0.708089	-3.349	-0.408694	-1.855
MECHANIC	0.128924	0.650	-0.161166	-0.950
COATING	-0.688885	-2.205	-0.009693	-0.027
PLUMBER	0.235692	1.302	0.263510	1.554
Type of employment:				
TEMPORARY	-1.149821	-5.753	-1.452177	-6.694
SHIFT	0.159423	0.813	0.809063	4.318
Labour market conditions:				
U/V-RATIO	-0.344437	-4.745	-0.274874	-3.554
ADVER VOL	-1.602561	-3.763	-1.047493	-2.574
Other parameters:				
SCALE	1.554549	22.471	2.223473	25.696
LOGDELTA	-0.252421	-2.493		

Mean log-likelihood -1.67589
Number of cases 2531
The constant parameter has been divided by ten.

TABLE II

Results with dependent lognormal variables

	Channel 1		Channel 2	
Parameters	Estimates	Est./s.e.	Estimates	Est./s.e.
CONSTANT	0.365080	21.170	0.479223	29.431
Region:				
METROPOL	0.477766	7.949	0.149391	2.434
REST	0.506175	7.960	0.341954	5.413
Industry:				
FOODBEV	-0.062584	-0.300	-0.613835	-3.585
PAPER	-0.490111	-3.096	-0.545158	-3.307
CONSTRUC	-0.332000	-4.982	-0.080473	-1.108
SERVICES	-0.104328	-0.588	-0.189184	-1.137
Occupation in the technical field:				
SUPERVISOR	0.414786	3.585	0.071627	0.621
BUILDER	-0.377309	-3.626	-0.672013	-6.344
FOODBEV	0.624787	2.749	0.564958	3.000
METALWARE	0.296172	2.455	0.078885	0.782
WELDER	-0.274377	-2.381	0.018515	0.156
FITTER	-0.299412	-2.592	-0.143942	-1.262
MECHANIC	0.075894	0.744	-0.092437	-1.039
COATING	-0.338177	-2.035	0.095135	0.521
PLUMBER	0.105069	1.135	0.131079	1.501
Type of employment:				
TEMPORARY	-0.507202	-4.808	-0.656252	-6.038
SHIFT	0.093784	0.906	0.362440	3.766
Labour market conditions:				
U/V-RATIO	-0.176529	-4.762	-0.121250	-3.068
ADVER VOL	-0.767484	-3.581	-0.474205	-2.307
Other parameters:				
SCALE	0.720212	20.947	1.111453	47.723
R	0.434752	2.451		

Mean log-likelihood -1.67240
Number of cases 2531
The constant parameter has been divided by ten.

FIGURE 1: Marginal effort and waiting costs and optimal effort level

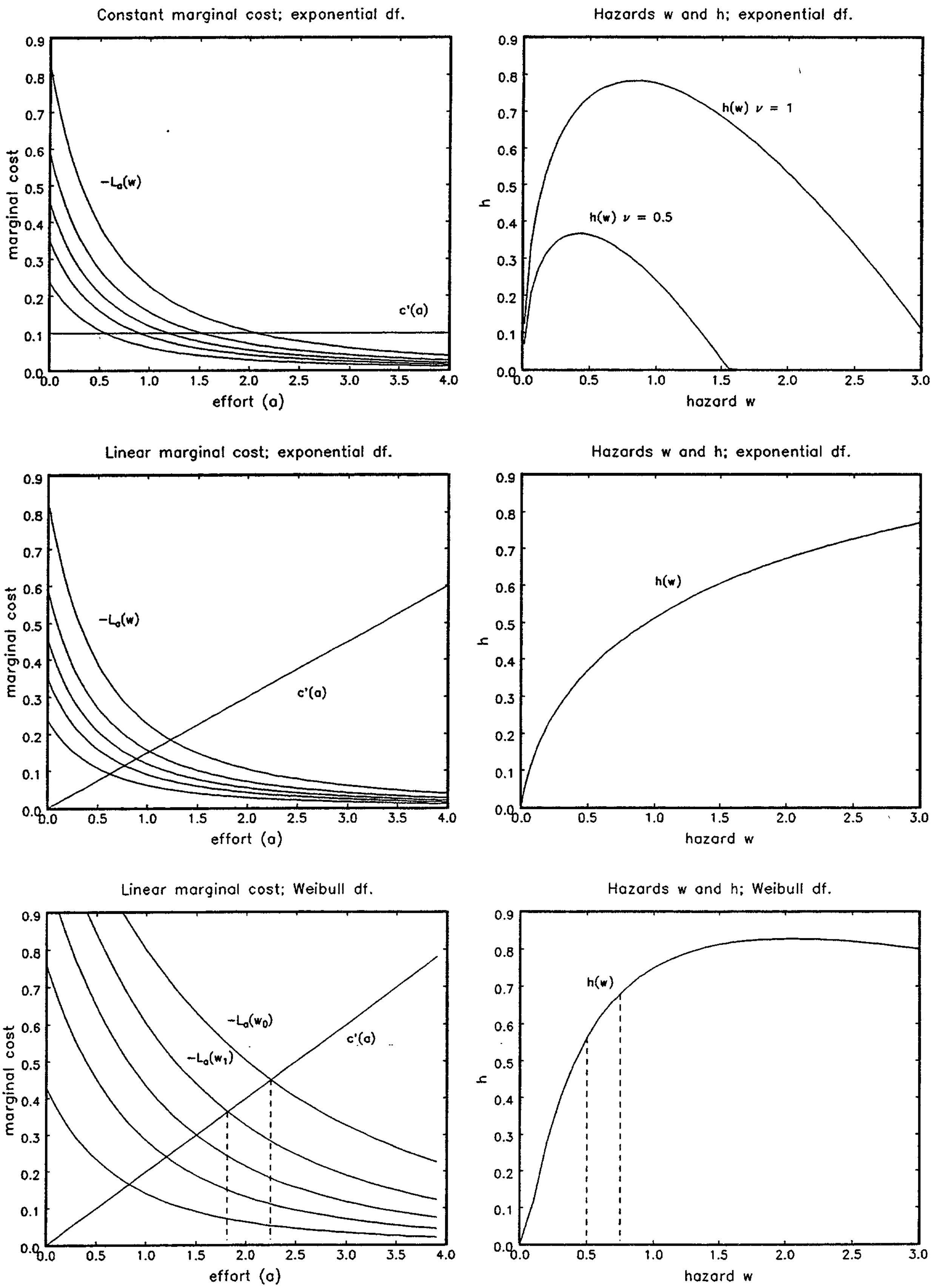

FIGURE 2: Partial hazards by channels

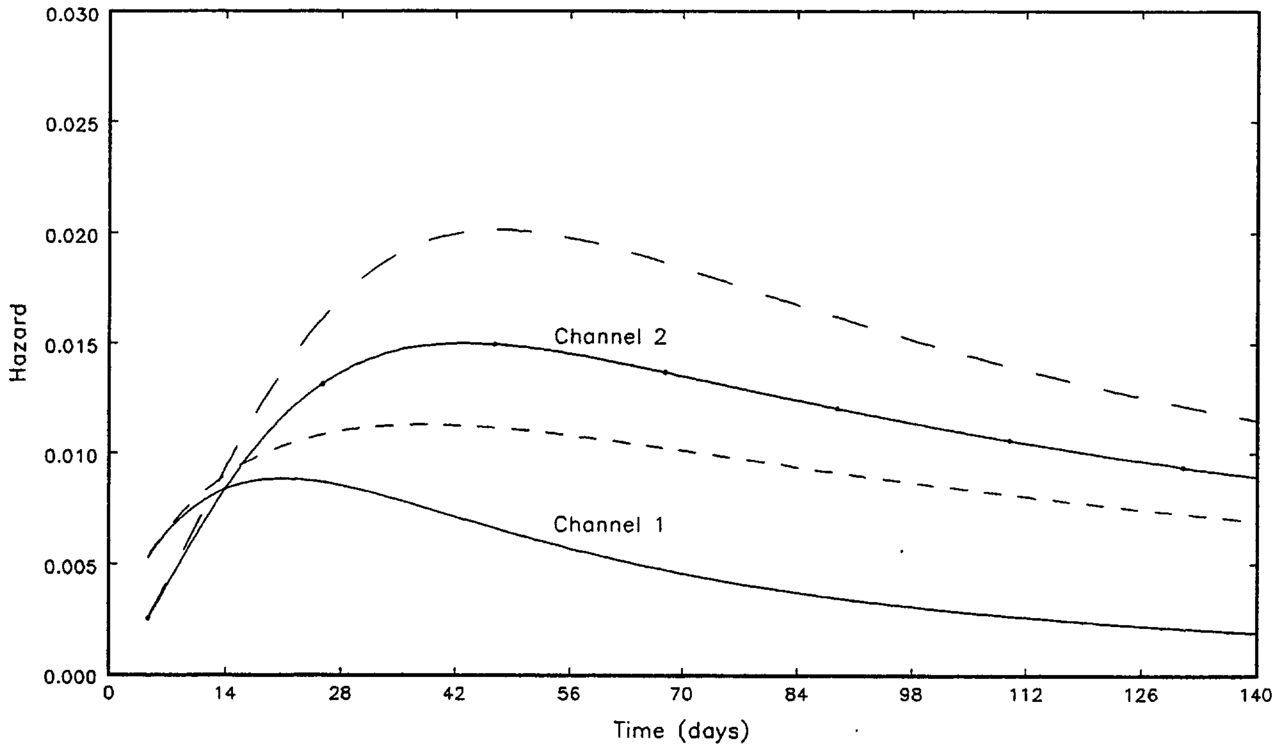

FIGURE 3: Cross ratio function with lognormal durations