Bureau
d'économie
théorique
et appliquée
(BETA)
UMR 7522

Documents
de travail

Documents
de travail

« Partial Likelihood Estimation of a Cox Model With Random Effects : an EM Algorithm Based on Penalized Likelihood »

Auteur

Guillaume HORNY

Document de travail n° 2006–**10**

Avril 2006

Faculté des sciences économiques et de gestion
Pôle européen de gestion et d'économie (PEGE)
61 avenue de la Forêt Noire
F-67085 Strasbourg Cedex

Secrétariat du BETA
Christine Demange
Tél. : (33) 03 90 24 20 69
Fax : (33) 03 90 24 20 70
demange@cournot.u-strasbg.fr
http://cournot.u-strasbg.fr/beta

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITÉ LOUIS PASTEUR STRASBOURG

Partial Likelihood Estimation of a Cox Model With Random Effects: an EM Algorithm Based on Penalized Likelihood

Guillaume Horny*

March 6, 2006

Abstract

The aim of this paper is to present a general EM algorithm to estimate Mixed Proportional Hazard models including more than one random effect, through partial likelihood. We assume only that the mixing distributions admit Laplace transforms. We show how to transform inference in a single complicated model in the estimation of MPH models involving only a single frailty, which are easily manageable. We then face on gamma unobserved heterogeneity. This choice is a weak assumption as the heterogeneity distribution among survivors converges to a gamma distribution, often quickly, for many types of unobserved heterogeneity distributions. The proposed approach can thus be used to estimate a wide class of models. We describe how to use the penalized partial likelihood within the EM algorithm, to improve speed and stability. The behaviour of the estimator on different clusterings and sample sizes is assessed through a Monte Carlo study. We also provide an application on the ratification of ILO conventions by developing countries over the period 1975-1995. Both the simulations and the empirical results indicate an important decrease in computing time. Furthermore, our procedure converges in settings where a standard EM algorithm does not.

Keywords: Random Effects, Duration analysis, Dynamic model

*BETA-*Theme*, University Louis Pasteur (Strasbourg I), 61 avenue de la Forêt Noire, 67085 Strasbourg Cedex, France (email: horny@cournot.u-strasbg.fr). The author thanks Gerard van den Berg and François Laisney for detailed comments on previous versions of this paper. He is also grateful to Bernhard Boockmann for kindly providing the data on the ILO conventions. He would like to thank Narayan Sastry for explanations on his algorithm, as well as participants at the seminar of the Tinbergen Institute, March 2005, Amsterdam and at the seminar of the ZEW, March 2006, Mannheim.

1 Introduction

In this paper, I propose a general framework to estimate mixed proportional hazard models (MPH, see Van den Berg, 2001, for a survey) with a fixed number of risks, where unobserved heterogeneity is modeled at different levels. This class of models describes a population stratified on different criteria, allowing unobserved characteristics located at different levels. If the random effects are nested, the sample is divided in clusters, sub-clusters and so on.

Many econometric studies in duration analysis account for unobserved heterogeneity. Indeed, it is not reasonable to assume that we observe all the determinants leading to a transition, and furthermore that the recorded explanatory variables are free of measurement error. Studies handling a single random effect are nowadays widespread, but unobserved heterogeneity is in general specified at a single level, while the possibility of omitted variables with group structure at different levels arises in several applications. Ignoring some of the unobserved heterogeneity can lead to substantial biases (see Pakes, 1983, Moulton, 1986, and Gouriéroux and Peaucelle, 1990, for some case studies in linear models) but modeling it at different levels is not an easy task and raises some awkward problems for inference. Indeed, it involves multidimensional integrals which typically do not admit analytical expressions.

A few studies in the biometric and demographic literature handle two levels of clustering. Manda and Meyer (2005) consider a model in discrete time, Yau (2001) and Sastry (1997) study two nested random effects with log-normal and gamma mixing distributions, respectively. In this last study, inference is based on the Expectation Maximization (EM) algorithm (see Dempster, Laird and Rubin, 1977, for the first general formulation). The EM algorithm is ideally suited for mixture models, due to their missing data structure. It is used in numerous studies involving the Cox model with one frailty, such as Clayton and Cuzick (1985), Gill (1985) and Parner (1997). However, this theoretical attractivity it balanced by many numerical drawbacks. First, convergence is usually slow, sometimes fails, is sensitive to the choice of the starting values and is time consuming. Indeed, Dempster et al. (1977) show that the algorithm is made of linear iterations and is thus slower than the usual methods, such as Newton-Raphson procedures, involving approximation by a quadratic function at each iteration. Furthermore, the rate of convergence of the algorithm depends on the amount of information in the sample and, as incompleteness of the data is usually large in duration analysis, convergence is typically slow. Second, two different cases of non-convergence are mentioned in the literature. Bolstad and Manda (2001) pointed out one case where the variance of the random effect becomes large enough to raise numerical issues. Lancaster (1990, p. 267) described the case of an unbounded likelihood with a variance of the unobserved heterogeneity tending to zero. Third, Ng, Krishnan and McLachlan (2004) emphasize the importance of the starting values and its consequence on the number of iterations and even convergence when the likelihood is unbounded and the starting values are close to the boundary. Step E can have no analytical solution and the expressions thus require an evaluation using numerical integration or Monte Carlo methods. This Monte Carlo EM (MCEM) algorithm is described in Wei and Tanner (1990), but simulations increase the computational cost and introduce a Monte Carlo error. Fourth, the EM algorithm typically requires a large number of iterations, even in easy models. These drawbacks lead Therneau, Grambsch and Pankratz (2003) to investigate inference in the one-gamma-frailty setting using penalized partial likelihood. They obtain an estimator equivalent to the one provided by the EM algorithm, but free of all these numerical problems.

In this paper, I present a general framework for inference using the EM algorithm in a MPH model involving I random effects ($I \geq 2$), as described in Section 2. This method has the advantage to transform inference in a single complicated model in the estimation of I MPH models, each with a single frailty, which are easily manageable. It is presented in Section 3. As shown in Abbring and Van den Berg (2001), the mixing distribution among survivors converges to a gamma distribution under some conditions. Their result requires the heterogeneity distribution to be regularly varying at 0, as defined in Feller (1971). It is a weak restriction as distributions such as the exponential, uniform, beta and all distributions with a mass point at 0 fulfill it. In Section 4, I thus describe how to make use of the penalized partial likelihood presented by Therneau et al. (2003). That is, we remain in the convenient theoretical framework of the EM algorithm while taking advantage of the numerical stability, simplicity and speed of the penalized likelihood. Using a small Monte Carlo study and data on the timing of ratification of ILO conventions, this method is compared in Sections 5 and 6 to the accelerated EM algorithm described in Sastry (1997).

2 Mixed Proportional Hazard model with random effects

Consider a frailty model belonging to the general family of MPH models, with the particularity that the unobserved heterogeneity term here is written as the product of I random effects, enabling us to handle multilevel clustering.

There can be, for example, two nested random effects and the model is then the one described in Bolstad and Manda (2001), Sastry (1997) and Yau (2001). In the studies cited above, the population is divided into clusters, each cluster is divided into subclusters and several individuals belong to the same subcluster. The underlying idea is that durations are correlated in some way, and a realization of a random effect is common to all observations in the same cluster or subcluster. One can also consider non-nested random effects as in the model described in Horny, Boockmann, Djurdjevic and Laisney (2005). The population is then clustered according to two different criteria which are not necessarily related in a hierarchical way. The study cited above concerns the ratification of ILO conventions and observations are clustered among conventions and among countries.

Consider a population stratified at I levels which are not necessarily nested. Let j ($j = 1, \ldots, J_i$) be the index of the groups defined at level i, and k be the cross-sectional unit index ($k = 1, \ldots, K$). The hazard function is written as:

$$\lambda_{ijk}(t) = \left(\prod_{i=1}^{I} v_{ij} \right) \lambda_0(t) \lambda_1[X_k(t), \beta], \tag{1}$$

where v_{ij} is a random effect defined a the stratification level i. The term λ_0 is called the “baseline hazard” as it is common to all observations, and depends only on the elapsed time. The function λ_1 is the “systematic part” of the hazard and is commonly specified so as to be multiplicative in the elements of $X_k(t)$. The explanatory variables can be time varying and we make the standard regularity assumption that the process $X_k(t)$ is absolutely continuous with respect to the Lebesgue measure. The function of the unobserved heterogeneity is here supposed to be the product of the random effects and is thus log-linear.

Each unit k belongs to I different clusters and each level defines J_i different groups. Notice that the v_{ij} can be unit-specific random effects if $J_i = K$. Let us denote by $h_i(v_{ij}; \alpha_i)$ the mixing distributions at stratification level i, where α_i is a level specific vector of parameters. Assume that $h_i(v_{ij}; \alpha_i)$ admit a Laplace transforms, or equivalently, suppose that all the moments exist up to the order ∞. A large variance of a random effect means a tighter positive association among units of the same group and greater differences between the groups defined at this level.

Conditionally on the random effects and explanatory variables, the observations are assumed independent. Let us denote by T the vector of durations and by d the vector of transition indicators. If we suppose the random effects to be independent, we obtain the likelihood by taking the product of the frailty densities with the conditional likelihood function (conditional on the random effects), which is the conditional hazard times the conditional survivor function:

$$
\begin{aligned}
L(T,d|X,v,\beta,\alpha) = \prod_{k=1}^{N} \Bigg\{ & \left[\left(\prod_{i=1}^{I} v_{ij} \right) \lambda_0(t_k)\lambda_1[X_k(t_k),\beta] \right]^{\delta_k} \\
& \exp\left[-\left(\prod_{i=1}^{I} v_{ij} \right) \int_0^{t_k} \lambda_0(u)\lambda_1[X_k(u),\beta]\,du \right] \prod_{i=1}^{I} h_i(v_{ij};\alpha_i) \Bigg\}. \quad (2)
\end{aligned}
$$

The log-likelihood is thus given by:

$$
\begin{aligned}
\ln L(T,d|X,v,\beta,\alpha) = \sum_{k=1}^{N} \Bigg[& \delta_k \sum_{i=1}^{I} \ln v_{ij} + \delta_k \ln \lambda_0(t_k)\lambda_1[X_k(t_k),\beta] \\
& - \left(\prod_{i=1}^{I} v_{ij} \right) \int_0^{t_k} \lambda_0(u)\lambda_1[X_k(u),\beta]\,du + \sum_{i=1}^{I} \ln h_i(v_{ij},\alpha_i) \Bigg]. \quad (3)
\end{aligned}
$$

3 Inference using the EM algorithm

The structure of the MPH model makes it ideally suited for the EM algorithm. Indeed, had the v_{ij} been observed, the evaluation of the maximum likelihood estimator would have been straightforward.

Durations specific to cluster j at stratification level i are denoted by T_{ij}. Non-censoring indicators are denoted by d_{ij}. For fixed $\alpha = (\alpha_1, \ldots, \alpha_I)$, we can ignore the first term of equation (3). Assuming all the random effects to be independent leads us to the part of the log-likelihood which depends only on β, and gives us the E step at iteration (q):

$$
\begin{aligned}
Q\big(\beta,\beta^{(q)},\alpha^{(q)}\big) = \sum_{k=1}^{N} \Bigg(& \delta_k \ln \lambda_0(t_k)\lambda_1[X_k(t_k,\beta)] \\
& - \prod_{i=1}^{I} \mathrm{E}_{\beta^{(q)},\alpha^{(q)}}[v_{ij}|T_{ij},d_{ij}] \int_0^{t_k} \lambda_0(u)\lambda_1[X_k(u),\beta]\,du \Bigg). \quad (4)
\end{aligned}
$$

We thus need to evaluate $\mathrm{E}_{\beta^{(q)},\alpha^{(q)}}[v_{ij}|T_{ij},d_{ij}]$ for all (i,j). Let us denote by $\mathcal{L}_1^{(l_{1ik})}$ the l_{1ik}'th derivative of the Laplace transform. Extending the approach presented in Parner (1997), we show in Appendix A the following result:

$$\mathrm{E}\left[v_{ij}|T_{ij},d_{ij}\right] = \frac{\mathrm{E}\left[v_{ij}^{1+l_{iij}}\xi_{(-i)}\right]}{\mathrm{E}\left[v_{ij}^{l_{iij}}\xi_{(-i)}\right]}, \tag{5}$$

where:

$$\xi_{(-i)} = \mathrm{E}\left(v_{Ij}^{l_{Iij}}\ldots\mathrm{E}\left(v_{(i+1)j}^{l_{(i+1)ij}}\mathrm{E}\left(v_{(i-1)j}^{l_{(i-1)ij}}\ldots\mathrm{E}\left(v_{2j}^{l_{2ij}}\mathcal{L}_1^{(l_{1ij})}\right)\right)\right)\right), \tag{6}$$

and $l_{ii'j}$ is the number of transitions observed in the subcluster defined by the intersection of the two clusters containing unit j an obtained by stratification levels i and i'. These equations admit an analytical solution only in the case of a single gamma random effect (see Clayton and Cuzick, 1985). With more than one gamma frailty, these expectations have to be computed using numerical procedures (see Sastry, 1997), as with a log-normal frailty (see Vu and Knuiman, 2002).

However, the E-step can be handled in a simple way. Considering as fixed effects the expectations of all the frailties except one, the function $Q(\beta, \beta^{(q)}, \alpha^{(q)})$ is the E step of a model with only one frailty. The inference can be conducted in an iterative way, alternating between I EM algorithms: during a first EM, we consider a model with a single random effect, where the expectations of the other effects are treated as offsets. In a second EM algorithm, we consider a model where the second random effect is the only source of unobserved heterogeneity, the other effects being this time treated as offsets equal to the estimates obtained in the first EM algorithm; and so on.

This leads us to the following algorithm at iteration (q), where $\theta^{(q)} = (\beta^{(q)}, \alpha^{(q)})$:

1. Set $\mathrm{E}_{\theta^{(q-1)}}[v_{ij}|T_{ij},d_{ij}]$ $\forall i = 2,\ldots,I$ as offsets, estimate model:

$$\lambda_{ij}(t) = v_{1j}\left(\prod_{i=2}^{I}\mathrm{E}_{\theta^{(q-1)}}[v_{ij}|T_{ij},d_{ij}]\right)\lambda_0(t)\lambda_1[X_k(t),\beta], \tag{7}$$

   and recover $(\alpha_1^{(q)}, \beta_1^{(q)}, v_{1j}^{(q)})$,

2. Set $\mathrm{E}_{\theta^{(q)}}[v_{1j}|T_{1j},d_{1j}]$ and $\mathrm{E}_{\theta^{(q-1)}}[v_{ij}|T_{ij},d_{ij}]$ $\forall i = 3,\ldots,I$ as offsets, es-

timate model :

$$\lambda_{ij}(t) = v_{2j}\mathrm{E}_{\theta^{(q)}}[v_{1j}|T_{1j}, d_{1j}]\left(\prod_{i=3}^{I}\mathrm{E}_{\theta^{(q-1)}}[v_{ij}|T_{ij}, d_{ij}]\right)\lambda_0(t)\lambda_1[X_k(t), \beta], \tag{8}$$

and recover $(\alpha_2^{(q)}, \beta_2^{(q)}, v_{2j}^{(q)})$,

3. Estimate in this way the other $(I-2)$ single frailty models,

4. Iterate until convergence.

Convergence is monitored with a distance between the $\beta_i^{(q)}$.

The strength of this approach is to transform a rather complicated problem of inference in a model with different random effects in I sub-problems, each requiring the estimation of a model with only one random effect using the EM algorithm. We present in the next section the approach used to estimate the sub-models with continuous mixing distributions.

4 Estimation of the MPH models with a continuous random effect

Numerous studies set $\lambda_1[X_k(t_k), \beta] = \exp[X_k(t_k)\beta]$. Considering gamma unobserved heterogeneity, Johansen (1983) shows that the partial likelihood is a likelihood where λ_0 is profiled out, the baseline hazard being recovered with the Breslow estimator (i.e. Nelson-Aalen estimator without covariate). Gill (1985), extend the approach to multiple spells and gamma heterogeneity and suggest to use the EM algorithm. The idea is detailed in this setting in Klein (1992), and Parner (1997) extends it to all shared frailty models with a known Laplace transform of the mixing distribution. In this section, the partial likelihood in a model with one random effect and an important result of Therneau et al. (2003) are briefly recalled.

Using the same notations as before, consider the model:

$$\lambda_{1jk}(t) = v_{1j}\lambda_0(t)\exp[X_k(t)\beta], \tag{9}$$

where j $(j = 1, \ldots, J_1)$ is the group index and v_{1j} is a random effect. Define the risk set R_k as the set of spells still not completed at any instant before t_k. The associated partial likelihood (see for example Cox, 1972, 1975) is:

$$L_{PL}(\beta, v) = \prod_{k=1}^{N}\left[\frac{v_{1j}\exp[X_k(t_k)\beta]}{\sum_{(l,m)\in R_k} v_{1l}\exp[X_m(t_m)\beta]}\right]^{\delta_k}. \tag{10}$$

The integrated baseline hazard at iteration q can be recovered with the Breslow estimator:

$$\widehat{\Lambda}_0^{(q)}(t) = \sum_{t_k<t} \frac{\delta_k}{\sum_{(l,m)\in R_k}\left(\prod_{i=1}^{I} v_{il}^{(q)}\right)\exp\left[X_m(t_m)\beta^{(q)}\right]}. \tag{11}$$

4.1 Inference considering gamma heterogeneity

Abbring and Van den Berg (2001) show that the mixing distribution among survivors converges to a gamma distribution for a broad class of heterogeneity distributions. This result justifies the choice of a gamma mixing distribution on a stronger basis than simply its analytic tractability. Assuming a finite number of classes leads to the alternative approach of finite mixture models. The semiparametric heterogeneity model, where the classes are latent, is characterized and applied in Heckman and Singer (1984*a*, 1984b). When components are known rather than unobserved, it is equivalent to using fixed effects or a set of indicators.

Assume the v_{1j} follow a gamma distribution with expectation 1 and variance $1/\alpha_1$, Therneau et al. (2003) demonstrate that Klein's (1992) solution can be obtained exactly by maximizing the following penalized partial likelihood:

$$\ln L_{PPL}(\beta, v, \alpha_1) = \ln L_{PL}(\beta, v) - \frac{1}{\alpha_1}\sum_{j=1}^{J_1}(\ln v_{1j} - v_{1j}), \tag{12}$$

In a general penalized likelihood setting, $1/\alpha_1$ is a smoothing parameter indicating the tradeoff between the fit to the data and smoothness of the penalized likelihood. The solution maximizing the penalized partial likelihood above is equivalent to the EM solution for an MPH model with a gamma shared frailty such as (9) in the partial likelihood approach. This result relies on the choice of the penalty function and does not hold if one uses as a penalty function the quantity $\int_0^\infty \lambda_0^{(2)}(u)^2 du$, where $\lambda_0^{(2)}(u)$ stands for the second derivative of the baseline hazard, as done for example in Rondeau, Commenges and Joly (2003) to ensure existence and uniqueness of the estimator, as shown in de Montricher, Tapia and Thompson (1975).

We implement the penalized partial likelihood maximization algorithm instead of the EM algorithm to estimate all the I sub-models. Each algorithm is organized in two loops, and I describe here the algorithm corresponding to the sub-model where only the frailty defined at level i is considered as a random effect. At iteration (q), the first loop maximizes the marginal log penalized partial likelihood (see Therneau et al., 2003) and returns $\alpha_i^{(q)}$. The second loop consider this $\alpha_i^{(q)}$ as fixed, uses a Newton-Raphson procedure to optimize the penalized partial likelihood and returns $(\beta_i^{(q)}, v_{i1}^{(q)}, \dots, v_{iJ_i}^{(q)})$. Once the maximum is reached, the $v_{ij}^{(q)}$ are passed to the fitting program of the second sub-model and so on for all clustering levels. All algorithms are iterated and the estimated random effects are stable once convergence is achieved.

5 Monte Carlo Experiments

The aim of this Monte Carlo study is to compare the computing time between the algorithm we propose and the accelerated EM described in Sastry (1997), for different sizes of groups and subgroups. Our algorithm is hereafter referred to as Expectation Maximization algorithm based on Penalized Likelihood (EMPL).

5.1 An accelerated EM algorithm

Considering a piecewise constant baseline hazard, Sastry (1997) shows that the result of the E step can be separated in three functions, the two first ones depending of the parameters of the mixing distributions and the last one on β and $\Lambda_0(t)$. This separation dictates the organization of his EM algorithm: after the E step at iteration (q), the functions depending on the parameters of the mixing distributions are optimized separately within their own EM algorithms. Once these sub-routines achieve convergence, step M of iteration (q) is carried out for all the parameters and then iteration $(q+1)$ starts. Using these sub-routines enables to achieve efficiency and speed gains, and Sastry (1997) does not need to implement further acceleration techniques such as the ones described for example in Louis (1982) or Meilijson (1989). We extend his work to partial likelihood with Johansen's (1983) result, stated here at the beginning of Section 4.

5.2 Sample design and starting values

For each setting, 200 samples were simulated. We consider samples of sizes 2000 with two levels of clustering, and we vary the number of spells per group and the number of spells per subgroup accordingly. The smallest groups we design contain 10 observations and the largest 100. Subgroup sizes go from 5, the largest value compatible with groups of size 10, to 1.

We set $\mathrm{E}(v_{1j})$=$\mathrm{E}(v_{2j})$=1, $\mathrm{Var}(v_{1j})$=$\mathrm{Var}(v_{2j})$=0.5, and consider a constant baseline hazard and no censoring. Two standard gaussian covariates were used with coefficients $\beta_1 = 1$ and $\beta_2 = -1$, and there is no constant as it is not identified in a partial likelihood setting.

As pointed in Ng et al. (2004), convergence and thus computing time of the EM algorithm is sensitive to the choice of the starting values. We set them at their values in a model without unobserved heterogeneity, that is, the starting values are 1 for v_i and w_{ij}, 0 for the variances, and the estimates of a standard Cox model for the coefficients. All the results were obtained with the R 2.0.1 software, inference using penalized partial likelihood calling the function 'coxph' of the package 'survival'.

A preliminary comment is that the EMPL and accelerated EM provide the same estimates for all samples. Thus we report computing times only.

5.3 Results

Table 1 reports the 25th, 50th and 75th centiles of the computing times. Since computing times depend on the coding, it is difficult to make a fair comparison of both approaches. We can however obtain some insight about computing efficiency. Computing times differ widely when the estimators are applied to samples with the same clustering. In most cases, the ratio of the longest computing time over the shortest for a sample design is equal to 100 for the EMPL and 500 for the accelerated EM.

Sizes of groups and subgroups have a mixed impact on the computing times of the EMPL algorithm. They decrease monotonically with subgroup size, but are not much influenced by the number of spells per group. Due to this, a data scheme with groups containing 10 spells and subgroups of 5 observations will be estimate much more quickly than a sample where groups contain 100 spells and subgroups 2 observations. By contrast, computing times for the accelerated EM algorithm have an inverted $\cup$ profile with a maximum around 4 spells per subclusters. The more subclusters there are, the more v_{2j} have to be evaluated at each iteration, which raises computing times until a threshold. But more v_{2j} implies also more information, which speeds up convergence after the threshold. The EMPL algorithm is globally fast, with computing time quantiles generally far under their counterparts for the EM algorithm. We notice it especially when there are less than 20 durations per group: in this case, the 75th quantiles for the EMPL computing time are under the 25th quantiles for the EM computing time. These results of the EMPL algorithm being quicker than the EM one do not hold in presence of a single spell per subgroup, except when there are 10 spells per group.

Table 1: Computing time (Var(v_{1j})=Var(v_{2j})=0.50)

Number of spells			Computing Time					
Total	Per group	Per subgroup	Q_{25}	EMPL Median	Q_{75}	Q_{25}	EM Median	Q_{75}
2000	100	5	18.28	21.97	28.55	54.53	110.33	187.47
		4	22.42	23.52	34.70	88.00	176.26	561.04
		3	29.81	32.23	47.50	30.92	76.09	200.52
		2	55.08	73.77	83.48	29.75	86.64	212.78
		1	120.11	172.93	181.05	29.77	41.22	56.24
	50	5	19.05	27.34	33.68	97.90	162.64	313.25
		4	24.26	35.37	37.65	93.04	184.12	466.52
		3	32.91	49.01	57.69	72.17	159.11	401.08
		2	68.44	77.97	84.30	30.76	94.49	227.39
		1	156.37	175.44	181.12	36.98	44.98	60.07
	40	5	15.81	22.09	29.95	86.68	158.91	385.12
		4	25.77	38.55	46.07	126.02	218.38	573.89
		3	34.44	50.69	51.76	72.17	158.32	401.08
		2	57.01	72.28	82.79	39.42	104.57	215.62
		1	144.12	154.66	171.28	30.13	37.30	43.66
	20	5	19.85	28.64	36.70	173.84	283.86	510.24
		4	38.24	44.08	57.77	260.18	445.03	723.32
		3	44.99	51.84	61.33	224.50	359.47	583.09
		2	58.05	66.14	78.95	120.48	250.82	456.63
		1	137.73	149.90	158.33	48.05	72.67	135.53
	10	5	17.89	32.22	45.24	511.42	721.90	1028.27
		4	48.38	55.85	78.56	726.54	936.61	1402.90
		3	54.49	58.14	76.83	675.61	954.51	1456.78
		2	74.42	80.00	89.08	520.65	858.98	1485.12
		1	151.45	162.25	180.65	137.42	229.07	311.00

6 Ratification of the International Labour Organization conventions

As an example, we apply the EMPL algorithm on data reporting the timing of the ratification of ILO conventions. The dataset is presented and analysed in Boockmann (2001). The survival time represents the time between the adoption of ILO conventions and their ratification by developing countries over the period 1975-1995. The data comprise 80 countries and 29 conventions for a total of 228 ratifications. The hazard function is written as: $\lambda_{2jk}(t) = v_{1j}v_{2j}\lambda_0(t)\exp[X_k(t)\beta]$, where v_{1j} is a convention effect and v_{2j} a country effect.

We estimate this model using the EM algorithm. The result of the E step is separated in two parts, as described in Subsection 5.1, the first one involving the parameters of the mixing distribution and the second one the coefficients. The first function is optimized using its own sub-EM algorithms and we notice that the variances of both mixing distributions rise along their iterations. The algorithm then collapses during the M step, after 8 hours and a half of computation, returning coefficients tending to $-\infty$.

By contrast, the EMPL does converge. The results are reported in Appendix B and are closed to the ones obtained in Horny et al. (2005) using a Bayesian approach based on partial likelihood with log-normal frailties. Computation took 66 seconds with the EMPL and 3 months using Gibbs sampling.

7 Conclusion

This paper proposes a modified EM algorithm in the general framework of an MPH model with I ($I \geq 2$) random effects where mixing distributions have a known Laplace transform. We next devote a particular attention to gamma heterogeneity. Assuming this mixing distribution is a weak requirement since the result of Abbring and Van den Berg (2001) shows convergence of the distribution among survivors to it. We describe furthermore how to use penalized likelihood to avoid the numerical problems inherent to the EM algorithm. We thus provide a modified EM algorithm which is not only fast, but also simple and stable. The methodology is semi-parametric as it relies on partial likelihood. It does not require a specification of the baseline hazard and thus extends Sastry (1997). We also provide a small Monte Carlo study and an illustration on data to compare the behaviour of our procedure with an accelerated EM.

The computing times are dramatically reduced by using the EMPL algorithm, thus not asking for speeding-up routines as the ones described in Louis (1982) or Meilijson (1989) to be implemented in moderate size samples. The case of a random effect defining groups of one spell is an exception and speed depends on the size of the groups defined by the other frailties. Furthermore, the EMPL does converge in some settings where the EM does not.

We suppose that all the random effects are continuous. This assumption can be questionable when the population at hand is divided only in a few groups at the more aggregated levels. With 2 levels of heterogeneity, one possibility is to switch to a fixed effect approach, stratifying the baseline hazard at the broader levels and defining a frailty at the finest one, as proposed by Xue and Brookmeyer (1996).

A Derivation of $\mathbf{E}_{\beta^{(q)},\alpha^{(q)}}[v_{ij}|T_{ij},d_{ij}]$

In this appendix, we derive the conditional expectations of v_{ij}. Let us denote by $\mathcal{C}_{ij}$ the set of the units in cluster j defined by stratification level i. We have:

$$f(v_{1j},\ldots,v_{Ij},T_{ij},d_{ij}) = \prod_{k\in\mathcal{C}_{ij}} \left[\left(\prod_{i=1}^{I} v_{ij}\right)\lambda_0(t_k)\lambda_1[X_k(t_k)]\right]^{\delta_k} \exp\left[-\left(\prod_{i=1}^{I} v_{ij}\right)\int_0^{t_k}\lambda_0(u)\lambda_1[X_k(u)]\,du\right]\prod_{i=1}^{I} h_i(v_{ij};\alpha_i). \quad (13)$$

Integrating over v_{1j}, we obtain:

$$f(v_{2j},\ldots,v_{Ij},T_{ij},d_{ij}) = \left(\prod_{k\in\mathcal{C}_{ij}} \left[\left(\prod_{i=2}^{I} v_{ij}\right)\lambda_0(t_k)\lambda_1[X_k(t_k)]\right]^{\delta_k}\prod_{i=2}^{I} h_i(v_{ij};\alpha_i)\right) \int_{\mathcal{V}_1} v^{l_{1ik}}\exp\left[-\left(\prod_{i=1}^{I} v_{ij}\right)\sum_{k\in\mathcal{C}_{ij}}\int_0^{t_k}\lambda_0(u)\lambda_1[X_k(u)]\,du\right] h_1(v;\alpha_1)\,dv, \quad (14)$$

where l_{1ik} is the number of transitions observed in the group obtained as the intersection of the two clusters defined by stratification levels 1 and i and containing unit k. We can rewrite this expression as:

$$f(v_{2j},\ldots,v_{Ij},T_{ij},d_{ij}) = \left(\prod_{k\in\mathcal{C}_{ij}} \left[\left(\prod_{i=2}^{I} v_{ij}\right)\lambda_0(t_k)\lambda_1[X_k(t_k)]\right]^{\delta_k}\prod_{i=2}^{I} h_i(v_{ij};\alpha_i)\right) (-1)^{l_{1ik}}\mathcal{L}_1^{(l_{1ik})}\left[\left(\prod_{i=2}^{I} v_{ij}\right)\sum_{k\in\mathcal{C}_{ij}}\int_0^{t_k}\lambda_0(u)\lambda_1[X_k(u)]\,du\right], \quad (15)$$

where $\mathcal{L}_1^{(l_{1ik})}$ is the l_{1ik}'th derivative of the Laplace transform of a non-negative random variable v_1, defined as:

$$\mathcal{L}_1(s) = \int_{\mathcal{V}} \exp(-sv)dH(v), \quad (16)$$

where $s \geq 0$. Integrating over v_{2j} and omitting the argument of the Laplace transform, we obtain:

$$f(v_{3j}, \ldots, v_{Ij}, T_{ij}, d_{ij}) = \left(\prod_{k \in \mathcal{C}_{ij}} \left[\left(\prod_{i=3}^{I} v_{ij} \right) \lambda_0(t_k) \lambda_1[X_k(t_k)] \right]^{\delta_k} \prod_{i=3}^{I} h_i(v_{ij}; \alpha_i) \right) \\ (-1)^{l_{1ik}} \mathrm{E}_2 \left(v_{2j}^{l_{2ik}} \mathcal{L}_1^{(l_{1ik})} \right). \tag{17}$$

By further integrations, we can show that:

$$f(v_{ij}, T_{ij}, d_{ij}) = \left(\prod_{k \in \mathcal{C}_{ij}} [v_{ij} \lambda_0(t_k) \lambda_1[X_k(t_k)]^{\delta_k} h_i(v_{ij}; \alpha_i) \right) (-1)^{l_{1ik}} \\ \mathrm{E}_I \left(v_{Ij}^{l_{Iik}} \ldots \mathrm{E}_{(i+1)} \left(v_{(i+1)j}^{l_{(i+1)ik}} \mathrm{E}_{(i-1)} \left(v_{(i-1)j}^{l_{(i-1)ik}} \ldots \mathrm{E}_2 \left(v_{2j}^{l_{2ik}} \mathcal{L}_1^{(l_{1ik})} \right) \right) \right) \right) \tag{18}$$

To make equations shorter, let us denote:

$$\xi_{(-i)} = \mathrm{E}_I \left(v_{Ij}^{l_{Iik}} \ldots \mathrm{E}_{(i+1)} \left(v_{(i+1)j}^{l_{(i+1)ik}} \mathrm{E}_{(i-1)} \left(v_{(i-1)j}^{l_{(i-1)ik}} \ldots \mathrm{E}_2 \left(v_{2j}^{l_{2ik}} \mathcal{L}_1^{(l_{1ik})} \right) \right) \right) \right). \tag{19}$$

Thus:

$$\mathrm{E}\left[v_{ij} | T_{ij}, d_{ij} \right] = \frac{\int_{\mathcal{V}_i} v f(v, T_{ij}, d_{ij}) \, dv}{\int_{\mathcal{V}_i} f(v, T_{ij}, d_{ij}) \, dv} \tag{20}$$

B Results for the ratification of ILO conventions

Table 2: Estimates of the β parameters

Variable	Bayes		EMPL	
	Coef.	S.d	Coef.	S.d
Cost				
Real GDP per capita[a]	**3.81**	1.40	**3.03**	1.39
Real GDP per capita, squared	**-3.19**	1.51	-2.41	1.51
No explicit update	**1.39**	0.27	**1.26**	0.37
Own past ratification if explicit update	**1.62**	0.36	**1.52**	0.38
Population[b]	-0.02	0.05	-0.03	0.05
Internal pressure				
Democracy	**0.34**	0.15	0.29	0.15
Left majority	**-0.69**	0.31	**-0.62**	0.31
Vote against convention:				
Government	-0.22	0.23	-0.08	0.24
Employers	**0.38**	0.20	0.28	0.21
External pressure				
Development aid[c]	**-7.65**	2.05	**-8.56**	2.16
Worldbank loans[c]	2.00	1.55	**3.15**	1.57
IMF credits[c]	3.96	1.95	3.68	1.98
Exports[c]	-0.79	1.30	0.27	1.06
Exports into industrialized countries[c]	-0.18	3.66	-2.54	3.80
Exports into industrialized countries[c] (non oil exporting countries)	-0.77	3.48	0.52	3.71
Non oil exporting country	0.25	0.68	-0.01	0.71

Note: Bold entries are significant at the 5% level. *a.* 1985 international prices, in \$10 000. *b.* hundred milions. *c.* percent of GDP.

References

Abbring, J., and Van den Berg, G. (2001), The unobserved heterogeneity distribution in duration analysis,, Technical report, Free University Amsterdam, Amsterdam.

Bolstad, W., and Manda, S. (2001), "Investigating child mortality in Malawi using family and community random effects: a Bayesian analysis," *Journal of the American Statistical Association*, 96, 12–19.

Boockmann, B. (2001), "The ratification of ILO conventions: a hazard rate analysis," *Economics and Politics*, 13, 281–309.

Clayton, D., and Cuzick, J. (1985), "Multivariate generalizations of the proportional hazards model," *Journal of the Royal Statistical Society. Series A*, 148, 82–108.

Cox, D. (1972), "Regression models and life tables," *Journal of the Royal Statistical Society. Series B*, 34, 187–220.

Cox, D. (1975), "Partial likelihood," *Biometrika*, 62(2), 269–276.

de Montricher, G., Tapia, R., and Thompson, J. (1975), "Nonparametric maximum likelihood estimation of probability densities by penalty function methods," *The Annals of Statistics*, 3, 1329–1348.

Dempster, A., Laird, N., and Rubin, D. (1977), "Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion)," *Journal of the Royal Statistical Society. Series B*, 39, 1–38.

Feller, W. (1971), *An introduction to probability theory and its applications*, Vol. 2, 2 edn, New-York: Wiley.

Gill, R. (1985), "Discussion of the paper by D. Clayton and J. Cuzick," *Journal of the Royal Statistical Society. Series A*, 148, 108–109.

Gouriéroux, C., and Peaucelle, I. (1990), "Hétérogénéité I. Étude des biais d'estimation dans le cas linéaire," *Annales d'Économie et de Statistique*, 17, 163–184.

Heckman, J., and Singer, B. (1984*a*), "Econometric duration analysis," *Journal of Econometrics*, 24, 63–132.

Heckman, J., and Singer, B. (1984*b*), "A method for minimizing the impact of distributional assumptions in econometric models for duration data," *Econometrica*, 52, 271–320.

Horny, G., Boockmann, B., Djurdjevic, D., and Laisney, F. (2005), Bayesian estimation of Cox models with non-nested random effects: An application to the ratification of ILO conventions by developing countries,, Technical Report 05-23, ZEW.

Johansen, S. (1983), "An extension of Cox's regression model," *International Statistical Review*, 51, 258–262.

Klein, J. (1992), "Semiparametric estimation of random effects using the Cox model based on the EM algorithm," *Biometrics*, 48, 795–806.

Lancaster, T. (1990), *The econometric analysis of transition data*, Cambridge: Cambridge University Press.

Louis, T. (1982), "Finding the observed Information matrix when using the EM algorithm," *Journal of the Royal Statistical Society. Series B*, 44, 226–233.

Manda, S., and Meyer, R. (2005), "Age at first marriage in Malawi: a Bayesian multilevel analysis using a discrete time model," *Journal of the Royal Statistical Society. Series A*, 168, 439–455.

Meilijson, I. (1989), "A fast improvement to the EM algorithm on its own terms," *Journal of the Royal Statistical Society. Series B*, 51, 127–138.

Moulton, B. (1986), "Random group effects and the precision of regression estimates," *Journal of Econometrics*, 32, 385–397.

Ng, S., Krishnan, T., and McLachlan, G. (2004), "Handbook of computational statistics,".

Pakes, A. (1983), "On group effects and errors in variables in aggregation," *The Review of Economics and Statistics*, 65, 168–173.

Parner, E. (1997), Inference in semiparametric frailty models, PhD thesis, University of Aarhus.

Rondeau, V., Commenges, D., and Joly, P. (2003), "Maximum penalized likelihood estimation in a gamma-frailty model," *Lifetime Data Analysis*, 9, 139–153.

Sastry, N. (1997), "A nested frailty model for survival data, with an application to the study of child survival in northeast Brazil," *Journal of the American Statistical Association*, 92, 426–435.

Therneau, T., Grambsch, P., and Pankratz, V. (2003), "Penalized survival models and frailty," *Journal of Computational and Graphical Statistics*, 12, 156–175.

Van den Berg, G. (2001), "Duration models: specification, identification and multiple durations," in *Handbook of Econometrics*, Vol. 5, Elsevier, Amsterdam: J. J. Heckman and E. Leamer (eds.), chapter 55, pp. 3381–3463.

Vu, H., and Knuiman, M. (2002), "A hybrid ML-EM algorithm for calculation of maximum likelihood estimates in semiparametric shared frailty model," *Computational Statistics and Data Analysis*, 40, 173–187.

Wei, G., and Tanner, M. (1990), "Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithm," *Journal of the American Statistical Association*, 85, 699–704.

Xue, X., and Brookmeyer, R. (1996), "Bivariate frailty model for the analysis of multivariate survival time," *Lifetime Data Analysis*, 2, 277–289.

Yau, K. (2001), "Multilevel models for survival analysis with random effects," *Biometrics*, 57, 96–102.

Documents de travail du BETA

2000–**01** *Hétérogénéité de travailleurs, dualisme et salaire d'efficience.*
Francesco DE PALMA, janvier 2000.

2000–**02** *An Algebraic Index Theorem for Non–smooth Economies.*
Gaël GIRAUD, janvier 2000.

2000–**03** *Wage Indexation, Central Bank Independence and the Cost of Disinflation.*
Giuseppe DIANA, janvier 2000.

2000–**04** *Une analyse cognitive du concept de « vision entrepreneuriale ».*
Frédéric CRÉPLET, Babak MEHMANPAZIR, février 2000.

2000–**05** *Common knowledge and consensus with noisy communication.*
Frédéric KŒSSLER, mars 2000.

2000–**06** *Sunspots and Incomplete Markets with Real Assets.*
Nadjette LAGUÉCIR, avril 2000.

2000–**07** *Common Knowledge and Interactive Behaviors : A Survey.*
Frédéric KŒSSLER, mai 2000.

2000–**08** *Knowledge and Expertise : Toward a Cognitive and Organisational Duality of the Firm.*
Frédéric CRÉPLET, Olivier DUPOUËT, Francis KERN, Francis MUNIER, mai 2000.

2000–**09** *Tie–breaking Rules and Informational Cascades : A Note.*
Frédéric KŒSSLER, Anthony ZIEGELMEYER, juin 2000.

2000–**10** *SPQR : the Four Approaches to Origin–Destination Matrix Estimation for Consideration by the MYSTIC Research Consortium.*
Marc GAUDRY, juillet 2000.

2000–**11** *SNUS–2.5, a Multimoment Analysis of Road Demand, Accidents and their Severity in Germany, 1968–1989.*
Ulrich BLUM, Marc GAUDRY, juillet 2000.

2000–**12** *On the Inconsistency of the Ordinary Least Squares Estimator for Spatial Autoregressive Processes.*
Théophile AZOMAHOU, Agénor LAHATTE, septembre 2000.

2000–**13** *Turning Box–Cox including Quadratic Forms in Regression.*
Marc GAUDRY, Ulrich BLUM, Tran LIEM, septembre 2000.

2000–**14** *Pour une approche dialogique du rôle de l'entrepreneur/managerdans l'évolution des PME : l'ISO comme révélateur ...*
Frédéric CRÉPLET, Blandine LANOUX, septembre 2000.

2000–**15** *Diversity of innovative strategy as a source of technological performance.*
Patrick LLERENA, Vanessa OLTRA, octobre 2000.

2000–**16** *Can we consider the policy instruments as cyclical substitutes ?*

Sylvie DUCHASSAING, Laurent GAGNOL, décembre 2000.

2001–**01** *Economic growth and CO2 emissions : a nonparametric approach.*
Théophile AZOMAHOU, Phu NGUYEN VAN, janvier 2001.

2001–**02** *Distributions supporting the first–order approach to principal–agent problems.*
Sandrine SPÆTER, février 2001.

2001–**03** *Développement durable et Rapports Nord–Sud dans un Modèle à Générations Imbriquées : interroger le futur pour éclairer le présent.*
Alban VERCHÈRE, février 2001.

2001–**04** *Modeling Behavioral Heterogeneity in Demand Theory.*
Isabelle MARET, mars 2001.

2001–**05** *Efficient estimation of spatial autoregressive models.*
Théophile AZOMAHOU, mars 2001.

2001–**06** *Un modèle de stratégie individuelle de primo–insertion professionnelle.*
Guy TCHIBOZO, mars 2001.

2001–**07** *Endogenous Fluctuations and Public Services in a Simple OLG Economy.*
Thomas SEEGMULLER, avril 2001.

2001–**08** *Behavioral Heterogeneity in Large Economies.*
Gaël GIRAUD, Isabelle MARET, avril 2001.

2001–**09** *GMM Estimation of Lattice Models Using Panel Data : Application.*
Théophile AZOMAHOU, avril 2001.

2001–**10** *Dépendance spatiale sur données de panel : application à la relation Brevets–R&D au niveau régional.*
Jalal EL OUARDIGHI, avril 2001.

2001–**11** *Impact économique régional d'un pôle universitaire : application au cas strasbourgeois.*
Laurent GAGNOL, Jean–Alain HÉRAUD, mai 2001.

2001–**12** *Diversity of innovative strategy as a source of technological performance.*
Patrick LLERENA, Vanessa OLTRA, mai 2001.

2001–**13** *La capacité d'innovation dans les regions de l'Union Européenne.*
Jalal EL OUARDIGHI, juin 2001.

2001–**14** *Persuasion Games with Higher Order Uncertainty.*
Frédéric KŒSSLER, juin 2001.

2001–**15** *Analyse empirique des fonctions de production de Bosnie–Herzégovine sur la période 1952–1989.*
Rabija SOMUN, juillet 2001.

2001–**16** *The Performance of German Firms in the Business–Related Service Sectors : a Dynamic Analysis.*
Phu NGUYEN VAN, Ulrich KAISER, François LAISNEY, juillet 2001.

2001–**17** *Why Central Bank Independence is high and Wage indexation is low.*
Giuseppe DIANA, septembre 2001.

2001–**18** *Le mélange des ethnies dans les PME camerounaises : l'émergence d'un modèle d'organisation du travail.*

Raphaël NKAKLEU, octobre 2001.

2001–**19** *Les déterminants de la GRH des PME camerounaises.*
Raphaël NK AKLEU, octobre 2001.

2001–**20** *Profils d'identité des dirigeants et stratégies de financement dans les PME camerounaises.*
Raphaël NKAKLEU, octobre 2001.

2001–**21** Concurrence Imparfaite, Variabilité du Taux de Marge et Fluctuations Endogènes.
Thomas SEEGMULLER, novembre 2001.

2001–**22** *Determinants of Environmental and Economic Performance of Firms : An Empirical Analysis of the European Paper Industry.*
Théophile AZOMAHOU, Phu NGUYEN VAN et Marcus WAGNER, novembre 2001.

2001–**23** *The policy mix in a monetary union under alternative policy institutions and asymmetries.*
Laurent GAGNOL et Moïse SIDIROPOULOS, décembre 2001.

2001–**24** *Restrictions on the Autoregressive Parameters of Share Systems with Spatial Dependence.*
Agénor LAHATTE, décembre 2001.

2002–**01** *Strategic Knowledge Sharing in Bayesian Games : A General Model.*
Frédéric KŒSSLER, janvier 2002.

2002–**02** *Strategic Knowledge Sharing in Bayesian Games : Applications.*
Frédéric KŒSSLER, janvier 2002.

2002–**03** *Partial Certifiability and Information Precision in a Cournot Game.*
Frédéric KŒSSLER, janvier 2002.

2002–**04** *Behavioral Heterogeneity in Large Economies.*
Gaël GIRAUD, Isabelle MARET, janvier 2002.
(Version remaniée du Document de Travail n°2001–**08**, avril 2001).

2002–**05** *Modeling Behavioral Heterogeneity in Demand Theory.*
Isabelle MARET, janvier 2002.
(Version remaniée du Document de Travail n°2001–**04**, mars 2001).

2002–**06** *Déforestation, croissance économique et population : une étude sur données de panel.*
Phu NGUYEN VAN, Théophile AZOMAHOU, janvier 2002.

2002–**07** *Theories of behavior in principal–agent relationships with hidden action.*
Claudia KESER, Marc WILLINGER, janvier 2002.

2002–**08** *Principe de précaution et comportements préventifs des firmes face aux risques environnementaux.*
Sandrine SPÆTER, janvier 2002.

2002–**09** *Endogenous Population and Environmental Quality.*
Phu NGUYEN VAN, janvier 2002.

2002–**10** *Dualité cognitive et organisationnelle de la firme au travers du concept de communauté.*
Frédéric CRÉPLET, Olivier DUPOUËT, Francis KERN, Francis MUNIER, février 2002.

2002–**11** *Comment évaluer l'amélioration du bien–être individuel issue d'une modification de la qualité du service d'élimination des déchets ménagers ?*
Valentine HEINTZ, février 2002.

2002–**12** *The Favorite–Longshot Bias in Sequential Parimutuel Betting with Non–Expected Utility Players.*
Frédéric KŒSSLER, Anthony ZIEGELMEYER, Marie–Hélène BROIHANNE, février 2002.

2002–**13** *La sensibilité aux conditions initiales dans les processus individuels de primo–insertion professionnelle : critère et enjeux.*
Guy TCHIBOZO, février 2002.

2002–**14** *Improving the Prevention of Environmental Risks with Convertible Bonds.*
André SCHMITT, Sandrine SPÆTER, mai 2002.

2002–**15** *L'altruisme intergénérationnel comme fondement commun de la courbe environnementale* ***à la*** *Kuznets et du développement durable.*
Alban VERCHÈRE, mai 2002.

2002–**16** *Aléa moral et politiques d'audit optimales dans le cadre de la pollution d'origine agricole de l'eau.*
Sandrine SPÆTER, Alban VERCHÈRE, juin 2002.

2002–**17** *Parimutuel Betting under Asymmetric Information.*
Frédéric KŒSSLER, Anthony ZIEGELMEYER, juin 2002.

2002–**18** *Pollution as a source of endogenous fluctuations and periodic welfare inequality in OLG economies.*
Thomas SEEGMULLER, Alban VERCHÈRE, juin 2002.

2002–**19** *La demande de grosses coupures et l'économie souterraine.*
Gilbert KŒNIG, juillet 2002.

2002–**20** *Efficiency of Nonpoint Source Pollution Instruments with Externality Among Polluters : An Experimental Study.*
François COCHARD, Marc WILLINGER, Anastasios XEPAPADEAS, juillet 2002.

2002–**21** *Taille optimale dans l'industrie du séchage du bois et avantage compétitif du bois–énergie : une modélisation microéconomique.*
Alexandre SOKIC, octobre 2002.

2002–**22** *Modelling Behavioral Heterogeneity.*
Gaël GIRAUD, Isabelle MARET, novembre 2002.

2002–**23** *Le changement organisationnel en PME : quels acteurs pour quels apprentissages ?*
Blandine LANOUX, novembre 2002.

2002–**24** *TECHNOLOGY POLICY AND COOPERATION : An analytical framework for a paradigmatic approach.*
Patrick LLERENA, Mireille MATT, novembre 2002.

2003–**01** *Peut–on parler de délégation dans les PME camerounaises ?*
Raphaël NKAKLEU, mars 2003.

2003–**02** *L'identité organisationnelle et création du capital social : la tontine d'entreprise comme facteur déclenchant dans le contexte africain.*
Raphaël NKAKLEU, avril 2003.

2003–**03** *A semiparametric analysis of determinants of protected area.*
Phu NGUYEN VAN, avril 2003.

2003–**04** *Strategic Market Games with a Finite Horizon and Incomplete Markets.*
Gaël GIRAUD et Sonia WEYERS, avril 2003.

2003–**05** *Exact Homothetic or Cobb–Douglas Behavior Through Aggregation.*
Gaël GIRAUD et John K.–H. QUAH, juin 2003.

2003–**06** *Relativité de la satisfaction dans la vie : une étude sur données de panel.*
Théophile AZOMAHOU, Phu NGUYEN VAN, Thi Kim Cuong PHAM, juin 2003.

2003–**07** *A model of the anchoring effect in dichotomous choice valuation with follow–up.*
Sandra LECHNER, Anne ROZAN, François LAISNEY, juillet 2003.

2003–**08** *Central Bank Independence, Speed of Disinflation and the Sacrifice Ratio.*
Giuseppe DIANA, Moïse SIDIROPOULOS, juillet 2003.

2003–**09** *Patents versus ex–post rewards : a new look.*
Julien PÉNIN, juillet 2003.

2003–**10** *Endogenous Spillovers under Cournot Rivalry and Co–opetitive Behaviors.*
Isabelle MARET, août 2003.

2003–**11** *Les propriétés incitatives de l'effet Saint Matthieu dans la compétition académique.*
Nicolas CARAYOL, septembre 2003.

2003–**12** *The 'probleme of problem choice' : A model of sequential knowledge production within scientific communities.*
Nicolas CARAYOL, Jean–Michel DALLE, septembre 2003.

2003–**13** *Distribution Dynamics of CO_2 Emissions.*
Phu NGUYEN VAN, décembre 2003.

2004–**01** *Utilité relative, politique publique et croissance économique.*
Thi Kim Cuong PHAM, janvier 2004.

2004–**02** *Le management des grands projets de haute technologie vu au travers de la coordination des compétences.*
Christophe BELLEVAL, janvier 2004.

2004–**03** *Pour une approche dialogique du rôle de l'entrepreneur/manager dans l'évolution des PME : l'ISO comme révélateur …*
Frédéric CRÉPLET, Blandine LANOUX, février 2004.

2004–**04** *Consistent Collusion–Proofness and Correlation in Exchange Economies.*
Gaël GIRAUD, Céline ROCHON, février 2004.

2004–**05** *Generic Efficiency and Collusion–Proofness in Exchange Economies.*
Gaël GIRAUD, Céline ROCHON, février 2004.

2004–**06** *Dualité cognitive et organisationnelle de la firme fondée sur les interactions entre les communautés épistémiques et les communautés de pratique..*
Frédéric CRÉPLET, Olivier DUPOUËT, Francis KERN, Francis MUNIER, février 2004.

2004–**07** *Les Portails d'entreprise : une réponse aux dimensions de l'entreprise « processeur de connaissances ».*
Frédéric CRÉPLET, février 2004.

2004–**08** *Cumulative Causation and Evolutionary Micro–Founded Technical Change : A Growth Model with Integrated Economies.*
Patrick LLERENA, André LORENTZ, février 2004.

2004–**09** *Les CIFRE : un outil de médiation entre les laboratoires de recherche universitaire et les entreprises.*
Rachel LÉVY, avril 2004.

2004–**10** *On Taxation Pass–Through for a Monopoly Firm.*
Rabah AMIR, Isabelle MARET, Michael TROGE, mai 2004.

2004–**11** *Wealth distribution, endogenous fiscal policy and growth : status–seeking implications.*
Thi Kim Cuong PHAM, juin 2004.

2004–**12** *Semiparametric Analysis of the Regional Convergence Process.*
Théophile AZOMAHOU, Jalal EL OUARDIGHI, Phu NGUYEN VAN, Thi Kim Cuong PHAM, Juillet 2004.

2004–**13** *Les hypothèses de rationalité de l'économie évolutionniste.*
Morad DIANI, septembre 2004.

2004–**14** *Insurance and Financial Hedging of Oil Pollution Risks.*
André SCHMITT, Sandrine SPAETER, septembre 2004.

2004–**15** *Altruisme intergénérationnel, développement durable et équité intergénérationnelle en présence d'agents hétérogènes.*
Alban VERCHÈRE, octobre 2004.

2004–**16** *Du paradoxe libéral–parétien à un concept de métaclassement des préférences.*
Herrade IGERSHEIM, novembre 2004.

2004–**17** *Why do Academic Scientists Engage in Interdisciplinary Research ?*
Nicolas CARAYOL, Thuc Uyen NGUYEN THI, décembre 2004.

2005–**01** *Les collaborations Université Entreprises dans une perspective organisationnelle et cognitive.*
Frédéric CRÉPLET, Francis KERN, Véronique SCHAEFFER, janvier 2005.

2005–**02** *The Exact Insensitivity of Market Budget Shares and the 'Balancing Effect'.*
Gaël GIRAUD, Isabelle MARET, janvier 2005.

2005–**03** *Les modèles de type Mundell–Fleming revisités.*
Gilbert KOENIG, janvier 2005.

2005–**04** *L'État et la cellule familiale sont–ils substituables dans la prise en charge du chômage en Europe ? Une comparaison basée sur le panel européen.*
Olivia ECKERT–JAFFE, Isabelle TERRAZ, mars 2005.

2005–**05** *Environment in an Overlapping Generations Economy with Endogenous Labor Supply : a Dynamic Analysis.*
Thomas SEEGMULLER, Alban VERCHÈRE, mars 2005.

2005–**06** *Is Monetary Union Necessarily Counterproductive ?*
Giuseppe DIANA, Blandine ZIMMER, mars 2005.

2005–**07** *Factors Affecting University–Industry R&D Collaboration : The importance of screening and signalling.*
Roberto FONTANA, Aldo GEUNA, Mireille MATT, avril 2005.

2005–**08** *Madison–Strasbourg, une analyse comparative de l'enseignement supérieur et de la recherche en France et aux États–Unis à travers l'exemple de deux campus.*
Laurent BUISSON, mai 2005.

2005–**09** *Coordination des négociations salariales en UEM : un rôle majeur pour la BCE.*
Blandine ZIMMER, mai 2005.

2005–**10** *Open knowledge disclosure, incomplete information and collective innovations.*
Julien PÉNIN, mai 2005.

2005–**11** *Science–Technology–Industry Links and the 'European Paradox' : Some Notes on the Dynamics of Scientific and Technological Research in Europe.*
Giovanni DOSI, Patrick LLERENA, Mauro SYLOS LABINI, juillet 2005.

2005–**12** *Hedging Strategies and the Financing of the 1992 International Oil Pollution Compensation Fund.*
André SCHMITT, Sandrine SPAETER, novembre 2005.

2005–**13** *Faire émerger la coopération internationale : une approche expérimentale comparée du bilatéralisme et du multilatéralisme.*
Stéphane BERTRAND, Kene BOUN MY, Alban VERCHÈRE, novembre 2005.

2005–**14** *Segregation in Networks.*
Giorgio FAGIOLO, Marco VALENTE, Nicolaas J. VRIEND, décembre 2005.

2006–**01** *Demand and Technology Determinants of Structural Change and Tertiarisation : An Input–Output Structural Decomposition Analysis for four OECD Countries.*
Maria SAVONA, André LORENTZ, janvier 2006.

2006–**02** *A strategic model of complex networks formation.*
Nicolas CARAYOL, Pascale ROUX, janvier 2006.

2006–**03** *Coordination failures in network formation.*
Nicolas CARAYOL, Pascale ROUX, Murat YILDIZOGLU, janvier 2006.

2006–**04** *Real Options Theory for Law Makers.*
Marie OBIDZINSKI, Bruno DEFFAINS, janvier 2006.

2006–**05** *Ressources, compétences et stratégie de la firme : Une discussion de l'opposition entre la vision Porterienne et la vision fondée sur les compétences.*
Fernand AMESSE, Arman AVADIKYAN, Patrick COHENDET, janvier 2006.

2006–**06** *Knowledge Integration and Network Formation.*
Müge OZMAN, janvier 2006.

2006–**07** *Networks and Innovation : A Survey of Empirical Literature.*
Müge OZMAN, février 2006.

2006–**08** *A.K. Sen et J.E. Roemer : une même approche de la responsabilité ?*
Herrade IGERSHEIM, mars 2006.

2006–**09** *Efficiency and coordination of fiscal policy in open economies.*
Gilbert KOENIG, Irem ZEYNELOGLU, avril 2006.

2006–**10** *Partial Likelihood Estimation of a Cox Model With Random Effects : an EM Algorithm Based on Penalized Likelihood.*
Guillaume HORNY, avril 2006.

La présente liste ne comprend que les Documents de Travail publiés à partir du 1er janvier 2000. La liste complète peut être donnée sur demande.
This list contains the Working Paper writen after January 2000, 1rst. The complet list is available upon request.