DISCUSSION PAPER SERIES

IZA DP No. 8083

Dependence Measures in Bivariate Gamma Frailty Models

Gerard J. van den Berg
Georgios Effraimidis

March 2014

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Dependence Measures in Bivariate Gamma Frailty Models

Gerard J. van den Berg
University of Mannheim,
IFAU Uppsala, VU University Amsterdam and IZA

Georgios Effraimidis
University of Southern Denmark

Discussion Paper No. 8083
March 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 8083
March 2014

ABSTRACT

Dependence Measures in Bivariate Gamma Frailty Models*

Bivariate duration data frequently arise in economics, biostatistics and other areas. In "bivariate frailty models", dependence between the frailties (i.e., unobserved determinants) induces dependence between the durations. Using notions of quadrant dependence, we study restrictions that this imposes on the implied dependence of the durations, if the frailty terms act multiplicatively on the corresponding hazard rates. Marginal frailty distributions are often taken to be gamma distributions. For such cases we calculate general bounds for two association measures, Pearson's correlation coefficient and Kendall's tau. The results are employed to compare the flexibility of specific families of bivariate gamma frailty distributions.

JEL Classification: C41, C51, C34, C33, C32, J64

Keywords: bivariate gamma distribution, duration models, competing risks, Kendall's tau, negative and positive quadrant dependence, Pearson's correlation coefficient, unobserved heterogeneity, survival analysis

Corresponding author:

Georgios Effraimidis
University of Southern Denmark
Department of Business and Economics
Campusvej 55
5230 Odense M
Denmark
E-mail: gef@sam.sdu.dk

* The helpful comments of Maarten Lindeboom, Richard Paap, Geert Ridder and participants at seminars at USC and the Tinbergen Institute Amsterdam and at a conference in Milan are gratefully acknowledged.

1 Introduction

To describe the distribution of time that a subject spends in a certain state of interest, it is common to use models in which the exit rate out of the state (or hazard rate) depends multiplicatively on some unobserved characteristics or frailty term. In the bivariate frailty model (Clayton, 1978), two duration variables are considered, each with its own hazard rate and its own frailty term. These durations may describe two parallel, possibly competing, durations of the same subject (e.g., duration until merger and duration until bankruptcy of a firm) or single durations of two subjects which belong to the same cluster (e.g., death times of twins). This model has several applications in biostatistics, economics, engineering and many other fields. See Van den Berg (2001) for an overview.

The popularity of the bivariate frailty model among practitioners derives in part from the fact that it allows for a convenient way to model dependence (conditional on observed characteristics) between the duration variables. Dependence between the two duration variables is caused by dependence between the unobserved characteristics that enter the underlying hazard rates. Knowledge of the dependence structure of the frailty terms helps us to determine the type of association between the duration variables. Therefore, it is worth studying which functional forms of the distribution of the frailty terms are not needlessly restrictive in their implications for the joint duration distribution.

Our contribution in this paper is twofold. First, we study the notions of negative quadrant dependence and positive quadrant dependence for the joint distribution of the duration variables as a function of the quadrant dependence of the frailty terms. Secondly, we study the dependence between the duration variables under the condition that the frailty terms are gamma distributed. This choice of distribution is particularly relevant for two reasons. The first reason is that it leads to convenient functional forms for the duration model and hence it is commonly implemented in standard statistical software. The second reason is that the frailty distribution among the subjects who are still in the state of interest as time proceeds converges under weak requirements to a gamma distribution (see Abbring and Van den Berg (2007)). Thus, in many cases the gamma distribution provides a good approximation of an unknown true distribution. This has further increased the popularity of the gamma frailty distribution. Several authors have proposed graphical and numerical procedures to check for the adequacy of the assumption of gamma distribution for the frailty term. For example, Cui and Sun (2004) propose a supremum-type test statistic, whose asymptotic critical values are calculated by Monte Carlo simulation, and apply a numerical method as well as a graphical approach to test the validity of the gamma assumption. Shih and Louis (1995) apply a graphical method to test the assumption of gamma frailty by calculating the average of the posterior mean of the frailty given the observed data.

To quantify the degree of dependence between the duration outcomes, we will employ two association measures, specifically Pearson's correlation coefficient and Kendall's tau. The former, which measures the strength of the linear relationship between two random variables, is commonly used in empirical analysis for statistical inference. The latter measures the strength of any monotonic relationship between two random variables and consequently it is characterized by the rank-invariant property.

We focus on negative as well as positive dependence between the duration variables. In biostatistical applications, the duration variables are usually positively dependent as the corresponding hazard rates share same unobserved or nonmeasurable characteristics (e.g., environmental, genetic). In social sciences, there are numerous examples of negative dependence between the duration variables. This can be explained as follows. If an element of the set of covariates is unobserved in the data, and if this element has a positive effect on one of the hazard rates and a negative effect on the other, then this leads to a negative dependence between the duration variables. In labour economics, for instance, consider an individual who is unemployed and faces two competing exits from the unemployment state: employment and dropping out of the labor force. If the unemployed individual is strongly motivated (which is not observed) to get a job, then the exit rate into employment will be negatively associated with the exit rate out of the labor force. In the latter case, this can be captured by a multiplicative frailty term for the exit rate to employment that can be thought to be increasing in motivation, and another multiplicative frailty term for the exit rate into nonparticipation that can be thought to be decreasing in motivation.

The results of this paper are useful for researchers who work with the bivariate gamma frailty model. First, we discuss which bivariate distributions for the frailty terms generate negative and/or positive dependence between the duration variables. Second, we calculate bounds for the Pearson's correlation coefficient, if the baseline hazard are of the Weibull form. Moreover, we derive results on the bounds for the Kendall's tau which are more general concerning the bivariate gamma frailty model, as we do not make use of any parametric assumption regarding the interaction of time and explanatory variables. Finally, we compare our findings with the results of Van den Berg (1997), who provides nonparametric bounds for these two measures and considers bounds for discrete and lognormal frailty distributions.

The rest of the paper is structured as follows. Section 2 briefly introduces the bivariate frailty model and discusses dependence properties of the joint survival function (equivalently, joint distribution) of the duration variables given the dependence structure of the distribution of the frailty terms. In Section 3, we discuss the properties of different bivariate distributions with gamma marginals which can be used for modelling the bivariate distribution of the two frailty terms. Section 4 focuses on the bounds for Pearson's correlation coefficient, and Section 5 studies the bounds for Kendall's tau. Section 6 concludes and discusses possible extensions. The mathematical proofs are deferred to Appendix A. In Appendix B, we consider the dependence properties of some popular bivariate copulas. For notational convenience, we will omit the transpose symbol for vectors throughout the paper.

2 Quadrant dependence in the bivariate frailty model

2.1 Model Framework

Let T_1 and T_2 represent the nonnegative stochastic durations of interest and X be a vector of observable characteristics with support $\mathcal{X} \subseteq \mathbf{R}^d$, where d is a finite positive integer number. Denote by $x \in \mathcal{X}$ the realization of X. In addition, introduce two frailty terms $V_1 \in \mathbf{R}_+$ and $V_2 \in \mathbf{R}_+$ that are independent of the vector X and directly affect the realization of T_1 and T_2, respectively. The random variables V_1 and V_2 capture unobserved or nonmeasurable time-invariant characteristics. The corresponding hazard rate of the duration variables $T_1|x, V_1$ and $T_2|x, V_2$ is expressed as follows:

$$\begin{aligned}\theta_1(t|x,V_1) &= \lambda_1(t,x)V_1,\\ \theta_2(t|x,V_2) &= \lambda_2(t,x)V_2,\end{aligned}\tag{1}$$

with $\lambda_1 : \mathbf{R}_+ \times \mathcal{X} \to (0,\infty)$ and $\lambda_2 : \mathbf{R}_+ \times \mathcal{X} \to (0,\infty)$. We shall assume that the functions $\lambda_1(.,x)$ and $\lambda_2(.,x)$ are integrable on bounded intervals of the positive real line, that is, the quantities $\Lambda_1(t,x) = \int_0^t \lambda_1(\omega,x)d\omega$ and $\Lambda_2(t,x) = \int_0^t \lambda_2(\omega,x)d\omega$ exist for each $(t,x) \in \mathbf{R}_+ \times \mathcal{X}$.

We denote by G the distribution of the bivariate random vector (V_1, V_2) and by G_1 and G_2 the marginal distribution of V_1 and V_2, respectively. The main assumption that will hold throughout this paper is $T_1 \perp T_2|x,V_1,V_2$. In words, the duration variables are stochastically independent of each other given the observable characteristics and the frailty terms. Let $i = 1,2$, and consider the survival functions $S_i(t|x,V_i) = \mathbf{P}(T_i > t|x,V_i)$ and $S(t_1,t_2|x,V_1,V_2) = \mathbf{P}(T_1 > t_1, T_2 > t_2|x,V_1,V_2)$. The specification (1) implies $S_i(t|x,V_i) = \exp(-\Lambda_i(t,x)V_i)$, and therefore $S(t_1,t_2|x,V_1,V_2) = \exp(-\Lambda_1(t_1,x)V_1 - \Lambda_2(t_2,x)V_2)$ when the conditional independence property $T_1 \perp T_2|x,V_1,V_2$ is used. Also, introduce the survival functions $S_i(t|x) = \mathbf{P}(T_i > t|x)$ and $S(t_1,t_2|x) = \mathbf{P}(T_1 > t_1, T_2 > t_2|x)$. The survival function of $T_i|x$ can be explicitly calculated by a mixture of exponential distributions in the following way:

$$S_i(t|x) = \int_{\mathbf{R}_+} \exp(-\Lambda_i(t,x)v)dG_i(v) = \mathcal{L}_{G_i}\left(\Lambda_i(t,x)\right), \quad i = 1,2, \tag{2}$$

where the generic symbol $\mathcal{L}$ denotes the Laplace Transform (LT) of the corresponding probability measure. Likewise, the survival function of $(T_1,T_2)|x$ can be represented by a mixture of bivariate exponential distributions as follows:

$$\begin{aligned}S(t_1,t_2|x) &= \int_{\mathbf{R}_+^2} \exp(-\Lambda_1(t_1,x)v_1 - \Lambda_2(t_2,x)v_2)dG(v_1,v_2)\\ &= \mathcal{L}_G\left(\Lambda_1(t_1,x), \Lambda_2(t_2,x)\right).\end{aligned}\tag{3}$$

If $V_1 \perp V_2$ we get $\mathcal{L}_G(s_1,s_2) = \mathcal{L}_{G_1}(s_1)\mathcal{L}_{G_2}(s_2)$ for all $(s_1,s_2) \in \mathbf{R}_+^2$ and thus we have, by (2) and (3), $T_1 \perp T_2|x$ for any $x \in \mathcal{X}$. On the other hand, if $T_1 \perp T_2|x$ for some $x \in \mathcal{X}$, then $V_1 \perp V_2$ by noting, in view of (1), that $\ln V_i = -\ln \Lambda_i(T_i,x) + \epsilon_i$ for $i = 1,2$, where ϵ_1, ϵ_2 are independent random variables with probability density function $f_i(\epsilon) = e^{\epsilon} \exp(-e^{\epsilon})$.

2.2 Quadrant dependence

We first begin with the definitions of negative quadrant dependence and positive quadrant dependence (Lehmann, 1966).

Definition 1 *An $\mathbf{R}^2_+$−valued bivariate random vector (W_1, W_2) and its distribution function are said to be negative (positive) quadrant dependent if*

$$\mathbf{P}(W_1 \leq w_1, W_2 \leq w_2) \leq (\geq) \mathbf{P}(W_1 \leq w_1)\mathbf{P}(W_2 \leq w_2) \text{ for all } (w_1, w_2) \in \mathbf{R}^2_+.$$

Equivalently, an $\mathbf{R}^2_+$−bivariate random vector (W_1, W_2) and its survival function are said to be negative (positive) quadrant dependent if

$$\mathbf{P}(W_1 > w_1, W_2 > w_2) \leq (\geq) \mathbf{P}(W_1 > w_1)\mathbf{P}(W_2 > w_2) \text{ for all } (w_1, w_2) \in \mathbf{R}^2_+.$$

In the sequel, we use the acronyms NQD and PQD for the terms negative quadrant dependent and positive quadrant dependent, respectively. These two dependence concepts are the weakest for describing the dependence structure between two random variables. In particular, the density function of a bivariate random vector is reverse rule of order two, the strongest notion of negative dependence, only if the underlying distribution function is NQD. Likewise, the density function of a bivariate random vector is totally positive of order two, the strongest concept of positive dependence, only if the corresponding distribution is PQD. [1]. Next, we recall the definition of the concordance ordering $\prec_C$ that can be found in Joe (1997).

Definition 2 *Suppose $\mathcal{P}^a$ and $\mathcal{P}^b$ are bivariate distribution functions on $\mathbf{R}^2_+$ or bivariate survival functions on $\mathbf{R}^2_+$ with specific marginals $\mathcal{P}_1$ and $\mathcal{P}_2$. If $\mathcal{P}^a(w_1, w_2) \leq \mathcal{P}^b(w_1, w_2)$ for all $(w_1, w_2) \in$*

[1] A function $f : A \subseteq \mathbf{R}^2 \mapsto \mathbf{R}_+$ is totally positive of order two if

$$f(\max(x_1, y_1), \max(x_2, y_2)) f(\min(x_1, y_1), \min(x_2, y_2)) - f(x_1, x_2) f(y_1, y_2) \geq 0$$

for $(x_1, x_2), (y_1, y_2) \in A$. In case the above inequality is reversed, the function$f$ is reverse rule of order two (Joe, 1997).

$\mathbf{R}_+^2$, *then we say that* $\mathcal{P}^b$ *is more concordant than* $\mathcal{P}^a$, *written as* $\mathcal{P}^a \prec_C \mathcal{P}^b$.

We first obtain the following result which states that any concordance ordering between two different distributions of (V_1, V_2) will result in the same concordance ordering between the corresponding survival functions of $(T_1, T_2)|x$.

Proposition 1 *Let* G^a *and* G^b *represent two different distributions of the random vector* (V_1, V_2) *with* $G^a \prec_C G^b$. *Also, denote by* S^a *and* S^b *the corresponding mixtures of bivariate exponential distributions as defined in* (3). *Then,* $S^a \prec_C S^b$ *for each* $x \in \mathcal{X}$.

An important remark about Proposition 1 is that its result can be extended to any arbitrary bivariate hazard model in which the $S(t_1, t_2|x, v_1, v_2)$ is a bounded, continuous and $2-$positive function in (v_1, v_2) for all $(t_1, t_2, x) \in \mathbf{R}_+^2 \times \mathcal{X}$ (see Appendix A). The next corollary directly follows from Proposition 1 by setting $G^a(v_1, v_2) \leq G_1(v_1)G_2(v_2) = G^b(v_1, v_2), (v_1, v_2) \in \mathbf{R}_+^2$, for the NQD result and $G^a(v_1, v_2) = G_1(v_1)G_2(v_2) \leq G^b(v_1, v_2)$, $(v_1, v_2) \in \mathbf{R}_+^2$ for the PQD result.

Corollary 1 *Let* T_1 *and* T_2 *be the duration variables that are generated by the bivariate frailty model* (1). *If* (V_1, V_2) *is NQD (PQD), then* $(T_1, T_2)|x$ *is NQD (PQD) for every* $x \in \mathcal{X}$.

2.3 Association measures for the duration variables

In Sections 4 and 5 we shall consider bounds for the values of Pearson's correlation coefficient and Kendall's tau, respectively. The former quantitatively describes the strength of the linear relationship between T_1 and T_2, whereas the latter is a rank correlation coefficient between T_1 and T_2. According to Corollary 1, the type of quadrant dependence of the random vector (V_1, V_2) determines the type of quadrant dependence of the random vector $(T_1, T_2)|x$ for any $x \in \mathcal{X}$ and thereby the sign of these two association measures.

Assuming that $\mathbf{E}(T_i|x) < \infty$ and $\mathbf{E}(T_i^2|x) < \infty$ for $x \in \mathcal{X}$ and $i = 1, 2$, the conditional on x Pearson's correlation coefficient between T_1 and T_2 is expressed as

$$\rho(T_1, T_2|x) = \frac{\text{Cov}(T_1, T_2|x)}{[\text{Var}(T_1|x)\text{Var}(T_2|x)]^{\frac{1}{2}}}. \tag{4}$$

By Hoeffding's identity we have

$$\text{Cov}(T_1, T_2|x) = \int_{\mathbf{R}_+^2} \left[S(t_1, t_2|x) - S_1(t_1|x)S_2(t_2|x)\right] dt_1 dt_2.$$

Therefore, if $(T_1, T_2)|x$ is NQD for all $x \in \mathcal{X}$ it will hold that $S(t_1, t_2|x) - S_1(t_1|x)S_2(t_2|x) \leq 0$ for all $(t_1, t_2, x) \in \mathbf{R}_+^2 \times \mathcal{X}$ and therefore $\rho(T_1, T_2|x) \leq 0$ for any $x \in \mathcal{X}$. The previous inequalities will be reversed in case $(T_1, T_2)|x$ is PQD.

The main drawback of Pearson's correlation coefficient is that it is not rank-invariant, that is, generally $\rho(T_1, T_2|x) \neq \rho(h_1(T_1), h_2(T_2)|x)$ for any nonlinear strictly monotone transformations h_1 and h_2. A measure that satisfies this property is the Kendall's tau which has attracted the interest of researchers who work on duration analysis (Wang et al., 2000; Martin and Betensky, 2005; Beaudoin et al., 2007; Oakes, 2008). To be more precise, consider two independent copies $(T_1^A, T_2^A)|x$ and $(T_1^B, T_2^B)|x$ of the bivariate random vector $(T_1, T_2)|x$. The value of $\tau(T_1, T_2|x)$ for any $x \in \mathcal{X}$ is calculated by the following difference

$$\tau(T_1, T_2|x) = \mathbf{P}\left[(T_1^A - T_1^B)(T_2^A - T_2^B) > 0|x\right] - \mathbf{P}\left[(T_1^A - T_1^B)(T_2^A - T_2^B) < 0|x\right],$$

which gives

$$\tau(T_1, T_2|x) = 2\mathbf{P}\left[(T_1^A - T_1^B)(T_2^A - T_2^B) > 0|x\right] - 1. \tag{5}$$

Clearly, it holds that $-1 \leq \tau(T_1, T_2|x) \leq 1$ for all $x \in \mathcal{X}$ and it is also easy to see that the value of $\tau(T_1, T_2|x)$ is equal to -1 $(+1)$ if and only if $T_2 = h(T_1)$, with h to be a strictly decreasing (increasing) transformation. Some further elaboration of (5) gives

$$\tau(T_1, T_2|x) = 4\int_{\mathbf{R}_+^2} S(t_1, t_2|x) dS(t_1, t_2|x) - 1. \tag{6}$$

Note that we have chosen to express $\tau(T_1, T_2|x)$ as a functional of $S(t_1, t_2|x)$ and not of $F(t_1, t_2|x)$, where $F(t_1, t_2|x) = 1 - S_1(t_1|x) - S_2(t_2|x) + S(t_1, t_2|x)$, as we find it more convenient for the analysis in the sequel. Hence, in case $(T_1, T_2)|x$ is NQD for all $x \in \mathcal{X}$, it will hold that $S(t_1, t_2|x) \leq S_1(t_1|x)S_2(t_2|x)$ for all $(t_1, t_2, x) \in \mathbf{R}_+^2 \times \mathcal{X}$ and by using the result of Theorem 2 of Tchen (1980) it can be readily shown that $\tau(T_1, T_2|x) \leq 0$ for any $x \in \mathcal{X}$. On the other hand, if $(T_1, T_2)|x$ is PQD the previous inequalities will go in the opposite direction.

3 Bivariate frailty distribution with gamma marginals

To derive bounds for the values of the two association measures, we assume $V_i \sim Gamma(k_i, \mu_i)$ for $i = 1, 2$, where the parameters k_i and μ_i are defined as shape parameter and scale parameter, respectively, and we assume that they are strictly positive. More precisely, the probability density of V_i is given by

$$g_i(v) = \frac{1}{\mu_i^{k_i}\Gamma(k_i)} v^{k_i-1} \exp(-\frac{v}{\mu_i}),\ \ v_i > 0, k_i > 0, \mu_i > 0,$$

where the Eulerian gamma function Γ is computed by $\Gamma(k) = \int_0^\infty \omega^{k-1} \exp(-\omega) d\omega$ for $k > 0$. In the next two subsections we shall discuss possible parameterizations of G and the dependence structure they induce on $(T_1, T_2)|x$.

3.1 Bivariate gamma distributions

Before proceeding to the description of two bivariate gamma distributions, recall that if T_1 and T_2 are generated by (1), then $S(t_1, t_2|x) = \mathcal{L}_G\left(\Lambda_1(t_1, x), \Lambda_2(t_2, x)\right)$ and $S_i(t|x) = \mathcal{L}_{G_i}\left(\Lambda_i(t, x)\right)$ for $i = 1, 2$. The first distribution that we study as a candidate for the parameterization of G is the double bivariate gamma Kotz et al. (2000), which has the following stochastic representation

$$V_i \ \ = \ \ \mu_i(V_0 + V_{0i}), \quad i = 1, 2, \tag{7}$$

with $V_0 \sim Gamma(k_0, 1)$ and $V_{0i} \sim Gamma(k_{0i}, 1)$ being independent gamma variates. The marginal distribution of V_i is gamma distribution with shape parameter k_0+k_{0i} and scale parameter μ_i.

Cherian (1941) studied the above distribution for $\mu_1 = \mu_2 = 1$ and $k_{01} = k_{02}$. The use of the double bivariate gamma distribution is widespread in applications in the field of biostatistics (Korsgaard and Andersen, 1998; Zhong and Li, 2002; Jonker et al., 2009) and demography (Yashin et al., 1995). By using Bayes' law we can deduce that the vector (V_1, V_2) is PQD and consequently, by Corollary 1, the vector $(T_1, T_2)|x$ is PQD for each $x \in \mathcal{X}$. There is an alternative way to view that $(T_1, T_2)|x$ is PQD for this case. The LT of the double gamma distribution is expressed as $\mathcal{L}_G(s_1, s_2) = \mathcal{L}_{G_0}(\mu_1 s_1 + \mu_2 s_2)\mathcal{L}_{G_{01}}(\mu_1 s_1)\mathcal{L}_{G_{02}}(\mu_2 s_2)$ for any $(s_1, s_2) \in \mathbf{R}^2_+$. Provided that $V_0 \sim Gamma(k_0, 1)$, it holds that $\mathcal{L}_{G_0}(\mu_1 s_1 + \mu_2 s_2) \geq \mathcal{L}_{G_0}(\mu_1 s_1)\mathcal{L}_{G_0}(\mu_2 s_2)$ for all $(s_1, s_2) \in \mathbf{R}^2_+$, from which it is straightforward to infer that $(T_1, T_2)|x$ is PQD. Finally, note that for the limiting case $k_{0i} \to 0$ we get that $\mathbf{P}(\mu_2 V_1 = \mu_1 V_2) \to 1$.

The second bivariate gamma distribution that we consider for modelling G is mostly known by its LT, which is expressed as follows

$$\mathcal{L}_G(s_1, s_2) = (1 + \mu_1 s_1 + \mu_2 s_2 + \mu_{12} s_1 s_2)^{-k}, \quad (s_1, s_2) \in \mathbf{R}^2_+, \tag{8}$$

with $k > 0$, $\mu_1 > 0$, $\mu_2 > 0$ and $\mu_1\mu_2 - \mu_{12} \geq 0$. The above LT corresponds to a bivariate gamma distribution with $V_1 \sim Gamma(k, \mu_1)$ and $V_2 \sim Gamma(k, \mu_2)$. Kotz et al. (2000) call it the Kibble and Moran bivariate distribution. This bivariate gamma distribution is used by Henderson and Shimakura (2003) who apply a Poisson-gamma model in longitudinal data to account for individual-random effects and within-individual serial correlation. The case $\mu_{12} = 0$ corresponds to $\mathbf{P}(\mu_2 V_1 = \mu_1 V_2) = 1$ and the case $\mu_1\mu_2 - \mu_{12} = 0$ corresponds to independence between V_1 and V_2. It is easy to verify that $\mathcal{L}_G(s_1, s_2) \geq \mathcal{L}_{G_1}(s_1)\mathcal{L}_{G_2}(s_2)$ for all $(s_1, s_2) \in \mathbf{R}^2_+$, and therefore the random vector $(T_1, T_2)|x$ is PQD for each $x \in \mathcal{X}$.

Parameterization of G by using one of the two above distributions is convenient: although the corresponding densities have quite complicated expressions, the LT for each distribution has closed form expression which in turn gives a closed form expression for $S(t_1, t_2|x)$ as well. The main drawback of using one of these two bivariate distributions is that the $(T_1, T_2)|x$ is PQD, and consequently the $\rho(T_1, T_2|x)$ and $\tau(T_1, T_2|x)$ will be nonnegative for all $x \in \mathcal{X}$. On the other hand, Børing (2009) develops a three-parameter bivariate gamma distribution that allows for negative as well as positive correlation between V_1 and V_2. Clearly, its advantage compared to the two previous gamma distributions is that it also allows for negative correlation between V_1 and V_2 and the same for T_1 and T_2. However, we cannot say anything about quadrant dependence as negative (positive) correlation between V_1 and V_2 does not necessarily imply negative (positive) quadrant dependence between these two random variables (and consequently between T_1 and T_2). In the next subsection, we consider the notion of one-parameter copula for parameterizing G and as we shall discuss, it is possible that both types of quadrant dependence can be attained for the random variables V_1 and V_2.

3.2 Copula with gamma marginals

The advantage of using the copula approach is that it allows us to separate the bivariate distribution G into the marginals G_1, G_2 and an $\mathbf{R}$-valued pure dependence parameter ψ which captures the level of dependence between V_1 and V_2. Nelsen (2006) provides a detailed exposition of the important concept of copula.

According to the celebrated Sklar's theorem (Sklar, 1959) and given that the distributions G_1, G_2 are continuous functions, there exists a unique copula $C_\psi : [0,1]^2 \to [0,1]$ such that $G(v_1, v_2) = C_\psi(G_1(v_1), G_2(v_2))$ for all $(v_1, v_2) \in \mathbf{R}_+^2$. It is not difficult to see that C_ψ is the distribution of the random vector $(G_1(V_1), G_2(V_2))$. Conversely, for any given bivariate distribution G we can construct the corresponding copula by considering the quantity $G(G_1^{-1}(v_1), G_2^{-1}(v_2))$, where $G_i^{-1}(v) = \inf\{\omega \in \mathbf{R} : G_i(\omega) \geq v\}$ for $i = 1, 2$. Hence, we have $\psi = k_0$ for the double bivariate gamma and $\psi = \mu_{12}$ for the Kibble and Moran bivariate distribution.

It is well-known that the following Frechet bounds apply:

$$\max\{G_1(v_1) + G_2(v_2) - 1, 0\} \leq C_\psi(G_1(v_1), G_2(v_2)) \leq \min\{G_1(v_1), G_2(v_2)\} \tag{9}$$

for every $(v_1, v_2) \in \mathbf{R}_+^2$. When $C_\psi(G_1(v_1), G_2(v_2)) = \max\{G_1(v_1)+G_2(v_2)-1, 0\}$ for each $(v_1, v_2) \in \mathbf{R}_+^2$, it holds that $G_1(V_1) + G_2(V_2) - 1 = 0$ with probability one, and the random variables V_1 and V_2 are called countermonotonic. When $C_\psi(G_1(v_1), G_2(v_2)) = \min\{G_1(v_1), G_2(v_2)\}$ for all $(v_1, v_2) \in \mathbf{R}_+^2$, it holds $G_1(V_1) = G_2(V_2)$ with probability one, and the random variables V_1 and V_2 are called comonotonic. Equivalently, if C_ψ equals the lower (upper) Frechet bound, the random variable V_1 is a strictly decreasing (increasing) function of V_2. Note that when $\mathbf{P}(\mu_2 V_1 = \mu_1 V_2) = 1$, which implies $k_1 = k_2$, the G coincides with the upper Frechet bound, and thus both of the two bivariate gamma distributions that were studied in the previous subsection allow, in the limit, this probabilistic behavior.

The family of bivariate copulas that we could use to parameterize C_ψ is, for instance, either the Archimedean family or the Farlie-Gumbel-Morgenstern (FGM) family. In Appendix B, we provide a discussion about the functional form and dependence properties of three Archimedean copulas, Clayton, Frank, Gumbel, and also the FGM copula. Note that the Clayton copula we describe in Appendix B is a simple extension of the copula introduced by Clayton (1978). The three aforementioned Archimedean copulas are quite flexible in terms of positive dependence between V_1 and V_2 (and consequently, between T_1 and T_2) in the sense that they can be, in the limit, equal to the upper Frechet bound (9). Regarding negative dependence, the Gumbel copula does not admit a representation such that V_1 and V_2 are negatively dependent. However, the Clayton copula and the Frank copula allow for negative dependence, with the Frank copula converging towards the lower Frechet bound (9) for limiting values of the dependence parameter ψ. Note that the Clayton copula equals the lower Frechet bound for some certain value of the parameter ψ; however, if ψ converges towards this particular value the copula does not converge to the lower Frechet bound. On the other hand, the FGM copula does allow for both negative and positive dependence. But, its shortcoming is that it does not allow for strong (either positive or negative) dependence, that is, for any values of the parameter ψ, the Frechet bounds (9) cannot be approached.

4 Pearson's correlation coefficient

In this section we focus our attention on Pearson's correlation coefficient under the assumption of $\lambda_i(t,x) = \alpha_i t^{\alpha_i - 1}\varphi_i(x)$ for $i = 1,2$, with $\alpha_i > 0, t \in \mathbf{R}_+$ and $\varphi_i : \mathcal{X} \to (0,\infty)$. Namely, the hazard rates of the bivariate frailty model (1) are expressed as

$$\begin{aligned} \theta_1(t|x,V_1) &= \alpha_1 t^{\alpha_1 - 1}\varphi_1(x)V_1, \\ \theta_2(t|x,V_2) &= \alpha_2 t^{\alpha_2 - 1}\varphi_2(x)V_2. \end{aligned} \tag{10}$$

The specification (10), which is widely known as the Weibull bivariate frailty model, is a special case of the bivariate frailty model $\theta_i(t|x,V_i) = \tilde{\lambda}_i(t)\varphi_i(x)V_i$, where $\tilde{\lambda}_i$ is called baseline hazard and

φ_i is known as regressor function.

Next, we recall that

$$\rho(T_1, T_2|x) = \frac{\mathrm{Cov}(T_1, T_2|x)}{[\mathrm{Var}(T_1|x)\mathrm{Var}(T_2|x)]^{\frac{1}{2}}}, \quad x \in \mathcal{X}.$$

The covariance and the variance formulas are given by

$$\mathrm{Cov}(T_1, T_2|x) = \mathbf{E}\left[\mathbf{E}(T_1 T_2|x, V_1, V_2)\right] - \prod_{i=1}^{2} \mathbf{E}\left[\mathbf{E}(T_i|x, V_i)\right] \tag{11}$$

and

$$\mathrm{Var}(T_i|x) = \mathbf{E}\left[\mathrm{Var}(T_i|x, V_i)\right] + \mathrm{Var}\left[\mathbf{E}(T_i|x, V_i)\right] \tag{12}$$

for $i = 1, 2$, where the outer expectations and variance in the right-hand side of the two above equations are taken with respect to the distribution of the frailty terms. The term $\mathbf{E}\left[\mathrm{Var}(T_i|x, V_i)\right]$ captures the autonomous variation, whereas the term $\mathrm{Var}\left[\mathbf{E}(T_i|x, V_i)\right]$ captures the variation due to the presence of the frailty term. Under specification (10), the variable $T_i|x, V_i$ follows a Weibull distribution with shape parameter α_i and scale parameter $(\varphi_i(x)V_i)^{-\frac{1}{\alpha_i}}$ and thus $\mathbf{E}(T_i|x, V_i)$ and $\mathrm{Var}(T_i|x, V_i)$ are proportional to $V_i^{-\frac{1}{\alpha_i}}$ and $V_i^{-\frac{2}{\alpha_i}}$, respectively. Denote by ρ_{12} the Pearson's correlation coefficient between $V_1^{-\frac{1}{\alpha_1}}$ and $V_2^{-\frac{1}{\alpha_2}}$. Assuming that

$$\mathbf{E}\left(V_i^{-\frac{1}{\alpha_i}}\right) < \infty, \ \mathbf{E}\left(V_i^{-\frac{2}{\alpha_i}}\right) < \infty \text{ for each } \alpha_i > 0,$$

we write

$$\rho_{12} = \frac{\mathbf{E}\left(V_1^{-\frac{1}{\alpha_1}} V_2^{-\frac{1}{\alpha_2}}\right) - \mathbf{E}\left(V_1^{-\frac{1}{\alpha_1}}\right)\mathbf{E}\left(V_2^{-\frac{1}{\alpha_2}}\right)}{\prod_{i=1}^{2}\left[\mathbf{E}\left(V_i^{-\frac{2}{\alpha_i}}\right) - \left[\mathbf{E}\left(V_i^{-\frac{1}{\alpha_i}}\right)\right]^2\right]^{\frac{1}{2}}}. \tag{13}$$

After doing some algebra we can rewrite $\rho(T_1, T_2|x)$ as follows

$$\rho(T_1, T_2|x) = \rho_{12}\prod_{i=1}^{2}\left[\delta(\alpha_i) + (\delta(\alpha_i) - 1)\frac{\left[\mathbf{E}\left(V_i^{-\frac{1}{\alpha_i}}\right)\right]^2}{\mathbf{E}\left(V_i^{-\frac{2}{\alpha_i}}\right) - \left[\mathbf{E}\left(V_i^{-\frac{1}{\alpha_i}}\right)\right]^2}\right]^{-\frac{1}{2}}, \tag{14}$$

with

$$\delta(\alpha_i) = \frac{\left[\Gamma(1 + \frac{1}{\alpha_i})\right]^2}{\Gamma(1 + \frac{2}{\alpha_i})}, \qquad \alpha_i > 0. \tag{15}$$

The function δ is a strictly decreasing function in α_i, with $\lim_{\alpha_i \to 0} \delta(\alpha_i) = \infty$ and $\lim_{\alpha_i \to \infty} \delta(\alpha_i) = 1$. One important observation from (14) is that the value of $\rho(T_1, T_2|x)$ for fixed α_1 and α_2 depends on the strength of the linear relationship between the random variables $V_1^{-\frac{1}{\alpha_1}}$ and $V_2^{-\frac{1}{\alpha_2}}$ and not between the random variables V_1 and V_2. The latter is a consequence of the nonlinearity of the model (10).

Recall from Section 2.1 that if (V_1, V_2) is NQD (PQD) the $(T_1, T_2)|x$ is NQD (PQD) for any $x \in \mathcal{X}$ and therefore the $\rho(T_1, T_2|x)$ is nonpositive (nonnegative). This works in formula (14) by way of the term ρ_{12}. In particular, if (V_1, V_2) is NQD (PQD) the $(V_1^{-\frac{1}{\alpha_1}}, V_2^{-\frac{1}{\alpha_2}})$ is NQD (PQD) as well due the monotonic relationship between V_i and $V_i^{-\frac{1}{\alpha_i}}$ for each $\alpha_i > 0$, which in turn implies that ρ_{12} is nonpositive (nonnegative).

Define for $(\alpha_1, \alpha_2) \in (0, \infty)^2$

$$b_l(\alpha_1, \alpha_2) = -\frac{1}{[\delta(\alpha_1)\delta(\alpha_2)]^{\frac{1}{2}} + [(\delta(\alpha_1) - 1)(\delta(\alpha_2) - 1)]^{\frac{1}{2}}} \tag{16}$$

$$b_u(\alpha_1, \alpha_2) = \frac{1}{[\delta(\alpha_1)\delta(\alpha_2)]^{\frac{1}{2}}}. \tag{17}$$

As shown by Van den Berg (1997), for Weibull baseline hazards and any arbitrary joint distribution function of the random vector (V_1, V_2) the bounds for $\rho(T_1, T_2|x)$ are the following:

$$b_l(\alpha_1, \alpha_2) < \rho(T_1, T_2|x) < b_u(\alpha_1, \alpha_2) \tag{18}$$

for each pair $(\alpha_1, \alpha_2) \in (0, \infty)^2$. The bounds are tight for certain bivariate distributions of (V_1, V_2)

with discrete support; that is, they are approached arbitrarily closely. Given that $\delta(\alpha_i)$ is strictly decreasing in α_i, it is obvious from the above result that the range of possible values of $\rho(T_1, T_2|x)$ is increasing in α_i and thus the extreme values -1 and 1 are possible to obtain for $\alpha_i \to \infty$. This result can be explained as follows: to obtain maximum correlation it is required that the first type of variation $\mathbf{E}\left[\mathrm{Var}(T_i|x,V_i)\right]$ be minimal relative to the second type of variation $\mathrm{Var}\left[\mathbf{E}(T_i|x,V_i)\right]$ for each $= 1, 2$, and that the correlation between $V_1^{-\frac{1}{\alpha_1}}$ and $V_2^{-\frac{1}{\alpha_2}}$ be maximal. For $\alpha_i \to \infty$ the first type of variation decreases and is dominated by the second type, and thus it is possible to obtain any value in the interval $(-1, 1)$. Reverse statement will hold for $\alpha_i \to 0$.

Given that $V_i \sim Gamma(k_i, \mu_i)$, it can be easily shown that

$$\mathbf{E}\left(V_i^{-\frac{1}{\alpha_i}}\right) = \frac{\Gamma\left(k_i - \frac{1}{\alpha_i}\right)}{\Gamma(k_i)}\mu_i^{-\frac{1}{\alpha_i}}, \ \mathbf{E}\left(V_i^{-\frac{2}{\alpha_i}}\right) = \frac{\Gamma\left(k_i - \frac{2}{\alpha_i}\right)}{\Gamma(k_i)}\mu_i^{-\frac{2}{\alpha_i}}, \tag{19}$$

and therefore the restriction $k_i > \frac{2}{\alpha_i}$ is imposed so that the the first two moments of $V_i^{-\frac{1}{\alpha_i}}$ are defined for $i = 1, 2$. Then we can express $\rho(T_1, T_2|x)$ as follows:

$$\rho(T_1, T_2|x) = \rho_{12}\prod_{i=1}^{2}\left[\delta(\alpha_i) + (\delta(\alpha_i) - 1)\frac{\Gamma^2\left(k_i - \frac{1}{\alpha_i}\right)}{\Gamma\left(k_i - \frac{2}{\alpha_i}\right)\Gamma\left(k_i\right) - \Gamma^2\left(k_i - \frac{1}{\alpha_i}\right)}\right]^{-\frac{1}{2}}. \tag{20}$$

In the next two subsections we shall investigate how the assumption of gamma distributed frailties affect the behavior of $\rho(T_1, T_2|x)$. In particular, our interest is in studying whether the lower and upper bound of (18) can be arbitrarily approached in case the distribution of (V_1, V_2) has gamma marginals.

4.1 Lower bound for the Pearson's correlation coefficient

We first fix our attention on the lower bound of the linear correlation coefficient. The next proposition establishes a nonsharp (i.e., not necessarily attained) lower bound for the $\rho(T_1, T_2|x)$.

Proposition 2 *Suppose T_1 and T_2 are the duration variables that are generated by the bivariate frailty model (10), with $(\alpha_1, \alpha_2) \in (0, \infty)^2, V_1 \sim Gamma(k_1, \mu_1)$ and $V_2 \sim Gamma(k_2, \mu_2)$. Then,* *the following inequality holds*

$$\rho(T_1, T_2|x) \geq b_{gl}(\alpha_1, \alpha_2), \quad x \in \mathcal{X},$$

with

$$b_{gl}(\alpha_1,\alpha_2) = \min_{\substack{k_1 > \frac{2}{\alpha_1} \\ k_2 > \frac{2}{\alpha_2}}} \left[\prod_{i=1}^{2} \frac{\Gamma(k_i)}{\Gamma\left(k_i + \frac{1}{\alpha_i}\right)} - \prod_{i=1}^{2} \frac{\Gamma\left(k_i - \frac{1}{\alpha_i}\right)}{\Gamma(k_i)} \right] \prod_{i=1}^{2} \left[\frac{\Gamma\left(k_i - \frac{2}{\alpha_i}\right)}{\Gamma(k_i)} - \frac{\Gamma^2\left(k_i - \frac{1}{\alpha_i}\right)}{\Gamma^2(k_i)} \right]^{-\frac{1}{2}}$$

$$\times \prod_{i=1}^{2} \left[\delta(\alpha_i) + (\delta(\alpha_i) - 1) \frac{\Gamma^2\left(k_i - \frac{1}{\alpha_i}\right)}{\Gamma\left(k_i - \frac{2}{\alpha_i}\right)\Gamma(k_i) - \Gamma^2\left(k_i - \frac{1}{\alpha_i}\right)} \right]^{-\frac{1}{2}}.$$

The next table lists the bounds $b_l(\alpha_1, \alpha_2)$ and $b_{gl}(\alpha_1, \alpha_2)$ for different values of α_1, α_2. To make the comparison between $b_l(\alpha_1, \alpha_2)$ and $b_{gl}(\alpha_1, \alpha_2)$ more transparent, all numbers have been rounded off to three decimal digits.

(α_1, α_2)	$b_l(\alpha_1, \alpha_2)$	$b_{gl}(\alpha_1, \alpha_2)$
$(0.5, 1)$	-0.175	-0.125
$(0.5, 2)$	-0.254	-0.233
$(1, 1)$	-0.333	-0.220
$(1, 2)$	-0.472	-0.366
$(1, 3)$	-0.535	-0.520
$(1.5, 2)$	-0.582	-0.397
$(2, 2)$	-0.647	-0.451
$(2, 3)$	-0.719	-0.580
$(4, 4)$	-0.860	-0.590
$(5, 5)$	-0.860	-0.599

Table 1: $b_l(\alpha_1, \alpha_2)$ and $b_{gl}(\alpha_1, \alpha_2)$ values.

In view of the results of Table 1, we can claim that the bound $b_{gl}(\alpha_1, \alpha_2)$ is generally closer to zero than the bound $b_l(\alpha_1, \alpha_2)$. These results reveal a limitation of the the bivariate Weibull gamma frailty model to fit data with relatively large negative dependence between the duration variables. Note that the bound $b_{gl}(\alpha_1, \alpha_2)$ is not expected to be tight as three successive inequalities were employed to derive it. In fact, there could be values of α_1, α_2 such that $b_l(\alpha_1, \alpha_2) > b_{gl}(\alpha_1, \alpha_2)$; however, this is clearly due to the use of the three inequalities as $b_l(\alpha_1, \alpha_2)$ covers all the bivariate distributions with support on $\mathbf{R}_+^2$ and trivially all the bivariate distributions with gamma marginals.

To improve the lower bound for the exponential case (i.e., $\alpha_1 = \alpha_2 = 1$) we carry out Monte Carlo simulation. For the exponential model we have $\rho(T_1, T_2|x) = \rho_{12}\left(\sqrt{k_1 k_2}\right)^{-1}$. For given marginals G_1 and G_2, ρ_{12} will be minimized if and only if the distribution of (V_1^{-1}, V_2^{-1}) is equal to the lower Frechet bound. However, due to the fact that V_i^{-1} is strictly decreasing transformation of V_i, the ρ_{12} will be minimized for fixed G_1 and G_2 if and only if $G(v_1, v_2) = \max\{G_1(v_1) + G_2(v_2) - 1, 0\}$ for each $(v_1, v_2) \in \mathbf{R}_+^2$. In case G is parameterized by the Frank copula, the lower Frechet bound can be approached very well for limiting values of the dependence parameter. To derive an estimation of the minimum value of ρ_{12} for fixed k_1 and k_2, we draw gamma random variables V_1 and V_2 by using the relationship $G_1(V_1) + G_2(V_2) - 1 = 0$.

For the study of the values of ρ_{12} and $\rho(T_1, T_2|x)$ we present two figures. The first figure shows values of ρ_{12} as a function of k_1 and k_2. Note that we have reversed the axes with the values of k_1 and k_2 so that we have a clearer picture.

The second figure displays the values of $\rho(T_1, T_2|x)$ as a function of k_1 and k_2.

The estimated value of the lower bound is about -0.14, which is clearly much closer to zero than the tight bound $-\frac{1}{3}$. From the two above graphs we can easily notice the two opposite effects of the value of the shape parameters on the values of ρ_{12} and $\rho(T_1, T_2|x)$. More precisely, ρ_{12} approaches arbitrarily closely the value -1 for large values of k_1 and k_2. However, large values of the shape parameters weaken the linear relationship between the duration variables as the variation of the random variable $T_i|x$-due to the presence of the frailty-is negligible with respect to the autonomous variation. To see this, consider for simplicity the case $k_1 = k_2 = k$. Then, we obtain $\mathbf{E}\left[\mathrm{Var}(T_i|x, V_i)\right] = ((k-1)k)^{-1} = O(k^{-2})$ and $\mathrm{Var}\left[\mathbf{E}(T_i|x, V_i)\right] = ((k-1)^2(k-2)) =$

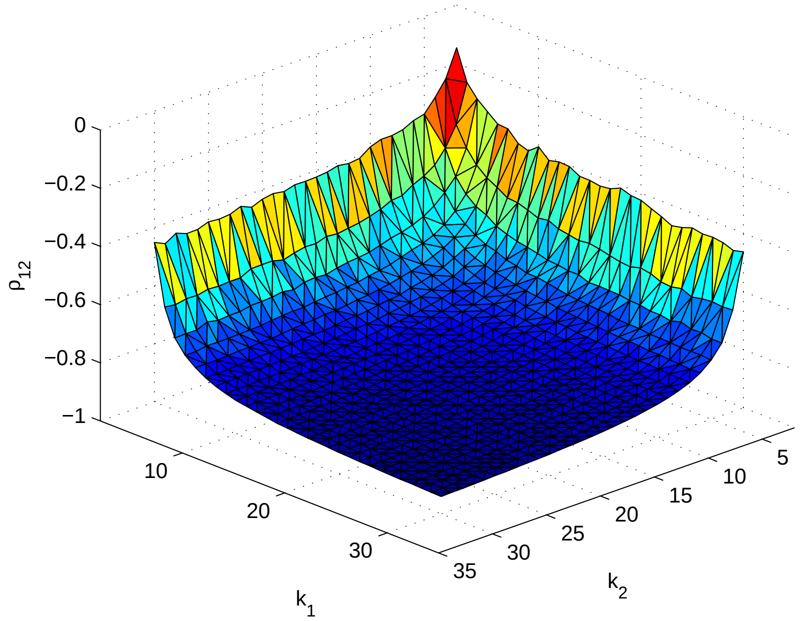

Figure 1: Plot of ρ_{12} as a function of k_1 and k_2, if $\alpha_1 = \alpha_2 = 1$.

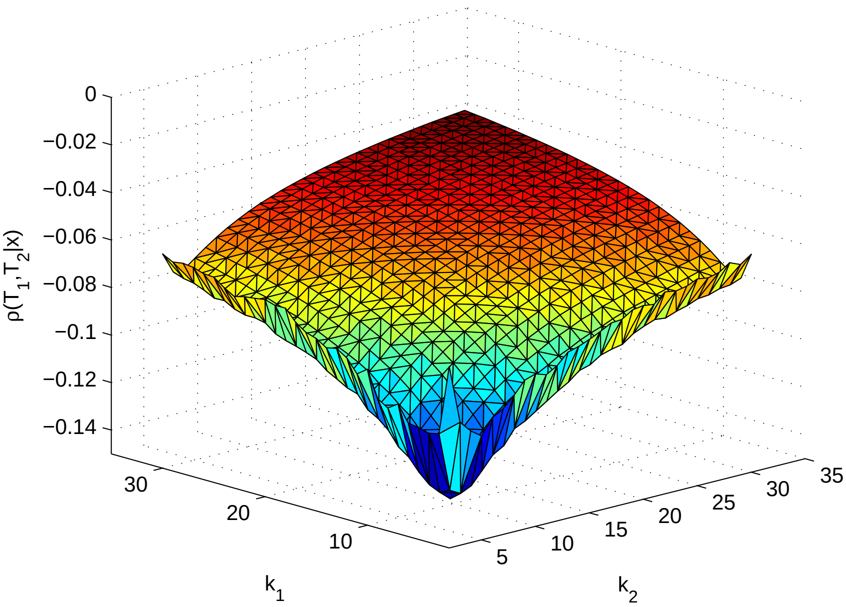

Figure 2: Plot of $\rho(T_1, T_2|x)$ as a function of k_1 and k_2, if $\alpha_1 = \alpha_2 = 1$.

$O(k^{-3}) = o(k^{-2})$ for $k \to \infty$ and $i = 1, 2$.

Next, we consider three other possible families of distributions for G with marginals different from gamma. In particular, Mardia (1970) shows that if the random vector (V_1^{-1}, V_2^{-1}) follows the Filon-Isserk bivariate Beta distribution, the $\rho(T_1, T_2|x)$ can attain any values in the interval $(-\frac{1}{3}, 0]$.

Moreover, Van den Berg (1997) shows that if $(V_i)^{-1} = \sum_{j=1}^{k} U_{ij}^2$ for $i = 1, 2$ and some finite positive integer k, where the vector (U_{1j}, U_{2j}) follows a bivariate normal distribution, the lower bound of $\rho(T_1, T_2|x)$ is about -0.23. Finally, Van den Berg (1997) shows that if $V_i = \exp(\eta_{i0} + \eta_{i1}\mathcal{N})$, where $\eta_{i0} \in \mathbf{R}$ and $\eta_{i1} \in \mathbf{R}\backslash\{0\}$ for $i = 1, 2$ and $\mathcal{N}$ is a normally distributed random variable, the lower bound of $\rho(T_1, T_2|x)$ is about -0.17. In view of these results and using as criterion the bounds for Pearson's correlation coefficient, the assumption that the distribution of (V_1, V_2) is characterized by gamma marginals seems quite restrictive for attaining large negative values.

4.2 Upper bound for Pearson's correlation coefficient

We now concentrate on the bivariate frailty model that has the property $\mathbf{P}(\mu_2 V_1 = \mu_1 V_2) = 1$, which in turn implies $G(v_1, v_2) = \min\{G_1(v_1), G_2(v_2)\}$ for all $(v_1, v_2) \in \mathbf{R}_+^2$ and $k_1 = k_2 = k$. Under the assumption of identical Weibull baseline hazards-that is, $\alpha_1 = \alpha_2 = \alpha$- we have $\rho_{12} \to 1$ for any $k > \frac{2}{\alpha}$. Also, for $k \to \frac{2}{\alpha}$ and given that $\lim_{k \to \frac{2}{\alpha}} \Gamma^2\left(k - \frac{2}{\alpha}\right) \to \infty$, we get by (20)

$$\rho(T_1, T_2|x) \to \frac{\left[\Gamma(1 + \frac{1}{\alpha})\right]^2}{\Gamma(1 + \frac{2}{\alpha})} = b_u(\alpha, \alpha).$$

Therefore, if $\alpha_1 = \alpha_2 = \alpha$ the upper bound of (18) can be arbitrarily approached in case G is equal either to one of the two bivariate gamma distributions of Section 3.1 or to one of the three Archimedean copulas described in detail in Appendix B.

Next, we turn our attention to the case $\alpha_1 \neq \alpha_2$ and $V_i \sim Gamma(k, \mu_i)$ for $i = 1, 2$; that is, $k_1 = k_2 = k$. Although imposing the assumption that both marginals have the same shape parameter may seem restrictive, it is rather general. In particular, it includes as special cases the bivariate frailty model in which $\mathbf{P}(\mu_2 V_1 = \mu_1 V_2) = 1$ that we described above for $\alpha_1 = \alpha_2$ and also the bivariate frailty model for which (V_1, V_2) is distributed according to the Kibble and Moran bivariate gamma distribution. The next proposition analytically establishes a nonsharp bound for this case that is strictly smaller than the nonparametric upper bound (18).

Proposition 3 *Let T_1 and T_2 be the duration variables that are generated by the bivariate frailty model (10) with $\alpha_1 > \alpha_2 > 0, V_1 \sim Gamma(k, \mu_1)$ and $V_2 \sim Gamma(k, \mu_2)$. Then the following* *inequality holds*

$$\rho(T_1, T_2|x) < b_{gu}(\alpha_1, \alpha_2), \quad x \in \mathcal{X},$$

with

$$b_{gu}(\alpha_1, \alpha_2) = \frac{1}{[\delta(\alpha_2)]^{\frac{1}{2}}} \left[\delta(\alpha_1) + (\delta(\alpha_1) - 1) \frac{\Gamma^2\left(\frac{2\alpha_1 - \alpha_2}{\alpha_1\alpha_2}\right)}{\Gamma\left(\frac{2(\alpha_1-\alpha_2)}{\alpha_1\alpha_2}\right)\Gamma\left(\frac{2}{\alpha_2}\right) - \Gamma^2\left(\frac{2\alpha_1-\alpha_2}{\alpha_1\alpha_2}\right)} \right]^{-\frac{1}{2}} < b_u(\alpha_1, \alpha_2).$$

The next table reports the bounds $b_u(\alpha_1, \alpha_2)$ and $b_{gu}(\alpha_1, \alpha_2)$ for different values of α_1, α_2, with $\alpha_1 > \alpha_2$. Like in the case with the lower bound, we have rounded all the numbers off to three decimal points.

(α_1, α_2)	$b_u(\alpha_1, \alpha_2)$	$b_{gu}(\alpha_1, \alpha_2)$
$(0.5, 0.25)$	0.049	0.037
$(0.75, 0.25)$	0.071	0.041
$(1, 0.5)$	0.289	0.204
$(2, 0.5)$	0.362	0.194
$(2, 1)$	0.627	0.469
$(5, 1)$	0.689	0.423
$(5, 2)$	0.864	0.704
$(10, 2)$	0.880	0.669
$(10, 5)$	0.968	0.921
$(20, 10)$	0.999	0.976

Table 2: $b_u(\alpha_1, \alpha_2)$ and $b_{gu}(\alpha_1, \alpha_2)$ values.

The reason that $b_{gu}(\alpha_1, \alpha_2) < b_u(\alpha_1, \alpha_2)$ is that the shape parameter k is bounded from below by the maximum between the values of the ratios $\frac{2}{\alpha_1}$ and $\frac{2}{\alpha_2}$ so that the first two moments of $V_i^{-\frac{1}{\alpha_i}}$ for $i = 1, 2$ are defined. Moreover, the bound of Proposition 3 is not attained as the gamma distribution is not closed under power transformation. In particular, if $V_1 \sim Gamma(k, \mu_1)$ the random variable $V_1^{\frac{\alpha_2}{\alpha_1}}$, for any fixed positive α_1, α_2 with $\alpha_1 \neq \alpha_2$, does not follow a gamma distribution, and this implies that we cannot have $\rho_{12} = 1$ such that V_1 and V_2 are gamma distributed. Hence, even if $\mathbf{P}(\mu_2 V_1 = \mu_1 V_2) = 1$ we will always have $\rho_{12} < 1$ for any fixed values of α_1, α_2, with $\alpha_1 \neq \alpha_2$.

5 Kendall's tau

We now proceed with the derivation of bounds for the range of values of the Kendall's tau as the results of Van den Berg (1997) do not directly carry over to the bivariate gamma frailty model. As explained in Section 2, for two independent copies, $(T_1^A, T_2^A)|x$ and $(T_1^B, T_2^B)|x$, of the bivariate random vector $(T_1, T_2)|x$ we have

$$\tau(T_1, T_2|x) = 2\mathbf{P}\left[(T_1^A - T_1^B)(T_2^A - T_2^B) > 0|x\right] - 1, \quad x \in \mathcal{X}. \tag{21}$$

In contrast to the Pearson's coefficient case, we will not assume anything about the range of values of the shape parameters. Also, we will not impose any condition on the functional form of λ_i except for the limiting result $\lim_{t\to\infty} \int_0^t \lambda_i(\omega, x)d\omega = \infty$ $(i = 1, 2)$. We will make use of the equality

$$\ln V_i = -\ln \Lambda_i(T_i, x) + \epsilon_i, \quad i = 1, 2, \tag{22}$$

with ϵ_1, ϵ_2 being independent random variables that have probability density function $f_i(\epsilon) = e^{\epsilon} \exp\left(-e^{\epsilon}\right)$. The above equation is an equivalent representation of (1). Also, recall that $S_i(t|x) = \mathcal{L}_{G_i}\left(\Lambda_i(t, x)\right)$ for $(t, x) \in \mathbf{R}_+ \times \mathcal{X}$. Provided that $V_i \sim Gamma(k_i, \mu_i)$, it follows $S_i(t|x) = (1 + \mu_i \Lambda_i(t, x))^{-k_i}$. Therefore, the stochastic duration T_i can be expressed in structural form as follows

$$T_i = \Lambda_i^{-1}\left(\frac{1}{\mu_i} U_i^{-\frac{1}{k_i}} - \frac{1}{\mu_i}, x\right), \quad U_i \sim Uniform(0, 1), \quad i = 1, 2. \tag{23}$$

We first focus on the lower bound of the values of $\tau(T_1, T_2|x)$. We assume that $G(v_1, v_2) = \max\{G_1(v_1) + G_2(v_2) - 1, 0\}$ for each $(v_1, v_2) \in \mathbf{R}_+^2$. This implies that $G_1(V_1) + G_2(V_2) - 1 = 0$ with probability one. Hence, V_2 is a strictly decreasing transformation of V_1 and we can write, by

(22),

$$\ln \Lambda_2(T_2, x) = \mathcal{H}(T_1, \epsilon_1, \epsilon_2, x), \tag{24}$$

where $\mathcal{H}(., \epsilon_1, \epsilon_2, x)$ is a strictly decreasing function, $\lim_{t\to\infty} \mathcal{H}(t, \epsilon_1, \epsilon_2, x) = h(t, x)$ for all $(\epsilon_1, \epsilon_2, x) \in \mathbf{R}^2 \times \mathcal{X}$, and $h(., x)$ is a strictly decreasing function. By using the rank-invariant property of Kendall's tau and combining (21) and (24), we have

$$\tau(T_1, T_2|x) = 2\mathbf{P}\left[(T_1^A - T_1^B)(\mathcal{H}(T_1^A, \epsilon_1^A, \epsilon_2^A, x) - \mathcal{H}(T_1^B, \epsilon_1^B, \epsilon_2^B, x)) > 0\right] - 1. \tag{25}$$

Clearly, $\tau(T_1, T_2|x)$ can be also written as follows

$$\begin{aligned} \tau(T_1, T_2|x) = 2\mathbf{P}&\left[\left\{(T_1^A - T_1^B) > 0\right\} \cap \left\{\mathcal{H}(T_1^A, \epsilon_1^A, \epsilon_2^A, x) - \mathcal{H}(T_1^B, \epsilon_1^B, \epsilon_2^B, x) > 0\right\}\right] + \\ &2\mathbf{P}\left[\left\{(T_1^A - T_1^B) < 0\right\} \cap \left\{\mathcal{H}(T_1^A, \epsilon_1^A, \epsilon_2^A, x) - \mathcal{H}(T_1^B, \epsilon_1^B, \epsilon_2^B, x) < 0\right\}\right] - 1. \end{aligned} \tag{26}$$

For $k_1 \to 0$ and $\mu_1 = O(k_1^{-1})$ we have $T_1^A \to \infty$ and $T_1^B \to \infty$ which yield $\{\mathcal{H}(T_1^A, \epsilon_1^A, \epsilon_2^A, x) - \mathcal{H}(T_1^B, \epsilon_1^B, \epsilon_2^B, x) > 0\} \to \{h(T_1^A, x) - h(T_1^B, x) > 0\} = \{T_1^A - T_1^B < 0\}$ and $\{\mathcal{H}(T_1^A, \epsilon_1^A, \epsilon_2^A, x) - \mathcal{H}(T_1^B, \epsilon_1^B, \epsilon_2^B, x) < 0\} \to \{h(T_1^A, x) - h(T_1^B, x) < 0\} = \{T_1^A - T_1^B > 0\}$. By making use of these limiting statements, it is obvious, by using (26), that $\tau(T_1, T_2|x) \to -1$.

To derive the conditions needed to be satisfied for the upper bound of the $\tau(T_1, T_2|x)$ values, we require that G be equal to the upper Frechet bound, namely, $G(v_1, v_2) = \min\{G_1(v_1), G_2(v_2)\}$ for each $(v_1, v_2) \in \mathbf{R}_+^2$. Under this scenario, $G_1(V_1) = G_2(V_2)$ with probability one. Thus, V_2 is a strictly increasing transformation of V_1 and therefore we can write, by (22),

$$\ln \Lambda_2(T_2, x) = \mathcal{Y}(T_1, \epsilon_1, \epsilon_2, x), \tag{27}$$

where $\mathcal{Y}(., \epsilon_1, \epsilon_2, x)$ is a strictly increasing function and $\lim_{t\to\infty} \mathcal{Y}(t, \epsilon_1, \epsilon_2, x) = y(t, x)$ for all $(\epsilon_1, \epsilon_2, x) \in \mathbf{R}^2 \times \mathcal{X}$, and $y(., x)$ is some strictly increasing function. Performing identical calculations to the ones of the previous paragraph we obtain

$$\tau(T_1,T_2|x) = 2\mathbf{P}\left[\left\{(T_1^A - T_1^B) > 0\right\} \cap \left\{\mathcal{Y}(T_1^A,\epsilon_1^A,\epsilon_2^A,x) - \mathcal{Y}(T_1^B,\epsilon_1^B,\epsilon_2^B,x) > 0\right\}\right] + 2\mathbf{P}\left[\left\{(T_1^A - T_1^B) < 0\right\} \cap \left\{\mathcal{Y}(T_1^A,\epsilon_1^A,\epsilon_2^A,x) - \mathcal{Y}(T_1^B,\epsilon_1^B,\epsilon_2^B,x) < 0\right\}\right] - 1. \tag{28}$$

For $k_1 \to 0$ and $\mu_1 = O(k_1^{-1})$ we obtain $T_1^A \to \infty$ and $T_1^B \to \infty$ which in turn gives $\{\mathcal{Y}(T_1^A,\epsilon_1^A,\epsilon_2^A,x) - \mathcal{Y}(T_1^B,\epsilon_1^B,\epsilon_2^B,x) > 0\} \to \{y(T_1^A,x) - y(T_1^B,x) > 0\} = \{T_1^A - T_1^B > 0\}$ and $\{\mathcal{Y}(T_1^A,\epsilon_1^A,\epsilon_2^A,x) - \mathcal{Y}(T_1^B,\epsilon_1^B,\epsilon_2^B,x) < 0\} \to \{y(T_1^A,x) - y(T_1^B,x) < 0\} = \{T_1^A - T_1^B < 0\}$. Given the equality $\mathbf{P}\left[(T_1^A - T_1^B) > 0|x\right] = \mathbf{P}\left[(T_1^A - T_1^B) < 0|x\right] = \frac{1}{2}$ for all $x \in \mathcal{X}$ and making use of (28), the limiting result $\tau(T_1,T_2|x) \to 1$ is obtained.

We summarize the above discussion to the next proposition.

Proposition 4 *Suppose T_1 and T_2 are the duration variables that are generated by the bivariate frailty model (1) with $V_1 \sim Gamma(k_1,\mu_1)$ and $V_2 \sim Gamma(k_2,\mu_2)$. Then the following double inequality holds:*

$$-1 < \tau(T_1,T_2|x) < 1, \quad x \in \mathcal{X}.$$

The extreme bounds -1 and 1 are tight in the sense that they can be approached arbitrarily closely. More precisely, if $G(v_1,v_2) = \max\{G_1(v_1) + G_2(v_2) - 1, 0\}$ for each $(v_1,v_2) \in \mathbf{R}_+^2$ and $k_1 \to 0$ with $\mu_1 = O(k_1^{-1})$, or $k_2 \to 0$ with $\mu_2 = O(k_2^{-1})$, we obtain $\tau(T_1,T_2|x) \to -1$. On the other hand, if $G(v_1,v_2) = \min\{G_1(v_1), G_2(v_2)\}$ for each $(v_1,v_2) \in \mathbf{R}_+^2$ and $k_1 \to 0$ with $\mu_1 = O(k_1^{-1})$, or $k_2 \to 0$ with $\mu_2 = O(k_2^{-1})$, then $\tau(T_1,T_2|x) \to 1$.

Therefore, by assuming gamma marginals for the distribution of (V_1,V_2) a necessary condition for approaching the lower bound of $\tau(T_1,T_2|x)$ is the distribution of (V_1,V_2) be equal to the Frank copula. On the other hand, the upper bound of $\tau(T_1,T_2|x)$ can be approached arbitrarily closely if the bivariate distribution is modelled by the two bivariate gamma distributions of Section 3.1 or one of the three Archimedean copulas presented in Appendix B. Note that if $\mathbf{P}(\mu_2 V_1 = \mu_1 V_2) = 1$, which clearly gives $G(v_1,v_2) = \min\{G_1(v_1), G_2(v_2)\}$ for each $(v_1,v_2) \in \mathbf{R}_+^2$, we will have $k_1 = k_2 \to 0$.

By applying results of Embrechts et al. (2002), we have that $\tau(T_1,T_2|x) \to -1$ if and only if $S(t_1, t_2|x) \to \max\{S_1(t_1|x) + S_2(t_2|x) - 1, 0\}$ for all $(t_1, t_2, x) \in \mathbf{R}_+^2 \times \mathcal{X}$, or equivalently, $S_1(T_1|x) + S_2(T_2|x) - 1 = 0$ for all $x \in \mathcal{X}$ with probability approaching one. On the other hand, $\tau(T_1, T_2|x) \to 1$ if and only if $S(t_1, t_2|x) \to \min\{S_1(t_1|x), S_2(t_2|x)\}$ for all $(t_1, t_2, x) \in \mathbf{R}_+^2 \times \mathcal{X}$, or equivalently, $S_1(T_1|x) = S_2(T_2|x)$ for all $x \in \mathcal{X}$ with probability approaching one. Hence, in view of Proposition 1, the condition in Proposition 4 that G is equal to the lower (upper) Frechet bound is indispensable. We should also point out here that $S(t_1, t_2|x)$ can be written in a copula form as a function only of $S_1(t_1|x)$ and $S_2(t_2|x)$ and not of x because

$$S(t_1, t_2|x) = \mathcal{L}_G\left(\mathcal{L}_{G_1}^{-1}(S_1(t_1|x)), \mathcal{L}_{G_2}^{-1}(S_2(t_2|x))\right), \quad (t_1, t_2, x) \in \mathbf{R}_+^2 \times \mathcal{X},$$

where $\mathcal{L}^{-1}$ denotes the inverse of the LT of the corresponding probability measure.

6 Conclusions

We examine the dependence structure in bivariate frailty models in which the duration variables are dependent by way of the frailty terms. We first show that if the distribution of the frailty terms is negative (positive) quadrant dependent, then the conditional, on observed characteristics, joint survival function of the duration outcomes is negative (positive) quadrant dependent as well. To quantify the level of dependence between the duration variables, we consider Pearson's correlation coefficient and Kendall's tau. We provide bounds for the range of values of these measures under the assumption of gamma distributed frailty terms. To model the dependence structure between the frailty terms, we can use either standard bivariate gamma distributions or copulas with gamma marginals. The former induce only positive dependence between the duration variables, whereas the latter can induce positive and/or negative dependence. Strong negative (positive) dependence between the duration outcomes can be generated by bivariate distributions of the frailty terms which can be, in the limit, equal to the lower (upper) Frechet bound.

We calculate bounds for the values of Pearson's correlation coefficient if the baseline hazards have a Weibull specification. Regarding the negative values, we analytically provide a nonsharp lower bound. We improve the lower bound for the exponential case by means of Monte Carlo simulation. The resulting lower bound is closer to zero than its nonparametric analogue which is derived by Van den Berg (1997). For positive values of Pearson's coefficient we show that the upper bound of Van den Berg (1997) can be approached arbitrarily closely in case the Weibull specifications are identical. Moreover, we provide an upper bound for different Weibull specifications which is strictly smaller than the nonparametric bound. The resulting bound cannot be attained due to the fact that the gamma distribution is not closed under power transformation.

In contrast to Pearson's correlation coefficient, Kendall's tau can take any value in the interval $(-1, 1)$ regardless of the functional form specification about the hazard rates. If the bivariate distribution of the frailty terms approaches the lower (upper) Frechet bound and the first moment of the frailty term(s) is finite, then the lower (upper) bound can be approached arbitrarily closely. In particular, we should impose the condition that one of the two shape parameters converges towards zero.

In terms of practical choices for functional forms, we make the following recommendations. First, if the interest is models that are able to capture a negative association between say two duration variables, and if the researcher wants to restrict him/herself to bivariate gamma frailty distributions, then the researcher should choose the bivariate gamma distribution that is based on the Frank copula. This specification allows for a larger range of negative associations than other specifications. Secondly, if the researcher wants to capture the largest possible range of negative associations between the duration variables regardless of the functional form of the frailty distribution, then the researcher should refrain from using bivariate gamma frailty distributions and instead adopt discrete frailty distributions. The latter are known to provide maximum flexibility in terms of association (Van den Berg, 2007).

A fruitful topic for future research is the study of bounds for the two association measures in bivariate duration models where the two duration variables are parallel and the realization of one of these two variables affects the hazard rate of the other. Moreover, a promising topic for investigation is the study of the range of values for local measures of dependence such as the cross-ratio function (Clayton, 1978). Finally, it is of practical relevance to consider the concepts of lower tail and upper tail dependence between the duration variables. In particular, if the data display dependence between extreme values of the duration variables, we should know which bivariate distributions for the frailty terms allow such a dependence pattern.

Appendix A

This appendix presents the mathematical proofs for the first three propositions in the main text.

Proof of Proposition 1. By definition

$$S^j(t_1,t_2|x)=\int_{\mathbf{R}_+^2} S(t_1,t_2|x,v_1,v_2)dG^j(v_1,v_2), \quad j=a,b, \tag{A-1}$$

where $S(t_1,t_2|x,v_1,v_2)=\exp\left(-\Lambda_1(t_1,x)v_1-\Lambda_2(t_2,x)v_2\right)$ for $(t_1,t_2,x,v_1,v_2)\in\mathbf{R}_+^2\times\mathcal{X}\times\mathbf{R}_+^2$. The integrand is a continuous bounded function in (v_1,v_2) for any $(t_1,t_2,x)\in\mathbf{R}_+^2\times\mathcal{X}$. Moreover, it holds that $\frac{\vartheta^2}{\vartheta v_1\vartheta v_2}S(t_1,t_2|x,v_1,v_2)>0$ for each $(t_1,t_2,x,v_1,v_2)\in\mathbf{R}_+^2\times\mathcal{X}\times(0,\infty)^2$ (i.e., the $S(t_1,t_2|x,v_1,v_2)$ is a $2-$positive function in v_1,v_2). Given that $G^a\prec_C G^b$, we obtain the inequality $S^a(t_1,t_2|x)\leq S^b(t_1,t_2|x)$ for all $(t_1,t_2,x)\in\mathbf{R}_+^2\times\mathcal{X}$ by Theorem 2 of Tchen (1980). Recall also that

$$S_i(t|x)=\int_{\mathbf{R}_+}\exp\left(-\Lambda_i(t,x)v\right)dG_i(v), \quad i=1,2. \tag{A-2}$$

Provided that G^a and G^b are characterized by the fixed marginals G_1 and G_2, it follows that the bivariate survival functions S^a and S^b are characterized by the same marginals, S_1 and S_2. This in turn implies that $S^a\prec_C S^b$ for each $x\in\mathcal{X}$. ■

Define for each $\varepsilon>0$ the digamma function

$$\psi(\varepsilon)=\frac{\Gamma'(\varepsilon)}{\Gamma(\varepsilon)} \tag{A-3}$$

and the polygamma function

$$\psi^{(n)}(\varepsilon)=\frac{d^n\psi(\varepsilon)}{d\varepsilon^n}, \quad n\in\mathbf{N}, \tag{A-4}$$

with $\psi^{(0)}(.)=\psi(.)$. Moreover, it holds that

$$\psi^{(n)}(\varepsilon)=(-1)^{n+1}\int_{\mathbf{R}_+}\frac{t^n}{1-e^{-t}}e^{-\varepsilon t}dt, \quad \varepsilon>0. \tag{A-5}$$

We state Lemma 1 which is needed for the proof of Proposition 2 and 3. Its simple proof, which makes use of (A-4) and (A-5), is omitted.

Lemma 1 *Let* $(\varepsilon_1, \varepsilon_2, \varepsilon_3) \in (0, \infty)^3$. *Then,*

$$\Gamma(\varepsilon_1)\Gamma(\varepsilon_1+\varepsilon_2+\varepsilon_3) - \Gamma(\varepsilon_1+\varepsilon_2)\Gamma(\varepsilon_1+\varepsilon_3) > 0.$$

Proof of Proposition 2. Recall that

$$\rho(T_1, T_2|x) = \rho_{12}\prod_{i=1}^{2}\left[\delta(\alpha_i) + (\delta(\alpha_i)-1)\frac{\Gamma^2\left(k_i - \frac{1}{\alpha_i}\right)}{\Gamma\left(k_i - \frac{2}{\alpha_i}\right)\Gamma(k_i) - \Gamma^2\left(k_i - \frac{1}{\alpha_i}\right)}\right]^{-\frac{1}{2}} \tag{A-6}$$

for $x \in \mathcal{X}$, where

$$\rho_{12} = \frac{\mathbf{E}\left(V_1^{-\frac{1}{\alpha_1}}V_2^{-\frac{1}{\alpha_2}}\right) - \mathbf{E}\left(V_1^{-\frac{1}{\alpha_1}}\right)\mathbf{E}\left(V_2^{-\frac{1}{\alpha_2}}\right)}{\prod_{i=1}^{2}\left[\mathbf{E}\left(V_i^{-\frac{2}{\alpha_i}}\right) - \left[\mathbf{E}\left(V_i^{-\frac{1}{\alpha_i}}\right)\right]^2\right]^{\frac{1}{2}}}. \tag{A-7}$$

Note that by Lemma 1 we get $\Gamma\left(k_i - \frac{2}{\alpha_i}\right)\Gamma(k_i) - \Gamma^2\left(k_i - \frac{1}{\alpha_i}\right) > 0$ for $\varepsilon_1 = k_i - \frac{2}{\alpha_i}$ and $\varepsilon_2 = \varepsilon_3 = \frac{1}{\alpha_i}$, with $k_i > \frac{2}{\alpha_i}$. Given also that $\delta(\alpha_i) > 1$ for each $\alpha_i > 0$, our problem reduces to bound from below the numerator of (A-7), for fixed marginals G_1, G_2, .

Denote by $\mathbf{E}_l$ the expectation with respect to the probability measure $\max\{G_1(v_1) + G_2(v_2) - 1, 0\}$. By using the formula for the covariance and employing Hoeffding's identity, we get

$$\mathbf{E}\left(V_1^{-\frac{1}{\alpha_1}}V_2^{-\frac{1}{\alpha_2}}\right) \geq \mathbf{E}_l\left(V_1^{-\frac{1}{\alpha_1}}V_2^{-\frac{1}{\alpha_2}}\right). \tag{A-8}$$

The mapping $\omega \mapsto (\omega)^{-1}$ is strictly convex and thus Jensen's inequality entails

$$\mathbf{E}_l\left(V_1^{-\frac{1}{\alpha_1}}V_2^{-\frac{1}{\alpha_2}}\right) \geq \left[\mathbf{E}_l\left(V_1^{\frac{1}{\alpha_1}}V_2^{\frac{1}{\alpha_2}}\right)\right]^{-1}, \tag{A-9}$$

which together with (A-8) implies

$$\mathbf{E}\left(V_1^{-\frac{1}{\alpha_1}}V_2^{-\frac{1}{\alpha_2}}\right) \geq \left[\mathbf{E}_l\left(V_1^{\frac{1}{\alpha_1}}V_2^{\frac{1}{\alpha_2}}\right)\right]^{-1} \tag{A-10}$$

For $G = \max\{G_1(v_1) + G_2(v_2) - 1, 0\}$, the random vector (V_1, V_2) is NQD, which in turn gives that the $(V_1^{\frac{1}{\alpha_1}}, V_2^{\frac{1}{\alpha_2}})$ is NQD as well due to the fact that $V_i^{\frac{1}{\alpha_i}}$ is strictly increasing transformation of V_i for $i = 1, 2$. Using again the formula of the covariance and Hoeffding's identity we get

$$\mathbf{E}_l \left(V_1^{\frac{1}{\alpha_1}} V_2^{\frac{1}{\alpha_2}} \right) \leq \mathbf{E} \left(V_1^{\frac{1}{\alpha_1}} \right) \mathbf{E} \left(V_2^{\frac{1}{\alpha_2}} \right). \tag{A-11}$$

Therefore, combining (A-10) and (A-11) we deduce

$$\mathbf{E} \left(V_1^{-\frac{1}{\alpha_1}} V_2^{-\frac{1}{\alpha_2}} \right) \geq \left[\mathbf{E} \left(V_1^{\frac{1}{\alpha_1}} \right) \mathbf{E} \left(V_2^{\frac{1}{\alpha_2}} \right) \right]^{-1}. \tag{A-12}$$

For $m < k_i$, the $m-$th moment of V_i and V_i^{-1} is given by

$$\mathbf{E} \left(V_i^m \right) = \frac{\Gamma \left(k_i + m \right)}{\Gamma \left(k_i \right)} \mu_i^m, \ \mathbf{E} \left(V_i^{-m} \right) = \frac{\Gamma \left(k_i - m \right)}{\Gamma (k_i)} \mu_i^{-m}. \tag{A-13}$$

Hence, use of the formulas (A-6), (A-7), (A-12) and (A-13) for $m = \frac{1}{a_i}$ and $m = \frac{2}{a_i}$ and some algebra yields the thesis of the proposition. ■

Proof of Proposition 3. For $i = 1, 2$ the ratio within the brackets in the formula of $\rho(T_1, T_2|x)$, see (A-6), can be rewritten as $\frac{1}{\mathcal{F}(y_i(k), \alpha_i^{-1}) - 1}$ for $y_i(k) = k - \frac{2}{\alpha_i}$, where $\mathcal{F}(\varepsilon_1, \varepsilon_2) = \frac{\Gamma(\varepsilon_1)\Gamma(\varepsilon_1 + 2\varepsilon_2)}{\Gamma^2(\varepsilon_1 + \varepsilon_2)}$, $\varepsilon_1 > 0, \varepsilon_2 > 0$. We first show that $\mathcal{F}(\varepsilon_1, \varepsilon_2)$ is strictly decreasing in ε_1 for each $\varepsilon_2 \in (0, \infty)$, which in turn will imply that $\mathcal{F}(y_i(k), \alpha_i^{-1})$ is strictly decreasing in k for any positive α_i. Taking the logarithm of $\mathcal{F}(\varepsilon_1, \varepsilon_2)$ and then differentiating with respect to ε_1, we obtain

$$\frac{\vartheta \log \mathcal{F}(\varepsilon_1, \varepsilon_2)}{\vartheta \varepsilon_1} = \psi(\varepsilon_1) + \psi(\varepsilon_1 + 2\varepsilon_2) - 2\psi(\varepsilon_1 + \varepsilon_2), \tag{A-14}$$

Differentiating $\frac{\vartheta \log \mathcal{F}(\varepsilon_1, \varepsilon_2)}{\vartheta \varepsilon_1}$ with respect to ε_2 it follows

$$\frac{\vartheta}{\vartheta \varepsilon_2} \left[\frac{\vartheta \log \mathcal{F}(\varepsilon_1, \varepsilon_2)}{\vartheta \varepsilon_1} \right] = 2\psi^{(1)}(\varepsilon_1 + 2\varepsilon_2) - 2\psi^{(1)}(\varepsilon_1 + \varepsilon_2). \tag{A-15}$$

Clearly, $\psi^{(2)}(\varepsilon) \leq 0$ for $\varepsilon > 0$, which in turn implies $\frac{\vartheta}{\vartheta \varepsilon_2} \left[\frac{\vartheta \log \mathcal{F}(\varepsilon_1, \varepsilon_2)}{\vartheta \varepsilon_1} \right] \leq 0$ for all $(\varepsilon_1, \varepsilon_2) \in (0, \infty)^2$ by using (A-5). Hence, given that $\frac{\vartheta \log \mathcal{F}(\varepsilon_1, 0)}{\vartheta \varepsilon_1} = 0$ it follows $\frac{\vartheta \log \mathcal{F}(\varepsilon_1, \varepsilon_2)}{\vartheta \varepsilon_1} \leq 0$ for all $(\varepsilon_1, \varepsilon_2) \in (0, \infty)^2$. Therefore, given that $\rho(T_1, T_2|x)$ is strictly decreasing in $\mathcal{F}(y_i(k), \alpha_i^{-1})$ for $\rho_{12} > 0$, it follows that it is strictly decreasing in k for every positive α_i, and consequently, for $\rho_{12}=1$, $k \to \max\{\frac{2}{\alpha_1}, \frac{2}{\alpha_2}\} = \frac{2}{\alpha_2}$ and by continuity of $\Gamma(.)$, the bound is obtained. By Lemma 1, we have $\log \mathcal{F}(\varepsilon_1,\varepsilon_2) > 0$ for all $\varepsilon_1 > 0, \varepsilon_2 > 0$ and thus $\Gamma\left(\frac{2(\alpha_1-\alpha_2)}{\alpha_1\alpha_2}\right)\Gamma\left(\frac{2}{\alpha_2}\right) - \Gamma^2\left(\frac{2\alpha_1-\alpha_2}{\alpha_1\alpha_2}\right) > 0$ for all $\alpha_1, \alpha_2 > 0$,with $\alpha_1 > \alpha_2$. Using also the property $\delta(\alpha_1) > 1$ for each $\alpha_1 > 0$, the inequality $b_{gu}(\alpha_1,\alpha_2) < b_u(\alpha_1,\alpha_2)$ is shown.

■

Appendix B

In this appendix we provide a brief discussion about the Archimedean family, the FGM family of copulas, and their corresponding properties. The Archimedean family is constructed according to $C_\psi(\omega_1,\omega_2) = \xi_\psi^{[-1]}(\xi_\psi(\omega_1)+\xi_\psi(\omega_2))$ with $\xi_\psi : [0,1] \to [0,\infty), \xi'_\psi(\omega) < 0, \xi''_\psi(\omega) > 0$ for each $\omega \in (0,1)$ and $\xi_\psi(1) = 0$. The function $\xi_\psi^{[-1]}(\omega)$ is called pseudo-inverse and is equal to $\xi_\psi^{-1}(\omega)$ if $\omega < \xi_\psi(0)$ and 0 elsewhere. In case $\xi_\psi^{-1}(\omega) = \xi_\psi^{[-1]}(\omega)$ for every $\omega \in [0,\infty)$, both the copula and the respective generator are called strict. The case of $\xi_\psi(\omega) = -\ln\omega$ corresponds to independence between the underlying random variables. Nelsen (2006) describes this important class of copulas. We first describe the three most popular copulas which belong to the Archimedean family.

Clayton Copula: For $\xi_\psi(\omega) = \frac{1}{\psi}(\omega^{-\psi}-1)$ we obtain the Clayton copula which is given by

$$C_\psi(\omega_1,\omega_2) = \max\left\{\left(\omega_1^{-\psi}+\omega_2^{-\psi}-1\right),0\right\}^{-\frac{1}{\psi}}, \ \psi \in [-1,\infty)\backslash 0. \tag{B-1}$$

If $\psi \in [-1,0)$ the C_ψ is NQD and for every $\psi \in (0,\infty)$ the C_ψ is PQD. Additionally, $C_{-1}(\omega_1,\omega_2)$ $= \max\{\omega_1+\omega_2-1,0\}$, $\lim_{\psi\to\infty} C_\psi(\omega_1,\omega_2) = \min\{\omega_1,\omega_2\}$ and $\lim_{\psi\to 0} C_\psi(\omega_1,\omega_2) = \omega_1\omega_2$ for every $(\omega_1,\omega_2) \in [0,1]^2$. Note that $\lim_{\psi\to -1} C_\psi(\omega_1,\omega_2) \neq \max\{\omega_1+\omega_2-1,0\}$, which implies that C_ψ is not right-continuous at -1.

Frank Copula: If we apply $\xi_\psi(\omega) = -\ln\frac{e^{-\psi\omega}-1}{e^{-\psi}-1}$ as generator, we get the Frank copula

$$C_\psi(\omega_1,\omega_2) = -\frac{1}{\psi}\ln\left[1+\frac{(e^{-\psi\omega_1}-1)(e^{-\psi\omega_2}-1)}{e^{-\psi}-1}\right], \ \psi \in (-\infty,\infty)\backslash 0. \tag{B-2}$$

For any $\psi \in (-\infty,0)$ the C_ψ is NQD and for every $\psi \in (0,\infty)$ the C_ψ is PQD. Additionally, $\lim_{\psi\to-\infty} C_\psi(\omega_1,\omega_2) = \max\{\omega_1+\omega_2-1,0\}$, $\lim_{\psi\to\infty} C_\psi(\omega_1,\omega_2) = \min\{\omega_1,\omega_2\}$ and $\lim_{\psi\to 0} C_\psi(\omega_1,\omega_2)$

$= \omega_1\omega_2$ for all $(\omega_1, \omega_2) \in [0,1]^2$.

Gumbel Copula: For $\xi_\psi(\omega) = (-\ln \omega)^\psi$ we get the Gumbel copula which is expressed as

$$C_\psi(\omega_1, \omega_2) = \exp\left[-\left((-\ln \omega_1)^\psi + (-\ln \omega_2)^\psi\right)^{\frac{1}{\psi}}\right],\ \psi \in [1, \infty). \tag{B-3}$$

The $C_\psi(\omega_1, \omega_2)$ is PQD for any $\psi \in (1, \infty)$. Moreover, $\lim_{\psi \to \infty} C_\psi(\omega_1, \omega_2) = \min\{\omega_1, \omega_2\}$, $C_1(\omega_1, \omega_2) = \omega_1\omega_2$ for any $(\omega_1, \omega_2) \in [0,1]^2$.

Finally, another copula that we could employ for parameterizing G is the Farlie-Gumbel-Morgenstern (FGM) copula.

Farlie-Gumbel-Morgenstern Copula: This family of distributions is expressed as

$$C_\psi(\omega_1, \omega_2) = \omega_1\omega_2 + \psi\omega_1\omega_2(1 - \omega_1)(1 - \omega_2),\ \psi \in [-1, 1]. \tag{B-4}$$

If $\psi \in [-1, 0)$ the C_ψ is NQD, if $\psi \in (0, 1]$ the C_ψ is PQD, and $C_0(\omega_1, \omega_2) = \omega_1\omega_2$ for any $(\omega_1, \omega_2) \in [0,1]^2$.

References

Abbring, J. H. and Van den Berg, G. J. (2007), "The unobserved heterogeneity distribution in duration analysis," *Biometrika*, 94, 87–99.

Beaudoin, D., Duchesne, T., and Genest, C. (2007), "Improving the estimation of Kendall's tau when censoring affects only one of the variables," *Computational Statistics & Data Analysis*, 51, 5743–5764.

Børing, P. (2009), "Gamma Unobserved Heterogeneity and Duration Bias," *Econometric Reviews*, 29, 1–19.

Cherian, F. (1941), "A bivariate gamma type distribution function," *Journal of the Indian Mathematical Society*, 5, 133–144.

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

Cui, S. and Sun, Y. (2004), "Checking for the gamma frailty distribution under the marginal proportional hazards frailty model," *Statistica Sinica*, 14, 249–267.

Embrechts, P., McNeil, A., and Straumann, D. (2002), "Correlation and dependence in risk management: properties and pitfalls," *Risk management: value at risk and beyond*, 176–223.

Henderson, R. and Shimakura, S. (2003), "A serially correlated gamma frailty model for longitudinal count data," *Biometrika*, 90, 355–366.

Joe, H. (1997), *Multivariate models and dependence concepts*, vol. 73, Chapman & Hall/CRC.

Jonker, M., Bhulai, S., Boomsma, D., Ligthart, R., Posthuma, D., and Van Der Vaart, A. (2009), "Gamma frailty model for linkage analysis with application to interval-censored migraine data," *Biostatistics*, 10, 187–200.

Korsgaard, I. R. and Andersen, A. H. (1998), "The additive genetic gamma frailty model," *Scandinavian Journal of Statistics*, 25, 225–269.

Kotz, S., Johnson, N., and Balakrishnan, N. (2000), *Continuous multivariate distributions: models and applications*, vol. 1, Wiley-Interscience.

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

Mardia, K. (1970), *Families of bivariate distributions*, Griffin London.

Martin, E. and Betensky, R. (2005), "Testing quasi-independence of failure and truncation times via conditional Kendall's tau," *Journal of the American Statistical Association*, 100, 484–492.

Nelsen, R. (2006), *An introduction to copulas*, Springer Verlag.

Oakes, D. (2008), "On consistency of Kendall's tau under censoring," *Biometrika*, 95, 997–1001.

Shih, J. H. and Louis, T. A. (1995), "Inferences on the association parameter in copula models for bivariate survival data," *Biometrics*, 1384–1399.

Sklar, A. (1959), "Fonctions de répartition à n dimensions et leurs marges." *Publications de l'Institut de Statistique de L'Université de Paris*, 8, 229–231.

Tchen, A. (1980), "Inequalities for distributions with given marginals," *The Annals of Probability*, 8, 814–827.

Van den Berg, G. J. (1997), "Association measures for durations in bivariate hazard rate models," *Journal of Econometrics*, 79, 221–245.

Van den Berg, G. J. (2001), "Duration models: specification, identification and multiple durations," *Handbook of econometrics*, 5, 3381–3460.

Wang, W., Wells, M. T., et al. (2000), "Estimation of Kendall's tau under censoring," *Statistica Sinica*, 10, 1199–1216.

Yashin, A., Vaupel, J., and Iachine, I. (1995), "Correlated individual frailty: an advantageous approach to survival analysis of bivariate data," *Mathematical Population Studies*, 5, 145–159.

Zhong, X. and Li, H. (2002), “An additive genetic gamma frailty model for two-locus linkage analysis using sibship age of onset data,” *Statistical Applications in Genetics and Molecular Biology*, 1, 1–24.