A Copula Duration Model with Dependent States and Spells

Lo, Simon M. S. ; Shi, Shuolin ; Wilke, Ralf

Document Version
Final published version

Published in:
Computational Statistics & Data Analysis

DOI:
10.1016/j.csda.2024.108104

Publication date:
2025

Computational Statistics and Data Analysis 204 (2025) 108104

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

A copula duration model with dependent states and spells

Simon M.S. Lo [a], Shuolin Shi [b], Ralf A. Wilke [b,*]

[a] United Arab Emirates University, Department of Innovation in Government and Society, United Arab Emirates
[b] Copenhagen Business School, Department of Economics, Denmark

ARTICLE INFO

Keywords:
Hierarchical copula
Frailty
Pseudo maximum likelihood estimation

ABSTRACT

A nested Archimedean copula model for dependent states and spells is introduced and the link to a classical survival model with frailties is established. The model relaxes an important restriction of classical survival models as the distributions of unobservable heterogeneities are permitted to depend on the observable covariates. Its modular structure has practical advantages as the different components can be separately specified and estimation can be done sequentially or separately. This makes the model versatile and adaptable in empirical work. An application to labour market transitions with linked administrative data supports the need for a flexible specification of the dependence structure and the model for the marginal survivals. The conventional Markov Chain Model is shown to give sizeably biased results in the application.

1. Introduction

The increased availability of longitudinal data generated by operations in firms and public administration enables the application of complex models for repeated durations and failure times in various disciplines such as biostatistics, engineering and social sciences. Although the content of this paper is relevant clearly beyond economics, we use it in this paper to pin down the model. For instance, there are repeated periods of unemployment, employment or illness. At the same time these correspond to different states and individuals are jumping between them. Multiple states and spells models are therefore regularly used in applied economic research in different setups such as duration models (e.g. Eberwein et al., 1997; Carrasco and Garcia-Perez, 2015) and Markov Chain Models (Magnac and Robin, 1994; Fougère and Kamionka, 2008). An important aspect of these models is that there are unobserved random factors, so-called frailties or unobserved heterogeneities, which induce correlations between states and spells. Copulas are a powerful tool for modelling dependencies (Trivedi and Zimmer, 2005; Nelsen, 2006; Durante and Sempi, 2016) and can be conveniently used to characterise dependencies in unobservables. In the context of duration or survival analysis, copulas have been introduced by Zheng and Klein (1995) and Carrière (1994, 1995) for modelling dependencies in the competing risks model. This paper presents the first multiple states and spells copula duration model that allows for dependencies between states and spells with the help of a nested or hierarchical Archimedean copula structure (Hofert and Mächle, 2011; Hofert, 2012). The use of copulas in competing risks duration models is becoming increasingly popular to model risk dependencies (e.g. Zheng and Klein, 1995; Rivest and Wells, 2001; Braekers and Veraverbeke, 2005; Sujica and Van Keilegom, 2018; Emura et al., 2020) or dependencies between multiple spells (Lo et al., 2020; Lipowski et al., 2021; Wang and Emura, 2021). This paper introduces a generalisation of these models. The suggested copula structure is motivated by dependencies between various frailty terms. It has some practical advantages compared to the classical approaches as the dependence structure and the marginal distributions are separately modelled. The model is therefore compatible

* Corresponding author.
E-mail addresses: losimonms@yahoo.com.hk (S.M.S. Lo), shuolinnemo@gmail.com (S. Shi), rw.eco@cbs.dk (R.A. Wilke).

https://doi.org/10.1016/j.csda.2024.108104
Received 26 April 2024; Received in revised form 19 November 2024; Accepted 21 November 2024

Available online 28 November 2024

with various copula and distribution models, which can be easily replaced in an application. Copula modelling is indeed not different from commonly used competing risks duration models such as the multivariate mixed proportional hazards model (MMPHM). In fact the latter has been shown to be a special case of the copula duration model (Lo et al., 2017). While Lo et al. (2017), Emura et al. (2019) and Wang et al. (2020) consider the link between frailties and the copula approaches for the single spell competing risks model, we establish it for a nested copula structure that allows for multiple spells and states. Another generalisation of our model is that the dependence structure is allowed to depend on observable covariates. Intuitively, this corresponds to the frailty distributions to depend on observable covariates, something that conventional duration models do not allow for. This feature is of practical importance as typical data structures are characterised by sizable correlations between variables. We demonstrate that misspecification of the dependence structure leads to inconsistent estimates. To directly relate our model to an existing model, we carefully work out that it is considerably less restrictive than the Markov Chain Model. We demonstrate nice finite sample properties of our suggested model with the help of Monte Carlo simulations. It is confirmed that a careful modelling of dependence structures gives more insightful results and avoids inconsistencies in particular when observables are correlated with frailties. To make our statistical model more easily accessible for practitioners, we provide a running example from labour economics. To demonstrate the applicability, we present an application to labour market data where we study the transition paths of older people from economic activity into inactivity around retirement age. We find that the dependence structure depends on observable covariates and the Markov Chain Model in comparison gives biased results, because its additional restrictions are not supported by the data. We make sample code available to ease the adoption of our model among practitioners.

The paper is structured as follows. Section 2 introduces the model and presents the theoretical results and the estimation approach. Section 3 contains the application to labour market data.

2. The model

We consider a longitudinal model on individual level for the time spent (duration) in states $r \in \{1,2\} = \mathcal{R}$. Individuals can switch back and forth between the two states and thus can re-enter the same state r multiple times denoted as $k = 1,2,...$ (repeated occurrences, multiple spells). The duration or transition time for the k'th spell in state r is T_{rk}. Suppose the initial state is $r = 1$ followed by sequential transitions between states $r = 1$ and $r = 2$. The corresponding sequence of duration variables is $T_{11}, T_{21}, T_{12}, T_{22}, T_{13}, T_{23}, T_{14}, ..., T_{rk_r}$ with k_r is the maximum number of repeated occurrence in state $r = 1,2$. The set $\mathcal{K}_r = \{1,2,...,k_r\}$ contains the repeated occurrences for state r. $\boldsymbol{X}_{rk}$ is a vector of observable covariates for the k'th spell in state r, which are recorded at the start of the spell. For simplicity of exposition, we present the model for $k_1 = k_2 = 2$, where any combination of $k_1 = 1,2$ and $k_2 = 1,2$ are nested in the presented model. The model can be easily generalised to $k_r > 2$, although its estimation becomes increasingly complicated as a range of higher order derivatives of the copulas are required. See Lo et al. (2020) in the context of multiple spells. Various other extensions to the model are introduced in Supplement S1. These include, lagged duration dependence, independent censoring, a discussion of an extension to covariates that vary within a spell and an extension to more than two states.

We illustrate the model with the help of a running example throughout this section to ease the reading. This example is also the relevant context for our empirical analysis in Section 3.

Running example. Low birth rates and an increased longevity in the developed countries lead to increasing pressure on the pension and social security systems. It is therefore of great interest to understand better the pathways from economic activity into inactivity of the older workforce coming closer to retirement age. In this context the two states are economically active (various forms of employment, short-term illness, unemployment, etc.) (state 1) and economically inactive (various forms of out of labour force, including long-term illness, retirement, etc.) (state 2). Let an individual be employed (state 1) initially, before becoming ill for a longer period. This is followed by a transition back to work until old-age retirement. In this example there are two spells of economic activity and one spell of inactivity, i.e. $k_1 = 2$, $k_2 = 1$. The recorded sequence of the duration variables is the employment duration (state 1) until becoming long-term ill T^*_{11}. This is followed by the duration of long-term illness (state 2) until reentering employment T^*_{21}. The final one is employment duration (state 1) until retirement (state 2) T^*_{12}. The covariates, such as educational background and wealth, are measured at the start of each spell ($\boldsymbol{X}_{11}, \boldsymbol{X}_{21}, \boldsymbol{X}_{12}$) and may consist of different variables across spells and states.

Let $\boldsymbol{T}_r$ be a stacked vector of all T_{rk} for any k given r, i.e. all multiple spells for risk r. Let $\boldsymbol{T}$ be a stacked vector of the T_{rk} for all r and k. The stacked vectors of covariates $\boldsymbol{X}$, $\boldsymbol{X}_r$, and $\boldsymbol{X}_k$ are defined analogously. The marginal survival function for each duration T_{rk} conditional on $\boldsymbol{X}_{rk}$ is

$$S_{rk}(t;\boldsymbol{x}_{rk}) = \Pr(T_{rk} > t|\boldsymbol{X}_{rk} = \boldsymbol{x}_{rk}) = \exp\left[-\int_0^t \lambda_{rk}(s;\boldsymbol{x}_{rk})ds\right] \quad (1)$$

for $r \in \mathcal{R}$, $k \in \mathcal{K}_r$, where $\lambda_{rk}(t;\boldsymbol{x}_{rk}) = \lim_{\epsilon\to 0} \Pr(t \le T_{rk} < t+\epsilon|T_{rk} \ge t; \boldsymbol{X}_{rk} = \boldsymbol{x}_{rk})/\epsilon$ is the marginal hazard. The joint survival function for all durations $\boldsymbol{T}$ conditional on $\boldsymbol{X}$ is

$$H(\boldsymbol{t};\boldsymbol{x}) = \Pr(\cap_{\{r\in\mathcal{R},k\in\mathcal{K}_r\}} T_{rk} > t_{rk}|\boldsymbol{X} = \boldsymbol{x}). \quad (2)$$

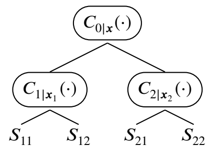


Fig. 1. The nested copula structure with two states ($r = 1, 2$) and two repeated occurrences ($k = 1, 2$) has three copulas C_l for $l = 0, 1, 2$.

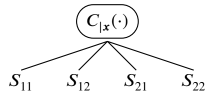


Fig. 2. One-layer copula structure with two states ($r = 1, 2$) and two repeated occurrences ($k = 1, 2$).

The joint survival function for different spells having the same original state r, i.e. $\boldsymbol{T}_r$, conditional on $\boldsymbol{X}_r$ is

$$H_r(\boldsymbol{t}_r; \boldsymbol{x}_r) = \Pr(\cap_{\{k \in \mathcal{K}_r\}} T_{rk} > t_{rk} | \boldsymbol{X}_r = \boldsymbol{x}_r), \text{ for } r \in \mathcal{R}. \tag{3}$$

H_r is therefore allowed to differ with $r = 1, 2$. In our application, the dependence between consecutive spells in economic activity states (state 1) can be characterised by the individual's attitude and willingness to work. The dependence between consecutive spells in economic inactivity (state 2), in contrast, is likely due to unobservable health or wealth related factors. While it is difficult to anticipate the factors behind these dependencies, there is no prior reason why H_1 and H_2 should be the same.

For any given joint distribution H_r and any given marginal distribution S_{rk}, there is a unique copula C_r that establishes the link between them (Sklar, 1959). For spell dependence for state r, the copula is $C_{r|\boldsymbol{x}_r}$, such that

$$H_r(\boldsymbol{t}_r; \boldsymbol{x}_r) = C_{r|\boldsymbol{x}_r}[S_{r1}(t_{r1}; \boldsymbol{x}_{r1}), S_{r2}(t_{r2}; \boldsymbol{x}_{r2})], \text{ for } r \in \mathcal{R}. \tag{4}$$

The advantage of using a copula to model the joint distribution is that $C_{r|\boldsymbol{x}_r}$ and S_{rk} are separate models that each depend on r and $\boldsymbol{X}_r$. This allows for convenient replacement of the models for the marginals and dependence structure.

H can be also written as a copula model. For this, we adopt a nested copula structure as considered by Hofert (2012) in the case of linking distribution functions. We phrase it here in terms of survival functions, where the mother copula $C_{0|\boldsymbol{x}}$ links H_1 and H_2 as given in (4):

$$\begin{aligned} H(\boldsymbol{t}; \boldsymbol{x}) &= C_{0|\boldsymbol{x}}[H_1(\boldsymbol{t}_1; \boldsymbol{x}_1), H_2(\boldsymbol{t}_2; \boldsymbol{x}_2)] \\ &= C_{0|\boldsymbol{x}}[C_{1|\boldsymbol{x}_1}(\cdot), C_{2|\boldsymbol{x}_2}(\cdot)]. \end{aligned} \tag{5}$$

The nested structure is illustrated in Fig. 1. The upper layer represents the dependence between the states (mother copula), while the lower layer describes the dependence between the multiple spells within a state (daughter copula) for each state. An alternative way to craft the dependence structure by means of reverting the two layers is presented in Supplement S1. C_l is characterised by a known functional form with unknown parameters θ_l for $l = 0, 1, 2$, which determine the strength or the degree of dependence. The dependence between the two marginal survivals that are linked by $C_{r|\boldsymbol{x}_r}$ for $r = 1, 2$ is therefore determined by θ_r. For instance, θ_1 determines the dependence between S_{11} and S_{12}. The dependence between the marginal survivals under two different $C_{r|\boldsymbol{x}_r}$, e.g. S_{11} and S_{21}, is determined by $C_{0|\boldsymbol{x}}$ and in particular by its parameter(s) θ_0. The dimension of θ_l is the number of parameters of C_l. For most popular copulas it is one and this is what we consider in the following. Depending on the chosen copulas, the ranges of the parameters are to be constrained. While from a theoretical point the copulas in (5) exist, some restrictions on them are required for practicality. In the following the model is restricted to the Archimedean family, one of the most popular classes of copulas. Archimedean copulas are popular because of their symmetry properties. This property facilitates convenient closed form solutions and therefore simplifies numerical analysis. For example in the special case when $C = C_0 = C_1 = C_2$ is an Archimedean copula, model (5) corresponds to a one layer Archimedean model

$$H(\boldsymbol{t}; \boldsymbol{x}) = C_{|\boldsymbol{x}}[S_{11}(t_{11}; \boldsymbol{x}_{11}), S_{12}(t_{12}; \boldsymbol{x}_{12}), S_{21}(t_{21}; \boldsymbol{x}_{21}), S_{22}(t_{22}; \boldsymbol{x}_{22})], \tag{6}$$

because of risk pooling (Lo and Wilke, 2010). See Fig. 2 for an illustration. Archimedean copulas are characterised by

$$C_{l|\boldsymbol{x}}(u_1, \cdots, u_m; \boldsymbol{x}) = \phi^{-1}_{\theta_l(\boldsymbol{x})}[\phi_{\theta_l(\boldsymbol{x})}(u_1(\boldsymbol{x})) + \cdots + \phi_{\theta_l(\boldsymbol{x})}(u_m(\boldsymbol{x}))] \tag{7}$$

for $l \in \{0, 1, 2\}$ and $\phi_{\theta_l}(u) \in [0, \infty)$ is the generator function for any $u \in (0, 1]$, which is strictly decreasing and completely monotonic in u. θ_l is the dependence parameter of copula C_l for $l \in \{0, 1, 2\}$. A higher absolute value of θ_l means a stronger dependence and $\theta_l = 0$ corresponds to independence. For reasons outlined below we restrict the model to $\theta_l \geq 0$ and therefore to positive dependence.

By allowing θ_l to be a function of $\boldsymbol{x}$, makes the frailty distribution dependent on $\boldsymbol{x}$. In the context of duration models there is a direct link between the Archimedean copula and frailty. Lo et al. (2017) and Wang et al. (2020) consider this for a competing risks model with dependencies between multiple competing risks but there are no multiple spells nor states in their models. By adopting Puzanova's (2011) mixture representation of the hierarchical Archimedean copula model, we show in the following that the hierarchical Archimedean copula structure in Model (5) is equivalent to a survival model with dependent frailties and that the mixing variables in our model correspond to these frailties.

Suppose there are three frailties V_0, V_1 and V_2. V_l has marginal distribution function $F_l(v_l)$ for $l = 0, 1, 2$. V_1 is the frailty common to the two spells under the same copula C_1 in the lower layer of Fig. 1, and it determines the dependency between S_{11} and S_{12}. Similarly, V_2 determines the dependence between S_{21} and S_{22}. V_0 is the frailty common to all spells in the second layer, and it governs the dependence amongst all these four spells under C_0. Since V_0 in the upper layer plays a role in the dependency in the lower layer, both distributions of V_1 and V_2 depend on V_0. Moreover, dependency across the two lower layers are driven merely by V_0, such that V_1 and V_2 are independent when V_0 is fixed. We denote the conditional distribution of V_r ($r = 1, 2$) given V_0 as $F_{r|0}(v_{r|0})$, where $V_{r|0}$ is the frailty V_r when V_0 is controlled. Let the joint distribution of V_0, V_1, and V_2 be $F(v_0, v_1, v_2)$ and $S^*_{rk}(t; v_r, \boldsymbol{x}) = Pr(T_{rk} \geq t | V_r = v_r, \boldsymbol{X}_{rk} = \boldsymbol{x}_{rk})$. Note that $S^*_{rk}(t; v_r, \boldsymbol{x})$ is different from $S_{rk}(t; \boldsymbol{x})$, as the former is the marginal survival conditional on the frailty (and $\boldsymbol{x}$), whereas the latter is the unconditional marginal survival that results from integrating $S^*_{rk}(t; v_r, \boldsymbol{x})$ with respect to the frailty distribution F_r. Goethals et al. (2008) show for the case $S^*_{rk}(t; v_r, \boldsymbol{x})$ is Weibull and F_r is a Gamma distribution that $S_{rk}(t; \boldsymbol{x})$ is not Weibull.

To establish the mathematical link between our copula model and the frailty model, we require additional restrictions on the latter. As a matter of fact these are the usual restrictions on the frailties in commonly used multiplicative frailty models such as the MMPHM (see Joe, 1997). We therefore directly connect our model to these popular models, but our copula model has the benefit of a modular structure which allows for convenient replacement of the models for S_{rk} for $r = 1, 2$, $C_{l|\boldsymbol{x}}$ for $l = 0, 1, 2$, and the dependence of C on $\boldsymbol{x}$.

Assumption 1. (i) [multiplicative frailty]

$$S^*_{rk}(t; v_r, \boldsymbol{x}_{rk}) = \exp[-v_r \Lambda^*_{rk}(t; \boldsymbol{x}_{rk})] \tag{8}$$

with $\Lambda^*_{rk}(t; \boldsymbol{x}_{rk})$ is the integrated hazard for this model and $r, k = 1, 2$.
(ii) [conditional independence I] Conditional on v_0, v_1 and v_2 are independent:

$$F(v_1, v_2 | v_0) = F_{1|0}(v_{1|0}) F_{2|0}(v_{2|0}), \tag{9}$$

while V_1 and V_2 depend on V_0.
(iii) [conditional independence II] T_{rk} for $r = 1, 2$ and $k = 1, 2$ are independent conditional on $\boldsymbol{X}$, V_0, V_1 and V_2 with

$$H(\boldsymbol{t}; \boldsymbol{x}) = \int_{v_0} \int_{v_1} \int_{v_2} \prod_{r=1}^{2} \prod_{k=1}^{2} S^*_{rk}(t_{rk}; v_r, \boldsymbol{x}_{rk}) dF(v_0, v_1, v_2; \boldsymbol{x}). \tag{10}$$

(iv) [Laplace transform I] $\phi^{-1}_{\theta_l}$ ($l \in \{0, 1, 2\}$) is the Laplace transform with respect to the frailty distribution $F_l(v_l; \boldsymbol{x}_l)$, such that for any $s > 0$

$$\phi^{-1}_{\theta_l}[s(\boldsymbol{x}_l)] = \int_{v_l} \exp(-v_l s) dF_l(v_l; \boldsymbol{x}_l). \tag{11}$$

(v) [Laplace transform II] $\phi^{-1}_{\theta_r|0}$ ($r \in \{1, 2\}$) is the Laplace transform with respect to $F_{r|0}(v_{r|0})$, and for any $s > 0$,

$$\phi^{-1}_{\theta_r|0}[s(\boldsymbol{x}_r)] = \int_{v_{r|0}} \exp(-v_{r|0} s) dF_{r|0}(v_{r|0}; \boldsymbol{x}_r) = \xi_r(s; \boldsymbol{x}_r)^{v_0} \tag{12}$$

for some function ξ_r.

Proposition 1. *The restrictions of Assumption 1 imply the hierarchical Archimedean copula structure for the joint survival function of T_{rk} for $r = 1, 2$ and $k = 1, 2$ given $\boldsymbol{x}$ given in (5):*

$$H(\boldsymbol{t}; \boldsymbol{x}) = C_{0|\boldsymbol{x}}\{C_{1|\boldsymbol{x}}[S_{11}(t_{11}; \boldsymbol{x}_{11}), S_{12}(t_{12}; \boldsymbol{x}_{12})], C_{2|\boldsymbol{x}}[S_{21}(t_{21}; \boldsymbol{x}_{21}), S_{22}(t_{22}; \boldsymbol{x}_{22})]\}. \tag{13}$$

The proof is given in Appendix A.1 Proposition 1 establishes the formal link between the copula and a dependent frailty model, which are frequently used in applied economic research (see e.g. Van den Berg, 2001). Modelling the dependencies with the help of copulas is practical as the copula and the marginal survival model are separate components which are compatible with a range models. In contrast to classical approaches such as the MMPHM, the dependence structure is allowed to depend on covariates. Although we use frailties to motivate the dependence structure in Model (5), the dependencies could also arise from other things such as simultaneity in outcomes. In the following we continue to use frailty as the motivating example to link our model better to the popular existing approaches. It is for this reason that we also maintain Assumption 1.

Frailty plays a role in the model through two channels. First, it acts as a multiplier on the hazard function, see (8). A higher value of frailty therefore implies a shorter duration. Second, the dependence between spells can be related to the frailty variance. Take the

Clayton copula as an example for which the frailty follows a Gamma distribution with parameters $\Gamma(1/\theta, \theta)$ (compare for example Lo et al., 2017). The mean of the frailty is one while its variance is θ. A greater variance therefore implies a greater spell dependence.

Running example *(continued)*. Suppose V_r is unobserved financial needs to ensure the desired amount of spending and r is inactivity. An individual with higher level of financial needs is expected to have a shorter inactivity spell. An individual with lower level of financial needs is expected to have a longer inactivity spell. The dependence of any two inactivity spells is also related to frailty. For higher levels of financial needs, both inactivity spells are expected to be shorter. With lower financial needs, both inactivity spells are expected to be longer. A larger variance of financial needs in the population translates into a stronger dependence between two inactivity spells. Low variation in frailty, in contrast, reduces the observed positive dependence. In the case there is no variation at all in frailty, i.e. a constant, there is no observed dependence.

Our example has illustrated that the interpretation of the dependence structure or its parameters θ_r is therefore directly linked to the distribution of unobservables or frailties v_r in $S^*_{rk}(t; v_r, \boldsymbol{x})$. The dependence is driven by the dispersion of the frailty. Copula dependence and the dispersion of the frailty are merely two different presentations of the same model.

There are the following restrictions on the dependence structure in model (13) to make it well defined. The nested Archimedean copula structure requires a regularity condition regarding the copula generators in the upper layer ϕ_{θ_0} and the lower layer ϕ_{θ_r} for $r = 1, 2$ such that $\phi_{\theta_0} \circ \phi^{-1}_{\theta_r}$ is an infinite differentiable monotone increasing function. Commonly used one-parameter Archimedean copula generators such as for Clayton, Frank, and Gumbel satisfy this technical property. There are also range restrictions on θ_l for $l = 0, 1, 2$ that can be more restrictive than the usual copula specific ranges. Firstly, the nested Archimedean copula structure implies that the dependence in the upper layer must be smaller than in the lower layer (Joe, 1997), i.e. $\theta_0 < \min\{\theta_1, \theta_2\}$. This can be conveniently added as a constraint to the empirical model. Alternatively, the copula layers can be reverted as outlined in Supplement S1. The nested structure also rules out that one can mix positive and negative dependence in the two layers. Secondly, we only consider positive dependence in this paper, because a mixture model by means of frailty terms permits only positive dependence. By doing so, we also obtain much clearer results for the comparative statics in Subsection 2.1. In consequence, the range of θ_l needs to be restricted to positive values by means of restrictions in the empirical model for those Archimedean copulas that technically allow for negative dependence.

V_l for $l = 0, 1, 2$ can be interpreted as omitted variables and for this reason they should be present in practically all empirical models. However, commonly used frailty models in duration analysis require the distributions of V_l to be independent of observable covariates (Wooldridge, 2010).

Running example *(continued)*. Suppose the observable covariate is wealth. Richer individuals normally have a higher variance of financial needs due to more heterogeneous life circumstances. Therefore the variance of frailty measured by θ is expected to be greater for wealthier individuals than for poorer. In other words, the distribution of frailty depends on covariates.

To accommodate the covariate dependence of the dependence structure, the parametric relationship

$$\theta_l(\boldsymbol{x}) = \exp(\alpha_{0l} + \boldsymbol{x}'\boldsymbol{\alpha}_l), \text{ for } l \in \{0, 1, 2\} \tag{14}$$

is used. $\theta_l(\boldsymbol{x}) > 0$ for all $\boldsymbol{x}$ and all l. The restriction $\theta_0(\boldsymbol{x}) < \min\{\theta_1(\boldsymbol{x}), \theta_2(\boldsymbol{x})\}$ requires $\alpha_{00} + \boldsymbol{x}'\boldsymbol{\alpha}_0 < \min\{\alpha_{01} + \boldsymbol{x}'\boldsymbol{\alpha}_1, \alpha_{02} + \boldsymbol{x}'\boldsymbol{\alpha}_2\}$ for all $\boldsymbol{x}$. As this is not guaranteed for all $\boldsymbol{x}$, we introduce a slightly modified functional form in Supplement S3 if required for the estimation.

It is more intuitive to consider a measure for the degree of dependence instead $\theta(\boldsymbol{x})$. Various measures of concordance, correlation and dependence including tail dependencies have been studied in the literature. We focus here on the commonly used Kendall's tau $\tau(\boldsymbol{x}) \in (0, 1)$ for any copula $C_{\boldsymbol{x}}$ (Joe, 1997). $\tau(\boldsymbol{x})$ is a function of $\theta(\boldsymbol{x})$ (Nelsen, 2006). In the case of independence, $\tau(\boldsymbol{x})$ is zero in expectation conditional on $\boldsymbol{x}$, $\tau(\boldsymbol{x}) = 1$ corresponds to perfect dependence.

Lastly, we study a scenario of misspecification of $\tau(\boldsymbol{x})$. In particular, we consider what happens when $\tau(\boldsymbol{x}) = \tau$ is incorrectly assumed. This is the default assumption of standard frailty models. It corresponds to specifying the unconditional copula $E_{\boldsymbol{x}}[C_{\boldsymbol{x}}] = \int_{\boldsymbol{x}} C_{\boldsymbol{x}} dF(\boldsymbol{x})$, where Kendall's τ is $\tau = 4E[\int_{\boldsymbol{x}} C_{\boldsymbol{x}} dF(\boldsymbol{x})] - 1$ in this case. The following Proposition establishes that the misspecified τ equals $E_{\boldsymbol{x}}[\tau(\boldsymbol{x})] = \tilde{\tau}$ for some standard copulas.

Proposition 2. *Let $\tau(\boldsymbol{x})$ be falsely assumed to be independent of $\boldsymbol{x}$. $\tau = \tilde{\tau}$ if the following relationship holds:*

$$\int_{[0,1]^2} \left(\int_{\boldsymbol{x}} C_{\boldsymbol{x}} dF(\boldsymbol{x}) \right) d \left(\int_{\boldsymbol{x}} C_{\boldsymbol{x}} dF(\boldsymbol{x}) \right) = \int_{\boldsymbol{x}} \int_{[0,1]^2} C_{\boldsymbol{x}} dC_{\boldsymbol{x}} dF(\boldsymbol{x}). \tag{15}$$

The proof is given in Appendix A.1. As condition (15) is technical, we check it with help of simulations for several commonly used copulas. We find that it holds for the Clayton and Frank copulas. For Gumbel, Joe, and Ali-Mikhail-Haq (AMH) copula, we find $\tau \approx \tilde{\tau}$. This suggests that falsely assuming $\tau(\boldsymbol{x}) = \tau$ corresponds approximately to modelling $E_{\boldsymbol{x}}[\tau(\boldsymbol{x})]$. If this suffices for the application, the model can be simplified.

2.1. Interpretation and comparative statics

In this subsection, we provide additional results that are useful for interpretation and comparative statics. For example, we elaborate how the modelled dependencies contribute to extracting economically meaningful results.

Running example *(continued).* The model in (5) allows for dependencies between the previous and current spell due to unobserved factors that are modelled by the copula structure. For example the length of the previous economic inactivity spell can play a role for the length of the subsequent activity period. Similarly, the length of the second economic activity spell can depend on the length of the first activity spell. The source of these dependencies are unobserved factors that play a role for the outcome in the various states and spells.

To translate this dependence into a meaningful concept we focus on the conditional survival and conditional hazard of terminating a spell given the length of a past spell. For illustration, we focus on the case that the previous spell exceeded a certain length. Here we distinguish between two spells in the same state and two spells in different states. Any two spells in state r depend on each other through the copula in the second layer C_r in Fig. 1. Therefore, the survival of T_{rk}, conditional on $T_{rl} \geq t_{rl}$, or equivalently, $S_{rl} < s_{rl}$ where $s_{rl} = S_{rl}(t_{rl})$, for any $k > l \in \mathcal{K}_r$ and for any $r \in \mathcal{R}$ is

$$\begin{aligned} S_{rk}(t_{rk}|S_{rl} < s_{rl}; \boldsymbol{x}_r) &= \Pr(T_{rk} \geq t_{rk}|T_{rl} \geq t_{rl}; \boldsymbol{x}_r) \\ &= \Pr(S_{rk} < s_{rk}|S_{rl} < s_{rl}; \boldsymbol{x}_r) \\ &= \frac{C_{r|\boldsymbol{x}_r}[S_{rk}(t_{rk}; \boldsymbol{x}_{rk}), S_{rl}(t_{rl}; \boldsymbol{x}_{rl})]}{S_{rl}(t_{rl}; \boldsymbol{x}_{rl})}. \end{aligned} \quad (16)$$

The conditional hazard for the k'th spell in state r conditional on the l'th spell in state r is

$$\begin{aligned} \lambda_{rk}(t_{rk}|S_{rl} < s_{rl}; \boldsymbol{x}_r) &= \lim_{\epsilon \to 0} \Pr(S_{rk} \in [s_{rk}, s_{rk} + \epsilon]|S_{rk} < s_{rk}; S_{rl} < s_{rl}; \boldsymbol{x}_r)/\epsilon \\ &= \frac{1}{C_{r|\boldsymbol{x}_r}} \frac{\partial C_{r|\boldsymbol{x}_r}}{\partial s_{rk}}. \end{aligned} \quad (17)$$

For any two spells in different states, the dependence is triggered by the copula in the first layer C_0 in Fig. 1. Hence, the conditional hazard of the k'th spell in state r conditional on the l'th spell in state $s \neq r$ with $k < l$ is

$$\lambda_{rk}(t_{rk}|S_{sl} < s_{sl}; \boldsymbol{x}) = \frac{1}{C_{0|\boldsymbol{x}}} \frac{\partial C_{0|\boldsymbol{x}}}{\partial s_{rk}}. \quad (18)$$

The following result is only shown for the Clayton copula with $\phi_{\theta_r(\boldsymbol{x})}(u) = (u^{-\theta_r(\boldsymbol{x})} - 1)/\theta_r(\boldsymbol{x})$ for $\theta_r(\boldsymbol{x}_r) > 0$ because of its analytical convenience. Although similar results can be likely derived for other copulas. We use the independence copula for $\theta_r(\boldsymbol{x}_r) = 0$.

Proposition 3. *Suppose C_r is Clayton with $\theta_r(\boldsymbol{x}_r) \geq 0$ for $r \in \{1, 2\}$. An increase in the degree of dependence leads often to a decrease in the conditional hazard. In particular,*

$$\frac{\partial \lambda_{rk}(t_{rk}|S_{rl} < s_{rl}; \boldsymbol{x})}{\partial \theta_r(\boldsymbol{x})} < 0$$

for all $r \in \{1, 2\}$ and $k > l$, whenever $s_{rl}^{1/(1-s_{rl}^{\theta_r(\boldsymbol{x})})} < s_{rk}$, and

$$\frac{\partial \lambda_{rk}(t_{rk}|S_{sl} < s_{sl}; \boldsymbol{x})}{\partial \theta_0(\boldsymbol{x})} < 0$$

for $r \in \{1, 2\}$ and $s \neq r$, whenever $s_{sl}^{1/(1-s_{sl}^{\theta_0(\boldsymbol{x})})} < s_{rk}$.

We prove Proposition 3 in Appendix A.1 and present several scenarios for which the inequality conditions in Proposition 3 hold. We also show with the help of a numerical exercise that the partial effect is predominantly negative over all combinations of $(s_{rk}, s_{rl}, \theta_r(\boldsymbol{x}_r))$. On the grounds of these findings one should expect in an application with moderate degree of dependence (i.e. Kendall's $\tau < 0.5$) that the conditional hazard decreases with the degree of dependence. This result is intuitive as the greater the dependence between spells, the smaller the probability of ending the k'th spell given that the previous l'th spell exceeded duration t_{rl} or t_{sl}. A model that assumes independence (such as the Markov Chain Model in Supplement S4) is therefore expected to overestimate the size of conditional hazards and the transition probabilities.

Several intermediate results are obtained in the proof of Proposition 3 that should be of wider interest. For this reason, we list them in what follows. The percentage change of the marginal survival for the k'th spell when the l'th spell is shorter (larger s_{rk} means shorter duration t_{rk}) is

$$\frac{\partial \log S_{rk}(s_{rk}|S_{rl} < s_{rl})}{\partial s_{rl}} < 0 \tag{19}$$

for $\theta_r(\boldsymbol{x}_r) > 0$. It means that if there is a positive dependence between spells through unobservables, a shorter duration in spell t_{rl} (larger s_{rl}) induces a shorter duration in the next spell t_{rk}. It is also shown that the partial derivative is zero in the case of independence ($\theta_r(\boldsymbol{x}_r) = 0$). The partial derivative of the log of equation (16) with respect to θ is

$$\frac{\partial \log S_{rk}(s_{rk}|S_{rl} < s_{rl})}{\partial \theta_r(\boldsymbol{x}_r)} > 0, \tag{20}$$

for $\theta_r(\boldsymbol{x}_r) > 0$. Since the conditional survival for the k'th spell given t_{rl} increases with the dependence between the spells, a shorter duration in spell t_{rl} is more likely followed by a shorter duration in the next spell t_{rk} when the positive dependence is stronger.

The partial derivative of the log of (17) with respect to s_{rl} is

$$\frac{\partial \log \lambda_{rk}(s_{rk}|S_{rl} \le s_{rl})}{\partial s_{rl}} > 0, \tag{21}$$

for $\theta_r(\boldsymbol{x}_r) > 0$. This means that when there are multiple spells with positive dependence, conditional on the event that the past spell survives up to t_{rl}, the hazard rate of the k'th spell decreases with t_{rl} for all t_{rk}. For $\theta_r(\boldsymbol{x}_r) = 0$, the hazard does not depend on s_{rl} and therefore the partial derivative is zero.

2.2. Estimation

An important advantage of crafting a model with the help of copulas is that it has a modular structure. This allows for convenient change or replacement of the different model components. The modular structure also enables separate or sequential estimation of model components, which reduces complexity. This is practically useful for example if joint estimation is numerically infeasible. It also brings greater flexibility in estimating the model components by means of various techniques. In particular, there are two sets of components. The first comprises of the copula functions with unknown parameters $\boldsymbol{\alpha}_l$ for $l = 0, 1, 2$ as given in equation (14). Here any Archimedean copula can be chosen and they can be different. The second set of components consists of the marginal survival functions

$$S_{rk}(t; \boldsymbol{x}_{rk}, \boldsymbol{\gamma}_{rk}) = \exp\left[-\int_0^t \lambda_{rk}(s; \boldsymbol{x}_{rk}, \boldsymbol{\gamma}_{rk}) ds\right] \tag{22}$$

with unknown parameters $\boldsymbol{\gamma}_{rk}$. Any model for $\lambda_{rk}(t; \boldsymbol{x}_{rk}, \boldsymbol{\gamma}_{rk})$ can be specified. Examples include the log-normal accelerated failure time model, log-logistic proportional odds model, odd-rate transformation model with Gompertz baseline hazard and proportional hazards model with Weibull baseline hazard, etc. An extension to semiparametric models for λ_{rk} such as the Cox proportional hazards model is possible, although joint estimation of all model components is numerically challenging as the partial likelihood presentation for the estimation of $\boldsymbol{\gamma}_{rk}$ is lost. In this case it is more convenient to use stepwise estimation as elaborated further below. The parametric piecewise-constant hazard model is a practical and flexible parametric alternative as it approximates the baseline hazard without strong restrictions.

To estimate the model in (5), it is useful to exploit that the likelihood is separable due to the two layers of the copula structure. The estimation can be done separately for each layer to reduce complexity. For the copula in the lower layer, the estimation bases on the joint distribution of observed durations in each state. For a random sample of $i = 1, \dots, n$ observations, the likelihood for C_r is

$$L_r(\boldsymbol{\alpha}_r, \boldsymbol{\gamma}_r; \boldsymbol{t}_r, \boldsymbol{x}_r) = \prod_{i=1}^n \left[\frac{\partial^2 C_r(s_{r1i}, s_{r2i}; \boldsymbol{x}_{ri}, \boldsymbol{\alpha}_r)}{\partial S_{r1} \partial S_{r2}} \times \prod_{k=1}^2 \frac{\partial S_{rk}(t_{rki}; \boldsymbol{x}_{rki}, \boldsymbol{\gamma}_{rk})}{\partial t_{rk}}\right], \tag{23}$$

for $r = 1, 2$. The likelihood for the upper layer of Model (5), i.e. C_0, bases on the nested structure of the copula model, in which C_0 is not only the copula for S_{1k} and S_{2k}, but also the copula for S_{11} and S_{22}, and for S_{12} and S_{21}. In this case, the likelihood is based on the derivative of $C_0(s_{11}, s_{21})$, $C_0(s_{11}, s_{22})$, $C_0(s_{12}, s_{21})$, $C_0(s_{12}, s_{22})$ w.r.t. all components of $\boldsymbol{t}$:

$$\begin{aligned} L_0(\boldsymbol{\alpha}_0, \boldsymbol{\gamma}; \boldsymbol{t}, \boldsymbol{x}) = \prod_{i=1}^n \Bigg\{ \prod_{k=1}^2 &\left[\frac{\partial^2 C_0(s_{1ki}, s_{2ki}; \boldsymbol{x}_i, \boldsymbol{\alpha}_0)}{\partial S_{1k} \partial S_{2k}} \times \prod_{r=1}^2 \frac{\partial S_{rk}(t_{rki}; \boldsymbol{x}_{rki}, \boldsymbol{\gamma}_{rk})}{\partial t_{rki}}\right] \\ &\times \frac{\partial^2 C_0(s_{11i}, s_{22i}; \boldsymbol{x}_i, \boldsymbol{\alpha}_0)}{\partial S_{11} \partial S_{22}} \times \frac{\partial S_{11}(t_{11i}; \boldsymbol{x}_{11i}, \boldsymbol{\gamma}_{11})}{\partial t_{11i}} \times \frac{\partial S_{22}(t_{22i}; \boldsymbol{x}_{22i}, \boldsymbol{\gamma}_{22})}{\partial t_{22i}} \\ &\times \frac{\partial^2 C_0(s_{12i}, s_{21i}; \boldsymbol{x}_i, \boldsymbol{\alpha}_0)}{\partial S_{12} \partial S_{21}} \times \frac{\partial S_{12}(t_{12i}; \boldsymbol{x}_{12i}, \boldsymbol{\gamma}_{12})}{\partial t_{12i}} \times \frac{\partial S_{21}(t_{21i}; \boldsymbol{x}_{21i}, \boldsymbol{\gamma}_{21})}{\partial t_{21i}} \Bigg\}. \end{aligned} \tag{24}$$

There are alternative approaches to writing the likelihood in (24). A simplified variant that incorporates the dependence across states for the same spell but ignores state dependence for different spells is given in Supplement S2. Supplement S3 outlines how restrictions on the range and order of $\theta_l(\boldsymbol{x})$ for $l = 0, 1, 2$ can be conveniently imposed.

Three maximum likelihood approaches are considered for estimation: full maximum likelihood estimation (FMLE), partial or pooled maximum likelihood estimation (PMLE), and pseudo maximum likelihood estimation (PsMLE). Because these are well established approaches we omit a detailed presentation including their large sample properties.

We define FMLE as separate one step estimation of $\boldsymbol{\alpha}_r$ and $\boldsymbol{\gamma}_r$ using (23) and $\boldsymbol{\alpha}_0$ and $\boldsymbol{\gamma}$ using (24). While being efficient as it models all dependencies, it is also numerically more challenging than ignoring them. In our numerical analysis, the estimation of FMLE takes from a couple of minutes up to 15 minutes, which suggests that it is practicable. PMLE corresponds to ignoring any dependence structure by only estimating $\boldsymbol{\gamma}$ with a model for (22). It is well known that PMLE of $\boldsymbol{\gamma}$ in the context of dependent data is consistent, although inefficient. For a detailed treatment of PMLE see Wooldridge (2010). While being numerically most convenient, the main disadvantage of PMLE is that $\boldsymbol{\alpha}$ in the dependence structure is not estimated, which is of major interest in our model and it is also required for performing comparative statics (compare subsection 2.1). PMLE is therefore not sufficient in this context. PsMLE is done in two steps in the context of copula models (Ruppert and Matteson, 2015, ch. 8). First, the parameters of the marginals are estimated and second, the copula parameters are estimated. $\boldsymbol{\gamma}$ is estimated by PMLE in the first step. Second step estimation of $\boldsymbol{\alpha}$ uses the likelihoods in (23) and (24) with pre-estimated $\boldsymbol{\gamma}$ from the first step. An extension to a semiparametric model for S_{rk} is straightforward in the context of PsMLE, as the usual Cox proportional hazards model can be fitted in the first step by ignoring the copula structure. Such an extension is much more complicated for FMLE as the numerical convenience of partialling out the nonparametric baseline hazard is being lost.

A standard approach in empirical research to modelling multiple transitions between states is the Markov Chain Model. We present in Supplement S4 how this approach substantially restricts models (1) and (5). Our suggested model can be directly used to test for the additional restrictions of the Markov Chain Model.

2.3. Finite sample performance

We generate data for different models and apply the estimation approaches of Subsection 2.2 to investigate the finite sample properties under correct and incorrect model specification. The full results and details are given in Supplement S5. We separately investigate robustness properties when the copula structure or the model for the marginal distribution is misspecified. Our results confirm previous findings for the single state model that the efficiency losses of PMLE are minor and choosing a wrong copula function only slightly affects the estimates (Lo et al., 2020, and Lipowski et al., 2021). We also consider scenarios, where θ depends on covariates. While PMLE retains consistency, FMLE looses this property under misspecification of $\theta(\boldsymbol{x})$. In particular, a misspecified $\theta(\boldsymbol{x})$ leads to inconsistent estimates of $\boldsymbol{\gamma}$. Based on these observations, it is recommended in practice to compare the estimated $\boldsymbol{\gamma}$ from FMLE with those from PMLE. When they are different, it would be evidence of misspecification and PsMLE is preferable.

In relation to modelling the marginal survivals, our simulations provide evidence of choosing a flexible model for a partially unknown marginal survival practically affects neither the estimates for the dependence parameter nor the marginal survival. In particular, the flexible piecewise-constant (PWCON) hazard model gives accurate estimates provided that the other aspects of the model are correctly specified. In the next section, we therefore use the PWCON and the semiparamaterix Cox proportional hazards (CoxPH) model for comparison as the first step estimators in pseudo MLE for an application to labour market transitions.

3. Application

To show that our suggested model gives insightful results, it is applied to administrative labour market data from Denmark. These comprise several linked administrative register for the population. They are provided by Statistics Denmark and contain a wealth of variables, including weekly observations of government transfer payments and monthly observations of employment statistics. The sample is restricted to individuals born in 1951 and resident in the capital region of Denmark in 2008. We construct durations in different labour market states for the period 2008 until 2016. The individuals in our sample attain the usual retirement age in 2016 with independent censoring at the end of 2016 due to end of the observation period. We group different labour market states into two states by following the classification system of register-based labour force statistics by Statistics Denmark (2019). State 1 is the economically active state, which contains employment and gross unemployment. State 2 is the economically inactive, or out of the labour force state, which contains all the other states. Individuals are therefore observed in two states with transitions between these two states. In line with our theoretical model we only consider the first two spells for each state for simplicity. In the end, we have 17,114 individuals with 19,188 spells in state 1 and 17,058 spells in state 2. A detailed composition of the spells in each state is shown in Table S6 in Supplement S7.

We estimate our suggested model by PsMLE and PMLE using PWCON and the popular CoxPH model in the first step for comparison. We also estimate the Markov Chain Model (compare Supplement S4) in comparison. We estimate two variants for the second step of PsMLE: one Clayton dependence structure with and one without covariate dependence. We include both time-constant variables and variables that vary across spells.

We report estimation results for selected variables of interest for PWCON and the Markov Chain Model in Tables 1 and 2. The remaining results are shown in Tables S7 and S8 in the supplementary material. The results for the CoxPH model are also only reported in Tables S9 and S10 in Supplement S7, because of their similarity to the PWCON results.

Table 1 shows the estimated coefficients of a subset of variables for the survival function. It is apparent that the PWCON and Markov Chain Model give partly sizably different estimates. As the Markov Chain Model assumes that hazards are the same for all spells, it is misspecified if variables have different partial effects for different spells. We can see that sometimes the coefficients of the Markov Chain Model are between those of the two spells of PWCON, and sometimes they are above or below the maximum or

Table 1
Selected estimation results of the survival function for all states and all spells.

Model	PsMLE with PWCON							
	state 1 - Economically active				state 2 - Inactivity			
	spell 1		spell 2		spell 1		spell 2	
Covariates	coef.	(s.e.)	coef.	(s.e.)	coef.	(s.e.)	coef.	(s.e.)
education: middle	0.068	(0.026)	-0.070	(0.047)	0.168	(0.038)	0.169	(0.057)
education: high	-0.207	(0.038)	-0.133	(0.063)	0.223	(0.047)	0.179	(0.071)
married	0.089	(0.021)	0.016	(0.037)	-0.080	(0.031)	-0.059	(0.047)
sickness benefit	0.306	(0.025)	0.213	(0.051)	0.048	(0.041)	-0.137	(0.062)
disposable income: high	-0.202	(0.027)	-0.032	(0.052)	0.005	(0.040)	0.036	(0.066)
pension deposit: high	-0.786	(0.050)	-0.223	(0.074)	0.010	(0.060)	0.069	(0.089)
wealth: high	0.293	(0.037)	0.210	(0.057)	-0.288	(0.047)	-0.179	(0.071)
Model	Markov Chain							
	state 1 - Economically active				state 2 - Inactivity			
	spell 1		spell 2		spell 1		spell 2	
Covariates	coef.	(s.e.)	coef.	(s.e.)	coef.	(s.e.)	coef.	(s.e.)
education: middle	0.045	(0.023)			0.233	(0.032)		
education: high	-0.132	(0.032)			0.286	(0.039)		
married	0.070	(0.018)			-0.109	(0.026)		
sickness benefit	0.252	(0.022)			-0.039	(0.034)		
disposable income: high	-0.161	(0.024)			0.093	(0.033)		
pension deposit: high	-0.442	(0.04)			0.224	(0.049)		
wealth: high	0.207	(0.031)			-0.334	(0.039)		

Notes: The estimates for the Markov Chain Model for the two spells in each state are the same.

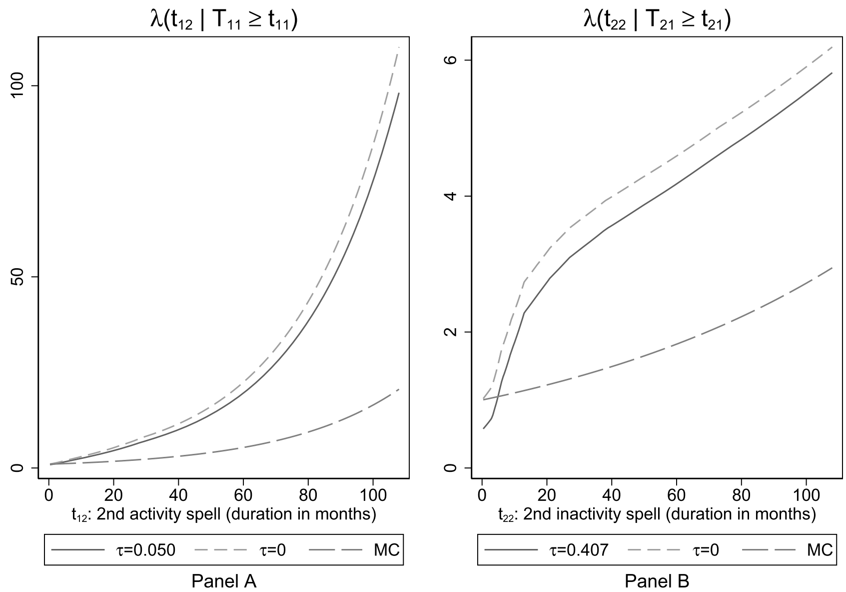
minimum of the two spells, suggesting misspecification of the hazard function. On the contrary, the estimation results of PWCON are very close to CoxPH in Tables S9, suggesting that the parametric specification of the hazard function in PWCON is not rejected by the semiparametric alternative. For most variables, the partial effects change sign across the two states. For example, higher income leads to longer activity spells and shorter inactivity spells, showing evidence of substitution effect between labour and leisure, while higher wealth leads to shorter activity spells and longer inactivity spells, showing evidence of an income effect between labour and leisure. Pension deposit consists of both private pension and employer-administered pension. From the results we can see that the partial effect of pension deposit is dominated by the substitution effect rather than the income effect, suggesting that the employment relationship plays a larger role in the composition of pension deposit.

Table 2 shows the estimated coefficients of the dependence structure that reflect spell dependence and state dependence. The corresponding PsMLEs on the grounds of the CoxPH model (Table S10) are again similar except for spell dependence in state 1, where there are numerical issues for PsMLE after CoxPH. Since the estimates of the first step of PsMLE for PWCON and CoxPH are very close, it could be that non-smoothness pattern of the CoxPH baseline hazards are the cause of these problems. It is apparent from Table 2 that $\theta_l(\boldsymbol{x}) > 0$ for $l = 0, 1, 2$. In particular, dependence between two inactivity spells is rather strong. We also estimate the model under additional constrains as the restriction $\theta_0(\boldsymbol{x}) \leq \theta_1(\boldsymbol{x})$ is not satisfied. Table S11 shows these results, which are the same except that $\theta_0(\boldsymbol{x}) = \theta_1(\boldsymbol{x})$. Our results also confirm that many covariates in the model for θ_l are significant. For example, the dependence between two inactivity spells is weaker for individuals with higher education level. It is therefore important to allow for this in the model. For the interpretation of the dependence estimates, we look at both the coefficient estimates and the conditional hazards. A positive coefficient means that the variable increases the degree of dependence. We can also see that the coefficients are largely different for different dependencies. It suggests that the underlying factors that drive the dependencies are different. For the dependence between two activity spells, education and marriage increases it, while illness decreases it. One underlying unobserved factor that drives dependence between activity spells is ability or motivation. People with higher education level tend to have higher variance of ability in different jobs. This makes dependence between activity spells stronger. Married people have a higher variance of household labour supply, which corresponds to stronger dependence. Individuals in poorer health condition have a lower variance of working capability, which leads to weaker dependence. For the dependence between two inactivity spells, education has a negative effect, while income, pension deposit and wealth have positive effects. One underlying unobserved factor that drives dependence of inactivity spells is the financial need to work. People with higher income or wealth have higher variance in budget constraints and therefore in the financial need to work, leading to greater dependence. The signs of the education variables are opposite from those for the active state. This could be related to the unobserved attitude or willingness to work. For example, individuals with higher educational level tend to be more homogeneous in terms of working attitude, leading to a weaker dependence. For the dependence between the two states, or between an activity spell and an inactivity spell, education and illness have negative effects, while income and pension deposit have positive effects. In this case, greater dependence means that individuals have more likely either long activity and inactivity spells, or short activity and inactivity spells. From Table 1 we can see that higher education leads to longer activity spells and shorter inactivity spells. Illness on the contrary leads to shorter economic activity and longer inactivity spells, which reduces the positive dependence between the two states.

Table 2
Selected estimation results for the dependence parameters $\boldsymbol{\alpha}_l$ and τ_l.

Model	PsMLE with PWCON and the Clayton copula					
	Spell dependence				State dependence	
	State 1		State 2			
Covariates	$\hat{\boldsymbol{\alpha}}_1$	(s.e.)	$\hat{\boldsymbol{\alpha}}_2$	(s.e.)	$\hat{\boldsymbol{\alpha}}_0$	(s.e.)
$\theta_l(\boldsymbol{x}_l)$						
education: middle	0.220	(0.322)	-0.167	(0.093)	-0.100	(0.063)
education: high	0.585	(0.378)	-0.712	(0.138)	-0.192	(0.073)
married	0.650	(0.711)	0.079	(0.072)	0.083	(0.059)
sickness benefit	-0.598	(0.424)	-0.119	(0.071)	-0.857	(0.131)
disposable income: high	0.311	(0.394)	0.134	(0.077)	0.723	(0.089)
pension deposit: high	-0.120	(0.521)	0.216	(0.092)	0.377	(0.073)
wealth: high	0.0693	(0.319)	0.141	(0.079)	-0.075	(0.060)
constant	-2.254	(0.870)	0.317	(0.073)	-1.130	(0.073)
$\hat{\tau}$	0.050		0.407		0.139	
$\theta_l(\boldsymbol{x}_l) = \theta_l$						
constant	-1.648	(0.082)	0.375	(0.038)	-0.718	(0.027)
$\hat{\tau}$	0.088		0.421		0.196	

Notes: First step: piecewise-constant baseline hazard model. Copula: Clayton. The estimated Kendall's tau is for the reference individual (basic education, unmarried, low disposable income, low or none pension deposit, low or negative wealth). Bootstrap standard errors.



Notes: MC refers to the Markov Chain Model.

Fig. 3. Conditional hazards for the second spell conditional on the first spell for two states.

To show how the dependence estimates contribute to analysing lagged duration dependence, we compute the conditional hazard in (17) for terminating the second spell given the duration of the first spell for the two states for a reference individual (basic education, unmarried, low disposable income, low or none pension deposit, low or negative wealth) and illustrate them in Fig. 3. The results in Table 2 provide evidence of different degrees of dependence in the two states for the reference individual. While it is rather low (Kendall-$\tau = 0.050$) for the activity state, it is sizable (Kendall-$\tau = 0.407$) for the inactivity state, highlighting the need for a flexible specification. By using our theory for the Clayton copula, we compute the conditional hazards according to (52) in Appendix A.1 as

$$\lambda_{r2}(t_{r2}|S_{r1} < s_{r1}; \boldsymbol{x}_2) = \lim_{\epsilon \to 0} \Pr(S_{r2} \in [s_{r2}, s_{r2} + \epsilon]|S_{r1} < s_{r1}; S_{r2} < s_{r2}; \boldsymbol{x}_r)/\epsilon$$
$$= \frac{1}{s_{r2}^{\theta_r+1} s_{r1}^{-\theta_r} + s_{r2} - s_{r2}^{\theta_r+1}}$$

for $r = 1, 2$. To better illustrate the results, we transform the survival probabilities to durations using the estimates from the first step PMLE with PWCON. In this way, we are able to report the conditional hazards as a function of durations, rather than survival probabilities. We do the same using the Markov Chain Model for comparison. Since the conditional hazard depends on the survival probability of both spells, we fix the duration of the first spell to its median level for both states and see how the conditional hazard changes with the duration of the second spell. Specifically, in Fig. 3 Panel A, we fix the duration of the first spell to be 72 months and the corresponding survival probability is 0.18. In Panel B, we fix the duration of the first spell to be 32 months and the corresponding survival probability is 0.65. For both states, the conditional hazards increase with the duration of the second spell for a fixed duration of the first spell. It means that the longer the second spell, the more likely the individual transits out of that spell given the durations of the two spells. Comparing the conditional hazards with positive dependence ($\tau > 0$) to those with assumed independence ($\tau = 0$), we find that in line with the theoretical result in Proposition 3, assuming independence overestimates the conditional hazards for all durations of the second spell. The Markov Chain Model also assumes independence, but it has an extra restriction that the marginal hazard is a constant. Comparing the conditional hazards with a flexible specification for the marginal hazards ($\tau = 0$ with PWCON) to those of the Markov Chain Model, we find that the Markov Chain Model considerably underestimates the conditional hazards. So, besides restrictions on the dependence parameter, restrictions on the marginal survival functions also matter for the conditional hazards. These comparative statics can also be computed for the relationship between any spell in the two states. It can be a useful tool to analyse the dependence between multiple states and spells. For example, with information on labour market history of an individual, firms can benefit from knowing how likely an employee will continue working for the firm, and job centres and policymakers can benefit from knowing how likely an unemployed or inactive person will become employed given her employment biography. Our results show that using a flexible specification for the marginal survival functions and allowing the spells to be dependent are vital to provide correct comparative statics and thus understand problems of practical importance.

4. Conclusions

We introduce a modular multiple states and spells duration model which allows for separate modelling of the dependencies and marginal distributions. This gives a high degree of flexibility to the empirical researcher as a range of models is compatible with the components. We adopt a nested Archimedean copula structure with different layers for spell and state dependencies. We establish the formal link to models with unobserved heterogeneities or frailties that are popular in applied economic research and we show that the Markov Chain Model is a restrictive special case of our model. Our model allows frailty distributions to depend on observable regressors, which is not permitted in the classical models. We suggest different routes of maximum likelihood estimation and show with the help of simulations that pseudo maximum likelihood is the most practical and robust route. We also show that the flexible piecewise-constant hazards model does a good job in approximating unknown hazards and gives almost the same results as the semiparametric Cox proportional hazards model in our application, while being numerical more stable. In our application to labour market transitions we find evidence of positive dependencies between spells and states and that the dependence structure is a function of covariates. A careful modelling of the dependence structure is therefore necessary to avoid inconsistencies. The Markov Chain Model is found to give substantially different results, which we explain by its set of stronger restrictions.

Data sets and computer code

Simulated data and sample STATA code for the model of this paper are available on github.com: https://github.com/ralfawilke/copula_spells_states.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article. The data access was funded by the Pension Research Centre (PeRCent), which receives base funding from the Danish pension funds and Copenhagen Business School.

Acknowledgements

We thank the anonymous reviewers for their helpful comments.

Appendix A

A.1. Proofs

Proof of Proposition 1. We begin by showing how the copula model can be derived from the shared frailty model in the lower layer (having the same risk r) when the multiple spells T_{r1} and T_{r2} are correlated due to a shared frailty denoted as V_r. From equation (8) in Assumption 1(i) the marginal survival function of the k-th spell with risk r conditional on frailty v_r is

$$S^*_{rk}(t; v_r, \boldsymbol{x}) = \exp[-v_r \Lambda^*_{rk}(t; \boldsymbol{x})]. \tag{25}$$

The conditional marginal survival S^*_{r1} and S^*_{r2} have the shared frailty v_r in common. The unconditional marginal survival function is derived by integrating (25) with respect to the distribution of V_r, i.e., $F_r(v_r; \boldsymbol{x})$, which is

$$S_{rk}(t;\boldsymbol{x}) = \int_{v_r} \exp[-v_r \Lambda^*_{rk}(t;\boldsymbol{x})] dF_r(v_r;\boldsymbol{x}) \\ = \phi^{-1}_{\theta_r(\boldsymbol{x})}[\Lambda^*_{rk}(t;\boldsymbol{x})]. \tag{26}$$

$\phi^{-1}_{\theta_r(\boldsymbol{x})}$ is the Laplace transform with respect to the distribution F_r. Inverting the function in (26), $\Lambda^*_{rk}(t;\boldsymbol{x})$ can be derived as

$$\Lambda^*_{rk}(t;\boldsymbol{x}) = \phi_{\theta_r(\boldsymbol{x})}[S_{rk}(t;\boldsymbol{x})]. \tag{27}$$

Given the conditional independent assumption in Assumption 1(iii), the joint distribution of the multiple spells for risk r (T_{r1} and T_{r2}) is derived as

$$\begin{aligned} H_r(\boldsymbol{t}_r;\boldsymbol{x}_r) &= \int_{v_r} S^*_{r1}(t;v_r,\boldsymbol{x}) S^*_{r2}(t;v_r,\boldsymbol{x}) dF_r(v_r;\boldsymbol{x}) \\ &= \int_{v_r} \exp[-v_r(\Lambda^*_{r1}(t;\boldsymbol{x}) + \Lambda^*_{r2}(t;\boldsymbol{x}))] dF_r(v_r;\boldsymbol{x}) \\ &= \phi^{-1}_{\theta_r(\boldsymbol{x})}[\Lambda^*_{r1}(t;\boldsymbol{x}) + \Lambda^*_{r2}(t;\boldsymbol{x})] \\ &= \phi^{-1}_{\theta_r(\boldsymbol{x})}[\phi_{\theta_r(\boldsymbol{x})}[S_{r1}(t;\boldsymbol{x})] + \phi_{\theta_r(\boldsymbol{x})}[S_{r2}(t;\boldsymbol{x})]]. \end{aligned} \tag{28}$$

The second equality comes from (25). The third equality comes from the definition of the Laplace transform. The fourth equality comes from (27).

Note that the last line in (28) is an Archimedean copula with the copula generator equal to $\phi_{\theta_r(\boldsymbol{x})}$, we can rewrite the joint distribution of the multiple spells with risk r in the second layer as

$$H_r(\boldsymbol{t}_r;\boldsymbol{x}_r) = C_{\theta_r(\boldsymbol{x})}(S_{r1}(t;\boldsymbol{x};), S_{r2}(t;\boldsymbol{x};)). \tag{29}$$

Equation (29) shows that the joint distribution of the unconditional marginal survival function S_{rk} can be modelled as a copula function. The dependency is driven by the shared frailty V_r present in the multiple spells T_{r1} and T_{r2}.

We, next, derive the joint distribution of the multiple spells from both risks in the first layer when these spells are correlated due to a shared frailty V_0. There are three frailties in this model (V_0, V_1, V_2) and their joint distribution is $F(v_0, v_1, v_2;\boldsymbol{x})$. The joint distribution of $(T_{11}, T_{12}, T_{21}, T_{22})$ is

$$\begin{aligned} H(\boldsymbol{t};\boldsymbol{x}) &= \int_{v_0}\int_{v_1}\int_{v_2} \prod_{r=1}^{2}\prod_{k=1}^{2} S^*_{rk}(t;v_r,\boldsymbol{x}) dF(v_0,v_1,v_2;\boldsymbol{x}) \\ &= \int_{v_0}\int_{v_1}\int_{v_2} \prod_{r=1}^{2}\prod_{k=1}^{2} \exp[-v_r\Lambda^*_{rk}(t;v_r,\boldsymbol{x})] dF(v_0,v_1,v_2;\boldsymbol{x}). \end{aligned} \tag{30}$$

The second equality comes from (25). Due to Assumption 1(ii), the conditional independent assumption of (v_1, v_2) given v_0, the joint distribution for the frailty is

$$F(v_0,v_1,v_2;\boldsymbol{x}) = F_{1|0}(v_{1|0};\boldsymbol{x}) F(v_{2|0};\boldsymbol{x}) F_0(v_0;\boldsymbol{x}).$$

The joint distribution in (30) can be rewritten as

$$\begin{aligned} H(\boldsymbol{t};\boldsymbol{x}) &= \int_{v_0}\int_{v_{1|0}}\int_{v_{2|0}} \prod_{r=1}^{2}\prod_{k=1}^{2} \exp[-v_r\Lambda^*_{rk}(t;v_r,\boldsymbol{x})] dF_{1|0}(v_{1|0};\boldsymbol{x}) dF_{2|0}(v_{2|0};\boldsymbol{x}) dF_0(v_0;\boldsymbol{x}) \\ &= \int_{v_0} \left\{ \int_{v_{1|0}} \exp[-v_1(\Lambda^*_{11}(t;v_1,\boldsymbol{x}) + \Lambda^*_{12}(t;v_1,\boldsymbol{x}))] dF_{1|0}(v_{1|0};\boldsymbol{x}) \right\} \\ &\quad \times \left\{ \int_{v_{2|0}} \exp[-v_2(\Lambda^*_{21}(t;v_2,\boldsymbol{x}) + \Lambda^*_{22}(t;v_2,\boldsymbol{x}))] dF_{2|0}(v_{2|0};\boldsymbol{x}) \right\} dF_0(v_0;\boldsymbol{x}). \end{aligned} \tag{31}$$

In (31), the item $\exp[-v_r\Lambda^*_{rk}(t;v_r,\boldsymbol{x})]$ is integrated with respect to the mixing distribution which is a conditional frailty distribution $F_{r|0}$. We have to derive the Laplace transform for this conditional distribution in terms of $\phi_{\theta_r(\boldsymbol{x})}$ and $\phi_{\theta_0(\boldsymbol{x})}$. This Laplace transform, by definition, is

$$\phi^{-1}_{\theta_{r|0}(\boldsymbol{x})}(s) = \int_{v_{r|0}} \exp[-v_{r|0}s] dF_{r|0}(v_{r|0};\boldsymbol{x}). \tag{32}$$

For this purpose, we rewrite $S_{rk}(t;\boldsymbol{x})$ in (26) as a mixture of conditional frailty distribution, i.e.,

$$S_{rk}(t;\boldsymbol{x}) = \int_{v_r} \exp[-v_r\Lambda^*_{rk}(t;\boldsymbol{x})]dF_r(v_r;\boldsymbol{x})$$
$$= \int_{v_0}\left\{\int_{v_{r|0}} \exp[-v_{r|0}\Lambda^*_{rk}(t;\boldsymbol{x})]dF_{r|0}(v_{r|0};\boldsymbol{x})\right\}dF_0(v_0;\boldsymbol{x}). \tag{33}$$

The inner bracket is the Laplace transform $\phi^{-1}_{\theta_{r|0}(\boldsymbol{x})}$ defined in (32). (33) can be rewritten as

$$S_{rk}(t;\boldsymbol{x}) = \int_{v_0} \phi^{-1}_{\theta_{r|0}(\boldsymbol{x})}[\Lambda^*_{rk}(t;\boldsymbol{x})]dF_0(v_0;\boldsymbol{x}). \tag{34}$$

Now, equate it with $S_{rk}(t;\boldsymbol{x})$ in (26) to arrive to the important observation that

$$\phi^{-1}_{\theta_r(\boldsymbol{x})}[s] = \int_{v_0} \phi^{-1}_{\theta_{r|0}(\boldsymbol{x})}[s]dF_0(v_0;\boldsymbol{x}). \tag{35}$$

Using (12) in Assumption 1(v), i.e., $\phi^{-1}_{\theta_{r|0}(\boldsymbol{x})}[s] = \xi_r(s;\boldsymbol{x}_r)^{v_0}$ for some function ξ_r, equation (35) becomes

$$\phi^{-1}_{\theta_r(\boldsymbol{x})}[s] = \int_{v_0} \xi_r(s;\boldsymbol{x}_r)^{v_0}dF_0(v_0;\boldsymbol{x})$$
$$= \int_{v_0} \exp[-v_0\log\xi_r(s;\boldsymbol{x}_r)]dF_0(v_0;\boldsymbol{x})$$
$$= \phi^{-1}_{\theta_0(\boldsymbol{x})}[-\log\xi_r(s;\boldsymbol{x}_r)]. \tag{36}$$

The second equality is the mathematical operation using exp and log. The third equality uses the definition of the Laplace transform. Inverting equation (36), we have

$$\phi_{\theta_0(\boldsymbol{x})}\circ\phi^{-1}_{\theta_r(\boldsymbol{x})}[s] = -\log\xi_r(s;\boldsymbol{x}_r).$$

After rearranging, it becomes

$$\xi_r(s;\boldsymbol{x}_r) = \exp[-\phi_{\theta_0(\boldsymbol{x})}\circ\phi^{-1}_{\theta_r(\boldsymbol{x})}(s)]. \tag{37}$$

Since $\phi^{-1}_{\theta_{r|0}(\boldsymbol{x})}[s] = \xi_r(s;\boldsymbol{x}_r)^{v_0}$, we obtain the expression of the Laplace transform for the conditional frailty distribution in terms of $\phi_{\theta_r(\boldsymbol{x})}$ and $\phi_{\theta_0(\boldsymbol{x})}$. Namely,

$$\phi^{-1}_{\theta_{r|0}(\boldsymbol{x})}(s) = \exp[-v_0\phi_{\theta_0(\boldsymbol{x})}\circ\phi^{-1}_{\theta_r(\boldsymbol{x})}(s)]. \tag{38}$$

For the final step, we use the following result that is obtained substituting the Laplace transform in (38) into the joint distribution in (31):

$$\int_{r|0} \exp(-v_r\Lambda^*_{rk}(t;v_r,\boldsymbol{x}))dF_{r|0}(v_{r|0};\boldsymbol{x}) = \phi^{-1}_{\theta_{r|0}(\boldsymbol{x})}[\Lambda^*_{rk}(t;v_r,\boldsymbol{x}))]$$
$$= \exp[-v_0\phi_{\theta_0(\boldsymbol{x})}\circ\phi^{-1}_{\theta_r(\boldsymbol{x})}[\Lambda^*_{rk}(t;v_r,\boldsymbol{x}))]]. \tag{39}$$

We now obtain

$$H(\boldsymbol{t};\boldsymbol{x}) = \int_{v_0}\left\{\int_{v_{1|0}} \exp[-v_1(\Lambda^*_{11}(t;v_1,\boldsymbol{x}) + \Lambda^*_{12}(t;v_1,\boldsymbol{x}))]dF_{1|0}(v_{1|0};\boldsymbol{x})\right\}$$
$$\times\left\{\int_{v_{2|0}} \exp[-v_2(\Lambda^*_{21}(t;v_2,\boldsymbol{x}) + \Lambda^*_{22}(t;v_2,\boldsymbol{x}))]dF_{2|0}(v_{2|0};\boldsymbol{x})\right\}dF_0(v_0;\boldsymbol{x})$$
$$= \int_{v_0}\left\{\exp[-v_0\phi_{\theta_0(\boldsymbol{x})}\circ\phi^{-1}_{\theta_1(\boldsymbol{x})}[\Lambda^*_{11}(t;v_1,\boldsymbol{x}) + \Lambda^*_{12}(t;v_1,\boldsymbol{x})]]\right\}$$
$$\times\left\{\exp[-v_0\phi_{\theta_0(\boldsymbol{x})}\circ\phi^{-1}_{\theta_2(\boldsymbol{x})}[\Lambda^*_{21}(t;v_2,\boldsymbol{x}) + \Lambda^*_{22}(t;v_2,\boldsymbol{x})]]\right\}dF_0(v_0;\boldsymbol{x})$$

$$= \int_{v_0} \left\{ \exp[-v_0 \sum_{r=1}^{2} \phi_{\theta_0(\boldsymbol{x})} \circ \phi^{-1}_{\theta_r(\boldsymbol{x})} [\Lambda^*_{r1}(t; v_r, \boldsymbol{x}) + \Lambda^*_{r2}(t; v_r, \boldsymbol{x})]] \right\} dF_0(v_0; \boldsymbol{x})$$

$$= \phi^{-1}_{\theta_0(\boldsymbol{x})} \left\{ \sum_{r=1}^{2} \phi_{\theta_0(\boldsymbol{x})} \circ \phi^{-1}_{\theta_r(\boldsymbol{x})} \left[\Lambda^*_{r1}(t; v_r, \boldsymbol{x}) + \Lambda^*_{r2}(t; v_r, \boldsymbol{x})\right] \right\}$$

$$= \phi^{-1}_{\theta_0(\boldsymbol{x})} \left\{ \phi_{\theta_0(\boldsymbol{x})} \circ \phi^{-1}_{\theta_1(\boldsymbol{x})} \left[\phi_{\theta_1(\boldsymbol{x})}(S_{11}) + \phi_{\theta_1(\boldsymbol{x})}(S_{12})\right] + \phi_{\theta_0(\boldsymbol{x})} \circ \phi^{-1}_{\theta_2(\boldsymbol{x})} \left[\phi_{\theta_2(\boldsymbol{x})}(S_{21}) + \phi_{\theta_2(\boldsymbol{x})}(S_{22})\right] \right\}$$

$$= C_{0|\boldsymbol{x}}[C_{1|\boldsymbol{x}}(S_{11}, S_{12}), C_{2|\boldsymbol{x}}(S_{21}, S_{22})], \tag{40}$$

where the second equality comes from (39). The third equality uses the property of the exp function. The fourth equality comes from the Laplace transform with respect to F_0. The fifth equality comes from (27). The last equality uses the following definition:

$$C_{r|\boldsymbol{x}}(S_{r1}, S_{r2}) = \phi_{\theta_0(\boldsymbol{x})} \circ \phi^{-1}_{\theta_r(\boldsymbol{x})} \left[\phi_{\theta_r(\boldsymbol{x})}(S_{r1}) + \phi_{\theta_r(\boldsymbol{x})}(S_{r2})\right].$$

The equations leading to (40) show that the joint distribution of all multiple states and multiple spells derived from the shared frailty model can be summarised by a nested Archimedean copula structure. We remark here that our proof shares ideas of Hasan and Braekers (2021) in the context of clustered data. Their model is, however, not the same as ours, since they only consider one frailty V_0, which is common to all layers. They do not derive the inner generator in their model using the frailties specific to the lower layer, i.e. V_1 and V_2. □

Proof of Proposition 2. $\tau(\boldsymbol{x})$ for copula $C_{\boldsymbol{x}}$ is defined as

$$\tau(\boldsymbol{x}) = 4 \int_{[0,1]^2} C(\boldsymbol{x}) dC(\boldsymbol{x}) - 1. \tag{41}$$

$\tilde{\tau} = E_{\boldsymbol{x}}[\tau(\boldsymbol{x})]$ is therefore

$$\tilde{\tau} = \int_x \left(4 \int_{[0,1]^2} C(\boldsymbol{x}) dC(\boldsymbol{x}) - 1 \right) dF(\boldsymbol{x})$$

$$= 4 \int_x \int_{[0,1]^2} C(\boldsymbol{x}) dC(\boldsymbol{x}) dF(\boldsymbol{x}) - 1. \tag{42}$$

The unconditional copula is $E_{\boldsymbol{x}}[C_{\boldsymbol{x}}] = \int_x C_{\boldsymbol{x}} dF(\boldsymbol{x})$, and hence its Kendall's τ is

$$\tau = 4 \int_{[0,1]^2} \left(\int_x C_{\boldsymbol{x}} dF(\boldsymbol{x}) \right) d \left(\int_x C_{\boldsymbol{x}} F(\boldsymbol{x}) \right) - 1. \tag{43}$$

Equation (42) and (43) are identical if

$$\int_{[0,1]^2} \left(\int_x C_{\boldsymbol{x}} dF(\boldsymbol{x}) \right) d \left(\int_x C_{\boldsymbol{x}} F(\boldsymbol{x}) \right) = \int_x \int_{[0,1]^2} C(\boldsymbol{x}) dC(\boldsymbol{x}) dF(\boldsymbol{x}). \quad \square \tag{44}$$

Proof of Proposition 3. To simplify the exposition and without loss of generality, we focus on the lower layer in Model (5) and state 1 with two multiple spells $k = 1, 2$. The marginal survival distributions are $S_{11} = S_1$ and $S_{12} = S_2$ for simplicity of notation. We omit the conditioning on $\boldsymbol{x}$ and do the derivations for the Clayton copula.

The Clayton copula has copula generator $\phi_\theta(s) = (s^{-\theta} - 1)/\theta : (0, 1] \to [0, \infty)$ with inverse $\phi^{-1}_\theta(t) = ([1 + \theta t]_+)^{-1/\theta} : [0, \infty) \to (0, 1]$, where $[s]_+ = \max\{0, s\}$ and $\theta \in [0, \infty)$. The joint distribution of S_1 and S_2 is then

$$C(s_1, s_2) = (s_1^{-\theta} + s_2^{-\theta} - 1)^{-1/\theta}. \tag{45}$$

The joint distribution increases with s_k for $k = 1, 2$, i.e.

$$\frac{\partial C}{\partial s_k} = \left(\frac{C}{s_k} \right)^{\theta+1} > 0, \tag{46}$$

because $C < s_k$ for all s_1 and s_2 and $\theta > 0$. Note that $(C/s_1)^\theta > 1$ when $\theta < 0$. For positive dependence, C has a lower bound which is $C > s_1 s_2$, we have therefore $\partial C/\partial s_k > s_{k'}^{\theta+1}$ for all k and $k' \neq k$. For independence ($\theta = 0$), $\partial C/\partial s_k|_{\theta=0} = s_{k'}$. This implies that the joint distribution with positively dependent spells is always greater than the joint distribution under independence. Specifically we compute the following:

$$\frac{\partial \log C}{\partial \theta} = \frac{1}{\theta}\left[\left(\frac{C}{s_1}\right)^{\theta} \log s_1 + \left(\frac{C}{s_2}\right)^{\theta} \log s_2 - \log C\right] = \frac{p}{\theta} > 0, \tag{47}$$

where $p > 0$. For $\theta > 0$, we know from the Fréchet bound that $s_1 s_2 < C(s_1, s_2) < \min\{s_1, s_2\}$. Put $C = (s_1^{-\theta} + s_2^{-\theta} - 1)^{-1/\theta}$ into the following equation:

$$\left(\frac{C}{s_1}\right)^{\theta} \log s_1 = \left(\frac{s_2^{\theta}}{s_1^{\theta} + s_2^{\theta} - s_1^{\theta} s_2^{\theta}}\right) \log s_1. \tag{48}$$

Assuming without loss of generality that $C < s_1 < s_2$ and $s_1 s_2 > 0$, we have

$$\begin{aligned} p &= \left(\frac{s_1^{\theta} \log s_1 + s_2^{\theta} \log s_2}{s_1^{\theta} + s_2^{\theta} - s_1^{\theta} s_2^{\theta}}\right) - \log C \\ &> \left(\frac{(s_1^{\theta} + s_2^{\theta}) \log s_1}{s_1^{\theta} + s_2^{\theta}}\right) - \log C \\ &> \log s_1 - \log C > 0. \end{aligned}$$

The joint distribution C therefore monotonically increases with θ for all s_1 and s_2. The meaning of this result in the context of our running example is that stronger positive dependence between multiple spells increases the likelihood for both spells not being terminated at t_1 and t_2. For example, economic activity is more likely not terminated in both spells.

The conditional survival (16) becomes

$$S_2(s_2 | S_1 \le s_1) = C / s_1. \tag{49}$$

Taking log and differentiate it with s_1 gives

$$\frac{\partial \log S_2(s_2 | S_1 \le s_1)}{\partial s_1} = \frac{1}{s_1}\left[\left(\frac{C}{s_1}\right)^{\theta} - 1\right] < 0. \tag{50}$$

The rate of change of the conditional survival equals $1/s_1$ multiplied by the second term. The inequality (50) holds because $C < s_1$ and $\theta > 0$. This shows the inequality (19) for the Clayton copula. For $\theta = 0$ the conditional survival does not change with t_1 as the spells are independent.

The partial derivative of the log of equation (49) with respect to θ is

$$\frac{\partial \log S_2(s_2 | S_1 \le s_1)}{\partial \theta} = \frac{\partial \log(C)}{\partial \theta} = \frac{p}{\theta} > 0, \tag{51}$$

which is identical to (47). This means that the conditional survival for the second spell changes with θ in the same way as the joint distribution of the two spells and it increases with θ.

The conditional hazard (17) in the case of the Clayton copula is

$$\lambda_2(s_2 | S_1 \le s_1) = \frac{1}{s_2}\left(\frac{C}{s_2}\right)^{\theta}. \tag{52}$$

For $\theta = 0$, it depends on s_2 but not s_1, which is obvious. Taking the log of (52) and differentiation with respect to s_1 gives

$$\frac{\partial \log \lambda_2(s_2 | S_1 \le s_1)}{\partial s_1} = \frac{\theta}{C}\left(\frac{C}{s_1}\right)^{\theta+1} > 0, \tag{53}$$

which is (21) for the Clayton copula. For $\theta = 0$, the hazard rate of s_2 does not depend on s_1.

Taking the log of (52) and differentiating with respect to θ gives

$$\frac{\partial \log \lambda_2(s_2 | S_1 \le s_1)}{\partial \theta} = \left(\frac{C}{s_1}\right)^{\theta} \log s_1 + \left(\frac{C}{s_2}\right)^{\theta} \log s_2 - \log s_2. \tag{54}$$

To find the condition to make (54) negative, rewrite it as

$$\frac{\partial \log \lambda_2(s_2 | S_1 \le s_1)}{\partial \theta} = \frac{s_2^{\theta} \log s_1 + s_2^{\theta}(s_1^{\theta} - 1) \log s_2}{s_1^{\theta} + s_2^{\theta} - s_1^{\theta} s_2^{\theta}} < 0.$$

This is satisfied for $\theta = 0$. For $\theta > 0$ it can be shown to be equivalent to $s_1^{\frac{1}{1-s_1^{\theta}}} < s_2$. The minimum of s_2 for which this inequality holds is $s_1^{1/(1-s_1^{\theta})}$. It has several interesting properties: it increases with θ and s_1; it is always smaller than s_1 and it converges to s_1 as $\theta \to \infty$; the difference between s_1 and $s_1^{1/(1-s_1^{\theta})}$, the quantity $s_1 - s_1^{1/(1-s_1^{\theta})}$, increases with s_1 and decreases with θ. These properties are illustrated in Fig. 4. Combining these properties with a numerical exercise, we have the following results:

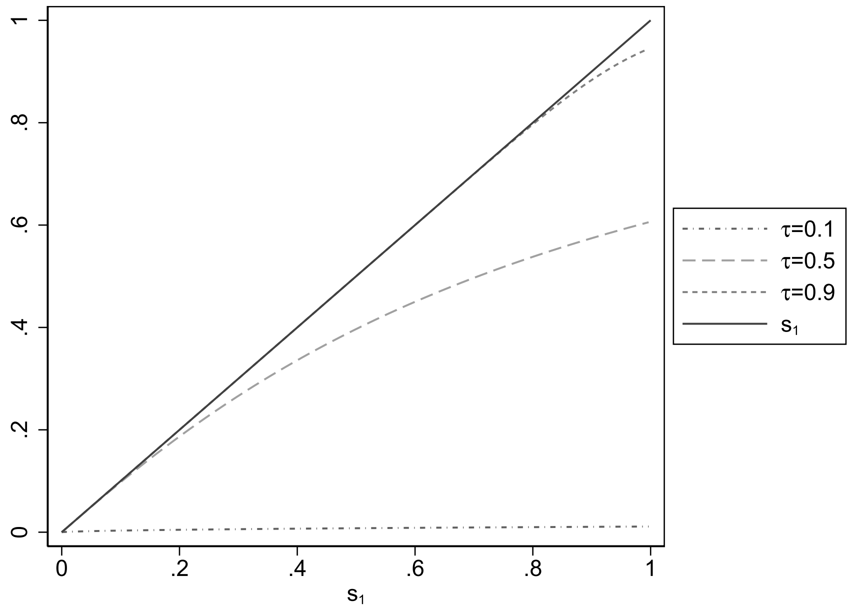


Fig. 4. Graphical illustration of the function $f(s_1) = s_1^{1/(1-s_1^\theta)}$ for three levels of dependency.

(i) $\frac{\partial \log \lambda_2(s_2|S_1 \le s_1)}{\partial \theta} < 0$ if $\theta = 0$ for all s_1 and s_2;
(ii) $\frac{\partial \log \lambda_2(s_2|S_1 \le s_1)}{\partial \theta} < 0$ if $s_2 > s_1$ for all s_1 and θ;
(iii) $\frac{\partial \log \lambda_2(s_2|S_1 \le s_1)}{\partial \theta} < 0$ if $s_2 \ge 0.01$ and $\tau \le 0.1$ for all s_1;
(iv) The interval $(s_1^{\frac{1}{1-s_1^\theta}}, s_1)$ is wide for many combinations of s_1 and θ.

In all, we show that $\partial \log \lambda_2(s_2|S_1 \le s_1)/\partial\theta < 0$ for most combinations of s_1, s_2, θ. Counterexamples include the trivial case $s_1 = 1$ or when τ increases and $s_2 < s_1$ as there are then more and more values of s_2 for which the sign is positive. □

Appendix B. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.csda.2024.108104.

References

Braekers, R., Veraverbeke, N., 2005. A copula-graphic estimator for the conditional survival function under dependent censoring. Can. J. Stat. 33, 429–447.
Carrasco, R., Garcia-Perez, J.I., 2015. Employment dynamics of immigrants versus natives: evidence from the boom bust period in Spain, 2000-2011. Econ. Inq. 53, 1038–1060.
Carrière, J.F., 1994. Dependent decrement theory. Trans. Soc. Actuar. 46, 45–74.
Carrière, J.F., 1995. Removing cancer when it is correlated with other causes of death. Biom. J. 37, 339–350.
Durante, F., Sempi, C., 2016. Principles of Copula Theory. CRC Press, Boca Raton, FL.
Eberwein, C., Ham, J.C., LaLonde, R.J., 1997. The impact of being offered and receiving classroom training on the employment histories of disadvantaged women: evidence from experimental data. Rev. Econ. Stud. 64 (4), 655–682.
Emura, T., Matsui, S., Rondeau, V., 2019. Survival Analysis with Correlated Endpoints Joint Frailty-Copula Models. JSS Research Series in Statistics. Springer, Singapore.
Emura, T., Shih, J., Ha, I.D., Wilke, R., 2020. Comparison of the marginal hazard model and the sub-distribution hazard model for competing risks under an assumed copula. Stat. Methods Med. Res. 29 (8), 2307–2327.
Fougère, D., Kamionka, T., 2008. Econometrics of individual labor market transitions. In: The Econometrics of Panel Data. Springer.
Goethals, K., Janssen, P., Duchateau, L., 2008. Frailty models and copulas: similarities and differences. J. Appl. Stat. 35 (9), 1071–1079.
Hasan, M.N., Braekers, R., 2021. Estimation of the association parameters in hierarchically clustered survival data by nested Archimedean copula functions. Comput. Stat. 36, 2755–2787.
Hofert, M., Mächle, M., 2011. Nested Archimedean copulas meet R: the nacopula package. J. Stat. Softw. 39 (9), 1–20.
Hofert, M., 2012. A stochastic representation and sampling algorithm for nested Archimedean copulas. J. Stat. Comput. Simul. 82 (9), 1239–1255.
Joe, H., 1997. Multivariate Models and Dependence Concepts. Monographs on Statistics and Applied Probability, vol. 37. Chapman & Hall, London, UK.
Lipowski, C., Lo, S.M.S., Shi, S., Wilke, R.A., 2021. Competing risks regression with dependent multiple spells: Monte Carlo evidence and an application to maternity leave. Jpn. J. Stat. Data Sci. 4, 953–981.
Lo, S.M.S., Wilke, R.A., 2010. A copula model for dependent competing risks. J. R. Stat. Soc., Ser. C, Appl. Stat. 59 (2), 359–372.
Lo, S.M.S., Stephan, G., Wilke, R.A., 2017. Competing risks copula models for unemployment duration: an application to a German Hartz-reform. J. Econom. Methods 6 (1), 1–20.
Lo, S.M.S., Mammen, E., Wilke, R.A., 2020. A nested copula duration model for competing risks with multiple spells. Comput. Stat. Data Anal. 150, 1–15.

Magnac, T., Robin, J.M., 1994. An econometric analysis of labour market transitions using discrete and tenure data. Labour Econ. 1 (3–4), 327–346.
Nelsen, R.B., 2006. An Introduction to Copulas, 2nd edition. Springer, New York.
Puzanova, N., 2011. A hierarchical Archimedean copula for portfolio credit risk modelling. Available at https://doi.org/10.2139/ssrn.2794067.
Rivest, L., Wells, M., 2001. A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. J. Multivar. Anal. 79, 138–155.
Ruppert, D., Matteson, D.S., 2015. Statistics and Data Analysis for Financial Engineering, 2nd ed. Springer Texts in Statistics. Springer, New York.
Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Stat. Univ. Paris 8, 229–231.
Statistics Denmark, 2019. Documentation of statistics for register-based labour force statistics 2019. Available at https://www.dst.dk/en/Statistik/dokumentation/documentationofstatistics/register-based-labour-force-statistics.
Sujica, A., Van Keilegom, I., 2018. The copula-graphic estimator in censored nonparametric location-scale regression models. Econom. Stat. 7, 89–114.
Trivedi, P.K., Zimmer, D.M., 2005. Copula modeling: an introduction for practicioners. Found Trends Econom. 1, 1–111.
Van den Berg, G.J., 2001. Duration models: specification, identification, and multiple duration. In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 5. North-Holland, Amsterdam, pp. 3381–3460.
Wang, Y.C., Emura, T., 2021. Multivariate failure time distributions derived from shared frailty and copulas. Jpn. J. Stat. Data Sci. 4, 1105–1131.
Wang, Y.C., Emura, T., Fan, T.H., Lo, S.M.S., Wilke, R.A., 2020. Likelihood-based inference for a frailty-copula model based on competing risks failure time data. Qual. Reliab. Eng. Int. 36 (5), 1622–1638.
Wooldridge, J., 2010. Cross Section and Panel Econometrics, 2nd edt. MIT Press.
Zheng, M., Klein, J.P., 1995. Estimates of marginal survival for dependent competing risks based on assumed copula. Biometrika 82 (1), 127–138.