A unique bond: Does losing your co-twin affect your remaining life-span?

Gerard J. van den Berg[*]
Bettina Drepper[†]

January 2012

Abstract

Twins share a unique bond that can lead to severe emotional stress and even health deterioration, once that bond is broken. In this paper we present new empirical evidence suggesting that the loss of the co-twin can shorten the remaining life-span of the surviving twin. In addition to this causal effect of bereavement, our model accounts for the influence of unobservable shared factors such as the genetic makeup and early childhood experiences that constitute a major source of the dependence between twin life-spans. Previous studies are limited to modeling exclusively one of these two channels of dependence. In this paper we present a new identification result on a symmetric version of the timing-of-events model of Abbring and Van den Berg (2003) which incorporates both of these dependencies and only requires limited covariate variation. Our empirical analysis is based on 9,268 twin pairs from the Danish Twin Registry. The estimated bereavement effect is decreasing in the age at the time when the loss is experienced and is more prominent in monozygotic twin pairs.

[*]Alexander von Humboldt Professor of Econometics and Empirical Economics, University of Mannheim; VU University Amsterdam; IFAU-Uppsala; IZA. Address: Department of Economics, University of Mannheim, L7, 3–5, 68131 Mannheim, Germany.

[†]University of Mannheim, L7, 3–5, 68131 Mannheim, Germany.

Acknowledgements: We acknowledge valuable comments and suggestions by Bo E. Honoré, Petyo Bonev, Georgios Effraimidis and Lena Janys. The Danish Twin Registry and Kaare Christensen kindly allowed us to use their data.

1 Introduction

The effects of losing a close member of the social network (bereavement) are studied in several fields of research. In particular, studies that look at bereavement among twins can be divided into two strands of literature. Psychologists and psychoanalysts are concerned with different psychological manifestations of bereavement and the different stages of grief (see Segal et al., 1995; Segal and Ream, 1998; Woodward, 1988). While this part of the literature is mainly focused on psychological effects, demographers, epidemiologists and human biologists are interested in developing a joint model of twin life-spans that correctly accounts for the complex dependence structure within twin pairs. Tomassini et al. (2002; 2001)[1] and Hougaard et al. (1992a) use survival models for each twin life-span and include the life-span of the co-twin as an exogenous time dependent covariate. Hougaard et al. point out the problem with this approach, i.e. their estimated effect does not only capture the effect of bereavement but also the dependence due to shared genetic factors. Other studies are focused on modeling this latter type of dependence which does not only reflect the influence of shared genetic factors but also the similarity of the environment in early childhood. The most elaborate approach in this field is based on a bivariate frailty model which specifies a flexible dependence structure between the frailty terms (see Yashin and Iachine, 1995a; Wienke et al., 2001; Hougaard et al., 1992a,b). In this paper we present a model which unites these two approaches of the twin mortality literature allowing us to estimate a causal bereavement effect while controlling for unobserved shared factors.

The twin studies by Segal et al. and Woodward document that the loss of a co-twin can cause severe emotional stress and that the grief intensity for an identical (monozygotic) twin is typically higher than that for other relatives or spouses (see Segal and Bouchard, 1993; Segal et al., 1995). Besides feelings of despair, depersonalization (numbness, shock), rumination (preoccupation with the deceased) and loss of control, bereaved twins also show symptoms like loss of appetite, loss of vigor and other physical symptoms (see Segal and Blozis, 2002). According to Selye's General Adaptation Syndrome, psychological stress is a major cause of disease because chronic stress causes long-term biochemical changes (see Selye, 1936, 1955).

[1] Tomassini et al. also use a model in which they match each bereaved twin to two not bereaved twins based on zygosity, age and sex and compare the two resulting hazard rates after the age when bereavement takes place. Note, that this method also ignores the endogeneity of the time of bereavement caused by shared genetic factors.

"Every stress leaves an indelible scar, and the organism pays for its survival after a stressful situation by becoming a little older." (Hans Selye)

Sanders (1999) integrates Selye's well-established general theory of stress in her Integrative Theory of Bereavement. She points out that besides the familiar stages of grief, patients also show physiological changes and consequent vulnerability to illness after bereavement (see Sanders, 1980).

The psychological manifestations of grief in general are well documented and twin studies have established the existence of a strong psychological reaction to the loss of the co-twin. Furthermore, the studies by Selye and Sanders show the existence of a direct link between emotional stress and health outcomes. However, no empirical study has clearly established the existence of a causal dependence between bereavement and mortality for the case of twins.

Previous studies are focused on the effects of conjugal bereavement (see Manor and Eisenbach, 2003; van den Berg et al., 2006; Bowling, 1987). These studies find convincing evidence that the loss of a spouse can severely affect mortality shortly after bereavement. However, it remains unclear wether the measured effect on mortality originates exclusively from emotional stress, since the loss of the spouse also greatly affects the everyday life of the surviving partner. In contrast to spouses, most twins have separate families and support systems. This suggests that a causal dependence between twin life-spans should be in large part attributed to emotional bereavement.

The first contribution to the twin bereavement literature focused on mortality outcomes was made by Hougaard et al. (1992a). They study the joint distribution of life times of twins, estimating two models. The first one is the Freund (1961) model, which uses a bivariate exponential distribution and as such can be interpreted as a simplified version of the Tomassini et al. models. The Freund model allows for a causal effect of the death of one twin on the hazard of the other. The second model does not include a bereavement effect but instead accounts for the influence of unobservable shared factors. This latter source of dependence is an important component in models of twin life-spans. Twins often experience a similar early childhood and while dizygotic twins share approximately 50% of their genetic material (Segal and Ream, 1998), monozygotic twins have an identical genetic makeup. Consequently, the identification of a causal bereavement effect among twins requires a model which simultaneously accounts for the effect of unobservable shared factors. Previous attempts in the twin literature such as the models of Hougaard et al. and Tomassini et al. are limited to modeling only

one of the two effects at a time but fail to unite both in one model.

In the econometric literature, dependence due to shared factors is often called a "selection effect" while the effect of an event on the hazard rate of another outcome is labeled a "causal effect". This terminology highlights the fact that, to empirically study the causal effect, the results may also reflect the other effect. Ignoring one type of dependence may lead to biased estimates of the other type. The studies by Tomassini et al. (2002) and Hougaard et al. (1992a) recognize that the different dependencies are suited to fit different aspects of the data, giving hope that the two effects can be distinguished empirically. In her discussion of Hougaard et al.'s paper, Flournoy (1992) argues that a super-model is needed which unites both models and incorporates both effects simultaneously. Related models have recently been used in empirical econometric studies on the effect of labor market programs on unemployment durations. Abbring and Van den Berg (2003a) demonstrate that bivariate versions of the mixed proportional hazard model can be identifies which take both dependencies into account.

In our application we use data on 9,268 twin pairs including mono- and dizygotic twins from the Danish Twin Registry. As our analysis exploits the timing of deaths, it is advantageous to observe as many exits as possible in the data. The Danish Twin Registry is one of the oldest existing twin datasets and allows us to use cohorts from 1870 to 1930, ensuring that 80.8% of these twins have died by 2004[2]. The drawback of using such old cohorts is the limited information available on the twin pairs. The Danish Twin Registry is designed as a medical dataset, providing very detailed information on the causes of death and exit dates on a daily basis but has only very limited information on other characteristics of the twins. In particular, the available covariates do not vary within same-sex twin pairs.[3] Unfortunately, the identification result of Abbring and Van den Berg relies on the assumption of sufficient covariate variation within the unit of interest, the twin pair in our case. Consequently, their identification result does not apply to our dataset. In Section 2 we present a new identification result for a symmetric[4] version of the timing-of-events model that does not rely on this assumption.[5]

Our semi-parametric identification result has a wider relevance for the empir-

[2]In 2004 our window of observation ends.

[3]The major part of the dataset comprises same-sex twin pairs, since less effort was put into following up on different-sex twins in the Danish Twin Registry.

[4]In contrast to the original timing-of-events model, the model we use here allows for treatment in both directions. Before the first exit occurs, both life-spans can potentially affect eachother.

[5]Note, that in our identification result we impose a multiplicative structure on the treatment effect function.

ical study of parallel systems and networks and for epidemiological research. The symmetric timing-of-events model describes a very general setting in which two parallel durations are connected due to observable as well as unobservable time-constant shared factors and at the same time the first exit potentially affects the survival of the other. In the most extreme case, the complete symmetry of our model allows for the two durations to be indistinguishable in terms of observable characteristics. So even if the durations can not be indexed (or the index 1,2 is completely uninformative) and the only observable covariates are characteristics of the duration pair (not individual characteristics), our identification result still applies. This result is relevant in cases of old datasets or for datasets which are limited for other reasons as well as in cases where the available covariates, creating the otherwise necessary variation within duration pairs, might be endogenous and therefore have to be excluded.

Our model allows to estimate a causal dependence between twin life-spans while controlling for the influence of shared genetic factors. However, the symmetric timing-of-events model can also be used to estimate the dependence between twin life-spans caused by shared genetic factors while controlling for a potential additional causal dependence. There is an extended field of research with the purpose to quantify the influence of genetic factors on mortality using data sets similar to ours. Our model allows to compare these different approaches and in our empirical analysis we show how the magnitude of the effects changes when excluding one of the two sources of dependence.

In Section 3 we shortly introduce the twin dataset from the Danish Twin Registry. For the purpose of our empirical analysis we impose some additional structure on the general symmetric timing-of-events model in Section 4 using functional forms that are well established in the twin mortality literature. Subsequently, our estimation results are presented in Section 5. We find a significant positive effect on the hazard rate of a bereaved adult twin that is decreasing in the age at which bereavement occurred. This effect is considerably stronger for monozygotic compared to dizygotic twin pairs. We sum up with a discussion of our results in Section 6 and a brief outlook in Section 7.

2 Model and identification result

In order to find a suitable model for the life-spans of twin pairs, we use a symmetric version of the timing-of-events model by (see Abbring and van den Berg, 2003a; Abbring and Heckman, 2007). This model can be expressed by the two individual hazard rates of durations T_1 and T_2 conditional on the observable vari-

ables x, the frailties V_1 and V_2 and the realized exit of the other duration.

Model A: *The hazard rates of* $T_1|(T_2 = t_2, x, V_1)$ *and* $T_2|(T_1 = t_1, x, V_2)$ *are given by*

$$\theta(t|T_2 = t_2, x, V_1) = \lambda(t)\phi(x)\delta(t, t_2, x)^{I(t>t_2)}V_1$$
$$\theta(t|T_1 = t_1, x, V_2) = \lambda(t)\phi(x)\delta(t, t_1, x)^{I(t>t_1)}V_2,$$

where the vector of frailties $V = (V_1, V_2)'$ *is assumed to be drawn from the bivariate distribution* $G(v_1, v_2)$ *and the bereavement effect function is multiplicative in two of its arguments* $\delta(t, t_k, x) = \delta_a(t - t_k)\delta_b(t_k, x)$.

Here, the first multiplicative term δ_a of the bereavement effect function describes the dependence of the bereavement effect on the time passed since the loss occurred and δ_b accounts for the dependence on the age at the time of bereavement as well as on observable variables x. $I(.)$ denotes an indicator function, which is 1 if its argument is true and 0 otherwise. The function $\lambda(t)$ models the duration dependence of the hazard rate and $\phi(x)$ holds the effect of the covariates.

Note, that given the observables x, Model A allows for two types of dependencies between duration T_1 and T_2. The first is reflected in the joint distribution of V_1 and V_2. In our application to twin life-spans this dependency is caused by unobserved shared factors such as the genetic makeup and early childhood experiences. The second type of dependence is incorporated via the bereavement effect function $\delta(t, t_k, x)$ that occurs as a multiplicative term in the hazard rate of duration j once duration k exits. Conditional on x and V, the variables T_1 and T_2 are only dependent through $\delta(t, t_k, x)$. Consequently, this factor can be given a causal interpretation as the effect of T_k on T_j. In contrast to the frailties V which reflect the influence of time constant unobserved characteristics, the bereavement effect accounts for the timing of events. Therefore, in Model A this effect can be seen as a local effect as it only affects the hazard rate of the surviving duration after the exit of the other occurs. In contrast, the influence of the unobservable factors V can be seen as a global effect as these factors influence the hazard rate of the two durations over the whole time interval $[0, \infty)$. It should also be pointed out that the setup in Model A rules out the existence of a third source of dependence in form of unobserved shared influences which vary over time. This also gives an intuition for the identifiability of Model A. By only allowing for one effect that is local (depends on the timing of events) and one that is global (reflects only time constant influences), both can be distinguished.

In contrast to Model A, in the original timing-of-events model of Abbring

and van den Berg (2003a) the functions λ and ϕ are allowed to differ across the two hazards and only the hazard of duration 1 can be directly affected by the exit of duration 2.[6] In their paper the authors already point out that their identification results can be straightforwardly extended to a setting in which the full distribution of $(T_1, T_2)|x$ is observable and both durations can potentially be treated by the exit of the other, similar to our setup. However, a different identification strategy is needed for the identification of Model A in which λ and in particular ϕ is the same in both hazards. The difficulty arises from this complete symmetry in the covariate effects $\phi(x)$. The latter implies that all covariates in the vector x enter both hazards with the same value and have the same effect.

The result by Abbring and van den Berg (2003a) uses the fact that until the first exit occurs, the two durations are competing risks. Therefore, their proof exploits an identification result of the mixed proportional hazard competing risk model (see Abbring and van den Berg, 2003b). In such a competing risks model, variation of the covariate effects across the two hazards is necessary in order to trace out the bivariate frailty distribution $G(v_1, v_2)$. In particular, the assumption is necessary that the function $(\phi_1(x), \phi_2(x))$ can attain all values over a nonempty open set $\Phi \subset (0, \infty)^2$ when x varies over $\mathcal{X}$.[7] Since in our symmetric setup it holds that $\phi_1(x) = \phi_2(x) = \phi(x)$, we cannot exploit this exogenous variation across the two hazards in our model.[8]

Although the original model of Abbring and van den Berg is more flexible compared to Model A, as it allows for different baseline hazards as well as different regression component functions across the two hazards, it is also more restrictive in the sense that it relies on sufficient variation of the covariate effects across the two durations. Therefore the symmetric case of Model A is not covered by their result. The main difference in terms of the identification strategy is the fact that while the original result of Abbring and van den Berg exploits the results from the mixed proportional hazard competing risk model, our identifica-

[6] The original model also does not need the assumption that the treatment effect can be separated into two multiplicative parts.

[7] If $\phi_j(x) = \mathrm{e}^{\beta_j^T x}$ then it would be sufficient that the vector x has two continuous covariates which affect the hazard rates of both risks but with different nonzero coefficients, and which are not perfectly collinear.

[8] For twins it is very unlikely that observable characteristics like sex or cohort will affect twin 1 systematically different compared to twin 2. In our dataset, twins are indexed according to their order of births. The firstborn has index 1 and the second index 2. But this information is extremely unreliable especially for the older cohorts. Note further, that since we use cohorts from 1870 onwards, we only have a very limited set of covariates available in our analysis none of which vary within same-sex twin pairs. Therefore, we can not rely on sufficient exogenous variation within twin pairs.

tion strategy relies on results from the simple mixed proportional hazard model. By imposing a multiplicative structure on the treatment effect function we are able to split the hazard rate into multiplicative parts reflecting the dependence on time t, observables x and unobservable influences V which is characteristic for a mixed proportional hazard model. We will exploit this structure at several steps throughout our proof.[9]

Assumption 1 *The vector x is k-dimensional with $1 \leq k < \infty$ and $\phi : \mathcal{X} \to U \subset (0, \infty)$. The set $\mathcal{X} \subset \mathbb{R}^k$ contains at least two values.*

Assumption 2 *$\delta_a : \mathbb{R}_+ \to (0, \infty)$ with $\lim_{s \downarrow 0} \delta_a(s) < \infty$ and for $\delta_b : [0, \infty) \times \mathcal{X} \to (0, \infty)$ it holds that $\nexists\ c \in (0, \infty)$ s.t. $\delta_b(0, x) = c\phi(x)^{-1}\ \forall x \in \mathcal{X}$.*

Assumption 3 *For the function $\lambda : [0, \infty) \to (0, \infty)$ it holds that for all $t \in (0, \infty)$ $\lim_{s \downarrow t} \lambda(s) < \infty$ and has integral*

$$\Lambda(t) := \int_0^t \lambda(\tau)\, d\tau < \infty, \quad \forall\ t \geq 0$$

and further

$$\tilde{\Lambda}(t, s) := \int_s^t \lambda(\tau)\delta_a(\tau - s)\, d\tau < \infty, \quad \forall\ \{(t, s) \in [0, \infty)^2 : t > s\}.$$

For some a priori chosen t_0, t_0^* and x_0, there holds that $\int_0^{t_0} \lambda(\tau)\, d\tau = 1$, $\int_0^{t_0^*} \lambda(\tau)\delta_a(\tau)\, d\tau = 1$ and $\phi(x_0) = 1$.

Assumption 4 *V is an $\mathbb{R}_+^2$-valued time-invariant random vector (V_1, V_2)' and is drawn from distribution G which does not depend on x and has a finite, positive mean. G is such that $P(V \in (0, \infty)^2) = 1$. Further, for all $(t, x) \in (0, \infty) \times \mathcal{X}$ $\lim_{s \downarrow t} E(V_j | T_j \geq s, T_k = t, x) = E(V_j | T_j \geq t, T_k = t, x)$.*

Assumption 5 *$\exists$ an open set $\Psi \in (0, \infty)^2$ with $t_1 > t_2\ \forall\ (t_1, t_2) \in \Psi$ s.t. at all points $(t_1, t_2) \in \Psi$ the function $\Delta(t_1, t_2, x) = \tilde{\Lambda}(t_1, t_2)\delta_b(t_2, x)$ is continuously*

[9]Note, that the identification results presented in this section can be straightforwardly extended to the case where the bereavement effect function differs between the two durations. So if the two spells can be distinguished in the data, it is possible to identify two separate bereavement effects $\delta_1(t, t_2, x)$ and $\delta_2(t, t_1, x)$. The first measures the effect of the exit of duration 1 on duration 2 and the other the effect of the exit of duration 2 on duration 1. However, in most applications including our twin model the effect of bereavement will be symmetric ($\delta_1(t, t_2, x) = \delta_2(t, t_1, x)$).

differentiable with respect to t_2.[10]

Note, that for Assumption 1 a single dummy variable x that does not need to vary across the two hazards suffices, provided that it has an effect. In that case $\phi(x)$ takes on only two values on $\mathcal{X}$. Recall, that in the case of the original timing-of-events model of Abbring and van den Berg (2003a) usually two continuous variables with different effects on the two hazards are needed in order to assure identification.[11] For our model the most limited case of covariate variation in form of a single dummy variable suffices.

Assumption 3 restricts the baseline hazard function to have existing limit $\lim_{s\downarrow t}\lambda(s) < \infty$ for all $t \in (0,\infty)$. Note, that this does not rule out the piecewise constant case as well as most functional forms. Further, since this property only has to hold for strictly positive values, this does not rule out forms with $\lim_{s\downarrow 0}\lambda(s) = \infty$ like the Weibull function. However, the initial jump of the bereavement effect has to have a finite limit. Consequently, this rules out functional forms for δ_a with $\lim_{s\downarrow 0}\delta_a(s) = \infty$. This leads to the following propositions:

Proposition 1 *If Assumptions 1-4 are satisfied, then the functions* $\lambda, \phi, \delta_a, \delta_b$ *from Model A are non-parametrically identified (up to a scaling factor) from the distribution of* $(T_1, T_2)|x$.

Note, that the identification of the function G is not included in Proposition 1. This leads to Proposition 2:

Proposition 2 *If Assumptions 1-5 are satisfied, then Model A which is characterized by the functions* $G, \lambda, \phi, \delta_a, \delta_b$ *is non-parametrically identified (up to a scaling factor) from the distribution of* $(T_1, T_2)|x$.

Proof of Proposition 1

Identification of λ *and* ϕ: Let Z be the minimum of the two durations T_1 and T_2. The survival function of $Z|x$ is given as (see Appendix A.1 for details)

$$S_Z(t|x) = \int_0^\infty \mathrm{e}^{-\Lambda(t)\phi(x)w}\,\mathrm{d}G_W(w), \text{ with } W = V_1 + V_2. \tag{1}$$

[10] Alternative assumption 5: The open set $\Psi \in (0,\infty)^2$ could also exist for $t_1 < t_2\ \forall\,(t_1,t_2) \in \Psi$ s.t. at all points $(t_1,t_2) \in \Psi$ the function $\Delta(t_2,t_1,x)$ is continuously differentiable with respect to t_1.

[11] The requirements for the vector of covariates x depends on the functional form assumptions for $\phi_1(x)$ and $\phi_2(x)$.

Equation 1 shows that due to the symmetry of Model A, the distribution of Z has a hazard rate of the mixed proportional form: $\theta_z(t|x,W) = \lambda(t)\phi(x)W$ with frailty $W = V_1 + V_2$ drawn from distribution G_W. Intuitively, the hazard of the first exit given x and W is the sum of the hazards of the two competing risks: $\theta_z(t|x,W) = \theta(t|T_2 \geq t, x, V_1) + \theta(t|T_1 \geq t, x, V_2) = \lambda(t)\phi(x)(V_1 + V_2)$. Note, that conditional on x and W the two competing risks of twin 1 and twin 2 to die at time t, given that the co-twin survives this age, are independent events. The results by Elbers and Ridder (1982), (also see Lancaster, 1990; van den Berg, 2001, for an overview)[12] on the identification of the mixed proportional hazard model imply that under Assumptions 1-4, the model in equation (1) characterized by the functions λ, ϕ and G_W is identified up to a scaling factor (see A.1 for details).

Identification of δ_a: The survival function of duration T_j given x and given that the exit of the other duration occurred at $T_k = 0$ can be expressed as follows

$$S(t|T_k = 0, x) = \int_0^\infty \mathrm{e}^{-\int_0^t \theta(\tau|T_k=0,x,V_j)\,d\tau}\,\mathrm{d}G_{V_j|T_k=0,x}(v_j),$$

where $\theta(t|T_k = 0, x, V_j) = \lambda(t)\phi(x)\delta_a(t-0)\delta_b(0,x)V_j$. Here, we make use of the subset $T_j|(T_k = 0, x)$ of the observable bivariate distribution $(T_1, T_2)|x$. One of the durations exits at time $T_k = 0$ and therefore the hazard of the other duration is affected by bereavement over the full interval $(0, \infty)$. Due to the multiplicative structure of the bereavement effect function, the distribution of $T_j|(T_k = 0, x)$ has a hazard rate of the mixed proportional form: $\theta(t|T_k = 0, x, V_j) = \tilde{\lambda}(t)\tilde{\phi}(x)V_j$ with $\tilde{\lambda}(t) = \lambda(t)\delta_a(t)$ and $\tilde{\phi}(x) = \phi(x)\delta_b(0,x)$. Again the results by Elbers and Ridder imply that under Assumptions 1-4 the mixed proportional hazard model defined by $\{\tilde{\lambda}, \tilde{\phi}, G_{V_j|T_k=0,x}\}$ is identified up to a scaling factor and since λ is known, this also identifies δ_a. Note, that one central assumption for the identifiability of a mixed proportional hazard model is the independence of observable variables and unobservable frailties. In Appendix A.2 we show that under Assumptions 1-4 the conditional frailty distribution $G_{V_j|T_k=0,x}$ does not depend on x. Further, Assumption 2 states that the functions $\phi(x)$ and $\delta_b(0,x)$ are not proportional assuring that the function $\hat{\phi}(x) = \phi(x)\delta_b(0,x)$ generates sufficient exogenous variation.

Identification of δ_b: In the following, we exploit information on the jump of the

[12] Also see Kortram et al. (1995) for the case of only two possible values for $\phi(x)$.

hazard rate at the moment of bereavement

$$\frac{\lim_{s\downarrow t}\theta(s|T_k=t,x)}{\theta(t|T_k=t,x)} = \frac{\phi(x)\delta_b(t,x)\lim_{s\downarrow t}\delta_a(s-t)\lambda(s)E(V_j|T_j\geq s,T_k=t,x)}{\phi(x)\lambda(t)E(V_j|T_j\geq t,T_k=t,x)}$$
$$= \delta_b(t,x)\lim_{s\downarrow t}\delta_a(s-t)\frac{\lim_{s\downarrow t}\lambda(s)}{\lambda(t)}. \quad (2)$$

Assumptions 2 and 3 assure the existence of $\lim_{s\downarrow t}\delta_a(s-t)$ and $\lim_{s\downarrow t}\lambda(s)$. With this, the second equality directly follows from Assumption 4 stating that $\lim_{s\downarrow t}E(V_j|T_j\geq s,T_k=t,x)=E(V_j|T_j\geq t,T_k=t,x)$. Note, that the left hand side of equation 2 is observable for all $(t,x)\in(0,\infty)\times\mathcal{X}$. Since $\lim_{s\downarrow t}\delta_a(s-t)$, $\lim_{s\downarrow t}\lambda(s)$ and $\lambda(t)$ are known from previous steps, we can trace out the function $\delta_b(t,x)$ over $(0,\infty)\times\mathcal{X}$.[13]

Proof of Proposition 2

Identification of G: Recall that the functions $\lambda,\phi,\delta_a,\delta_b$ in Model A are identified under Assumptions 1-4. The only function that remains unknown is the bivariate frailty distribution $G(v_1,v_2)$. The density $f(t_1,t_2|x)$ for $t_1>t_2$ can be expressed as follows (see Appendix A.3.1)

$$f(t_1,t_2|x)=c(t_1,t_2,x)\partial^2_{s_1,s_2}\mathcal{L}_G\big(\phi(x)(\Lambda(t_2)+\Delta(t_1,t_2,x)),\phi(x)\Lambda(t_2)\big), \quad (3)$$

with $c(t_1,t_2,x)=\lambda(t_1)\lambda(t_2)\phi(x)^2\delta_a(t_1-t_2)\delta_b(t_2,x)$ and $\Delta(t_1,t_2,x)=\tilde{\Lambda}(t_1,t_2)\delta_b(t_2,x)$. Note, that all functions on the right hand side of equation 3 are identified except the cross derivative of the bivariate Laplace transformation $\partial^2_{s_1,s_2}\mathcal{L}_G(s_1,s_2)$, with arguments $s_1=\phi(x)(\Lambda(t_2)+\Delta(t_1,t_2,x))$ and $s_2=\phi(x)\Lambda(t_2)$. The Laplace transformation $\mathcal{L}_G(s_1,s_2)$ is known to be a completely monotone function. This property implies that it's cross derivative $\partial^2_{s_1,s_2}\mathcal{L}_G(s_1,s_2)$ is also completely monotone (see Appendix A.3.2). Since completely monotone functions are real analytic and real analytic functions are uniquely determined by their values on a nonempty open set, the function $\partial^2_{s_1,s_2}\mathcal{L}_G(s_1,s_2)$ can be identified on its whole support $S=[0,\infty)^2$ if we know all it's values on a nonempty open set. In Appendix A.3.3 we show that under Assumption 5 the function $(\phi(x)(\Lambda(t_2)+\Delta(t_1,t_2,x)),\phi(x)\Lambda(t_2))$ attains all values on a nonempty open set $\Upsilon\subset(0,\infty)^2$ when t_1 and t_2 vary over $\Psi\subset(0,\infty)^2$ with $t_1>t_2$.[14] This identifies $\partial^2_{s_1,s_2}\mathcal{L}_G(s_1,s_2)$ and the integral

[13]Note, that $\delta_b(0,x)$ is already known from the last identification step.

[14]Note, that if $t_1<t_2\ \forall\ (t_1,t_2)\in\Psi$ then the same reasoning can be applied to the function $(\phi(x)\Lambda(t_1),\phi(x)(\Lambda(t_1)+\Delta(t_2,t_1,x)))$, which then holds the arguments of $\partial^2_{s_1,s_2}\mathcal{L}_G$ in the equation above for the case $t_1<t_2$.

$\int_0^{s1}\int_0^{s2}\partial^2_{s_1,s_2}\mathcal{L}_G(u_1,u_2)\,\mathrm{d}u_1\mathrm{d}u_2$ then gives us $\mathcal{L}_G$. Due to the uniqueness of the Laplace transformation this also determines $G(v_1,v_2)$.

3 Data

In our empirical analysis we use data from the Danish Twin Registry. When the registry was first established in 1954, the goal was to follow up on all same-sex twins born since 1870 and surviving as twins until at least the age of 6. However, there is some selectivity in the very early cohorts. Twins who died young are less likely to be included in the sample. Further, most of the information on characteristics is only available for twins who survived January 1st 1943. Therefore, we restrict attention to twin pairs still alive at this date. This is not a serious limitation since we are particularly interested in the effects of bereavement at higher ages. We use cohorts from 1870 to 1930, assuring that we observe the exit of most twins before January 1st 2004 when our window of observation ends. The registry contains some different-sex twin pairs, but most effort was put into following up on same-sex and in particular monozygotic twin pairs. We refer to Skytthe et al. (2002) for detailed descriptions of the registry and the way it has been collected.

As a result, our sample includes 2,806 monozygotic and 6,462 dizygotic twin pairs of which 1,219 are different sex twin pairs. All twins are born between 1873 and 1930 and in all pairs both twins have survived until at least January 1st 1943. The birth and death dates and the resulting individual lifetime durations are observed in days. Individuals still alive on January 1st 2004 or who have emigrated have right-censored durations. Overall the death date is observed for 80% of the individuals in our sample. For 94.4% of this group we observe the death cause. The death cause is classified according to the International Classification of Diseases(ICD) system, versions 5-8, at the 3-digit level. These are grouped into 12 categories, of which the following groups are of specific interest: 'cardiovascular' (32.42%, death due to cardiovascular malfunctions or diseases), 'apoplexy' (14.13%), 'cancer' (26.03%, death due to malignant neoplasms), 'suicide' (1.03%), 'accidents' (3.7%) and 'other' (including death due to tuberculosis, other infectious diseases, diseases of the respiratory, digestive or uro-genital system).

For each twin pair in our sample we observe zygosity, sex, year of birth, season of birth and region of birth. Note, that except sex, none of the available covariates vary within the twin pair. In the previous section we showed that our model does not rely on this kind of variation. The information on zygosity is very accurate

with a misclassification rate below 5% (see Holm, 1983; Lykken, 1978). We use an indicator for being born in Copenhagen in our analysis to distinguish between rural and urban areas in Denmark. The additional distinctions between small towns and rural areas outside of Copenhagen turned out to be uninformative in our empirical analysis.

Besides having one of the oldest existing twin datasets in the world, the country of Denmark is especially suited for mortality studies using individual lifetime data over a long time interval. At the beginning of our window of observation in the 1870ties, Denmark compared to the rest of Europe already had a quite well established/comprehensive health care system. This is of particular importance for our purposes as a functioning health care system dampens economic shocks which twin pairs are exposed to over their whole life. There were also no major epidemics between 1870 and 2004 in Denmark. Recent studies have compared international mortality levels for 1918 and found that for the 1918–1919 worldwide influenza pandemic Denmark stands out as the country with the lowest levels of excess mortality (see Canudas-Romo and Erlangsen, 2008; Ansart et al., 2009). Furthermore, Denmark remained neutral in both World Wars and although it was occupied by Germany during the Second World War, casualties were negligible compared to the rest of Europe. In summary, lifetime data from Denmark from the 1870ties up to today provides a dataset that is little affected by economic or direct health shocks compared to the rest of Europe.

4 Model of twin life-spans

For the estimation of our twin model we impose additional structure on the symmetric timing-of-events model (Model A) introduced in Section 2. In particular, for each twin pair the vector of frailties (V_1, V_2) is assumed to be drawn from a Cherian bivariate Gamma distribution. This distribution is often used in lifetime models for twins (see Yashin and Iachine, 1995b; Wienke et al., 2001, 2002) and allows for an interpretation of the individual frailty term as the sum of a shared twin pair-specific term $\tilde{V}_0$ and an individual-specific term $\tilde{V}_1$

$$V_j = \tilde{V}_0 + \tilde{V}_j \ \textit{for}\ j \in 1, 2.$$

Here, each term $\tilde{V}_1, \tilde{V}_2$ and $\tilde{V}_0$ is independently drawn from a Gamma distribution. With this structure, the bivariate Gamma distribution of (V_1, V_2) has identical marginal distributions which reflects the symmetry of life-spans within twin pairs. Their mean is normalized to one and as a result, the joint distribution of (V_1, V_2) can be fully described by two parameters: the variance σ^2 of V_j and correlation

ρ of V_1 and V_2. The latter equals the fraction of the total variance of V_j that the two twins have in common $\rho = \frac{Var(\tilde{V}_0)}{Var(\tilde{V}_0+\tilde{V}_j)}$. Recall, that our sample includes monozygotic (MZ) as well as dizygotic (DZ) twin pairs. We estimate separate parameters for both types of zygosity: $\sigma^2_{MZ},\ \rho_{MZ}$ and $\sigma^2_{DZ},\ \rho_{DZ}$.

In the following we denote the two life-spans of each twin pair by the vector of random variables (T_1, T_2) and their realizations by (t_1, t_2). T_1 and T_2 are assumed to be independently censored from the right.[15] The twin life-spans follow a distribution given by the bivariate survival function $S(t_1, t_2|x) = P(T_1 > t_1, T_2 > t_2|x)$. Note, that for all twin pairs for which the minimum of T_1 and T_2 is censored, we do not observe the exact time of bereavement. The first exit could occur any time within the interval $(min\{t_1, t_2\}, \infty)$. Consequently, we have to integrate over the respective interval to account for the occurrence of all possible exit times. This leads to a survival function of the form[16]

$$S(t_1, t_2|x) = \begin{cases} S^*(t_1, t_1|x) - \int_{t_2}^{t_1} S_{t_2}(t_1, \tau|x)\, d\tau & \text{, for } t_1 \geq t_2 \\ S^*(t_2, t_2|x) - \int_{t_1}^{t_2} S_{t_1}(\tau, t_2|x)\, d\tau & \text{, for } t_1 < t_2 \end{cases}$$

$$\begin{aligned} \textit{with} \quad S^*(t_1, t_2|x) = &(1 + \sigma^2\phi(x)[\Lambda(t_1) + \Lambda(t_2)])^{-\frac{\rho}{\sigma^2}} \\ &(1 + \sigma^2\phi(x)\Lambda(t_1))^{-\frac{(1-\rho)}{\sigma^2}} (1 + \sigma^2\phi(x)\Lambda(t_2))^{-\frac{(1-\rho)}{\sigma^2}} \end{aligned}$$

and partial derivatives $S_{t_j}(t_1, t_2|x) = \frac{\partial S(t_1,t_2|x)}{\partial t_j}$, for $(j = 1, 2)$.

In the data section it was already mentioned that our dataset only includes twin pairs for which both twins are still alive on January 1st 1943. This truncation of data has to be reflected in the likelihood function as well. For this purpose we denote the age of twin j on January 1st 1943 by $t_{j,age43}$. This leads to the survival function conditional on both twins surviving January 1st 1943

$$S(t_1, t_2|T_1 > t_{1,age43}, T_2 > t_{2,age43}, x) = S(t_1, t_2|x)S(t_{1,age43}, t_{2,age43}|x)^{-1}$$

[15] Recall that for a small part of the twin pairs in our sample at least one twin is still alive at the end of the observation window on 1st January 2004. However, the censoring durations only depend on the cohort of the twin pair and not on the life-spans of the two twins. This assures independent censoring in our data.

[16] The integrals $\int_{t_2}^{t_1} S_{t_2}(t_1, \tau|x)\, d\tau$ and $\int_{t_1}^{t_2} S_{t_1}(\tau, t_2|x)\, d\tau$ are approximated with numerical integration methods.

With this we can derive the likelihood contribution of a twin pair

$$\begin{aligned} L(t_1, t_2, c_1, c_2|x) = [\ & c_1c_2S(t_1, t_2|x) - c_1(1-c_2)S_{t_2}(t_1, t_2|x) \\ & -(1-c_1)c_2S_{t_1}(t_1, t_2|x) + (1-c_1)(1-c_2)S_{t_1,t_2}(t_1, t_2|x)\ ] \\ & S(t_{1,age43}, t_{2,age43}|x)^{-1}. \end{aligned}$$

Here, c_1 and c_2 denote the censoring indicators for T_1 and T_2 and $S_{t_1,t_2}(t_1, t_2|x) = \frac{\partial^2 S(t_1,t_2|x)}{\partial t_1 \partial t_2}$. Note, that due to the specific functional form of the Cherian bivariate Gamma distribution the likelihood function has a closed form[17]. The functional forms of S, S_{t_1}, S_{t_2} and S_{t_1,t_2} and their derivation is presented in Appendix B.

For the purpose of our empirical analysis we also impose additional structure on the functions $\phi, \lambda, \delta_a, \delta_b$ of Model A. The logarithm of $\delta_a(t)$ is specified as piecewise constant with three time intervals after bereavement occurred: first year after bereavement, second to fourth and after the fourth year. The function $ln(\delta_b(t_k, x)) = \delta_l^{age} + \beta_\delta^\intercal x$ models the dependence of the bereavement effect on the twin's age at the time t_k she/he experiences the loss of the co-twin and the dependence on observable characteristics x such as sex and zygosity. Here, p indicates the age interval in which the loss occurred : ages below 65, 66 to 79 and above 80. The covariate effects enter the hazard as $\mathrm{e}^{\beta^\intercal x}$, which is the standard choice in mixed proportional hazard models and the duration dependence function λ is assumed to follow a flexible version of the Gompertz function, i.e. $\lambda(t) = e^{\alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3}$.

We choose a flexible baseline hazard which includes the Gompertz function as a special case for $\alpha_2 = \alpha_3 = 0$ which is often used in mortality models and is known to give an acceptable fit. Specifying the correct functional form for the baseline hazard is of particular importance in our analysis in which we are measuring the impact of intermediate events in a lifetime. If for instance the baseline hazard function were specified to be too restrictive in terms of the slope at higher ages this lack of flexibility would be reflected in the causal bereavement effect which in most cases occurs at higher ages. Note, that we use a very wide range of cohorts 1870 to 1930 in our analysis. For these cohorts, the aging process has evidently changed over time. In particular, the life expectancy at higher ages has increased drastically between 1870 and 1930 (see Gavrilov and Nosov, 1985). In order to account for this change in the shape of the duration dependence function we estimate separate sets of parameters α_{c1}, α_{c2} and α_{c3} for three different cohort groups $c \in \{1, 2, 3\}$: 1873 - 1899, 1900 - 1915 and 1916 - 1930.

[17]The only exception are the integrals over the interval of all possible bereavement times for censored twin pairs.

With this structure we can express our model in terms of the logarithm of the hazard rates of twin 1 and 2 conditional on observable and unobservable variables x and V and the realization of the other duration

$$\begin{aligned}\ln\theta(t|T_2 = t_2, x, V_1) &= \alpha_{c1}t + \alpha_{c2}t^2 + \alpha_{c3}t^3 + \beta^{'}x + I(t > t_2)(\delta_q^t + \delta_p^{age} + x^{'}\delta^x) + \ln(V_1)\\ \ln\theta(t|T_1 = t_1, x, V_2) &= \alpha_{c1}t + \alpha_{c2}t^2 + \alpha_{c3}t^3 + \beta^{'}x + I(t > t_1)(\delta_q^t + \delta_p^{age} + x^{'}\delta^x) + \ln(V_2).\end{aligned} \tag{4}$$

Here, $\delta_q^t, \delta_p^{age}, \delta^x$ are parameters that model the effect of bereavement. The indicator for the three time intervals after bereavement is denoted by $q = 1, 2, 3$ and $p = 1, 2$ is the indicator for the three age groups at which bereavement occurs, ages below 65 being the reference group.

5 Empirical Analysis

5.1 Residual Life Expectancies

One advantage of modeling twin life-spans on the individual level is the possibility to make predictions about residual life expectancies depending on when bereavement is experienced. Expected residual lifetimes are relevant for health care policy and are frequently calculated in the demographic and gerontological literature. The expected residual lifetime at age s are computed as (see Lancaster, 1990)

$$E(s) = \frac{\int_s^\infty S(t|x)\,dt}{S(s|x)}.$$

In Tables 2a and 2b the residual life expectancies for male, female, monozygotic and dizygotic twins implied by the estimates of Model IV (Table 1) are presented. A male monozygotic twin who has reached the age of 65 and has lost his co-twin at the age of 60 will life on average 11.22 remaining years. If he never had to experience this loss he would live on average 2 years longer. A very similar pattern is observed for female twins. Since the dependence of the bereavement effect on sex was insignificant we set this effect to zero in our calculations in Tables 2a and 2b.

5.2 Estimation Results

In our empirical analysis we estimate four different bivariate survival models (Table 1: Models I-IV). The different models allow the comparison of our approach to previous models used in the twin mortality literature.

Model II is a correlated frailty model that does not include a bereavement effect. It represents the strand in the epidemiological literature that models the influence of shared genetic factors by allowing for a dependence between frailty terms (see Yashin and Iachine, 1995a; Wienke et al., 2001). Note, that in these models a potential causal dependence between twin life-spans is ignored. In Model (4) this corresponds to the case of $\delta_a = \delta^{age} = \delta^x = 0$.

Model I on the other hand is a bivariate survival model in which the only dependence between twin life-spans is modeled via a bereavement effect. In fact, it does not allow for any frailty distribution ($\sigma^2 = 0$ in Model (4)). It represents the approach in the twin bereavement literature where bereavement is modeled as an exogenous event ignoring the influence of shared genetic factors (see Hougaard et al., 1992a; Tomassini et al., 2002).

Finally, Model III is an application of the Symmetric Timing-of-Events Model to twin life-spans which accounts for both, the influence of shared genetic factors as well as a causal dependence between twin life-spans (Model (4) with $\delta^{age} = \delta^x = 0$ [18]).

In Model IV we include a more flexible bereavement effect function allowing besides zygosity for a dependence on sex and the age at bereavement (Model (4)).

When comparing the estimates of the Correlated Gamma frailty distribution in Model II to the ones from Model III, one finds considerably higher estimates of the variance and the correlation parameters in Model II. This is true for the frailty distribution of monozygotic (σ^2_{MZ}, ρ_{MZ}) as well as for the one of dizygotic (σ^2_{DZ}, ρ_{DZ}) twin pairs. Especially the correlation between frailties which reflects the influence of shared genetic factors decreases strongly (around 30%) when including the bereavement effect in Model III. It is clear from this comparison that the estimated correlation in Model II does not only reflect the time-invariant influence of shared genetic factors but also captures some time-varying influences such as a causal dependence between twin life-spans.

In Model I we find relatively high estimates for the bereavement effect. These estimates would imply that a monozygotic male twin who is 75 years old and has lost his co-twin at the age of 70 would die on average 2.2 years earlier compared to when he would never have to experience this loss. These high estimates are not surprising since they do not only capture a bereavement effect but also the influence of shared genetic factors. In Model III and IV we control for this influence and find considerably lower estimates (28% less in terms of residual life expectancy in Model IV). This illustrates how strongly the estimates of the bereavement effect are biased in the presence of unobserved shared genetic fac-

[18] $\delta^x = 0$ The only exception is the dependence of the bereavement effect on zygosity.

tors when the model fails to control for them. Considering these results, it also becomes clear to what extend previous empirical studies have overestimated a bereavement effect for twins.

Note, that we do not report the estimated parameters of the baseline hazard function in Table 1. In Figure 1 the function $\lambda(t) = e^{\alpha_{c1}t+\alpha_{c2}t^2+\alpha_{c3}t^3}$ is plotted over the age interval 0 - 120 for the three cohort groups $c = 1, 2, 3$ implied by the estimated parameters in Model IV. Evidently, younger cohorts have a considerably lower mortality hazard at higher ages compared to the older cohorts. This change in the aging process over time is known as the late-life mortality deceleration (see Gavrilov and Nosov, 1985).

In Model IV, we find a highly significant positive effect of being male (0.513) reflecting the shorter life expectancy for males compared to females. When comparing monozygotic male twins to monozygotic female twins this estimate implies a higher residual life expectancy of 2.38 years for females at the age of 75 (see Tables 2a and 2b). Being born in spring has a weakly significant positive effect on the mortality hazard which is in line with the findings of Doblhammer (2004). If a twin is born in Copenhagen this increases mortality considerably (21.4% of the effect of being male). This could be due to a greater exposition to diseases, pollution or other risk factors in urban areas. Note, that eventhough dizygotic twins are known to live slightly longer than monozygotic twins, we find a positive effect on mortality for dizygotic twins compared to monozygotic twins. However, in studies which restirct attention to twins who suvived infancy, this result is not surprising. Identical twins face a higher infant mortality risk compared to fraternal twins leading to a selective sample which overrepresents healthy identical twins .

We estimate a piecewise constant bereavement effect in Model IV, accounting for three different time intervals after bereavement occurred: the fist year after the loss, second to fourth year and after four years. The overall positive effect is highly significant and slightly decreases over time. Further, the size of the bereavement effect strongly depends on zygosity (-.23) but not on the sex of the twin. The size of the effect is almost twice the size for monozygotic compared to dizygotic twins. This large difference is in line with the findings from psychological studies (see Segal and Bouchard, 1993; Segal et al., 1995). They conduct studies with bereaved twins and construct measures of grief intensities for monzygotic and dizygotic twins. Overall, they document grief intensities of monozygotic twins which are twice as large as the grief intensities observed for dizygotic twins. In Model IV the bereavement effect function also depends on the age at bereavement. We distinguish the ages before 65 and above 80, while ages 66 to 79 constitute

the reference group. Evidently, there is a decrease of the effect of bereavement in the age at which the loss occurrs. In particular, the effect of losing your co-twin after the age of 80 is relatively small, with a decrease in residual life expectancy of 0.58 years (for age 85, monozygotic males, see Table 2b).

Table 1: ***Estimation Results:*** *Four Bivariate Survival Models*

	(Model I)		(Model II)		(Model III)		(Model IV)	
	Bivariate Model with BE		**Corr. Frailty Model no BE**		**Corr. Frailty Model with BE**		**Corr. Frailty Model with BE (ext.)**	
Variable	Estimate	St.Error	Estimate	St.Error	Estimate	St.Error	Estimate	St.Error
Covariates:								
male	.3982***	(.0167)	.6006***	(.0291)	.5318***	(.0299)	.513***	(.0301)
log(birth year)	-.1065***	(.0236)	-.2441***	(.0505)	-.1845***	(.0402)	-.1837***	(.0391)
spring	.0281	(.0186)	.0531*	(.0285)	.0428*	(.0248)	.0415*	(.0243)
Copenhagen	.0909***	(.0243)	.1313***	(.0371)	.1165***	(.0324)	.1144***	(.0318)
dizygotic	.1553***	(.0243)	.0472	(.0343)	.1103***	(.0352)	.1096***	(.0347)
Bereavement effect:								
first year	.5209***	(.0546)	-	-	.3838***	(.0779)	.4203***	(.09)
second to fourth year	.5224***	(.04)	-	-	.3816***	(.0714)	.409***	(.0844)
after four years	.5534***	(.0333)	-	-	.4037***	(.0774)	.3893***	(.0903)
dizygotic	-.2807***	(.0364)	-	-	-.24***	(.0752)	-.23***	(.0793)
male	-	-	-	-	-	-	.0353	(.0377)
age at ber. below 65	-	-	-	-	-	-	.1171***	(.0386)
age at ber. above 80	-	-	-	-	-	-	-.1253***	(.0503)
Corr. Gamma frailty:								
variance monozygotic	-	-	.6055***	(.1193)	.4059***	(.183)	.3894***	(.1957)
dizygotic	-	-	.5681***	(.1225)	.3474***	(.202)	.3424***	(.2129)
correlation monozygotic	-	-	.8676	(-)	.584	(-)	.5083	(-)
dizygotic	-	-	.4483	(-)	.3183	(-)	.1585	(-)

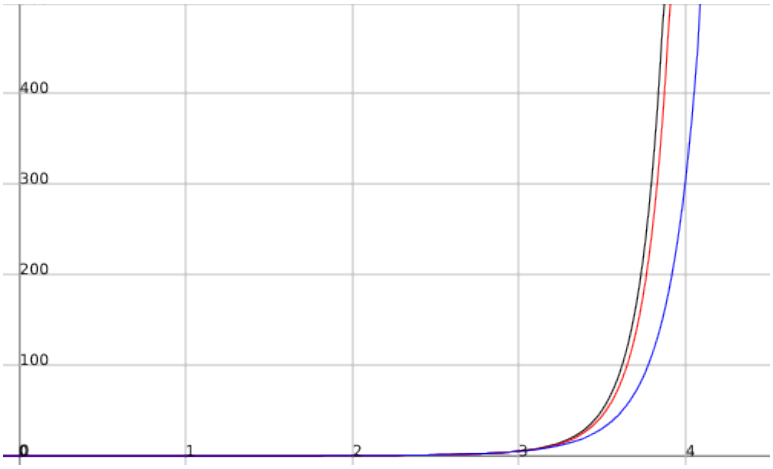

Figure 1: Plot of baseline hazard functions of Model IV

black: cohort group 1873 - 1899
red: cohort group 1900 - 1915
blue: cohort group 1916 - 1930

Table 2a: ***Residual Life Expectancies*** *(in years)*

Monozygotic Males

Age	No Bereav.	Experienced Bereav. at age			
		60	70	80	90
65	14.22	11.22	0	0	0
75	7.85	5.82	6.26	0	0
85	3.31	2.27	2.48	2.73	0
95	.91	.58	.65	.72	.72

Table 2b: ***Residual Life Expectancies*** *(in years)*

Monozygotic Females

Age	No Bereav.	Experienced Bereav. at age			
		60	70	80	90
65	17.48	14.26	0	0	0
75	10.23	7.88	8.4	0	0
85	4.67	3.32	3.61	3.93	0
95	1.41	.92	1.01	1.13	1.13

Dizygotic Males

Age	No Bereav.	Experienced Bereav. at age			
		60	70	80	90
65	13.55	11.91	0	0	0
75	7.38	6.27	6.73	0	0
85	3.06	2.49	2.72	2.98	0
95	.83	.65	.72	.8	.8

Dizygotic Females

Age	No Bereav.	Experienced Bereav. at age			
		60	70	80	90
65	16.77	15.01	0	0	0
75	9.7	8.41	8.95	0	0
85	4.35	3.62	3.92	4.26	0
95	1.29	1.02	1.12	1.25	1.25

6 Discussion

The structure of the symmetric timing-of-events model (Model A in Section 2) imposes some implicit assumptions on the underlying process generating the pairs of twin life-spans. Since identification of the model exploits the timing of the loss, a key assumption in the case of our twin model is that the event of losing your co-twin at some age t does not affect your own mortality hazard prior to that date. In the duration literature this assumption is called 'No Anticipation'. In our application to bereavement among twins this terminology can be misleading. In the case of some diseases, a twin will learn about the increased risk of dying of his co-twin when he is diagnosed with a severe illness before the actual loss occurs. However, this only constitutes a problem in terms of our model if his own mortality hazard reacts prior to the loss. Even if some of the psychological symptoms of grief may already manifest at an early stage when the co-twin is diagnosed, the actual event of bereavement only takes place when the other person is suddenly not part of the bereaved twin's life anymore. The exact timing of this loss is in most cases not anticipated. Nevertheless, in the case of some severe longterm illnesses the process of bereavement might to some degree already take place during the last stage of illness and the additional effect of loss will be small. In this case, our model would underestimated the true bereavement effect.[19] In light of this, one should interpret our estimated effect as the effect of actual bereavement, meaning the effect of physically losing the co-twin.

In the symmetric timing-of-events model all unobserved shared factors causing a dependence between the two life-spans of the twin pair are assumed to be time-invariant influences. This means that our model accounts for all unobserved shared factors such as the genetic makeup or early childhood experiences as long as their influence on the mortality hazard is time-invariant. But some genetic dispositions manifest themselves more strongly during a certain stage in your life, leading to an increased mortality hazard. This additional source of dependence between twin life-spans would lead to an upward biased bereavement effect. We can investigate this problem further by exploiting the detailed information on death causes available in our dataset. In summary, it is conceivable that unobserved time-varying shared influences are partly responsible for the dependence between twin life-spans which our model can not capture. However, our model

[19]Consider the case in which a twin who's co-twin is diagnosed with a terminal illness is so severely affected by this anticipated loss that he will himself die before his co-twin. This very extreme case would constitute a problem for our model since in this case anticipation would cause the estimated bereavement effect to capture a causal effect that is reverse.

controls for the major source of dependence that are time-invariant shared influences.

An additional source for unobserved time-varying shared variation are events that affect the health of both twins at the same time during their life. However, twins typically have their own family and support systems and they usually don't life in the same area. Further, living in Denmark during the period 1870 to 2004 reduces the probability of being exposed to shocks on the national level such as major wars or epidemics. Additionally, the impact of health shocks is dampened by a well established health care system (see Section 3 for more details).

7 Conclusion

The contribution of this paper is twofold. First, we show that the symmetric version of the timing-of-events model (Model A) can be identified in the case of very limited covariate variation by imposing a multiplicative structure on the bereavement effect function. More specifically, the only exogenous variation that we exploit can be generated by a single dummy variable that does not need to vary between the two durations. The identification results of this symmetric model have wider relevance for the empirical study of parallel systems and networks and for epidemiological research. Besides the application to twin life-spans our model can be applied to any symmetric bivariate duration model in which the dependence between durations is caused by two effects: the influence of time-invariant shared factors and a causal effect. In particular, our results allow the estimation of a model in which the two durations are not distinguishable from each other in any way. Even in the case where the index of duration 1 and 2 is completely uninformative and the only available covariates are characteristics of the pair of durations (not individual characteristics), our identification result still applies. Such completely symmetric systems can be found in many different fields of research.

Second, our empirical analysis is the first approach to model twin life-spans by uniting two models that have been previously used in this strand of literature. With our model we are able to disentangle both of the effects that were only addressed separately in previous twin studies: the causal effect of bereavement and the influence of time-constant shared factors. Ignoring the influence of shared factors has particularly strong consequences in the case of twins since their life-spans are closely connected due to their shared genetic makeup and early childhood.

References

J. H. Abbring and J. J. Heckman. Econometric evaluation of social programs, part iii. In *Handbook of Econometrics, edition 1, volume 6*. Elsevier, 2007.

Jaap H. Abbring and G. J. van den Berg. The nonparametric identification of treatment effects in duration models. *Econometrica*, 71(5):1491–1517, 09 2003a.

Jaap H. Abbring and Gerard J. van den Berg. The identifiability of the mixed proportional hazards competing risks model. *Journal Of The Royal Statistical Society Series B*, 65(3):701–710, 2003b.

S. Ansart, C. Pelat, P. Boelle, F. Carrat, A. Flahault, and A. Valleron. Mortality burden of the 1918-1919 influenza pandemic in europe. *Influenza and Other Respiratory Viruses*, 3:99–106, 2009.

Ann Bowling. Mortality after bereavement: A review of the literature on survival periods and factors affecting survival. *Social Science & Medicine*, 24(2):117–124, January 1987.

Vladimir Canudas-Romo and Annette Erlangsen. Denmark: the lowest excess mortality during the influenza pandemic of 1918. 2008.

G. Doblhammer. *The late life legacy of very early life*. Springer, Berlin, 2004.

Chris Elbers and Geert Ridder. True and spurious duration dependence: The identifiability of the proportional hazard model. *Review of Economic Studies*, 49(3):403–09, 1982.

Nancy Flourney. Discussion of "assessment of dependence in the life times of twins". In J.P. Klein and P.K. Goel, editors, *Survival Analysis: State of the Art*, pages 91–97. Kluwer Academic Publishers, 1992.

John E. Freund. A bivariate extension of the exponential distribution. *Journal of the American Statistical Association*, 56(296):971–977, 1961.

L. A. Gavrilov and V. N. Nosov. A new trend in human mortality decline: derectangularization of the survival curve. *Age*, 8:93, 1985.

N. V. Holm. *The use of twin studies to investigate causes of diseases with complex etiology with a focus on cancer*. PhD thesis, Odense University, 1983.

Philip Hougaard, Bent Harvald, and Niels V. Holm. Assessment of dependence in the life times of twins. In J.P. Klein and P.K. Goel, editors, *Survival Analysis: State of the Art*, pages 77–90. Kluwer Academic Publishers, 1992a.

Philip Hougaard, Bent Harvald, and Niels V. Holm. Measuring the similarities between the lifetimes of adult danish twins born between 1881-1930. *Journal of the American Statistical Association*, 87(417):17–24, March 1992b.

R. A. Kortram, A. C. M. van Rooij, A. J. Lenstra, and G. Ridder. Constructive identification of the mixed proportional hazards model. *Statistica Neerlandica*, 49(3):269–281, 1995.

Tony Lancaster. *The econometric analysis of transition data*. Cambridge: Cambridge University Press, 1990.

D. T. Lykken. The diagnosis of zygosity in twins. *Behavior Genetics*, 8:437–473, 1978.

Orly Manor and Zvi Eisenbach. Mortality after spousal loss: are there socio-demographic differences? *Social Science & Medicine*, 56(2):405 – 413, 2003.

Catherine M. Sanders. A comparison of adult bereavement in the death of a spouse, child, and parent. *Journal of Death and Dying*, 10(4):303–322, 1980.

Catherine M. Sanders. *Grief: The Mourning After: Dealing with Adult Bereavement, 2nd Edition*. Wiley & Sons, Chichester, 1999.

Nancy L. Segal and S. A. Blozis. Psychobiological and evolutionary perspectives on coping and health characteristics following loss: a twin study. *Twin Research*, 5:175–87, 2002.

Nancy L. Segal and Thomas J. Bouchard. Grief intensity following the loss of a twin and other relatives: test of kinship genetic hypotheses. *Human Biology*, 65:87–105, 1993.

Nancy L. Segal and Sarah L. Ream. Decrease in grief intensity for deceased twin and non-twin relatives: an evolutionary perspective. *Personality and Individual Differences*, 25(2):317 – 325, 1998.

Nancy L. Segal, Steven M. Wilson, Thomas J. Bouchard, and Dinah G. Gitlin. Comparative grief experiences of bereaved twins and other bereaved relatives. *Personality and Individual Differences*, 18(4):511 – 524, 1995.

Hans Selye. A syndrome produced by diverse nocuous agents. *Nature*, 138:32, 1936.

Hans Selye. Stress and disease. *Science*, 122(3171):625–631, 1955.

A. Skytthe, K. Kyvik, N.V. Holm, J.W. Vaupel, and K. Christensen. The danish twin registry: 127 birth cohorts of twins. *Twin Research*, 5:352–357, 2002.

C. Tomassini, F. C. Billari, A. Rosina, A. Skytthe, and K. Christensen. Born together-die together. live together-die together. the role of the partner and of the co-twin on longevity at very old ages. *Genus*, LVII(3-4):63–82, 2001.

C. Tomassini, A. Rosina, F. C. Billari, A. Skytthe, and K. Christensen. The effect of losing the twin and losing the partner on mortality. *Twin Research*, 5:210–217, 2002.

Gerard J. van den Berg. Duration models: specification, identification and multiple durations. *Handbook of Econometrics*, 5:3381–3460, 2001.

Gerard J. van den Berg, Maarten Lindeboom, and France Portrait. Conjugal bereavement effects on health and mortality at advanced ages. IZA Discussion Paper, October 2006.

D. V. Widder. *The Laplace Transform*. Princeton University Press, Princeton, 1946.

A. Wienke, N.V. Holm, A. Skytthe, and A.I. Yashin. The heritability of mortality due to heart diseases: a correlated frailty model applied to danish twins. *Twin Research*, 4:266–74, August 2001.

Andreas Wienke, Kaare Christensen, Axel Skytthe, and Anatoli I. Yashin. Genetic analysis of cause of death in a mixture model of bivariate lifetime data. *Statistical Modelling*, 2:89–102, 2002.

Joan Woodward. The bereaved twin. *Acta Geneticae Medicae et Gemellologiae*, 37:173–80, 1988.

A. I. Yashin and I. A. Iachine. Genetic analysis of durations: Correlated frailty model applied to survival of danish twins. *Genetic Epidemiology*, 12:529–538, 1995a.

A. I. Yashin and I. A. Iachine. How long can humans live? lower bound for biological limit of human longevity calculated from danish tiwn data using

correlated frailty model. *Mechanisms of Ageing and Development*, 80:147–169, 1995b.

A Appendix to the proof of propositions 1 and 2

A.1 Identification of λ and ϕ

The survival function of $Z|x$ with $Z = \min\{T_1, T_2\}$ is derived as follows

$$\begin{aligned} P(Z > t|x) &= P(T_1 > t, T_2 > t|x) \\ &= \int_0^\infty \int_0^\infty P(T_1 > t|x, V_1) P(T_2 > t|x, V_2) \, \mathrm{d}G(v_1, v_2) \\ &= \int_0^\infty \int_0^\infty \mathrm{e}^{-\phi(x)\Lambda(t)(V_1+V_2)} \, \mathrm{d}G(v_1, v_2) \\ &= \int_0^\infty \mathrm{e}^{-\phi(x)\Lambda(t)W} \, \mathrm{d}G_W(w) \qquad \textit{,with } W = V_1 + V_2. \end{aligned} \tag{5}$$

Note, that for the second equality we exploit that before the first exit occurs no bereavement effect will cause a dependence between T_1 and T_2. Consequently, conditional on x and V the events $(T_1 > t)$ and $(T_2 > t)$ are independent. We further use Assumption 4 which implies $G(v_1, v_2|x) = G(v_1, v_2)$.

In the following, we will discuss some of the assumptions used by Elbers and Ridder (1982) for the identification of a mixed proportional hazard model in view of the model given in equation 5. Assumption 1 assures sufficient covariate variation in form of at least one dummy variable.[20] Further, the distribution of W has to be independent of x and has a positive and finite mean. Assumption 4 assures the independence of (V_1, V_2) and x. From this the independence of $W = V_1 + V_2$ directly follows. Similarly, as V_1 and V_2 are assumed to have finite positive mean, so does W.

A.2 Identification of δ_a

We consider the following hazard rate of mixed proportional form:

$$\theta(t|T_k = 0, x, V_j) = \tilde{\lambda}_j(t)\tilde{\phi}_j(x)V_j \;\; \textit{with } \tilde{\lambda}_j(t) = \lambda(t)\delta_a, \;\; \tilde{\phi}_j(x) = \phi(x)\delta_b(0, x), \tag{6}$$

where the frailties V_j are drawn from $G_{V_j|T_k=0,x}$ for $j, k \in \{1, 2\}$ and $j \neq k$. One necessary assumption for the identifiability of this mixed proportional hazard model is that the frailty distribution does not depend on x. Note, that in the

[20] Also see Kortram et al. (1995) for the case of only two possible values for $\phi(x)$.

above model, the frailties V_j are drawn from a conditional distribution. Therefore, we need to show that $G_{V_j|T_k=0,x}$ does not depend on x. The conditional density of $V_j|(T_k = 0, x)$ is given by:

$$\begin{aligned} f(v_j|T_k = 0, x) &= \frac{\theta_k(0|x, V_j)S_k(0|x, V_j)f(v_j|x)}{\theta_k(0|x)S_k(0|x)} \\ &= \frac{\int_0^\infty \lambda(0)\phi(x)v_k \,\mathrm{d}G(v_k|x, V_j)f(v_j|x)}{\int_0^\infty \lambda(0)\phi(x)v_k \,\mathrm{d}G(v_k|x)} \\ &= \frac{E(V_k|x, V_j)f(v_j|x)}{E(V_k|x)}. \end{aligned} \tag{7}$$

According to Assumption 4 (V_1, V_2) are independent of x. Therefore, equation 7 simplifies to

$$f(v_j|T_k = 0, x) = \frac{E(V_k|V_j)f(v_j)}{E(V_k)}. \tag{8}$$

From equation 8 it also follows that the distribution of $(V_j|T_k = 0)$ for $j, k \in \{1, 2\}$ and $j \neq k$ has a positive and finite mean, since $G(v_1, v_2)$ has this property.

A.3 Identification of G

A.3.1 Derivation of a mixing distribution

The density $f(t_1, t_2|x)$ for $t_1 > t_2$ can be expressed as follows

$$\begin{aligned} f(t_1, t_2|x) &= \int_0^\infty \int_0^\infty f(t_1|T_2 = t_2, x, V_1)f(t_2|x, V_2) \,\mathrm{d}G(v_1, v_2) \\ &= c(t_1, t_2, x) \int_0^\infty \int_0^\infty V_1 V_2 e^{-\phi(x)(\Lambda(t_2)+\Delta_1(t_1,t_2,x))V_1} e^{-\phi(x)\Lambda(t_2)V_2} \,\mathrm{d}G(v_1, v_2) \\ &= c(t_1, t_2, x)\partial^2_{s_1,s_2}\mathcal{L}_G\big(\phi(x)(\Lambda(t_2) + \Delta_1(t_1, t_2, x)), \phi(x)\Lambda(t_2)\big), \end{aligned}$$

with $c(t_1, t_2, x) = \lambda(t_1)\lambda(t_2)\phi(x)^2\delta_a(t_1-t_2)\delta_b(t_2, x)$, $\Delta(t_1, t_2, x) = \tilde{\Lambda}(t_1, t_2)\delta_b(t_2, x)$ and bivariate Laplace transformation $\mathcal{L}_G$ with cross derivative $\partial^2_{s_1,s_2}\mathcal{L}_G$.

A.3.2 Complete monotonicity

Definition 1 *Let Ω be a nonempty open set in $\mathbb{R}^n$. A function $f : \Omega \to \mathbb{R}$ is absolutely monotone if it is nonnegative and has nonnegative continuous partial derivatives of all orders. f is completely monotone if $f \circ m$ is absolutely monotone, where*

$m : x \in \{\omega \in \mathbb{R}^n : -\omega \in \Omega\} \to -x$.[21]

Note, that this definition states that a function f is completely monotone if it's derivatives of all orders exist, and if these derivatives are continuous and have switching signs for each order (starting with a positive first derivative). From this definition it follows directly that if a function f is completely monotone then all derivatives of second order of f will also be completely monotone. Since the bivariate Laplace transformation $\mathcal{L}_G(s_1, s_2)$ is known to be a completely monotone function, it directly follows from Definition 1 that the cross derivative of $\mathcal{L}$ given by $\partial^2_{s_1,s_2}\mathcal{L}_G(s_1, s_2) = \frac{\partial^2 \mathcal{L}_G(s_1,s_2)}{\partial s_1 \partial s_2}$ is also completely monotone.

A.3.3 Tracing out the Laplace transformation

The function $f : \mathbb{R}^2_+ \to \mathbb{R}^2_+$ is given by $f(t_1, t_2) = (\phi(x)(\Lambda(t_2)+\Delta(t_1, t_2, x)), \phi(x)\Lambda(t_2))$. It maps the vector (t_1, t_2) on the vector of arguments of the Laplace transformation (s_1, s_2), with $s_1 = \phi(x)(\Lambda(t_2) + \Delta(t_1, t_2, x))$ and $s_2 = \phi(x)\Lambda(t_2)$. In the following we will show that we can vary (t_1, t_2) on an open set such that $f(t_1, t_2)$ will also attain all values in a nonempty open set. Under Assumption 5 (with $t_1 > t_2 \; \forall \; (t_1, t_2) \in \Psi$) it holds that at all points (t_1, t_2) in the open set Ψ the first derivatives of f exist and are continuous and f has Jacobian

$$J_f(t_1, t_2) = \begin{bmatrix} \phi(x)\lambda(t_1)\delta(t_1, t_2, x) & \phi(x)(\lambda(t_2) + \frac{\partial\Delta(t_1,t_2,x)}{t_2}) \\ 0 & \phi(x)\lambda(t_2) \end{bmatrix}.$$

Note, that the determinant of J_f is given by $det(J_f(t_1, t_2)) = \phi(x)^2\lambda(t_1)\lambda(t_2)\delta_1(t_1, t_2, x)$, and since under Assumptions 1-4 the functions $\phi, \lambda, \delta_a, \delta_b$ can only attain strictly positive (and finite) values on Ψ, it follows that $det(J_f(t_1, t_2)) \neq 0 \; \forall \; (t_1, t_2) \in \Psi$. Assumption 5 assures that $\frac{\partial\Delta(t_1,t_2,x)}{t_2}$ exists and is continuous on Ψ. Therefore, on the nonempty open set Ψ the function $f(t_1, t_2)$ is continuously differentiable with invertible Jacobian J_f. From the Inverse-Function Theorem it directly follows that there exists an nonempty open set $\Upsilon \subset (0, \infty)^2$ such that the function $f(t_1, t_2)$ attains all values in Υ when t_1 and t_2 vary over $\Psi \subset (0, \infty)^2$.

[21] For $n = 1$ this definition reduces to the familiar definitions in Widder (1946).

B Appendix to the derivation of the likelihood function

In the following the functional forms of S, S_{t_1}, S_{t_2} and S_{t_1,t_2} are derived. We start with the survival function $S(t_1, t_2|x) = P(T_1 > t_1, T_2 > t_2|x)$:

$$S(t_1, t_2|x) = \begin{cases} S^*(t_1, t_1|x) - \int_{t_2}^{t_1} S_{t_2}(t_1, \tau|x)\, d\tau & \text{, for } t_1 \geq t_2 \\ S^*(t_2, t_2|x) - \int_{t_1}^{t_2} S_{t_1}(\tau, t_2|x)\, d\tau & \text{, for } t_1 < t_2 \end{cases}$$

Recall, that in the case when the first exit is not observable due to censoring we have to integrate over all possible exit times. The resulting integrals $\int_{t_2}^{t_1} S_{t_2}(t_1, \tau|x)\, d\tau$ and $\int_{t_1}^{t_2} S_{t_1}(\tau, t_2|x)\, d\tau$ are approximated with numerical integration methods. Here, $S^*(t_1, t_2|x)$ denotes the survival function in the absence of a bereavement effect

$$\begin{aligned} S^*(t_1, t_2|x) &= \iint_0^\infty P(T_1 > t_1|x, V_1) P(T_2 > t_2|x, V_2)\, \mathrm{d}G(v_1, v_2) \\ &= \iiint_0^\infty \mathrm{e}^{\phi(x)\Lambda(t_1)(\tilde{V}_0+\tilde{V}_1)} \mathrm{e}^{\phi(x)\Lambda(t_2)(\tilde{V}_0+\tilde{V}_2)}\, \mathrm{d}G(\tilde{v}_0) dG(\tilde{v}_1) dG(\tilde{v}_2) \\ &= \int_0^\infty \mathrm{e}^{\phi(x)[\Lambda(t_1)+\Lambda(t_2)]\tilde{V}_0}\, \mathrm{d}G(\tilde{v}_0) \int_0^\infty \mathrm{e}^{\phi(x)\Lambda(t_1)\tilde{V}_1}\, \mathrm{d}G(\tilde{v}_1) \int_0^\infty \mathrm{e}^{\phi(x)\Lambda(t_2)\tilde{V}_2}\, \mathrm{d}G(\tilde{v}_2) \\ &= (1 + \sigma^2\phi(x)[\Lambda(t_1) + \Lambda(t_2)])^{-\frac{\rho}{\sigma^2}} (1 + \sigma^2\phi(x)\Lambda(t_1))^{-\frac{(1-\rho)}{\sigma^2}} (1 + \sigma^2\phi(x)\Lambda(t_2))^{-\frac{(1-\rho)}{\sigma^2}}. \end{aligned}$$

The equalities 2-4 follow from the assumption that $G(v_1, v_2)$ is a Cherian bivariate Gamma distribution with independent terms $\tilde{V}_0, \tilde{V}_1, \tilde{V}_2$ drawn from Gamma distributions $\tilde{V}_0 \sim \Gamma(\rho\sigma^{-2}, \sigma^{-2})$ and $\tilde{V}_1, \tilde{V}_2 \sim \Gamma((1-\rho)\sigma^{-2}, \sigma^{-2})$.

In the following S_{t_j} is derived. For this purpose we define the functions g_a, g_b and g_c

$$\begin{aligned} g_a(s_1, s_2, x) &= 1 + \sigma^2\phi(x)[\Lambda(s_2) + \Delta(s_1|s_2, x)] \\ g_b(s_1, s_2, x) &= 1 + \sigma^2\phi(x)[2\Lambda(s_2) + \Delta(s_1|s_2, x)] \\ g_c(s, x) &= 1 + \sigma^2\phi(x)\Lambda(s). \end{aligned}$$

with $\Delta(s_1|s_2, x) = \int_{s_2}^{s_1} \lambda(u)\delta^t(u - s_2)\delta^{age,x}(s_2, x)\ \mathrm{d}u$.

We can now derive $S_{t_j}(t_j, t_k|x) = \frac{\partial S(t_j, t_k|x)}{\partial t_j} = -P(T_j = t_j, T_k > t_k|x)$. Let

$t_j \geq t_k$ with $j,k \in \{1,2\}, j \neq k$

$$\begin{aligned}
S_{t_k}(t_j,t_k|x) &= \iint_0^\infty P(T_j > t_j|T_k = t_k, x, V_j) P(T_k = t_k|x, V_k)\, \mathrm{d}G(v_j, v_k) \\
&= \phi(x)\lambda(t_k) \\
&\quad \iiint_0^\infty (\tilde{V}_0 + \tilde{V}_k) \mathrm{e}^{\phi(x)[\Lambda(t_k)+\Delta(t_j|t_k,x)](\tilde{V}_0+\tilde{V}_j)} \mathrm{e}^{\phi(x)\Lambda(t_k)(\tilde{V}_0+\tilde{V}_k)}\, \mathrm{d}G(\tilde{v}_0) dG(\tilde{v}_j) dG(\tilde{v}_k) \\
&= \phi(x)\lambda(t_k) g_b(t_j,t_k,x)^{-(\frac{\rho}{\sigma^2}+1)} g_c(t_k,x)^{-(\frac{(1-\rho)}{\sigma^2})} g_a(t_j,t_k,x)^{-(\frac{(1-\rho)}{\sigma^2}+1)} \\
&\quad [\rho g_a(t_j,t_k,x) + (1-\rho) g_b(t_j,t_k,x)].
\end{aligned}$$

This yields

$$S_{t_j}(t_j,t_k|x) = \begin{cases} \frac{\partial S^*(t_j,t_j|x)}{\partial t_j} + \int_{t_k}^{t_j} S_{t_1,t_2}(t_1,\tau|x)\, d\tau & \text{, for } t_j > t_k \\ \phi(x)\lambda(t_k) g_b(t_j,t_k,x)^{-(\frac{\rho}{\sigma^2}+1)} g_c(t_k,x)^{-(\frac{(1-\rho)}{\sigma^2})} & \\ g_a(t_j,t_k,x)^{-(\frac{(1-\rho)}{\sigma^2}+1)} [\rho g_a(t_j,t_k,x) + (1-\rho) g_b(t_j,t_k,x)] & \text{, for } t_j \leq t_k. \end{cases}$$

Finally, $S_{t_1,t_2}(t_1,t_2|x) = \frac{\partial^2 S(t_1,t_2|x)}{\partial t_1 \partial t_2} = P(T_1 = t_1, T_2 = t_2|x) = f^*(\max\{t_1,t_2\}, \min\{t_1,t_2\})$ with

$$\begin{aligned}
f^*(t_j,t_k) &= \phi(x)^2 \lambda(t_j)\lambda(t_k) \delta^t(t_j - t_k) \delta^{age,x}(t_k,x) \\
&\quad g_b(t_j,t_k,x)^{-(\frac{\rho}{\sigma^2}+2)} g_a(t_j,t_k,x)^{-(\frac{(1-\rho)}{\sigma^2}+1)} g_c(t_k,x)^{-(\frac{(1-\rho)}{\sigma^2}+1)} \\
&\quad [\rho(\rho+\sigma^2) g_a(t_j,t_k,x) g_c(t_k,x) + \rho(1-\rho) g_b(t_j,t_k,x) g_c(t_k,x) \\
&\quad \rho(1-\rho) g_b(t_j,t_k,x) g_a(t_j,t_k,x) + (1-\rho)^2 g_b(t_j,t_k,x)^2].
\end{aligned}$$