DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 30, ARTICLE 58, PAGES 1591–1620
PUBLISHED 21 MAY 2014
http://www.demographic-research.org/Volumes/Vol30/58/
DOI: 10.4054/DemRes.2014.30.58

Research Article

Multistate event history analysis with frailty

Govert E. Bijwaard

This publication is part of the Special Collection on "Multistate Event History Analysis," organized by Guest Editors Frans Willekens and Hein Putter.

Table of Contents

Demographic Research: Volume 30, Article 58
Research Article

Multistate event history analysis with frailty

Govert E. Bijwaard [1]

Abstract

BACKGROUND
In survival analysis a large literature using frailty models, or models with unobserved heterogeneity, exists. In the growing literature and modelling on multistate models, this issue is only in its infant phase. Ignoring frailty can, however, produce incorrect results.

OBJECTIVE
This paper presents how frailties can be incorporated into multistate models, with an emphasis on semi-Markov multistate models with a mixed proportional hazard structure.

METHODS
First, the aspects of frailty modeling in univariate (proportional hazard, Cox) and multivariate event history models are addressed. The implications of choosing shared or correlated frailty is highlighted. The relevant differences with recurrent events data are covered next. Multistate models are event history models that can have both multivariate and recurrent events. Incorporating frailty in multistate models, therefore, brings all the previously addressed issues together. Assuming a discrete frailty distribution allows for a very general correlation structure among the transition hazards in a multistate model. Although some estimation procedures are covered the emphasis is on conceptual issues.

RESULTS
The importance of multistate frailty modeling is illustrated with data on labour market and migration dynamics of recent immigrants to the Netherlands.

[1] Netherlands Interdisciplinary Demographic Institute (NIDI), PO Box 11650, 2502 AR The Hague, The Netherlands; Phone: (+31) 70 3565 224; Fax: (+31) 70 3647187; E-Mail: bijwaard@nidi.nl.

1. Introduction

Demographers are increasingly interested in understanding life histories or the individual life course, with a focus on events, their sequence, ordering and transitions that people make from one state of life to another. A multistate model describes the transitions people experience as life unfolds. When people may change among a set of multiple states and/or may experience repeated changes through time, a multistate event history model, also known as multistate lifetable and increment-decrement life tables, is a proper choice. Typical examples of such processes in demography include migration, (Rogers 1975; 1995), changes in marital status and other life course processes, (Courgeau and Lelièvre 1992 and Willekens 1999). Many other demographic applications of the multistate models exist. Multistate models are also common in medicine and economics. In medicine, the states can designate conditions such as healthy, diseased and death. For an overview of the use of multistate models in biostatistics, see a.o. Commenges (1999), Hougaard (2000), and Putter, Fiocco, and Geskus (2007). In economics the main application of multistate models has been labour force dynamics; see Flinn and Heckman (1983), Van den Berg (2001) and, Fougère and Kamionka (2008). Poverty dynamics and recidivism are other important applications of multistate models. The methodology of multistate models is discussed in several books; the most important are Andersen et al. (1993), Hougaard (2000), and Aalen, Borgan, and Gjessing (2008).

In our empirical application we focus on the return decision of labour migrants and its relation to labour market dynamics. Many migrants only stay temporarily in the host country. On the one hand, return migration is seen as planned and part of optimal decision making to maximize total utility over the whole life cycle, where return migration is motivated by locational preference for the home country. On the other hand, return migration is seen as unplanned and the result of failure either due to imperfect information about the host country in terms of labor market prospects or the cost of living, or the inability to fulfil the migration plans in terms of target savings. In both cases, return behaviour is intrinsically related to the timing of labour market changes of the individual migrant. Migrants who become unemployed are more prone to leave, but when they find a new job again they are more prone to stay, see Bijwaard, Schluter, and Wahba (2014). Migrants who are employed in high paying jobs have a lower probability of becoming unemployed and can accumulate more savings while working. When these migrants have reached target savings they are more prone to leave, see Bijwaard and Wahba (2014). Labour market dynamics may also be affected by the labour market history. These factors suggest that return migration behaviour of labour migrants should be modeled by a multistate model.

The basic parameters of a multistate model are the transition hazard rates or intensities. These intensities may depend on the time spent in a particular state (semi-Markov models) and on observed characteristics. Many multistate models assume that the intensities are homogeneous, conditional on these observed factors. Unfortunately, it is hardly ever possible to include all relevant factors, either because the researcher does not know all the relevant factors, or because it is not possible to measure all the relevant factors. Ignoring such unobserved heterogeneity or frailty may have a large impact on inference in multistate models. The duration dependence, the effect of the length of the duration in a particular state on the exit rate out of this state, will be biased towards a more declining effect of the duration when frailty is ignored. The effect of covariates on the transition rates will be biased towards zero when frailty is ignored. For univariate event history data, also called survival data or duration data, a large literature on models with frailty exits, e.g. Van den Berg (2001), Duchateau and Janssen (2008) and Wienke (2011). In the multistate literature, the issue of including frailty is only in its infant phase. The few articles that deal with frailty in multistate models are Pickles and Crouchley (1995), Govindarajulu et al. (2011) and Putter and van Houwelingen (2011).

The purpose of this article is to provide an overview of frailty modeling for multistate event history models. We assume that the frailty for all members, just as the effect of observed characteristics, enters the intensity multiplicatively. Thus, we only consider the Cox model in continuous time with frailty, in econometrics called the Mixed Proportional Hazard model, and its multivariate extensions.

The outline of the paper is as follows. In Section 2. we start with a discussion on the issues involved in frailty in univariate survival models. Multistate models extend the univariate survival models in two dimensions: (1) Several people that may experience an event may be grouped in clusters. Note that this is conceptually equal to the situation in which for one individual several processes are followed simultaneously. In both cases we have parallel events. (2) People may experience multiple periods of the same type or recurrent events. In both dimensions, frailties can be independent, shared, or correlated. We also discuss these issues separately for parallel and for recurrent event data. In Section 3. frailties in a multistate setting are addressed, combining the knowledge of the preceding section on incorporating frailties in models for parallel data and in models for recurrent data. In Section 4. we illustrate the importance of incorporating frailty in a semi-Markov multistate model with data on labour market and migration dynamics of recent immigrants to the Netherlands. In Section 5. we briefly discuss identification issues in multistate frailty models. Section 6. summarizes the findings.

2. Frailty in event history models

2.1 Frailty in univariate event history models

The simplest multistate model is a univariate survival model, which considers the transition from 'alive' to 'dead', or return migration, in our running example. The observation for a given individual will in this case consist of a random variable T, representing the time from a given origin (time 0) to the occurrence of the event 'death'. The distribution of T may be characterized by the survival function $S(t) = \Pr(T > t)$. We can also characterize the distribution of T by its hazard rate $\lambda(t)$, which is the transition intensity from state 'alive' to state 'dead', i.e. the instantaneous probability per time unit of going from 'alive' to 'dead'. The hazard rate provides a full characterization of the distribution of T, just like the distribution function, the survival function, and the density function.

A typical feature of event history analysis is the inability to observe complete event histories. A common problem is that by the end of the observation period, some individuals still have not yet experienced the event of interest. This kind of incomplete observation is known as right-censoring. The hazard function is usually the focal point of analysis. A major advantage of using the hazard function as the basic building block is that it is invariant to independent censoring. The most common model for the hazard rate is the Cox or proportional hazard (PH) model, with hazard rate

$$\lambda(t|X) = \lambda_0(t)\exp(\beta' X), \tag{1}$$

where $\lambda_0(t)$ is called the baseline hazard or duration dependence and it is a function of t alone.

In a Mixed Proportional Hazard (MPH) model it is assumed that all unmeasured factors and measurement errors can be captured in a multiplicative random term, the frailty V. The hazard rate becomes

$$\lambda(t|X, V) = V\lambda_0(t)\exp(\beta' X). \tag{2}$$

This model was independently developed by Vaupel, Manton, and Stallard (1979) and by Lancaster (1979). The frailty $V > 0$ is time-independent and independent of the observed characteristics X.

The impact of frailty in event history models differs substantially from the impact of frailty in linear regression models. In ordinary regression models, unobserved heterogeneity leads to more variability of the response compared to the case when the variables are included. In event history data, however, the increased variability implies a change in the hazard function. When the hazard rate exhibits positive duration dependence, ignoring frailty will make this duration dependence less pronounced or even negative. When the hazard rate exhibits negative duration dependence, ignoring frailty will make this negative duration dependence stronger. Another consequence of ignoring frailty is that the effect of a covariate is biased towards zero.

The most commonly used frailty distributions are the Gamma frailty distribution, the log-normal frailty distribution, and the discrete frailty distribution. For more details on these and other frailty distributions, like the Power Variance Function family of frailty distributions that includes the important Inverse Gaussian and Stable frailty distributions I refer to Hougaard (2000) and Wienke (2011).

The Gamma distribution is the most widely applied frailty distribution. From an analytical and computational view it is a very convenient distribution. The closed form expressions for the unconditional survival and hazard are easy to derive. The link with random effects or mixed models makes the log-normal model very attractive. A disadvantage is the lack of closed form expressions. But with increasing computer power the numerical solution of the integrals involved is not an issue anymore.

The discrete frailty model, in which it is assumed that the population consists of two, or more, latent sub populations, which are homogeneous within, is a finite mixture model. For example we may have (1) a high risk subpopulation that leaves fast, and (2) a low risk subpopulation that leaves slowly, but the class identification for each individual is unknown. In econometrics such frailty models are commonly applied in survival analysis.

2.2 Frailty in multivariate event history models

There are two typical ways multivariate event history data can arise. The first situation of multivariate event history data is *parallel event history data*, in which for one individual several processes are followed simultaneously. A typical parallel events data example is the competing risks model of different causes of death or of exits from employment to either unemployment or non-participation. Data in which several individuals that may experience an event are grouped in a cluster are conceptually the same with parallel data. Examples of such data include twin and family studies. The second situation of multivariate event history data is *recurrent/repeated events* which arises when several events of the same type are registered for each individual, for instance child birth to a woman, or periods of unemployment. In univariate event history models frailty captures the possible heterogeneity due to unobserved covariates. In a multivariate setting frailty can also be used to model associations between events, but events from different clusters are considered to be independent.

A key point for an MPH model is conditional independence, that is conditional on the frailty v the survival times are independent. In the multivariate setting we continue to assume the MPH structure and the conditional independence. In principle the frailty might be independent for each event. Then the analysis does not differ from the analysis in a univariate setting. Here we consider the more interesting cases of (1) shared frailty and (2) correlated frailty.

The shared frailty approach assumes that within a cluster, the value of the frailty term is constant over time, and common to all individuals in the cluster. Examples include: the frailty in the transition from employment to unemployment is equal to the frailty in the transition from employment to non-participation, and the frailty in the death rate for all members of one family are the same. This common term, creates dependence between event times within a cluster. This dependence is always positive. The shared frailty model, first introduced by Clayton (1978), dominates the literature on multivariate survival models, see Hougaard (2000), Therneau and Grambsch (2000) and Duchateau and Janssen (2008), among others.

Despite the similarity between individual frailty and shared frailty conceptually they are different. In the univariate case, the frailty variance σ^2 is a measure of unobserved heterogeneity, while in a shared frailty multivariate case, the frailty variance is a measure of correlation between event times within a cluster.

The shared frailty model, with a shared Gamma frailty distribution as most popular choice, has some important limitations; see Xue and Brookmeyer (1996) for an extensive discussion. First, the assumption that the frailty is the same for all members in the cluster is often inappropriate. Second, shared frailty models only induce positive association within clusters. However, in some situations, the event times for individuals within the same cluster are negatively associated. For example, the reduction in the risk of dying from one disease may increase the risk of dying from another disease. Third, the dependence between survival times within a cluster is based on marginal distributions of event times. This leads to a symmetric relationship between all possible pairs within a cluster. It also limits the interpretation of the variance of the shared frailty model as a measure of association between event times within a cluster, and not as a measure of unobserved heterogeneity. Correlated frailty models allow more flexibility.

In a correlated frailty model the frailties of individuals within a cluster are correlated but not necessarily shared. It enables the inclusion of additional correlation parameters and associations are no longer forced to be the same for all pairs of individuals within a cluster. We consider three different ways of generating correlated frailties: (*i*) additive frailty in which the frailty is the sum of a cluster-specific and an individual-specific component; (*ii*) nested frailty, in which the frailty is the multiplication of a cluster-specific and an individual-specific component; and (*iii*) joint modeling of the member specific frailties within a cluster. In all three cases the conditional survival still has an MPH structure.

The *additive frailty* model based on a correlated gamma distribution was introduced by Yashin, Vaupel, and Iachine (1995). The model has a very convenient representation of the survival function in closed form. It consists of a bivariate model clustered event history model with additive gamma frailty. Each frailty is constructed by adding two components, one common to both and one individual specific. A consequence of the model structure is that when the values of the variances of the two individual terms differ a lot the correlation cannot be very large. Another disadvantage of the additive correlated gamma frailty is that estimation of the model becomes very complex with increasing cluster size.

The *nested frailty* model assumes that the clustering of the event times occurs at multiple levels. In family studies, where we have a hierarchical clustering by family and individual, this models seems appropriate. Sastry (1997a) suggested a nested frailty model with two hierarchical levels in which the frailty of member j in a particular cluster is $V_j = W_0 \cdot W_j$ with W_0 and W_j are mutually independent unit-mean gamma distributed random variables with variance η_0 and η_1. Thus, within each cluster the frailty is composed of a cluster-specific component common to all cluster members times an individual-specific component that are mutually independent. The unconditional survival for this nested gamma frailty has a complicated form, but estimation is possible using an EM-algorithm (Sastry 1997a; 1997b), a Bayesian procedure (Manda 2001) or penalized likelihood methods (Rondeau, Commenges, and Joly 2003).

A very flexible way to allow for correlated frailties is by modeling the *joint frailty* distribution directly. The correlated log-normal frailty model, first applied by Xue and Brookmeyer (1996), is especially useful in modeling dependence structures. The distribution can be obtained by assuming a multivariate normal distribution on the logarithm of the frailty vector. However, the log-normal distribution does not have analytical solutions for the unconditional joint survival and hazards, and the number of integrals to evaluate for calculating them increases with the dimension of the multivariate normal distribution.

Assuming a joint discrete frailty distribution is another way to allow for correlation between the frailties. However, in an unstructured discrete frailty model the number of (additional) parameters increases fast. This dimensional burden can be reduced by assuming a factor loading specification, e.g. 2-factor loading model in which $V_j = \exp(\alpha_{j1}W_1 + \alpha_{j2}W_2)$ with W_1 and W_2 are binary mutually independent variables on $(-1, 1)$ with $p_k = Pr(W_k = 1)$.

2.3 Frailty in recurrent events models

Another extension of the univariate survival model is that an individual can experience the same event several times, e.g. become repeatedly unemployed. Reviews of models for such recurrent event data appeared in Cook and Lawless (2007). Recurrent data can be represented in different ways depending on the timescale that is used; (i) gap time, or clock reset time; (ii) total time, or clock forward time, see e.g. Kelly and Lim (2000).

Related to the choice of the time scale are the *risk-interval* and the *risk set*. The risk interval corresponds to the time interval where an individual is at risk of experiencing an event. The risk set is the collection of individuals which are at risk at a certain point in time. In the gap-time representation, time at risk starts at 0 after an event and ends at the time of the next event. Hence, time is reset to zero after each event. In the total-time formulation, the length of the time at risk is the same as in the gap-time representation. The difference is that the starting time of the at-risk period is not reset to zero after an event but it is put equal to the actual time since the beginning of the observation period.

Based on the choice of the risk set, the three most common approaches to recurrent events are the independent increment model of Andersen and Gill (1982), the marginal model of Wei, Lin, and Weissfeld (1989), and the conditional model of Prentice, Williams, and Peterson (1981); see Kelly and Lim (2000) for a comparison. For the marginal and conditional models, each occurrence of the event is modeled as a separate event, while the independent increment model assumes that the occurrence of an event of one individual is independent of the number and timing of previous events. The independent increments model is usually defined in total time, but it can also be formulated in gap time. This model assumes that the gap times are generated from a renewal process. In essence, the marginal model treats the consecutive event times as if they come from an unordered competing risk setting, with the number of occurrences at the number of competing events. The marginal model can only be formulated in total time. The conditional model assumes that an individual cannot be at risk for the second occurrence of an event until the event has occurred for the first time.

Nielsen et al. (1992) discuss how to include frailty in the independent increments model. In Chapter 9 of his seminal book, Hougaard (2000) also discusses shared frailty models for recurrent events in which the frailty is shared over time. Frailty models specially designed for recurrence data are considered in detail in Oakes (1992), Duchateau et al. (2003) and Bijwaard, Franses, and Paap (2006). For recurrent events the frailty variation is not a group variation, but a variation between individuals, and the variation described by the hazard function is not an individual variation but a variation within individuals. The interpretation of the frailty variance also depends on the time scale and risk sets used.

The arguments for correlated frailty models also apply for recurrent events. An extension of the correlated frailty model that is particularly relevant for recurrent events is the time-dependent frailty model. Often events that occur close in time are highly correlated, while events that occur further apart are less correlated. To model such kind of serial dependence, Yau and McGilchrist (1998) define a dynamic frailty model that assumes that the frailties on subsequent intervals follow an autocorrelation process of order one.

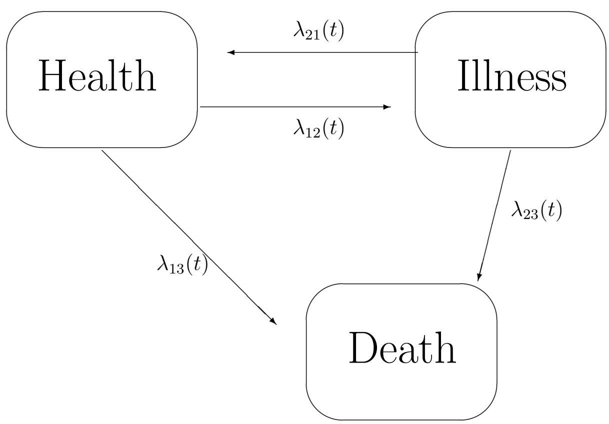
3. Frailty in multistate models

A multistate model is defined as a stochastic process, which at any point in time occupies one of a set of discrete states. The class of multistate models includes both recurrent and multivariate event history data. For example, in labour force dynamics, multiple (un)employment periods are recurrent events and the states are multivariate because from a state of employment an individual can either become unemployed or leave the labour market. In that respect special cases of a multistate model are the multivariate parallel and recurrent models in the previous section. Thus, including frailty in multistate models follows the lines of the previous sections. Before introducing multistate models with frailty we explain the main concepts of multistate models.

3.1 Multistate model concepts

The most commonly applied multistate model in biostatistics is the illness-death model (Putter, Fiocco, and Geskus 2007). This model is depicted in Figure 1. In this class of models individuals start out healthy, the initial state 1. From a healthy state they may become ill (state 2) or they may die (state 3). Ill individuals may die or recover and become healthy again. Most concepts of multistate event history analysis can be explained using this simple model.

Figure 1: Illness-death model



Multistate modeling is closely related to Markov chain theory and many of its terms originate from the theory of Markov chains and processes. Most multistate models have three states: the initial state(s), the states in which an individual can enter the study; absorbing states, states that represent an endpoint from which the individual cannot leave or one is not interested in what happens after this state has been reached; intermediate or transient states are all other states. In an illness-death model death is an absorbing state and illness is an intermediate state. The multistate event history model is defined in hazard/transitions rates. We denote the hazard of making a transition from state i to state j $(i \neq j)$ at t by $\lambda_{ij}(t)$.

Just as for recurrent events, the choice of the time scale in a multistate model has important implications for the analysis. In a total time representation the event times, t, correspond to the time since the individual entered the initial state. The time keeps moving forward, both when intermediate events occur or when the individual returns to the initial state. In a gap time representation, the event times correspond to the time since the entry in state i. The time is reset to zero each time an individual makes a transition. Gap time is also called sojourn time, clock reset time and backward recurrence time. The time scale chosen has implications for the risk set that defines who are at risk for a particular event, for a transition within a multistate framework.

Another important choice in relation to the risk set is whether a Markov model is assumed. In a Markov model the transition rate only depends on the state an individual is in, not on the time an individual has been in that state nor on any other events that occurred before entering that state. Thus, multistate models in gap time representation cannot be Markov models as the time scale itself depends on the history before the current state has been reached. If it is assumed the gap times depend on the history of the process only through the present state, the resulting multistate model is a Markov renewal model. In a semi-Markov model the transition rate from one state to another state also depends on the time the individual has spent in that particular state. A semi-Markov model in total time representation is also possible, but only with an additional time scale, measuring the sojourn time. In a Markov model the transition rate may depend on the time since entry. This is called a time-inhomogeneous Markov model. Including such time dependence in a semi-Markov model implies an additional time scale, measuring the time since entry. We restrict to uni-time scale multistate models with a semi-Markov model in gap time representation. Thus, we assume duration dependence in the transition rates. The duration dependence measures the effect of the length of stay in a particular state. It is rather straightforward to extend this model to let the transition rates depend on the order or occurrence of the particular state, recurrent events effects, by allowing for occurrence specific duration dependence and for occurrence specific covariate effects. We restrict to mixed proportional hazard type transition rates and assume that, conditionally on the value of the frailty, the semi-Markov property holds.

3.2 Frailty in the illness-death model

We illustrate the choices involved in including frailty in multistate models by using the illness-death model. In the illness-death model we have four transitions rates

$$
\begin{aligned}
\lambda_{12}(t|X_{12}, v_{12}) &= v_{12}\lambda_{012}(t)\exp(\beta_{12}X_{12}) && \text{from healthy to ill} && (3)\\
\lambda_{13}(t|X_{13}, v_{13}) &= v_{13}\lambda_{013}(t)\exp(\beta_{13}X_{13}) && \text{from healthy to death} && (4)\\
\lambda_{21}(t|X_{21}, v_{21}) &= v_{21}\lambda_{021}(t)\exp(\beta_{21}X_{21}) && \text{from ill to healthy} && (5)\\
\lambda_{23}(t|X_{23}, v_{23}) &= v_{23}\lambda_{023}(t)\exp(\beta_{23}X_{23}) && \text{from ill to death,} && (6)
\end{aligned}
$$

where X_{12}, X_{13}, X_{21} and X_{23} are the observed individual characteristics. The included covariates might be different for each transition. The baseline hazards $\lambda_{012}(\cdot), \ldots, \lambda_{023}(\cdot)$ depend on the sojourn time in the state, and might be equal for the same origin state. Here we focus on the choice of the frailty distribution. When all the frailties are mutually independent the model reduces to two independent competing risks models. From the healthy state, the competing states the individual can move to are illness and death. From the illness state the individual can either move to healthy or to death. In both cases the competing risks are uncorrelated and the frailty variance is a measure of unobserved heterogeneity within the origin-destination combination.

An illness-death model with shared frailty model by origin state implies equal frailties from the healthy state $v_1 = v_{12} = v_{13}$ and equal frailties from the illness state $v_2 = v_{21} = v_{23}$. The frailty variance is in this case a measure of correlation between events times from either healthy to illness or to death or from illness to healthy or to death.

Concerning correlated frailty models, we have the choice between three different ways of generating correlation between the linked transitions: an additive frailty model, a nested frailty model or a joint frailty model. With correlation based on origin states, we have two sets of mutually independent correlated frailties, the frailties of the healthy state, v_{12} and v_{13}, and the frailties of the illness state, v_{21} and v_{23}. With correlation based on destination states, we have three sets of mutually independent correlated frailties: the frailty to the healthy state, v_{21}, the frailty to the illness state, v_{21} and the frailties to the death state, v_{13} and v_{23}. When all four frailties are correlated, we can use four-dimensional frailty models. Note that the additive gamma frailty models become very complex for four dimensional frailties. For the discrete correlated frailty models, a two-factor loading model specification would leave the parameter space manageable without putting too much restriction on the correlations. An advantage of the discrete model is that it also allows for negative correlations among the unobserved factors. In the prototypical application of the illness-death model in biostatistics, describing the transitions of patients, this possibility might sound redundant, as factors increasing the rate into illness usually also increase the death rate. However, when the illness-death model is applied to socio-economic transitions, restricting to positive correlation can be very restrictive. For example, labour migrants who are more likely to become unemployed ("ill") are often less likely to find a new job again ("healthy").

So far we have ignored possible recurrent behaviour in the model. In the illness-death model, only the health and illness state might be recurrent. But the transition rates to death from these two states may also change with reoccurence. Of course, a simple way to allow for such dependence is to include as an additional covariate the number of times an individual has been in the state. Recurrence may also affect frailties. When the frailties are shared over the occurrences, the possible models are basically the same as mentioned above, with the only difference is that the baseline duration and the regression function are stratified by occurrence. When the frailties are independent over the recurrence, i.e. each recurrence has a separate frailty, the model is just a repeated version of the model above. When allowing a more flexible correlation the possible frailty structures becomes rather large. In principle an extension of the autocorrelated frailty model of Yau and McGilchrist (1998) to the illness-death model is possible. We assume that the frailty is shared or independent over the occurrences.

3.3 General multistate models with frailty

For general multistate models beyond the simple illness-death model, many alternative correlation structures for the frailties are possible. In principle, a multistate model has three dimensions; the origin states, the destination states and the recurrent events of a particular state. The hazard from state i to state j ($i \neq j$) for the k^{th} time is

$$\lambda_{ijk}(t|X, V_{ijk}) = V_{ijk}\lambda_{ijr0}(t)\exp(\beta'_{ijk}X_{ijk}).$$

Of course, it is allowed to put restrictions on the duration dependence, on the observed characteristics or on the effect of the observed characteristics on the hazard. For example, the duration dependence might be shared for all exits of one origin state, $\lambda_{ijr0}(t) = \lambda_{ir0}(t)$, the observed characteristics might be shared over all recurrent events, $X_{ijk} = X_{ij}$, and the effect of these factors might only depend on the destination state $\beta_{ijk} = \beta_j$. Here we only focus on the correlation structure of the frailties.

Table 1: Possible correlation structures of frailty in a multistate model

Origin-destination structure	Recurrent structure: independent	Recurrent structure: shared
Fully independent	$\rho(v_{ijk}, v_{irk}) = 0$ $\rho(v_{ijk}, v_{mjk}) = 0$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 0$	$\rho(v_{ijk}, v_{irk}) = 0$ $\rho(v_{ijk}, v_{mjk}) = 0$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 1$
Shared over origin	$\rho(v_{ijk}, v_{irk}) = 1$ $\rho(v_{ijk}, v_{mjk}) = 0$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 0$	$\rho(v_{ijk}, v_{irk}) = 1$ $\rho(v_{ijk}, v_{mjk}) = 0$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 1$
Shared over destination	$\rho(v_{ijk}, v_{irk}) = 0$ $\rho(v_{ijk}, v_{mjk}) = 1$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 0$	$\rho(v_{ijk}, v_{irk}) = 0$ $\rho(v_{ijk}, v_{mjk}) = 1$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 1$
Fully shared	$\rho(v_{ijk}, v_{irk}) = 1$ $\rho(v_{ijk}, v_{mjk}) = 1$ $\rho(v_{ijk}, v_{mrk}) = 1$ $\rho(v_{ijk}, v_{ijg}) = 0$	$\rho(v_{ijk}, v_{irk}) = 1$ $\rho(v_{ijk}, v_{mjk}) = 1$ $\rho(v_{ijk}, v_{mrk}) = 1$ $\rho(v_{ijk}, v_{ijg}) = 1$
Correlated over origin	$\rho(v_{ijk}, v_{irk}) = \rho_{ij,ir}$ $\rho(v_{ijk}, v_{mjk}) = 0$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 0$	$\rho(v_{ijk}, v_{irk}) = \rho_{ij,ir}$ $\rho(v_{ijk}, v_{mjk}) = 0$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 1$
Correlated over destination	$\rho(v_{ijk}, v_{irk}) = 0$ $\rho(v_{ijk}, v_{mjk}) = \rho_{ij,mj}$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 0$	$\rho(v_{ijk}, v_{irk}) = 0$ $\rho(v_{ijk}, v_{mjk}) = \rho_{ij,mj}$ $\rho(v_{ijk}, v_{mrk}) = 0$ $\rho(v_{ijk}, v_{ijg}) = 1$
Fully correlated	$\rho(v_{ijk}, v_{irk}) = \rho_{ij,ir}$ $\rho(v_{ijk}, v_{mjk}) = \rho_{ij,mj}$ $\rho(v_{ijk}, v_{mrk}) = \rho_{ij,mr}$ $\rho(v_{ijk}, v_{ijg}) = 0$	$\rho(v_{ijk}, v_{irk}) = \rho_{ij,ir}$ $\rho(v_{ijk}, v_{mjk}) = \rho_{ij,mj}$ $\rho(v_{ijk}, v_{mrk}) = \rho_{ij,mr}$ $\rho(v_{ijk}, v_{ijg}) = 1$

With independent frailty, we have for each origin-destination-recurrence pair, v_{ijk}, an independent frailty, and all frailties are uncorrelated. This is, for example, the case when all transitions in a labour-dynamics return migration multistate model are only correlated through observed characteristics. Table 1 provides the possible tractable correlation structures when changing the dependence in all three dimensions. The first column gives the correlation structure when the frailties are independent over the recurrences and the second column when the frailties are shared over the recurrences. When the frailties are shared over recurrences, the correlation between two frailties of different occurrence, $\rho\big(v_{ijk}, v_{ijg}\big)$, is one. When the frailty is shared over the origin state and independent over recurrences, i.e. all destinations from one origin share the same frailty but not over recurrences, then the frailty distribution only depends on the origin state i and the correlation between two frailties from the same origin, $\rho\big(v_{ijk}, v_{irk}\big)$, is one. For example, this amounts to dependence of the transition hazards to unemployment, non-participation and abroad from employment for a particular employment period, but independence of these hazards for different employment periods. When the frailty is shared over the destination state and independent over recurrences, i.e. hazards to one particular destinations share the same frailty but not over recurrences, then the frailty distribution only depends on the destination state j and the correlation between two frailties to the same destination, $\rho\big(v_{ijk}, v_{mjk}\big)$, is one. When the frailty is shared over origin, destination and recurrence states we have only one frailty value for each individual and therefore the correlation is one.

When the frailties are correlated over origin states the frailties from the same origin i to different destinations j and r are correlated and depend on the destination state, i.e. $\rho\big(v_{ijk}, v_{irk}\big) = \rho_{ij,ir}$. When the frailties are correlated over destination states the frailties to the same destination j from different origins i and m are correlated and depend on the origin state, i.e. $\rho\big(v_{ijk}, v_{mjk}\big) = \rho_{ij,mj}$. When the frailties are correlated over both origin and destination states the correlation is defined in all possible origin-destination combinations, e.g. $\rho\big(v_{ijk}, v_{mrg}\big) = \rho_{ij,mr}$. The first situation, correlated over origin, implies that the hazard from employment to unemployment, from employment to non-participation and from employment to living abroad are correlated through the frailty term, while the hazards from unemployment to either employment, non-participation or moving abroad are uncorrelated with the out-off-employment hazards. In the second situation, correlated over destination, the hazards to employment, from unemployment, from non-participation and from abroad are all correlated through frailty. In the third situation, full correlation, the hazards from employment, from unemployment, from non-participation and from living abroad to all other states are all correlated through the frailty.

Multistate models with correlated frailty can become very complex, with many (frailty) parameters. Of course, the choice of the correlation structure in a multistate model depends on data availability and on the questions one wants to answer. For high dimensional multistate models, including correlated frailties also extends the parameters space. By using factor loading models, it is possible to reduce the dimension of correlated frailty distribution a great deal, without loosing much flexibility. Consider a M-factor loading model with frailty

$$v_{ijk} = \prod_{m=1}^{M} \exp\Big(\alpha_{ijk}^{(m)} \cdot W_{ijk}^{(m)}\Big), \tag{7}$$

where $W_{ijk}^{(1)}, \ldots, W_{ijk}^{(M)}$ are M binary variables mutually independent on $(-1, 1)$ with $p_{ijk} = Pr(W_{ijk} = 1)$. For example, in a 2-factor loading model each frailty can attain four different values, $\{e^{\alpha_1+\alpha_2}, e^{-\alpha_1+\alpha_2}, e^{\alpha_1-\alpha_2}, e^{-\alpha_1-\alpha_2}\}$. In general, an M-factor loading model allows for 2^M possible values for each frailty. In Table 2 we display the restrictions on the factor loading and number of factors implied by the alternative correlation structures of Table 1.

Consider, for example, a discrete frailty 2-factor loading model. Table 2 shows that when a separate model is defined for each origin, i.e. separate W_i's for each origin, the frailty is shared over recurrence and correlated or shared over origin. The frailty is shared over the origin states when the factor loadings are the same for each destination, $\alpha_{ij}^{(m)} = \alpha_i^{(m)}$. Similarly, when we have a 2 factor model with factor loadings depending on the origin and destination state, the frailties are fully correlated over the origin and destination states. Note that for these factor loading models, the shared frailty models are nested in the correlated frailty models. A fully shared model is nested in a fully correlated model, and a shared over the origin model is nested in a correlated over origin states model. This implies that testing the equality of the relevant α's is a test on correlated versus shared frailties.

Table 2: Restrictions implied by correlation structure on factor loadings and number of factors for a discrete frailty factor loading model

Origin-destination structure	Recurrent structure independent	shared
Fully independent	$\alpha_{ijk}^{(m)} = \alpha_{ijk}^{(m)}$ $W_{ijk}^{(m)} = W_{ijk}^{(m)}$	$\alpha_{ijk}^{(m)} = \alpha_{ij}^{(m)}$ $W_{ijk}^{(m)} = W_{ij}^{(m)}$
Shared over origin	$\alpha_{ijk}^{(m)} = \alpha_{ik}^{(m)}$ $W_{ijk}^{(m)} = W_{ik}^{(m)}$	$\alpha_{ijk}^{(m)} = \alpha_{i}^{(m)}$ $W_{ijk}^{(m)} = W_{i}^{(m)}$
Shared over destination	$\alpha_{ijk}^{(m)} = \alpha_{jk}^{(m)}$ $W_{ijk}^{(m)} = W_{jk}^{(m)}$	$\alpha_{ijk}^{(m)} = \alpha_{j}^{(m)}$ $W_{ijk}^{(m)} = W_{j}^{(m)}$
Fully shared	$\alpha_{ijk}^{(m)} = \alpha_{k}^{(m)}$ $W_{ijk}^{(m)} = W_{k}^{(m)}$	$\alpha_{ijk}^{(m)} = \alpha^{(m)}$ $W_{ijk}^{(m)} = W^{(m)}$
Correlated over origin	$\alpha_{ijk}^{(m)} = \alpha_{ijk}^{(m)}$ $W_{ijk}^{(m)} = W_{ik}^{(m)}$	$\alpha_{ijk}^{(m)} = \alpha_{ij}^{(m)}$ $W_{ijk}^{(m)} = W_{i}^{(m)}$
Correlated over destination	$\alpha_{ijk}^{(m)} = \alpha_{ijk}^{(m)}$ $W_{ijk}^{(m)} = W_{jk}^{(m)}$	$\alpha_{ijk}^{(m)} = \alpha_{ij}^{(m)}$ $W_{ijk}^{(m)} = W_{j}^{(m)}$
Fully correlated	$\alpha_{ijk}^{(m)} = \alpha_{ijk}^{(m)}$ $W_{ijk}^{(m)} = W_{k}^{(m)}$	$\alpha_{ijk}^{(m)} = \alpha_{ij}^{(m)}$ $W_{ijk}^{(m)} = W^{(m)}$

Notes: Origin state $i \neq m$, destination state $j \neq r$ and recurrent event $k \neq g$. Factor discrete model: $v_{ijk} = \prod_{m=1}^{M} \exp\big(\alpha_{ijk}^{(m)} \cdot W_{ijk}^{(m)}\big)$ with $W = \{-1, 1\}$ and $\Pr(W_{ijk}^{(m)} = 1) = p_{ijk}^{(m)}$.

4. Empirical illustration

The labour market performance of immigrants has received ample attention in the empirical literature. Neglected, however, is the question as to what extent the labour market performance affects the return migration of migrants. Labour market transitions and return migration of immigrants are intertwined, and should, therefore, be analysed in conjunction. We address these issues by using a unique administrative panel for the entire population of recent migrants and estimate a multistate model on the labour market and migration dynamics of these migrants.

To this end we use administrative data from the Netherlands, where we observe all immigrants who have entered the country between 1999 and 2007. Our data comprise the entire *population* of immigrants who entered during our observation window of 1999–2007, and after merging in other administrative registers, we obtain a panel. In addition to the date of entry and exit, the administration also records the migration motive of the individual. Either the motive is coded according to the visa status of the immigrant, or the immigrant reports the motive upon registration in the population register. See Bijwaard (2010) for an extensive descriptive analysis of the various migration motives. Here we focus exclusively on 94,270 labour migrants, which comprise about 23% of all non-Dutch immigrants in the age group 18–64 years.

This immigration register is linked by Statistics Netherlands to the Municipal Register of Population (Gemeentelijke Basisadministratie, GBA) and to their Social Statistical Database (SSD). The GBA contains basic demographic characteristics of the migrants, such as age, gender, marital status and country of origin. From the SSD we have information, on a monthly basis, on the labour market position, income, employment sector, housing and household situation.

4.1 Four states model for labour market and migration dynamics

We are interested, per se, in the labour market and the migration dynamics, the timing of the transitions and the time between transitions. Since we observe immigrants from the time they enter until the end of our observation window, and since we focus on those employed immigrants at entry, an immigrant potentially faces different risks of exiting his/her first state of employment and multiple durations. We define the following four states: (e) Employed in the Netherlands; (u) Unemployed and receiving benefits in the Netherlands; (n) Out of the labour market (and not receiving benefits = non-participating) in the Netherlands (NP); (a) Living abroad. Table 3 reports the observed transitions among these three labour market states and the living abroad state. Note that by the end of the observation window, 1-1-2008, all migrants are categorized in one of the four states.

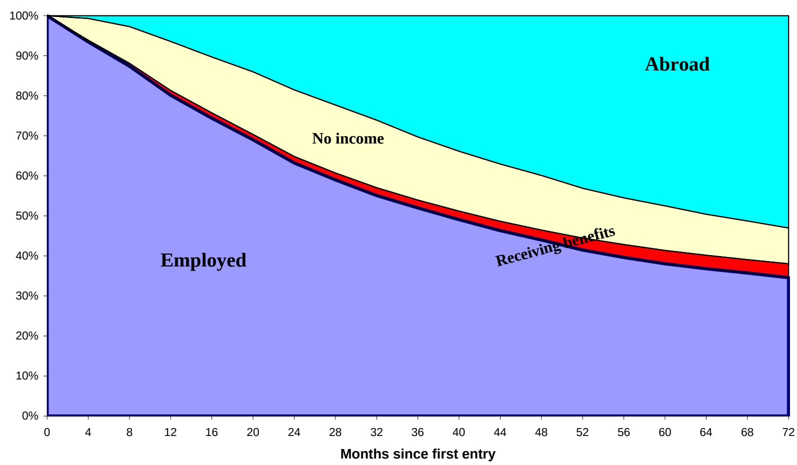
Table 3: Spell dynamics of the labour migrants (# 94,270)

		Percentage ending in			
	# of spell	employed	UI	NP	Abroad
Employed	124058	43%	5%	42%	10%
Unemployed (UI)	11898	49%	14%	33%	4%
Non-participation (NP)	56559	38%	9%	14%	39%
Abroad	45578	6%	0.2%	2%	92%

Source: Statistics Netherlands, based on own calculations.

Many employment spells are still continuing by the end of the observation period. For the majority of employment spells that end in a transition the migrant leaves the labour market, becomes non-participating. Many non-participation spells end abroad, while many unemployed return to employment. A third of the unemployed receiving benefits leave the labour market. Very few migrants leave the country directly from a state of unemployment. When a migrant leaves the country, they usually remain abroad; they are still abroad at the end of the observation period.

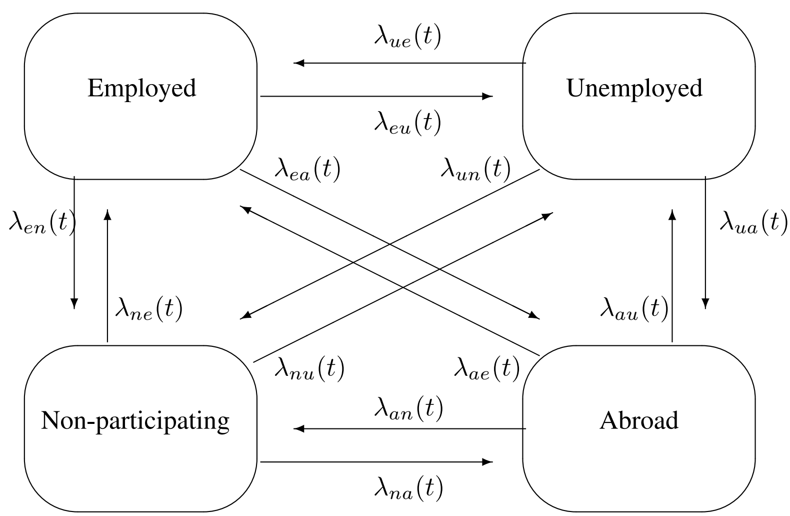
Figure 2: Development of SES of labour immigrants arriving in 1999



By definition all labour migrants start in the employed state at entry. Soon after arrival some migrants move to the other states. Some may return and some may move on to another state. But the migrant is always in one of the four states. In Figure 2 we depict the raw development over time of the distribution over the four states for the 1999-entry cohort, the labour migrants who arrived in 1999. The proportion abroad continuously increases. Six years after arrival more than 50% of the migrants have left the country. The remaining migrants are either employed or non-participating. Only a few migrants get unemployment benefits, possibly because they have not yet gained any benefit rights in the Netherlands.

We view the migrant behavior as a semi-Markov process, with individuals moving between the four states. These states are mutually exclusive and exhaust all possible destinations.[2] A migrant may leave a state $i = \{e, u, n, a\}$ for any of the other destination states. The 4-state multistate model is depicted in Figure 3.

Figure 3: Multistate model for labour-migration dynamics



[2] The mortality rate for the age range 18–64 is small enough to ignore deaths.

We use a competing risks hazard model for each origin-destination pair. We assume that recurrence of the states does not directly affect the frailty or the duration dependence, they are shared over recurrence. Recurrence influence on the hazard is captured by including the labour-migration history in the covariates. Define the random variables T_{ij} that describe the time since entry in i for a transition from i to j. We assume a (mixed) proportional hazard model for which the intensity for the transition from j to k is:

$$\lambda_{ij}(t|\overline{X}_{ij}(t), v_{ij}) = v_{ij}\lambda_{0ij}(t)\exp\bigl(\beta'_{ij}X_{ij}(t)\bigr) \tag{8}$$

where $\overline{X}_{ij}(t) = \{X_{ij}(s)|0 \leq s \leq t\}$ is the sample path of the observed characteristics up to time t, which is, without loss of generality, assumed to be left continuous. For the baseline duration $\lambda_{0ij}(t)$ we assume that it is piecewise constant on eleven intervals; every six months and beyond five years. Let the intervals $I_m(t) = I(t_{m-1} \leq t < t_m)$ for $m = 1, \ldots, M+1$ with $t_0 = 0$ and $t_{M+1} = \infty$ be the intervals on which we define the piecewise constant intensity. Then, the baseline intensity is $\lambda_{0ij}(t) = e^{\beta_{0ij}} \cdot \sum_{m=1}^{M+1} e^{\alpha_{mij}} I_m(t)$, with $\alpha_{1ij} = 0$. Thus β_{0ij} determines the intensity in the first interval. The α's determine the difference in intensity at each interval compared to this first interval. The baseline intensity for a duration of $t \in [t_{m-1}, t_m)$ is higher than the baseline intensity to leave for a duration of $t < t_1$ if $\alpha_{mij} > 0$ and lower if $\alpha_{mij} < 0$.

We use three different frailty models: (1) a PH model, a model without frailty (PH); (2) uncorrelated MPH model with a two-point discrete frailty (MPH); and (3) a two-factor loading correlated frailty over the origin state (correlated MPH). The covariates included in the model refer to demographic (gender, age-dummies, martial status and age of children), country of origin-dummies, individual labour market characteristics (monthly income, employment sector-dummies), labour market history and migration history. We control for business cycle conditions by including the national unemployment rate, both at the moment of first entry to the country and the time-varying monthly rate. The unemployment rate at entry captures the 'scarring effect' of migrants, while the running unemployment rate captures the impact of the business cycle on the transition intensities. With the abundant information on the migrants, the model contains many parameters. We used maximum likelihood estimation in STATA to estimate all the coefficients.[3] Here we only discuss the parameter estimates of the transition from employed to abroad, $\lambda_{ea}(t)$ and, focus on the differences induced by the alternative frailty assumptions.[4]

[3] The code is available upon request. The standard errors are calculated using the outer-product of the gradient vector in the estimated parameter vector. Other alternative estimation procedures for event history models with frailty are: the Expectation-Maximization (EM) algorithm, penalized partial likelihood, and Bayesian Markov Chain Monte Carlo methods.

[4] All estimation results are available from the author. Related multistate models on similar data are estimated by Bijwaard (2009) and Bijwaard and Wahba (2014). Bijwaard (2009) discusses calculation of the marginal

Table 4: Parameter estimates transition from employed to abroad, $\lambda_{ea}(t)$

	Independent		Correlated
	PH	MPH	MPH
female	-0.314^{**}	-0.377^{**}	-0.382^{**}
self-employed	-0.781^{**}	-0.976^{**}	-1.118^{**}
income $<$ 1000	0.692^{**}	0.917^{**}	0.971^{**}
income 1000–2000	0.155^{**}	0.206^{**}	0.237^{**}
income 3000–4000	0.178^{**}	0.182^{**}	0.174^{**}
income 4000–5000	0.152^{**}	0.169^{**}	0.163^{**}
income 5000–6000	0.306^{**}	0.321^{**}	0.334^{**}
income $>$ 6000	0.319^{**}	0.349^{**}	0.376^{**}
married	-0.147^{**}	-0.175^{**}	-0.220^{**}
divorced	-0.297^{**}	-0.383^{**}	-0.427^{**}
repeated entry	0.382^{**}	-0.388^{**}	-0.346^{**}
repeated unemployment	-0.376^{**}	-0.303^{*}	-0.419^{**}
Unemployment rate at entry	0.101^{**}	0.108^{**}	0.107^{**}
Unemployment rate	0.060^{**}	0.039^{**}	0.040^{**}
α_2 (6–12 months)	0.641^{**}	0.792^{**}	0.848^{**}
α_3 (12–18 months)	0.694^{**}	0.945^{**}	1.036^{**}
α_4 (18–24 months)	0.834^{**}	1.151^{**}	1.269^{**}
α_5 (24–30 months)	0.648^{**}	1.032^{**}	1.175^{**}
α_6 (30–36 months)	0.750^{**}	1.198^{**}	1.363^{**}
α_7 (36–42 months)	0.417^{**}	0.924^{**}	1.112^{**}
α_8 (42–48 months)	0.480^{*}	1.038^{**}	1.248^{**}
α_9 (48–54 months)	0.352^{**}	0.966^{**}	1.200^{**}
α_{10} (54–60 months)	0.263^{**}	0.921^{**}	1.180^{*}
α_{11} ($>$ 60 months)	-0.095^{*}	0.669^{**}	0.987^{*}
constant (β_0)	-6.118^{**}	-5.937^{**}	-5.850^{**}

Source: Statistics Netherlands, based on own calculations.

The estimated duration dependence and covariate effects of $\lambda_{ea}(t)$ are reported in Table 4. As expected, ignoring frailty biases the hazard of leaving the country towards

effects of the MPH in a multistate model with uncorrelated frailties. Bijwaard and Wahba (2014) estimate and discuss an extension of the multistate model of this paper in which the wage earned while employed is also correlated with the transition rates.

negative duration dependence. According to the PH model, the hazard of leaving the country from a state of employment five years after entry, α_{11}, has returned to the level of the hazard in the first six months, while according to the MPH model the hazard is then almost twice as high and according to the correlated MPH model, even 2.7 times as high. Thus, in the model without frailty, the migrants seem to become less prone to leave the longer they are in the country, while in the models with frailty this is much less the case. Allowing for correlation among all the three competing frailties starting in the employed state; employed to abroad, v_{ea}, employed to unemployed, v_{eu}, and employed to non-participation, v_{en}, increases the hazard duration dependence, the α's, of leaving even more.

The estimated duration dependence implies that the intensity of leaving increases with the duration of employment up till 3 years. After 3 years of employment in the host country, the intensity to leave slightly decreases. Including frailty has for some covariates a substantial effect. The effect of repeated entry even changes sign when allowing for frailty. Most covariate effects become more pronounced after allowing for frailty.[5]

Table 5 shows the comparison of the models. Of course, extending the parameter space of a model can never decrease the likelihood. We therefore include the AIC and BIC, which both penalize the number of included parameters. The results clearly show that the multistate model has many parameters: 786 using a PH model, 810 using an MPH model and 814 using a correlated MPH model. We can conclude that the correlated MPH model is the preferred model, because it has the lowest AIC and, except for the transition from abroad, the lowest BIC.

[5] We also carried out formal tests on equivalence of the parameters in the PH versus MPH model and on the equivalence of the parameters in the MPH versus the correlated MPH model. From these tests we can conclude that allowing for frailty significantly, on a 5%-level, changes the covariate effects and the duration dependence. Further allowing for correlation among the frailties only significantly changes the duration dependence.

Table 5: Model comparison

	# of parameters	log-likelihood	AIC	BIC
From *Employed*				
PH	249	–359860	720218	722641
MPH	255	–359538	719586	722066
cMPH[a]	256	–359252	719016	721506
From *Unemployed*				
PH	165	–37427	75183	76402
MPH	171	–37310	74962	76224
cMPH[a]	172	–37290	74925	76195
From *Non-participation*				
PH	183	–205288	410942	412579
MPH	189	–199903	400184	401875
cMPH[a]	190	–199724	399829	401528
From *Abroad*				
PH	189	–25474	51326	52975
MPH	195	–25465	51321	53023
cMPH[a]	196	–25459	51311	53021

Notes: [a] Correlated MPH.

Another test on the MPH and correlated MPH models is to check the significance of the variance and the correlation. For both the uncorrelated and the correlated frailty model we find that the frailty variance is significant on a 5%-level. In the correlated model we also find significant positive correlation between the three competing frailties from the employed state, see Table 6.[6] This implies that employed migrants who are more prone to become unemployed or non-participating are also more prone to leave. For each origin state we also tested whether the α's, the factor loadings, of each factor are equal. When the factor loadings are equal it implies that a shared frailty model, which restricts the correlation to one, is sufficient. However, for all four origin states, we reject this hypothesis.

[6] These correlations are derived from the 2 factor loading model and the standard errors are calculated using the delta method.

Table 6: Correlation between the unobserved heterogeneity terms

	v_{eu}	v_{en}	v_{ue}	v_{un}
v_{eu}	–			
v_{en}	0.999^{**}	–		
v_{ea}	0.940^{**}	0.939^{**}		
v_{ue}			–	
v_{un}			-0.997^{**}	–
v_{ua}			-0.470^{**}	0.539^{*}
	v_{ne}	v_{nu}	v_{ae}	v_{au}
v_{ne}	–			
v_{nu}	0.996^{**}	–		
v_{na}	-0.685^{**}	-0.735^{**}		
v_{ae}			–	
v_{au}			0.994^{**}	
v_{an}			0.312^{*}	0.214^{*}

Notes: $^{*}p < 0.05$ and $^{**}p < 0.01$.

These results indicate that for these data of recent labour migrants to the Netherlands it is important to include frailties and to allow these frailties to be correlated. More details on the data and on other analyses using these data on return behaviour and labour market transitions can be found in Bijwaard (2010), Bijwaard (2009), Bijwaard and Wahba (2014) and Bijwaard, Schluter, and Wahba (2014).

5. Identification issues in multistate frailty models

Associated with frailty models is a general identification problem of the logical possibility of decomposing the individual contributions to the average survival probability of the baseline duration dependence, the unobserved frailty, and the observed characteristics, given the observed data. Identifiability refers to the ability to uniquely estimate the parameters of the duration dependence, the regression function and the frailty distribution. More specifically, if the proportional hazard model were not identified, then it would be logically impossible to separate the individual contributions of duration dependence and frailty. In the econometric literature, the case of the univariate MPH model has been investigated in detail. Elbers and Ridder (1982) and Heckman and Singer (1984b) have established the identification of the MPH model under certain conditions, for an overview see Van den Berg (2001). The most important assumptions here are that the frailty has a finite mean and we have some exogenous variation in the observed characteristics. Ridder and Woutersen (2003) show that bounding the duration dependence hazard away from 0 and ∞ at the start is also sufficient for nonparametric identification of the MPH model, and with this assumption the finite mean assumption can be discarded.

Honoré (1993) shows that both the frailty distribution and the duration dependence are identified with multivariate event history data under much weaker assumptions. All shared frailty models are identified without additional information, such as observed covariates or parametric assumptions about the duration dependence. Furthermore, the duration dependence may depend on observed covariates in an unspecified way, and the frailty and the observed covariates may be dependent. This identifiability property holds for a broader class of frailty models, including correlated frailty models.

A caveat of multistate data is that such data is more sensitive to censoring. With univariate event history data, many types of censoring can be captured by standard adjustments to the likelihood function, see Andersen et al. (1993) and Klein and Moeschberger (2003). With sequential events, either recurrent or from different types, one has to be more careful. Consider two consecutive events with time t_1 and t_2, and where the data are subject to right-censoring at a fixed time after the starting point or the first event. Then the moment at which t_2 is right-censored is not independent from t_2 itself. For example, individuals with a large value of frailty will, on average, have a short time until the first event. As a result the time until the second event will start relatively early. This implies that the time until the second event will often be censored after a relatively longer period (or not censored at all). Thus, t_2 and its censoring probability are both affected by frailty. It may also happen that the process or some of the processes are not observed from the origin. With left-censoring, not to be confused with left-truncation, the analysis is more complicated, see Heckman and Singer (1984a) and Commenges (2002). In a Markov multistate model, defined in the time since the start of the process, in which censoring is independent of previous events and uses the same time scale, the censoring issues are similar to censoring issues in univariate event history models, see Andersen et al. (1993) and Aalen, Borgan, and Gjessing (2008).

A cautionary note should be given that for all these situations, identification is only possible when the model is a correctly specified mixed proportional hazards model. It is impossible to distinguish between a misspecified proportional hazard model and a correctly specified mixed proportional hazards model, see Putter and van Houwelingen (2011) for a discussion.

6. Summary and concluding remarks

This article has provided an overview of multistate event history models with frailty, with an emphasis on semi-Markov multistate models with a mixed proportional hazard structure. The literature on this subject is continuing and growing, and with the increased computer power the complexity of the models will not discourage researchers from using them. We have shown that ignoring frailty can have a large impact on the parameters of interest for the transition hazards, the duration dependence and the effect of observed covariates on the hazard. We discuss how different correlation structures of the frailties in a multistate model can be achieved.

Obviously, I did not intend to cover exhaustively all aspects of multistate frailty models. Many issues we did not address receive ample attention in the literature. An important observation is that the literature is highly segmented into mathematical research, biostatistical research, econometric research and demographic research. Although different terms are used, the problems addressed are similar, and the solutions are often very similar too. I advocate looking beyond the borders of your own research discipline to grasp the knowledge of the other fields.

References

Aalen, O.O., Borgan, Ø., and Gjessing, H.K. (2008). *Survival and Event History Analysis*. New York: Springer-Verlag. doi:10.1007/978-0-387-68560-1.

Andersen, P.K., Borgan, Ø., Gill, R.D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. New York: Springer-Verlag. doi:10.1007/978-1-4612-4348-9.

Andersen, P.K. and Gill, R.D. (1982). Cox's regression model for counting processes: A large sample study. *Annals of Statistics* 10(4): 1100–1120. doi:10.1214/aos/1176345976.

Bijwaard, G.E. (2009). Labour market status and migration dynamics. Discussion Paper No. 4530, IZA.

Bijwaard, G.E. (2010). Immigrant migration dynamics model for The Netherlands. *Journal of Population Economics* 23(4): 1213–1247. doi:10.1007/s00148-008-0228-1.

Bijwaard, G.E., Franses, P.H., and Paap, R. (2006). Modeling purchases as repeated events. *Journal of Business & Economic Statistics* 24(4): 487–502. doi:10.1198/073500106000000242.

Bijwaard, G.E., Schluter, C., and Wahba, J. (2014). The impact of labour market dynamics on the return–migration of immigrants. *Review of Economics and Statistics*, forthcoming.

Bijwaard, G.E. and Wahba, J. (2014). Do high-income or low-income immigrants leave faster? *Journal of Development Economics* 108: 54–68. doi:10.1016/j.jdeveco.2014.01.006.

Clayton, D. (1978). A model for the association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65(1): 141–151. doi:10.1093/biomet/65.1.141.

Commenges, D. (1999). Multi-state models in epidemiology. *Lifetime Data Analysis* 5(4): 315–327. doi:10.1023/A:1009636125294.

Commenges, D. (2002). Inference for multi-state models from interval-censored data. *Statistical Methods in Medical Research* 11(2): 167–182. doi:10.1191/0962280202sm279ra.

Cook, R.J. and Lawless, J.F. (2007). *The Statistical Analysis of Recurrent Events*. New York: Springer-Verlag.

Courgeau, D. and Lelièvre, E. (1992). *Event History Analayis in Demography*. Oxford: Clarendon Press.

Duchateau, L. and Janssen, P. (2008). *The Frailty Model*. New York: Springer-Verlag.

Duchateau, L., Janssen, P., Kezic, I., and Fortpied, C. (2003). Evolution of recurrent asthma event rate over time in frailty models. *Applied Statistics* 52(3): 355–363. doi:10.1111/1467-9876.00409.

Elbers, C. and Ridder, G. (1982). True and spurious duration dependence: The identifiability of the proportional hazard model. *Review of Economic Studies* 49(3): 403–409. doi:10.2307/2297364.

Flinn, C.J. and Heckman, J.J. (1983). Are unemployment and out of the labor force behaviorally distinc labor force states? *Journal of Labor Economics* 1(1): 28–42. doi:10.1086/298002.

Fougère, D. and Kamionka, T. (2008). Econometrics of individual labor market transitions. In: Mátyás, L. and Sevestre, P. (eds.). *The econometrics of panel data, Fundamentals and recent developments in theory and practice*. Princeton: Princeton University Press: 865–905.

Govindarajulu, U.S., Lin, H., Lunetta, K.L., and D'Agostino, R.B. (2011). Frailty models: Applications to biomedical and genetic studies. *Statistics in Medicine* 30(22): 2754–2764. doi:10.1002/sim.4277.

Heckman, J.J. and Singer, B. (1984a). Econometric duration analysis. *Journal of Econometrics* 24(1–2): 63–132. doi:10.1016/0304-4076(84)90075-7.

Heckman, J.J. and Singer, B. (1984b). The identifiability of the proportional hazard model. *Review of Economic Studies* 51(2): 231–241. doi:10.2307/2297689.

Honoré, B.E. (1993). Identification results for duration models with multiple spells. *Review of Economic Studies* 60(1): 241–246. doi:10.2307/2297821.

Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. New York: Springer-Verlag. doi:10.1007/978-1-4612-1304-8.

Kelly, P.J. and Lim, L.Y. (2000). Survival analysis for recurrent event data: An application to childhood infectious diseases. *Statistics in Medicine* 19(1): 13–33. doi:10.1002/(SICI)1097-0258(20000115)19:1<13::AID-SIM279>3.0.CO;2-5.

Klein, J.P. and Moeschberger, M.L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data (2nd edition)*. New York: Springer-Verlag.

Lancaster, T. (1979). Econometric methods for the duration of unemployment. *Econometrica* 47(4): 939–956. doi:10.2307/1914140.

Manda, S.O.M. (2001). A comparison of methods for analysing a nested frailty model to child survival in Malawi. *Australian & New Zealand Journal of Statistics* 43(1): 7–16. doi:10.1111/1467-842X.00150.

Nielsen, G.G., Gill, R.D., Andersen, P.K., and Sørensen, T.A.I. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics* 19: 25–43.

Oakes, D.A. (1992). Frailty models for multiple event times. In: Klein, J.P. and Goel, P.K. (eds.). *Survival Analysis: State of the Art*. Dordrecht: Kluwer: 371–379. doi:10.1007/978-94-015-7983-4_22.

Pickles, A. and Crouchley, R. (1995). A comparison of frailty models for multivariate survival data. *Statistics in Medicine* 14(13): 1447–1461. doi:10.1002/sim.4780141305.

Prentice, R.L., Williams, B.J., and Peterson, A.V. (1981). On the regression analysis of multivariate failure time data. *Biometrika* 68(2): 373–379. doi:10.1093/biomet/68.2.373.

Putter, H., Fiocco, M., and Geskus, R.B. (2007). Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine* 26(11): 2389–2430. doi:10.1002/sim.2712.

Putter, H. and van Houwelingen, H.C. (2011). Frailties in multi-state models: Are they identifiable? Do we need them? *Statistics Methods in Medical Research* in press. doi:10.1177/0962280211424665.

Ridder, G. and Woutersen, T. (2003). The singularity of the efficiency bound of the mixed proportional hazard model. *Econometrica* 71(5): 1579–1589. doi:10.1111/1468-0262.00460.

Rogers, A. (1975). *Introduction to Multiregional Mathematical Demography*. New York: Wiley.

Rogers, A. (1995). *Multiregional Demography: Principles, Methods and Extensions*. New York: Wiley.

Rondeau, V., Commenges, D., and Joly, P. (2003). Maximum penalized likelihood estimation in a gamma–frailty model. *Lifetime Data Analysis* 9(2): 139–153. doi:10.1023/A:1022978802021.

Sastry, N. (1997a). Family-level clustering of childhood mortality risks in northeast Brazil. *Population Studies* 51(3): 245–261. doi:10.1080/0032472031000150036.

Sastry, N. (1997b). A nested frailty model for survival data, with an application to the study of child survival in northeast Brazil. *Journal of the American Statistical Associ-*

ation 92(438): 426–435. doi:10.1080/01621459.1997.10473994.

Therneau, T. and Grambsch, P. (2000). *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag. doi:10.1007/978-1-4757-3294-8.

Van den Berg, G.J. (2001). Duration models: Specification, identification, and multiple duration. In: Heckman, J. and Leamer, E. (eds.). *Handbook of Econometrics, Volume V*. Amsterdam: North-Holland: 3381–3460.

Vaupel, J.W., Manton, K.G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16(3): 439–454. doi:10.2307/2061224.

Wei, L.J., Lin, D.Y., and Weissfeld, L. (1989). Regression analysis of multivariate failure time data by modeling marginal distributions. *Journal of the American Statistical Association* 84(408): 1065–1073. doi:10.1080/01621459.1989.10478873.

Wienke, A. (2011). *Frailty Models in Survival Analysis*. Boca Raton: Chapman & Hall/CRC.

Willekens, F.J. (1999). Life course: Models and analysis. In: Dykstra, P.A. and van Wissen, L.J.G. (eds.). *Population Issues: An Interdisciplinary Focus*. New York: Plenum Press: 23–51. doi:10.1007/978-94-011-4389-9_2.

Xue, X. and Brookmeyer, R. (1996). Bivariate frailty model for the analysis of multivariate survival time. *Lifetime Data Analysis* 2(3): 277–290. doi:10.1007/BF00128978.

Yashin, A.I., Vaupel, J.W., and Iachine, I.A. (1995). Correlated individual frailty: An advantageous approach to survival analysis of bivariate data. *Mathematical Population Studies* 5(2): 145–159. doi:10.1080/08898489509525394.

Yau, K.K.W. and McGilchrist, C.A. (1998). ML and REML estimation in survival analysis with time dependent correlated frailty. *Statistics in Medicine* 17(11): 1201–1213. doi:10.1002/(SICI)1097-0258(19980615)17:11<1201::AID-SIM845>3.0.CO;2-7.