Intensity-based premium evaluation for unemployment insurance products

Francesca Biagini[a,1,*], Andreas Groll[a], Jan Widenmann[a,2]

[a] *Department of Mathematics, University of Munich, Theresienstraße 39, 80333 Munich, Germany.*

Abstract

We present a flexible premium determination method for insurance products, in particular for unemployment insurance products. The price is determined with the real-world pricing formula and under the assumption that the employment-unemployment progress of an insured person follows an $\mathbb{F}$-doubly stochastic Markov chain. The stochastic intensity processes are estimated for the German labor market, using Cox's proportional hazards model with time-dependent covariates on a sample of integrated labor market biographies. The estimation procedure is based on a counting process framework with stochastic compensators, which we show to be naturally connected to the class of $\mathbb{F}$-doubly stochastic Markov chains. Based on the statistical analysis, the prices are computed using Monte Carlo simulations.

Keywords: Unemployment insurance, intensity-based model, $\mathbb{F}$-doubly stochastic Markov chain, Cox proportional hazards model, benchmark approach
2010 MSC: 62N01, 62N02, 62N05, 62P05
JEL: J11, J64, J65, C14

1. Introduction

The debt crisis in the Euro zone with Cyprus, Greece, Ireland, Portugal and Spain having applied for financial assistance of the European Stability Mechanism (ESM) or the European Financial Stability Facility (EFSF), respectively, constitutes an immense challenge for Europe and the rest of the world. One of the core structural problems of the affected countries is a troubled labor market

*Corresponding author. Tel.: +49 (0)89 2180 4492
Email addresses: `francesca.biagini@math.lmu.de` (Francesca Biagini), `andreas.groll@math.lmu.de` (Andreas Groll), `jan.widenmann@math.lmu.de` (Jan Widenmann)

[1] The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no [228087].
[2] The Ph.D. position of Jan Widenmann at the University of Munich is gratefully supported by Swiss Life Insurance Solutions AG.

with unemployment rates around 25%. Beside the financial drawbacks for the unemployed people, these high levels of unemployment rates burden the public unemployment insurance systems as well as the private insurance sector, which has started to introduce special products against unemployment. The demand on modern, well elaborated and tested mathematical models for premium determination and risk mitigation for these kind of insurance products, but also other insurance products in general, is high and an ongoing field in actuarial research, see e.g. Bacinello et al. (2009), Biagini and Schreiber (2012), Biagini and Widenmann (2012, 2013) or Møller (1998). In times, where all sectors of an economy are closely connected, one main issue in this context is to consider the insurance market as part of a hybrid market, consisting among others of stocks, equities, commodities, fixed income and insurance products, all influenced by micro- and macro-economic factors. Hence, the correlations and dependencies of models for (unemployment) insurance products and other sources of randomness in hybrid markets need to be investigated thoroughly.

In this context, the present paper aims to introduce a flexible premium determination framework for unemployment insurance products, particularly for so called payment protection insurances (PPIs): given some underlying payment obligation of the insured person, e.g. a loan, the insurance company will take over the instalments during an unemployment period. In this way, the financial challenges for the insured persons and the credit default risk for the creditor are both reduced at the same time.

Our pricing model is generally based on a two-state switching process with state space $\{1, 2\}$ ($1 \widehat{=}$ employed, $2 \widehat{=}$ unemployed), generated by two stochastic intensity processes. Generally speaking, the intensity of a transition from employment to unemployment at time t characterizes the conditional instantaneous probability at time t for an employee to become unemployed, given the currently available information. The intensity of a transition from unemployment to employment is interpreted analogously. In regard to the aforementioned dependencies of the model in hybrid markets, we assume the intensity processes to be driven by individual-related as well as macro- and micro-economic covariate processes, see Equation (1) in Section 2.

An adequate class for the two-state switching process, which allows for stochastic intensity processes is the class of $\mathbb{F}$-doubly stochastic Markov chains, introduced by Jakubowski and Niewęgłowski (2010). It extends the notion of classical Markov chains.

As general pricing rule we adopt the real-world pricing formula, see Platen and Heath (2007). A first model, using $\mathbb{F}$-doubly stochastic Markov chains and the benchmark approach for pricing PPIs, is proposed by Biagini and Widenmann (2012). However, the intensities there are assumed to be random, but not varying over time. In the present paper we extend this approach in order to address the more realistic case of stochastic intensity processes. However, in this case it is generally not possible to obtain an analytical expression for the insurance premium which will be here computed by using Monte Carlo simulations.

In order to calibrate the price for the unemployment insurance products to real data, we esti-

mate the intensity processes using Cox's proportional hazards model, see Cox (1972, 1975) and Andersen et al. (1993). Our data set is provided by the "Institut für Arbeitsmarkt- und Berufsforschung" (IAB) and contains a sample of integrated labor market biographies, including the duration of employment and unemployment periods between 1975-2008 of more than 1.5 million German individuals as well as several useful socio-demographic covariates, such as age, nationality, educational level, regional details, etc. In order to reflect additional dependencies of the intensity processes of macro-economic factors, we also incorporate further covariates such as time series for MSCI-world returns and German unemployment rates.

An advantage of using Cox's proportional hazards model is the availability of adequate implementations, see for example the R-packages corresponding to de Wreede et al. (2010), Jackson (2011) or Aalen et al. (2004); a Bayesian approach is proposed by Kneib and Hennerfeind (2008) and implemented in the software BayesX. Here, a major difficulty is to adequately operationalize the data set with regard to the software packages.

Technically, the implemented estimators are based on the theory on multivariate counting processes and their compensators, where the counting process is assumed to count subsequent jumps of the same kind of an underlying multi-state switching process. Given the martingale property of the compensated counting process, estimators for the compensators can be derived. In the present paper we extend the existing theory for (classical) Markov chains[3] and show that the class of $\mathbb{F}$-doubly stochastic Markov chains is the natural candidate for the underlying multi-state switching processes corresponding to the multivariate counting processes with stochastic compensators of the form given by Cox's proportional hazards model. This relation can easily be extended to general multiplicative hazard models as given in Andersen et al. (1993).

In order to test the obtained estimation results, we apply conventional goodness-of-fit methods. The results generally show adequate performance of the estimated model parameters. Moreover, we introduce a further, non-standard method for testing the applicability of the obtained parameters with respect to prediction, by comparing actual and simulated jump times for selected paths from the data set. The results here show good predictive power, which implicates robustness of the Monte Carlo simulations to compute the premiums. A sensitivity analysis of the insurance premiums also confirms these findings.

Our approach, therefore, represents a flexible premium determination tool for unemployment insurance products since it takes into account various risk factors. Moreover, it can be easily adapted to model and estimate stochastic intensities and dependence structures in many other different applications of financial and actuarial practice as well as from other fields.

The rest of the paper is structured as follows. In Section 2 we introduce the form of the considered

[3]The classical time-inhomogeneous Markov chains generally have deterministic matrix-valued intensity functions of time. The corresponding counting processes also have deterministic compensators.

unemployment insurance products and derive the pricing formula. The connection between the multivariate counting processes of Cox's proportional hazards model and the class of $\mathbb{F}$-doubly stochastic Markov chains is established in Section 3. Additionally, the data set is described and estimation results for the intensity processes are presented. The Monte Carlo simulation procedure is explained in detail in Section 4 followed by a sensitivity analysis of the insurance premiums.

2. Unemployment insurance contracts

We now give a brief overview about the characteristics of unemployment insurance contracts. The product's basic idea is that the insurance company compensates to some extend the financial deficiencies, which an unemployed insured person is exposed to. We only consider contracts with deterministic, a priori fixed claim payments c_i which possibly take place at predefined payment dates T_i, $i = 1, ..., N$. Hence, the randomness of the claims is only due to their occurrence and not to their amount. As a practical example, one could think of PPI products against unemployment, which are linked to some payment obligation of an obligor to its creditor. The claim amount here is defined by the (a priori known) instalments, which are paid at predefined payment dates.
In order to conclude the insurance contract, the insured person must be employed at least for a certain period before the beginning of the contract. We therefore assume that all insured persons are employed at $t = 0$ almost surely. The contract's exclusion clauses define three time periods, specifying whether the insured person is entitled to receive a claim payment or not:

- The *waiting period* starts with the beginning of the contract. If an insured person becomes unemployed at any time of this period, she is not entitled to receive any claim payments during this unemployment period.

- The *deferment period* starts with the first day of unemployment. An unemployed insured person is not entitled to receive claim payments until the end of this period.

- The third period is comparable with the waiting period and is called the *requalification period*. The requalification period starts, if the insured person is re-employed after a stage of unemployment during the contract's duration. If the insured person becomes unemployed (again) at any time of the requalification period, she is not entitled to receive any claim payment during the whole unemployment period.

For existing unemployment insurance contracts, the waiting, deferment and requalification periods currently vary from three to twelve months.
We now provide the quantitative framework for determining the premium of unemployment insurance contracts. Note that on the market, there exist several modifications of the described insurance contract. Moreover, on both financial and insurance markets there exist similar products that can be investigated analogously. In these cases, it may be necessary to modify the following framework

accordingly. As the described procedure is based on Monte Carlo simulations, this can be done easily by adjusting the respective code segments from case to case.

Let $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ be a complete, filtered probability space, where the filtration $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]}$ is assumed to fulfill the usual conditions of completeness and right-continuity, see Protter (2003), with some arbitrary time horizon $T > 0$. Moreover, let $X = (X(t))_{t\in[0,T]}$ be a right-continuous stochastic process with state space $\{1, 2\}$. Here, the state "1" shall express that an insured person is employed and the state "2" that she is unemployed. Hence, X represents the employment-unemployment progress of a person in time. We denote by $\mathbb{F}^X$ the natural filtration generated by X, i.e. $\mathcal{F}_t^X = \sigma(X(u) : u \leq t)$ for all $t \in [0, T]$, and assume $\mathbb{G} = \mathbb{F}^X \vee \mathbb{F}^Z$. Here, $\mathbf{Z} = \big((Z^1(t), ..., Z^p(t))\big)^\intercal_{t\in[0,T]}$ is a p-variate predictable process of covariates, representing both individual-related and micro- as well as macro-economic risk factors, influencing the model as we will see more accurately in Section 3. We then assume X to follow an $\mathbb{F}^Z$-doubly stochastic Markov chain[4] with matrix valued intensity process $\mathbf{\Psi} = (\Psi_t)_{t\in[0,T]}$, whose entries $\alpha_{j,k}(t)$, $j, k \in \{1, 2\}$, $j \neq k$, at time t admit the representation

$$\alpha_{j,k}(t) = \alpha^{jk}(Z^1(t), ..., Z^p(t)), \tag{1}$$

for measurable functions $\alpha^{jk} : (\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p)) \to (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. We define the jump times of X from one state to the other as

- $\tau_0 := 0$
- $\tau_n := \inf\{\tau_{n-1} < t \leq T : X(t) \neq X(t-)\} \qquad n \geq 1\ ,$

where $X(t-) := \lim_{s\nearrow t} X(s)$. Moreover, we use W, D, R to express the length of the waiting, deferment and requalification period, respectively. We think of $t = 0$ as the beginning of the contract and therefore set $X(0) = 1$ $\mathbb{P}$-a.s.

We now describe the payoff of the contract. The insurance company has to pay the claim to the amount of c_i at time T_i if the following conditions are satisfied:

- $W < \tau_1 \leq T_i - D$
  The first jump τ_1 to unemployment of the insured person must occur *after* the waiting period W. Moreover, at least the deferment period D must lie between τ_1 and the payment date T_i, i.e. $T_i - \tau_1 \geq D$.
  and

- $T_i < \tau_2$
  The insured person must not have jumped back to employment *before* the payment date T_i.

OR for $j \geq 2$

[4]For the definition and properties of $\mathbb{F}^Z$-doubly stochastic Markov chains, we refer to AppendixA.

- $\tau_{2j-1} - \tau_{2j-2} > R$
  At least the requalification period R must lie in between a jump τ_{2j-2} to employment and the next jump τ_{2j-1} back to unemployment.
  and

- $W < \tau_{2j-1} \leq T_i - D$
  Any jump to unemployment τ_{2j-1} must occur after the waiting period W. Moreover, at least the deferment period D must lie between τ_{2j-1} and the payment date T_i, i.e. $T_i - \tau_{2j-1} \geq D$.
  and

- $\tau_{2j} > T_i$
  Before the payment date T_i the insured person must not have jumped back to employment.

Based on this, the random insurance claim C_i at the payment date T_i is modeled as

$$C_i(\omega) := c_i \mathbb{1}_{\{W<\tau_1 \leq T_i - D, \tau_2 > T_i\} \cup \bigcup_{j=2}^{\infty} \{\tau_{2j-1} - \tau_{2j-2} > R, W < \tau_{2j-1} \leq T_i - D, \tau_{2j} > T_i\}}(\omega) \ .$$

Note that due to the definition of the jump times the events $\{W < \tau_{2j-1} \leq T_i - D\} \cap \{\tau_{2j} > T_i\}$, $j \geq 1$, are disjoint. Therefore, we can rewrite the random insurance claim as

$$C_i(\omega) := c_i \mathbb{1}_{\{W<\tau_1 \leq T_i - D, \tau_2 > T_i\}} + \sum_{j=2}^{\infty} \mathbb{1}_{\{\tau_{2j-1} - \tau_{2j-2} > R, W < \tau_{2j-1} \leq T_i - D, \tau_{2j} > T_i\}}(\omega) \ .$$

In order to determine the premium for this kind of insurance contracts, we apply the real-world pricing formula[5] as described in Biagini and Widenmann (2012). For a motivation of the use of this approach for pricing in hybrid markets, we refer to Remark AppendixB.4 of AppendixB. With the real-world pricing formula (B.2), the price $P_t(C_i)$ of C_i at time $t \in [0, T_i]$ is hence given as

$$\begin{aligned} P_t(C_i) = S_t^{\delta^*} c_i \Bigg(& \mathbb{E}\left[\frac{1}{S_{T_i}^{\delta^*}} \mathbb{1}_{\{W<\tau_1 \leq T_i - D, \tau_2 > T_i\}} \middle| \mathcal{G}_t \right] \\ & + \sum_{j=2}^{\infty} \mathbb{E}\left[\frac{1}{S_{T_i}^{\delta^*}} \mathbb{1}_{\{\tau_{2j-1} - \tau_{2j-2} > R, W < \tau_{2j-1} \leq T_i - D, \tau_{2j} > T_i\}} \middle| \mathcal{G}_t \right] \Bigg) \ , \end{aligned}$$

where S^{δ^*} represents the $\mathbb{P}$-numéraire portfolio[6], used as benchmark for the market and as discounting factor.
Now we can sum up over all payment dates T_i to receive the (overall) insurance premium P_t at

[5] We introduce and motivate the real-world pricing formula in AppendixB.
[6] For the definition of the $\mathbb{P}$-numéraire portfolio, we refer again to AppendixB.

time $t \in [T_{k-1}, T_k)$, $k \geq 1$,[7] given by

$$P_t = \sum_{i=k}^{N} S_t^{\delta^*} c_i \Bigg(\mathbb{E}\left[\frac{1}{S_{T_i}^{\delta^*}} \mathbb{1}_{\{W<\tau_1 \leq T_i - D, \tau_2 > T_i\}} \Big| \mathcal{G}_t \right] + \sum_{j=2}^{\infty} \mathbb{E}\left[\frac{1}{S_{T_i}^{\delta^*}} \mathbb{1}_{\{\tau_{2j-1}-\tau_{2j-2}>R, W<\tau_{2j-1} \leq T_i - D, \tau_{2j} > T_i\}} \Big| \mathcal{G}_t \right] \Bigg) .$$

An insurance company is mainly interested in the insurance premium P_0 at time $t = 0$. Hence, for our further calculations, we only consider

$$P_0 = \sum_{i=1}^{N} c_i \Bigg(\mathbb{E}\left[\frac{1}{S_{T_i}^{\delta^*}} \mathbb{1}_{\{W<\tau_1 \leq T_i - D, \tau_2 > T_i\}} \right] + \sum_{j=2}^{\infty} \mathbb{E}\left[\frac{1}{S_{T_i}^{\delta^*}} \mathbb{1}_{\{\tau_{2j-1}-\tau_{2j-2}>R, W<\tau_{2j-1} \leq T_i - D, \tau_{2j} > T_i\}} \right] \Bigg) . \tag{2}$$

In order to determine (2), it is necessary to investigate the joint distributions of $S_{T_i}^{\delta^*}$, $i = 1, ..., N$, and the *sojourn-times* $\tau_j - \tau_{j-1}$, $j \geq 1$. Note that a major advantage of the pricing formula (2) is that these joint distributions can be directly investigated under the real-world probability measure. A first model in this context is given in Biagini and Widenmann (2012), where the underlying jump process X is considered to follow a special type of $\mathbb{F}$-doubly stochastic Markov chain with random but not time-varying intensity matrix. Here we extend this approach in order to address the more realistic case of stochastic intensity processes. Since in this setting it is in general not possible to obtain an analytical expression for P_0 as in Biagini and Widenmann (2012), we compute the insurance premium (2) by using Monte Carlo approximations and solve the problem of describing dependencies among $\mathbb{P}$-numéraire portfolio and sojourn-times as follows. By Bielecki and Rutkowski (2004) and by considering the insured person to be employed at the contract's beginning, the jump times $(\tau_j)_{j \in \mathbb{N}}$ are defined in terms of the intensities as

- $\tau_0 = 0$,
- $\tau_j = \inf\{\tau_{j-1} \leq t \leq T : e^{-\int_{\tau_{j-1}}^{t} \alpha_{1+((j+1) \bmod 2), 2-((j+1) \bmod 2)}(s) ds} \leq U_j\}$, $\quad j \geq 1$, $\qquad$ (3)

where $(U_j)_{j \geq 1}$ is a sequence of mutually independent uniformly distributed random variables on $[0, 1]$, which are also chosen to be independent of the intensity processes. Since under our model assumption (1) the covariate processes drive the randomness of the intensities, they hence also determine the sequence $(\tau_j)_{j \in \mathbb{N}}$ of the jump times. In particular, if the $\mathbb{P}$-numéraire portfolio S^{δ^*} or its monthly returns, respectively, are considered as one of the covariate processes, we intrinsically

[7] We set $T_0 := 0$ here.

obtain the required dependency structure between S^{δ^*} and the sequence of jump times. The other covariates can be any risk factors of interest such as individual related characteristics like sex, educational level, job sector, etc. or other external risk factors like monthly unemployment rates. In this way, we obtain a highly flexible premium determination framework for unemployment insurance products, which can be adjusted according to various characteristics of the insured person and to external micro- and macro-economic risk factors.
This method can also be applied to capture the dependency structures affecting intensity processes in other kinds of applications.
In the following section we apply the well-known and elaborated Cox's proportional hazards model to estimate the intensity processes on a real data set provided by the IAB.

3. Estimation method of the intensity processes

The development of suitable statistical models for general multi-state switching processes with applications e.g. in biomedicine or econometrics, in which time-to-event variables are analyzed, has been an important task in statistics for a long time. Recently, there have been great efforts also on the implementational side, see Aalen et al. (2004), Kneib and Hennerfeind (2008), Simon et al. (2011), Jackson (2011) and de Wreede et al. (2010). The underlying estimators, implemented in these software packages, are based on the theory of multivariate counting processes and their compensators, see Andersen et al. (1993). This is due to the fact that any multi-state switching process generates a multivariate counting process by counting successively the jumps of the same type over time. An important question in this context is how the properties of the multi-state switching process and the counting process relate to each other. It is well known for a classical Markov chain that the compensator of the corresponding counting process is determined by the Markov chain's matrix-valued (deterministic) intensity function, see e.g. Andersen et al. (1993). Any estimator for the deterministic compensator is therefore implicitly an estimator for the intensity function of the Markov chain. The derivation of suitable estimators for a counting process' compensator is based on a (partial) likelihood function, which can be derived very generally for any compensator of a multivariate counting process. Hence, it is natural to extend this method to elaborate statistical models for general multivariate counting processes with stochastic compensators of various forms, see Andersen et al. (1993) or Therneau and Grambsch (2000). The remaining question is then whether it is still possible to characterize properties of the underlying multi-state switching processes appropriately and derive e.g. estimators for their intensities. Although estimators for stochastic compensators are used frequently and are often claimed to provide stochastic intensity processes for some underlying multi-state switching process, a thorough investigation on the relations between these processes is, to the best of our knowledge, missing in the statistical literature so far.

We now fill in this gap and show in Section 3.1 how stochastic compensators of the form given in Cox's proportional hazards model are naturally linked to a sub-class of $\mathbb{F}$-doubly stochastic Markov chains and their intensities by using the results of Theorem AppendixA.5. Therefore, the estimators, which are implemented in the above mentioned software packages, also represent suitable estimators for the matrix-valued stochastic intensity processes of the underlying $\mathbb{F}$-doubly stochastic Markov chains.

3.1. Counting processes for $\mathbb{F}$-doubly stochastic Markov chains

The most frequently used model in survival analysis which is able to incorporate covariates belongs to the class of *multiple hazard models* and is known as the *Cox model*, see Cox (1972, 1975), implemented in the R-function `coxph` from the R-package `survival`, see Therneau (2012). The function fits a Cox proportional hazards regression model including time dependent covariates, time dependent strata, and multiple events per subject. The `survival` package uses the estimators as derived in Andersen et al. (1993), which are based on the general theory of multivariate counting processes and their compensators or, more precisely, on the compensators' densities with respect to Lebesgue's measure in the absolutely continuous case.

We now describe Cox's proportional hazards model as given in Andersen et al. (1993, Section VII.2.) for the special case when successive jumps of n individuals are counted over time and any individual can perform two types of jump: from employment to unemployment and vice versa.

On the probability space $(\Omega, \mathcal{G}, \mathbb{P})$, we consider a $2n$-variate counting process

$$\mathbf{N} = (\mathbf{N}(t))_{t\in[0,T]} = \left(N_1^{12}(t), N_1^{21}(t), ..., N_n^{12}(t), N_n^{21}(t)\right)^{\intercal}_{t\in[0,T]} \ ,$$

where, $N_i^{jk}(t)$, $i \in \{1, ..., n\}$, $j, k \in \{1, 2\}$, $j \neq k$, is supposed to count the respective jumps of the i-th observed person from employment to unemployment, denoted by the superscript "12", or from unemployment to employment, denoted by the superscript "21", respectively, up to time t. The $2n$-variate process

$$\mathbf{\Lambda}(\boldsymbol{\theta}) = (\mathbf{\Lambda}(t, \boldsymbol{\theta}))_{t\in[0,T]} = \left(\Lambda_1^{12}(t, \boldsymbol{\theta}), \Lambda_1^{21}(t, \boldsymbol{\theta}), ..., \Lambda_n^{12}(t, \boldsymbol{\theta}), \Lambda_n^{21}(t, \boldsymbol{\theta})\right)^{\intercal}_{t\in[0,T]}$$

is assumed to be the compensator of $\mathbf{N}$ with respect to the filtration $\hat{\mathbb{G}} = (\hat{\mathcal{G}}_t)_{t\in[0,T]}$, with $\hat{\mathcal{G}}_t = \mathcal{F}_t^N \vee \bigvee_{i=1}^n \mathcal{I}^i \vee \mathcal{A}$. Here, $\mathcal{F}_t^N = \sigma(\mathbf{N}(u) : u \leq t)$ is the σ-algebra, generated by $\mathbf{N}$ up to time t, $\mathcal{I}^i$ is a σ-algebra related to individual i, and $\mathcal{A}$ some arbitrary σ-algebra. The filtration $\hat{\mathbb{G}}$, interpreted as the level of information at any time $t \in [0, T]$, is hence determined by the information $\mathcal{F}_t^N$, generated by the observations of $\mathbf{N}$ up to and including time t, by additional i-specific information $\mathcal{I}^i$ of all individuals $i \in \{1, ..., n\}$, which is not changing over time, and by some other information

$\mathcal{A}$, neither changing from individual to individual nor over time[8]. Note that the compensator is also supposed to depend on a parameter vector $\boldsymbol{\theta}$.

A desirable feature for statistical models is then that for every person $i \in \{1, ..., n\}$, the i-th component of $\boldsymbol{\Lambda}(\boldsymbol{\theta})$, i.e. the 2-variate process $\boldsymbol{\Lambda}_i(\boldsymbol{\theta}) = (\Lambda_i^{12}(t, \boldsymbol{\theta}), \Lambda_i^{21}(t, \boldsymbol{\theta}))^\intercal_{t\in[0,T]}$, is given by the compensator of the counting processes $\mathbf{N}_i = (N_i^{12}(t), N_i^{21}(t))^\intercal_{t\in[0,T]}$ with respect to the reduced filtration $\hat{\mathbb{G}}^i = (\hat{\mathcal{G}}_t^i)_{t\in[0,T]}$ with $\hat{\mathcal{G}}_t^i = \mathcal{F}_t^{N_i} \vee \mathcal{I}^i \vee \mathcal{A}$. A sufficient condition for this to hold is that for every $t \in [0, T]$, the family $\left(\mathcal{F}_t^{N_i} \vee \mathcal{I}^i\right)_{i=1,...,n}$ is conditionally independent given $\mathcal{A}$. This yields a particular robustness of the statistical model since in this case the estimation of the compensator is independent of the number of observed individuals.

For every $i \in \{1, ..., n\}$, $j, k \in \{1, 2\}$, $j \neq k$, $t \in [0, T]$, we assume the component $\Lambda_i^{jk}(t, \boldsymbol{\theta})$ of $\boldsymbol{\Lambda}(\boldsymbol{\theta})$ to have a density, i.e. a non-negative $\hat{\mathbb{G}}$-predictable process $(\lambda_i(t, \boldsymbol{\theta}))_{t\in[0,T]}$, such that

$$\Lambda_i^{jk}(t, \boldsymbol{\theta}) = \int_0^t \lambda_i^{jk}(s, \boldsymbol{\theta}) ds \ , \quad t \in [0, T] \, .$$

Moreover we assume the densities to be of Cox's multiplicative form

$$\lambda_i^{jk}(t, \boldsymbol{\theta}) = Y_i^{jk}(t) \alpha_0^{jk}(t, \phi) e^{(\boldsymbol{\beta}^{jk})^\intercal \mathbf{Z}_i(t)} \ , \tag{4}$$

where for $j, k \in \{1, 2\}$, $j \neq k$, Y_i^{jk} indicates, whether individual i is at risk for a jump jk just before time t; α_0^{jk} is the baseline-hazard function, unique for all individuals $i \in \{1, ..., n\}$; $\boldsymbol{\beta}^{jk} = \left(\beta_1^{jk}, ..., \beta_p^{jk}\right)^\intercal$ is a parameter vector; $\boldsymbol{\theta}^\intercal = (\phi, (\boldsymbol{\beta}^{12})^\intercal, (\boldsymbol{\beta}^{21})^\intercal)$ is the vector collecting all unknown parameters, where ϕ is considered to be some nuisance parameter and $\mathbf{Z}_i = (\mathbf{Z}_i(t))_{t\in[0,T]} = \left(Z_i^1(t), ..., Z_i^p(t)\right)^\intercal_{t\in[0,T]}$ is the p-dimensional, predictable process of i-specific covariate processes as introduced in Section 2. Note that by (4) the densities depend only on $\boldsymbol{\theta}^{jk} := (\phi, (\boldsymbol{\beta}^{jk})^\intercal), j, k \in \{1, 2\}, j \neq k$, and consequently we can write $\Lambda_i^{jk}(t, \boldsymbol{\theta}^{jk})$ instead of $\Lambda_i^{jk}(t, \boldsymbol{\theta})$.

We now assume $\mathcal{A} = \mathcal{F}_T^Z$, where $\mathbf{Z}(t) = [Z_i^l(t)]_{i\in\{1,...,n\}, l\in\{1,...,p\}}$ is the $n \times p$-(design) matrix valued stochastic process of all i-specific covariates $\mathbf{Z}_i, i \in \{1, \ldots, n\}$.

Now we provide the connection of this counting process setting with the class of $\mathbb{F}$-doubly stochastic Markov chains. As for every individual $i \in \{1, ..., n\}$ the consecutive jumps N_i^{jk}, $j, k \in \{1, 2\}, j \neq k$, from unemployment to employment and vice versa are counted, one implicitly observes over time the associated employment-unemployment progress of the i-th person, given by the right-continuous

[8]The authors in Andersen et al. (1993) usually consider σ-algebras $\mathcal{F}_t^N$ and $\mathcal{A} = \mathcal{F}_0$, where $\mathcal{F}_0$ is supposed to contain information known at time $t = 0$. For the descriptions in this section, however, this setting needed to be extended. This further generalization has no consequence on the estimation results.

$\{1,2\}$-valued stochastic process $X_i = (X_i(t))_{t\in[0,T]}$. If we put $H_i^j(t) := \mathbb{1}_{\{X_i(t)=j\}}$, we have

$$N_i^{jk}(t) = \int_0^t H_i^j(u-)dH_i^k(u)\ , \quad j,k \in \{1,2\}, j \neq k,$$

and $\mathcal{F}_t^{X_i} = \sigma(X_i(0)) \vee \mathcal{F}_t^{\mathbf{N}_i}$ for all $t \in [0,T]$.

An immediate consequence of Theorem AppendixA.5 is then that every X_i follows an $\mathbb{F}^Z$-doubly stochastic Markov chain with 2×2-matrix valued intensity process $\mathbf{\Psi}^i$ with entries

$$\alpha_{j,k}^i(t,\boldsymbol{\theta}) = \alpha_0^{jk}(t,\phi)e^{(\boldsymbol{\beta}^{jk})^\intercal \mathbf{Z}_i(t)}\ , \tag{5}$$

if and only if the compensators' densities of the counting processes N_i^{jk} are given as in (4) with $Y_i^{jk}(t) = H_i^j(t-)$ for $j,k \in \{1,2\}, j \neq k$.

Moreover, since $\mathcal{F}_t^{X_i} = \sigma(X_i(0)) \vee \mathcal{F}_t^{N_i}$, we have that for every $t \in [0,T]$ the family $(\mathcal{F}_t^{X_i})_{i=1,\dots,n}$ is conditionally independent given $\mathcal{F}_T^Z$, if and only if the family $(\mathcal{F}_t^{N_i} \vee \mathcal{I}^i)_{i=1,\dots,n}$ is conditionally independent given $\mathcal{F}_T^Z$, where $\mathcal{I}^i = \sigma(X_i(0))$.

Hence, the estimation schemes for the individual counting processes $\mathbf{N}_i$ and their compensators $\mathbf{\Lambda}_i(\boldsymbol{\theta})$, described in Section 3.2, also provide estimators for the intensity matrix $\mathbf{\Psi}^i$ of the underlying 2-state switching processes X_i if and only if for every individual $i \in \{1,\dots,n\}$, the underlying employment-unemployment progress X_i is an $\mathbb{F}^Z$-doubly stochastic Markov chain and the family $(X_i)_{i=1,\dots,n}$ is conditionally independent given $\mathcal{F}_T^Z$.

3.2. Estimation for Cox's proportional hazards model

In the following we explain the estimation procedure for the regression parameters $\boldsymbol{\beta}^{12}$ and $\boldsymbol{\beta}^{21}$, as well as the underlying cumulative hazard functions in Cox's proportional hazard model. Following Andersen et al. (1993), we do not suppose any specific form for the underlying baseline hazard functions $\alpha_0^{jk}(t,\phi), j,k \in \{1,2\}$, $j \neq k$, appearing in (4). In particular, we assume no dependence of α_0^{jk} on a parameter ϕ and hence, we write $\alpha_0^{jk}(t)$ instead of $\alpha_0^{jk}(t,\phi)$.

According to Andersen et al. (1993), for $j \neq k \in \{1,2\}$ and for fixed $\boldsymbol{\beta}^{jk}$, maximum likelihood estimates of the integrated underlying baseline hazard functions

$$A_0^{jk}(t) = \int_0^t \alpha_0^{jk}(u)du \tag{6}$$

are obtained as the so-called *Nelson-Aalen* estimators

$$\hat{A}_0^{jk}(t,\boldsymbol{\beta}^{jk}) = \int_0^t \frac{J^{jk}(u)}{S_0^{jk}(\boldsymbol{\beta}^{jk},u)} dN_\bullet^{jk}(u),$$

where $S_0^{jk}(\boldsymbol{\beta}^{jk},t) := \sum_{i=1}^n \exp\{(\boldsymbol{\beta}^{jk})^\intercal \mathbf{Z}_i(t)\} H_i^j(t-)$ and $J^{jk}(t) = \mathbb{1}_{\{Y_\bullet^{jk}(t)>0\}}$ with $Y_\bullet^{jk}(t) = \sum_{i=1}^n Y_i^{jk}(t)$ and $N_\bullet^{jk}(t) = \sum_{i=1}^n N_i^{jk}(t)$. Note that $J^{jk}(t)$ indicates if at least one individual is in the risk of a

transition of type "jk" at time t. Otherwise, the function $S_0^{jk}(\boldsymbol{\beta}^{jk}, t)$ would be zero. Then the log *Cox partial likelihood*

$$\log L(\boldsymbol{\beta}^{jk}) = \sum_{i=1}^{n} \int_0^T (\boldsymbol{\beta}^{jk})^\intercal \mathbf{Z}_i(t) dN_i^{jk}(t) - \int_0^T \log S_0^{jk}(\boldsymbol{\beta}^{jk}, t) dN_\bullet^{jk}(t),$$

is maximized with respect to $\boldsymbol{\beta}^{jk}$, yielding $\hat{\boldsymbol{\beta}}^{jk}$. The *Breslow estimator* for (6) is finally given by $\hat{A}_0^{jk}(t, \hat{\boldsymbol{\beta}}^{jk})$.

3.3. Description of the data set

The sample of integrated labor market biographies (Stichprobe der integrierten Arbeitsmarktbiographien - SIAB) is a 2% sample of the population of the integrated employment biographies (IEB) of the IAB. We got access to the regional SIAB (Version 1975-2008), which is a file for scientific use in factual anonymous form that covers the employment and unemployment histories of 1,515,463 individuals in a total of 34,862,777 lines of data. The IEB comprises all individuals who showed one of the following statuses at least once during the observation period:

- employment subject to social security (recorded from 1975 onwards),
- marginal part-time employment (recorded from 1999 onwards),
- receipt of benefits in accordance with Social Code Book III (recorded from 1975 onwards) or Social Code Book II (recorded from 2005 onwards),
- registered with the Federal Employment Agency as a jobseeker (recorded from 2000 onwards),
- planned or actual participation in an employment or training measure (recorded from 2000 onwards).

Note that all statuses are depicted exact to the day. In general, the data have the following structure: each row represents an observation of a certain individual for a certain time interval, revealing some information about the working status of the individual together with several additional covariates concerning its educational level, sex, age, etc. An overview and a short description of all variables contained in the data is presented in Table 1, for more details see Dorner et al. (2011). The information regarding the working status allows a unique classification (in accordance with the German labor law) for each observation, whether an individual is employed or unemployed during the corresponding time interval.

[7]The variables *occup* and *eco* are not known during periods of unemployment. Hence, both variables are not considered in the intensity process for the transition from unemployment to employment.

[10]The 18 levels include the 16 German federal states themselves. Individuals with *no information* and individuals working in more than one federal state at the same time, are summarized in category *rest*.

Variable	Description
id	Identification number of individual
Tstart	Starting point of the time interval (in days)
Tstop	Endpoint of the time interval (in days)
work	Dummy variable for employment status (1=unemployed, 0=employed)
status	Dummy variable indicating whether a transition occurred or not; in the latter case, the observation is censored (1=yes, 0=no)
trans	Binary variable indicating for which transition an observation is at risk (1=employed to unemployed, 2=unemployed to employed)
sex	Dummy variable for gender (1=female, 0=male)
birth	Year of birth
natio	Nationality (1=German, 0=not German)
occup[9]	Occupational status and working hours (categorical, 8 levels)
eco[9]	Economic activity (categorical, 13 levels)
state	Federal state, where the individual is working (categorical, 18 levels[10])
educ	Educational level summarizing school education, school-leaving qualification and vocational training (categorical, 9 levels)
wage	Daily wage (in Euro; if more than one source at a time the average is used)
com	Dummy variable for commuter status (1=commuter, 0=no commuter)
msci	Monthly returns of the MSCI World Index (in %)
urate	Annual/monthly (federal state-specific) unemployment rate (in %)

Table 1: Description of the variables of the integrated German labor market data for 1975-2008.

In order to be able to apply the R-routine `coxph` on the SIAB data, several preparations are required that we describe in the following remark.

Remark 3.1.

1. *Since the data come from several records of different sources and are merged in the SIAB, the time intervals of observations belonging to one individual often overlap. Hence, in a first step we break down overlapping time intervals into disjoint intervals.*
2. *Information coming from different contemporaneous records of one individual are updated according to the most actual or crucial information. For example, if at least one of several contemporaneous records (coming from different sources) classifies an individual as a commuter, the individual is classified as a commuter for the corresponding time interval.*
3. *The metric variable* wage *is averaged over contemporaneous records; categorical variables with similar levels, coming from contemporaneous records, such as e.g.* educational level (educ), economic activity (eco) *or* occupational status and working hours (occup)*, are merged.*
4. *As some of the variables are categorical with many factor levels, we have to aggregate some of them. For example for the variables* economic activity (eco) *and* occupational status and working hours (occup) *only factor levels with at least 1,000,000 members are considered; factor levels with fewer members are summarized in the category* rest
5. *After aggregation of factor levels, still 29,778,075 lines of data, with no overlapping time intervals within individuals, remain.*
6. *In addition to individual-related factors, two macro-economic variables are considered, the* monthly returns of the MSCI World Index[11] *and the* annual/monthly unemployment rate[12]

[11]Source: webpage MSCI - A clear view of risk and return; MSCI index performance world

[12]Source: webpage Bundesagentur für Arbeit, time series for unemployment since 1950 by structural features. Up to 1991, the annual unemployment rates for the federal states of West Germany are given, from 1991 onwards the unemployment rates of all federal states of Germany are given on a monthly basis.

id	Tstart	Tstop	status	trans	sex	state	...	msci	urate
25	7763	7793	0	1	0	Sachsen	...	0.027	18.3
25	7793	7854	0	1	0	Sachsen	...	0.027	18.6
25	7854	7874	1	1	0	Sachsen	...	0.080	18.3
25	7874	7935	0	2	0	Sachsen	...	0.018	17.7
25	7967	7998	0	2	0	Sachsen	...	-0.086	17.8
25	8003	8044	1	2	0	Sachsen	...	-0.067	17.8
25	8044	8073	0	1	0	Sachsen	...	0.072	20.4
25	8073	8078	0	1	0	Sachsen	...	0.033	20.4
37	5829	5844	0	1	1	Bayern	...	0.019	5.7
37	5844	5875	0	1	1	Bayern	...	0.034	5.1
37	5875	5903	0	1	1	Bayern	...	0.091	4.7
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 2: Structure of the SIAB data, exemplarily for a fictitious extraction.

(separately for the German federal states) for the period from 1975 to 2008. Since in the fitting procedure for the estimation of (5) *all covariates have to be constant during each time interval, it has to be checked for every observation, if one of the two external variables has changed. In this case, the time interval has to be split into smaller ones where the external covariates are constant. Although both external variables only change monthly, still, this increases the data set to contain more than 200,000,000 lines. Hence, we restrained to a data sample with the maximum size that can be handled by the R-routine, ending up with 150,000 individuals.*

To illustrate the structure of the data, a short fictitious extraction is presented in Table 2.

3.4. Estimation results for the intensity processes

In the following we fit two independent Cox models for both transitions, including all covariates from Table 1. As the influence of the three metric covariates *birth*, *wage* and *urate* may be non-linear, we model their effect via cubic polynomials. For the metric regressor *msci* some special treatment becomes necessary. As the splitting procedure for *msci* creates almost exclusively time intervals, within which the variable *msci* is equal for all observations, singularities are found in the design matrix, with the consequence that the `coxph` function reports an error message. One possible solution to this problem is to include an interaction effect between the variables *state* and *msci*. Note that we do not face this problem for the variable *urate* as the time series are available separately for the German federal states. In the framework of the `coxph` package, the fit of the models can then be obtained by the following R-command, exemplarily for transition "12":

```
R> cox.obj <- coxph(Surv(Tstart, Tstop, status) ~ as.factor(sex) + birth
                 + I(birth^2) + I(birth^3) + as.factor(natio)
                 + as.factor(occup) + as.factor(eco) + as.factor(educ)
                 + wage + I(wage^2) + I(wage^3) + as.factor(com)
                 + as.factor(state) + as.factor(state):msci
                 + urate + I(urate^2) + I(urate^3),
                 data = sample, method = "breslow")
```

variable	$\hat{\beta}^{12}$
sex:woman	-0.290
natio:german	-0.345
natio:non-german	-0.483
occup:skilled worker	-0.509
occup:untrained employee	-0.547
occup:employee	-0.930
occup:rest	-1.250
occup:part-time employee (insured)	-1.167
occup:trainee	-1.807
occup:part-time employee (not insured)	-1.889
⋮	⋮

Table 3: Estimated linear effects, exemplarily for transition "12" (employment to unemployment).

Figure 1: Estimated non-linear effects of *birth, wage* and *urate* exemplarily for transition "12" (employment to unemployment).

In Table 3, we present a selection of the estimated parameters $\hat{\boldsymbol{\beta}}^{12}$, exemplarily for transition "12" (employment to unemployment). All other estimated linear effects for both transitions are presented in AppendixC. The estimated cubic effects of the metric covariates *birth, wage* and *urate*, again exemplarily for transition "12", are illustrated in Figure 1. Especially for the variable *birth*, the effect is clearly non-linear. We find that for increasing date of birth the hazard rate from (5) has a cubic form: first it slightly decreases for older individuals, then increases for younger individuals before it decreases again for very young individuals. In addition, higher wages substantially reduce the hazard rate. Furthermore, as expected, an increasing unemployment rate in the federal state, where an individual is employed, increases its instantaneous probability of a transition into unemployment.

In the two upper plots of Figure 2 we illustrate the cumulative hazards for transition "12" (left) and

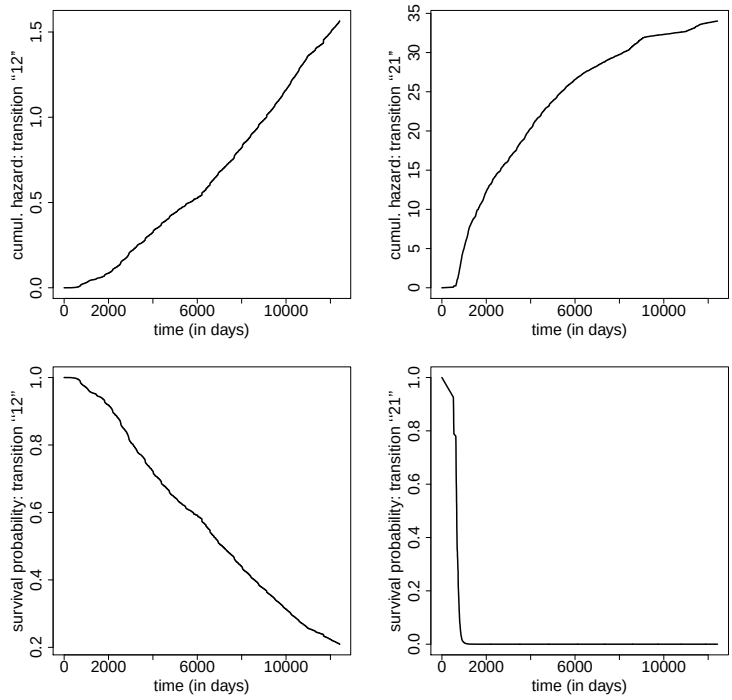

Figure 2: Cumulative hazards (top) and the survival probabilities (bottom) for transition "12" (employment to unemployment, left) and for transition "21" (unemployment to employment, right) for a "mean observation" based on the Nelson-Aalen estimates.

for transition "21" (right) for an average individual, based on the Nelson-Aalen estimates. Here, a "mean observation" represents an observation with constant covariate realizations obtained as the covariates averages over the used sample. We can see that as jumps from employment into unemployment occur much less often than vice versa and since individuals remain in employment usually much longer than in unemployment, the cumulative hazard for transition "21" (right) is much steeper than the one for transition "12" (left). The opposite effect is also observed for the survival functions, which are shown in the two lower plots of Figure 2.

3.5. Goodness-of-fit analysis

The estimation results and their graphical illustrations indicate that the model is reasonable and that the results coincide with intuitive anticipations. Nevertheless, statistically well-founded goodness-of-fit criteria are desirable. In this context, we first check the overall adequacy of the estimated model based on the Cox-Snell residuals, see Cox and Snell (1968). The Cox-Snell residual plots are shown in Figure 3 for both transitions. It is seen that the validity of the model is more

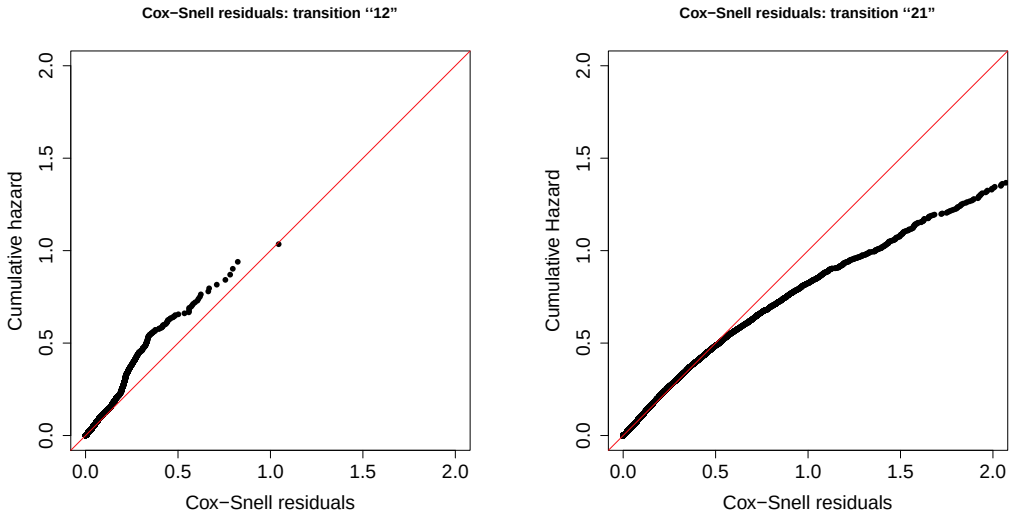

Figure 3: Cox-Snell residuals for transition "12" (left) and "21" (right).

disputable for transition "21" (unemployment to employment), since the estimated cumulative hazard of these residuals diverges from the bisecting line. One major reason might be that the two covariates *occup* and *eco* are not available for this transition and also the realizations of other variables, such as *state*, *wage* or *com*, are less reliable for individuals in the state "2" (unemployment), as they are predominantly extrapolated from the preceding employment period.

A test on the validity of the proportional hazards assumption versus the alternative of time-varying coefficients is provided by Grambsch and Therneau (1994). We show exemplarily the test results for the categorical covariate *state* and for transition "21" in Table 4. For some categories, e.g. for *Bayern* and *Berlin*, the test results indicate significant non-proportionality. This can also be graphically illustrated by plots of the scaled Schoenfeld residuals, see Schoenfeld (1982), versus a smoothed coefficient estimate, see Figure 4. Note that the left-continuous version of the Kaplan-Meier survival curve (without covariates) is used to scale the survival times. In general, we found violations of the proportional hazards assumption for several covariates and for both transitions. This indicates that several effects may vary over time and hence, models with predictors of a more complex structure could be worth of consideration. On the other hand, with regard to the large quantity of covariates it is not surprising that not all of them are consistent with the proportional hazards assumption. However, in our setting we decided to first ignore time-dependent effects, since our main objective to obtain a good prediction of the underlying jump times is achieved in a very satisfactory way in our statistical analysis, as explained below.

	rho	chisq	p
state:Baden-Wuerttemberg	-0.0047	2.4387	0.1184
state:Bayern	-0.0089	9.1932	0.0024
state:Berlin	0.0087	8.4570	0.0036
state:Brandenburg	0.0057	3.5889	0.0582
state:Bremen	0.0043	2.0088	0.1564
state:Hamburg	0.0083	7.7138	0.0055
state:Hessen	-0.0025	0.7137	0.3982
state:Mecklenburg-Vorpommern	0.0062	4.2591	0.0390
state:Niedersachsen	0.0053	3.3315	0.0680
state:Nordrhein-Westfalen	0.0048	2.8329	0.0924
state:Rheinland-Pfalz	0.0006	0.0424	0.8368
state:Saarland	0.0015	0.2595	0.6105
state:Sachsen	0.0059	3.9180	0.0478
state:Sachsen-Anhalt	0.0068	5.1429	0.0233
state:Schleswig-Holstein	0.0042	2.0377	0.1534
state:Thueringen	0.0053	3.1070	0.0780
state:sonst	0.0004	0.0134	0.9080

Table 4: Test results for testing the proportional hazards assumption

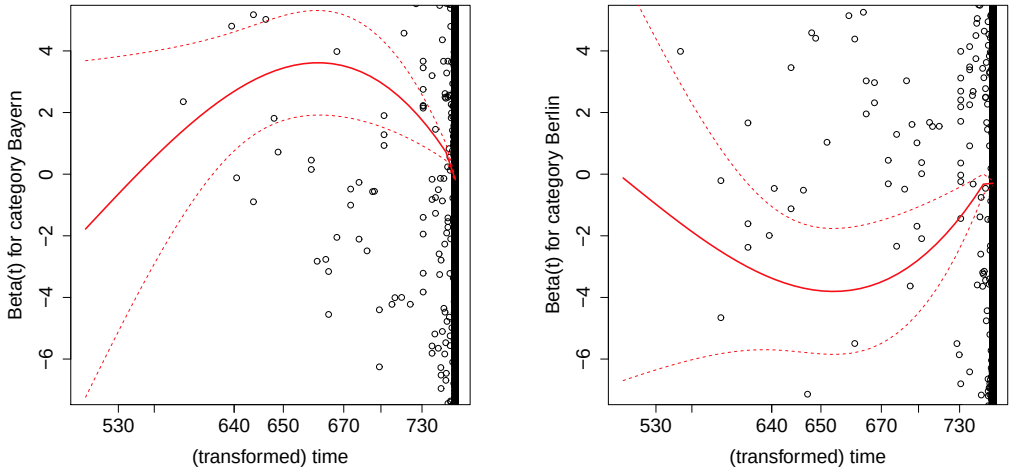

Figure 4: Smoothed time-varying coefficients for the state categories *Bayern* (left) and *Berlin* (right) together with scaled Schoenfeld residuals, exemplarily for transition "21".

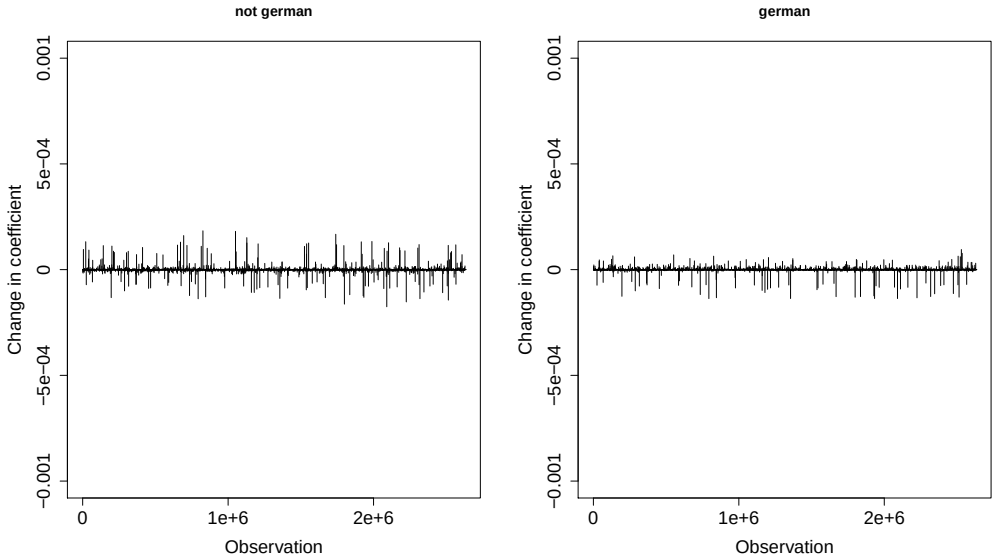

Figure 5: Approximate changes of the coefficient vector if single observations were dropped, exemplarily for the coefficients of the variable *natio* and transition "12".

Next, we investigate the robustness of the models. Due to the large size of our sample, the coefficient estimates are highly robust against slight changes in the data (critical limit is given by the range $[-0.1, 0.1]$). The approximate changes of the coefficient vector if single observations were dropped are illustrated in Figure 5, exemplarily for the coefficients of the variable *natio* and transition "12". In general, we found similar results for all other covariates and both transitions.

Finally we want to check the performance of our model in predicting transititon times by comparing true observed jump times with those estimated by the model. This is achieved by using a new goodness-of-fit method as follows.

Based on two different specifications for the estimated intensity processes, we can simulate realizations of a series of jump times for several individuals from our sample, by applying the simulation schemes for the jump times as given in (3). Compare also Bielecki and Rutkowski (2004, Section 11.3.1) or Jakubowski and Niewęgłowski (2010). The first specification uses the time-varying estimates of the baseline intensity functions as obtained from the R-routine. For individual i and $j, k \in \{1, 2\}, j \neq k$, we have

$$\hat{\alpha}^i_{j,k}(t) \;=\; \hat{\alpha}^{jk}_0(t) \exp((\hat{\boldsymbol{\beta}}^{jk})^\intercal \mathbf{Z}_i(t)). \qquad \text{(Technique (a))}$$

However, the extrapolation of the time-varying estimated baseline hazards, as it would be necessary for Technique (a), is generally a delicate task. Hence we consider a second specification based on

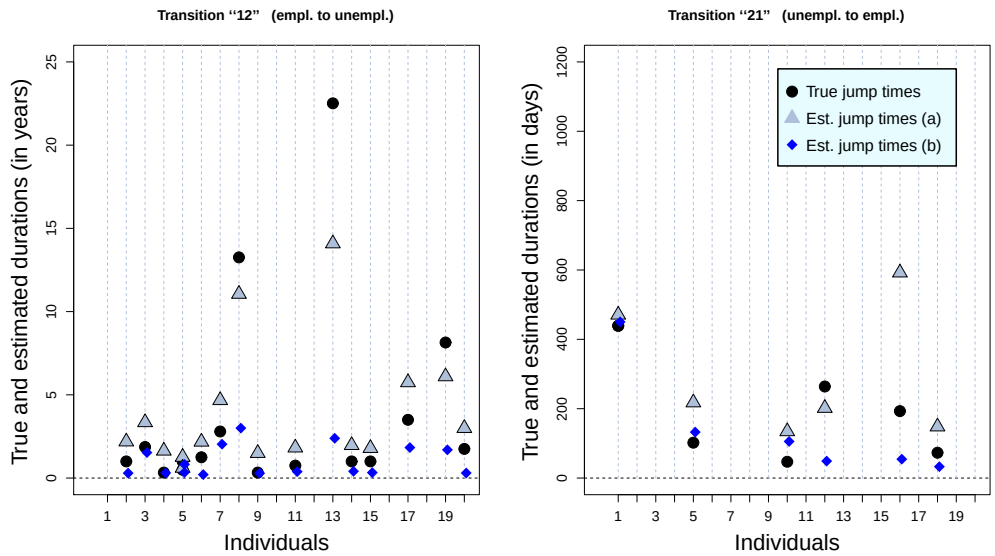

Figure 6: Goodness-of-fit results: predicted durations in the states "1" (employment, left) and "2" (unemployment, right) for 20 individuals, both for Technique (a) and (b) and each based on $1,000$ simulation runs, together with true durations.

constant baseline intensity estimates $\hat{\bar{\alpha}}_0^{jk}$, which are computed as the weighted mean of the time-varying baseline intensity estimates. For $j, k \in \{1, 2\}, j \neq k$, we obtain

$$\hat{\alpha}_{j,k}^i(t) \;=\; \hat{\bar{\alpha}}_0^{jk} \exp((\hat{\boldsymbol{\beta}}^{jk})^\intercal \mathbf{Z}_i(t)). \qquad \text{(Technique (b))}$$

Technique (b) hence provides a more natural and simple way to predict jump times. Figure 6 shows the estimated and true durations in the states "1" (employment, left) and "2" (unemployment, right) for 20 individuals, both for Technique (a) and (b). It is seen that for both techniques the predicted jump times are reasonable and quite close to the true ones. We recognize that except for extraordinary long durations in the two states, Technique (b) with $\hat{\bar{\alpha}}_0^{12} = .00076$ and $\hat{\bar{\alpha}}_0^{21} = .00417$, performs quite well and is therefore used for the simulations in Section 4.

Since the goodness-of-fit tests have shown the robustness of our analysis and the good predictive power of the model for jump times, we postpone further investigations of the proportional hazards assumption and possible extensions of the model to further works.

4. Premium determination for unemployment insurance contracts

Taking advantage of the statistical analysis of Section 3, we are now able to compute the premium for the unemployment insurance contracts as introduced in Section 2. In order to approximate the

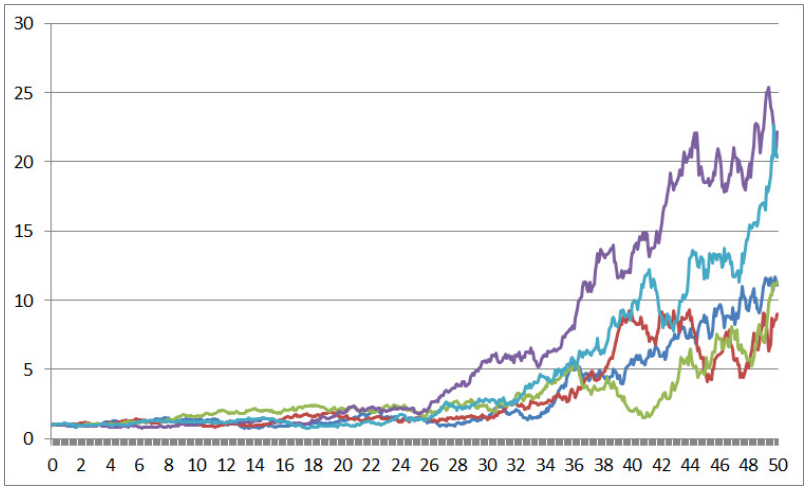

Figure 7: Simulated paths of the $\mathbb{P}$-numéraire portfolio over a time horizon of $T = 50$ years.

unemployment insurance premium (2) through Monte Carlo simulations, we need to simulate the underlying $\mathbb{F}$-doubly stochastic Markov chain with its corresponding jump times and the discounting factor, given by the $\mathbb{P}$-numéraire portfolio.

For the $\mathbb{P}$-numéraire portfolio we use a simulation scheme, described in detail in Ignatieva and Platen (2012) and Platen and Rendek (2011). The underlying assumption is that the market follows the structure of a minimal market model as introduced in Platen and Heath (2007). In this setting, the $\mathbb{P}$-numéraire portfolio, discounted with e^{rt}, is a squared Bessel process of dimension four and can be expressed as the sum of four independent time-transformed Wiener processes, see Platen and Heath (2007). Given the estimators for the model parameters in Ignatieva and Platen (2012), we simulate $N = 10000$ paths of the $\mathbb{P}$-numéraire portfolio and choose $r = 2\%$. Five exemplary paths of the $\mathbb{P}$-numéraire portfolio are shown in Figure 7.

In order to obtain realizations of the $\mathbb{F}$-doubly stochastic Markov chain X, we first simulate the intensity processes $\hat{\alpha}_{1,2}$ and $\hat{\alpha}_{2,1}$ of the Markov chain's corresponding intensity matrix by assuming the concrete multiplicative structure $\hat{\alpha}_{1,2}(t) = \hat{\bar{\alpha}}_0^{12} \exp\left((\widehat{\boldsymbol{\beta}}^{12})^\intercal \mathbf{Z}(t)\right)$ and $\hat{\alpha}_{2,1}(t) = \hat{\bar{\alpha}}_0^{21} \exp\left((\widehat{\boldsymbol{\beta}}^{21})^\intercal \mathbf{Z}(t)\right)$, introduced as Technique (b) in Section 3, with $\hat{\bar{\alpha}}_0^{12} = 0.0007611199$ and $\hat{\bar{\alpha}}_0^{21} = 0.004170514$ the mean values of the estimated baseline hazards. Here, $\widehat{\boldsymbol{\beta}}^{12}$ and $\widehat{\boldsymbol{\beta}}^{21}$ are the estimated parameter vectors and $\mathbf{Z}$ is the vector of investigated covariates of Section 3.

For the sake of simplicity, we choose most of those covariates, which may vary over time, like *wage*, *urate* or *state*, as being constant over time, although they could also be simulated according to some stochastic transition process. The only time-varying factor included in our simulations is the process of the monthly returns of our simulated paths of the $\mathbb{P}$-numéraire portfolio. In this setting

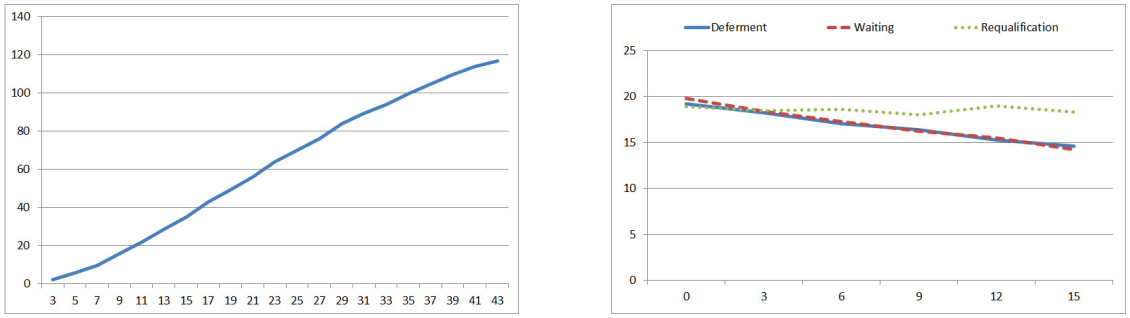

Figure 8: Simulated insurance premiums for different maturities from $T = 3$ to $T = 43$ years (left) and for different levels of deferment, waiting and requalfication periods (right), verying from 0 to 15 months.

we also assume for the simulations that the covariates $Z^1, ..., Z^p$ are independent. Note that this hypothesis is not required to perform the statistical analysis in Section 3.
Based on the resulting intensity processes we then simulate realizations of the series of jump times, underlying an $\mathbb{F}$-doubly stochastic Markov chain, as given in (3).
We then obtain realizations of an unemployment insurance contract's claim payments by testing if the insured person fulfills the criteria of receiving a claim payment at the respective payment dates. For the Monte Carlo approximations we perform this scheme of simulating the claim payments $N = 10,000$ times.
In the following we present several plots where we tested the simulated insurance premium against variation of its defining factors. To be more precise, the insurance premium (2) depends among others on

(i) contractual specifications, i.e. the maturity T, the deferment period D, the waiting period W and the requalification period R,

(ii) on the categorial characteristics of the insured person, i.e. the covariate categories *sex*, *natio*, *occup*, *eco*, *state*, *educ* and *commuter*,

(iii) on the characteristics of the metric covariates *birth*, *wage*, *urate* and *msci*, i.e. the monthly returns of the simulated $\mathbb{P}$-numéraire portfolio, for which the MSCI can be taken as a proxy.

In order to test the sensitivity of the premium against the variation of one of these factors, we fix all other factors at the mean levels of the sample, and let only the factor of interest vary.
Figure 8 shows the simulated insurance premiums for different maturities from $T = 3$ to $T = 43$ years and for different levels of the deferment, waiting and requalification periods from 0 to 15 months, respectively. Both plots are coherent with the intuitive anticipations, for example that the insurance premium has to decline if the claim-excluding time periods increase. Note, however, that

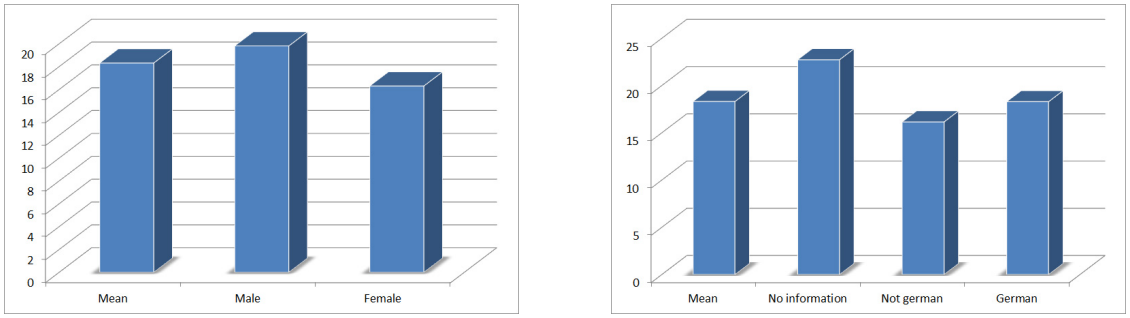

Figure 9: Simulated insurance premium for the different categories of *sex* (left) and of the nationality *natio* (right).

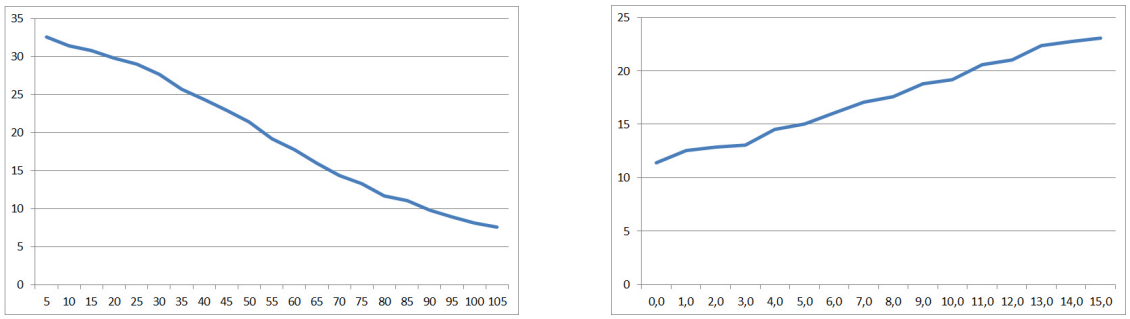

Figure 10: Simulated insurance premiums for different levels of *wage* (left) of a (time-constant) *urate* (right).

the effects of the deferment and waiting period are stronger than the effects of the requalification period.

Figure 9 shows the insurance premiums for the different levels of the covariates *sex* and *natio*. Based on our estimation results, the insurance premium for women and foreign employees are lower than for men or German employees.

Figure 10 shows the simulated insurance premiums for different levels of the covariates *wage* and *urate*.

All other simulation results can be found in AppendixC.

5. Conclusion

Here we establish an innovative and flexible premium determination framework for unemployment insurance products which incorporates both static and stochastic covariate processes in an effective way. This allows us to adjust the insurance premiums according to individual characteristics of the insured person as well as to micro- and macro-economic risk factors.

Within the framework of Cox's proportional hazards model the intensities are estimated on a data set provided by the IAB. We show that the class of $\mathbb{F}$-doubly stochastic Markov chains provides appropriate multi-state switching processes, underlying the counting process framework of the estimation procedure. Several goodness-of-fit methods indicate the adequateness of the model assumptions and the estimated results. In particular, with a new goodness-of-fit method we show the good predictive power of the results for the jump times of the underlying multi-state switching process. Further improvements of the model could be to consider other (time-dependent) covariates and a direct parametrization of the likelihood function. We then model dependencies directly in the structure of the stochastic intensity processes and calculate the fair insurance premiums with Monte Carlo arguments. This evaluation method, based on the real-world pricing formula and on stochastic intensities can be easily adapted to other financial and insurance products traded on the market. Furthermore, our approach for estimating stochastic intensities and the dependence structure of jump times can be used for a wide spectrum of applications. Therefore, we are confident to have provided a flexible and widely employable framework which can be applied beyond the scopes of this paper.

AppendixA. $\mathbb{F}$-doubly stochastic Markov chains

We introduce here the concepts and basic properties of $\mathbb{F}$-doubly stochastic Markov chains, as given in Jakubowski and Niewęgłowski (2010). As seen in the paper, these kind of stochastic processes are very adequate for modeling time-dependent random covariate effects since they provide matrix-valued stochastic intensity processes.
Let $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ be a complete, filtered probability space, where the filtration $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]}$ is assumed to fulfill the usual conditions of completeness and right-continuity, see Protter (2003), with some arbitrary time horizon $T > 0$. Moreover, let $X = (X(t))_{t\in[0,T]}$ be a right-continuous stochastic process with state space $\{1, 2\}$. We denote by $\mathbb{F}^X$ the natural filtration generated by X, i.e. $\mathcal{F}_t^X = \sigma(X(u) : u \leq t)$ for all $t \in [0, T]$, and consider the filtration $\mathbb{G}$ to be the enlargement of $\mathbb{F}^X$ through some reference filtration $\mathbb{F}$, i.e. we assume $\mathcal{G}_t = \mathcal{F}_t^X \vee \mathcal{F}_t$ for all $t \in [0, T]$. Furthermore, we set $\widetilde{\mathcal{G}}_t = \mathcal{F}_t^X \vee \mathcal{F}_T$, $t \in [0, T]$.

Definition AppendixA.1. *A process X is called an $\mathbb{F}$-doubly stochastic Markov chain with state space $\{1, 2\}$, if there exists a family of stochastic 2×2-matrices $\boldsymbol{P}(s, t) = [p_{j,k}(s, t)]_{j,k\in\{1,2\}}$ for $0 \leq s \leq t \leq T$ such that*

(1) the matrix $\boldsymbol{P}(s, t)$ is $\mathcal{F}_t$-measurable, and $\boldsymbol{P}(s, .)$ is progressively measurable for any given $s \in [0, t]$,

(2) for every $j, k \in \{1, 2\}$ we have

$$\mathbb{1}_{\{X(s)=j\}}\mathbb{P}(X(t) = k \mid \widetilde{\mathcal{G}}_s) = \mathbb{1}_{\{X(s)=j\}}p_{j,k}(s, t) \ .$$

The process $\boldsymbol{P}$ is called the conditional transition probability process *of X.*

Note that Definition AppendixA.1 generalizes the notion of a classical continuous time Markov chain which itself is an $\mathbb{F}$-doubly stochastic Markov chain with $\mathcal{F}_t = \{\emptyset, \Omega\}$ for all $t \in [0, T]$. Other examples for $\mathbb{F}$-doubly stochastic Markov chains are compound Poisson processes or Cox processes, see Jakubowski and Niewęgłowski (2010). Since the reference filtration $\mathbb{F}$ is not specified, an $\mathbb{F}$-doubly stochastic Markov chain is not necessarily a Markov process according to the usual definition. Another property, which makes the class of $\mathbb{F}$-doubly stochastic Markov chains interesting for applications, is that they have matrix-valued stochastic intensity processes in the following sense.

Definition AppendixA.2. *An $\mathbb{F}$-doubly stochastic Markov chain X with state space $\{1, 2\}$ is said to have an intensity, if there exists an $\mathbb{F}$-adapted 2×2-matrix-valued stochastic process $\boldsymbol{\Psi} = (\boldsymbol{\Psi}(t))_{t\in[0,T]} = \big([\alpha_{j,k}(t)]_{j,k\in\{1,2\}}\big)_{t\in[0,T]}$ such that*

1) $\boldsymbol{\Psi}$ is integrable, i.e.

$$\int_0^T \sum_{j\in\{1,2\}} |\alpha_{j,j}(s)|ds < \infty \ . \tag{A.1}$$

2) $\boldsymbol{\Psi}$ satisfies the following conditions:

$$\alpha_{j,k}(t) = -\alpha_{j,j}(t) \geq 0 \quad \forall j,k \in \{1,2\}, j \neq k,\ t \in [0,T] \tag{A.2}$$

$$\boldsymbol{P}(v,t) - \mathbb{I} = \int_v^t \boldsymbol{\Psi}(u)\boldsymbol{P}(u,t)du \quad \forall 0 \leq v \leq t \leq T \quad \textit{(Kolmogorov backward equation)}$$

$$\boldsymbol{P}(v,t) - \mathbb{I} = \int_v^t \boldsymbol{P}(v,u)\boldsymbol{\Psi}(u)du \quad \forall 0 \leq v \leq t \leq T \quad \textit{(Kolmogorov forward equation)}$$

A process $\boldsymbol{\Psi}$, satisfying the above conditions, is called an intensity of the $\mathbb{F}$-doubly stochastic Markov chain X.

Not every $\mathbb{F}$-doubly stochastic Markov chain has an intensity, but there are sufficient conditions for the existence.

Theorem AppendixA.3. *Let X be an $\mathbb{F}$-doubly stochastic Markov chain with conditional transition probability process $\boldsymbol{P}$. Assume that*

1) $\boldsymbol{P}$ as a matrix-valued mapping

$$\boldsymbol{P} : ([0,T]^2 \times \Omega, \mathcal{B}([0,T]^2) \otimes \mathcal{G}) \to ([0,1]^{2\times 2}, \mathcal{B}([0,1]^{2\times 2})$$

is measurable.

2) there exists a version of $\boldsymbol{P}$, which is continuous in s and in t.

3) for every $t \in [0,T]$, the following limit exists almost surely

$$\boldsymbol{\Psi}(t) := \lim_{h \searrow 0} \frac{\boldsymbol{P}(t,t+h) - \mathbb{I}}{h} ,$$

and is integrable.

Then $\boldsymbol{\Psi}$ is the intensity of X.

Proof. See Jakubowski and Niewęgłowski (2010, Theorem 3.12). □

The following theorem gives a useful result for applications, namely that for every given 2×2-matrix-valued stochastic process $\widetilde{\boldsymbol{\Psi}}$ with a particular structure, there exists an $\mathbb{F}$-doubly stochastic Markov chain X, having intensity $\widetilde{\boldsymbol{\Psi}}$.

Theorem AppendixA.4. *Let $(\widetilde{\boldsymbol{\Psi}}(t))_{t\in[0,T]}$ be an $\mathbb{F}$-adapted 2×2 matrix-valued stochastic process, satisfying conditions* (A.1) *and* (A.2)*. Then there exists an $\mathbb{F}$-doubly stochastic Markov chain X with intensity $(\widetilde{\boldsymbol{\Psi}}(t))_{t\in[0,T]}$.*

Proof. See Jakubowski and Niewęgłowski (2010, Theorem 4.8). □

For $j \in \{1,2\}$ let

$$H^j(t) := \mathbb{1}_{\{X(t)=j\}}$$

be the indicator function for X, being in state j at time t, and denote by $\mathbf{H}(t) = (H^1(t), H^2(t))^\intercal$ the corresponding 2-variate vector. Moreover, for $j, k \in \{1, 2\}, j \neq k$, let $N^{jk} = (N^{jk}(t))_{t\in[0,T]}$ with

$$N^{jk}(t) := \int_0^t H(u-)^j dH^k(u) \ ,$$

define the counting processes of the jumps of X from state j to k, denoted by the superscript "jk" up to time t.
The following theorem provides a martingale characterization of $\mathbb{F}$-doubly stochastic Markov chains that is the core connection between $\mathbb{F}$-doubly stochastic Markov chains and counting processes, underlying the estimation procedure, given in Section 3.

Theorem AppendixA.5. *Let $X = (X(t))_{t\in[0,T]}$ be a stochastic process with state space $\{1, 2\}$ and $\mathbf{\Psi} = (\mathbf{\Psi}(t))_{t\in[0,T]}$ be a 2×2-matrix-valued process, satisfying* (A.1) *and* (A.2)*. The following conditions are equivalent:*

i) X is an $\mathbb{F}$-doubly stochastic Markov chain.

ii) For $j \in \{1, 2\}$, the processes $M^j = (M^j(t))_{t\in[0,T]}$ with

$$M^j(t) := H^j(t) - \int_{]0,t]} \alpha_{X(u),j}(u)du$$

are $\widetilde{\mathbb{G}}$-local martingales.

iii) For $j, k \in \{1, 2\}, j \neq k$, the processes $M^{jk} = (M^{jk}(t))_{t\in[0,T]}$ with

$$M^{jk}(t) := N^{jk}(t) - \int_{]0,t]} H^j(u)\alpha_{j,k}(u)du$$

are $\widetilde{\mathbb{G}}$-local martingales.

iv) The vector-valued process $\boldsymbol{L} = (\boldsymbol{L}(t))_{t\in[0,T]}$, defined by

$$\boldsymbol{L}(t) := \boldsymbol{Q}(0, t)^\intercal \boldsymbol{H}(t) \ ,$$

where $\boldsymbol{Q}(0, t)$ is the unique solution to the random integral equation

$$d\boldsymbol{Q}(0, t) = -\mathbf{\Psi}(t)\boldsymbol{Q}(0, t)dt, \ \ \boldsymbol{Q}(0, 0) = \mathbb{I},$$

is a $\widetilde{\mathbb{G}}$-local martingale.

Proof. See Jakubowski and Niewęgłowski (2010, Theorem 4.1). □

AppendixB. Real-world pricing

Here we summarize the main concepts of the theory of real-world pricing. All fundamental results concerning the benchmark approach can be found in Platen and Heath (2007) for jump diffusion and Itô process driven markets and in Platen (2004) for a general semimartingale market.

We consider a frictionless market model in continuous time, which is set up on the complete, filtered probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$, with arbitrary time horizon $T > 0$ and $d+1$ non-negative adapted tradable (primary) security account processes, denoted by $S^j = (S^j(t))_{t\in[0,T]}$, $j \in \{0, 1, ..., d\}$, $d \geq 1$. The security account process S^0 is often taken as a riskless bank account in the domestic currency. We write $\mathbf{S} = (\mathbf{S}(t))_{t\in[0,T]} = (S^0(t), S^1(t), ..., S^d(t))^\intercal_{t\in[0,T]}$ for the $(d+1)$-dimensional random vector process, consisting of the $d+1$ assets, and assume that $\mathbf{S}$ is a càdlàg semimartingale. Let $L(\mathbf{S})$ denote the space of $\mathbb{R}^{d+1}$-valued predictable strategies $\boldsymbol{\delta} = (\boldsymbol{\delta}(t))_{t\in[0,T]}$ for which the corresponding gains and losses, respectively, from trading in the assets, i.e. $\int_0^t \boldsymbol{\delta}^\intercal(u) \cdot d\mathbf{S}(u)$, exist for all $t \in [0, T]$. The portfolio value $S^\delta(t)$ at time $t \in [0, T]$ is then given by

$$S^\delta(t) = \boldsymbol{\delta}^\intercal(t)\mathbf{S}(t) = \sum_{j=0}^{d} \delta^j(t)S^j(t)\,.$$

A strategy $\boldsymbol{\delta} \in L(\mathbf{S})$ is called *self-financing*, if changes in the portfolio value are only due to changes in the assets and not due to in- or outflow of money, i.e. if

$$dS^\delta(t) = \boldsymbol{\delta}^\intercal(t)d\mathbf{S}(t)\ .$$

We write $\mathcal{V}_x^+$ ($\mathcal{V}_x$) for the set of all strictly positive (non-negative) and self-financing portfolios S^δ with initial capital $S^\delta(0) = x$. We now introduce the notion of the $\mathbb{P}$-numéraire portfolio.

Definition AppendixB.1. *A portfolio $S^{\delta^*} \in \mathcal{V}_1^+$ is called* $\mathbb{P}$-numéraire portfolio *if every non-negative portfolio $S^\delta \in \mathcal{V}_x$, discounted (or benchmarked) with S^{δ^*}, forms a $(\mathbb{G}, \mathbb{P})$-supermartingale for every $x \geq 0$. In particular, we have*

$$\mathbb{E}\left[\frac{S^\delta(\sigma)}{S^{\delta^*}(\sigma)}\,\middle|\,\mathcal{G}_\tau\right] \leq \frac{S^\delta(\tau)}{S^{\delta^*}(\tau)} \qquad a.s. \tag{B.1}$$

for all stopping times $0 \leq \tau \leq \sigma \leq T$.

If a $\mathbb{P}$-numéraire portfolio exists, it is unique, as can be easily seen with the help of the supermartingale property and Jensen's inequality, see Becherer (2001).

We now assume the existence of the $\mathbb{P}$-numéraire portfolio $S^{\delta^*} \in \mathcal{V}_1^+$ in our market model. This obviously depends on the model specifications of the market. However, it is a rather weak hypothesis, since existence has been proven for most settings of nowadays practical interest, see Becherer (2001), Karatzas and Kardaras (2007) or Platen and Heath (2007). Regarding estimation and calibration of the $\mathbb{P}$-numéraire portfolio, it is shown in Platen and Heath (2007) that every sufficiently diversified market portfolio yields a good proxy for it, which is why we choose time series of the MSCI in our estimation procedure in Section 3.

With the existence of the $\mathbb{P}$-numéraire portfolio and the corresponding supermartingale property

(B.1), (strong) arbitrage opportunities[13] are excluded, see Platen (2004).
In this setting derivative pricing can be performed under the probability $\mathbb{P}$ by using the fundamental property (B.1) of the $\mathbb{P}$-numéraire portfolio as follows.

Definition AppendixB.2. *A portfolio process $S^\delta = (S^\delta(t))_{t\in[0,T]}$ is called* fair*, if its benchmarked value process, i.e. its portfolio, discounted with the $\mathbb{P}$-numéraire portfolio, $\widehat{S}^\delta(t) := \frac{S^\delta(t)}{S^{\delta^*}(t)}, t \in [0,T]$ forms a $(\mathbb{G},\mathbb{P})$-martingale.*

A T-contingent claim C is given as a $\mathcal{G}_T$-measurable random variable with $\mathbb{E}\left[\frac{|C|}{S^{\delta^*}(T)}\right] < \infty$. According to Definition AppendixB.2, it is natural to define the so called *real-world pricing formula* for a T-contingent claim C as follows:

Definition AppendixB.3. *For a T-contingent claim C the* fair price *$P_t(C)$ of C at time $t \in [0,T]$ is given by*

$$P_t(C) := S^{\delta^*}(t)\mathbb{E}\left[\frac{C}{S^{\delta^*}(T)}\Big|\mathcal{G}_t\right] . \tag{B.2}$$

Due to Definition AppendixB.1, (B.2) represents the minimal price for C.

Remark AppendixB.4. *Since price determination occurs under the real-world measure and the $\mathbb{P}$-numéraire portfolio represents an intrinsic benchmark indicator of financial and economic conditions, the real-world pricing method appears to be most natural to be applied to the evaluation of hybrid financial insurance products. More reasons to apply real-world pricing are the following:*

1. *there is a natural connection with classical premium determination via* (B.2)*;*
2. *there is no need of introducing a risk neutral measure that may be not defined for hybrid markets;*
3. *real-world pricing is consistent with (asymptotic) utility indifference pricing in a very general setting;*
4. *the real-world pricing formula* (B.2) *for $t = 0$ also represents the minimal price for initiating several optimal hedging schemes like (local) risk-minimizing strategies in incomplete markets.*

For further details, we refer to the discussions in Biagini (2011) and Biagini and Widenmann (2012).

[13] A non-negative, self-financing portfolio $S^\delta \in \mathcal{V}_x$ permits *(strong) arbitrage*, if

$$\mathbb{P}(S^\delta(\tau) = 0) = 1 \qquad \text{and} \qquad \mathbb{P}(S^\delta(\sigma) > 0 \,|\, \mathcal{G}_\tau) > 0$$

for some stopping times $0 \le \tau \le \sigma \le T$ and some initial capital x. A given market model is said to be *arbitrage-free*, if there exists no non-negative portfolio $S^\delta \in \mathcal{V}_x$ of the above kind.

AppendixC. Estimation and simulation results

variable	$\hat{\beta}^{12}$	$\hat{\beta}^{21}$
sex:woman	-0.290	0.083
natio:german	-0.345	-0.208
natio:non-german	-0.483	-0.291
occup:skilled worker	-0.509	-
occup:untrained employee	-0.547	-
occup:employee	-0.930	-
occup:part-time employee (insured)	-1.167	-
occup:trainee	-1.807	-
occup:part-time employee (not insured)	-1.889	-
occup:rest	-1.250	-
birth	0.073	0.564
I(birth^2)	0.000	-0.003
I(birth^3)	-0.014	0.047
eco:retail	-0.169	-
eco:transport and communications	-0.370	-
eco:business oriented services	-0.171	-
eco:household oriented services	-0.021	-
eco:educational, social and health facility	-0.138	-
eco:public administration, social insurance	-0.182	-
eco:basic material and good production	-0.146	-
eco:structural steel, light-metal and machine engineering	-0.504	-
eco:steel forming, vehicle- and equipment engineering	-0.396	-
eco:consumer goods industry	-0.172	-
eco:main construction trades	0.597	-
eco:rest	-0.011	-
educ:abitur	0.878	2.460
educ:abitur and job training	0.679	0.828
educ:job training	1.365	1.860
educ:senior technical college (Fachhochschule)	0.869	1.051
educ:college/university	0.617	0.827
educ:elementary or secondary modern school, secondary school leaving certificate or similar	0.778	2.679
educ:elementary or secondary modern school, secondary school leyving certificate or similar with job training	0.582	2.787
educ:no education	2.853	2.495
wage	-0.854	1.380
I(wage^2)	-0.064	-0.148
I(wage^3)	0.052	-0.083
com:commuter	0.333	0.144
com:no commuter	0.090	0.147
state:Baden-Wuerttemberg	-1.362	-0.168
state:Bayern	-0.994	0.023
state:Berlin	-1.451	-0.077
state:Brandenburg	-1.533	0.002
state:Bremen	-1.354	-0.074
state:Hamburg	-1.410	-0.153
state:Hessen	-1.259	-0.157
state:Mecklenburg-Vorpommern	-1.454	0.009
state:Niedersachsen	-1.108	-0.018
state:Nordrhein-Westfalen	-1.382	-0.207
state:Rheinland-Pfalz	-1.201	-0.064
state:Saarland	-1.459	-0.236
state:Sachsen	-1.541	0.066
state:Sachsen-Anhalt	-1.582	0.017
state:Schleswig-Holstein	-1.096	0.034
state:Thueringen	-1.536	0.132
state:rest	-1.531	-0.407
urate	0.057	-0.022
I(urate^2)	0.002	-0.001
I(urate^3)	-0.000	0.000

Note that for the federal state, we have to adjust not only the covariate *state* but also *urate* and *msci*. For the unemployment rates, we take the contemporary valid unemployment rate and for *msci* the monthly returns of the simulated paths of the $\mathbb{P}$-numéraire portfolio.

variable	$\hat{\beta}^{12}$	$\hat{\beta}^{21}$
state:no info:msci	0.346	-1.317
state:Baden-Wuerttemberg:msci	-1.342	-0.489
state:Bayern:msci	0.214	0.416
state:Berlin:msci	1.102	-0.532
state:Brandenburg:msci	-0.090	-0.113
state:Bremen:msci	-1.860	-0.785
state:Hamburg:msci	-0.650	-0.865
state:Hessen:msci	-0.699	-0.960
state:Mecklenburg-Vorpommern:msci	-0.829	0.113
state:Niedersachsen:msci	0.046	-0.365
state:Nordrhein-Westfalen:msci	-0.957	-1.019
state:Rheinland-Pfalz:msci	-0.235	-1.277
state:Saarland:msci	-1.138	-1.370
state:Sachsen:msci	-0.285	-0.535
state:Sachsen-Anhalt:msci	-0.041	-0.294
state:Schleswig-Holstein:msci	-0.682	-0.077
state:Thueringen:msci	NA	NA
state:rest	NA	NA

Table C.5: Estimated linear effects.

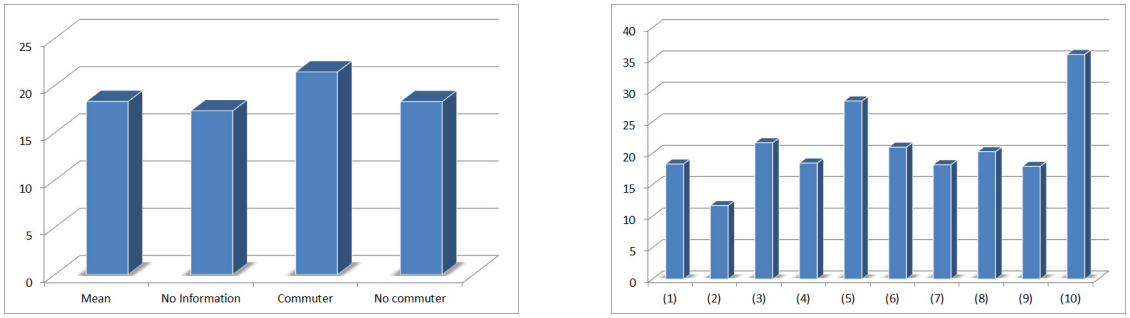

Figure C.11: Simulated insurance premiums for the different levels of *commuter* (left) and of *educ* (right), with (1) Sample mean, (2) No information, (3) Abitur, (4) Abitur and job training, (5) Job training, (6) Senior technical college (Fachhochschule), (7) College/University, (8) Elementary or secondary modern school, secondary school leaving certificate or similar, (9) Elementary or secondary modern school, secondary school leyving certificate or similar with job training, (10) No Education.

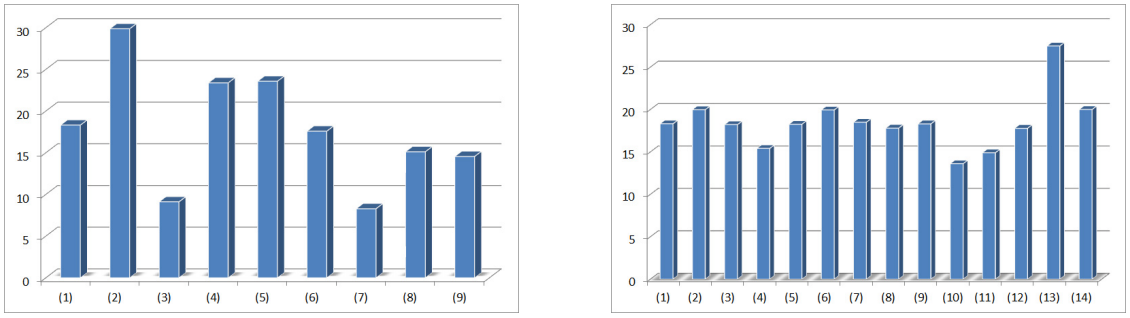

Figure C.12: Simulated insurance premiums for the different levels of *occup* (left) with (1) Sample mean, (2) No information, (3) In education, (4) Un-/semi-skilled worker, (5) Skilled worker, (6) Employee, (7) Part-time employee without unemployment insurance, (8) Part-time employee with unemployment insurance, (9) Others, and of *eco* (right) with (1) Sample mean, (2) No information, (3) Retail, (4) Transport and communications, (5) Business oriented services, (6) Household oriented services, (7) Educational, social and health facility, (8) Public administration, social insurance, (9) Basic material and good production, (10) Structural steel, light-metal and machine engineering, (11) Steel forming, vehicle- and equipment engineering, (12) Consumer goods industry, (13) Main construction trades, (14) Others.

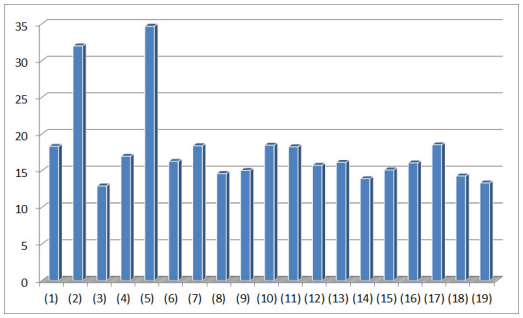

Figure C.13: Simulated insurance premiums for the different levels of *state* with (1) Sample mean, (2) No information, (3) Baden-Württemberg, (4) Bayern, (5) Berlin, (6) Brandenburg, (7) Bremen, (8) Hamburg, (9) Hessen, (10) Mecklenburg-Vorpommern (11) Niedersachsen, (12) Nordrhein-Westfalen, (13) Rheinland-Pfalz, (14) Saarland, (15) Sachsen, (16) Sachsen-Anhalt, (17) Schleswig-Holstein, (18) Thüringen, (19) Others.

O. O. Aalen, J. Fosen, H. Weedon-Fekjær, Ø. Borgan, and E. Husebye. Dynamic analysis of multivariate failure time data. *Biometrics*, 60:764–773, 2004.

P. K. Andersen, Ø. Borgan, R. D. Gill, and N. Keiding. *Statistical Models Based on Counting Processes*. Springer Series in Statistics, Springer, New York, 2nd edition, 1993.

A. Bacinello, E. Biffis, and P. Millossovich. Pricing life insurance contracts with early exercise features. *Journal of Computational and Applied Mathematics*, 233:27–35, 2009.

D. Becherer. The numeraire portfolio for unbounded semimartingales. *Finance and Stochastics*, 5: 327–341, 2001.

F. Biagini. Evaluating hybrid products: the interplay between financial and insurance markets. *Proceedings of Ascona 2011: Seventh Seminar on Stochastic Analysis, Random Fields and Applications*, 2011. In press.

F. Biagini and I. Schreiber. Risk-minimization for mortality derivates. *SIAM Journal on Financial Mathematics*, 2012. In press.

F. Biagini and J. Widenmann. Pricing of unemployment insurance products with doubly-stochastic Markov chains. *International Journal of Theoretical & Applied Finance*, 15(4):1250025–1 – 1250025–32, 2012.

F. Biagini and J. Widenmann. Risk minimization for insurance products via $\mathbb{F}$-doubly stochastic Markov chains. Work in progress, Ludwig-Maximilians-University of Munich, 2013.

T.R. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging*. Springer, 2004.

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

D. R. Cox. Partial likelihood. *Biometrics*, 62:269–275, 1975.

D. R. Cox and E. J. Snell. A general definition of residuals (with discussion). *Journal of the Royal Statistical Society series B*, 30:248–275, 1968.

L. C. de Wreede, M. Fiocco, and H. Putter. The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. *Computer Methods and Programs in Biomedicine*, 99:261–274, 2010.

M. Dorner, M. König, and S. Seth. FDZ-Datenreport: Sample of integrated labour market biographies; regional file 1975-2008 (siab-r 7508). Technical report, Research Data Centre fo the German Federal Employment Agency (BA) and the Institut for Employment Research (IAB), 2011.

P. Grambsch and T. Therneau. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81:515–526, 1994.

K. Ignatieva and E. Platen. Estimating the diffusion coefficient function for a diversified world stock index. *Computational Statistics and Data Analysis*, 56:1333–1349, 2012.

C. H. Jackson. Multi-state models for panel data: The msm package for r. *Journal of Statistical Software*, 38(8):1–29, 2011.

J. Jakubowski and M. Niewęgłowski. A class of $\mathbb{F}$-doubly stochastic Markov chains. *Electronic journal of probability*, 15:1743–1771, 2010.

I. Karatzas and C. Kardaras. The numéraire portfolio in semimartingale financial models. *Finance and Stochastics*, 11(4):447–493, 2007.

T. Kneib and A. Hennerfeind. Bayesian semiparametric multi-state models. *Statistical Modelling*, 8:169–198, 2008.

T. Møller. Risk-minimizing hedging strategies for unit-linked life insurance contracts. *ASTIN Bulletin*, 28:17–47, 1998.

E. Platen. A benchmark framework for risk management. In *Stochastic Processes and Applications to Mathematical Finance, Proceedings of the Ritsumeikan Intern. Symposium,*: 305–335, 2004.

E. Platen and D. Heath. *A Benchmark Approach to Quantitative Finance.* Springer, 2007.

E. Platen and R. Rendek. Simulation of diversified portfolios in a continuous financial market. QFRC Working Paper, University of Sydney, 2011.

P. Protter. *Stochastic integration and differential equations*. Springer, 2003.

D. Schoenfeld. Partial residuals for the proportional hazards regression model. *Biometrika*, 69: 239–241, 1982.

N. Simon, J. Friedman, T. Hastie, and R. Tibshirani. Regularization paths for Cox's proportional hazards model via coordinate descent. *Journal of Statistical Software*, 39(5):1–13, 2011.

T. Therneau. *A Package for Survival Analysis in S*, 2012. R package version 2.36-14.

T. Therneau and P. Grambsch. *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, 2000.