

Lapse tables for lapse risk management in insurance: a competing risk approach

Xavier Milhaud*, Christophe Dutang†

May 31, 2017

Abstract

This paper deals with the crucial problem of modelling policyholders' behaviours in life insurance. We focus here on the surrender behaviours and model the contract lifetime through the use of survival regression models. Standard models fail at giving acceptable forecasts for the timing of surrenders because of too much heterogeneity, whereas the competing risk framework provides interesting insights and more accurate predictions. Numerical results follow from using F&G model (Fine & Gray (1999)) on an insurance portfolio embedding Whole Life contracts: through backtests, this framework reveals to be quite efficient and recovers the empirical lapse rate trajectory by aggregating individual predicted lifetimes. These results could be particularly useful to design future insurance product. Moreover, this setting allows to calibrate experimental lapse tables, simplifying the lapse risk management for operational teams.

Keywords: life insurance, lifetime, surrender, lapse, competing risks, cumulative incidence functions.

1 Introduction

Lapses correspond to “the expiration of all rights and obligations under an insurance contract if the policyholder fails to comply with certain obligations required to uphold those”¹. In terms of financial consequences, lapse risk is one of the biggest risks to consider for life insurers. Lapses strongly affect insurers' Asset and Liabilities Management (ALM) since they trigger unexpected cash flows, and modify the insurers' commitments through changes in contractual guarantees.

In this paper, we aim to provide an accurate prediction for the timing of policyholder's lapse. More precisely, we focus on surrenders, where surrenders are part of the underwriting risk module in the Solvency II directive and are defined¹ as “the (premature) termination of an insurance contract by the policyholder”. In this case, the insurer has to pay the policyholder or its beneficiary the *surrender value* (or *cash value*) which is contractually agreed. Lapses thus include surrenders, but also other causes of termination: death, default on premium payments, maturity, or else. That being

*Université de Lyon, Université Lyon 1, ISFA, Laboratoire SAF, 50 Avenue Tony Garnier, F-69007 Lyon.

†Laboratoire Manceau de Mathématiques, Université du Maine, Avenue Olivier Messaien, F-72000 Le Mans.

¹See http://ec.europa.eu/finance/insurance/docs/solvency/impactassess/annex-c08d_en.pdf

said, it seems natural to consider a *competing risk approach* since these causes of lapse are mutually exclusive (Laurent et al. (2016)). For sure, someone who dies cannot surrender her contract, or a default on premium payments has nothing to do with contractual maturity. Surprisingly, there is no such approach in the literature concerning the lapse modelling within a regression framework, though it is a standard tool to model prepayments in the banking sector. Many empirical studies show that the surrender rate drives the lapse rate trajectory in life insurance companies: indeed this trajectory basically results from policyholders' decisions, but do not significantly depend on other involuntary events (e.g. death) that remain quite stable in proportion over time. That is the reason why this paper focuses on surrender behaviours.

In practice, it is crucial for insurers to model not only the surrender decision but also its timing, because initial expenses to issue the contract as well as associated management costs cannot be recovered in case of early surrenders. This means that the insurer must pay a particular attention to the product design to avoid this, but more on that in Section 2 through our real-life analysis. Based on this idea, we adopt a survival analysis framework combined to a competing risk approach. To the best of our knowledge, lapse risk has never been modeled this way: the novelty of our work lies in the consideration of the *subdistribution approach* to model the *contract survival probability* related to surrenders. There exist many other statistical approaches available in the literature to model surrenders: generalized linear models (GLM) were often used in the past (Cox & Lin (2006), Milhaud (2013)), as well as financial mathematics techniques (to price the surrender option, see Bacinello (2005), Bacinello et al. (2008)) or temporal series and cointegration (Kuo et al. (2003)). However, they all suffer from “unacceptable” drawbacks: a selection bias is introduced when using GLM on multiple periods, the assumption of agent's rationality is very strong in the second case, and the heterogeneity of policyholders' behaviours cannot be captured within the last approach. To assess the financial consequences of lapses on the insurer balance sheet, please refer to the interesting papers by Buchardt (2014) and Buchardt et al. (2015) (competing risks appear therein, but there is no focus on individual risk factors that impact the surrender rate).

In this competing risk framework, the *surrender* is our cause of interest. In this perspective, we develop a survival regression model which allows to predict the *individual contract lifetime before surrender* given a set of risk factors affecting the propensity to terminate the contract. Most of time, practitioners suggest to distinguish structural from temporary surrenders: the firsts relate to surrenders due to personal projects (e.g. purchase a car, buy a new house, which is not independent from the policyholder's age for instance), whereas others are triggered by adverse scenarios concerning the conditions surrounding the insurance contract (e.g. bad macroeconomic conditions, firm reputation). In this view, the model has to integrate various effects, from idiosyncratic risk factors (Outreville (1990)) to external information (e.g. a financial market index linked to the insurance contract profitability, see Kim (2005) or Russell et al. (2013)). Roughly speaking, the model should be flexible enough to take into account the heterogeneity of policyholders' behaviours facing different situations. In this context, we show hereafter that the F&G model (Fine & Gray (1999)) provides satisfying results and enables to get accurate predictions of contract lifetime depending on policyholder's characteristics, contract features and financial environment.

The paper is organized as follows: Section 2 presents the portfolio and the insurance product under consideration. Some general descriptive statistics are provided. In Section 3, we explain the differences between the cause-specific and the subdistribution approaches in a competing risk framework: despite being the most famous one, the former approach tends to be unadapted to our context. To illustrate this, we give a practical example which shows that the estimation of the surrender intensity can be strongly biased because of the misestimation of intensities related to other (rare) events. Then we

build an *optimal model* in Section 4, which is validated in Section 5 through classical test statistics applied to the learning data. Furthermore, we check for its prediction power by two methodologies: i) we compare predicted and observed surrender rates (computed from the aggregation of individual lifetimes), and ii) we compare predicted and observed survival probabilities by risk profile. Finally and in the same spirit as CSO tables for mortality in US, we propose to build multidimensional experimental surrender tables in Section 6.

2 Context and portfolio under study

In this section, we first describe the general context corresponding to our study and give some particular features of the product under consideration. Our simulated dataset is largely inspired by a real-life portfolio provided by a private insurer operating in US.

Life insurance in US can roughly be divided into two families: either Term Life or Whole Life contracts. Term Life contracts concerns temporary insurance contracts which guarantees a capital against the death of the insured (e.g. a mortgage life insurance) and does not accumulate a cash value. On the contrary, Whole Life insurance deals with permanent insurance contracts which accumulates a cash value and terminates at the death of the policyholder. Our data deals with Whole Life policies, where the policyholder may voluntary cancel her contract (possibly with penalties in case of early withdrawals).

Typically, Whole Life insurance comes along with a minimum death benefit (*nominal value*), a maturity benefit on the cash value (the minimum between the nominal and the economic value of the unit-linked account), and sometimes with *riders* which guarantee extra cash amount in case of predetermined events (e.g. accidental death). The underwriting of such policies assumes the payment of a contractual premium. That premium can be paid periodically (e.g. every month), or as a lump sum at the beginning. This particular feature may have an impact on the customer behaviour: either attachment to the contract because of the regular contribution, or tiredness. The premium depends on at least four features: the nominal value, the policyholder’s age and gender, and the tobacco consumption.

Expenses and commissions are added up to the pure premium. The corresponding rates depend mainly on the distribution channel: brokers and tied-agents are not similarly compensated. In Table 1, we present the commission and the expense rates generally used for Whole Life contracts sold by tied-agents. Notably, the commission rate is very high in the first year of the contract and falls to zero after the tenth year. Therefore, the tied-agent may have an incentive to force customers to move to a new contract with higher commissions for him. It is also important to mention that the expense rates vary depending on the premium frequency. The general rule is: the more frequent the premium is paid, the higher the expenses.

Type (nominal > 10000\$)	1 st year	2-10 years	11+ years
Commission	50%	4%	0%
Management expense	0%	2%	3%
Total	50%	6%	3%

Table 1: Commissions and expense as percentage of the annual premium.

As already mentioned, the policyholder may choose to surrender her life insurance policy in order to immediately retrieve the cash value. However, the policyholder faces potential penalties. For example in France, there are 35% of tax on capital gains of the contract if the withdrawal happens before the fourth year, 15% between the fifth and the seventh year, and only 7.5% above an increasing threshold afterwards. In US (our setting), surrender charges also apply if a withdrawal occurs during the surrender period (from 10 to 15 years). Here, the surrender charge is 100% on the cash value in the first three years and then linearly decreases as time elapses. This feature is a clear incentive to keep the contracts in force at least three years, and then as long as possible.

Each year and after deduction of expenses and commissions, the premium remainder is devoted to a saving account. This accumulated amount is called the *total cash value*. It is the sum of three components: the guaranteed cash value, the dividends, and the terminal dividends minus potential penalties. The value of dividends is not guaranteed and depends on the financial performance of the insurance company. Whatever the value, the dividend option gives extra money to the policyholder punctually. Non-terminal dividends may be used to pay cash to the policyholder, to accumulate on interest on the dividend value or to buy paid-up options (i.e. new riders: increasing the guaranteed value, decreasing the premium, increasing the death benefit). These possibilities are incentives to increase the loyalty of customers. Unfortunately, in our dataset, non terminal dividends are unknown. Terminal dividend may be known at underwriting, see Table 2. Since the terminal dividends are zero before the twelfth year, the customer has another incentive to keep the policy in force for a long time. These features of Whole Life insurance are common to both American and European markets, and can be seen as a *profit benefit option*.

When there is a premium payment default, multiple termination options are possible. Either the contract may be transformed to a temporary contract paid-up extended Term insurance (so that the cash value is used to buy a Term Life insurance policy), or the contract may be extended to a reduced-paid up insurance so that the nominal value gets lower.

Regulation and taxation also impacts the surrender value and by consequence the customer choice. In US, the TAMRA law was initially designed to prevent policyholders from using life insurance contracts to gain huge tax relief. In particular, it sets a maximum amount of premiums paid during the first seven years. Otherwise, the life insurance policy is considered as a modified endowment contract. In addition, any change in the contractual nominal value or the inclusion of an extra rider renews this seven-year period.

Let us now say a few words about the individual data, and give some descriptive statistics. The portfolio contains detailed information on the 29,317 Whole Life policies, all sold from the tied-agent channel between January 1995 and December 2008. Table 3 provides usual statistical indicators on the main variables: most of them are categorical, such as the gender or the payment frequency. Only two information (also called further covariates or risk factors) are continuous: the annual premium, and the last observed Dow Jones (DJ) quarterly variation. **Both have been standardized** for future statistical modelling.

Underwriting age	Policy age (years)									
	[0, 11]	12	13	14	15	16	17	18	19	≥ 20
[0, 39]	-	-	-	1%	1.5%	2%	2.5%	3%	3.5%	4%
[40, 44]	-	-	1%	1.5%	2%	2.5%	3%	3.5%	4%	4.5%
≥ 45	-	1%	1.5%	2%	2.5%	3%	3.5%	4%	4.5%	5%

Table 2: Dividend rates at the termination date.

Variable	Statistics	Comments
Issue Date	between 01/01/1995 and end of 2008	for policies not terminated in Dec. 2008, we have no information: fixed right censored
Time duration (in quarters)	min: 0.01; max: 62.09; mean: 30.26; std: 18.78	unknown if policy not terminated
Gender	male: 50.05%; female: 49.95%	no missing value
Payment frequency	infra annual: 61.37%; annual: 23.44%; other (supra annual): 15.19%	Infra-annual: monthly, quarterly, semi-annual.
Risk state	non smoker: 63.01%; smoker: 36.99%	no missing value
Underwriting age	young: 47.46% middle: 34.04% old: 18.50%	between 0 and 34 years old between 35 and 54 years old between 55 and 84 years old
Living place	east coast: 20.62%; west coast: 4.60%; other: 74.78%	no missing value
Annual premium	min: -1.07; median: -0.30; max: 12.13 mean: \$560.88 ; std: \$526.5870	this variable has been standardized. (in original dollar scale)
DowJones Index variation	min: -4.53; median: -0.38; max: 2.43 mean: 0.001781 ; std: 0.049413	this variable has been standardized. (in original scale)
Accidental death rider	yes: 16.42%; no: 83.58%	
Termination cause	0: 49.06% 1: 38.22% 2: 12.72%	in force surrender cancellation (other causes: death, term,...)
Death indicator	0: 95.62% 1: 4.38%	alive dead

Table 3: List of main variables in our database.

To visualize the data, we use traditional boxplots where the width of the boxplot is proportional to the exposure of each variable category in the whole dataset. In Figure 1, we represent the boxplot of the duration with respect to the termination cause (i.e. no termination, surrender, cancellation). We logically observe that the duration of the policy is much longer for in-force policies than those that have been terminated. In average the contract lifetime is close to 14 quarters for surrendered contracts, whereas half of the contracts experience at least 45 quarters of duration. These statements are in line with the aforementioned product features.

Similar boxplots can be obtained from other categorical covariates, but are not illustrated here since the differences in terms of duration for each category are not as convincing as for the termination cause. Instead, we present in Table 4 the results of the Kruskal-Wallis test (χ^2 statistic) for independence between the lifetime and the categorical variable under study. All variables except the living place of the policyholder are statistically not independent to the surrender decision. It thus makes sense to analyze them in the coming modeling section, yet the living place should not appear to be relevant.

	Surrender	Termination	Acc Rider	Gender	Under. Age	Living place	Risk state	Prem. Freq.
p-value (%)	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	0.0126	0.0001	38.32	0.0157	$< 10^{-4}$
KW statistic	7626.30	11211.89	27.62	14.70	27.59	1.92	14.29	111.05
deg. of freedom	1	2	1	1	2	2	1	2

Table 4: Results of the Kruskal-Wallis rank sum test.

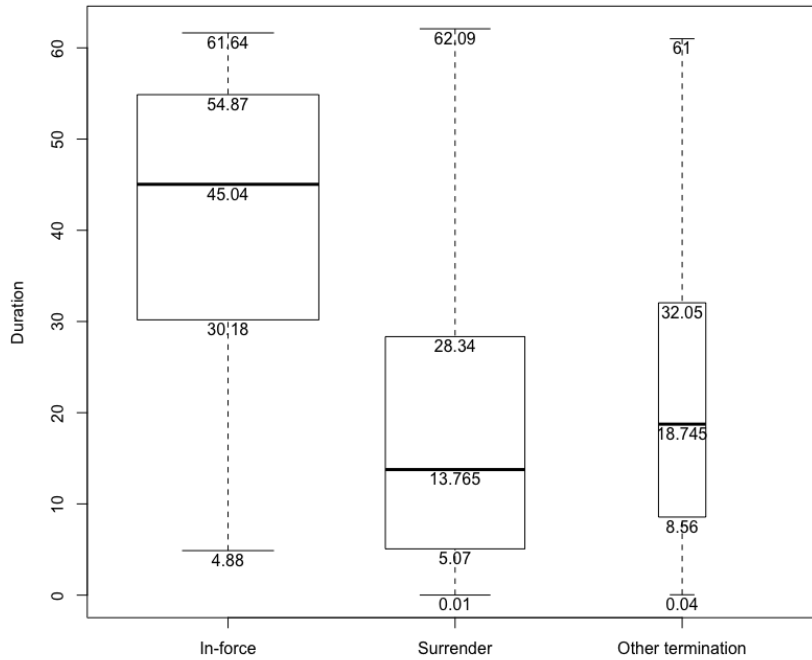


Figure 1: Boxplots of duration (in quarters) depending on the termination cause.

In Figure 2, we plot the quarterly lapse and surrender rates all along the observation period (from 1995 to 2008). Notice that the two trajectories are similar and have a strong correlation, which shows that other events than surrenders (e.g. deaths) have a stationary impact on the lapse rate. The barplots in the background represent the number of in-force policies: the biggest exposure in the portfolio happens in 2006.

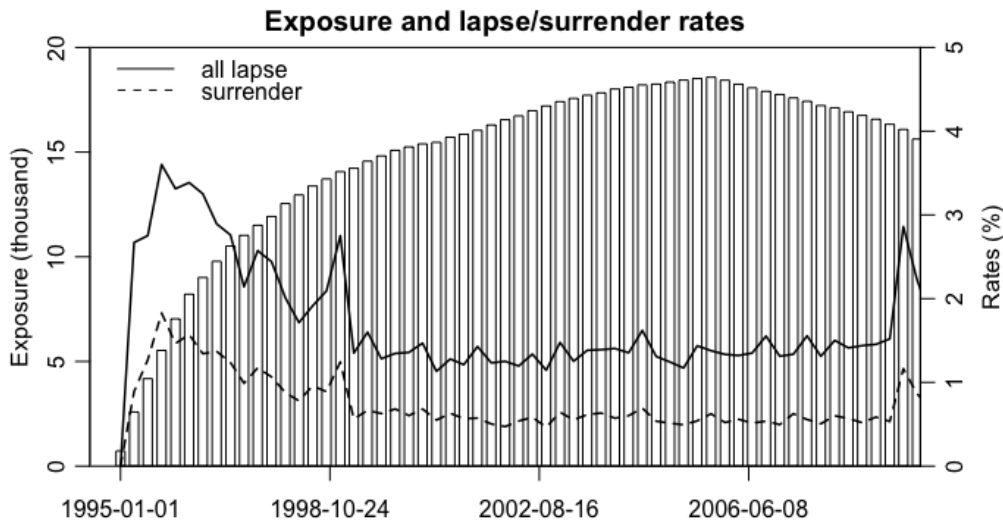


Figure 2: Exposure and exit rates over time, on the **whole sample**. The exposure (left y-axis) refers to the number of in-force policies for each quarter, and the right y-axis represents quarterly rates.

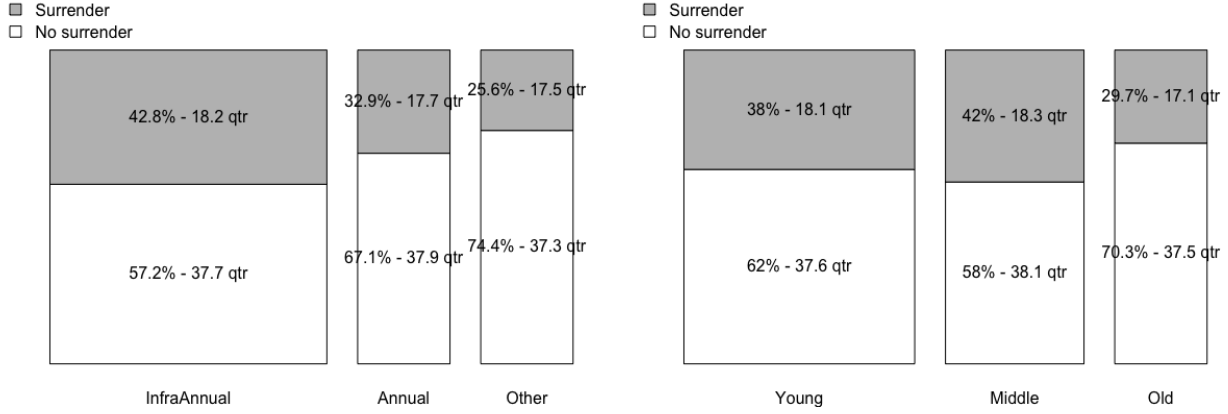


Figure 3: Surrender rates depending on premium frequency and policyholder's age.

In Figure 3, we analyze the effect of two risk factors on the surrender decision: the payment frequency and the underwriting age. We have chosen these two variables because they have the highest statistic value in Table 4. For each category, we also report the surrender rates (in percentage) and the mean of the contract lifetime (in quarters). Graphically, the effect is clear: both the payment frequency and the age have a significant impact. For instance, people that pay premiums more than once a year are more likely to surrender, and middle-aged policyholders exhibits the highest surrender rates. This last statement is not surprising since people from 35 to 54 years old usually have more frequent liquidity needs for personal projects. However, despite that they tend to surrender more than the others, they also have higher mean contract lifetime.

3 Survival analysis: cause-specific or subdistribution approach?

Recall that our goal is to predict individual contract lifetimes. In order to achieve this, there exist two standard approaches in survival analysis: the cause-specific one, and the subdistribution one. The goal of this section is twofold: to present the probabilistic details of both models, and to make a choice between both, based on the specificity of our data.

Let us denote by T the random variable of the contract lifetime. In a survival analysis, the distribution of T is generally specified by its hazard function, or its survival function. Indeed, the survival function defined as $S_T(t) = P(T > t)$ characterizes the distribution, as well as the hazard rate defined as $\lambda_T(t) = -S'_T(t)/S_T(t)$. Typically, for an exponential distribution of rate λ , $\lambda_T(t) = \lambda$; and for a Weibull distribution of scale λ and shape k , $\lambda_T(t) = (k/\lambda) (t/\lambda)^{k-1}$. The survival function can be derived from the hazard function by $S_T(t) = \exp(-\int_0^t \lambda_T(s) ds)$ since $\lambda_T(t) = -d \ln S_T(t)/dt$.

In practice lifetimes are may be censored, that is we observe the random variable $Y = \min(T, C)$ where T is the variable of interest and C is the censoring variable. C may be random or deterministic depending on the type of studies: hereafter, C is considered to be a random variable. Assuming that C is independent from T , it is easy to show that $\lambda_T(t) = \lambda_Y(t) + \lambda_C(t)$. So a straightforward conclusion is that random censoring without the information whether T is censored or not leads to false estimation of $\lambda_T(t)$ when estimating $\lambda_Y(t)$. However, when we know if T is censored or not, the

couple $(T, \mathbb{1}_{T \leq C})$ has the targeted hazard rate

$$\lambda_T(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt \mid T \geq t)}{dt}.$$

Hence, there is no particular bias when dealing with censoring as long as we know which data are censored.

Let us now consider the counting process framework. A counting process $(N_t)_t$ is a càdlàg stochastic process adapted to a filtration with $N_0 = 0$, $N_t < +\infty$ a.s. and jumps of size 1. The well-known Doob-Meyer decomposition, see Fleming & Harrington (2013), states that $N_t = \Lambda_t + M_t$ where $(\Lambda_t)_t$ is called the compensator, a non decreasing predictable process and $(M_t)_t$ is a local martingale w.r.t. the filtration such that $E(N_t) = \Lambda_t$. In the absolute continuous case, the compensator has the following form $\Lambda_t = \int_0^t \lambda(s) ds$ where $\lambda(t)$ is a predictable process known as the *intensity*. The very special case of a constant intensity $\lambda(t) = \lambda$ leads to the Poisson process with exponential rate jumps.

A process of particular interest linked to the couple $(T, \mathbb{1}_{T \leq C})$ is $N_t = \mathbb{1}_{T \leq t, T \leq C}$, that jumps if the time of interest T is below t and non censored. The at-risk process defined as $R_t = (\mathbb{1}_{T \geq t})_t$ indicates if neither the event nor the censoring occurs before t . (N_t, R_t) is the stochastic process counterpart of static random variables $(T, \mathbb{1}_{T \leq C})$. Studying n i.i.d. replicates $(T_i, \mathbb{1}_{T_i \leq C_i})_{i=1, \dots, n}$ leads to the empirical version of the counting processes by summing over all individuals:

$$N_t = \sum_{i=1}^n \mathbb{1}_{T_i \leq t, T_i \leq C_i}, \quad R_t = \sum_{i=1}^n \mathbb{1}_{\min(T_i, C_i) \geq t}.$$

One can show that the compensator of N_t is $\Lambda_t = \int_0^t R_s \lambda_T(s) ds$. The well-known nonparametric Nelson-Aalen estimator of Λ_t is defined by $\hat{\Lambda}_t = \int_0^t (1/R_s) \mathbb{1}_{R_s > 0} dN_s$. By the central limit theorem, the asymptotic distribution of $\hat{\Lambda}_t$ can be obtained.

In a multi-cause framework, see e.g. Martinussen & Scheike (2006), we suppose that at the failure time we know the cause of the failure. That is we define J_T the type of failure among $\{1, \dots, J\}$. The process $(J_t)_t$ starting with $J_0 = 0$ is a continuous-time random process that jumps at T into a state in $\{1, \dots, J\}$. From a multi-state model point-of-view, the multi-cause model is a special case where 0 is the initial state and $\{1, \dots, J\}$ are absorbing states, see Figure 4.

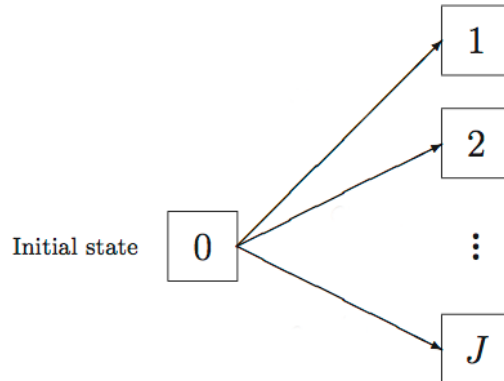


Figure 4: Competing risk model and multistate representation, for $j = 1, 2, \dots, J$.

Let us define *cause-specific* hazard rates for $j \in \{1, \dots, J\}$ as

$$\lambda_{T,j}(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt, J_T = j \mid T \geq t)}{dt}. \quad (1)$$

By the formula of total probability, we can retrieve the overall hazard rate by summation of Equation (1): $\lambda_{T,1}(t) + \dots + \lambda_{T,J}(t) = \lambda_T(t)$, and recover the overall survival distribution of T by

$$P(T > t) = 1 - F_T(t) = S_T(t) = \exp\left(-\int_0^t (\lambda_{T,1}(s) + \dots + \lambda_{T,J}(s)) ds\right).$$

In practice we are interested not in S_T but in the following probability $F_{T,j}(t) = P(T \leq t, J_T = j)$, the so-called *cumulative incidence function* (CIF) for the cause j . Due to the event $J_T = j$, this is not a proper cumulative distribution function since $F_{T,j}(t) \rightarrow P(J_T = j)$ as $t \rightarrow +\infty$.

For a continuous distribution of T , we characterize it as the integral of a (improper) density $F_{T,j}(t) = \int_0^t f_{T,j}(s) ds$. By conditioning on $T \geq t$, the improper density is obtained as the following limit:

$$f_{T,j}(s) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt, J_T = j)}{dt} = \lambda_{T,j}(t) S_T(t).$$

In other words, $f_{T,j}(t)$ is the product of the cause-specific hazard rate and the probability to survive up to time t . Therefore, the CIF is

$$F_{T,j}(t) = \int_0^t \lambda_{T,j}(s) \exp\left(-\int_0^s \lambda_T(u) du\right) ds. \quad (2)$$

Hence, in this framework, the CIF of cause j depends on **all other causes** via the global survival function, which makes the interpretation of the effects of covariates quite tricky since some effects come from the overall hazard rate $\lambda_T(t)$. The CIF has the good property to be interpretable and summable $P(T \leq t) = F_{T,1}(s) + \dots + F_{T,J}(s)$, unlike to the function $1 - \exp\left(-\int_0^t \lambda_{T,j}(u) du\right)$.

This approach, called the *cause-specific approach*, thus requires to estimate the hazard rates (1) of all causes so as to estimate the CIF (2) of cause j .

The concurrent methodology to estimate the CIF of a single cause is possible by considering a new competing risk process. Let us assume that cause 1 is our cause of interest. We define τ as

$$\tau = T \times \mathbb{1}_{J_T=1} + \infty \times \mathbb{1}_{J_T \neq 1}.$$

The distribution of τ is the same as T for $J_T = 1$, $P(\tau \leq t) = F_{T,1}(t)$ and a mass point at infinity $1 - F_{T,1}(\infty)$, probability to observe other causes ($J_T \neq 1$) or not to observe any failure. The hazard rate of τ can be written as

$$\lambda_\tau(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt, J_T = 1 \mid \{T \geq t\} \cup \{T \leq t, J_T \neq 1\})}{dt}. \quad (3)$$

Hence the CIF for cause 1 is computed as

$$F_{T,1}(t) = 1 - \exp\left(-\int_0^t \lambda_\tau(s) ds\right). \quad (4)$$

Therefore the estimation of the CIF (4) does not depend on the estimation of other causes' hazard rates. This second approach is called the *subdistribution approach*, and often leads to different effects of covariates on the cause-specific hazard function on one side and on the corresponding CIF on the other side (Gray (1988), Pepe (1991)). Without censoring, the process $N_\tau(t) = \mathbb{1}_{\tau \leq t} = \mathbb{1}_{T \leq t, J_T=1}$ has the compensator

$$\Lambda_\tau(t) = \int_0^t \lambda_\tau(s) (\mathbb{1}_{T \geq s} + \mathbb{1}_{T \leq s, J_T \neq 1}) ds.$$

Taking into account censoring, the nonparametric Nelson-Aalen estimator of Λ_τ can be adapted to the competing risk framework by updating the at-risk process R_t and the counting process N_t accordingly. That is, the numerator is the count of non-censored times $T_i \leq C_i$ of interest $J_{T_i} = 1$ while the denominator is the count of non-exited individuals and exited individuals of other causes ($J_{T_i} \neq 1$):

$$\widehat{\Lambda}_\tau(t) = \mathbb{1}_{\max_i T_i > t} \sum_{i=1}^n \frac{\mathbb{1}_{T_i \leq t \leq C_i, J_{T_i}=1}}{R_{T_i}}, \quad R_t = \sum_{i=1}^n (1 - \mathbb{1}_{T_i \leq t \leq C_i, J_{T_i}=1}),$$

with $Y_i = \min(T_i, C_i)$.

To illustrate the differences between these two approaches, we consider a randomly selected subset of our database, and extract the three main causes of termination: surrender, death, and other causes grouped in a single category (i.e. $J_T \in \{1, 2, 3\}$). Our cause of interest is the surrender cause, denoted by $J_T = 1$. Other causes are tagged as competing risks. On this subset, we fit a simple regression model via the cause-specific approach by considering one single explanatory variable: the risk state. In addition, we fit the same regression model (with the risk state as covariate, related to the smoker status) via the subdistribution approach. By construction, the cause-specific approach will provide an estimation of three CIFs $P(T \leq t, J_T = 1), P(T \leq t, J_T = 2), P(T \leq t, J_T = 3)$ unlike the subdistribution approach which only provides $P(T \leq t, J_T = 1)$.

Estimations are plotted in Figure 5. The solid line corresponds to the CIF of non-smoker policyholders, while the dashed line represents the CIF of smokers. Colored curves stand for the regression model, whereas the black curves correspond to the non-parametric estimation. On the middle and the right-hand graphs, we observe that the CIF estimated by the cause-specific approach are particularly badly estimated: there is a significant difference with the non-parametric estimation. On the left-hand graph we observe that the estimated CIF are graphically close, irrespectively of the approach. In order to differentiate these approaches numerically, we compute error statistics² in Table 5. In overall, the CIF for non-smoker are better estimated than the CIF for smoker (errors three times lower). On the two competing risks (causes 2 and 3), the errors are particularly high: from 8 to 90 times higher than the error for cause 1. Regarding the cause of interest, the estimation via the subdistribution approach is slightly better than with the cause-specific approach. This is probably due to uncertain estimations of hazard rates for other causes used when computing the CIF. The difference would probably be much higher if there were very few events for one of the different causes of lapse: compensation would be experienced, which should lead to poor estimation of individual hazard rates. In conclusion to these results, we opt for the subdistribution approach in the coming modelling section.

Risk state	Sum of absolute relative error		Sum of squared relative error	
	non-smoker	smoker	non-smoker	smoker
Cause-spec. $P(T \leq t, J_T = 1)$	11.53	45.66	0.69	76.08
Cause-spec $P(T \leq t, J_T = 2)$	83.40	184.73	44.59	156.91
Cause-spec $P(T \leq t, J_T = 3)$	37.68	135.98	11.46	448.04
Subdist. $P(T \leq t, J_T = 1)$	9.71	39.80	0.62	68.86

Table 5: Goodness of fit statistics: summary of errors related to the CIF estimation.

²See Appendix A for a definition.

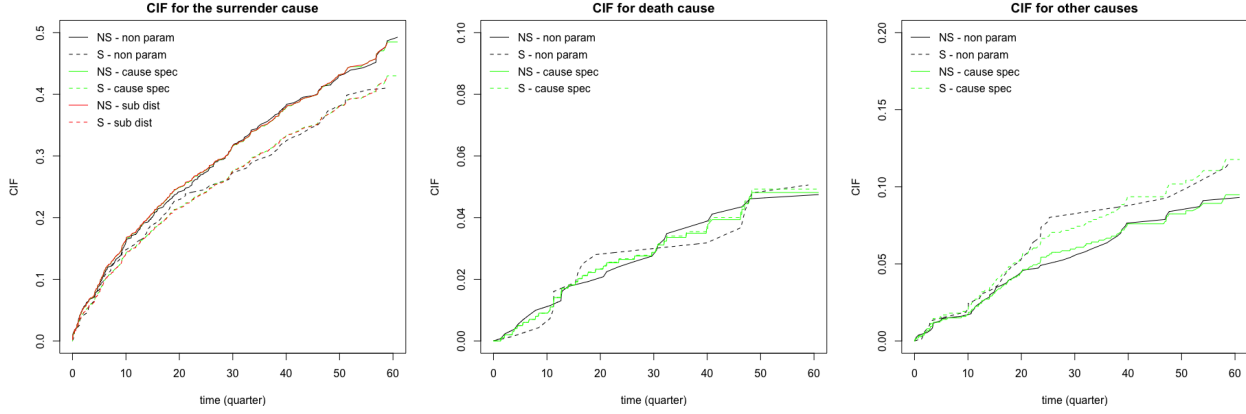


Figure 5: Estimation of the cumulative incidence functions (NS: non-smoker, S: smoker).

4 Model selection

To build the model, the dataset is splitted into two subsamples. Two third of the data, the learning sample, are used to fit the competing risk model. The last third, representing the test sample, will enable us to assess the performance and the robustness of the selected model in Section 5.

First, let us introduce formally the F&G proportional hazards model, see Fine & Gray (1999). As previously explained, the principle of the subdistribution approach lies in considering a new and artificial *at-risk population* (this idea already existed in the context of cure models). Integrating covariates to take into account the individual characteristics of the population (through the vector \mathbf{X}), the subdistribution hazard of this model for surrenders (cause 1 from (4)) is defined by

$$\begin{aligned} \lambda_{\tau,1}(t; \mathbf{X}) &:= \lim_{dt \rightarrow 0} \frac{1}{dt} P(t \leq T < t + dt \cap J_T = 1 \mid \{T \geq t\} \cup \{T \leq t \cap J_T \neq 1\}, \mathbf{X}) \\ &= \lambda_{10}(t) \exp(\mathbf{X}(t)' \beta), \end{aligned} \quad (5)$$

where $\lambda_{10}(t)$ is a completely unspecified nonnegative function (risk of the reference profile, more details further), $\mathbf{X}(t)' = (X_1(t), \dots, X_k(t))$ is the vector of the k (time-varying or not) covariates, and $\beta' = (\beta_1, \dots, \beta_k)$ is the vector of regression coefficients to be estimated³.

This specification looks like the well known Cox model. However, recall that the *at-risk population* is different to take into account the competing risk framework, which has an impact on the estimation of $\lambda_{10}(t)$ and β . The expression of the hazard $\lambda_{\tau,1}(t; \mathbf{X})$ yields to the following formula for the CIF of surrenders:

$$F_{T,1}(t; \mathbf{X}) = P(T \leq t, J_T = 1 \mid \mathbf{X}) = 1 - \exp\left(-\int_0^t \lambda_{10}(s) \exp(\mathbf{X}^T(s)\beta) ds\right). \quad (6)$$

In order to select an optimal model and make some comparisons at the end, we start by estimating a fully nonparametric model (whose regression coefficients are all time-varying). Then, step by step, we introduce parametric terms through the consideration of constant effect on the response for some covariates. The selection of introduced parametric terms is made based on statistical tests that measure the significance of the effects and their type (time-varying versus constant coefficients β). Typically, the supremum test and the Kolmogorov-Smirnov test are used. In the fully nonparametric model,

³, denotes the transpose.

all the regression coefficients are time-varying. On the opposite, the case of only constant regression coefficients is the semiparametric model by Fine & Gray (1999), see Equation (5). This procedure allows us to check that considering only constant effects is not too restrictive from a modelling viewpoint. All the estimations are performed in R (R Core Team (2017)) thanks to the packages `timereg` (Martinussen & Scheike (2006), Scheike & Zhang (2011)) and `cmprsk` (Gray (2014)).

To start with, all available covariates of the database were initially inputted in the modelling. That is to say *accidental death rider*, *gender*, *premium frequency*, *risk state*, *underwriting age*, *living place*, *annual premium*, and *Dow Jones index* (more precisely the relative variation of this index in the last observed quarter). The results show that one category of the living place, people living on the *West Coast*, has a non significant effect on the lifetime before surrender (see Table 12 in Appendix B). We thus aggregate this category with another one (*East Coast*), and estimate once again the model. Considering *living place* as a risk factor is still not relevant (see Table 13, still in Appendix B). The estimation of the model without this variable leads to satisfying results in terms of global significance (Table 14), despite concerns about some effects that should not be considered as time-varying. Indeed, it is clear from Table 15 in Appendix B that some covariates seem to have a constant effect (constant associated regression coefficients) on the lifetime before surrender.

Looking at the outputs of previous statistical tests, we now estimate an “intermediate” model in which some of the covariates have a constant effect on the response. These covariates are the *underwriting age* and the *gender*. After having fitted this new model, results show that other covariates should also be considered as having constant effects: this is the case of the category *annual* of the *premium frequency*, as well as the category *smoker* for the *risk state*. Finally, this process ends up with a model having constant effects (*gender*, *risk state*, category *annual* of *premium frequency*, category *middle* of *underwriting age*) and time-varying effects (category *other* of *premium frequency*, category *old* of *underwriting age*, *accidental death rider*, *annual premium* and *Dow Jones*). Tables 16 and 17 in Appendix B provide the numerical outputs of this model. The fact that half of the covariates seem to have a constant effect comforts us in the opportunity to apply the F&G model further, where F&G model is less flexible but more parsimonious (and is thus likely to provide more robust predictions).

Regarding time-varying effects of this intermediate model, we plot in Figure 6 the intercept and the regression coefficients of the five corresponding variables (namely *death rider*, *premium frequency*, *underwriting age*, *annual premium* and *Dow Jones variation*). Notice that the effect of the variation of the Dow Jones Index tends to explode for highest durations. This was also the case with the fully nonparametric approach. The confidence we can have in this estimation is low: the highest the duration, the widest the confidence interval. This is due to the fact that there are very few events observed for such durations. In practice, this means that the propensity to surrender due to the variation of the Dow Jones Index is likely to be largely overestimated for highest durations.

Let us now focus on the estimation of the F&G proportional hazards model of Equation (5). Beginning with all available covariates, we fit the model by a backward stepwise approach. As in the fully nonparametric modelling, the *living place* is not significant (associated p-values for each category exceed 20%). This is in line with our conclusions following descriptive statistics in Section 2. The next step thus consists in performing the estimation without this covariate, still embedding only parametric terms. This procedure comes up with a model that is globally statistically significant, and where every covariate also has a statistically significant effect. In other words, this is one of the best model in this model class. The results of the F&G model are stored in Table 6.

All the regression coefficients in Table 6 are estimated in comparison to a reference profile, i.e. a

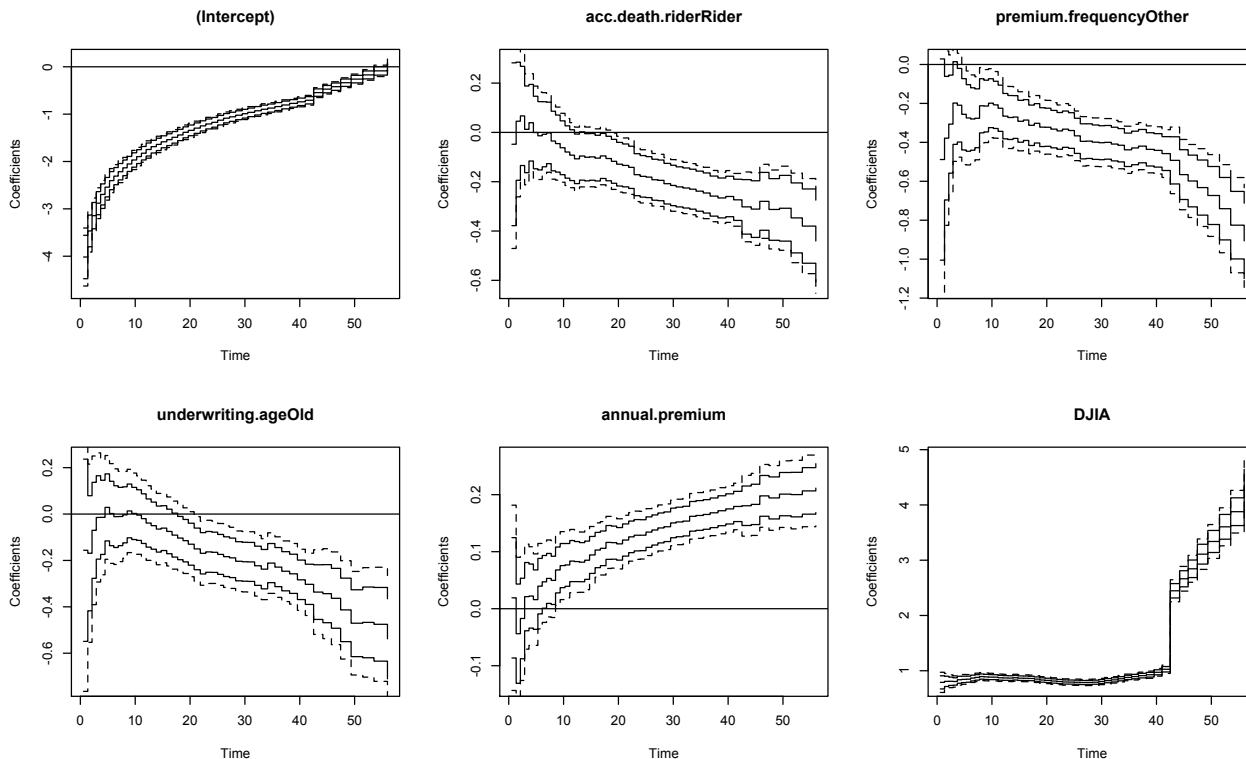


Figure 6: Estimation of time-varying regression coefficients in the intermediate model.

Covariate	Coefficient	Standard Error	Gray's test statistic	p-value
Accidental Death Rider - Yes	-0.191	0.0371	-5.16	2.5×10^{-7}
Gender - Female	-0.0784	0.0256	-3.06	2.2×10^{-3}
Premium Frequency - Annual	-0.222	0.0316	-7.03	2.13×10^{-12}
Premium Frequency - Other	-0.395	0.0398	-9.94	0
Risk State - Smoker	-0.136	0.0269	-5.04	4.7×10^{-7}
Underwriting Age - Middle	0.121	0.0293	4.12	3.83×10^{-5}
Underwriting Age - Old	-0.2	0.0389	-5.14	2.73×10^{-7}
Annual Premium	0.142	0.0116	12.30	0
Dow Jones variation	0.889	0.0186	47.90	0

Table 6: Estimated constant regression coefficients in the F&G model.

policyholder with the following characteristics in our case: a young non-smoker male with an infra-annual premium frequency, without the rider of accidental death, and whose annual premium equals \$560.8 and last relative DJIA variation equals 0.178 % (cf Table 3).

Table 6 gives some interesting insights in terms of interpretation. Firstly, people with an accidental death rider tend to have a lower propensity to surrender whatever the current lifetime of their contract, and this is also the case for those having a low premium frequency. The lower the frequency, the lower the intensity of surrender. Practitioners' intuition is in line with the aforementioned statements. Secondly, women, as well as smokers, seem to have a lower propensity to surrender. This is less intuitive, and there is no well-known consensus on the sense of impact for these risk factors. Anyway, looking at the absolute values of regression coefficients, these effects seem to be less pronounced

than the others. Thirdly, speaking about the underwriting age, people from 35 to 54 years old are more likely to surrender than others: this makes sense since this age range corresponds to a period in which people often invest in personal projects (this remark was already made in the descriptive analysis). Fourthly, the coefficient related to annual premium indicates that the behaviour of richer policyholders differs from the others: they are more likely to surrender their contract. In an extreme case we know that the richest people have personal advisors, which obviously plays a big role on their behaviour. Finally, the Dow Jones has the most prominent impact: although averaged as compared to its maximum value in the intermediate model ($\simeq 4$, see Figure 6), a rise in Dow Jones variation has a deep impact on the surrender intensity, making it increase significantly. Once again, it seems natural given the type of contract under study. Indeed these contracts are not index-linked, which means that their owners do not directly benefit from such an increase. This could cause further frustration and trigger some surrenders.

5 Assessment of the predictive power

Statistically speaking, we have validated the usual tests to ensure the significance of the selected models. We now have to validate their predictive power on a new independent test set. Adopting a backtesting approach, we compare the predictions given by the three selected models (only time-varying regression coefficients, mix of time-varying and constant coefficients, and the F&G model with only constant effects) to real-life observations.

To begin with, one has to check whether the main characteristics of the independent test set are similar to those of the learning set. There are 9772 policyholders in this validation sample, which represents one third of the full dataset. The proportions of each category for all other covariates are similar in both subsamples, which ensure that the subsamples were randomly built. The censoring rate equals 49.77%, whereas there was 49.06% of lapsed contracts in the learning sample. Therefore, more than half of the contracts in the test sample have already experienced one of the events that triggers a lapse: death, maturity, default on premium payments, surrender, and so on. Among those 4908 lapses, 3628 were surrenders and 444 were deaths. Here, the subdistribution approach should reveal useful since only 4.54% of events correspond to death: a survival model for this cause would thus be quite tricky to fit, affecting the estimation of intensities related to other causes, and causing compensations to get a good final fit on the overall hazard rate.

5.1 Overall quality of the predictions

Firstly, recall that the model was fitted on a quarterly basis for the estimation of lifetimes. For each quarter and each individual i , one computes the probability that the policyholder makes the decision to terminate her contract in this period. To do so, Equation (6) is useful: indeed, we want to estimate the conditional probability

$$P(d_1 < T_i \leq d_2, J_{T_i} = 1 \mid \mathbf{X}_i, T_i > d_1), \quad (7)$$

where d_1 and d_2 are durations that respectively correspond to the beginning and end of the period, for the i^{th} policy under study. For instance, consider a policy issued on the 1st of May, 2005, and say that each year is divided into four periods with the following quarters: from 01-01-XXXX to 03-31-XXXX, from 04-01-XXXX to 06-30-XXXX, from 07-01-XXXX to 09-30-XXXX, and from 10-01-XXXX to 12-31-XXXX. To compute the individual propensity to surrender in the third quarter of 2005, d_1 and

d_2 would respectively equal $(2/3)$ and $(1 + 2/3)$. Concretely, Equation (7) yields to

$$\begin{aligned} P(d_1 < T_i \leq d_2, J_{T_i} = 1 \mid \mathbf{X}_i, T_i > d_1) &= \frac{P(d_1 < T_i \leq d_2, J_{T_i} = 1, T_i > d_1 \mid \mathbf{X}_i)}{P(T_i > d_1, J_{T_i} = 1 \mid \mathbf{X}_i)} \\ &= \frac{F_{T,1}(d_2; \mathbf{X}_i) - F_{T,1}(d_1; \mathbf{X}_i)}{1 - F_{T,1}(d_1; \mathbf{X}_i)}. \end{aligned} \quad (8)$$

The estimation of all the quantities in Equation (8) results from the fitted model, see Equation (5) where estimators given in Table 6 replace the parameters. Indeed, the CIF for surrenders is linked to the surrender hazard rate thanks to Equation (4). The prediction of the surrender rate, denoted further by \hat{r}_t , is then deduced by summation of estimated probabilities within each period (and divided by the exposure on the same period):

$$\hat{r}_t = \frac{1}{n_t} \sum_{i \in \mathcal{R}_t} \hat{P}(t < T_i \leq t + 1, J_{T_i} = 1 \mid \mathbf{X}_i, T_i > t), \quad (9)$$

where \mathcal{R}_t denotes the population at-risk at the beginning of quarter t , and n_t its size.

Figure 7 illustrates this aggregation. The solid line is the observed surrender rate. Notice that the predictions given by the three built models are similar until the beginning of 2005. This is roughly because the estimators of regression coefficients in all three models are comparable for durations below 40 quarters (or equivalently ten years). However, as soon as the duration of the contract exceeds this threshold, there is a huge difference when taking into account the effects due to the variations of the Dow Jones Index (see the last graph in Figure 6, and the discussion about the *intermediate model* in Section 4). In particular, an increase of the DJ index may have a (very) big impact on the propensity to surrender for “oldest” contracts, both in the *intermediate* and *nonparametric* models. That explains the sudden increase of the surrender rate before the financial crises (a period in which there was an important rise of the DJ index), and then its decrease in 2008. The effect is not negligible since more than one third of the contracts, exactly 3519, have a duration higher than 40 quarters in the test sample. On the contrary, the F&G model (green dot-dashed line) moderates this impact since the associated coefficient remains constant for every duration and consequently for the whole period. In reality, it seems that the F&G model works much better in terms of predictions: indeed the temporary surrenders before the financial crisis were largely overestimated in the *nonparametric* and *intermediate models*, whereas the predicted surrender rate by F&G still has the same pattern as the observed one at the end of the observed period.

Another way to compare those predictions is to introduce a numerical indicator representing the difference between the models: the *Area Under the Trajectory* (AUT). This area could easily be linked to another indicator concerning the liquidity risk faced by the insurer, knowing that the latter is forced to give back immediately to the policyholder her surrender value. Table 7 summarizes these information, and provides some measures of errors within each model. Relative errors are first computed w.r.t. the whole observed trajectory, and then w.r.t. the experienced surrender rates⁴. As expected, the F&G model is better than others, but it tends to slightly underestimate the global risk.

To sum up, the overall quality of the F&G model is rather satisfying when considering the timing of individual surrender decisions. In an Asset and Liabilities Management perspective, this is an important result since it gives precise information about the future cash-flows of the insurer. Nonetheless,

⁴see Appendix A for details.

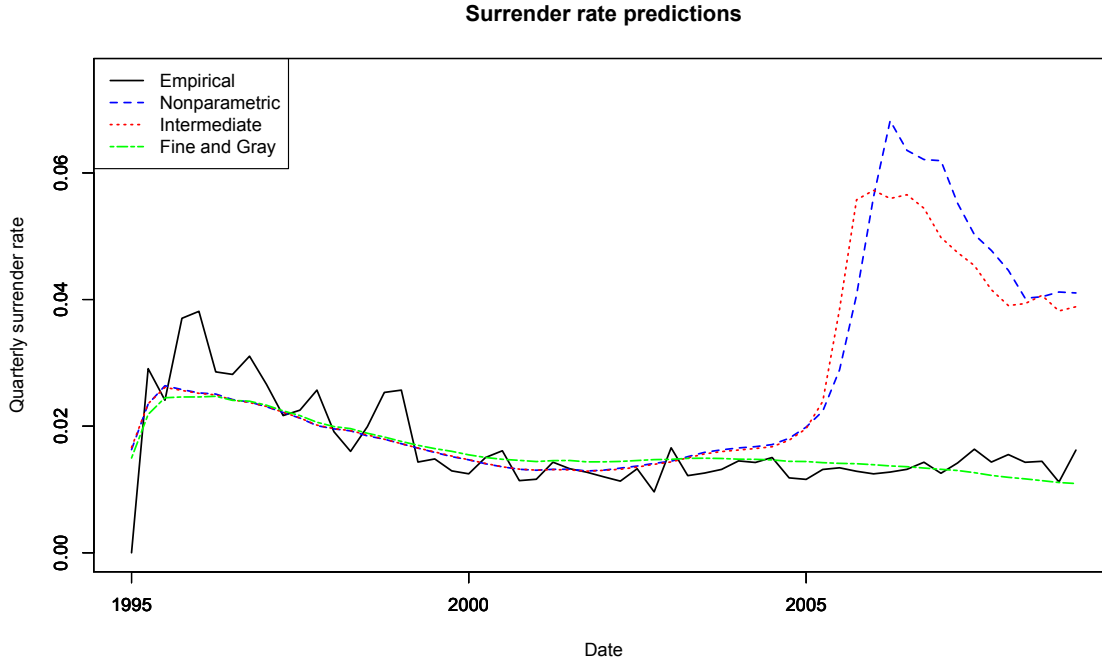


Figure 7: Comparison between the observed and predicted surrender rates obtained by different models on the test sample, from 01-01-2005 to 12-31-2008.

	Empirical surrender rate	Nonparametric model	Intermediate model	F&G model
Area under the trajectory (AUT)	0.94	1.47	1.42	0.92
Absolute relative error of AUT	0%	60.3%	55.5%	2.1%
Mean abs. rel. error of rates		0.844	0.787	0.155
Mean sq. rel. error of rates		2.122	1.748	0.036

Table 7: Areas and errors of the observed and estimated curves for the surrender rate (see Figure 7).

this method of aggregation is not comprehensive enough to check for the quality of the predictions individually. Concretely one would like to identify the cases where the model fails, i.e. detect the risk profiles for which the predicted probabilities to surrender for a given contract duration are significantly different from the observed ones. This is the aim of the coming section.

5.2 Prediction by policyholder’s profile

One focuses here on predicting the surrender rates by risk profile, still in the test set. The estimated surrender rate for a given profile is once again computed by Equation (9), but the at-risk population \mathcal{R}_t becomes the set of policyholders present at the beginning of quarter t and having a given set of characteristics.

A first attempt when selecting profiles is to make the exhaustive combination of all explanatory variables. This results in a list of 72 profiles: however, many profiles represent a tiny proportion in

Pf. #	Gender	Prem. Freq.	Risk state	UW age	Premium	DJIA	Death rider	Percent.
1	Male, Female	Infra, Other	NonSmoker	Young	\$560.88	0.178 %	Yes, No	22.24 %
2	Male, Female	Annual	NonSmoker	Young	\$560.88	0.178 %	Yes, No	6.74 %
3	Male, Female	Infra, Other	Smoker	Young	\$560.88	0.178 %	Yes, No	12.55 %
4	Male, Female	Annual	Smoker	Young	\$560.88	0.178 %	Yes, No	4.32 %
5	Male, Female	Infra, Other	NonSmoker	Middle	\$560.88	0.178 %	Yes, No	16.4 %
6	Male, Female	Annual	NonSmoker	Middle	\$560.88	0.178 %	Yes, No	5.18 %
7	Male, Female	Infra, Other	Smoker	Middle	\$560.88	0.178 %	Yes, No	10.05 %
8	Male, Female	Annual	Smoker	Middle	\$560.88	0.178 %	Yes, No	2.88 %
9	Male, Female	Infra	NonSmoker	Old	\$560.88	0.178 %	Yes, No	7.69 %
10	Male, Female	Annual	NonSmoker	Old	\$560.88	0.178 %	Yes, No	3.07 %
11	Male, Female	Infra	Smoker	Old	\$560.88	0.178 %	Yes, No	4.33 %
12	Male, Female	Annual	Smoker	Old	\$560.88	0.178 %	Yes, No	1.63 %

Table 8: Table of selected profiles for the analysis. Common and differing risk factors are reported, with associated proportions in the test population.

the test set (from 0.08187 % to 7.726 %). In a second attempt to select some relevant risk profiles, we have chosen to group together some of these profiles. We thus consider fewer covariates on which to differentiate profiles, or fewer categories for categorical covariates. Looking at results stored in Table 6, we select profiles according to three highly significant risk factors : *premium frequency*, *risk state* and *underwriting age*. We obtain 18 risk profiles, among which the 12 ones listed in Table 8. Note that the proportion of each selected profile in the validation population is given in the last column. Among these profiles, we select the three most common profiles (#1, 3, 5) and the three least common profiles (#8, 10, 12), see bolded numbers in Table 8. Let us note that the three most common profiles represent 51.2% of the test set while the three least common ones only stand for 7.6 % of the size. From a marketing perspective, the three most common profiles would be chosen as the first targets of customer relation management (advertisement campaign), while the three least common profiles would be the last targets.

For the six selected profiles, we plot the estimated and observed surrender rates in Figures 8 and 9. In the first series in Figure 8, we observe that the estimated surrender rates are relatively stable before Quarter 40. After that, the predictions by *nonparametric* and semi-parametric (*intermediate*) are particularly volatile, unlike the F&G model which provides stable outputs (this was already depicted in Figure 7).

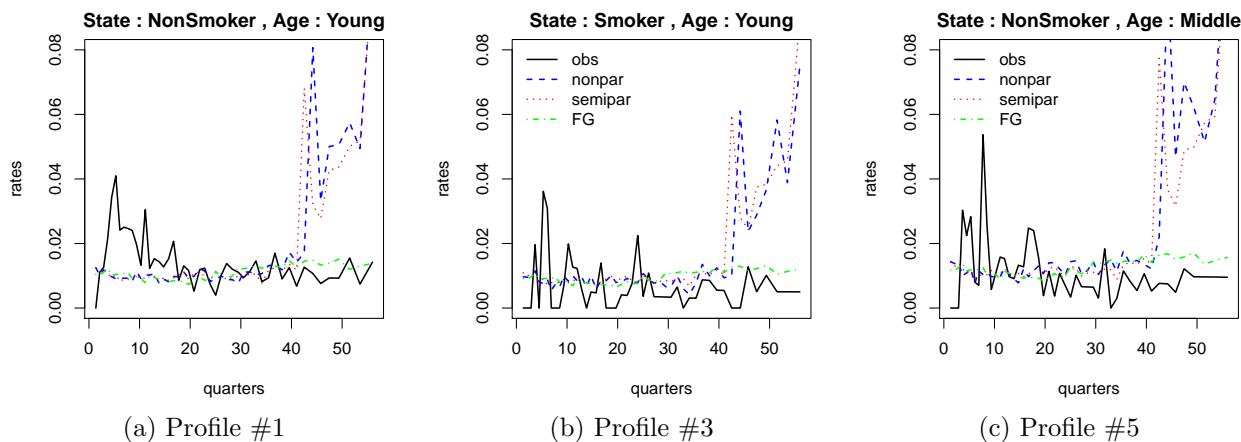


Figure 8: Common features: non annual premium frequency, male or female, with or without rider.

Figures 8a and 8b differ from the risk state, respectively non-smoker and smoker. The observed rates are naturally more volatile in 8b since the associated population is smaller, but the predicted rates are almost the same. Figures 8a and 8c differ from the underwriting age, respectively young and middle. Predicted and observed rates are slightly decreasing before Quarter 20.

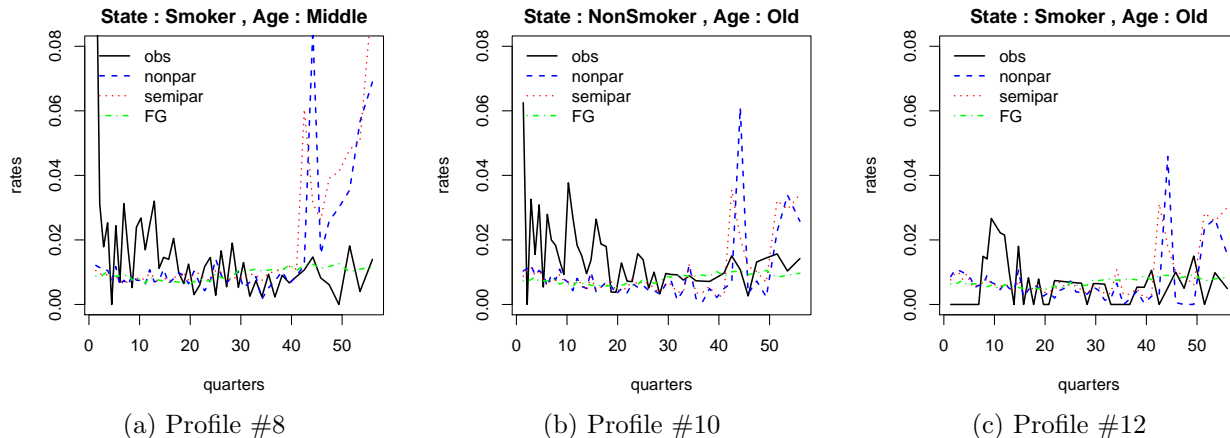


Figure 9: Common profile features: annual premium frequency, male or female, with or without rider.

For the three Figures 9a, 9b, 9c, the observed rates (solid lines) are very erratic with a slight decreasing trend for 9a, 9b. The predictions by the *nonparametric* and semiparametric (*intermediate*) models (respectively dashed and dotted lines) are more stable than in Figure 8, even after Quarter 40. In fact, the regression coefficients associated to these risk profiles in both models compensates more than with other profiles the effect of the DJIA variation. As always, the F&G model (dot-dashed line) provides stable estimated surrender rates, due to its constant effect of DJIA variation. Clearly, it seems in Figure 9c that the F&G model overestimates the surrender rates, but one has to keep in mind that the profile #12 only represents a small part (1.63%) of the test sample.

Furthermore, let us complete this graphical analysis with the assessment of mean absolute relative errors and mean squared errors for the six selected profiles (see Appendix A for details). Results are stored in Table 9. Irrespective of the metrics considered, the F&G model remains the best model for all the profiles. However, the errors of the *nonparametric* and *intermediate* models are less pronounced for the three least exposed profiles than for others (particularly for profile #10). This is in line with what was observed in Figures 8 and 9.

Notice also that the best predictions of the F&G model are obtained for two antagonistic profiles #1 and #12, respectively non-annual premium, non smoker, young and annual premium, smoker, old profiles. So the F&G model does not need a large number of individuals to provide reasonable and fair predictions, and is relatively robust to changes of risk factors.

Finally, we select the individuals in the validation sample according to their underwriting year. In Figure 10, we plot the observed and predicted surrender rates for such cohorts. The same conclusions apply here: it seems that whatever the underwriting year, the predictions for every models are pretty right before Quarter 40 but fail in the *nonparametric* and *intermediate* cases for longer durations (see Figures 10a, 10b and 10c). Once again, the coefficient related to the Dow Jones variation must be responsible for these unrobust results.

Profile		Mean absolute relative error			Mean squared relative error		
Num.	Perc.	nonparametric	intermediate	F&G	nonparametric	intermediate	F&G
1	22.24%	1.045	0.973	0.411	3.884	3.099	0.263
3	12.55%	1.762	1.939	0.923	11.858	14.331	1.329
5	16.40%	1.907	1.652	0.836	11.992	8.329	1.388
8	2.88%	1.167	1.352	0.762	5.672	5.967	1.235
10	3.07%	0.656	0.659	0.470	0.893	0.718	0.379
12	1.63%	0.922	0.706	0.426	2.864	1.470	0.243

Table 9: Relative errors for the six selected profiles.

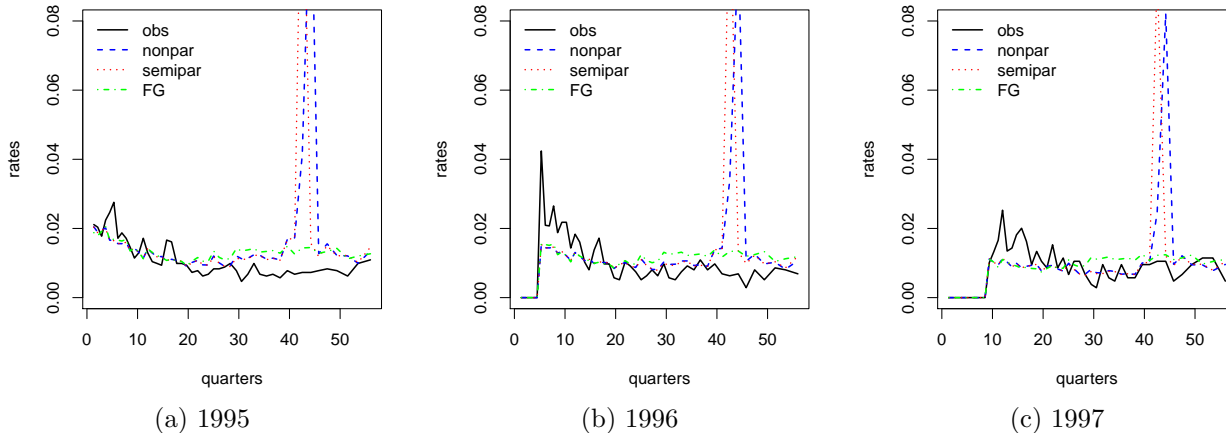


Figure 10: Surrender rates prediction by underwriting year.

6 Experimental surrender tables

In this section, the aim is to propose an experimental table based on historical data. In the same spirit as experimental mortality tables, this tool could help operational teams to manage the surrender risk in a day-to-day task and make some easier ALM predictions.

In our context, the table provides the surrender rate by duration of the contract (in month), for the 14 first years (our experience here). It corresponds to surrender rates for the reference profile of the portfolio, which means that some adjustments have then to be made depending on the profile of the policyholder. In other words, it represents the average surrender risk for the current portfolio composition. Table 10 is easily deduced from the estimators of Table 6 and formulas of the CIF, and has to be updated on a regular basis (each time the underlying model is updated, once a year in practice for instance).

Recall that the reference profile is the following: a young non-smoker male with an infra-annual premium frequency, and without the rider of accidental death. Corresponding CIFs are plotted in solid lines in Figures 11a, 11b, 11c, for the F&G model. In order to make some assumptions about the expected contract lifetime for other profiles, Table 11 gives the multiplicative coefficients to apply on the mean risk represented by Table 10. It is straightforward, thanks to Taylor expansions, to compute these coefficients thanks to the ratio of (1-estimated CIF) for two individuals differing only on one given risk factor. For a given covariate X_k , this k^{th} coefficient is roughly equal to $(\exp(\hat{\beta}_k))$: in the case of a continuous covariate, this is valid for a 1-unit increase of the covariate, otherwise this is the

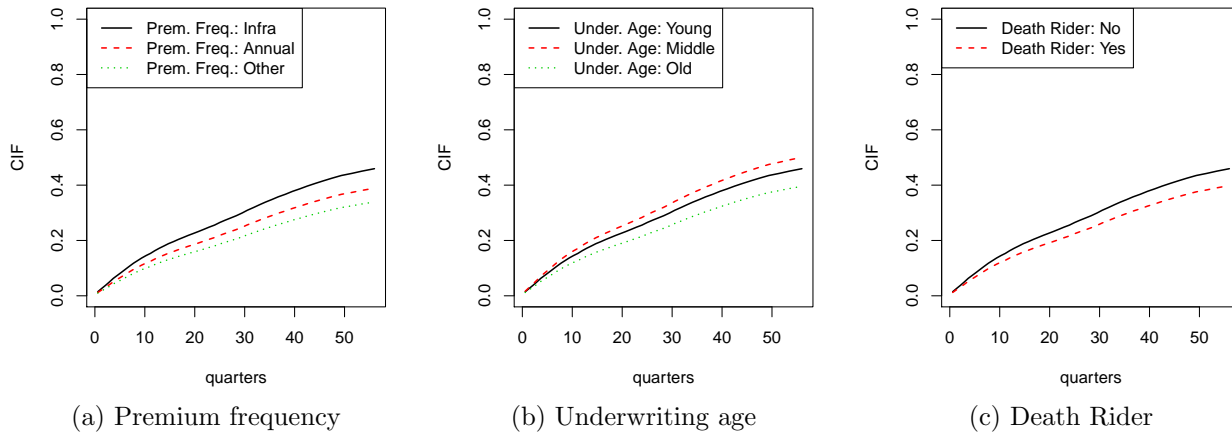


Figure 11: Comparison of CIF for the Fine & Gray model

coefficient to apply when changing the category from the reference one to the one under study. As this coefficient does not depend on time (here the duration of the contract), it can be applied to all the rates in Table 10 without any difference. To sum up, surrender rates have to be corrected if the portfolio composition evolves, or if the (economic and financial) context changes.

# Months	Rate (%)											
1-12	0.8500	0.7665	0.4990	0.5199	0.5052	0.5057	0.5155	0.5505	0.5640	0.5904	0.5738	0.4907
13-24	0.4824	0.4858	0.4741	0.4927	0.5071	0.5097	0.4925	0.5104	0.5097	0.5034	0.5070	0.4349
25-36	0.4368	0.4398	0.4429	0.4311	0.3982	0.3928	0.3839	0.3373	0.3514	0.4092	0.4476	0.4342
37-48	0.3966	0.3970	0.3878	0.3470	0.3615	0.3665	0.3752	0.3766	0.3103	0.3063	0.3085	0.3293
49-60	0.3267	0.3390	0.3125	0.3085	0.2993	0.3154	0.3177	0.3187	0.2878	0.2924	0.2946	0.3006
61-72	0.2912	0.3024	0.3163	0.3200	0.3223	0.3168	0.3323	0.3347	0.3212	0.2930	0.3045	0.3162
73-84	0.4045	0.3940	0.4078	0.3591	0.3536	0.3439	0.3409	0.3338	0.3336	0.3263	0.3413	0.3411
85-96	0.3479	0.3970	0.3844	0.4001	0.4273	0.4621	0.4470	0.4317	0.3857	0.3872	0.3770	0.4048
97-108	0.4153	0.4155	0.4024	0.4041	0.4042	0.4058	0.3728	0.3681	0.3695	0.3708	0.3875	0.3860
109-120	0.4014	0.4014	0.3470	0.3248	0.3368	0.3395	0.3691	0.3800	0.3941	0.3973	0.3780	0.3232
121-132	0.3355	0.3367	0.3394	0.3650	0.3566	0.3693	0.3707	0.3555	0.3236	0.3363	0.3374	0.3403
133-144	0.3229	0.3257	0.3267	0.3295	0.3170	0.3231	0.3190	0.3200	0.3107	0.3220	0.3092	0.3014
145-156	0.3041	0.2945	0.3042	0.3069	0.2229	0.2092	0.2167	0.2172	0.2177	0.2110	0.2312	0.2425
157-168	0.2449	0.2364	0.2442	0.2467	0.2327	0.2078	0.2173	0.2160	0.2164	0.2169	0.2118	0.2104

Table 10: Surrender rates for each month of lifetime for the reference profile in the portfolio.

Covariate	Coefficient to be applied	Type of effect	How much?
Accidental Death Rider - Yes	0.8261	decrease risk	$\simeq 17\%$
Gender - Female	0.9245	decrease risk	$\simeq 8\%$
Premium Frequency - Annual	0.8009	decrease risk	$\simeq 20\%$
Premium Frequency - Other	0.6736	decrease risk	$\simeq 33\%$
Risk State - Smoker	0.8728	decrease risk	$\simeq 13\%$
Underwriting Age - Middle	1.1286	increase risk	$\simeq 13\%$
Underwriting Age - Old	0.8187	decrease risk	$\simeq 18\%$
Annual Premium (+1 unit)	1.1525	increase risk	$\simeq 15\%$
Dow Jones variation (+1 unit)	2.4326	increase risk	$\simeq 143\%$

Table 11: Coefficients to use to adjust the risk from the reference profile.

7 Conclusion

In this paper, one raises the question about the timing of surrenders. This is key in a risk management perspective since initial costs for the insurer cannot be recovered in case of early lapses. Competing risk models, and more precisely Fine & Gray’s subdistribution approach, reveals to be quite efficient and appropriate. Results show that the trajectory of the surrender rate is well predicted in the test sample, and the regression framework allows to identify some crucial impacts coming from certain risk factors.

All along this work, an important assumption is made on the different causes of lapses: we consider that those causes are independent and mutually exclusive. For future research, an extension could thus be to incorporate some correlation between two (or more) causes. Practically speaking, it makes sense to introduce this effect: for instance, an old person who needs money to pay for medical expenses related to a deterioration of her health would be more likely to surrender her contract. This way, surrender and death are two correlated causes of lapse.

References

- Bacinello, A. R. (2005), ‘Endogenous model of surrender conditions in equity-linked life insurance’, *Insurance: Mathematics and Economics* **37**, 270–296.
- Bacinello, A. R., Biffis, E. & P., M. (2008), ‘Pricing life insurance contracts with early exercise features’, *Journal of Computational and Applied Mathematics* .
- Buchardt, K. (2014), ‘Dependent interest and transition rates in life insurance’, *ime* **55**, 167–179.
- Buchardt, K., Moller, T. & Schmidt, K. (2015), ‘Cash flows and policyholder behaviour in the semi-markov life insurance setup’, *Scandinavian Actuarial Journal* **2015**, 660–688.
- Cox, S. H. & Lin, Y. (2006), Annuity lapse rate modeling: tobit or not tobit?, *in* ‘Society of actuaries’.
- Fine, J. & Gray, R. (1999), ‘A proportional hazards model for the subdistribution of a competing risk’, *Journal of the American Statistical Association* **94**(446), 496–509.
- Fleming, T. & Harrington, D. (2013), *Counting Processes and Survival Analysis*, Wiley.
- Gray, B. (2014), *cmprsk: Subdistribution Analysis of Competing Risks*. R package version 2.2-7.
URL: <https://CRAN.R-project.org/package=cmprsk>
- Gray, R. (1988), ‘A class of k-sample tests for comparing the cumulative incidence of a competing risk’, *The Annals of Statistics* **16**, 1141–1154.
- Kim, C. (2005), ‘Modeling surrender and lapse rates with economic variables’, *North American Actuarial Journal* pp. 56–70.
- Kuo, W., Tsai, C. & Chen, W. (2003), ‘An empirical study on the lapse rate: the cointegration approach’, *Journal of Risk and Insurance* **70**, 489–508.
- Laurent, J., Norberg, R. & Planchet, F. (2016), *Modelling in life insurance - A management perspective*, EAA Series, Springer.

- Martinussen, T. & Scheike, T. (2006), *Dynamic Regression Models for Survival Data*, Springer.
- Milhaud, X. (2013), ‘Exogenous and endogenous risk factors management to predict surrender behaviours’, *ASTIN Bulletin* **43**(3), 373–398.
- Outreville, J. F. (1990), ‘Whole-life insurance lapse rates and the emergency fund hypothesis’, *Insurance: Mathematics and Economics* **9**, 249–255.
- Pepe, M. (1991), ‘Inference for events with dependent risks in multiple endpoint studies’, *Journal of the American Statistical Association* **86**, 770–778.
- R Core Team (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: <http://www.R-project.org>
- Russell, D., Fier, S., Carson, J. & Dumm, R. (2013), ‘An empirical analysis of life insurance policy surrender activity’, *Journal of Insurance Issues* **36**(1), 35–37.
- Scheike, T. & Zhang, M.-J. (2011), ‘Analyzing competing risk data using the r timereg package’, *Journal of Statistical Software* **38**, 1–15.
URL: <http://www.jstatsoft.org/v38/i02/>

A Error definition

Consider a set of predicted values $\hat{x}_1, \dots, \hat{x}_n$ for a set of observations x_1, \dots, x_n . Let $p = 1$ or 2 . We define the i th squared/absolute relative error as

$$e_i^p = \begin{cases} \frac{|\hat{x}_i - x_i|^p}{x_i} & \text{if } x_i \neq 0, \\ 0 & \text{if } x_i = 0 \end{cases}$$

Therefore, the mean and the sum of relative errors are respectively given by

$$\frac{\sum_{i=1}^n e_i^p}{\sum_{i=1}^n 1_{x_i \neq 0}}, \quad \sum_{i=1}^n e_i^p.$$

The area under the trajectory for points x_1, \dots, x_n is obtained

$$AUT(x_1, \dots, x_n) = h \sum_{i=1}^n x_i,$$

for a step h of the computing grid.

B Additional regression outputs of Section 4

Here is the test for non-significant effects in the competing risk framework, based on the supremum-test (only nonparametric terms in this modelling):

Covariate	Supremum test	p-value: $H_0 : \beta(t) = 0$
(Intercept)	6.34	0
acc.death.riderRider	6.97	0
genderFemale	3.31	0.04
premium.frequencyAnnual	7.62	0
premium.frequencyOther	10.60	0
risk.stateSmoker	5.64	0
underwriting.ageMiddle	5.32	0
underwriting.ageOld	7.09	0
living.placeEastCoast	2.51	0.02
living.placeWestCoast	2.01	0.44
annual.premium	12.20	0
DJIA	16.40	0

Table 12: Significance tests in the fully nonparametric model.

Now, let us present the results of significance when removing the category *WestCoast* of the covariate *living place*, still in the fully nonparametric model:

Covariate	Supremum test	p-value: $H_0 : \beta(t) = 0$
(Intercept)	6.34	0
acc.death.riderRider	6.97	0
genderFemale	3.30	0
premium.frequencyAnnual	7.60	0
premium.frequencyOther	10.5	0
risk.stateSmoker	5.62	0
underwriting.ageMiddle	5.34	0
underwriting.ageOld	7.08	0
living.placeCoasts	2.05	0.26
annual.premium	12.20	0
DJIA	16.40	0

Table 13: Significance tests in the second step of the fully nonparametric model.

Finally, here are the results of significance in the optimized fully nonparametric model:

Covariate	Supremum test	p-value: $H_0 : \beta(t) = 0$
(Intercept)	6.28	0
acc.death.riderRider	6.97	0
genderFemale	3.29	0
premium.frequencyAnnual	7.61	0
premium.frequencyOther	10.5	0
risk.stateSmoker	5.68	0
underwriting.ageMiddle	5.32	0
underwriting.ageOld	7.1	0
annual.premium	12.20	0
DJIA	16.40	0

Table 14: Significance tests in the optimized fully nonparametric model.

Within this modelling, we can test for time invariant effects:

The following table illustrates the global significance of the optimized intermediate model:

Covariate & category	Kolmogorov-Smirnov (KS) test	p-value KS test H_0 : constant effect	Cramer Von Mises (CVM) test	p-value CVM test H_0 : constant effect
(Intercept)	2.79	0	49.3	0
acc.death.riderRider	0.288	0	0.962	0
genderFemale	0.0958	0.26	0.058	0.26
premium.frequencyAnnual	0.122	0.06	0.209	0.02
premium.frequencyOther	0.55	0	2.44	0
risk.stateSmoker	0.123	0.34	0.22	0.08
underwriting.ageMiddle	0.105	0.56	0.0664	0.42
underwriting.ageOld	0.338	0	1.31	0
annual.premium	0.187	0	0.201	0
DJIA	3.5	0	98.9	0

Table 15: Tests for time invariant effects in the optimized fully nonparametric model.

Covariate	Supremum test	p-value: $H_0 : \beta(t) = 0$
(Intercept)	25.40	0
acc.death.riderRider	6.44	0
premium.frequencyOther	10.30	0
underwriting.ageOld	6.85	0
annual.premium	13.30	0
DJIA	44.30	0

Table 16: Significance tests in the optimized intermediate model.

And the results of the optimized *intermediate* model:

Covariate & category:	Indicators depending on the type of effect:			
Nonparametric terms:	KS test	p-value KS test	CVM test	p-value CVM test
(Intercept)	2.5	0	36.1	0
acc.death.riderRider	0.266	0.02	0.867	0
premium.frequencyOther	0.527	0	2.1	0
underwriting.ageOld	0.368	0	1.55	0
annual.premium	0.177	0	0.197	0
DJIA	3.480	0	99.4	0
Parametric terms:	Coefficient	Standard error	Statistic value	p-value
genderFemale	-0.0819	0.0267	-3.07	2.14×10^{-3}
premium.frequencyAnnual	-0.211	0.0329	-6.41	1.45×10^{-10}
risk.stateSmoker	-0.133	0.028	-4.75	2.03×10^{-6}
underwriting.ageMiddle	0.1240	0.0304	4.09	4.27×10^{-5}

Table 17: Optimized intermediate model: regression coefficients for the parametric terms, and tests for time invariant effects for remaining covariates (corresponding coefficients are functions, see Figure 6).