MRIC | Working Papers

www.mric.lmu.de/research/wp

# Estimating extreme cancellation rates in life insurance

Francesca Biagini

Tobias Huber

Johannes Jaspersen

Andrea Mazzon

Munich Risk and Insurance Center

Working Paper 33

October 26, 2020

Estimating extreme cancellation rates in life insurance
Francesca Biagini, Tobias Huber, Johannes Jaspersen, and Andrea Mazzon
MRIC Working Paper No. 33
October 2020

# ABSTRACT

This paper assesses the risk of a mass lapse event in life insurance. The rarity of the event and the complexity of policyholder behavior, make the risk assessment of such a scenario difficult. Using a simulation study, we evaluate how different estimation methods can assess the scenario when using panel data at the company level. We then use the best performing method to estimate the probability distribution function of a mass cancellation event in the U.S. and Germany. We identify dependencies of the event on company and country characteristics, which so far are not taken into account by regulating agencies. We also find that the current mass lapse scenario in Solvency II has no empirical foundation for the German market. We show that an empirically valid scenario leads to a significantly lower solvency capital requirement for the average German life insurer.

**Keywords:** Extreme Value Theory · Dynamic Peaks Over Threshold · Life Insurance · Mass Cancellation

**JEL Classification:** C14 · G18 · G22 · G32

Francesca Biagini
Munich Risk and Insurance Center
Workgroup Financial and Insurance Mathematics
Ludwig-Maximilians-Universität München
80539 Munich
biagini@math.lmu.de

Tobias Huber (*Corresponding Author*)
Munich Risk and Insurance Center
Munich School of Management
Ludwig-Maximilians-Universität München
80539 Munich
tobias.huber@lmu.de

Johannes G. Jaspersen
Institut für Versicherungsbetriebslehre
Economics and Management
Leibniz Universität Hannover
30159 Hannover
jgj@ivbl.uni-hannover.de

Andrea Mazzon
Munich Risk and Insurance Center
Workgroup Financial and Insurance Mathematics
Ludwig-Maximilians-Universität München
80539 Munich
mazzon@math.lmu.de

We are thankful for helpful comments from participants at the 2017 annual meeting of the German Insurance Science Association, the 2017 annual meeting of the American Risk and Insurance Association, the 2018 International Congress of Actuaries, and seminar participants at Temple University. We are indebted to Alexander Braun, Evan Eastman, Pierre Joos, Lu Li, Thorsten Moenig, Joëlle Näger, Andreas Richter, Jochen Ruß, Günter Schwarz, and Richard Vierthauer for valuable comments.

# 1 Introduction

Rare events with extreme consequences often have a profound influence on economies and agents within them. The attacks of 9/11 lead to considerable changes in many economic sectors such as the airline industry. Hurricanes like Katrina and Sandy shaped how we see flood protection and insurance. Financial crises like the great depression and the financial crisis of 2009 influence our considerations of financial regulation and economic risks in general. For organizations, extreme events can similarly shape existence. One such extreme event for life insurance companies is the occurrence of a mass cancellation event – that is a large portion or even majority of the policyholders canceling their life insurance policy abruptly. However, even though policyholder cancellation behavior has received considerable attention in the academic literature (e.g., Kuo et al., 2003; Eling and Kiesenbauer, 2013), the question how to model mass cancellation scenarios has only been considered sparsely so far. This is surprising because the possibility of mass cancellations has a large effect on insurance companies' asset liability management and leads to one of the largest financial reserves in the European risk management framework Solvency II (EIOPA, 2011).

The omission in the literature can in part be explained by the data requirements associated with estimating extreme cancellation events. Extreme cancellation rates are rare events so that data sourced from only one insurer are not sufficient to estimate this tail risk. In this paper, we overcome this challenge by employing an estimation approach provided by Chavez-Demoulin et al. (2016), which is based on extreme value theory and can be applied to panel data. We consider cancellation rates as realizations of random variables but do not make any assumptions about their distribution functions, except that they are continuous.[1] The method then enables us to estimate the unknown probability distribution functions and associated risk measures for extreme cancellation events from readily available panel data at the company level.

Dependent on contract characteristics, cancellation can either negatively or positively affect an insurer's profit. Unexpected changes in the level of cancellation rates can lead to liquidity problems, the loss of expected future profits, and unbalanced initial expenses (Kuo et al., 2003; Eling and Kiesenbauer, 2013). Life insurers are thus interested in assessing their exposure to this risk as accurately as possible. U.S. life insurers can potentially face very high cancellation rates, partic-

[1]The continuity assumption enables us to obtain results on the far end tail of the cancellation rates' distribution functions under technically mild assumptions. The literature shows that the conditions for the existence of a limit distribution for discrete distribution functions are strong (Nadarajah and Mitov, 2002; Davison and Huser, 2015; Hitz et al., 2017). Our continuity assumption allows us to consider a more flexible framework. A justification for the continuity assumption is given by the following reasons: (1) We observe no evidence for a discontinuous distribution function in the histograms of cancellation rates stemming from two different countries (U.S. and Germany) between 1996 and 2018 (see Figure 1 and Figure 4). (2) The commonly assumed sources (e.g., unemployment rate or interest rate) of cancellation rates are modeled as continuous random variables.

ularly after premium guarantee periods expire (SOA and LIMRA, 2012, p.29). The possibility of such an extreme event needs to be taken into consideration for a company's asset liability management. In other situations, however, cancellation can increase an insurer's profit as policies are usually front-loaded and a cancellation lets insurers reap the early profits without having to incur expected losses in the later part of the policy's life cycle (Gottlieb and Smetters, 2019). In each case, understanding cancellation behavior is crucial for life insurers in order to ensure an adequate asset liability management.

For the European market, the importance of understanding the cancellation behavior of policyholders has become pivotal with the introduction of the new regulatory framework Solvency II. The framework's solvency capital requirement (SCR) has a great sensitivity towards lapse risk (EIOPA, 2011, p.67). In the standard model of the regulation framework, the so-called mass lapse shock leads companies to have solvency capital requirements in the hundreds of millions (Old Mutual, 2016; UNIQUA, 2017), because insurers have to apply a scenario where 40% of all policyholders cancel their contract in the same year. This assumption neither has an empirical justification for the assumed size of the shock, nor for the fact that the same shock is assumed for all different national markets in Europe. Our model offers a way to use empirical assumptions appropriate for the individual national insurance market in the lapse risk module, which can help limiting the risk of possible over- or underreserving. Previous literature on cancellation events in general has used two approaches – estimating cancellation rates from data directly (e.g., Eling and Kiesenbauer, 2013; Knoller et al., 2016) or calibrating economic models based on specific assumptions about macroeconomic factors and policyholder behavior (Albizzati and Geman, 1994; Bacinello, 2003; Consiglio and De Giovanni, 2010). These approaches differ in their interpretation. Empirical approaches often make fewer assumptions but seldom allow for a causal interpretation of the results. Model-based approaches identify causal relationships but assume specific relationships between policyholder behavior and macroeconomic factors, which may or may not hold. For extreme cancellation events, some theoretical modeling approaches exist (Loisel and Milhaud, 2011; Barsotti et al., 2016), but as of yet no empirical estimations have been reported.

Applying the estimation method to data sourced from U.S. life insurers demonstrates the applicability of our approach and shows the importance of the product type for high cancellation rates in the U.S. market. While for companies with a high share of term life policies mass cancellation scenarios with rates up to 50% can be adequate, for companies with a high share of permanent life insurance policies a mass cancellation scenario is estimated to be approximately 20%. Furthermore, we discuss the calibration of Solvency II's mass cancellation scenario by using German data. Our results show that cancellation rates of 20–25% are reflecting a mass cancellation

scenario in the German life insurance market. These values raise skepticism whether the arbitrarily chosen cancellation rate of 40% for this scenario in Solvency II is adequate for the German life insurance market. In both analyzed markets, the severity of the extreme cancellation event is decreasing in the size of the insurers' portfolio and increasing in the amount of new business. This encourages further research on other company-specific variables. The combination of these results with the fact that the mass cancellation scenario can depend on the predominant product type sold in a market calls into question whether a uniform mass cancellation scenario for all European life insurance markets is appropriate. European life insurance markets differ both in their company characteristics (Insurance Europe, 2019) and in their dominant products (Standard and Poors, 2018) and might thus differ in their mass cancellation scenario as well.

We see the contribution of our paper as complementary to rather than competing with model-based approaches. While our approach does not rely on assumptions about policyholder behavior, we are only able to estimate a cancellation rate distribution, but are unable to identify the causal relationship that leads to the cancellation rate. Model-based approaches can use our results to validate their behavioral assumption by observing whether their models produce similar cancellation rate distributions. Alternatively, our results can be directly used for models of cancellation behavior.

In the following, we first summarize the empirical estimation method and provide a simulation study to examine the approach in a panel data context. We then apply the estimation procedure to U.S. and German data. The next section provides implications for the modeling of extreme cancellation rates and for insurance regulation. The paper ends with concluding remarks and provides directions for future research.

# 2 Methodology

## 2.1 Overview

Our utilized approach is based on considering cancellation rates as a stochastic variable. The main hypotheses prevailing in the literature to explain cancellation behavior are based on macroeconomic variables (such as the interest rate or unemployment rate).[2] As macroeconomic variables can be seen as stochastic processes, and changes in macroeconomic variables lead to changes in cancellation rates, cancellation rates can also be modeled stochastically. However, a major prob-

[2]This is mostly based on two concepts: the interest rate hypothesis and the emergency fund hypothesis. Evidence for the former can be found in Schott (1971), Pesando (1974), and Kuo et al. (2003). Evidence for the latter is reported in Dar and Dodds (1989), Outreville (1990), and Kiesenbauer (2012).

lem to assess the underlying distribution function is posed by the lack of sufficient extreme data. We thus employ the peaks over threshold (POT) method, which exploits the existence of a natural candidate for the conditional distribution function above a high level.

In this section, we first summarize the POT method as it is described in Embrechts et al. (1997) and McNeil et al. (2015). In its original form, the method can only be applied to independent and identically distributed (i.i.d.) data. For our purposes, this would require a long time series of cancellation rates for a single company. Such data, however, are not reliably available. We are limited to using panel data of multiple companies over shorter periods of time. These data feature a dependence structure along two dimensions. Firstly, macroeconomic variables affect all companies at the same time so that it can be assumed that observed cancellation rates within the same time period are correlated. Secondly, companies can have idiosyncratic factors leading to higher or lower cancellation rates, which will make the observations within a company serially correlated over time. To address these issues, this section thus continues by summarizing a dynamic version of the POT method, which was introduced by Coles (2001) and extended by Chavez-Demoulin et al. (2016). The dynamic POT method takes into account the dependency structure of our panel data by modeling the parameters of the estimated probability distribution dependent on time and on a quantitative company-level variable (in our case: the portfolio size). Finally, this section provides a simulation study to evaluate the dynamic POT method in a panel data context.

### 2.2 Peaks over threshold method

We consider a panel of $n$ insurance companies over $T$ periods with cancellation rates $y_{j,s}$, modeled as realizations of the random variables $Y_{j,s}$ ($j = 1, 2, \ldots, n$; $s = 1, 2, \ldots, T$). We first make the strong assumption that all random variables $Y_{j,s}$ are i.i.d. with cumulative distribution function $F$. The classic POT method then allows to reliably estimate the risk of extreme cancellation rates without knowing the common underlying distribution function $F$.

The intuition is to split the distribution function $F$ into two parts based on a "suitably large" threshold $u > 0$. Below the threshold, we are given enough data so that the empirical distribution provides a good fit. Above the threshold, the Pickands-Balkema-de Haan Theorem (see e.g., Balkema and de Haan, 1974; Pickands, 1975) enables us to approximate the excess distribution function of $F$ over $u$ by the generalized Pareto distribution (GPD) function.

We define the *excess distribution function* $F_u$ of $Y_{1,1}$ (all $Y_{j,s}$ are i.i.d.) over a threshold $u$, $0 < u <$

$z_F := \sup\{z \in \mathbb{R} : F(z) < 1\}$, by the conditional probability:

$$F_u(z) := \mathbb{P}(Y_{1,1} - u \leq z | Y_{1,1} > u) = \frac{F(z+u) - F(u)}{1 - F(u)}, \quad 0 \leq z \leq z_F - u. \tag{2.1}$$

Rearranging this equation and defining $\bar{F} := 1 - F$ and $\bar{F}_u := 1 - F_u$, we obtain a method for estimating the far end tail of $F$ by estimating $\bar{F}(u)$ and $\bar{F}_u(z)$:

$$\bar{F}(z+u) = \bar{F}(u) \cdot \bar{F}_u(z). \tag{2.2}$$

We approximate $\bar{F}(u)$ by the empirical distribution function $\bar{F}_N(u) := 1 - F_N(u)$ and approximate $\bar{F}_u(z)$ by a GPD.[3] This distribution function is dependent on a shape parameter $\xi \in \mathbb{R}$ and a scale parameter $\beta > 0$ and is given by:

$$G_{\xi,\beta}(z) = \begin{cases} 1 - \left(1 + \frac{\xi z}{\beta}\right)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp(-\frac{z}{\beta}), & \xi = 0, \end{cases} \tag{2.3}$$

with $z \geq 0$ if $\xi \geq 0$ and $z \in [0, -\beta/\xi]$ if $\xi < 0$. Since the realizations $y_{j,s}$ are i.i.d., $\bar{F}_u$ can be approximated by $\bar{G}_{\widehat{\xi},\widehat{\beta}(u)} := 1 - G_{\widehat{\xi},\widehat{\beta}(u)}$ with estimates $\widehat{\xi}$ and $\widehat{\beta}(u)$ of the threshold-independent shape parameter $\xi$ and the threshold-dependent scale parameter $\beta(u)$ (McNeil et al., 2015).

The estimates $\widehat{\xi}$ and $\widehat{\beta} = \widehat{\beta}(u)$ can be computed by the method of maximum likelihood applied to the set of data $y_{j,s_j}$ with $s_j \in S_j$, where $S_j \subseteq \{1, 2, \ldots, T\}$ denotes the subset of points in time at which the excesses over the threshold $u$ for company $j$ occur. In addition, a maximum likelihood estimate of $\bar{F}_N(u)$ is given by $N_u/N$, where $N_u$ denotes the total number of excesses and $N := nT$ designates the total number of cancellation rates. Given these estimates and a threshold $u$, the estimated quantile $\widehat{q}(\alpha)$ of the unknown distribution function $F$ at the confidence level $\alpha$ is equal to:

$$\widehat{q}(\alpha) = u + \frac{\widehat{\beta}}{\widehat{\xi}} \left[ \left( \frac{1-\alpha}{N_u/N} \right)^{-\widehat{\xi}} - 1 \right]. \tag{2.4}$$

The choice of the threshold $u$ is an important component of the POT method: if the threshold is too low, the exceedances include also non-extreme events and the estimates are biased. In contrast, if the threshold is too high, the number of exceedances is poor and the variance in the statistical estimation is large. To select a threshold, the goodness of fit of empirical excesses over

[3]The empirical distribution function $F_N$ of $F$ given $N = nT$ i.i.d. observations from $n$ companies over $T$ periods is defined as: $F_N(z) := \frac{1}{N} \sum_{j=1}^{n} \sum_{s=1}^{T} \mathbb{1}_{(y_{j,s} \leq z)}, z \geq 0$.

the chosen threshold $u$ to a parametric GPD model can for example be evaluated through a Q-Q plot. In such a graph the quantiles of the log-transformed excesses over $u$ are plotted against the theoretical quantiles of an exponential distribution. If a straight line is observed, then it is empirically confirmed that the GPD provides a good fit of the data. The optimal threshold is chosen as the smallest $u$ for which a good fit is observed.

For commonly available panel data the assumptions of i.i.d. random variables are usually violated. In our later applications, we thus employ a generalization of the POT method for non-i.i.d. data, which enables us to let the frequency and severity of the excesses be dependent on covariates. This generalization is called *dynamic POT method* and was introduced by Coles (2001) for parametric dependencies and extended by Chavez-Demoulin et al. (2016) to semi-parametric and nonparametric dependencies.

### 2.3 Dynamic peaks over threshold method

Again, $\xi$ and $\beta$ are parameters of the GPD modeling the size of the excesses over a large threshold. The number of excesses is assumed to follow a non-homogeneous Poisson process with rate function $\lambda$. The dynamic POT method allows the frequency parameter $\lambda$ and the severity parameters $\xi$ and $\beta$ to be dependent on covariates, which avoids making the assumption that the cancellation rates are realizations of i.i.d. random variables (Coles, 2001; Chavez-Demoulin et al., 2016). In our applications, we use two covariates: the first covariate $x$ represents the insurers' portfolio size and the second covariate $s$ represents time. With the reparameterization $\nu := \log(\beta(1+\xi))$ and general measurable functions $g_\xi$, $g_\nu$, $g_\lambda$: $\mathbb{R}_{\geq 0} \to \mathbb{R}$ and $h_\xi$, $h_\nu$, $h_\lambda$: $[1, T] \to \mathbb{R}$, this leads to the following generalized additive models for the parameters of the dynamic POT method:[4]

$$\xi = \xi(x, s) = g_\xi(x) + h_\xi(s), \tag{2.5}$$

$$\nu = \nu(x, s) = g_\nu(x) + h_\nu(s), \tag{2.6}$$

$$\lambda = \lambda(x, s) = g_\lambda(x) + h_\lambda(s). \tag{2.7}$$

In the above equations, $g_\xi$, $g_\nu$, $g_\lambda$ and $h_\xi$, $h_\nu$, $h_\lambda$ are either linear or smooth functions. In their combination they can model a parametric, semi-parametric, or nonparametric dependence of the parameters on the corresponding covariates. We utilize natural cubic splines for the smooth functions and determine their degrees of freedom and thus their smoothness based on Akaike's infor-

[4]The reparameterization $\nu := \log(\beta(1 + \xi))$ guarantees the convergence of the simultaneous fitting procedure for the two parameters $\xi$ and $\beta$ (Chavez-Demoulin et al., 2016). This reparameterization requires $\xi > -1$, which is fulfilled in our later applications.

mation criterion (AIC) value (see Section 3.1.3 and Chavez-Demoulin et al. (2016, p.746) for more details).

Pooling the $N_u$ excesses $y_{j,s_j}$ under the consideration of the corresponding portfolio size covariate $x = x_{j,s_j}$ and time covariate $s = s_j$, enables us to obtain estimates for $\xi(x, s)$ and $\nu(x, s)$ by employing a penalized maximum likelihood estimation provided by Chavez-Demoulin et al. (2016). Moreover, we can directly estimate the rate $\rho(x, s) := \lambda(x, s)/N(x, s)$ of the non-homogeneous Poisson process through a logistic regression model, where $N(x, s)$ stands for the total number of observations for a fixed covariate $x$ and time point $s$. Subsequently, the estimated $\alpha$-quantile $\widehat{q}(\alpha)$ for a threshold $u$ in dependence of the covariates $x$ and $s$ is given by:

$$\widehat{q}(\alpha)(x, s) = u + \frac{\widehat{\beta}(x, s)}{\widehat{\xi}(x, s)} \left[ \left( \frac{1 - \alpha}{\widehat{\rho}(x, s)} \right)^{-\widehat{\xi}(x,s)} - 1 \right]. \tag{2.8}$$

The applicability of the dynamic POT method has for example been shown in the papers by Chavez-Demoulin et al. (2016), Embrechts et al. (2018), Hambuckers et al. (2018a), and Hambuckers et al. (2018b). Each of these papers applies the dynamic POT method to operational losses and uses covariates to control for the underlying heterogeneity across time and companies. Before we apply the method to estimate extreme cancellation rates, the next section provides a simulation study to examine the dynamic POT method in a panel data context.

## 2.4 Simulation study

### 2.4.1 Overview

We now assess the suitability of the introduced methodology for applications to panel data. For this, we aim to achieve two goals. First, we compare three different approaches with respect to their accuracy: In case (1) we perform the classic POT estimation, in (2) we apply the dynamic POT to the companies' time series jointly, meaning that we take into account the panel structure of the data, and in (3) we employ the dynamic POT estimation to each company's time series separately. Second, we focus on threshold selection. As we have heterogeneity in the parameters across time and within companies, we would optimally choose time-dependent and company-specific thresholds. Due to limited data availability, we are often restricted to choosing one threshold based on all the pooled cancellation rate observations. We therefore examine the effect of such a simplification in the threshold selection on the estimation's accuracy in the dynamic POT method by comparing two approaches: In case (2.1) we choose a joint empirical threshold based on pooling all data, and

in (2.2) we choose an individual threshold for each company separately.

### 2.4.2 Approach

We consider $n = 1,000$ insurance companies over $T$ years ($T = 20, 40, 80$) and simulate cancellation rates for each company $j = 1, 2, \ldots, n$ at time $s = 1, 2, \ldots, T$ according to a log-normal distribution $\mathcal{L}(\mu_{j,s}, \sigma_{j,s})$ with parameters $\mu_{j,s} \in \mathbb{R}$ and $\sigma_{j,s} > 0$. Each company $j$ has at time $s$ a company-specific covariate $x_{j,s} = x_{j,1}$, which varies from company to company but is constant over time. The parameters of the log-normal distribution are dependent on this quantitative covariate and are also affected by a time-specific covariate $t_{j,s} = t_{1,s}$, which is constant across companies but varies over time. The covariates $x_{j,1}$ are chosen equidistantly between 0.2 and 0.4 for the first 900 companies. Additionally, the covariate of company 901 is equal to 0.7 and increases in equidistant steps to 0.8 until company 1,000. These 100 companies, thus, have markedly higher cancellation rates compared to the other companies. The time specific covariate $t_{1,s}$ is equal to 0 at $s = 1$ and increases in equidistant steps to 0.1 until $s = T$. Overall, the parameters are dependent on the company-specific covariate and time-specific covariate in the following way:[5]

$$\begin{aligned} \mu_{j,s} &= x_{j,1} + t_{1,s} - 3.5, \\ \sigma_{j,s} &= 1.1 \cdot x_{j,1} + 2 \cdot t_{1,s} + 0.1. \end{aligned} \tag{2.9}$$

Given the simulated cancellation rates and a threshold $u$, we apply the POT method and fit the excesses at time $s_j$ over this threshold to a GPD. Since the cancellation rate observations follow a log-normal distribution with the parameters $\mu_{j,s}$ and $\sigma_{j,s}$, we can calculate the exact quantile $q_{j,s}(\alpha)$ for a chosen confidence level $\alpha = 99.5\%$ by inverting the log-normal distribution. This enables us to define an error measure to evaluate the accuracy of estimating the $\alpha$-quantile with the POT method. We run the simulation $R = 1,000$ times and obtain in each run $k = 1, 2, \ldots, R$ estimates $\widehat{q}_{j,s_{j,k}}(\alpha)$ of the true quantile $q_{j,s_{j,k}}(\alpha)$ at the corresponding time $s_{j,k} \in S_{j,k} \subset \{1, 2, \ldots, T\}$ of the excess. We define the error measure as the mean relative difference between the estimated and true quantile and take the average over all $R$ runs:

$$err = \frac{1}{R} \sum_{k=1}^{R} \left( \frac{1}{\sum_{i=1}^{n} |S_{i,k}|} \sum_{j=1}^{n} \sum_{s_{j,k} \in S_{j,k}} \frac{|\widehat{q}_{j,s_{j,k}}(0.995) - q_{j,s_{j,k}}(0.995)|}{q_{j,s_{j,k}}(0.995)} \right). \tag{2.10}$$

[5]The specifications in Equation (2.9) are chosen so that the summary statistics of the simulated cancellation rates ($Y_{j,s} \sim \mathcal{L}(\mu_{j,s}, \sigma_{j,s})$) are similar to the cancellation rates observed in our empirical settings introduced below.

### 2.4.3 Results

Table 1 reports the results of the error measure for the different estimation approaches (classic POT, simultaneous dynamic POT with joint empirical threshold, simultaneous dynamic POT with individual empirical threshold, one-by-one dynamic POT) and for different length of the time series ($T = 20, 40, 80$). All these approaches are applicable only for a "suitably large" threshold. We do not have information on which threshold is high enough and therefore choose for all approaches four different thresholds equal to the $0\%$-, $30\%$-, $60\%$-, and $90\%$-quantile of the corresponding cancellation rates. For the classic POT or the dynamic POT with a joint empirical threshold, choosing the threshold equal to a specific quantile means that we have one threshold for all companies. In contrast, for the dynamic POT with individual empirical threshold and for the one-by-one dynamic POT, this means that we have unique thresholds $u_j$ for each company equal to the chosen quantile based on their corresponding cancellation rates. We thus compare for the simultaneous dynamic POT two threshold selection approaches, one ignoring and one acknowledging heterogeneity in individual thresholds.

We begin by focusing on the results for a time series of 40 years (see column 3 in Table 1). First, we compare the two versions of the simultaneous dynamic POT and find that the joint empirical threshold selection approach performs better than the individual empirical threshold approach in all cases with the exception of the 0%-quantile. The individual threshold is based on the empirical quantile of only 40 observations and is therefore affected by outliers. This makes the joint empirical threshold selection approach superior to the individual empirical threshold selection approach. Second, we compare the simultaneous dynamic POT with joint empirical threshold to the one-by-one dynamic POT estimation and find marked differences in the estimation's accuracy. The one-by-one dynamic POT estimation is limited to data within a single time series, which allows a reliable estimation only for many data points. In the case of high thresholds, applying the dynamic POT estimation to each company's time series separately is not possible due to the limited data availability. In cases when the one-by-one dynamic POT estimation is applicable, it leads to a less accurate estimation than the simultaneous dynamic POT with joint empirical threshold. Third, our results in Table 1 demonstrate that the simultaneous dynamic POT always performs better than the classic POT. The classic POT method has limited applicability in the case of data with heterogeneity across time and companies, which is commonly present in company-level panel data.

The results observed for a time series of 40 years are representative for those of time series with a different length (see column 2 and 4 in Table 1). Additionally, we find that the estimation's accuracy improves with the length of the time period for all estimation approaches (with the ex-

ception of the 0%-quantile). This effect is small for the classic POT estimation, considerably larger for the simultaneous dynamic POT estimation, and largest for the one-by-one dynamic POT estimation (see for example the $30\%$-quantile). In principle, the one-by-one dynamic POT estimation will likely get better than the simultaneous dynamic POT estimation if the number of time periods is sufficiently large. In applications, however, such long time series are usually not available or at least not reliably available due to changes in exogenous factors (such as, for example, regulation). In the range of time periods with reliable data, our results in Table 1 support the use of the simultaneous dynamic POT with joint empirical threshold in our empirical settings introduced below.

Table 1: Simulation study results

| | 20 years | 40 years | 80 years |
|---|---|---|---|
| (1) Classic POT | | | |
| 0%-quantile | 0.310 | 0.309 | 0.309 |
| 30%-quantile | 0.859 | 0.854 | 0.853 |
| 60%-quantile | 0.926 | 0.924 | 0.923 |
| 90%-quantile | 0.773 | 0.774 | 0.772 |
| (2) Dynamic POT: simultaneous estimation | | | |
| (2.1) Joint empirical threshold | | | |
| 0%-quantile | 0.310 | 0.310 | 0.308 |
| 30%-quantile | 0.047 | 0.038 | 0.033 |
| 60%-quantile | 0.050 | 0.040 | 0.034 |
| 90%-quantile | 0.062 | 0.049 | 0.040 |
| (2.2) Individual empirical threshold | | | |
| 0%-quantile | 0.056 | 0.054 | 0.155 |
| 30%-quantile | 0.058 | 0.046 | 0.038 |
| 60%-quantile | 0.077 | 0.062 | 0.052 |
| 90%-quantile | 0.191 | 0.172 | 0.160 |
| (3) Dynamic POT: one-by-one estimation | | | |
| 0%-quantile | 0.550 | 0.467 | 0.494 |
| 30%-quantile | – | 0.390 | 0.230 |
| 60%-quantile | – | – | 0.259 |
| 90%-quantile | – | – | – |

*Notes*: The table displays for different estimation approaches (classic POT, simultaneous dynamic POT, one-by-one dynamic POT) and for different length of the time series ($T = 20, 40, 80$) the error in estimating the 99.5%-quantile of the underlying log-normal distribution as defined in equation (2.10). For all approaches we choose four different thresholds equal to the $0\%$-quantile, $30\%$-quantile, $60\%$-quantile, and $90\%$-quantile of all cancellation rates. For the simultaneous dynamic POT we additionally compare two threshold selection approaches, one ignoring and one acknowledging heterogeneity in individual thresholds (joint empirical threshold, individual empirical threshold).

# 3 Empirical analysis

## 3.1 U.S. individual life insurance market

### 3.1.1 Motivation

We employ the dynamic POT model to data from U.S. life insurers to demonstrate the applicability of our approach. Of all countries, the U.S. has the most readily available, reliable, and extensive data. Using only U.S. data does not allow us to make a general statement about life insurance markets but will give us a first indication of the effect of company-level covariates on extreme cancellation events. Moreover, it enables us to calibrate an empirically justified mass cancellation scenario for the U.S. market. U.S. life insurers can potentially face very high cancellation rates, particularly after premium guarantee periods expire (SOA and LIMRA, 2012, p.29). To analyze this risk, we use available panel data at the company level.

### 3.1.2 Company-level, panel data

Our panel data include statistics on individual life insurance, collected from U.S. insurers' annual statement reports by the National Association of Insurance Commissioners (NAIC). The U.S. individual life insurance market distinguishes between two types of policies: Term insurance provides coverage for a predetermined period while permanent insurance provides coverage for the whole life. Permanent insurance is either traditional whole life, universal life, variable life, or variable universal life insurance (ACLI, 2018). The data provide direct written premiums as well as detailed portfolio information (e.g., policies in force, policies lost, policies issued) of 617 life insurers between 1996 and 2018. The reporting date of all the variables is the end of the year.

While the data do not include cancellation rates directly, we are able to calculate cancellation rates for life insurer $j$ in year $s$ by dividing the number of policies that were canceled in $s$ by the number of policies that were in force in $s$:

$$cancel_{j,s}^{US} := \frac{lap_{j,s}^{US} + sur_{j,s}^{US}}{lost_{j,s}^{US} + inForce_{j,s}^{US}} = \frac{lap_{j,s}^{US} + sur_{j,s}^{US}}{inForceProxy_{j,s}^{US}}. \tag{3.1}$$

In this definition $lap^{US}$ denotes the contracts that were canceled as result of nonpayment of premium and $sur^{US}$ designates contracts that were canceled and a cash surrender value was paid to the policyholder. To capture all policies in force during the year ($inForceProxy^{US}$) we add to the policies in force at the end of the year ($inForce^{US}$) all policies that were lost during the year

($lost^{US}$).

We select our data to avoid reporting inconsistencies and to restrict the analysis to U.S. life insurance companies with a focus on primary insurance business. For each step of the data selection process, Table 2 outlines the number of company-year observations and the number of companies kept with each data cleaning step.[6] The full sample includes 12,639 company-year observations of 617 life insurers between 1996 and 2018. In the first selection step we keep companies that do exhibit at least one positive direct written premium between 1996 and 2018. In step two and three, we keep companies with the share of policies assumed stemming from reinsurance or coinsurance activities never exceeding 20% and keep companies that never exhibit a share of revived policies greater than 2%. These variables used in the sample selection process are given by the following ratios:

$$shareAssumed_{j,s}^{US} := assumed_{j,s}^{US} / inForceProxy_{j,s}^{US}, \tag{3.2}$$

$$shareRevived_{j,s}^{US} := revived_{j,s}^{US} / inForceProxy_{j,s}^{US}. \tag{3.3}$$

In the full sample, 97% of company-year observations exhibit a share of assumed policies smaller than 20%, and 97% of company-year observations exhibit a share of revived policies smaller than 2%. The first two steps of the data selection process, thus, restrict the analysis to active insurers with a focus on primary insurance business rather than a significant reinsurance activity. Step three addresses the revival of lapsed policies. Since lapsed policies can be revived as long as five years after the lapse has happened, we are unable to identify whether a high number of lapses is the consequence of a high number of revivals in the subsequent year. The remaining sample after these three data selection steps represents 67% of aggregate direct written premiums and 35% of aggregate polices of the full sample.

Table 3 reports summary statistics of our company-level panel data in the selected sample. The reporting date of all the variables shown here is the end of the year. Direct written premiums are reported in thousand U.S. dollars, while variables on the portfolio are reported in number of policies. The descriptive summary of cancellation rates is in line with reported cancellation rates by SOA and LIMRA (2012) and ACLI (2018). We display the histogram of the cancellation rates in the selected sample in Figure 1. While the median of cancellation rates is equal to 4.79%, the histogram also shows cancellation rates above 30%. Since we are interested in a reliable estimate of the 99.5%-quantile of the cancellation rates, employing the dynamic POT method seems particularly suitable given the few observations in the tail. To take the dependency structure of our panel

[6] Each permutation of the sample selection steps listed in Table 2 leads to the same selected sample.

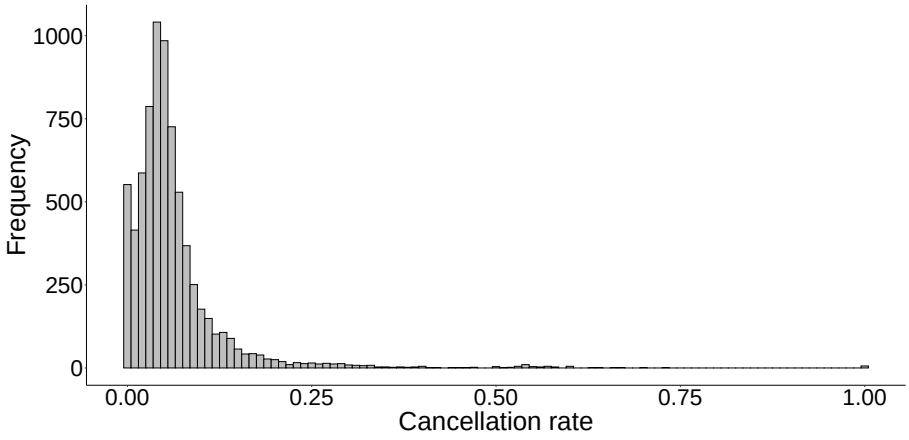
Table 2: U.S. data: Sample selection process

| Step | Description | Observations remaining | Companies remaining |
|---|---|---|---|
| 0 | All company-year observations between 1996 and 2018 | 12,639 | 617 |
| 1 | Keep companies exhibiting at least one positive direct written premium between 1996 and 2018 | 12,010 | 574 |
| 2 | Keep companies with $shareAssumed^{US}$ never exceeding 20% | 8,945 | 423 |
| 3 | Keep companies with $shareRevived^{US}$ never exceeding 2% | 7,336 | 348 |
| **Selected sample** | | **7,336** | **348** |

*Notes*: The table describes the steps of the sample selection process and outlines the number of company-year observations and the number of companies kept with each data cleaning step.

data into account, we will use the number of policies in force ($x = inForceProxy^{US}$) and the year ($s = 1996, 1997, \ldots, 2018$) as covariates in the estimation. We additionally include the following control variables: (1) We monitor new business activity by the share of issued policies in terms of existing policies ($shareIssued^{US}$). (2) We track the focus of a life insurance company on term or permanent life insurance policies by the share of term life policies in force between 1996 and 2018 ($shareTerm^{US}$). (3) We control for changes in market interest rates by the first-difference of U.S. government bond yields with a duration of 10 years ($interest^{US}$).

Figure 1: U.S. data: Histogram of cancellation rates




*Notes:* The figure displays the distribution of the 7,336 cancellation rates between 1996 and 2018.

Table 3: U.S. data: Summary statistics

| Variable | Description | N | Mean | Pctl(25) | Median | Pctl(75) |
|---|---|---|---|---|---|---|
| $dwp^{US}$ ($000) | direct written premium | 7,336 | 223,581 | 844 | 14,492 | 96,903 |
| $lap^{US}$ | policies lapsed | 7,336 | 15,174 | 51 | 1,075 | 6,159 |
| $sur^{US}$ | policies surrendered | 7,336 | 5,742 | 29 | 647 | 3,390 |
| $lost^{US}$ | policies lost | 7,336 | 29,446 | 387 | 3,592 | 16,189 |
| $issued^{US}$ | policies issued | 7,336 | 24,671 | 15 | 1,543 | 12,594 |
| $assumed^{US}$ | policies assumed | 7,336 | 2,952 | 0 | 0 | 0 |
| $revived^{US}$ | policies revived | 7,336 | 692 | 0 | 16 | 281 |
| $inForce_T^{US}$ | term life in force | 7,336 | 87,778 | 42 | 3,129 | 32,616 |
| $inForce_P^{US}$ | permanent life in force | 7,336 | 192,437 | 1,888 | 28,051 | 106,109 |
| $inForce^{US}$ | total in force | 7,336 | 290,120 | 3,962 | 42,199 | 169,331 |
| $shareTerm^{US}$ | share of term policies | 7,336 | 25.45% | 0.74% | 10.86% | 39.89% |
| $shareIssued^{US}$ | share of issued policies | 7,336 | 9.05% | 0.29% | 4.95% | 11.01% |
| $interest^{US}$ | interest rate changes | 23 | 0.15pp | -0.35pp | 0.11pp | 0.43pp |
| $cancel^{US}$ | cancellation rates | 7,336 | 6.27% | 2.89% | 4.79% | 7.23% |

*Notes*: The table reports summary statistics of the variables used in the U.S. analysis of extreme cancellation rates. The data include 348 life insurers between 1996 and 2018.

### 3.1.3 Model selection

The model specification consists of two steps: (1) the threshold selection and (2) the specification of the models for the frequency and severity of extreme cancellation rates. The specification is done via backward induction. For each threshold equal to the deciles of all cancellation rate observations, we perform the model selection for the severity of extreme cancellation rates individually. Based on the chosen model for each threshold, we then select the lowest threshold with a good model fit according to a Q-Q plot. We describe the frequency and severity model selection process for the ex-post selected threshold of the 50%-quantile of all cancellation rates, corresponding to $u = 4.79\%$ and 3,668 excesses. We test whether the classic or dynamic POT method is adequate in our analysis based on likelihood ratio tests (LRT) and Akaike's information criterion (AIC) value.[7] For the selection of the correct model specification within the dynamic POT method, we again base our selection on LRTs and AIC values. Whenever we introduce a nonparametric dependence of the dynamic POT parameters on our covariates, we follow Chavez-Demoulin et al. (2016, p.753)

[7] A likelihood ratio test can reject the model A in favor of model B if model A is nested in model B. The test is calculated as $LR = -2[\ell(A) - \ell(B)]$. $LR$ is chi-square distributed with the degrees of freedom being equal to the additional number of parameters in B when compared to A. Additionally, we compare models based on their AIC value, where among several models the one with lowest AIC value is preferred (see for example Hambuckers et al. (2018b) for a discussion on AIC values in generalized additive models).

and determine the number of knots of the utilized natural cubic spline according to the AIC.

We begin with the specification of the frequency parameter $\rho$. As shown in Table 4 we consecutively expand the model and compare the specifications based on a LRT. Our baseline model consists of a constant term and the control variables of Section 3.1.2. We denote this model by $\rho = Z\vec{\vartheta}_\rho$ with a matrix $Z$ consisting of ones and the controls. First, we test the null hypothesis of this baseline model (see Model (1)) against the alternative of a parametric inclusion of the covariate number of policies $\rho = Z\vec{\vartheta}_\rho + \beta_1 x$ (see Model (2)) via a LRT. At a 1% significance level we cannot reject the null hypothesis, meaning that the two likelihoods of Model (1) and Model (2) do not differ by more than sampling error. In the next step, we find that a nonparametric inclusion of the number of policies (see Model (3)) does not lead to a marked improvement compared to the parametric inclusion of the number of policies.[8] Subsequently, we include time in a parametric way (see Model (4)) and test it against Model (1). The LRT indicates a significant difference in the likelihoods of Model (1) and Model (4). Finally, we vary the degrees of freedom $df$ of the natural cubic spline $h_\rho^{(df)}(s)$, $df \in 1, 2, \ldots, 10$, to obtain the model with the lowest AIC value. Following Chavez-Demoulin et al. (2016, p.753) we plot the AIC value over the degrees of freedom ($df \in 1, 2, \ldots, 10$), which supports $df = 5$ since an additional increase in the degrees of freedom does not lead to a smaller AIC.[9] We then test Model (4) against this nonparametric inclusion of time, which is however not supported by the LRT test at a 1% significance level.

Repeating this procedure for the severity parameters $\xi$ (see the models (6)–(10)) and $\nu$ (see the models (11)–(15)), leads to the following model specification:

$$\rho(x, s) = Z\vec{\vartheta}_\rho + \beta_2 s, \tag{3.4}$$

$$\xi(x, s) = Z\vec{\vartheta}_\xi + \gamma_1 x, \tag{3.5}$$

$$\nu(x, s) = Z\vec{\vartheta}_\nu + \delta_1 x + \delta_2 s. \tag{3.6}$$

Given the model specification of the ex-post selected threshold equal to the 50% of all cancellation rates, we now turn to the threshold selection (see Table 5). We start the analysis at the 0%-quantile and proceed in steps of 10 percentage points. For each threshold, we specify the model according to the steps described in Table 4. Subsequently, we estimate the model and assess the model's goodness of fit in a Q-Q plot, in which we plot the model's residuals against the exponential distribution. The chosen threshold is the smallest threshold so that we observe a good fit in the Q-Q plot (see Panel (a) of Figure 3). This analysis leads us to select the 50%-quantile as our threshold.

[8] A df-AIC plot (similar to Figure B1 in the Online Appendix B) facilitates $df = 1$ for the function $g_\rho^{(df)}(x)$, which corresponds to the linear model specification of Model (2).

[9] Figure B1 in the Online Appendix B displays the corresponding df-AIC plot.

Table 4: U.S. data: Model selection

| $\mathbf{u = 50\%}$**-quantile,** $\mathbf{N_u = 3,668}$ **excesses** | | | LRT | | AIC | Selection |
|---|---|---|---|---|---|---|
| *Frequency parameter* $\rho$ | | | | | | |
| (1) | $\rho = Z\vec{\vartheta}_\rho$ | | | | 10,203 | |
| (2) | $\rho = Z\vec{\vartheta}_\rho + \beta_1 x$ | | (1) : (2) | x $(p = 0.03)$ | 10,201 | |
| (3) | $\rho = Z\vec{\vartheta}_\rho + g_\rho^{(1)}(x)$ | | | – | – | |
| (4) | $\rho = Z\vec{\vartheta}_\rho + \beta_2 s$ | | (1) : (4) | ✓ $(p < 0.00)$ | 10,100 | ✓ |
| (5) | $\rho = Z\vec{\vartheta}_\rho + h_\rho^{(5)}(s)$ | | (4) : (5) | x $(p = 0.01)$ | 10,095 | |
| *Severity parameter* $\xi$ | | | | | | |
| (6) | $\xi = Z\vec{\vartheta}_\xi$ | $\nu = Z\vec{\vartheta}_\nu$ | | | 18,008 | |
| (7) | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu$ | (6) : (7) | ✓ $(p < 0.00)$ | 9,493 | ✓ |
| (8) | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(1)}(x)$ | $\nu = Z\vec{\vartheta}_\nu$ | | – | – | |
| (9) | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu$ | (7) : (9) | x $(p > 0.99)$ | 9,747 | |
| (10) | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x + h_\xi^{(1)}(s)$ | $\nu = Z\vec{\vartheta}_\nu$ | | – | – | |
| *Severity parameter* $\nu$ | | | | | | |
| (11) | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu$ | | | 22,710 | |
| (12) | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu + \delta_1 x$ | (11) : (12) | ✓ $(p < 0.00)$ | 22,686 | |
| (13) | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu + g_\nu^{(1)}(x)$ | | – | – | |
| (14) | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu + \delta_1 x + \delta_2 s$ | (12) : (14) | ✓ $(p < 0.00)$ | 22,505 | ✓ |
| (15) | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu + h_\nu^{(1)}(s)$ | | – | – | |

*Notes*: The table displays the model selection for the frequency parameter $\rho$ and the severity parameters $\xi$ and $\nu$. For each parameter we consecutively expand the model by including the company covariate and the time covariate first parametrically and then nonparametrically. In the fourth column we report the result of a likelihood ratio test (LRT) at a 1% significance level. We denote a significant difference in the likelihood of two models by the symbol "✓' and an insignificant by the symbol "x". We also provide the corresponding p-value in parentheses. Whenever a df-AIC plot supports that the number of knots of the utilized natural cubic spline is equal to one (corresponding to linearity), we do not perform a LRT and indicate this by the symbol "–". The fifth column provides Akaike's information criterion (AIC) value for each model to indicate the model complexity. In the final column we mark the selected model by the symbol "✓".

Table 5: U.S. data: Threshold selection

| **Threshold** $\boldsymbol{u}$ | $\boldsymbol{N_u}$ | **Model for** $\boldsymbol{\xi}$ | **Model for** $\boldsymbol{\nu}$ | **Q-Q plot** |
|---|---|---|---|---|
| 0%-quantile | 7,336 | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + h_\xi^{(3)}(s)$ | $\nu = Z\vec{\vartheta}_\nu + h_\nu^{(6)}(s)$ | xxx |
| 10%-quantile | 6,602 | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu + \delta_2 s$ | xx |
| 20%-quantile | 5,869 | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu + \delta_2 s$ | xx |
| 30%-quantile | 5,135 | $\xi = Z\vec{\vartheta}_\xi + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu + \delta_2 s$ | x |
| 40%-quantile | 4,402 | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu + \delta_1 x + \delta_2 s$ | x |
| 50%-quantile | 3,668 | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu + \delta_1 x + \delta_2 s$ | ✓ |

*Notes*: The table describes the threshold selection procedure. For different selected thresholds $u$, the table provides the number of observations $N_u$ above the chosen threshold and reports the chosen model specification for the severity parameters $\xi$ and $\nu$. The last column indicates the observed goodness of fit in a Q-Q plot based on the symbol "x" for bad quality and "✓" for good quality of the fit.

### 3.1.4 Results

We estimate the model as specified in equations (3.4)–(3.6). Table 6 provides the coefficients obtained from estimating the generalized additive models.[10] The estimated coefficient for the time

Table 6: U.S. data: Generalized additive model output

| | *Dependent variable:* | | |
|---|---|---|---|
| | $\rho$ | $\xi$ | $\nu$ |
| **Parametric coefficients** | coeff. (std. err.) | coeff. (std. err.) | coeff. (std. err.) |
| Policy covariate (in m) | | $-0.068^{***}$ | 0.020 |
| | | (0.008) | (0.043) |
| Time | $-0.010^{***}$ | | $-0.009$ |
| | (0.001) | | (0.011) |
| Share of new business | $0.561^{***}$ | $-1.019^{***}$ | $4.055^{***}$ |
| | (0.041) | (0.057) | (0.554) |
| Share of term life | $0.232^{***}$ | $0.284^{***}$ | $0.445^{*}$ |
| | (0.020) | (0.041) | (0.239) |
| Interest rate changes | 0.938 | 0.024 | $-2.761$ |
| | (1.104) | (2.282) | (13.517) |
| Intercept | $20.286^{***}$ | $0.201^{***}$ | 13.742 |
| | (1.757) | (0.018) | (21.781) |
| Observations | 7,336 | 3,668 | 3,668 |
| Log Likelihood | $-5{,}050$ | $-5{,}186$ | $-11{,}253$ |

*Notes*: The table displays the output of the generalized additive model estimation for the parameters $\rho$, $\xi$, and $\nu$ according to the equations (3.4)–(3.6). The following significance codes are used: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level.

variable in the model for $\rho$ demonstrates that the likelihood of excesses decreases over time. This is also shown in Panel (a) of Figure 2, which provides boxplots of the parameter $\rho$ for each year. Table 6 and Panel (b) of Figure 2 furthermore indicate a negative effect of the portfolio size on the parameter $\xi$. In Panel (b) of Figure 2 we also display boxplots for the estimated parameters $\xi$ for each decile of the policy covariate rather than showing the estimated parameters $\xi$ for each value of the policy covariate. In most cases, the shape parameter is positive and implies a support between zero and positive infinity whereas cancellation rates are bounded between zero and one. Even though our model implies an unbounded support of the GPD, the probability of excesses

[10] Figure B2 in the Online Appendix B additionally displays the parameters in dependence of the covariates time and number of policies, fixing all other variables equal to their median values.

above one is small.[11] Finally, Table 6 reports the estimated coefficient for the parameter $\nu$ and Panel (c) and Panel (d) of Figure 2 display the corresponding boxplots of the estimates for $\beta$ in the corresponding year and in the decile of the policy covariate respectively.

For each boxplot of Figure 2, we also provide pointwise two-sided 95% confidence intervals for the displayed median. While we calculate asymptotic confidence intervals based on the standard errors for the parameter $\rho$, we employ a bootstrap procedure for obtaining confidence intervals for $\xi$ and $\beta$. We compute bootstrap confidence intervals by slightly adjusting the approach presented in Chavez-Demoulin et al. (2016), which is based on the post-blackend bootstrap of Chavez-Demoulin and Davison (2005). First, we resample the model's residuals with replacement in groups that exhibit the same decile of portfolio size, decile of share of new business, decile of share of term life business, interest rate change, and same year. Second, we calculate new excesses based on the original estimates for $\xi$ and $\beta$. Third, we apply the original model with fixed covariates and control variables to obtain new estimates for the severity parameters $\xi$ and $\beta$. We repeat this procedure 1,000 times and obtain pointwise two-sided 95% confidence intervals for each parameter estimate by calculating the 2.5% and 97.5% empirical quantiles of the corresponding 1,000 estimated values.

Given the estimated parameters, we assess the model's goodness of fit and calculate 99.5%-quantiles of the cancellation rate observations and corresponding pointwise two-sided 95% confidence intervals (see Figure 3). Panel (a) of Figure 3 displays a Q-Q plot, in which we plot the selected model's residuals against the exponential distribution. The good fit observed in this plot supports our model selection. Furthermore, Panel (b) provides a boxplot of the estimated 99.5%-quantiles for each year between 1996 and 2018. We observe a slight negative trend in the median of 99.5%-quantiles over time, indicating that the risk of high cancellation rates decreased over time (consistent with the coefficients of the time parameter in Table 6 and the boxplot in Panel (a) of Figure 2). The boxplots in each year furthermore show substantial heterogeneity in 99.5%-quantiles. This suggests that life insurers are exposed heterogeneously to high cancellation rates.

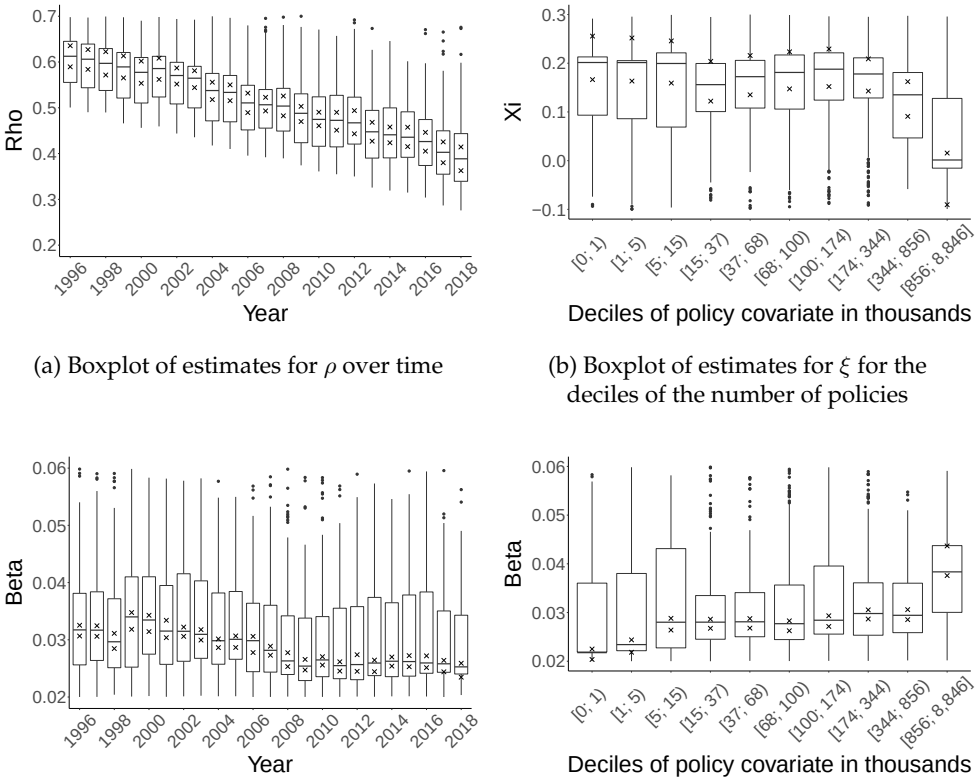
The heterogeneity is further affirmed by examining the 99.5%-quantiles in dependence of the number of policies. We provide boxplots of these estimates in Panel (c) for the deciles of the number of policies. In line with Table 6 and Figure 2, Panel (c) indicates that companies with fewer policies in force exhibit higher 99.5%-quantiles. In contrast to small life insurance companies, which exhibit 99.5%-quantiles as high as 40% (see the upper quartile in the boxplot), the companies in the highest decile of the policy covariate exhibit 99.5%-quantiles of around 20%. High

[11] Based on the estimates for the shape parameter and the scale parameter, we calculated the probability that cancellation rates exceed one in our model. Given all these excess probabilities, we found that the empirical 95%-quantile of the excess probabilities is smaller than 0.2%.

cancellation rates are additionally dependent on the share of new business in dependence of existing business and the product portfolio (see Panel (d) and Panel (e)). We find that a high share of new business activity results in higher cancellation rates, reaching levels of approximately 60% in the highest decile of this variable. In Panel (e), we additionally observe that a portfolio with a high share of term life business is more exposed to high cancellation rates compared to a portfolio consisting mainly of permanent insurance. While life insurance companies with predominantly permanent life insurance policies exhibit 99.5%-quantiles of 20%, we find that life insurers, which mainly focus on term life policies have 99.5%-quantiles up to 50%.

Previous literature models extreme cancellation rates as a result of changes in interest rates (Loisel and Milhaud, 2011; Barsotti et al., 2016). Panel (f) displays boxplots of the 99.5%-quantiles dependent on the interest rate changes in U.S. government bonds with a duration of 10 years. At least in our observed range of interest rate changes we do not identify any pattern between changes in interest rates and high cancellation rates. In addition to Panel (f), Table 6 shows that the coefficients of the interest rate variable are insignificant.

Figure 2: U.S. data: Parameter estimates




(a) Boxplot of estimates for $\rho$ over time

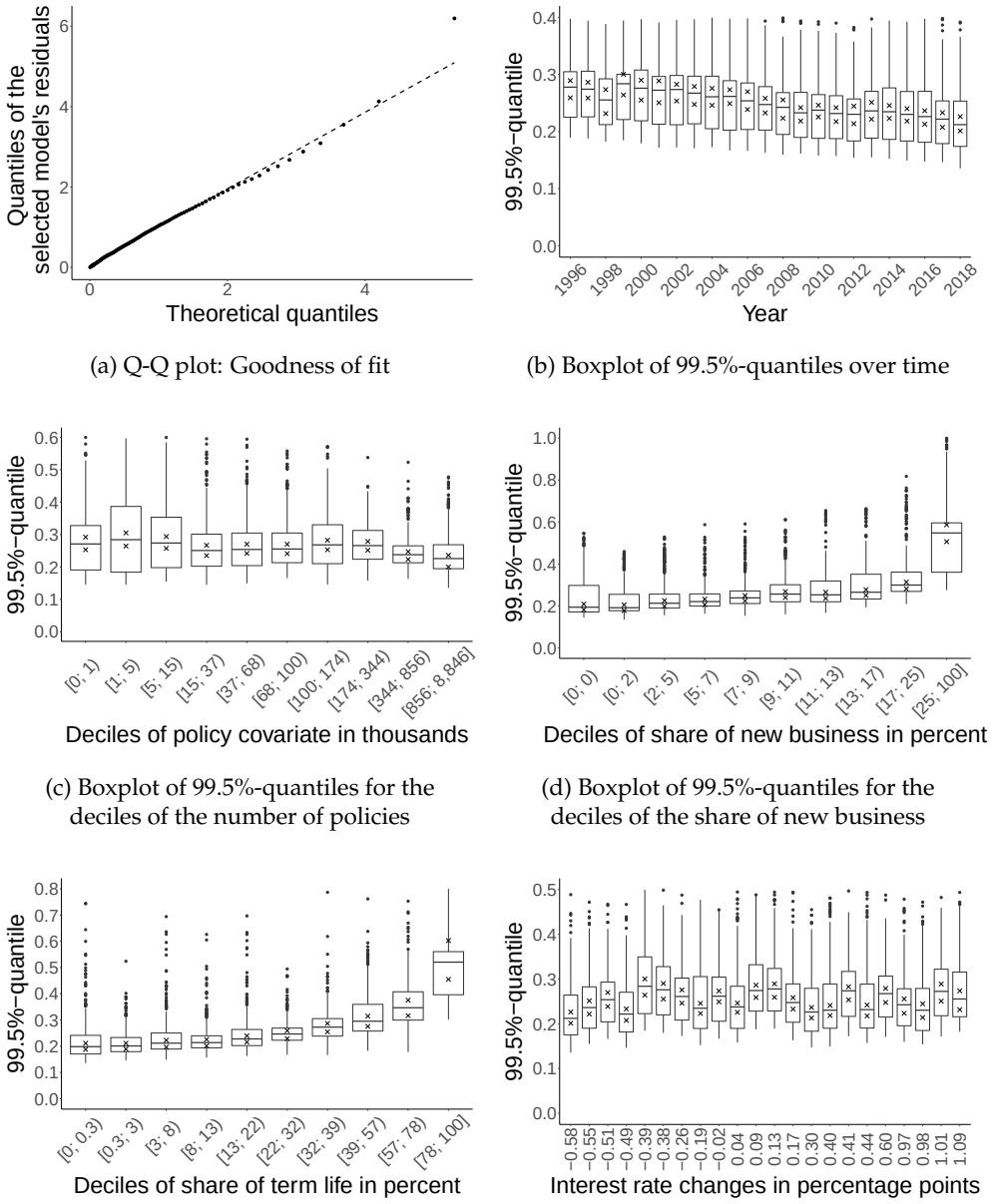
(b) Boxplot of estimates for $\xi$ for the deciles of the number of policies

(c) Boxplot of estimates for $\beta$ over time

(d) Boxplot of estimates for $\beta$ for the deciles of the number of policies

*Notes:* The figure displays boxplots of the parameter estimates for $\rho$, $\xi$, and $\beta$ in dependence of time and the deciles of the policy covariate. The symbol "x" additionally denotes pointwise two-sided 95% confidence intervals for the median of the parameter estimates shown in the corresponding boxplot.

Figure 3: U.S. data: Estimation results




(a) Q-Q plot: Goodness of fit

(b) Boxplot of 99.5%-quantiles over time

(c) Boxplot of 99.5%-quantiles for the deciles of the number of policies

(d) Boxplot of 99.5%-quantiles for the deciles of the share of new business

(e) Boxplot of 99.5%-quantiles for the deciles of the share of term life

(f) Boxplot of 99.5%-quantiles for the interest rate changes

*Notes:* The figure displays a goodness of fit plot in Panel (a) and estimation results in Panel (b)–(f). The goodness of fit plot is a Q-Q plot based on the selected model's residuals. Panel (b) provides for each year between 1996 and 2018 a boxplot of estimated 99.5%-quantiles. Panel (c) shows the boxplot of estimated 99.5%-quantiles for the deciles of the policy covariate, Panel (d) for the deciles of the share of new business, Panel (e) for the deciles of the share of term life, and Panel (f) for the interest rate changes. The estimation results additionally include pointwise two-sided 95% confidence intervals for the median of the 99.5%-quantiles indicated by the symbol "x" in the corresponding boxplot.

## 3.2 German life insurance market

### 3.2.1 Motivation

Cancellation risk is the second most important risk factor for European life insurers and is therefore explicitly addressed in European insurance regulation (EIOPA, 2011). Within its standard formula, Solvency II requires life insurers to calculate additional capital demand for adverse cancellation events such as a permanent increase or decrease of future cancellation rates. In addition, life insurers have to apply a one-time shock mass lapse scenario with cancellation rates equal to 40%. Interestingly, this level of 40% has not been predicted by utilizing data but is only based on expert judgment. In contrast to the expert judgment based approach in Solvency II's standard model, our approach can use data to calibrate an empirically justified mass cancellation scenario. For this purpose, we again use available panel data at the company level.

### 3.2.2 Company-level, panel data

The German Federal Financial Supervisory Authority (BaFin) collects information on all German life insurers subject to the Solvency II regulation. German life insurers predominantly offer the products endowment, term life, occupational disability, annuity, unit linked insurance, and group life insurance.[12] While we are unable to distinguish between these product categories, Eling and Kiesenbauer (2013) use policyholder-level data and find that in contrast to other markets, the German life insurance market exhibits only small differences in cancellation rates across product categories. Our data provide gross earned premiums, detailed portfolio information (e.g., amount of insurance in force, amount of insurance issued), and cancellation rates of 122 life insurers between 1996 and 2018. In contrast to the U.S. analysis, we do not apply any filters because the BaFin database only includes active primary insurers and policies can only be revived as long as one month. Table 7 reports summary statistics of our company-level panel data and Figure 4 provides a histogram of the 2,344 cancellations rates. While the median of cancellation rates in both the German and the U.S. data is around 4.7%, U.S. cancellation rates exhibit a higher variability. The interquartile range of cancellation rates is 4.34% for the U.S. and 2.83% for Germany. Even though German cancellation rates are less variable compared to the U.S. market, Figure 4 also shows

[12] An endowment policy either pays the sum insured in the case of death or if the insured person is still at life at the end of the contract period. A term life policy pays the sum insured if the insured person dies within the contract period. An occupational disability policy pays an annuity until retirement if the insured person is no longer able to work. An annuity policy pays an annuity until the insured person's death starting at a predefined age. Special forms of annuities are Riester and Rürup policies (see Börsch-Supan and Wilke (2004) for more details). Finally, a unit linked policy pays the sum insured if the insured person is still alive at the end of the contract period. The accrued interest is dependent on the performance of stocks or funds.

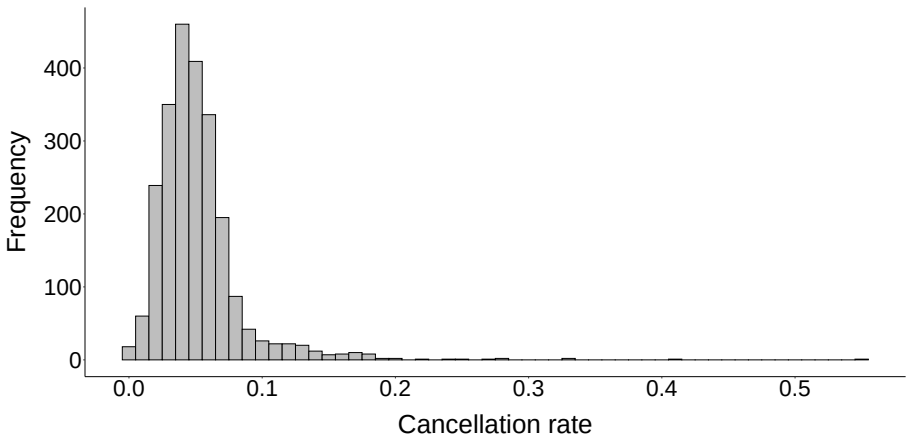
cancellation rates above 20% in the German market. Since we are again interested in a reliable estimate of the 99.5%-quantile of the cancellation rates, employing the dynamic POT method seems particularly suitable given the few observations in the tail. Due to data availability, we will use the gross earned premium instead of the number of policies as covariate in the German estimation. We additionally include time as a covariate and the following control variables: (1) We monitor new business activity by the share of issued policies in terms of existing policies ($shareIssued^{DE}$). (2) We control for changes in market interest rates by the first-difference of German government bond yields with a duration of 10 years ($interest^{DE}$).

Table 7: German data: Summary statistics

| Variable | Description | N | Mean | Pctl(25) | Median | Pctl(75) |
|---|---|---|---|---|---|---|
| $gep^{DE}$ (EUR) | gross earned premium | 2,344 | 720 m | 63 m | 216 m | 779 m |
| $inFore^{DE}$ (EUR) | amount of insurance in force | 2,344 | 23 bn | 2 bn | 9 bn | 27 bn |
| $issued^{DE}$ (EUR) | amount of insurance issued | 2,344 | 2 bn | 92 m | 525 m | 2 bn |
| $shareIssued^{DE}$ | share of issued policies | 2,344 | 11.34% | 5.02% | 8.59% | 13.60% |
| $interest^{DE}$ | interest rate changes | 23 | -0.25pp | -0.54pp | -0.33pp | 0.07pp |
| $cancel^{DE}$ | cancellation rates | 2,344 | 5.15% | 3.35% | 4.67% | 6.18% |

*Notes*: The table reports summary statistics of the variables used in the German analysis of extreme cancellation rates. The data include 122 life insurers between 1996 and 2018.

Figure 4: German data: Histogram of cancellation rates




*Notes:* The figure displays the distribution of the 2,344 cancellation rates between 1996 and 2018.

### 3.2.3 Model selection

In the German analysis, we select a threshold equal to the 65%-quantile of all cancellation rates, corresponding to $u = 5.48\%$ and 821 excesses. Again, the specification of the threshold is done via backward induction. For each threshold equal to the deciles of all cancellation rates, we perform the model selection for the severity of extreme cancellation rates. We then choose the lowest threshold with a good model fit according to a Q-Q plot. As in the U.S. analysis, we specify the model for the frequency parameter $\rho$ and severity parameters $\xi$ and $\nu$ based on likelihood ratio tests and Akaike's information criterion value (see Table 8).

Table 8: German data: Model selection

| $\mathbf{u = 65\%}$**-quantile,** $\mathbf{N_u = 821}$ **excesses** | | | | LRT | | AIC | Selection |
|---|---|---|---|---|---|---|---|
| *Frequency parameter* $\rho$ | | | | | | | |
| (1) | $\rho = Z\vec{\vartheta}_\rho$ | | | | | 3,139 | |
| (2) | $\rho = Z\vec{\vartheta}_\rho + \beta_1 x$ | | (1) : (2) | ✓ | $(p < 0.01)$ | 3,131 | |
| (3) | $\rho = Z\vec{\vartheta}_\rho + g_\rho^{(1)}(x)$ | | | – | | – | |
| (4) | $\rho = Z\vec{\vartheta}_\rho + \beta_1 x + \beta_2 s$ | | (2) : (4) | ✓ | $(p < 0.00)$ | 2,969 | |
| (5) | $\rho = Z\vec{\vartheta}_\rho + \beta_1 x + h_\rho^{(2)}(s)$ | | (4) : (5) | ✓ | $(p < 0.00)$ | 2,959 | ✓ |
| *Severity parameter* $\xi$ | | | | | | | |
| (6) | $\xi = Z\vec{\vartheta}_\xi$ | $\nu = Z\vec{\vartheta}_\nu$ | | | | 2,201 | |
| (7) | $\xi = Z\vec{\vartheta}_\xi + \gamma_1 x$ | $\nu = Z\vec{\vartheta}_\nu$ | (6) : (7) | x | $(p > 0.99)$ | 2,875 | |
| (8) | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x)$ | $\nu = Z\vec{\vartheta}_\nu$ | (6) : (8) | ✓ | $(p < 0.00)$ | 2,136 | |
| (9) | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu$ | (8) : (9) | ✓ | $(p < 0.00)$ | 2,077 | ✓ |
| (10) | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + h_\xi^{(1)}(s)$ | $\nu = Z\vec{\vartheta}_\nu$ | | – | | – | |
| *Severity parameter* $\nu$ | | | | | | | |
| (11) | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu$ | | | | 7,563 | |
| (12) | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu + \delta_1 x$ | (11) : (12) | ✓ | $(p < 0.00)$ | 7,342 | ✓ |
| (13) | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu + g_\nu^{(1)}(x)$ | | – | | – | |
| (14) | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu + \delta_1 x + \delta_2 s$ | (12) : (14) | x | $(p > 0.99)$ | 7,427 | |
| (15) | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu + \delta_1 x + h_\nu^{(1)}(s)$ | | – | | – | |

*Notes*: The table displays the model selection for the frequency parameter $\rho$ and the severity parameters $\xi$ and $\nu$. For each parameter we consecutively expand the model by including the company covariate and the time covariate first parametrically and then nonparametrically. In the fourth column we report the result of a likelihood ratio test (LRT) at a 1% significance level. We denote a significant difference in the likelihood of two models by the symbol "✓' and an insignificant by the symbol "x". We also provide the corresponding p-value in parentheses. Whenever a df-AIC plot supports that the number of knots of the utilized natural cubic spline is equal to one (corresponding to linearity), we do not perform a LRT and indicate this by the symbol "–". The fifth column provides Akaike's information criterion (AIC) value for each model to indicate the model complexity. In the final column we mark the selected model by the symbol "✓".

We begin with the specification of the frequency parameter $\rho$ and our baseline is again a model

$\rho = Z\vec{\vartheta}_\rho$ with a matrix $Z$ consisting of ones and the control variables of Section 3.2.2. We test the null hypothesis of this baseline model (see Model (1)) against the alternative of a parametric inclusion of the covariate gross earned premium $\rho = Z\vec{\vartheta}_\rho + \beta_1 x$ (see Model (2)) via a LRT. At a 1% significance level we can reject the null hypothesis of a constant model. While we cannot reject the null hypothesis of Model (2) when comparing it to a nonparametric inclusion of the gross earned premium (see Model (3)), a parametric inclusion of time is supported (see Model (4)). Finally, a df-AIC plot and a LRT test facilitate to specify the parameter $\rho$ as a natural cubic spline function of time with two degrees of freedom (see Model (5)). We repeat this procedure for the severity parameters $\xi$ (see the models (6)–(10)) and $\nu$ (see the models (11)–(15)) and finally obtain the following model specification:[13]

$$\rho(x, s) = Z\vec{\vartheta}_\rho + \beta_1 x + h_\rho^{(2)}(s), \tag{3.7}$$

$$\xi(x, s) = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + \gamma_2 s, \tag{3.8}$$

$$\nu(x, s) = Z\vec{\vartheta}_\nu + \delta_1 x. \tag{3.9}$$

We now return to the threshold selection (see Table 9). The chosen threshold is the smallest threshold so that we observe a good fit in the Q-Q plot (see Panel (a) of Figure 6). This analysis supports to select the 65%-quantile as our threshold.[14]

Table 9: German data: Threshold selection

| **Threshold $u$** | **$N_u$** | **Model for $\xi$** | **Model for $\nu$** | **Q-Q plot** |
|---|---|---|---|---|
| 10%-quantile | 2,110 | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(4)}(x)$ | $\nu = Z\vec{\vartheta}_\nu$ | xxx |
| 20%-quantile | 1,881 | $\xi = Z\vec{\vartheta}_\xi$ | $\nu = Z\vec{\vartheta}_\nu + h_\nu^{(2)}(s)$ | xxx |
| 30%-quantile | 1,641 | $\xi = Z\vec{\vartheta}_\xi$ | $\nu = Z\vec{\vartheta}_\nu + \delta_1 x + \delta_2 s$ | xx |
| 40%-quantile | 1,407 | $\xi = Z\vec{\vartheta}_\xi + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu + g_\nu^{(3)}(x) + \delta_2 s$ | xx |
| 50%-quantile | 1,172 | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(3)}(x) + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu + h_\nu^{(2)}(s)$ | x |
| 55%-quantile | 1,055 | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(3)}(x)$ | $\nu = Z\vec{\vartheta}_\nu + h_\nu^{(2)}(s)$ | x |
| 60%-quantile | 938 | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x)$ | $\nu = Z\vec{\vartheta}_\nu + \delta_2 s$ | x |
| 65%-quantile | 821 | $\xi = Z\vec{\vartheta}_\xi + g_\xi^{(2)}(x) + \gamma_2 s$ | $\nu = Z\vec{\vartheta}_\nu + \delta_1 x$ | ✓ |

*Notes*: The table describes the threshold selection procedure. For different selected thresholds $u$, the table provides the number of observations $N_u$ above the chosen threshold and reports the chosen model specification for the severity parameters $\xi$ and $\nu$. The last column indicates the observed goodness of fit in a Q-Q plot based on the symbol "x" for bad quality and "✓" for good quality of the fit.

[13]The dependence of the parameter $\xi$ on time is unusual since we observe in most applications that the shape parameter does not depend on time (see e.g., Chavez-Demoulin et al., 2016). However, in our empirical setting, this does not lead to a switch between finite/infinite first moment of the GPD over time.

[14]We start the analysis at the 0%-quantile and proceed in steps of 10 percentage points. After the 50%-quantile we start moving in 5 percentage points increments in order to still have a reasonably large sample size.

### 3.2.4 Results

We estimate the model as specified in equations (3.7)–(3.9). Table 10 provides the parametric coefficients and the degrees of freedom of the smooth terms obtained from estimating the generalized additive models.[15]

Table 10: German data: Generalized additive model output

| | *Dependent variable:* | | |
|---|---|---|---|
| | $\rho$ | $\xi$ | $\nu$ |
| **Parametric coefficients** | coeff. (std. err.) | coeff. (std. err.) | coeff. (std. err.) |
| Policy covariate (in bn) | $-0.010^{*}$ | | $-0.397^{***}$ |
| | (0.006) | | (0.077) |
| Time | | $0.026^{***}$ | |
| | | (0.003) | |
| Share of new business | $0.290^{***}$ | $-0.208$ | $2.463^{*}$ |
| | (0.078) | (0.155) | (1.442) |
| Interest rate changes | $-1.282$ | 1.420 | $-12.083$ |
| | (1.629) | (2.472) | (23.818) |
| Intercept | $0.322^{***}$ | $-51.932^{***}$ | $-4.022^{***}$ |
| | (0.015) | (6.878) | (0.311) |
| **Smooth terms** | df | df | df |
| Policy covariate (in bn) | | $2^{***}$ | |
| Time | $2^{***}$ | | |
| Observations | 2,344 | 821 | 821 |
| Log Likelihood | $-1{,}479$ | $-860$ | $-2{,}670$ |

*Notes*: The table displays the output of the generalized additive model estimation for the parameters $\rho$, $\xi$, and $\nu$ according to the equations (3.7)–(3.9). The following significance codes are used: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level.

The estimated coefficient for the policy covariate in the model for $\rho$ demonstrates that the likelihood of excesses decreases with increasing portfolio size. This is also shown in Panel (b) of Figure 5, which provides boxplots of the parameter $\rho$ for each decile of the gross earned premium. Panel (a) of Figure 5 additionally displays a negative relationship between the likelihood of excesses over time. While the parameter $\rho$ decreases over time, Table 10 and Panel (c) of Figure 5 indicate a positive relationship between $\xi$ and time. Consistent with the U.S. analysis, Table 10 and Panel

[15] Figure B3 in the Online Appendix B additionally displays the parameters in dependence of the covariates time and gross earned premium, fixing all other variables equal to their median values.

(d) of Figure 5 furthermore show a negative effect of the portfolio size on the parameter $\xi$. Again, even though positive values of $\xi$ imply an unbounded support of the GPD, the probability of excesses above one is small.[16] Finally, Table 10 reports the estimated coefficient for the parameter $\nu$ and Panel (e) and Panel (f) of Figure 5 display the corresponding boxplots of the estimates of $\beta$ in the corresponding year and in the decile of the policy covariate respectively.

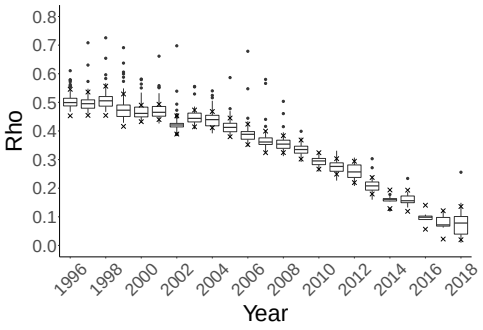
Given the estimated parameters, we again assess the model's goodness of fit and calculate 99.5%-quantiles of the cancellation rate observations and corresponding pointwise two-sided 95% confidence intervals (see Figure 6). Panel (a) of Figure 6 shows the model's goodness of fit in a Q-Q plot and supports our model selection. Panel (b) reports boxplots for the estimated 99.5%-quantiles for each year between 1996 and 2018. The median of 99.5%-quantiles never exceeds 22% (the corresponding upper bound of the confidence interval never exceeds 25%) and exhibits a decreasing trend over this time period. The higher level of extreme cancellation rates at the beginning of the observation period is likely a result of the dot-com bubble. During this period, some German life insurers faced solvency issues and thus a decline in trust in their company (see e.g., the company "Mannheimer Leben", Protektor, 2004). In comparison to the U.S. analysis we see, however, less heterogeneity in 99.5%-quantiles for each year.

Furthermore, Panel (c) of Figure 6 displays boxplots of the 99.5%-quantiles dependent on the deciles of the gross earned premium. Panel (c) indicates that the level of extreme cancellation rates is decreasing in the gross earned premium (in line with Table 10 and Figure 5). While the companies in the lowest decile of the policy covariate exhibit 99.5%-quantiles of about 20–25%, this risk lies at around 10% for life insurers in the largest decile of the policy covariate. As in the U.S. analysis, Panel (d) demonstrates that a high share of new business activity results in higher cancellation rates, which are estimated to get as high as 35%. Additionally, in Panel (e) we observe, in line with the U.S. analysis, no pattern between changes in interest rates and high cancellation rates. This is also consistent with the insignificant coefficients of the interest rate variable in Table 10.

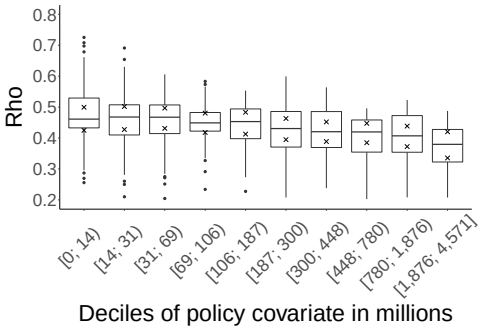
Our dynamic POT estimation furthermore enables us to calculate the confidence level that corresponds to a cancellation rate of 40%. In each year between 1996 and 2018, the median of the implied confidence levels lies above 99.9%. We thus conclude that Solvency II's mass lapse assumption corresponds to the 99.9%-quantile rather than the 99.5%-quantile, on which the solvency capital requirement in European insurance regulation is usually based on.

[16] Based on the estimates for the shape parameter and the scale parameter, we calculated the probability that cancellation rates exceed one in our model. Given all these excess probabilities, we found that the empirical 95%-quantile of the excess probabilities is smaller than 0.2%.

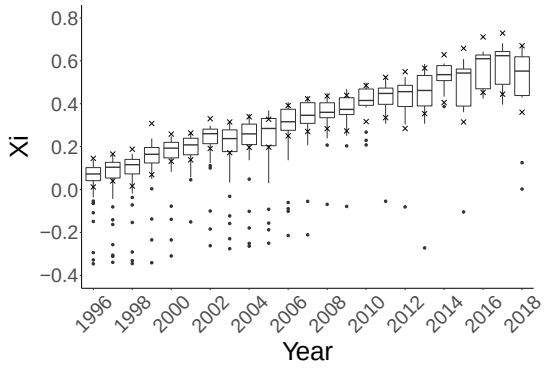
Figure 5: German data: Parameter estimates



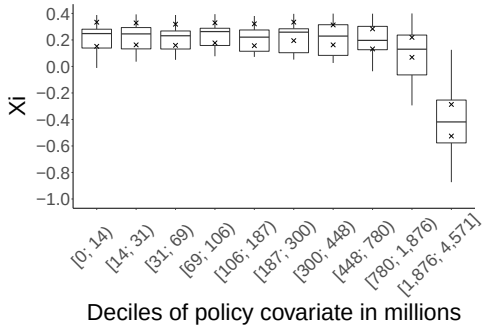

(a) Boxplot of estimates for $\rho$ over time



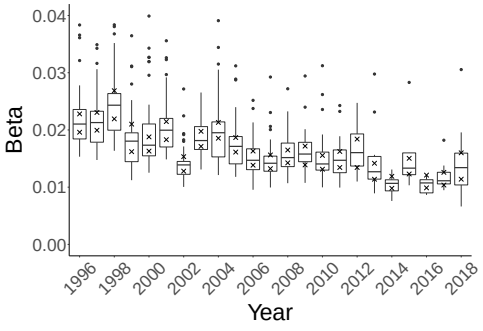

(b) Boxplot of estimates for $\rho$ for the deciles of the gross earned premium



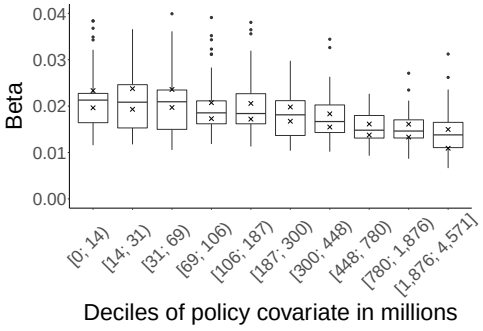

(c) Boxplot of estimates for $\xi$ over time




(d) Boxplot of estimates for $\xi$ for the deciles of the gross earned premium



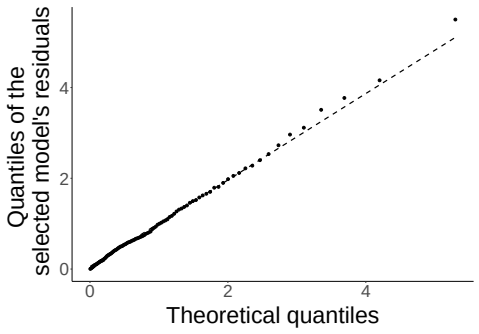

(e) Boxplot of estimates for $\beta$ over time



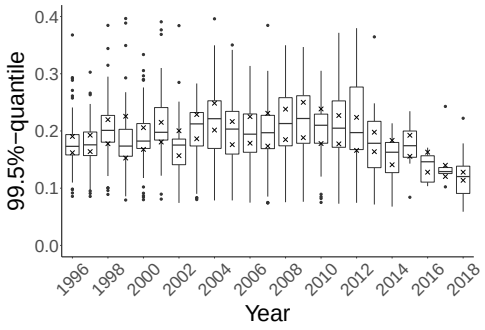

(f) Boxplot of estimates for $\beta$ for the deciles of the gross earned premium

*Notes:* The figure displays boxplots of the parameter estimates for $\rho$, $\xi$, and $\beta$ in dependence of time and the deciles of the policy covariate. The symbol "x" additionally denotes pointwise two-sided 95% confidence intervals for the median of the parameter estimates shown in the corresponding boxplot.

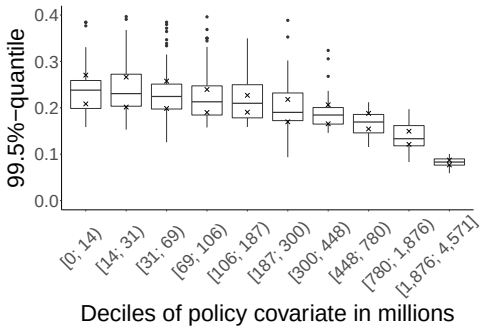
Figure 6: German data: Estimation results



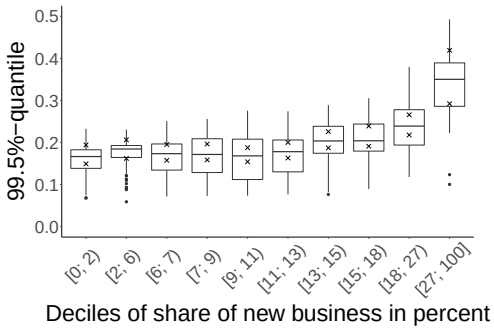

(a) Q-Q plot: Goodness of fit



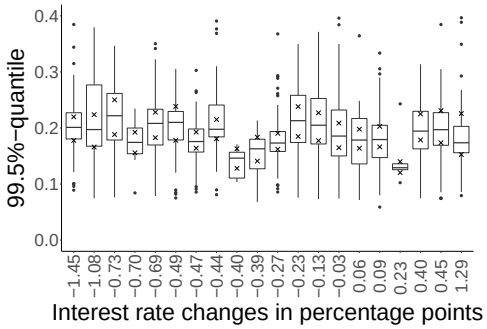

(b) Boxplot of 99.5%-quantiles over time




(c) Boxplot of 99.5%-quantiles for the deciles of the gross earned premium




(d) Boxplot of 99.5%-quantiles for the deciles of the share of new business




(e) Boxplot of 99.5%-quantiles for the interest rate changes

*Notes:* The figure displays a goodness of fit plot in Panel (a) and estimation results in Panel (b)–(e). The goodness of fit plot is a Q-Q plot based on the selected model's residuals. Panel (b) provides for each year between 1996 and 2018 a boxplot of estimated 99.5%-quantiles. Panel (c) shows the boxplot of estimated 99.5%-quantiles for the deciles of the policy covariate, Panel (d) for the deciles of the share of new business, and Panel (e) for the interest rate changes. The estimation results additionally include pointwise two-sided 95% confidence intervals for the median of the 99.5%-quantiles indicated by the symbol "x" in the corresponding boxplot.

# 4 Discussion

## 4.1 Implications for the modeling of extreme cancellation rates

Our results allow several conclusions. Although the U.S. and the German life insurance market are different in terms of their offered products, we find that the effects of the covariates (portfolio size, time) and control variables (share of new business, interest rate changes) are consistent across both markets. This demonstrates the validity of the employed estimation method. In both the U.S. and the German life insurance market, we find that the risk of extreme cancellation rates is decreasing in the size of the company's portfolio. We additionally observe a positive relationship between new business activity and extreme cancellation rates. As policies in early policy years are subject to higher cancellation rates, a high share of new business raises the likelihood and severity of an extreme cancellation event.

We also discuss the effect of changes in interest rates on extreme cancellation rates. Previous literature models extreme cancellation rates as a result of increasing interest rates (Loisel and Milhaud, 2011; Barsotti et al., 2016). Our analysis shows no pattern between changes in interest rates and high cancellation rates. However, interest rates are mainly decreasing in our time period, which does not allow us to test the effect of increases in interest rates on extreme cancellation rates. The U.S. analysis furthermore indicates the importance of the product portfolio on cancellation rates, at least in certain markets. While life insurance companies with predominantly permanent life insurance policies exhibit 99.5%-quantiles of 20%, we see that life insurers, which mainly focus on term life policies, have 99.5%-quantiles up to 50%. Renshaw and Haberman (1986), Cerchiara et al. (2009) and Milhaud et al. (2011) find that this also holds at normal cancellation rate levels using data from Scotland, Italy, and Spain respectively. In the German data we are, unfortunately, unable to distinguish between product categories. However, Eling and Kiesenbauer (2013) show that in contrast to other markets, the German life insurance market exhibits only small differences in cancellation rates across product categories. Still, a thorough analysis of the German market by product type is a promising area for future research.

Interestingly, our above results identify dependencies of a mass cancellation event on company characteristics, which so far have not been taken into account in Solvency II. The adverse cancellation rate scenarios used to calculate the solvency capital requirement do not distinguish between small or large insurance companies, so-called run-off companies without new business or companies that exhibit a high share of new business, nor are these scenarios dependent on the product type. We additionally find that the current scenario of 40% for a mass cancellation event in

Solvency II's standard model has, at least for the German life insurance market, no empirical foundation. According to our analysis, a cancellation rate of 40% corresponds to the 99.9%-quantile. The 99.5%-quantile, on which the solvency capital requirement in European insurance regulation is usually based on, lies at approximately 20–25% (see the median of 99.5%-quantiles and the corresponding upper bound of the confidence interval in Panel (b) of Figure 6).

### 4.2 Implications for the solvency capital requirement calculation

The solvency capital requirement for European life insurers is sensitive to changes in cancellation rates and is especially sensitive to a mass cancellation scenario (EIOPA, 2011). The South African insurer Old Mutual (2016) reports that the solvency capital requirement regarding its business in Europe for a mass cancellation event is equal to 500 million pounds. The Austrian life insurer UNIQUA (2017) amounts the capital demand for this risk to be 262 million Euros. European insurance regulation defines the required capital for an adverse scenario such as a mass cancellations event as the reduction in expected future profits due to deviations from the company's best-estimate assumptions. In the following, we illustrate the sensitivity of a term life policy's expected profitability to a mass cancellation event.[17]

The expected profitability, called technical provision (TP), at time $\tau$ is defined as the difference between the expected present value of future cash outflows at time $\tau$ and the expected present value of future cash inflows at time $\tau$. To calculate the technical provision under Solvency II, we discount expected future cash flows by discount factors, which can be calculated from the spot rates provided by EIOPA (2019). For example, a negative technical provision at contraction ($\tau = 0$) means that the contract generates overall expected future profits for the insurer. In our illustration we consider a 40-year-old insured, who purchases a term life policy with a contract period of 20 years and a sum insured of 150,000 Euros (death benefit).[18] The insurer expects cash inflows in terms of a level premium at time $\tau = 0, 1, \ldots, 19$ and is obliged to pay the specified death benefit at time $\tau = 1, 2, \ldots, 20$ if the insured dies. Based on the above assumptions and using mortality rates and a guaranteed interest rate of $0.9\%$ according to the German Actuarial Association, the policy's premium is equal to 495.37 Euros. During the contract period, the policyholder can exercise the option to cancel the policy, which terminates future premium payments as well as the death benefit. In the calculation of expected future profits, an insurer assumes best-estimate can-

[17] The Online Appendix A provides the corresponding actuarial calculations in detail.

[18] In Germany, people purchase a home on average at the age of 40 and usually purchase a term life policy to hedge the risk of an early death, which leaves the family members with the debt incurred by building loans. Moreover, the average period for building loans is 20 years and the average sum insured for term policies in 2018 was 150,000 Euros (SOEP, 2019; BaFin, 2019).

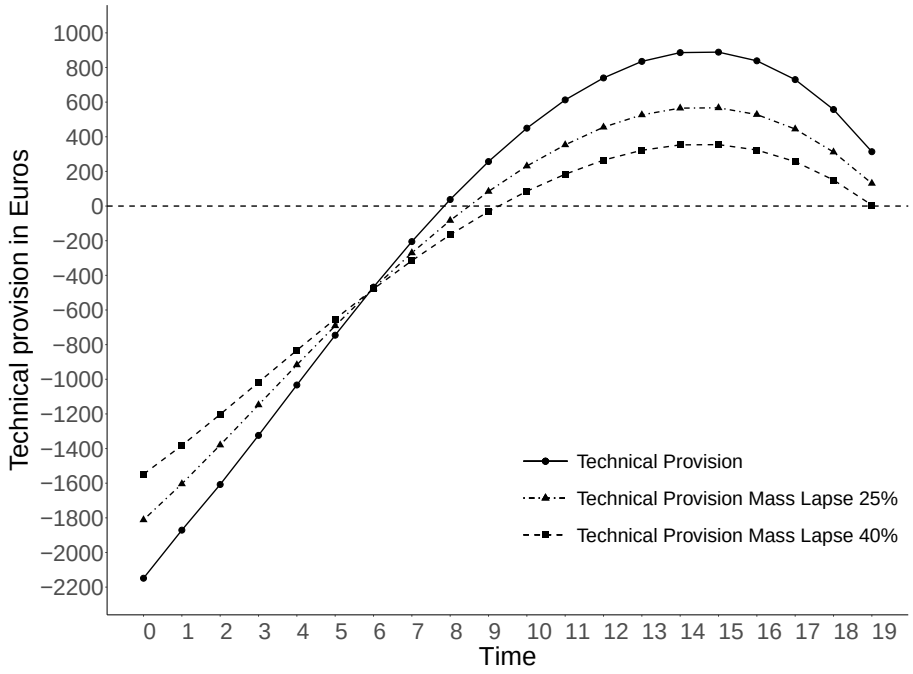
cellation rates for each year within the contract period. We use best-estimate cancellation rates, which are typical for a term life policy with a contract period of 20 years in the German market (Milbrodt and Helbig, 2008). Cancellation rates are high in the first policy years and are decreasing afterwards.[19] Figure 7 shows the expected profitability of the policy under best-estimate cancellation assumptions as well as two mass cancellation scenarios: (1) We consider Solvency II's 40% mass lapse shock and (2) calculate the expected profitability with a 25% mass lapse shock as it is supported by our analysis in Section 3.2.4.

According to European insurance regulation we calculate the expected profitability at time $\tau$ under a mass cancellation event by taking the best-estimate cancellation rates but replacing the cancellation rate at time $\tau + 1$ by the mass lapse shock (40% according to Solvency II and 25% according to our calculation). The figure displays that for the first eight contract years the technical provision (see the solid line) is negative, meaning that the policy generates an expected future profit for the insurer. Afterwards, the technical provision is positive, indicating that the policy generates an expected future loss. It is thus most profitable for the insurance company if the insured does not terminate her term life policy during the first contract years but cancels it as soon as the technical provision becomes positive in later contract years. Following this logic, it is reasonable that a mass lapse scenario (see the dashed lines) leads to a higher technical provision during the first six and to a lower technical provision after this period when compared to the best-estimate scenario.

Given the technical provision over time, we can do a back-of-the-envelope calculation of the effect of a mass lapse event on the capital demand for a life insurance portfolio. First, we build a portfolio of term life policies by assuming a constant new business volume of 6,500 policies over the next twenty years. This figure corresponds to the average number of issued term life policies in Germany in 2018 (BaFin, 2019). At time $\tau = 0$ we have 6,500 policies. Considering the above mortality and cancellation rates we expect that out of these 6,500 policies, 6,119 policies are left in the portfolio at time $\tau = 1$, then 5,690 policies at time $\tau = 2$, and finally 3,204 policies at time $\tau = 19$. Second, we calculate the expected profitability of the term life policy under best-estimate cancellation rates and under a mass lapse scenario for each point in time. For example, the expected profitability at time $\tau = 0$ is 2,149 Euros ($TP(0) = -2,149$ in Figure 7). Assuming a mass cancellation scenario with a shock level of 40% reduces the profitability of this policy to 1,549 Euros and with a shock level of 25% to 1,812 Euros. Performing this calculation for each point in time $\tau = 0, 1, \ldots, 19$ provides the capital demand for the above portfolio consisting of term life policies. The capital demand is equal to 12,310,895 Euros with a mass lapse shock level

[19] Column 2 in Table A1 in the Online Appendix A displays these best-estimate cancellation rates.

Figure 7: Sensitivity of the technical provision to Solvency II's mass lapse scenario




*Notes:* The figure displays the expected profitability of a term life policy under best-estimate cancellation assumptions and two mass cancellation scenarios: (1) We consider Solvency II's 40% mass lapse shock and (2) calculate the expected profitability with a 25% mass lapse shock as it is supported by our analysis in Section 3.2.4.

of 40% and 6,920,217 Euros with a mass lapse shock level of 25%. Therefore, reducing the shock level for the mass lapse scenario from 40% to 25% can reduce the capital demand for lapse risk by five million Euros for this portfolio.[20] This illustration and the figures reported by Old Mutual (2016) and UNIQUA (2017) demonstrate that the assumption of a 40% mass lapse shock in the cancellation risk model results in a high capital demand. We thus conclude that the overstatement of the mass lapse event in Solvency II leads to significant overreserving compared to empirically more justifiable scenarios.

[20]Solvency II defines the capital demand for a portfolio as maximum capital demand required under the following three scenarios: (1) Mass lapse shock of 40% (2) 50% increase in best-estimate cancellation rates and (3) 50% decrease in best-estimate cancellation rates. In our example, the 25% mass lapse shock still requires more capital than a 50% increase in best-estimate cancellation rates and approximately the same capital as a 50% decrease in best-estimate cancellation rates. All details of the calculation are provided in the Online Appendix A.

# 5 Conclusion

We contribute to the literature by assessing the risk of a mass cancellation scenario in life insurance. As extreme cancellation rates are rare events, data sourced from only one insurer do not provide enough observations to assess this tail risk. We thus use the dynamic peaks over threshold method developed by Chavez-Demoulin et al. (2016) to take quantitative covariates into account. We apply this approach to U.S. data and show that, depending on product type, cancellation rates up to 50% are a good assumption for a mass cancellation scenario in this market. Furthermore, we give implications for European insurance regulation. We discuss the appropriateness of Solvency II's mass lapse scenario by using German data. Cancellation rates in the range of 20–25% are reflecting the risk of a mass cancellation scenario in the German market. This calls into doubt whether the current scenario of a 40% cancellation rate in Solvency II's standard model is empirically justified. In both the U.S. and Germany, the mass cancellation scenario is dependent on the portfolio size, which provides some validity to the use of (partial) internal models when assessing reserves for mass cancellation scenarios. Lastly, the marked differences in the estimated mass cancellation scenarios between the U.S. companies that predominantly sell term life insurance and those that predominantly sell permanent life insurance indicate that the product type has an effect on the appropriate mass cancellation scenario, at least in certain markets. Since national life insurance markets in Europe greatly differ with regard to their dominant products (Standard and Poors, 2018), it can thus be expected that they differ with respect to their appropriate mass cancellation scenario, as well.

Our work leaves some directions for further research. Even though Eling and Kiesenbauer (2013) report only minor differences in the cancellation rates of different product types in the German market, future research should nevertheless employ the dynamic POT method on a richer data set to analyze the mass cancellation scenario by product type in the German market. Given the differences by product type in the U.S. analysis, different mass cancellation scenarios for different product types could be adequate in the German market, as well. Since such data are not publicly available in a panel structure of the entire market, cooperation with a regulating entity such as BaFin or EIOPA is likely required. Additionally, the lack of available public data for time periods with increasing interest rates allows us to only draw conclusions about economies with falling or stagnant interest rates. Further research for other time periods could give an estimate for the influence of interest rate development on mass cancellation risk. Lastly, the approach of modeling the mass cancellation scenario equally for all European life insurance markets is called into question in our results. An empirical study using data from several different national insurance markets in Europe, would be able to test it directly.

# References


ACLI (2018). Life insurance fact book. American Council of Life Insurers, https://www.acli.com/posting/rp18-007.

Albizzati, M.-O. and H. Geman (1994). Interest rate risk management and valuation of the surrender option in life insurance policies. Journal of Risk and Insurance 61(4), 616–637.

Bacinello, A. (2003). Fair valuation of a guaranteed life insurance participating contract embedding a surrender option. Journal of Risk and Insurance 70(3), 461–487.

BaFin (2019). Statistik der Erstversicherungsunternehmen – Lebensversicherung. Bundesanstalt für Finanzdienstleistungsaufsicht, https://www.bafin.de/DE/PublikationenDaten/Statistiken/Erstversicherung/erstversicherung_artikel.html.

Balkema, A. and L. de Haan (1974). Residual life time at great age. Annals of Probability 2(5), 792–804.

Barsotti, F., X. Milhaud, and Y. Salhi (2016). Lapse risk in life insurance: Correlation and contagion effects among policyholders' behaviors. Insurance: Mathematics and Economics 71, 317–331.

Börsch-Supan, A. and C. Wilke (2004). The German public pension system: How it was, how it will be. NBER Working Paper No. 10525.

CEIOPS (2009). CEIOPS' advice for level 2 implementing measures on Solvency II: Standard formula SCR – Article 109 c life underwriting risk. Committee of European Insurance and Occupational Pensions Supervisors, https://register.eiopa.europa.eu/CEIOPS-Archive/Documents/Advices/CEIOPS-L2-Final-Advice-on-Standard-Formula-Life-underwriting-risk.pdf.

Cerchiara, R., M. Edwards, and A. Gambini (2009). Generalized linear models in life insurance: Decrements and risk factor analysis under Solvency II. Giornale dell'Istituto degli Attuari 72, 100–122.

Chavez-Demoulin, V. and A. C. Davison (2005). Generalized additive modelling of sample extremes. Journal of the Royal Statistical Society: Series C (Applied Statistics) 54(1), 207–222.

Chavez-Demoulin, V., P. Embrechts, and M. Hofert (2016). An extreme value approach for modeling operational risk losses depending on covariates. Journal of Risk and Insurance 83(3), 735–776.

Coles, S. (2001). An introduction to statistical modeling of extreme values. London: Springer.

Consiglio, A. and D. De Giovanni (2010). Pricing the option to surrender in incomplete markets. Journal of Risk and Insurance 77(4), 935–957.

Dar, A. and C. Dodds (1989). Interest rates, the emergency fund hypothesis and saving through endowment policies: Some empirical evidence for the U.K. Journal of Risk and Insurance 56(3), 415–433.

DAV (2018). Herleitung der Sterbetafel DAV2008T für Lebensversicherungen mit Todesfallcharakter. Deutsche Aktuarvereinigung, https://aktuar.de/unsere-themen/lebensversicherung/sterbetafeln/2018-10-05_DAV-Richtlinie_Herleitung_DAV2008T.pdf.

Davison, A. C. and R. Huser (2015). Statistics of extremes. Annual Review of Statistics and its Application 2, 203–235.

EIOPA (2011). EIOPA report on the fifth quantiative impact study (QIS5) for Solvency II. European Insurance and Occupational Pensions Authority, https://register.eiopa.europa.eu/Publications/Reports/QIS5_Report_Final.pdf.

EIOPA (2019). Risk-free interest rate term structures. European Insurance and Occupational Pensions Authority, https://www.eiopa.europa.eu/tools-and-data/risk-free-interest-rate-term-structures-0_en.

Eling, M. and D. Kiesenbauer (2013). What policy features determine life insurance lapse? An analysis of the German market. Journal of Risk and Insurance 81(2), 241–269.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). Modelling extremal events. Berlin et al.: Springer.

Embrechts, P., K. Mizgier, and X. Chen (2018). Modeling operational risk depending on covariates: An empirical investigation. Journal of Operational Risk 13(3), 17–46.

Gottlieb, D. and K. Smetters (2019). Lapse-based insurance. Working Paper, https://cpb-us-w2.wpmucdn.com/sites.wustl.edu/dist/c/547/files/2019/12/Lapse-Based-Insurance.pdf.

Hambuckers, J., A. Groll, and T. Kneib (2018a). Understanding the economic determinants of the severity of operational losses: A regularized generalized pareto regression approach. Journal of Applied Econometrics 33(6), 898–935.

Hambuckers, J., T. Kneib, R. Langrock, and A. Silbersdorff (2018b). A markov-switching generalized additive model for compound poisson processes, with applications to operational losses models. Quantitative Finance 18(10), 1679–1698.

Hitz, A., R. Davis, and G. Samorodnitsky (2017). Discrete extremes. arXiv preprint 1707.05033.

Insurance Europe (2019). European insurance life industry database. https://www.insuranceeurope.eu/sites/default/files/assets/DatabaseMarch2019_Life.xlsx.

Kiesenbauer, D. (2012). Main determinants of lapse in the German life insurance industry. North American Actuarial Journal 16(1), 52–73.

Knoller, C., G. Kraut, and P. Schoenmaekers (2016). On the propensity to surrender a variable annuity contract: An empirical analysis of dynamic policyholder behavior. Journal of Risk and Insurance 83(4), 979–1006.

Kuo, W., C. Tsai, and W.-K. Chen (2003). An empirical study on the lapse rate: The cointegration approach. Journal of Risk and Insurance 70(3), 489–508.

Loisel, S. and X. Milhaud (2011). From deterministic to stochastic surrender risk models: Impact of correlation crises on economic capital. European Journal of Operational Research 214(2), 348–357.

McNeil, A. J., R. Frey, and P. Embrechts (2015). Quantitative risk management: Concepts, techniques and tools. Princeton: Princeton University Press.

Milbrodt, H. and M. Helbig (2008). Mathematische Methoden der Personenversicherung. Berlin et al.: Walter de Gruyter.

Milhaud, X., S. Loisel, and V. Maume-Deschamps (2011). Surrender triggers in life insuance: What main features affect surrender behavior in a classical economic context? Bulletin Français d'Actuariat 22, 5–48.

Nadarajah, S. and K. Mitov (2002). Asymptotics of maxima of discrete random variables. Extremes 5(3), 287–294.

Old Mutual (2016). Old mutual group's Solvency II and economic capital results. https://www.oldmutual.com/docs/default-source/investor-relations-files/plc/results/2016/solvency-ii-and-economic-capital-results.pdf?sfvrsn=1311c9a2_0.

Outreville, J. (1990). Whole life lapse rates and the emergency fund hypothesis. Insurance: Mathematics and Economics 9, 249–255.

Pesando, J. (1974). The interest sensibility of the flow of fund through life insurance companies: An economic analysis. The Journal of Finance 29(4), 1105–1121.

Pickands, J. (1975). Statistical inference using extreme order statistics. Annals of Probability 3(1), 119–131.

Protektor (2004). Geschäftsbericht 2003. https://www.protektor-ag.de/de/wp-content/uploads/sites/2/2015/07/protektor_geschaeftsbericht_2003.pdf.

Renshaw, A. and S. Haberman (1986). Statistical analysis of life assurance lapses. Journal of the Institute of Actuaries 113(3), 459–497.

Schott, F. (1971). Disintermediation through policy loans at life insuance companies. The Journal of Finance 26(3), 719–729.

SOA and LIMRA (2012). U.S. individual life insurance persistency. Society of Actuaries and Life Insurance Marketing and Research Association, https://www.soa.org/files/research/exp-study/research-2007-2009-us-ind-life-pers-report.pdf.

SOEP (2019). German socio-economic panel, data for years 1984–2017, version 34. doi: 10.5684/soep. v34.

Standard and Poors (2018). European life insurers are playing the long game with product shifts. https://www.allnews.ch/sites/default/files/files/European%20Life%20Insurers_22%20Feb%202018.pdf.

UNIQUA (2017). Group economic capital report. http://www.uniqagroup.com/gruppe/versicherung/media/files/UNIQA_ECR_Report_2017_12042018.pdf.

# Appendix A Suitable for online publication: Actuarial calculations

This section provides the details of the capital demand calculation of Section 4.2. We consider a 40-year-old insured, who purchases a term life policy with a contract period of 20 years and a sum insured of $SI = 150,000$ Euros (death benefit). The insurer expects cash inflows in terms of a level premium at time $\tau = 0, 1, \ldots, 19$ and is obliged to pay the specified death benefit at time $\tau = 1, 2, \ldots, 20$ if the insured dies. Following the German Actuarial Association, we use a guaranteed interest rate of $0.9\%$ and mortality rates $\{q_{40}, q_{41}, \ldots, q_{59}\}$ according to the mortality table DAV 2008 T (DAV, 2018). For example, the mortality rate $q_{40}$ denotes the probability that a person, who is 40 years old, dies before reaching the age 41. The mortality table DAV 2008 T provides both first-order mortality rates $\{q_{40}^{(1)}, q_{41}^{(1)}, \ldots, q_{59}^{(1)}\}$ and second-order mortality rates $\{q_{40}^{(2)}, q_{41}^{(2)}, \ldots, q_{59}^{(2)}\}$. In German life insurance, first-order mortality rates (prudent mortality rates) are used in the premium calculation and second-order mortality rates (best-estimate mortality rates) in the expected profitability calculation (DAV, 2018). Given the above assumptions, the annual level premium is equal to:[21]

$$P = \frac{{}_{|20}A_{40}}{\ddot{a}_{40,\overline{20|}}} \cdot SI = 495.37 \text{ Euro}. \tag{A.1}$$

In equation (A.1), the symbol ${}_{|20}A_{40}$ denotes the present value of a death benefit of 1, payable at the end of the year in which the insured, who is 40 years old at the beginning of the contract, dies (provided that the insured's death takes place within the contract period of 20 years). Additionally, the symbol $\ddot{a}_{40,\overline{20|}}$ designates the present value of an instantaneous life annuity with contract duration 20 years payable to an insured, who is 40 years old at the beginning of the contract.

During the contract period the policyholder can exercise the option to cancel the policy, which terminates future premium payments as well as the death benefit. In the calculation of expected future profits, an insurer assumes best-estimate cancellation rates for each year within the contract period. Milbrodt and Helbig (2008) provide best-estimate cancellation rates $\{cancel_1, cancel_2, \ldots, cancel_{20}\}$ for a German life insurance policy with a contract period of 20 years. Table A1 displays the cancellation rates under best-estimate assumptions as well as in Solvency II's different cancellation rate stress scenarios: (1) Mass lapse shock of 40%, (2) 50% increase in best-estimate cancellation rates, and (3) 50% decrease in best-estimate cancellation rates. Additionally, the table

[21] For simplicity, we do not consider acquisition and administrative costs and do not incorporate surplus distributions (interest gains, mortality gain, expense savings) for the policyholder in our calculations. Thus, the expected profitability of the term life policy for the insurer is higher than in reality. However, this is no drawback for our illustration since the solvency capital requirement calculation is based on changes in profits rather than actual profits. Additionally, the consideration of acquisition or administrative costs would increase the capital demand for insurers because early policy cancellations usually lead to unbalanced cost, which are allocated to the entire contract period.

reports a mass lapse scenario with a level of 25%.

Table A1: Best-estimate cancellation rates and alternative cancellation scenarios

| Policy Year | BE cancellation rates | Mass lapse 40% | Mass lapse 25% | BE cancellation rates (+50%) | BE cancellation rates (-50%) |
|---|---|---|---|---|---|
| 1 | 0.0580 | 0.4000 | 0.2500 | 0.0870 | 0.0290 |
| 2 | 0.0693 | 0.0693 | 0.0693 | 0.1040 | 0.0346 |
| 3 | 0.0567 | 0.0567 | 0.0567 | 0.0851 | 0.0284 |
| 4 | 0.0478 | 0.0478 | 0.0478 | 0.0717 | 0.0239 |
| 5 | 0.0435 | 0.0435 | 0.0435 | 0.0653 | 0.0217 |
| 6 | 0.0419 | 0.0419 | 0.0419 | 0.0629 | 0.0209 |
| 7 | 0.0373 | 0.0373 | 0.0373 | 0.0559 | 0.0186 |
| 8 | 0.0328 | 0.0328 | 0.0328 | 0.0492 | 0.0164 |
| 9 | 0.0296 | 0.0296 | 0.0296 | 0.0444 | 0.0148 |
| 10 | 0.0269 | 0.0269 | 0.0269 | 0.0403 | 0.0134 |
| 11 | 0.0243 | 0.0243 | 0.0243 | 0.0364 | 0.0121 |
| 12 | 0.0208 | 0.0208 | 0.0208 | 0.0312 | 0.0104 |
| 13 | 0.0257 | 0.0257 | 0.0257 | 0.0386 | 0.0128 |
| 14 | 0.0234 | 0.0234 | 0.0234 | 0.0351 | 0.0117 |
| 15 | 0.0233 | 0.0233 | 0.0233 | 0.0350 | 0.0117 |
| 16 | 0.0231 | 0.0231 | 0.0231 | 0.0346 | 0.0115 |
| 17 | 0.0227 | 0.0227 | 0.0227 | 0.0341 | 0.0114 |
| 18 | 0.0223 | 0.0223 | 0.0223 | 0.0335 | 0.0112 |
| 19 | 0.0218 | 0.0218 | 0.0218 | 0.0327 | 0.0109 |
| 20 | 0.0308 | 0.0308 | 0.0308 | 0.0462 | 0.0154 |

*Notes*: The table displays best-estimate (BE) cancellation rates sourced from Milbrodt and Helbig (2008) and Solvency II's three cancellation rate scenarios: (1) Mass lapse shock of 40%, (2) 50% increase in best-estimate cancellation rates, and (3) 50% decrease in best-estimate cancellation rates. The table additionally reports a mass lapse scenario with a level of 25%.

The expected profitability, called technical provision (TP), at time $\tau$ is defined as the difference between the expected present value of future cash outflows at time $\tau$ and the expected present value of future cash inflows at time $\tau$. To calculate the technical provision under Solvency II, we discount the expected future cash flows by discount factors $DF(i)$, which can be calculated from the spot rates provided by EIOPA (2019). We calculate these discount factors, which correspond to the spot rates $spot(0,i)$ prevailing as of December 2019 for the interval $[0,i]$, $i = 0, 1, \ldots, 20$ by:

$$DF(i) = \frac{1}{(1 + spot(0,i))^i}. \tag{A.2}$$

Assuming a constant term structure over time and second-order mortality rates according to DAV 2008 T, the technical provision $TP(\tau)$, $\tau = 0, 1, \ldots, 19$, can be calculated by taking into

account the policyholder's transition probabilities due to cancellation and death:

$$\begin{aligned} TP(\tau) = SI \cdot \sum_{i=1}^{20-\tau} DF(i) \prod_{j=1}^{i-1} \left(1 - q^{(2)}_{40+j+\tau-1}\right) \prod_{k=1}^{i} \left(1 - cancel_{k+\tau}\right) \cdot q^{(2)}_{40+i+\tau-1} \\ - P \cdot \sum_{i=0}^{20-\tau-1} DF(i) \prod_{j=1}^{i} \left(1 - q^{(2)}_{40+j+\tau-1}\right)\left(1 - cancel_{j+\tau}\right), \quad \tau = 0, 1, 2, \ldots, 19. \end{aligned} \tag{A.3}$$

We calculate the expected profitability of the term life policy under best-estimate cancellation rates for each point in time. Afterwards, we replace our best-estimate cancellation rates by Solvency II's cancellation rate scenarios and calculate the changes in the expected profitability compared to the best-estimate scenario. Table A2 reports the corresponding values. For example, Solvency II's mass lapse scenario with a cancellation rate of 40% leads to an increase in the technical provision by 600 Euros in comparison to the best-estimate assumption and thus decreases the expected profitability by 600 Euros.

Based on this example, we can do a back-of-the-envelope calculation of the effect of a mass lapse event on the capital demand for a life insurance portfolio. First, we build a portfolio of term life policies by assuming a constant new business volume of 6,500 policies over the next twenty years, which corresponds to the average number of issued term life policies in Germany in 2018. At time $\tau = 0$ we have 6,500 policies. Considering the above mortality and cancellation rates we expect that out of these 6,500 policies, 6,119 policies are left in the portfolio at time $\tau = 1$, then 5,690 policies at time $\tau = 2$, and finally 3,204 policies at time $\tau = 19$ (see column 2 in Table A2). Second, we determine for each policy which of the stress scenarios lead to a reduction in expected profits (CEIOPS, 2009). The mass lapse shock reduces expected profits for policies between $\tau = 0$ and $\tau = 5$, the permanent increase shock reduces expected profits for policies at $\tau = 0$ and $\tau = 1$, and the permanent decrease shock negatively affects expected profits between $\tau = 2$ and $\tau = 19$. Multiplying the changes (see the bold values in Table A2 with the corresponding number of policies (see column 2 in Table A2), we obtain the capital demand for each shock scenario:

$$SCR_{MassLapse40} = 6500 \cdot 600 + 6119 \cdot 489 + \cdots + 4874 \cdot 93 = 12\,310\,895, \tag{A.4}$$

$$SCR_{LapseUp} = 6500 \cdot 89 + 6119 \cdot 40 = 823\,260, \tag{A.5}$$

$$SCR_{LapseDown} = 5690 \cdot 51 + 5362 \cdot 89 + \cdots + 3204 \cdot 13 = 7\,319\,514. \tag{A.6}$$

The solvency capital requirement for cancellation risk is then defined as maximum of the capital

demand under the three scenarios:

$$SCR_{Lapse} = \max\left(SCR_{MassLapse40}, SCR_{LapseUp}, SCR_{LapseDown}\right) = 12\,310\,895. \tag{A.7}$$

Reducing the mass lapse shock from a level of 40% to a level of 25%, leads to a capital demand for the mass lapse shock equal to:

$$SCR_{MassLapse25} = 6500 \cdot 337 + 6119 \cdot 267 + \cdots + 4874 \cdot 54 = 6\,920\,217. \tag{A.8}$$

This reduces the capital demand for cancellation risk by five million to $SCR_{Lapse} = 7\,319\,514$.

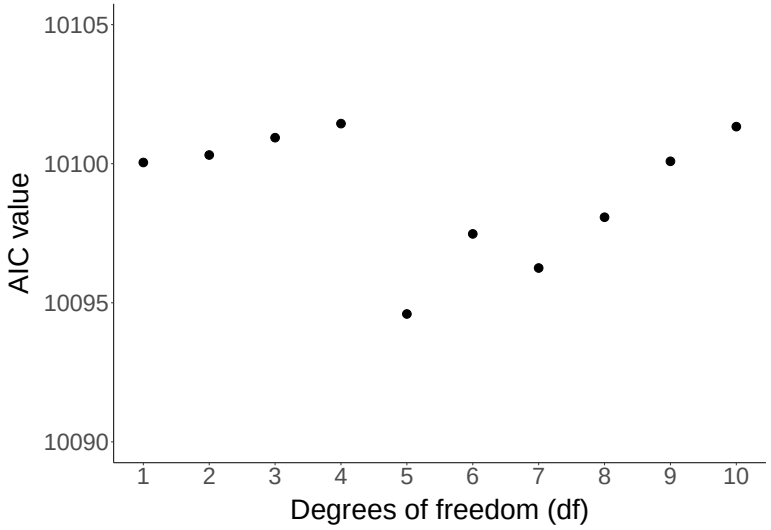
Table A2: Technical provision under different cancellation scenarios

| **Time** | **No. Policies** | **TP** | **ΔTP (mass lapse 40%)** | **ΔTP (mass lapse 25%)** | **ΔTP (rates +50%)** | **ΔTP (rates -50%)** |
|---|---|---|---|---|---|---|
| 0 | 6,500 | -2149 | **600** | **337** | **89** | -56 |
| 1 | 6,119 | -1871 | **489** | **267** | **40** | -4 |
| 2 | 5,690 | -1607 | **405** | **228** | -13 | **51** |
| 3 | 5,362 | -1324 | **306** | **176** | -52 | **89** |
| 4 | 5,101 | -1033 | **200** | **116** | -80 | **114** |
| 5 | 4,874 | -746 | **93** | **54** | -100 | **131** |
| 6 | 4,663 | -468 | -10 | -6 | -115 | **141** |
| 7 | 4,483 | -205 | -110 | -65 | -123 | **145** |
| 8 | 4,329 | 38 | -204 | -121 | -125 | **144** |
| 9 | 4,193 | 257 | -288 | -172 | -124 | **139** |
| 10 | 4,072 | 450 | -364 | -219 | -119 | **132** |
| 11 | 3,964 | 613 | -429 | -259 | -112 | **123** |
| 12 | 3,872 | 740 | -474 | -284 | -105 | **113** |
| 13 | 3,763 | 835 | -513 | -309 | -92 | **98** |
| 14 | 3,664 | 886 | -533 | -321 | -80 | **84** |
| 15 | 3,567 | 889 | -534 | -321 | -66 | **69** |
| 16 | 3,472 | 839 | -515 | -310 | -51 | **53** |
| 17 | 3,380 | 730 | -473 | -285 | -37 | **38** |
| 18 | 3,291 | 558 | -407 | -246 | -24 | **24** |
| 19 | 3,204 | 314 | -308 | -183 | -13 | **13** |

*Notes*: The table displays the development of the technical provision (TP) of a term life policy over time under the different cancellation scenarios given in Table A1.

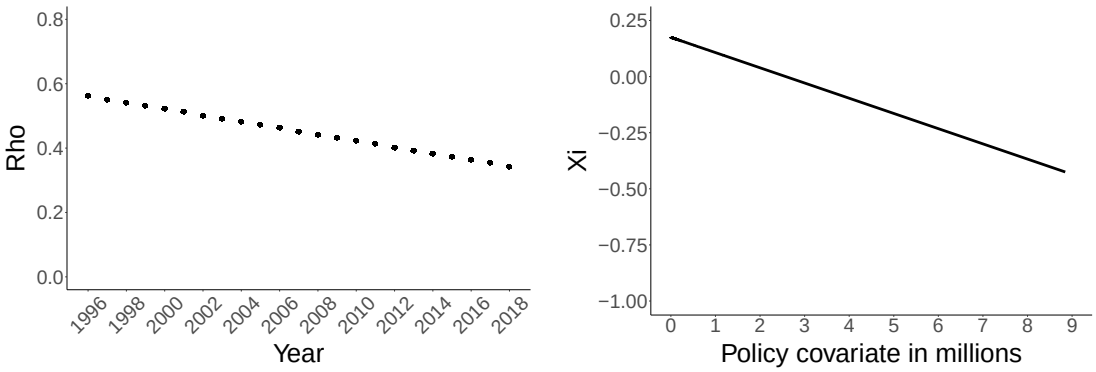
# Appendix B Suitable for online publication: Additional graphical representations

Figure B1: U.S. data: Degrees of freedom selection in the model specification of the frequency parameter




*Notes:* The figure displays for varying degrees of freedom $df$ of the natural cubic spline $h_\rho^{(df)}(s)$, $df \in 1, 2, \ldots, 10$ the AIC value of the model $\rho = Z\vec{\vartheta}_\rho + h_\rho^{(df)}(s)$. The graph supports $df = 5$ since an additional increase in the degrees of freedom does not lead to a smaller AIC.

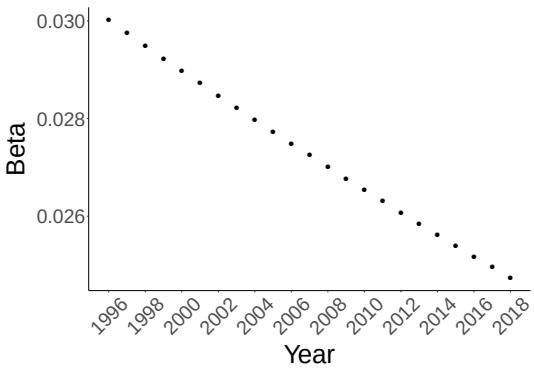
Figure B2: U.S. data: Parameters as functions of the covariates



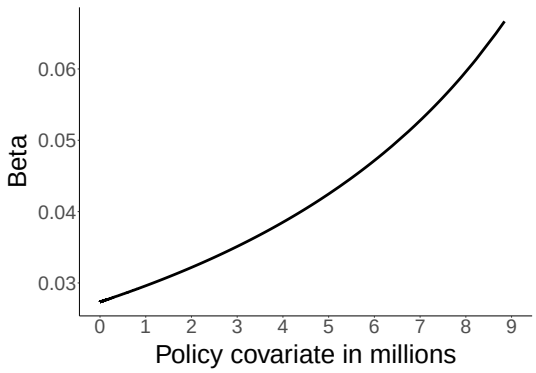

(a) $\rho$ as a function of time




(b) $\xi$ as a function of the number of policies



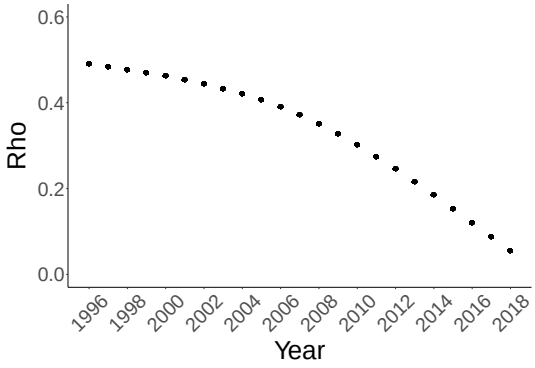

(c) $\beta$ as a function of time



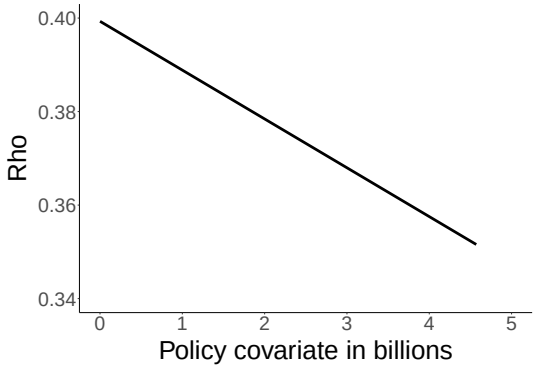

(d) $\beta$ as a function of the number of policies

*Notes:* The figure displays fitted values for the parameters $\rho$, $\xi$, and $\beta$ according to the model given in Table 6 for varying covariate values (number of policies or time), fixing all other variables equal to their median values.

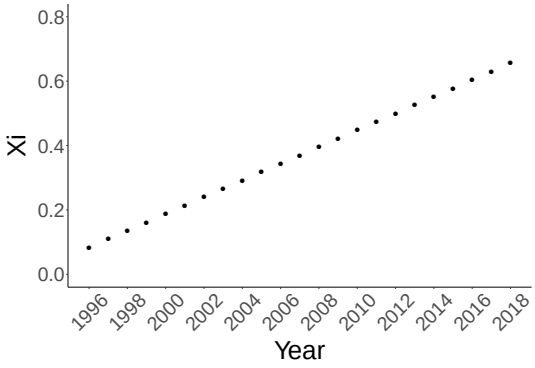
Figure B3: German data: Parameters as functions of the covariates



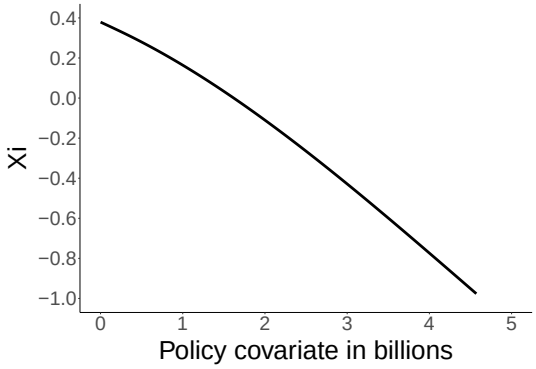

(a) $\rho$ as a function of time



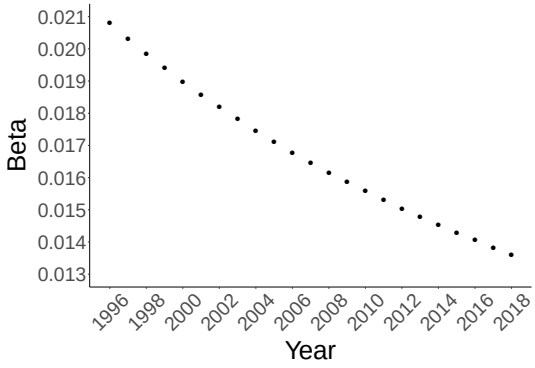

(b) $\rho$ as a function of the gross earned premium



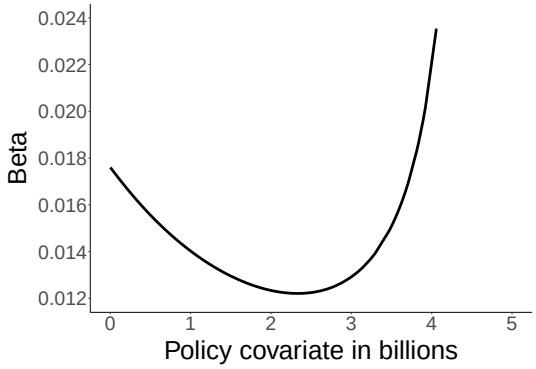

(c) [$\xi$ as a function of time




(d) $\xi$ as a function of the gross earned premium




(e) $\beta$ as a function of time




(f) $\beta$ as a function of the gross earned premium

*Notes:* The figure displays fitted values for the parameters $\rho$, $\xi$, and $\beta$ according to the model given in Table 10 for varying covariate values (gross earned premium or time), fixing all other variables equal to their median values.