Estimating Extreme Cancellation Rates In Life Insurance*

Francesca Biagini[†] Tobias Huber[‡ §] Johannes G. Jaspersen[‡] Andrea Mazzon[†]

May 12, 2019

Abstract

Extreme cancellation rates can severely distort life insurers' liquidity and profitability. Due to the rarity of the event and the complexity of policyholder behavior, the risk assessment of such a scenario is difficult. We introduce an estimation method that can utilize panel data on the company level to estimate the probability distribution function of a mass cancellation event as long as this function is continuous. The panel structure is taken into account by including annual fixed effects and company-level covariates. We demonstrate the method using cancellation rates from U.S. life insurers. We also apply it to German data and reveal difficulties in the European insurance regulation framework Solvency II. Our method allows risk managers and regulators to estimate extreme cancellation rates. In particular, we discuss the (in-)adequacy of the mass lapse scenario assumed in Solvency II, which can lead companies to have solvency capital requirements in the hundreds of millions.

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

JEL Classification: C14 · G18 · G22 · G32

*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, Lu Li, Thorsten Moenig, Joëlle Näger, Andreas Richter, Jochen Ruß, Günter Schwarz, and Richard Vierthauer for valuable comments.

[†]Workgroup Financial and Insurance Mathematics, Munich Risk and Insurance Center, Ludwig-Maximilians-Universität München.

[‡]Institute for Risk Management and Insurance, Munich Risk and Insurance Center, Ludwig-Maximilians-Universität München.

[§]*Corresponding Author*. tobias.huber@lmu.de, +49 89 2180 3883.

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 with the data requirements associated with estimating extreme cancellation events. When considering a single life insurance company, extreme cancellation rates are very rare. We thus face the problem of having many cancellation rate data under normal circumstances but only a few situations in which high cancellation rates are observed. In this paper, we offer an avenue to overcome this challenge. Our approach is based on dynamic extreme value theory by Chavez-Demoulin et al. (2016) but we extend it so that it can be applied to panel data. We consider cancellation rates as realizations of random variables but do not make any assumptions about their distribution function, except that it is continuous.[1] Our method then allows us to use the generalized Pareto distribution to model the excesses over a chosen high threshold for cancellation rates.[2] This enables us to estimate a probability distribution and associated risk measures for extreme cancellation events from readily available panel data on the company level. In addition, our extension of dynamic extreme value theory allows parameters of the probability distribution to be dependent on continuous variables. This provides the possibility to test whether extreme cancellation events are correlated with company-level covariates.

[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 2017 (see Figures 1 and 5). Additionally, the commonly assumed sources (e.g., unemployment rate or interest rate) of cancellation rates are modeled as continuous random variables. These two reasons justify our continuity assumption regarding the cancellation rates' distribution function.

[2] According to the Pickands-Balkema-de Haan theorem (Balkema and de Haan, 1974; Pickands, 1975) this distribution function provides a good fit of the unknown conditional distribution function.

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 acutely interested in assessing their exposure to this risk as accurately as possible. U.S. life insurers can potentially face very high cancellation rates, particularly 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, 2016). 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 the so called lapse risk module (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 consider 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.

Using our model on data from U.S. life insurance demonstrates the applicability of our approach and shows that cancellation rates between 17% and 23% are reflecting the risk of a mass cancellation scenario. Subsample analyses show that this risk markedly depends on the type of product sold by the insurance company. For companies with a high share of term life policies, for example, mass cancellation scenarios with rates up to 37% can be adequate. Furthermore, we use the developed model to discuss the calibration of Solvency II's mass cancellation scenario. For this, we apply the model to German data and obtain that cancellation rates of 13–21% are reflecting a mass cancellation scenario in the German life insurance market. These values show that the arbitrarily chosen cancellation rate of 40% for this scenario in Solvency II might not be 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. This not only demonstrates how our method can test for correlations between company-level covariates and extreme cancellation events, but also encourages further research on other company-specific covariates. The combination of this result 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 and its correlates, but are unable to identify the causal relationship that causes 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 and correlations with company-level covariates. Alternatively, models can use our results by directly calibrating their model on our derived distribution.

The remainder of this paper is structured as follows. The next section develops our empirical model, which we apply to data sourced from the U.S. individual life insurance market in the third section. The fourth section provides implications of our approach for insurance regulation, exemplified on the European regulatory framework Solvency II. The paper ends with concluding remarks and provides directions for further research.

2 Method Development

2.1 Approach

Our approach is based on considering the amount of company-wide policy cancellations 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).[3] 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. A major problem to assess the underlying distribution function is, however, posed by the lack of sufficient extreme cancellation rates. Since the empirical distribution function is unable to assess the risk of extreme cancellation rates appropriately, we will use extreme value theory. One of the most widely used approaches is the so-called 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 will first summarize the POT method as it is described in Embrechts et al. (1997) and Embrechts et al. (2005). 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 such 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. We thus introduce the dynamic POT method in Section 2.3 and its empirical implementation in Section 2.4. This method allows making the parameters of the estimated probability distribution dependent on categorical variables so that we can model the time dependency of the cancellation rates using dummy variables for each year. Categorical variables are, however, unfit to model the company dependency in this case, because we have too few observations of extreme events for reliable estimates. We can circumvent this issue by modeling the extreme cancellation rates to be dependent on a continuous company-level variable, in this case portfolio size. While this does not capture all idiosyncratic factors of companies regarding the cancellation

[3] This is mostly based on the 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).

rates, the company size variable is highly correlated over time, such that it will capture part of the dependency structure. We introduce the necessary extensions to the dynamic POT method in Section 2.5 and name this model extended dynamic POT.

2.2 Peaks Over Threshold Method

Let Y be a random variable on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with cumulative probability distribution function F. Furthermore, let $z_F := \sup\{z \in \mathbb{R} : F(z) < 1\}$ be the finite or infinite right endpoint of F. Then the *excess distribution function* F_u of Y over a threshold u, $0 < u < z_F$, is defined by the conditional probability

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

With $\bar{F} = 1 - F$ and $\bar{F}_u = 1 - F_u$, the latter relation can also be written as

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

We now introduce the *Generalized Pareto Distribution (GPD)*, since it is the natural candidate for the approximation of this excess distribution as stated by a classic result we report below. The GPD with shape parameter $\xi \in \mathbb{R}$ and scale parameter $\beta > 0$ is given by

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

where the support of the GPD is dependent on the parameter ξ:

$$supp\,(G_{\xi,\beta}) = \begin{cases} z \geq 0, & \text{if } \xi \geq 0, \\ 0 \leq z \leq -\beta/\xi, & \text{if } \xi < 0. \end{cases} \tag{2.4}$$

The following theorem (Balkema and de Haan, 1974; Pickands, 1975) motivates the role of the GPD for modeling the excess distribution F_u of an unknown continuous distribution function F.[4]

Theorem 1 *Let Y be a random variable on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with continuous distribution function F. Moreover, let z_F denote the finite or infinite right endpoint of F and let $u > 0$ be a threshold.*

[4] From now on, we use the notation given in Embrechts et al. (2005, p.277).

There exists a (positive-measurable) function $\beta(u)$ such that

$$\lim_{u \to z_F} \sup_{0 \le z < z_F - u} \left| F_u(z) - G_{\xi,\beta(u)}(z) \right| = 0, \tag{2.5}$$

if and only if F lies in the maximum domain of attraction of the generalized extreme value distribution H_ξ (in symbol: $F \in MDA(H_\xi)$), where

$$H_\xi(z) = \begin{cases} \exp\left\{-(1+\xi z)^{-1/\xi}\right\}, & \xi \neq 0, \\ \exp\{-\exp\{-z\}\}, & \xi = 0. \end{cases} \tag{2.6}$$

Proof: See Balkema and de Haan (1974) or Pickands (1975). □

Given an i.i.d. time series, the classic POT method relies on the above theorem in order to split an unknown distribution function F into a part below a "suitable" threshold, where enough data are given so that the empirical distribution provides a good fit, and another part where the generalized Pareto distribution is used to model the excesses over this threshold. Assume that $y_{t'_1}, \dots, y_{t'_n} \in \mathbb{R}$ are i.i.d. realizations at the time points $0 \le t'_1 \le \cdots \le t'_n \le T$. Furthermore, suppose for their common distribution function $F \in MDA(H_\xi)$ for some $\xi \in \mathbb{R}$, and fix u large enough.[5] Then, equation (2.2) provides a method for estimating the far end tail of F by estimating $\bar{F}_u(z)$ as well as $\bar{F}(u)$ and subsequently calculating $\bar{F}(z+u)$ as multiplication of both estimates. A natural estimator for $F(u) = 1 - \bar{F}(u)$ is given by the empirical distribution function

$$F_n(u) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{y_{t'_i} > u\}}. \tag{2.7}$$

Additionally according to Theorem 1, $\bar{F}_u$ can be approximated by $\bar{G}_{\hat{\xi},\hat{\beta}(u)} := 1 - G_{\hat{\xi},\hat{\beta}(u)}$, where $\hat{\xi}$ and $\hat{\beta}(u)$ are estimates of ξ and $\beta(u)$ as given in equation (2.3). Here, we use that the realizations are independent and identically distributed.

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. The goodness of fit of empirical excesses over the chosen threshold u to a parametric GPD model can be evaluated through a Q-Q plot. In such a graph the quantiles of

[5] As remarked in Embrechts et al. (2005, p.267), note that "essentially all the common continuous distributions of statistics or actuarial science are in $F \in MDA(H_\xi)$ for some value ξ". Hence, in the sequel we assume that $F \in MDA(H_\xi)$ for some ξ. For a detailed specification of this family of distribution functions, we refer to Fisher and Tibbett (1928).

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.

After having chosen a threshold u, the parameters $\hat{\xi}$ and $\hat{\beta} = \hat{\beta}(u)$ can be computed via the maximum likelihood method applied to the set of data belonging to $\{y_{t_1}, y_{t_2}, \ldots, y_{t_{n_u}}\}$ where $\{t_1, t_2, \ldots, t_{n_u}\} \subseteq \{t'_1, t'_2, \ldots, t'_n\}$ denotes the subset of points in time t_i, $i = 1, 2, \ldots, n_u$, at which the n_u excesses over the threshold u occur. By using the estimated parameters $\hat{\xi}$ and $\hat{\beta}$ we are able to calculate the value at risk $\widehat{\text{VaR}}_\alpha$ as α-quantile of the estimated distribution function:

$$\widehat{\text{VaR}}_\alpha = u + \frac{\hat{\beta}}{\hat{\xi}} \left[\left(\frac{1-\alpha}{\hat{\bar{F}}_n(u)} \right)^{-\hat{\xi}} - 1 \right]. \tag{2.8}$$

In equation (2.8) a maximum likelihood estimator of $\bar{F}_n(u) := 1 - F_n(u)$ is calculated as $\hat{\bar{F}}_n(u) = n_u/n$, where the number of exceedances n_u is given by:

$$n_u := \sum_{i=1}^{n} \mathbf{1}_{\{y_{t'_i} > u\}}. \tag{2.9}$$

More generally, the number of exceedances n_u is supposed to follow a Poisson process with intensity $\lambda > 0$ (see e.g., Embrechts et al., 1997, p.167).

To compute one estimated value of the parameters ξ and $\beta(u)$ from the observations, the assumptions of independence and identical distribution of the random variables are required. For commonly available panel data these assumptions are certainly questionable. Thus, we choose for our later applications a generalization of the POT method for non-i.i.d. time series. This generalization is called *dynamic POT method* and was introduced in Coles (2001) and Chavez-Demoulin et al. (2016). It allows the parameters to depend on covariates, which can also include time.

2.3 Dynamic Peaks Over Threshold Method

The main feature of the dynamic POT method is the introduction of a dependence of the parameters ξ, β and λ on covariates. Again, ξ and β are parameters of the GPD distribution of excesses over a large threshold and λ is the intensity of the Poisson process, which characterizes the number of excesses. The dependence of these parameters on covariates can be parametric, semi-parametric or non-parametric. We first describe this methodology as reported in Chavez-

Demoulin et al. (2016) who analyze exceedances of operational losses over a large threshold. In Section 2.5 we extend their approach for our purposes.

Denote by x the factor level of a given covariate (for example the level of a macroeconomic factor or a specific business line of a company) and let t designate time. The number of exceedances is assumed to follow a non-homogeneous Poisson process with rate function $\lambda = \lambda(t, x)$. This is a standard generalized additive model, which leads to an estimate for λ as a function of x and t.[6] Moreover, suppose that

$$\xi = \xi(x, t) = g_\xi(x) + h_\xi(t), \tag{2.10}$$

$$\nu = \nu(x, t) = g_\nu(x) + h_\nu(t), \tag{2.11}$$

where

$$\nu := \log((1 + \xi)\beta). \tag{2.12}$$

Here g_ξ, g_ν denote functions in the factor levels of the covariate and $h_\xi, h_\nu : [0, T] \to \mathbb{R}$ are general measurable functions. The choice to re-parameterize the GPD parameter β as a function of ν is due to a technical reason: in this way, the convergence of the simultaneous fitting procedure for the two parameters is guaranteed.[7] From the estimation of ν, one can compute

$$\beta = \beta(x, t) = \frac{\exp(\nu(x, t))}{1 + \xi(x, t)}. \tag{2.13}$$

The estimators $\widehat{g}_\xi$, $\widehat{h}_\xi$, $\widehat{g}_\nu$ and $\widehat{h}_\nu$ of g_ξ, h_ξ, g_ν and h_ν are obtained from the observed vectors $\bar{z}_i = (t_i, x_i, y_{t_i})$, $i = 1, \ldots, n_u$, where $0 \le t_1 \le \cdots \le t_{n_u} \le T$ denote the excess times, x_i, $i = 1, \ldots, n_u$, are the observed covariates, and y_{t_i}, $i = 1, \ldots, n_u$, designate the corresponding realizations of the excesses over the threshold u. Note that u can be chosen with the same procedure as described in Section 2.2.

[6] As pointed out in Chavez-Demoulin et al. (2016), for example, such models can be fit in `R` with the function `gam(..., family=poisson)` from the package `mgcv`.

[7] Note that this re-parameterization guarantees $\beta > 0$, but it is only valid for $\xi > -1$. However, as stated in Chavez-Demoulin et al. (2016), this is in general a rather weak assumption for applications.

2.4 Penalized Maximum Likelihood Estimation And Risk Measures

To fit reasonably smooth functions $\widehat{h}_\xi$ and $\widehat{h}_\nu$ to the observations $\bar{z}_i = (t_i, x_i, y_{t_i})$, $i = 1, \ldots, n_u$, we use the *penalized log-likelihood* as derived in Chavez-Demoulin et al. (2016):

$$\ell^p(g_\xi, h_\xi, g_\nu, h_\nu; \bar{z}_1, \ldots, \bar{z}_{n_u}) = \ell^r(\bar{\xi}, \bar{\nu}; \bar{y}) - \gamma_\xi \int_0^T h_\xi''(t)^2 dt - \gamma_\nu \int_0^T h_\nu''(t)^2 dt. \tag{2.14}$$

Here, we set $\bar{y} = (y_{t_1}, \ldots, y_{t_{n_u}})$. The second and third term in the above penalized log-likelihood ℓ^p avoid overfitting through the choice of the smoothing parameters $\gamma_\xi, \gamma_\nu \in \mathbb{R}$.[8] The first term in equation (2.14) is the re-parametrized log-likelihood ℓ^r and is given by:

$$\ell^r(\bar{\xi}, \bar{\nu}; \bar{y}) = \sum_{i=1}^{n_u} \ell^r(\xi_i, \nu_i; y_{t_i}) = \sum_{i=1}^{n_u} \ell^r(g_\xi(x_i) + h_\xi(t_i), g_\nu(x_i) + h_\nu(t_i); y_{t_i}). \tag{2.15}$$

Finally, by equation (2.12) we can trace the re-parametrized log-likelihood ℓ^r back to the log-likelihood for a single observation $z \in \mathbb{R}$ in the classic POT:

$$\ell(\xi, \beta; z) = \begin{cases} -\log(\beta) - (1 + 1/\xi)\log(1 + \xi z/\beta), & \text{if } \xi > 0, z \geq 0 \text{ or } \xi < 0, z \in [0, -\beta/\xi), \\ -\log(\beta) - z/\beta, & \text{if } \xi = 0, \\ -\infty, & \text{otherwise.} \end{cases} \tag{2.16}$$

The functions g_ξ, h_ξ, g_ν and h_ν can be estimated by using *natural cubic splines*. As reported in Chavez-Demoulin et al. (2016), for a natural cubic spline h with knots $s_1, \ldots, s_m$, $m \in \mathbb{N}$, it holds

$$\int_0^T h''(t)^2 dt = \bar{h}^\top K \bar{h}, \tag{2.17}$$

where $\bar{h} = (h(s_1), \ldots, h(s_m))$ and K is a symmetric (m, m)-matrix of rank $m - 2$ only depending on the knots $s_1 < \cdots < s_m$. The penalized log-likelihood in equation (2.14) can thus be written as

$$\begin{aligned} \ell^p(g_\xi, h_\xi, g_\nu, h_\nu; \bar{z}_1, \ldots, \bar{z}_{n_u}) &= \ell^r(\bar{\xi}, \bar{\nu}; \bar{y}) - \gamma_\xi \bar{h}_\xi^\top K_\xi \bar{h}_\xi - \gamma_\nu \bar{h}_\nu^\top K_\nu \bar{h}_\nu \\ &= \sum_{i=1}^{n_u} \ell^r(g_\xi(x_i) + h_\xi(t_i), g_\nu(x_i) + h_\nu(t_i); y_{t_i}) - \gamma_\xi \bar{h}_\xi^\top K_\xi \bar{h}_\xi - \gamma_\nu \bar{h}_\nu^\top K_\nu \bar{h}_\nu, \end{aligned} \tag{2.18}$$

with $\bar{h}_\xi = (h_\xi(t_1), \ldots, h_\xi(t_{n_u}))$ and $\bar{h}_\nu = (h_\nu(t_1), \ldots, h_\nu(t_{n_u}))$.

[8] Note that larger values of the smoothing parameters lead to smoother fitted curves. A subsection of Chavez-Demoulin et al. (2016) is dedicated to the selection of appropriate smoothing parameters.

These formulas are the basis of a backfitting algorithm for estimating the functions g_ξ, h_ξ, g_ν and h_ν, to obtain the parameters $\hat{\xi}$ and $\hat{\nu}$ (and thus $\hat{\beta}$) for every value of (x,t). The algorithm is described in Appendix A.2 of Chavez-Demoulin et al. (2016), and relies on an iterative weighted least squares procedure. In the first step, all the entries of the vectors $\hat{\underline{\xi}}$ and $\hat{\underline{\nu}}$ are equal, and they correspond to the values computed with the classic POT method. In each following step, the values of $\hat{\underline{\xi}}$ and $\hat{\underline{\nu}}$ are updated through Newton's algorithm to fit the functions g_ξ, h_ξ, g_ν and h_ν to these new values.

Based on the estimates $\hat{\xi}$ and $\hat{\beta}$ for a given covariate level x and time point t, we can then compute estimates of the risk measure VaR depending on x and t for a fixed threshold u. Equation (2.8) now becomes

$$\widehat{\mathrm{VaR}}_\alpha(x,t) = u + \frac{\hat{\beta}(x,t)}{\hat{\xi}(x,t)}\left[\left(\frac{1-\alpha}{\hat{\lambda}(x,t)/n(x,t)}\right)^{-\hat{\xi}(x,t)} - 1\right], \tag{2.19}$$

where $n(x,t)$ stands for the total number of observations for a fixed covariate x and time point t. In applications one directly estimates the rate $\rho(x,t) := \lambda(x,t)/n(x,t)$ through a logistic regression model (Chavez-Demoulin et al., 2016, p.746).

2.5 Dynamic Peaks Over Threshold Method With Quantitative Covariates

The analysis in Chavez-Demoulin et al. (2016) takes the covariate x as a business line of a firm, or as the level of a given characteristic of a company. For example, companies can be classified as large (L), medium (M) or small (S) looking at the number of employees. In this case x represents the firm's classification and takes values in $\{S,M,L\}$. Here, the factor x is not a quantitative variable but rather an indicator. For this reason, g_ξ and g_ν are not functions from the set of real numbers and do not have to be smoothed.

In the present work we extend this approach and let x be a quantitative variable. Later in our applications x represents the insurers' portfolio size.[9] To compute $\hat{\beta}(x,t)$ and $\hat{\nu}(x,t)$ for such a covariate x, we can still apply the algorithm of Chavez-Demoulin et al. (2016) with a changed log-likelihood: in this case, the functions g_ξ and g_ν in the equations (2.10)–(2.11) are defined on $\mathbb{R}$. Hence, we need to introduce two more smoothing parameters $\delta_\xi, \delta_\nu \in \mathbb{R}$. Furthermore, let the covariates for the excesses be given by $0 \le x_1 \le \cdots \le x_{n_u} \le S$. Then the penalized log-likelihood

[9] In the U.S. analysis we will use the number of policies in the insurers portfolio. In the German analysis we use due to data availability the gross earned premium of the insurance companies.

becomes

$$\begin{aligned}\ell^p(g_\xi, h_\xi, g_\nu, h_\nu; \bar{z}_1, \ldots, \bar{z}_{n_u}) = &\ell^r(\bar{\xi}, \bar{\nu}; \bar{y}) - \gamma_\xi \int_0^T h''_\xi(t)^2 dt - \gamma_\nu \int_0^T h''_\nu(t)^2 dt \\ &- \delta_\xi \int_0^S g''_\xi(x)^2 dx - \delta_\nu \int_0^S g''_\nu(x)^2 dx,\end{aligned} \tag{2.20}$$

with $\ell^r(\bar{\xi}, \bar{\nu}; \bar{y})$ as given in equation (2.15).

To the best of our knowledge, this is the first application of the dynamic POT method where a quantitative variable is used in addition to the time variable. We use a vector of time dummies and a covariate that differentiates companies from one another. In this way, we avoid making the assumption that cancellation rates for companies in the same group are i.i.d. over time. Moreover, the effect of pooling provides higher global information: even if every event (t_i, x_i, y_{t_i}) is a priori treated separately, the computation of the parameters $(\hat{\xi}_i, \hat{\beta}_i)$ is also affected by the observations with covariates close to x_i. For this reason, the methodology works best in regions where the values of the covariate x are not too sparse. We demonstrate the feasibility of estimating extreme cancellation events with this procedure in the next section and apply it to the European regulatory framework Solvency II in Section 4.

3 Demonstration

3.1 Data Selection Process

The National Association of Insurance Commissioners (NAIC) provides information on individual life insurance[10] from U.S. insurers' annual statement reports. Individual life insurance policies are either term or permanent insurance policies. Term insurance provides coverage for a predetermined period while permanent provides coverage for the whole life. Permanent insurance is either traditional whole life, universal life, variable life, or variable universal life insurance (ACLI, 2018). For each of the 713 life insurers and each year between 1996 and 2017 we collect data on direct written premiums as well as several variables from the exhibit of life insurance. Summary statistics of these data are reported in Table 1. 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 the other variables are reported in number of policies. For the extended dynamic POT we will use the

[10] Three types of life insurance dominate the U.S. market: Individual, group, and credit life insurance. Individual life insurance is separately underwritten for individuals and accounted for 42% of all number of life insurance policies in force at the end of 2017 (ACLI, 2018).

total of number of policies as one covariate of the estimation since it represents the size of a life insurance company and thus the propensity to be subject to high cancellation rates.[11]

Table 1: U.S. Data: Summary Statistics

Variable	**Description**	**N**	**Mean**	**Pctl(25)**	**Median**	**Pctl(75)**
dwp ($000)	direct written premium	14,546	164,593	16	2,650	47,081
lap	policies lapsed	14,566	23,561	1	417	4,257
sur	policies surrendered	14,571	4,830	0	188	2,168
$lost$	policies lost	14,561	48,735	40	1,669	12,251
$issued$	policies issued	14,560	15,682	0	177	5,583
$assumed$	policies assume	14,349	33,280	0	0	0
$revived$	policies revived	14,565	1,868	0	4	179
$inForce_T$	term life in force	14,568	228,654	0	811	20,038
$inForce_P$	permanent life in force	14,567	162,080	48	9,314	75,240
$inForce$	total in force	14,570	399,515	480	16,410	122,329
$cancel_{Full}$	cancellation rates (Full)	12,478	0.0650	0.0275	0.0484	0.0764
$cancel_{Selected}$	cancellation rates (Selected)	6,513	0.0557	0.0344	0.0491	0.0697

Notes: The table reports summary statistics of the variables used in the analysis of extreme cancellation rates in the U.S. individual life insurance industry. Data is comprised of 713 life insurance companies between the years of 1996 and 2017. The index $Full$ indicates the full sample, while the index $Selected$ indicates the sample after the data selection process. The number of cancellation rates in the full sample is lower than for the other variables. This is due to some observations exhibiting no insurance in force and no policies lost such that the cancellation rate calculated as in equation (3.1) is not defined.

To calculate cancellation rates for life insurer i in year t we divide the number of policies that were canceled in t by the number of policies that were in force in t:

$$cancel(i,t) := \frac{lap(i,t) + sur(i,t)}{lost(i,t) + inForce(i,t)}. \tag{3.1}$$

In this definition lap denotes the contracts that were canceled as result of nonpayment of premium and $surr$ designates contracts that were canceled and a cash surrender value was paid to the policyholder. To capture all policies in force during the year we add to the policies in force at the end of year all policies that were lost during the year. Policies lost comprises both lapsed and surrendered policies as well as policies that dropped out of the life insurer's portfolio due to other reasons (e.g., death or maturity). We denote the approximation for the life insurance policies in

[11] Panel (a) and Panel (c) in Figure 8 in the Online Appendix show the dependency of the frequency and size of excesses over a threshold of 14.5% on the number of policies. We observe a decreasing trend for the frequency and size of excesses in dependence of the number of policies.

force during the year by $inForceProxy$:

$$inForceProxy(i,t) := lost(i,t) + inForce(i,t). \tag{3.2}$$

We select our data to avoid inconsistencies and to restrict the analysis to representative U.S. life insurance companies. For each step of the data selection process, Table 2 reports the number of cancellation rates and the number of companies in the remaining sample after the respective step has been applied.

Table 2: U.S Data: Sample Selection Process

Step	**Description**	**Number of cancellation rates**	**Number of companies**
0	Full sample	12,478	713
1	Keep companies that have at least one positive direct written premium between 1996 and 2017	11,541	581
2	Keep data points with $shareOfAssumed$ less than or equal to 20%	11,295	579
3	Keep data points with $inForceProxy$ greater than or equal to 10,000	7,508	424
4	Keep data points with $shareOfIssued$ less than or equal to 20%	6,910	411
5	Keep companies with $shareOfRevived$ never being greater than 5%	6,513	384

Notes: The table describes the steps of the sample selection process and provides the number of cancellation rates and the number of companies in the remaining sample after the respective step has been applied.

The full sample includes 12,478 cancellation rates of 713 life insurance companies between 1996 and 2017.[12] In the first selection step we keep only companies that do exhibit at least one positive direct written premium between 1996 and 2017. Similarly, the second step removes data points with a share of assumed policies

$$shareOfAssumed(i,t) := assumed(i,t)/inForceProxy(i,t), \tag{3.3}$$

exceeding 20%. Here, $assumed$ denotes additions to the portfolio due to reinsurance or coinsurance activities. The first two steps therefore control for coinsurance and reinsurance activity of the life insurers as our main focus lies on primary insurers.

[12] The number of cancellation rates (12,478) is aggregated over all life insurance companies and all years. The number of cancellation rates in the full sample is considerably lower than the other variables' number of observations. This is due to a high share of observations exhibiting no insurance in force and no policies lost.

Thirdly, we keep only cancellation rates if the proxy of insurance in force ($inForceProxy$) is greater than or equal to 10,000 policies. Moreover, cancellation rates are high in the first contract years such that a high share of new business could lead to an upwards bias of cancellation rates. New business activity is tracked by the policies that were issued during the year. In step four we therefore keep data points if the share of issued policies

$$shareOfIssued(i,t) := issued(i,t)/inForceProxy(i,t), \tag{3.4}$$

does not exceed 20%. Step five addresses the revival of lapsed policies. Because lapsed policies can be revived as long as five years after the lapse has happened, we are unable to track whether a high number of lapses can be explained by a high number of revivals in the subsequent year. Thus, we keep only companies that never exhibit a share of revived policies

$$shareOfRevived(i,t) := revived(i,t)/inForceProxy(i,t), \tag{3.5}$$

greater than 5%. The remaining sample after these five data selection steps still represents 90% of aggregate direct written premiums and 55% of aggregate polices of the full sample.[13] A descriptive summary of both cancellation rates in the full sample ($cancel_{Full}$) and the remaining 6,513 cancellation rates ($cancel_{Selected}$) is given in Table 1. The statistics are in line with reported cancellation rates by SOA and LIMRA (2012) as well as ACLI (2018). We display the histogram of the cancellation rates in the final sample in Figure 1.

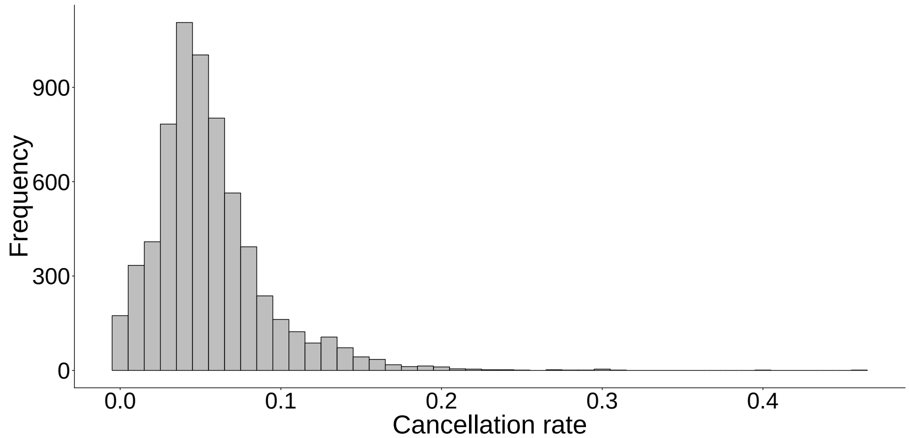
3.2 Model Selection

The first step of the model specification is the threshold selection. As explained in Section 2.2 we compare Q-Q plots of the empirical data for different thresholds. The threshold $u = 14.5\%$ is the smallest threshold such that we observe a good fit of the empirical data in the Q-Q plot. This also supports the choice of the GPD for modeling the excesses over this threshold. Panel (a) in Figure 3 provides the corresponding Q-Q plot. 158 cancellation rates, which correspond to 2% of all rates, lie above the chosen threshold of $u = 14.5\%$. The classic POT method pools all observations of excesses and estimates single parameters ξ, β, and λ for the whole population. The extended dynamic POT, though, uses the list of 158 triples consisting of excess, time of excess (covariate) and number of policies (covariate) as input.[14] It then estimates for each triple indi-

[13] The direct written premiums and number of policies are aggregated over time and companies.

[14] Panel (c) in Figure 8 in the Online Appendix shows that the excesses over the threshold exhibit a decreasing trend in dependence of the number of policies.

Figure 1: Histogram Of U.S. Cancellation Rates



Notes: The figure displays the distribution of the 6,513 cancellation rates between 1996 and 2017 used in the main analysis. The sample of cancellation rates was selected from the full data by the procedure described in Table 2.

vidual parameters ξ, β, and λ. As outlined in Chavez-Demoulin et al. (2016) we can estimate the rate $\rho(x,t) := \lambda(x,t)/n(x,t)$ by running a logistic regression rather than estimating λ directly. We observe that the frequency of an exceedance over the threshold is dependent on both the number of policies and time.[15] To select an appropriate model for the parameters ξ and β we choose different dependencies of these parameters on the covariates number of policies and time. Figure 2 provides the different specifications that we consider. Based on these specifications we perform the corresponding estimation. In each case we then plot the quantiles of the estimation's residuals against the exponential distribution and analyze the model specification's goodness of fit. In the U.S. market we find the best fit for model specification F. The Q-Q plots for all considered model specifications are given in Figure 2. If Q-Q plots are slightly better than the finally chosen model specification we base or decision on how well the model specification exhibits a decreasing trend of 99.5%-quantiles in the covariate regarding the number of policies. Finally, we specify our model parameters as follows:

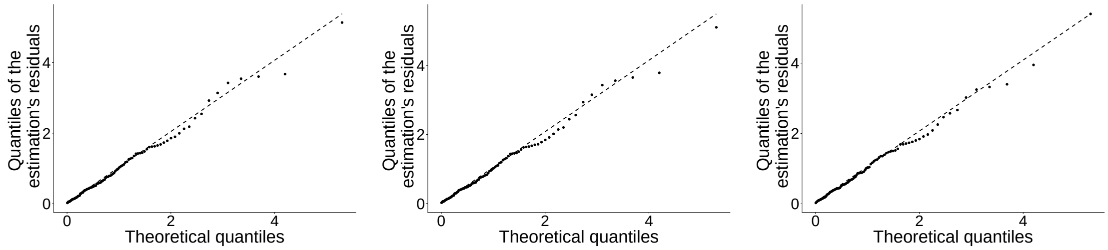
$$\rho(x,t) := \beta_0 + \beta_1 \cdot x + \beta_2 \cdot t, \tag{3.6}$$

$$\xi(x,t) := c_\xi + g_\xi(x) + h_\xi(t), \tag{3.7}$$

$$\nu(x,t) := c_\nu + g_\nu(x) + h_\nu(t). \tag{3.8}$$

[15] Plotting the frequency of an exceedance over the threshold shows a decreasing trend in dependence of the deciles of the company size covariate (see Panel (a) in Figure 8 in the Online Appendix).

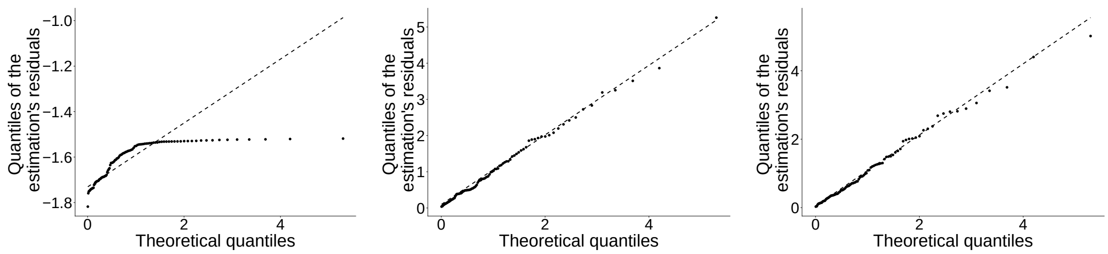
Figure 2: U.S. Data: Goodness Of Fit Based On Residuals



(a) Model A:
$\xi(x,t) := c_\xi$, $\nu(x,t) := c_\nu$

(b) Model B:
$\xi(x,t) := c_\xi + g_\xi(x)$,
$\nu(x,t) := c_\nu$

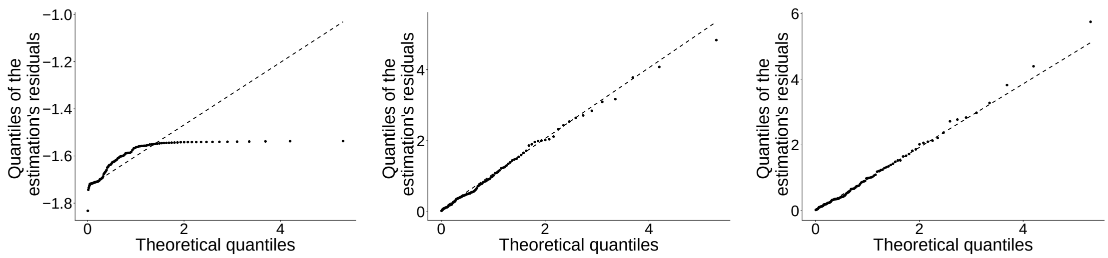
(c) Model C:
$\xi(x,t) := c_\xi + g_\xi(x) + h_\xi(t)$,
$\nu(x,t) := c_\nu$



(d) Model D:
$\xi(x,t) := c_\xi + g_\xi(x) + h_\xi(t)$,
$\nu(x,t) := c_\nu + g_\nu(x)$

(e) Model E:
$\xi(x,t) := c_\xi + g_\xi(x) + h_\xi(t)$,
$\nu(x,t) := c_\nu + h_\nu(t)$

(f) Model F:
$\xi(x,t) := c_\xi + g_\xi(x) + h_\xi(t)$,
$\nu(x,t) := c_\nu + g_\nu(x) + h_\nu(t)$



(g) Model G:
$\xi(x,t) := c_\xi + g_\xi(x)$,
$\nu(x,t) := c_\nu + g_\nu(x)$

(h) Model H:
$\xi(x,t) := c_\xi + g_\xi(x)$,
$\nu(x,t) := c_\nu + h_\nu(t)$

(i) Model I:
$\xi(x,t) := c_\xi + g_\xi(x)$,
$\nu(x,t) := c_\nu + g_\nu(x) + h_\nu(t)$

Notes: The figure displays graphical analyses of the goodness of fit of different model specifications (model A–I) for the estimation of the GPD based on the U.S. data. Each individual panel shows a Q-Q plot with the theoretical quantiles of the exponential distribution on the x-axis and the estimation's residuals on the y-axis. The better a model's goodness of fit, the more closely the data in the Q-Q plot resembles the displayed dashed line. The chosen model's Q-Q plot for the U.S. analysis is depicted in Panel (f).

3.3 Results

We estimate the model as specified in the equations (3.6)–(3.8). The Panels (b), (c), (d) in Figure 3 give the results of our estimation. According to Section 2.4 we avoid overfitting by penalizing the maximum likelihood estimation. We do not specify this penalization term ex-ante but rather use cross-validation so that the penalization term is specified by the data themselves.[16] Panel (b) shows the 99.5%-quantile based on the classic POT method as well as the median of 99.5%-quantiles for the extended dynamic POT. The figure also depicts 90% confidence intervals of the extended dynamic POT estimation based on bootstrapping. It shows that median of 99.5%-quantiles for the extended dynamic POT lies between 17% and 23% and can over- and underestimate the classic POT's estimate of 20%. Additionally, Panel (c) provides a boxplot of the estimated 99.5%-quantiles for each year between 1996 and 2017 and supports the quality of the 90% confidence interval as reasonable upper bound for extreme cancellation rates. Furthermore, Panel (d) shows boxplots of the 99.5%-quantile dependent on the deciles of the number of policies. The figure indicates that the level of extreme cancellation rates is indeed decreasing in the number of policies.[17]

To investigate the robustness of our results, we perform a sensitivity analysis with respect to the choice of the threshold u and the choice of 10,000 policies as lower bound in the data selection process.[18] Table 3 shows no major changes in both the classic and dynamic predictions for the 99.5%-quantile of cancellation rates. However, we cannot control for decreasing interest rates during the observed time period. This can potentially underestimate the mass cancellation risk. Previous literature often models extreme cancellation rates as a result of increasing interest rates (Loisel and Milhaud, 2011; Barsotti et al., 2016).

Cancellation rates of life insurance policies often differ with the type of the policy.[19] To understand how the product portfolio influences our results, we perform a subsample analysis and distinguish between those life insurers that exhibit a high share of term life insurance and those that exhibit a high share of permanent life insurance. Starting with the sample of the main analysis, we keep only data points with a share of term life insurance business higher than 50% and repeat our estimations.[20] As can be seen in Panel (a) of Figure 4, we find that a portfolio with a high share of term life business is more exposed to high cancellation rates compared to the whole sample. The median of 99.5%-quantiles of cancellation rates based on the extended dynamic POT lies between 18% and 37% for this sample.[21] Furthermore, Panel (b) of Figure 4 shows that a port-

[16] This procedure is implemented in `gam(...)` from the package `mgcv`.

[17] The decreasing trend of 99.5%-quantile in the number of policies becomes even more apparent if we look at it for each year individually (see Figure 9 in the Online Appendix).

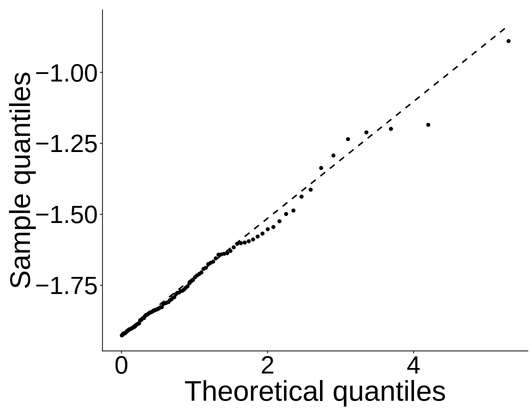
[18] Additionally, we perform a robustness analysis with the log-transformed company size covariate in Figure 11 in the Online Appendix. The results are close to the results of Figure 3.

[19] See, e.g., Renshaw and Haberman (1986), Cerchiara et al. (2009) and Milhaud et al. (2010) for empirical evidence on this.

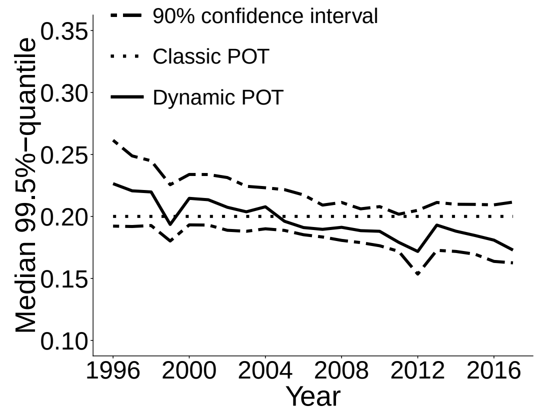
[20] Changing this threshold from 50% to 70% does not change the results. However, the sample size is then no longer sufficient to calculate confidence intervals.

[21] For this analysis we choose a threshold of $u = 12\%$.

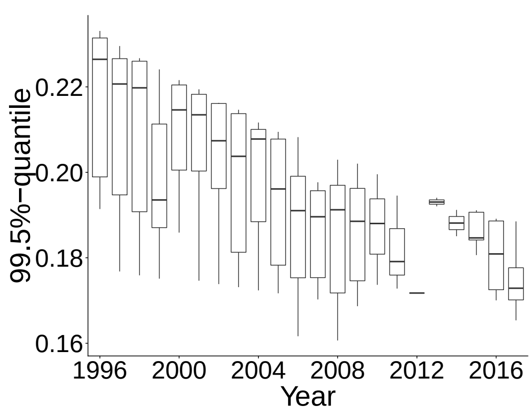
Figure 3: U.S. Data: Sample Q-Q Plot And Estimation Results



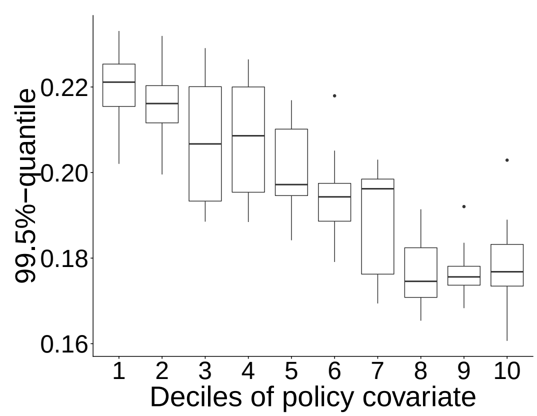
(a) Q-Q plot of the observed cancellation rates above the chosen threshold



(b) Median 99.5%-quantile of the extended dynamic POT (with 90% confidence intervals) and classic POT estimation



(c) Boxplot of 99.5%-quantile estimations over time



(d) Boxplot of 99.5%-quantile estimations for the deciles of the number of policies

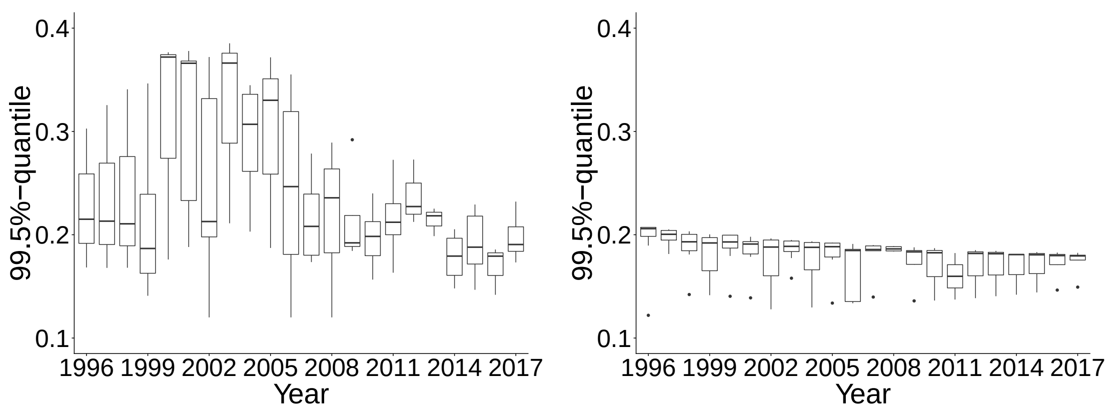
Notes: The figure displays the Q-Q plot of the observed cancellation rates in Panel (a) and estimation results in Panel (b)–(d). The Q-Q plot in Panel (a) is based on the empirical sample and justifies the chosen threshold and the use of the GPD as excess distribution. In contrast to Panel (a), the Q-Q plot in Panel (f) of Figure 2 is based on the predicted residuals. Panel (b) shows the median of predicted 99.5%-quantiles for the cancellation rates based on the classic and the extended dynamic POT method. For each year we have for each triple of excess, time and company size covariate a sample of estimated 99.5%-quantiles for the cancellation rates. The median of this sample of quantiles is depicted in this panel. Additionally, it displays 90% confidence intervals of the extended dynamic POT method. The confidence intervals are bootstrapped with 100 draws. Panel (c) provides for each year between 1996 and 2017 a boxplot of estimated 99.5%-quantiles. Panel (d) shows the boxplot of estimated 99.5%-quantiles for the deciles of the company size covariate. In both cases we again make use of the sample of estimated 99.5%-quantiles for the cancellation rates for each triple of excess, time and company size covariate (here: number of policies).

Table 3: U.S. Data: Sensitivity Analysis

Threshold policies	**u**	n_u	$1 - \frac{n_u}{n}$	Q_C	Q_D
0	0.145	323	0.9653	0.2648	0.2029–0.3176
1,000	0.145	249	0.9697	0.2319	0.2091–0.2551
5,000	0.145	189	0.9738	0.2108	0.1966–0.2329
10,000	0.100	626	0.9039	0.2058	0.1791–0.2367
10,000	0.120	374	0.9426	0.2020	0.1738–0.2291
10,000	**0.145**	**158**	**0.9757**	**0.2000**	**0.1718–0.2264**
10,000	0.160	95	0.9854	0.2001	0.1993–0.1997
10,000	0.180	56	0.9914	0.2002	0.1804–0.1859
20,000	0.145	119	0.9793	0.1828	0.1586–0.2060
100,000	0.145	67	0.9799	0.1803	0.1617–0.1826

Notes: The table provides a sensitivity analysis with respect to step 3 in the data selection process and the choice of the threshold u. n_u denotes the number of observations above the threshold and $1 - \frac{n_u}{n}$ gives the share of observations not utilized for the estimation. Q_C is the 99.5%-quantile resulting from the classic POT method. Q_D gives the range of the 99.5%-quantiles of resulting from the extended dynamic POT method.

Figure 4: U.S. Data: Subsample Analysis



(a) Term life: boxplot of 99.5%-quantile estimations over time

(b) Permanent life: boxplot of 99.5%-quantile estimations over time

Notes: The figure displays estimation results of the subsample analysis. Panel (a) provides for each year between 1996 and 2017 a boxplot of estimated 99.5%-quantiles for the subsample of companies with mainly term life policies in their portfolio. Panel (b) shows the same figure for the subsample of companies with mainly permanent life policies.

folio consisting mainly of permanent insurance is less exposed to a mass cancellation scenario. We restrict our sample to data with a share of permanent insurance higher than 95% and find that the median of 99.5%-quantiles of extreme cancellation rates here lies between 16% and 21%.[22]

[22] For this analysis we choose a threshold of $u = 11\%$.

4 Application: Calibrating Solvency II's Mass Lapse Scenario

4.1 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 explicitly requires to calculate additional capital demand for adverse cancellation events such as a permanent increase or decrease of future cancellation rates. Additionally, to calculate their solvency capital requirement 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. Studies point out that a mass cancellation shock is difficult to calibrate as only poor data is available but a shock cancellation rate of 40% seems too high (CEA, 2009; KPMG, 2015). Additionally, the risk of a mass cancellation scenario might differ between countries, which is not reflected in this harmonization of insurance regulation (IAIS, 2016). Finally, such a mass cancellation scenario is costly to insurers and lets them seek for capital relief through reinsurance of this risk (Risk.net, 2015). 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. The Austrian life insurer UNIQUA (2017) amounts the capital demand for this risk to be €262 million.

In contrast to the expert judgment based approach in the Solvency II standard model, our approach can use data to calibrate an empirically justified mass cancellation scenario. In this way we address the question whether the current scenario of 40% is justified. For this purpose, we again use available panel data on the company level. Of the European countries, Germany has the most readily available and reliable data. Using only German data does not allow us to make a general statement about the European life insurance market. However, the European market is highly decentralized into national clusters. As such, the German data will give us a reasonable estimate of the German life insurance market. In addition, by analyzing the impact of company-level covariates on the mass cancellation scenario, we can obtain a first indication whether the assumption of an equal mass cancellation scenario for all national life insurance markets in Europe is justified.

4.2 Data Selection Process

The German Federal Financial Supervisory Authority (BaFin) provides information on life insurance from the insurers' annual statement in the time period 1996–2017. For each life insurance company and each year we collect gross earned premiums and cancellation rates. Cancellation rates given on a company level are averaged over all product categories offered by the life insurer. German life insurers predominantly offer the products endowment, term life, occupational disability, annuity, unit linked insurance, and group life insurance.[23] While we are unable to distinguish between these product categories Eling and Kiesenbauer (2013) find that in contrast to other markets, the German life insurance market exhibits only small differences in cancellation rates between different product categories.[24]

We again apply filters to the data to avoid inconsistencies and to restrict the analysis to representative German life insurance companies. Specifically, we keep only life insurers with greater than or equal to 60 million Euro gross earned premiums. This step reduces the number of cancellation rates from 2,283 to 1,708. A descriptive summary of both cancellation rates in the full sample ($cancel_{Full}$) and the remaining 1,708 cancellation rates ($cancel_{Selected}$) is given in Table 4. A histogram of the cancellation rates in the final sample is given in Figure 5.

Table 4: German Data: Summary Statistics

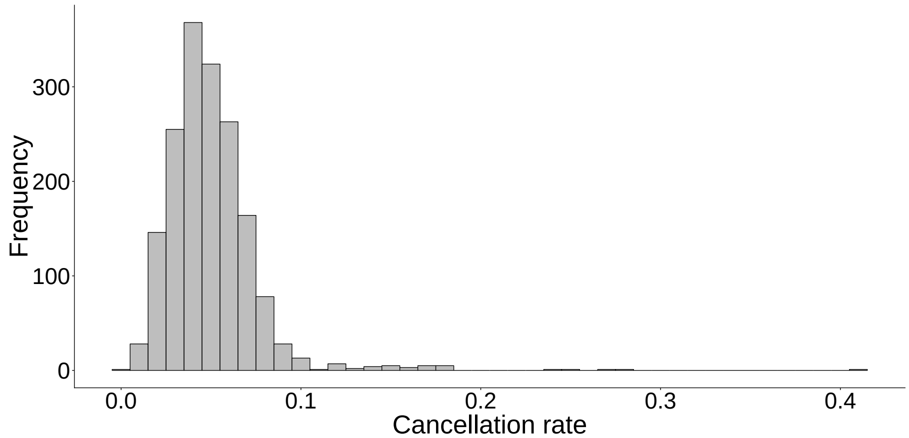
Variable	**Description**	**N**	**Mean**	**Pctl(25)**	**Median**	**Pctl(75)**
gep	gross earned premium	2,295	697 m	56 m	208 m	746 m
$cancel_{Full}$	cancellation rates (Full)	2,283	0.0520	0.0340	0.0471	0.0623
$cancel_{Selected}$	cancellation rates (Selected)	1,708	0.0501	0.0350	0.0470	0.0610

Notes: This table reports summary statistics of the variables used in the analysis of extreme cancellation rates in the German life insurance industry. Data is comprised of 125 life insurance companies between the years of 1996 and 2017. The index $Full$ indicates the full sample, while the index $Selected$ indicates the sample after the data selection process. For small life insurance companies one of the variables gep and $cancel_{Full}$ is sometimes missing such that the number of observations for these variables does not coincide.

[23] An endowment policy is contracted for a specified time and 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 is contracted for a specified time and pays the sum insured if the insured person dies within the contract period. An occupational disability policy is contracted for a specified time and pays an annuity until retirement if the insured person is no longer able to work. An annuity policy is contracted for a specified time and 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 on Riester and Rürup). Finally, a unit linked policy is contracted for a specified time and 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.

[24] However, a thorough analysis of the German market by product type could be a promising area for future research.

Figure 5: Histogram Of German Cancellation Rates



Notes: The figure displays the distribution of the 1,708 cancellation rates between 1996 and 2017 used in the main analysis. The sample of cancellation rates was selected from the full data by limiting the analysis to insurance companies with greater than or equal to 60 million Euro gross earned premium.

4.3 Model Selection

A Q-Q plot[25] supports the GPD and the choice of $u = 7.7\%$ as a threshold (see Panel (a) in Figure 7). As before the classic POT method pools all observations of excesses while the extended dynamic POT takes time and company-specific characteristics into account. Due to data availability we here use, in contrast to the estimation with U.S. data, the gross earned premium instead of the number of policies as covariate.[26] As in the U.S. market we observe that the frequency of an exceedance over the threshold is dependent on both the company size covariate and time.[27] We then run a logistic regression for $\rho(x,t) := \lambda(x,t)/n(x,t)$. Additionally, we choose the parameter ξ to be only dependent on the gross earned premium and the parameter ν be dependent on gross earned premium and time. This specification is a result of comparing Q-Q plots of the estimation's residuals for different model specifications for these parameters. Figure 6 provides a comparison of Q-Q plots for different model specifications of the parameters ξ and ν. The finally chosen model

[25] We plot the logarithm of excesses over the threshold $u = 7.7\%$ (y-axis) against the exponential distribution as reference distribution (x-axis).

[26] Panel (d) in Figure 8 in the Online Appendix shows that the excesses over the threshold $u = 7.7\%$ exhibit a decreasing trend in dependence of the premium covariate.

[27] Plotting the frequency of an exceedance over the threshold shows a decreasing trend in dependence of the deciles of the premium covariate (see Panel (b) in Figure 8 in the Online Appendix).

specification is as follows:

$$\rho(x,t) := \beta_0 + \beta_1 \cdot x + \beta_2 \cdot t, \tag{4.1}$$

$$\xi(x,t) := c_\xi + g_\xi(x), \tag{4.2}$$

$$\nu(x,t) := c_\nu + g_\nu(x) + h_\nu(t). \tag{4.3}$$

4.4 Results

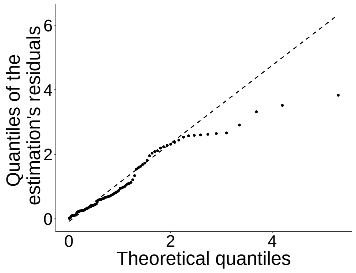
We estimate the model as specified in the equations (4.1)–(4.3). The estimation results are shown in Figure 7. Panel (b) shows the 99.5%-quantile based on the classic POT method as well as the median of 99.5%-quantiles for the extended dynamic POT. The figure also depicts 90% confidence intervals of the extended dynamic POT estimation based on bootstrapping. We see that the median of 99.5%-quantiles for the extended dynamic POT lies between 13% and 21% and can over- and underestimate the classic POT's estimate of 19%. Again, Panel (c) reports boxplots for the 99.5%-quantiles for each year between 1996 and 2017. It provides information about the variability of the risk measure for the cancellation rates. Finally, in Panel (d) we can observe a decreasing trend of the 99.5%-quantiles in the gross earned premium.[28]

Our results allow three conclusions. Firstly, the current scenario for a mass cancellation event in Solvency II's standard model has, at least for the German life insurance market, no empirical foundation. Given the large reserves which can be required due to the scenario, this fact should lead to further investigation by the regulatory agencies. Secondly, the result that the 99.5%-quantile of the cancellation rates depends on the size of the insurance company indicates that idiosyncratic factors of life insurance companies influence the mass cancellation scenario. This gives a certain validity for the use of (partial) internal models to calculate the mass cancellation scenario. Lastly, the marked difference between the 99.5%-quantiles of differently sized life insurance companies combined with the fact that insurance companies differ greatly between different European insurance markets (Insurance Europe, 2019) indicate that modeling the mass cancellation scenario equally in all European insurance markets might be questionable.

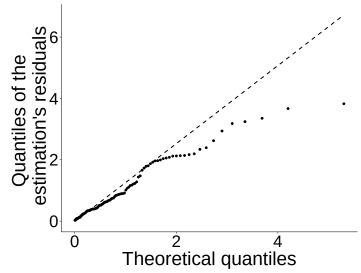
To analyze the robustness of our results and thus our conclusions, we perform a sensitivity analysis with respect to the choice of 60 million Euro gross earned premium as a lower bound in the data selection process. Additionally, we investigate the effect of variations in the choice of the

[28] Again, the decreasing trend of 99.5%-quantile in the company size covariate becomes even more apparent if we look at it for each year individually (see Figure 10 in the Online Appendix).

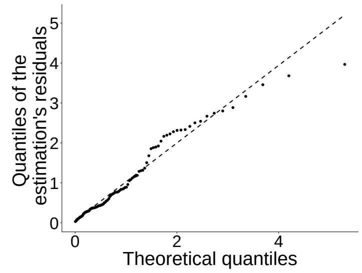
Figure 6: German Data: Goodness Of Fit Based On Residuals



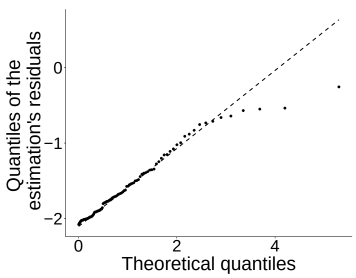
(a) Model A:
$\xi(x,t) := c_\xi$, $\nu(x,t) := c_\nu$



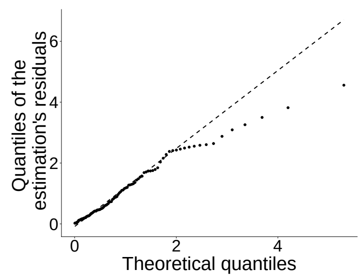
(b) Model B:
$\xi(x,t) := c_\xi + g_\xi(x)$,
$\nu(x,t) := c_\nu$



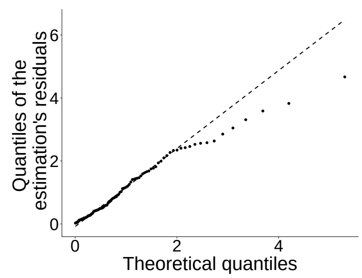
(c) Model C:
$\xi(x,t) := c_\xi + g_\xi(x) + h_\xi(t)$,
$\nu(x,t) := c_\nu$



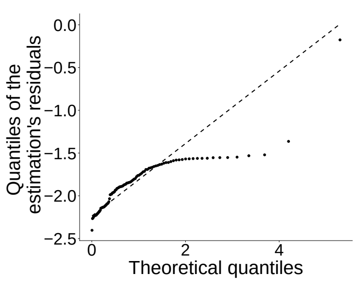
(d) Model D:
$\xi(x,t) := c_\xi + g_\xi(x) + h_\xi(t)$,
$\nu(x,t) := c_\nu + g_\nu(x)$



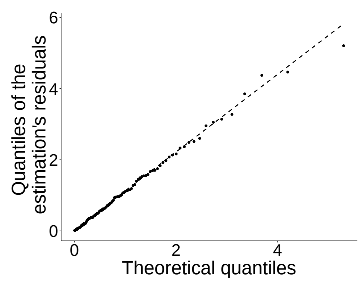
(e) Model E:
$\xi(x,t) := c_\xi + g_\xi(x) + h_\xi(t)$,
$\nu(x,t) := c_\nu + h_\nu(t)$



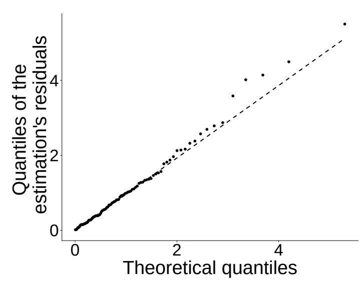
(f) Model F:
$\xi(x,t) := c_\xi + g_\xi(x) + h_\xi(t)$,
$\nu(x,t) := c_\nu + g_\nu(x) + h_\nu(t)$



(g) Model G:
$\xi(x,t) := c_\xi + g_\xi(x)$,
$\nu(x,t) := c_\nu + g_\nu(x)$



(h) Model H:
$\xi(x,t) := c_\xi + g_\xi(x)$,
$\nu(x,t) := c_\nu + h_\nu(t)$



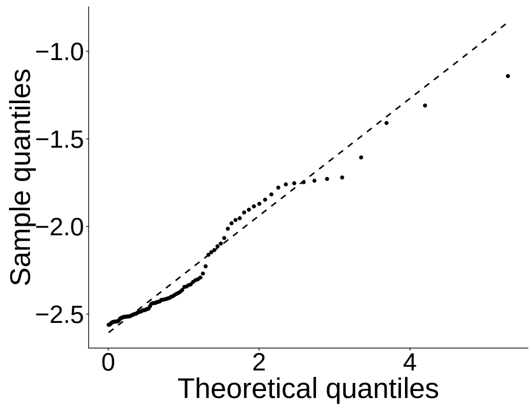
(i) Model I:
$\xi(x,t) := c_\xi + g_\xi(x)$,
$\nu(x,t) := c_\nu + g_\nu(x) + h_\nu(t)$

Notes: The figure displays graphical analyses of the goodness of fit of different model specifications (model A–I) for the estimation of the GPD based on the German data. Each individual panel shows a Q-Q plot with the theoretical quantiles of the exponential distribution on the x-axis and the estimation's residuals on the y-axis. The better a model's goodness of fit, the more closely the data in the Q-Q plot resembles the displayed dashed line. The chosen model's Q-Q plot for the German analysis is depicted in Panel (i).

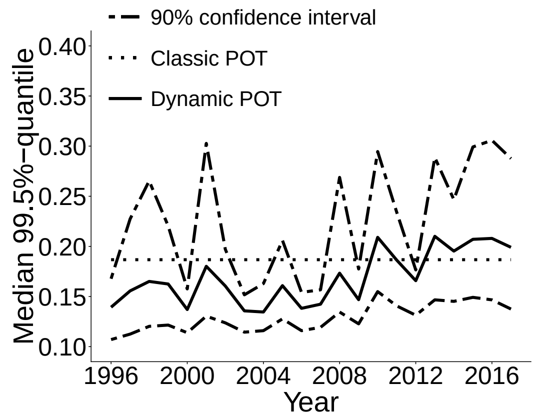
threshold u.[29] Table 5 shows no major changes in both the classic and extended dynamic predic-

[29] Furthermore, we perform a robustness analysis with the log-transformed company size covariate in Figure 12 in the Online Appendix. The results are close to the results of Figure 7.

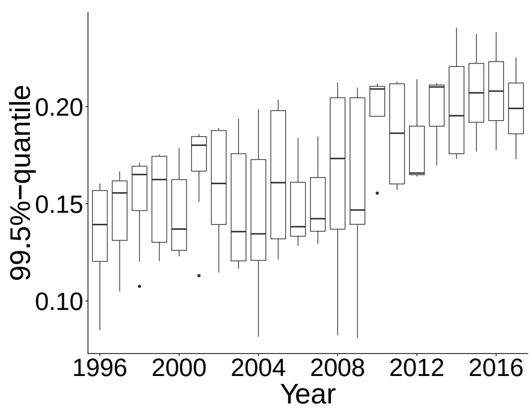
Figure 7: German Data: Sample Q-Q Plot And Estimation Results



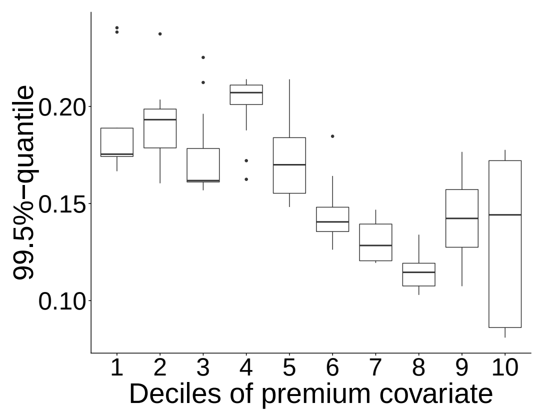
(a) Q-Q plot of the observed cancellation rates above the chosen threshold



(b) Median 99.5%-quantile of the extended dynamic POT (with 90% confidence intervals) and classic POT estimation



(c) Boxplot of 99.5%-quantile estimations over time



(d) Boxplot of 99.5%-quantile estimations for the deciles of the gross earned premium

Notes: The figure displays the Q-Q plot of the observed cancellation rates in Panel (a) and estimation results in Panels (b)–(d). The Q-Q plot in Panel (a) is based on the empirical sample and justifies the chosen threshold and the use of the GPD as excess distribution. In contrast to Panel (a), the Q-Q plot in Panel (i) of Figure 6 is based on the predicted residuals. Panel (b) shows the median of predicted 99.5%-quantiles for the cancellation rates based on the classic and the extended dynamic POT method. For each year we have for each triple of excess, time and company size covariate a sample of estimated 99.5%-quantiles for the cancellation rates. The median of this sample of quantiles is depicted in this panel. Additionally, it displays 90% confidence intervals of the extended dynamic POT method. The confidence intervals are bootstrapped with 100 draws. Panel (c) provides for each year between 1996 and 2017 a boxplot of estimated 99.5%-quantiles. Panel (d) shows the boxplot of estimated 99.5%-quantiles for the deciles of the company size covariate. In both cases we again make use of the sample of estimated 99.5%-quantiles for the cancellation rates for each triple of excess, time and company size covariate (here: gross earned premium).

tions for the 99.5%-quantile of cancellation rates. As in the U.S. data, we have mostly observed decreasing interest rates between 1996 and 2017 in Germany. Therefore, our estimation could po-

Table 5: German Data: Sensitivity Analysis

Threshold premium	**u**	n_u	$1-\frac{n_u}{n}$	Q_C	Q_D
0	0.077	252	0.8895	0.2214	0.1678–0.2354
30 m	0.077	166	0.9128	0.1958	0.1506–0.2048
60 m	0.060	448	0.7372	0.1675	0.1187–0.1634
60 m	0.070	228	0.8663	0.1800	0.1433–0.1767
60 m	**0.077**	**134**	**0.9214**	**0.1867**	**0.1345–0.2101**
60 m	0.090	61	0.9642	0.2223	0.1177–0.2114
60 m	0.100	42	0.9754	0.2211	0.1441–0.2290
100 m	0.077	108	0.9274	0.1727	0.1369–0.1837
150 m	0.077	94	0.9277	0.1832	0.1237–0.1682

Notes: The table provides a sensitivity analysis with respect to the data selection process and the choice of the threshold u. n_u denotes the number of observations above the threshold and $1-\frac{n_u}{n}$ gives the share of observations not utilized for the estimation. Q_C is the 99.5%-quantile resulting from the classic POT method. Q_D gives the range of the 99.5%-quantiles of resulting from the extended dynamic POT method.

tentially underestimate the risk of a mass cancellation event for periods with rising interest rates.

5 Conclusion

We contribute to the literature by providing a model to assess the risk of a mass cancellation scenario in life insurance. As extreme cancellation rates are rare events, cancellation rates sourced from only one insurer do not provide enough data to assess this tail risk. We thus develop a model that can be applied to a panel of insurance companies. The model is an extension of the dynamic Peaks Over Threshold model by Chavez-Demoulin et al. (2016), which can take continuous covariates into account.

We apply this non-parametric approach to U.S. data and show that, depending on product type, cancellation rates up to 37% 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 13–21% are reflecting the risk of a mass cancellation scenario in the German life insurance 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 company 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 whole life insurance indicate that the product type has an effect on the appropriate mass cancellation scenario. Because 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 use our estimation procedure on a richer dataset to analyze the mass cancellation scenario by product type in the German market. Given the results of the product type subsample analysis for the U.S. market, 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. A more direct test of this modeling approach would, however, require data from several different national insurance markets in Europe.[30]

Our model enables risk managers to assess the risk of extreme cancellation rates and regulators to investigate the appropriateness of Solvency II's mass lapse scenario. However, this is not the only possible application of the methodology. The extension to the dynamic Peaks Over Threshold method allows for the use of continuous covariates in any data set of companies. This can be used to calibrate probability distributions of any kind of extreme event from data of companies or individuals.

[30] Note that different national markets likely report their cancellation rates according to different national reporting standards. An analysis of multiple markets would thus have to be limited to companies following international reporting standards or would again require data from the European insurance regulator EIOPA.

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.

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 Series, https://www.nber.org/papers/w10525.pdf.

CEA (2009). Additional input to CEA comments on Ceiops' CP49 on the mass lapse risk calibration. Comité Européen des Assurances.

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., 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.

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

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., R. Frey, and A. McNeil (2005). Quantitative risk management: Concepts, techniques and tools. Princeton: Princeton University Press.

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

Fisher, R. and L. Tibbett (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Mathematical Proceedings of the Cambridge Philosophical Society 24(2), 180–190.

Gottlieb, D. and K. Smetters (2016). Lapse-based insurance. Working Paper, https://faculty.wharton.upenn.edu/wp-content/uploads/2016/11/Insurance41.pdf.

IAIS (2016). Compiled comments on risk-based global insurance capital. International Association of Insurance Supervisors, https://www.iaisweb.org/page/supervisory-material/insurance-capital-standard/file/63187/section-6-8-lapse-risk-public.

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.

KPMG (2015). Technical practices survey. https://assets.kpmg/content/dam/kpmg/pdf/2016/04/TPS_2015.pdf.

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.

Milhaud, X., S. Loisel, and V. Maume-Deschamps (2010). Surrender triggers in life insuance: What main features affect surrender behavior in a classical economic context? Bulletin Français d'Actuariat 22, 5–48.

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.

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

Risk.net (2015). Insurers turn to reinsurers for lapse risk cover. https://www.risk.net/asset-management/insurance/2427991/insurers-turn-reinsurers-lapse-risk-cover.

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

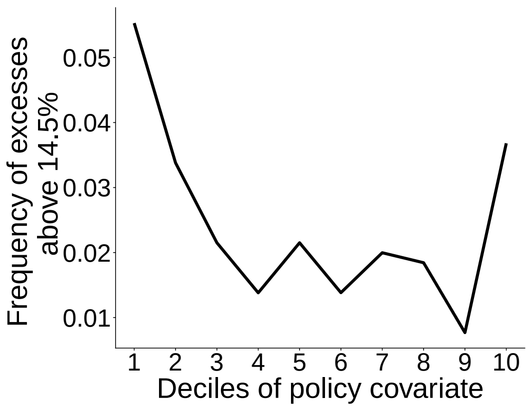
SOA and LIMRA (2012). U.S. individual life insurance persistency. Society of Actuaries and Life Insurance Marketing and Research Association, https://www.soa.org/files/research/exp-study/research-2007-2009-us-ind-life-pers-report.pdf.

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.

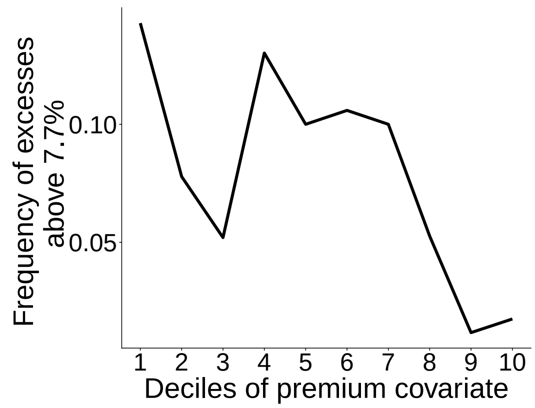
UNIQUA (2017). Group economic capital report. http://www.uniqagroup.com/gruppe/versicherung/media/files/UNIQA_ECR_Report_2017_12042018.pdf.

Appendix A Suitable For Online Publication: Additional Graphical Representations

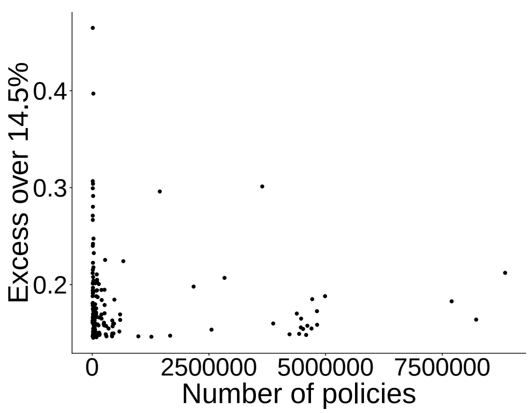
Figure 8: Excesses Over Chosen Threshold In Dependence Of The Company Size Covariate



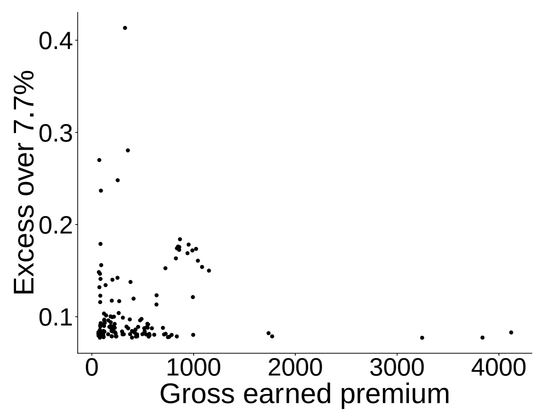
(a) U.S. data: excess probability in dependence of the number of policies



(b) German data: excess probability in dependence of the gross earned premium



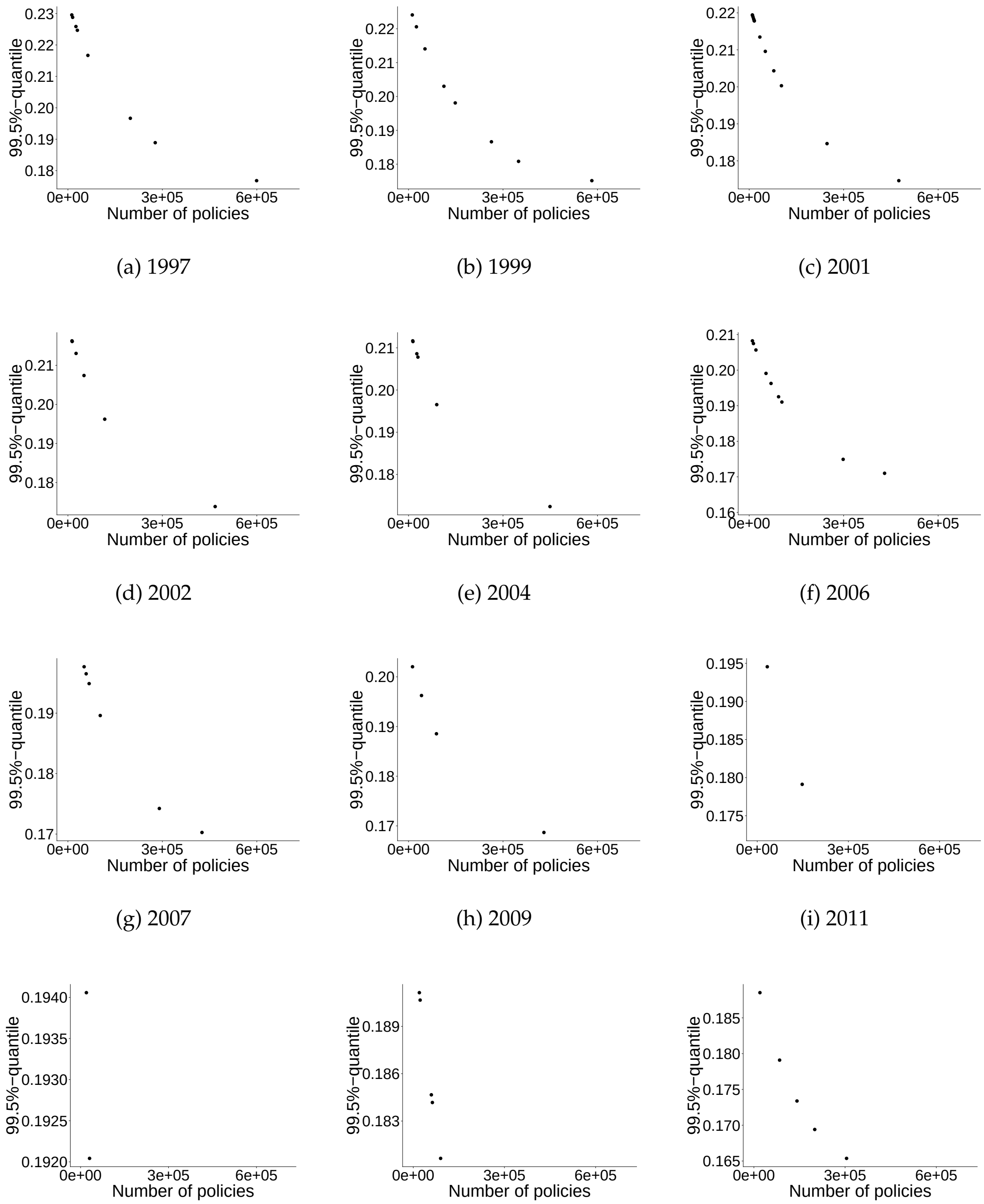
(c) U.S. data: excesses in dependence of the number of policies



(d) German data: excesses in dependence of the gross earned premium

Notes: The figure displays support for the use of the company size (either number of policies in the U.S. analysis or gross earned premium in the German analysis) as one covariate in the analysis. Panel (a) shows the frequency of excesses over the threshold in dependence of the deciles of the company size covariate in the U.S. analysis. Panel (c) shows the excesses over the threshold in dependence of the company size covariate. Panels (b) and (d) provide the same figures for the German analysis.

Figure 9: U.S. Data: 99.5%-Quantile In Dependence Of Policy Covariate Per Year



(a) 1997 (b) 1999 (c) 2001

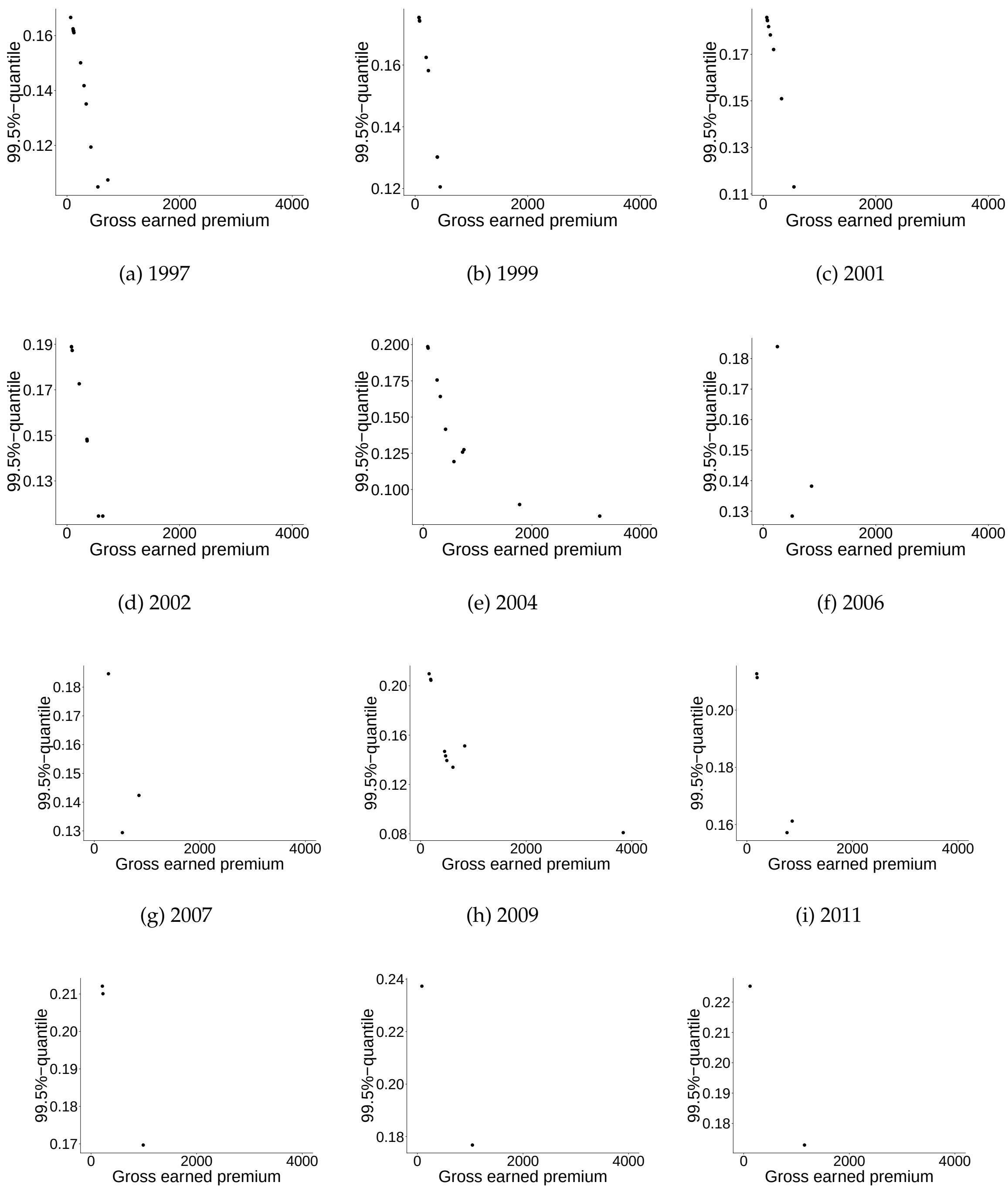
(d) 2002 (e) 2004 (f) 2006

(g) 2007 (h) 2009 (i) 2011

(j) 2013 (k) 2015 (l) 2017

Notes: The figure displays predicted 99.5%-quantiles for the cancellation rates in dependence of the company size covariate (here: number of policies) for selected years between 1996 and 2017. It shows that the estimated quantile of the cancellation rates exhibits a decreasing trend in dependence of the company size covariate.

Figure 10: German Data: 99.5%-Quantile In Dependence Of Premium Covariate Per Year



(a) 1997 (b) 1999 (c) 2001

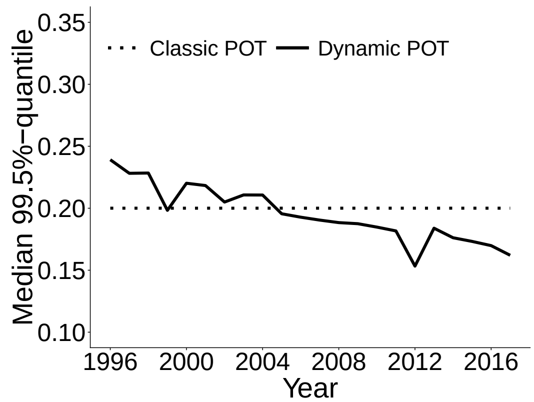
(d) 2002 (e) 2004 (f) 2006

(g) 2007 (h) 2009 (i) 2011

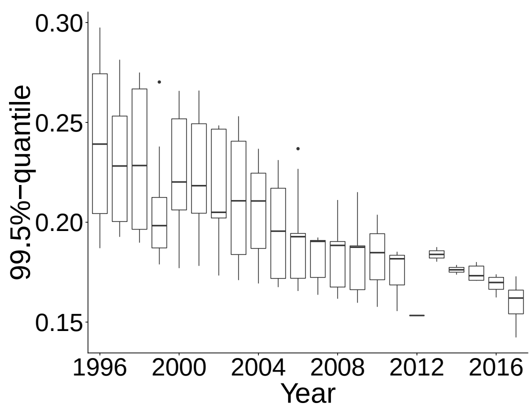
(j) 2013 (k) 2015 (l) 2017

Notes: The figure displays predicted 99.5%-quantiles for the cancellation rates in dependence of the company size covariate (here: gross earned premium) for selected years between 1996 and 2017. It shows that the estimated quantile of the cancellation rates exhibits a decreasing trend in dependence of the company size covariate.

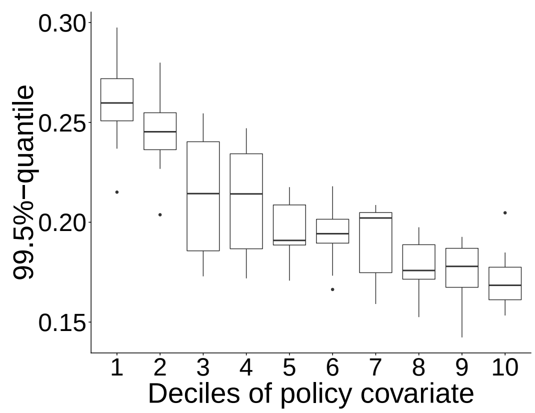
Figure 11: U.S. Data: Analysis With Logarithm Of Policy Covariate



(a) Median 99.5%-quantile of the extended dynamic POT and classic POT estimation



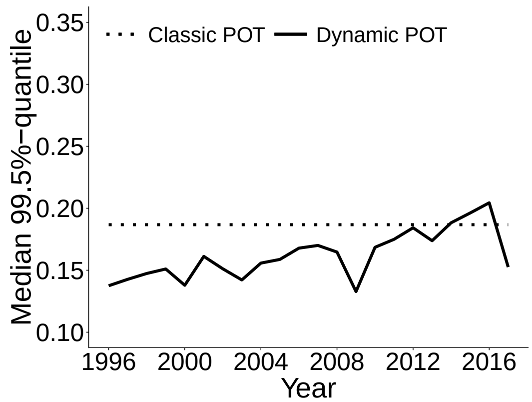
(b) Boxplot of 99.5%-quantile estimations over time



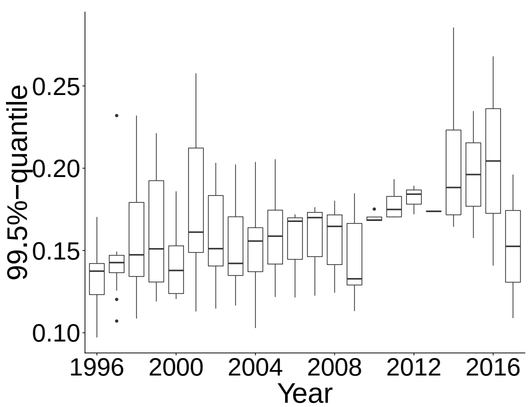
(c) Boxplot of 99.5%-quantile estimations for the deciles of the log-transformed number of policies

Notes: The figure displays estimation results for the robustness analysis with a log-transformed company size covariate. Panel (a) shows the median of predicted 99.5%-quantiles for the cancellation rates based on the classic and the extended dynamic POT method. Panel (b) provides for each year between 1996 and 2017 a boxplot of estimated 99.5%-quantiles. Panel (c) shows the boxplot of estimated 99.5%-quantiles for the deciles of the company size covariate (here: logarithm of number of policies).

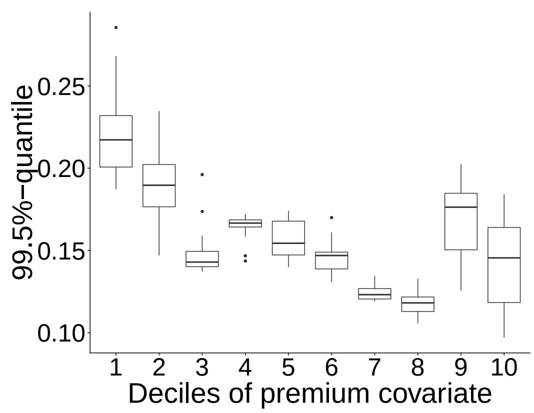
Figure 12: German Data: Analysis With Logarithm Of Premium Covariate



(a) Median 99.5%-quantile of the extended dynamic POT and classic POT estimation



(b) Boxplot of 99.5%-quantile estimations over time



(c) Boxplot of 99.5%-quantile estimations for the deciles of the log-transformed gross earned premium

Notes: The figure displays estimation results for the robustness analysis with a log-transformed company size covariate. Panel (a) shows the median of predicted 99.5%-quantiles for the cancellation rates based on the classic and the extended dynamic POT method. Panel (b) provides for each year between 1996 and 2017 a boxplot of estimated 99.5%-quantiles. Panel (c) shows the boxplot of estimated 99.5%-quantiles for the deciles of the company size covariate (here: logarithm of gross earned premium).