

# Kundenverhalten in der Lebensversicherung – automatisiertes Lernen aus Kundendaten

KI Insurance-Tagung

- Dr. Johannes Schupp
- 17.09.2024



# Einleitung

## Motivation zur Modellierung von VN-Optionen

- Projektion der künftigen Zahlungsströme auf Basis möglichst **realitätsnaher Annahmen** entscheidend für Asset-Liability-Management
- (Einseitige) Recht der Versicherungsnehmer einer Lebensversicherung, diese Zahlungsströme zu verändern, z.B. durch Storno, (Teil-)Rückkauf, Beitragsanpassungen
- relevant unter Solvency II, z.B. SCR für Stornorisiken
- Es handelt sich damit also um eine originäre **aktuarielle Aufgabe**.
- In der Praxis häufig Aufteilung der Daten nach VN-Option und einzelnen Teilbestände, z.B. je Produktart:
- meist jedoch weitere Informationen verfügbar und relevant → Verbesserung der Prognosegüte möglich

### Typischer Ansatz: granulare Teilmodelle

	Klassik	FLV	Riester	BU
■ Storno	■	■	■	■
■ Rückkauf	■	■	■	■
■ Beitragsfreistellung	■	■	■	■

### Vertragsspezifika

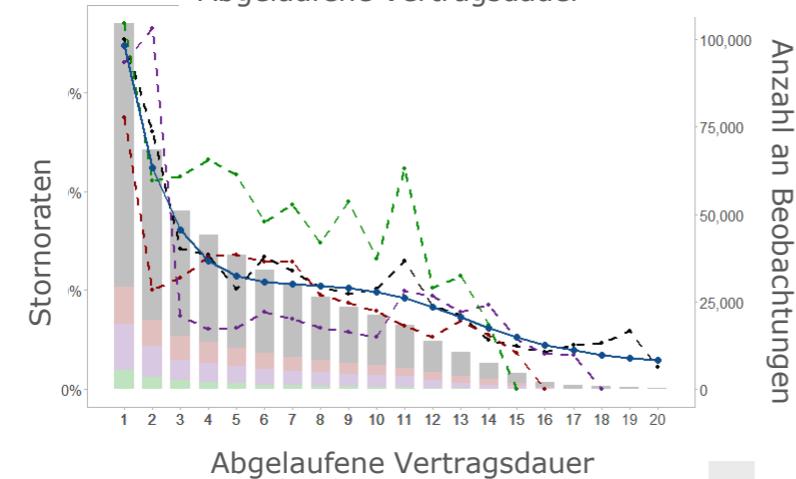
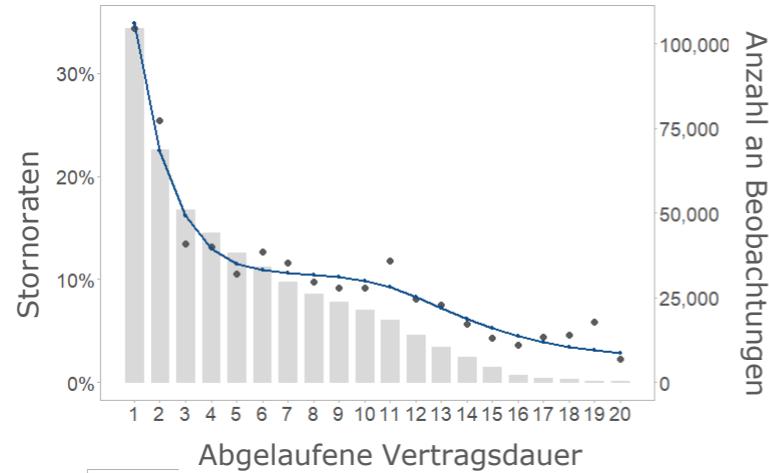
- Alter
- Geschlecht
- Vertragsdauer
- Versicherungssumme



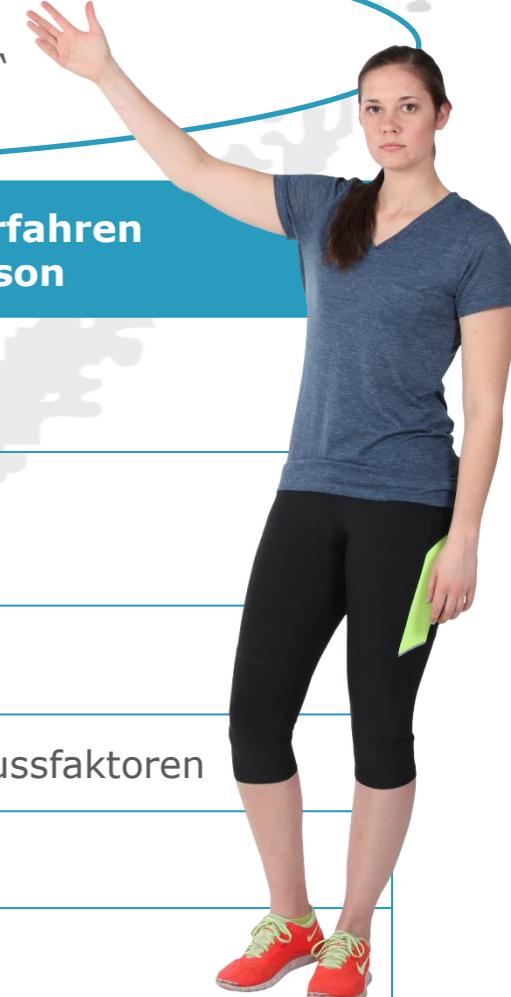
# Einleitung

## Whittaker-Henderson

Aktuarin: „Wir benötigen ein robustes, einfach interpretierbares und schnell kalibriertes Modell.“



übliches Vorgehen	univariate Glättungsverfahren z.B. Whittaker-Henderson
Verarbeitung vieler Vertragsinformationen	— sehr eingeschränkt
Ausnutzen von Gemeinsamkeiten von Teilbeständen	— nein
Aufwand für die Erstellung	+ sehr gering
Interpretierbarkeit	+ interpretierbare Einflussfaktoren
Robustheit	+ sehr robust
Prognosegüte	— schlecht



# Einleitung

## Black-Box Ansätze

Data Scientist: „Wir haben tolle neue Datenanalysetools!  
Unsere Modelle haben viel mehr Potenzial!“

KI-Ansatz	Beispielsweise neuronale Netze
Verarbeitung vieler Vertragsinformationen	+ möglich
Ausnutzen von Gemeinsamkeiten von Teilbeständen	+ ja
Aufwand für die Erstellung	- inklusive Fine-Tuning sehr hoch
Interpretierbarkeit	- nein
Robustheit	- wenig robust
Prognosegüte	+ in der Regel gut



# Einleitung

## Lasso basierte Modelle

Data Scientist: „Lass uns (gemeinsam) ein Modell entwickeln mit dem Besten aus beiden Perspektiven!“

übliches Vorgehen	Univariate Glättungsverfahren, z.B. Whittaker-Henderson
Verarbeitung vieler Vertragsinformationen	– sehr eingeschränkt
Ausnutzen von Gemeinsamkeiten von Teilbeständen	– nein
Aufwand für die Erstellung	+ sehr gering
Interpretierbarkeit	+ interpretierbare Einflussfaktoren
Robustheit	+ sehr robust
Prognosegüte	– schlecht

## Lasso-Ansatz

- Verarbeitung vieler Vertragsinformationen +
- Ausnutzen von Gemeinsamkeiten von Teilbeständen +
- Aufwand für die Erstellung +
- Interpretierbarkeit +
- Robustheit +
- Prognosegüte +

### Key-Features des Lasso

- gleichzeitige Modellkalibrierung und Variablenselektion
- multivariat, performant, datengetrieben, automatisiert und mit guter Prognosegüte
- erkennt vielfältige Strukturen innerhalb der Kovariablen



Beispielsweise neuronale Netze	
+ möglich	
+ ja	
– inklusive Fine-Tuning sehr hoch	
– nein	
– wenig robust	
+ in der Regel gut	

# Methodik und Ergebnisse

Grundidee des Lasso: Erweiterung des GLM

- GLM: gemeinsame Betrachtung mehrerer Merkmale ( $X = X_1, \dots, X_m$ ) um die Abhängigkeit einer Zielgröße ( $Y$ ) von **interpretierbaren** Einflussgrößen ( $\beta$ ) zu modellieren:

$$E(Y|X_1, \dots, X_m) = g^{-1}(X\beta)$$

- Innerhalb jedes Merkmals (z.B.  $X_1 = \text{Alter}$ ) können die Ausprägungen über weitere Strukturen modelliert werden.
- Beim Lasso werden zusätzliche Regularisierungsterme bei der Schätzung ergänzt:

$$\operatorname{argmin}_{\beta} \quad -\log(L(\beta, x, y)) + \lambda \sum_{j=1}^m g(\beta_j)$$

**Regularisierungsfaktor:**  $\lambda \geq 0$

Steuerung zwischen Bestrafung und Anpassungsgüte

**Bestrafungsart:**

Bestrafung der Parameter z.B.:  
 $g(\beta_j) = \sum_{i=1}^{p_j} |\beta_{i,j}|$

- Durch geschickte Wahl von  $g$  können Strukturen und Muster innerhalb von Merkmalen automatisiert erkannt werden (Designentscheidung).
- Ergebnisse und mögliche Designentscheidungen werden anhand eines Datensatzes eines Pan-Europäischen Lebensversicherers veranschaulicht.

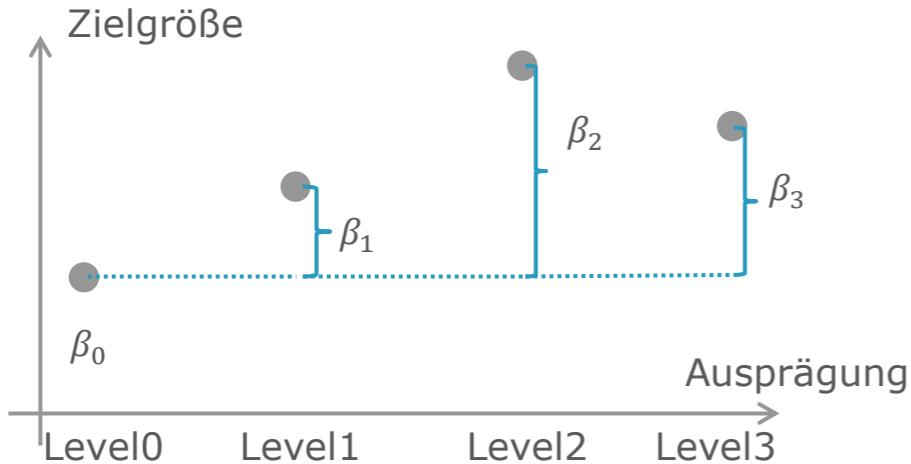
# Methodik und Ergebnisse

## Regular Lasso für nominale Merkmale

### Regular Lasso

### Methodik

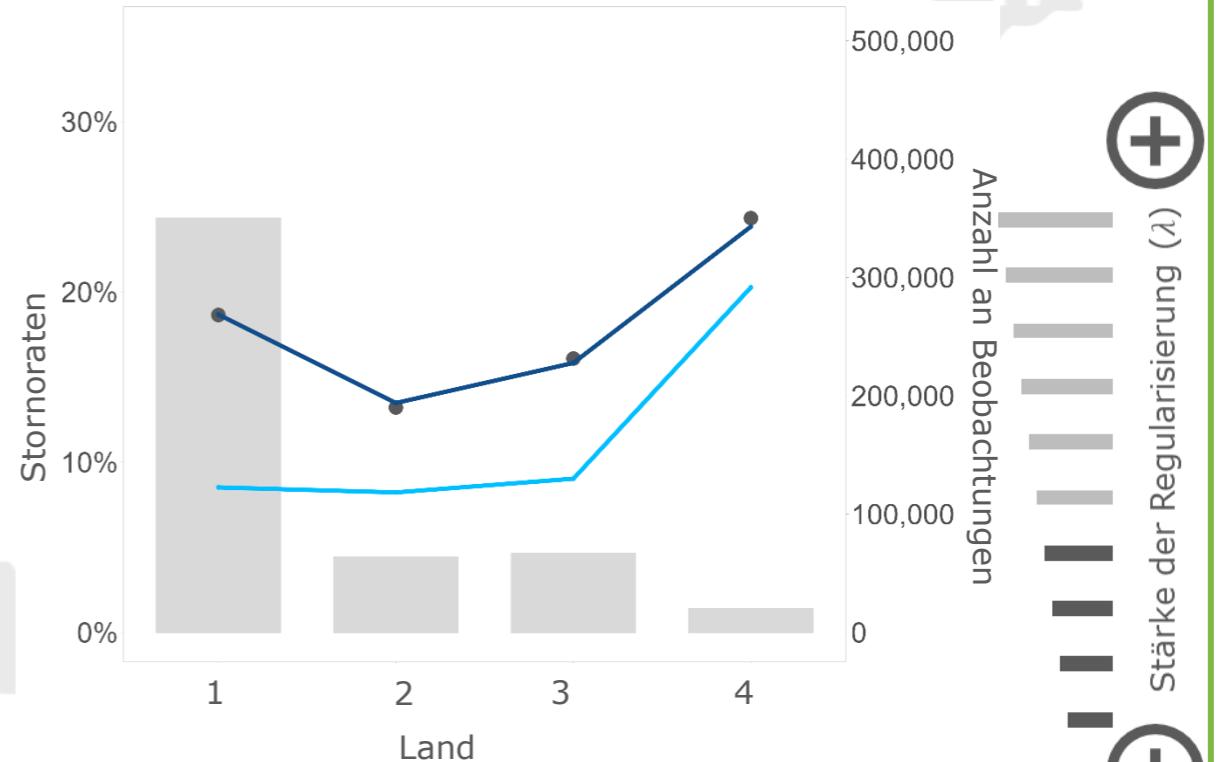
$$g_R(\beta_j) = \sum_{i=1}^{p_j} |\beta_{i,j}|$$



### Mit dieser Ausgestaltung beantwortbare Fragen:

### Ergebnisse

- Welche Ausprägungen können über den Intercept modelliert werden und welche nicht?



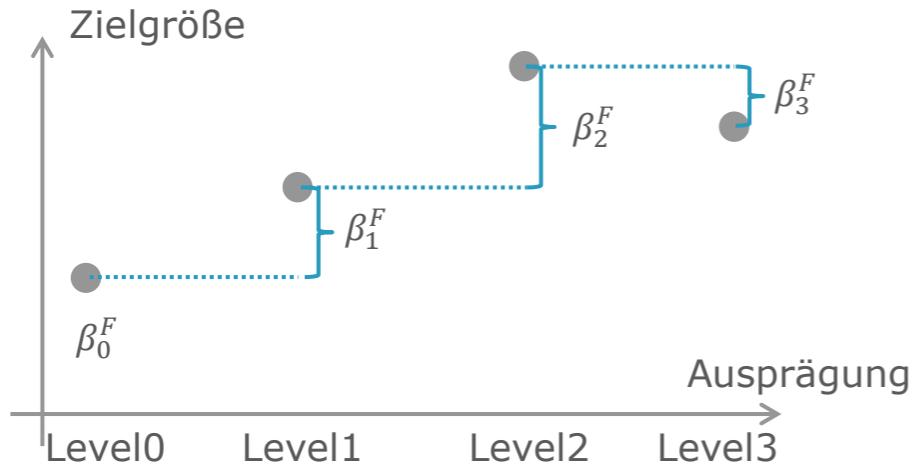
# Methodik und Ergebnisse

## Fused Lasso für ordinale Merkmale

### Fused Lasso

### Methodik

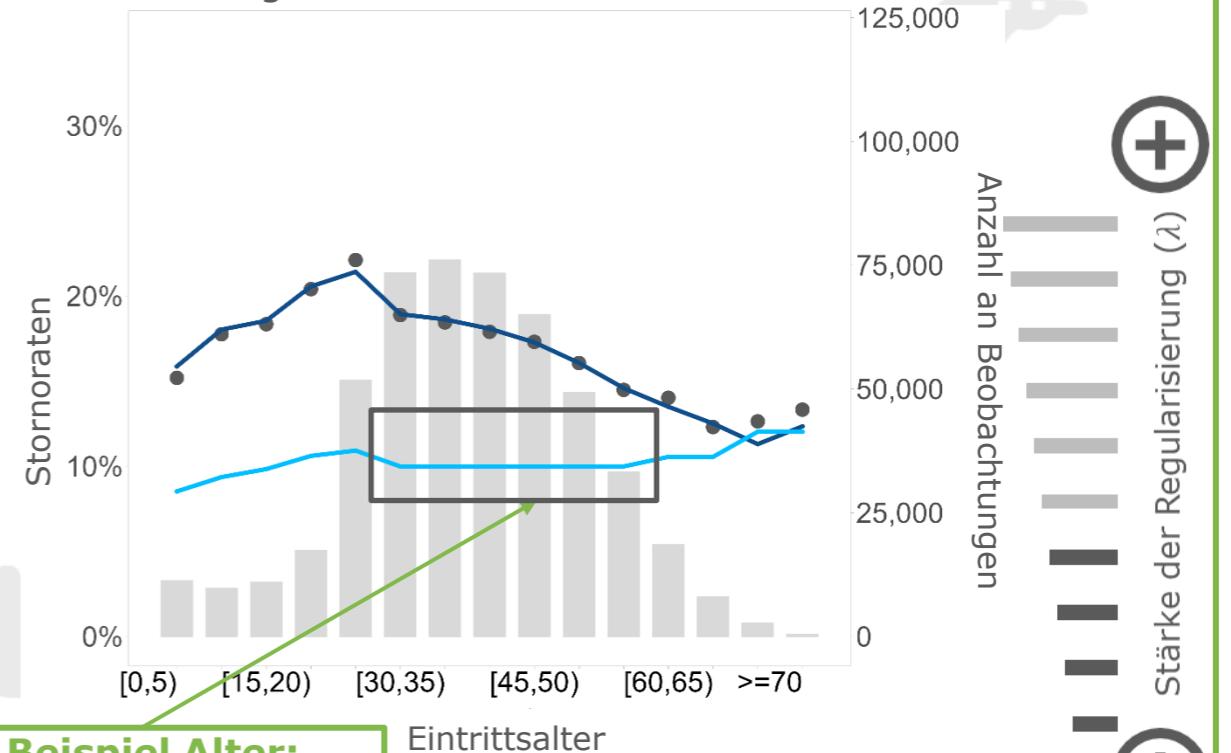
$$g_F(\beta_j) = |\beta_{j,1}| + \sum_{i=2}^{p_j} |\beta_{j,i} - \beta_{j,i-1}| =: \sum_{i=1}^{p_j} |\beta_{j,i}^F|$$



### Mit dieser Ausgestaltung beantwortbare Fragen:

### Ergebnisse

- Welche benachbarten Ausprägungen können zusammengefasst werden?



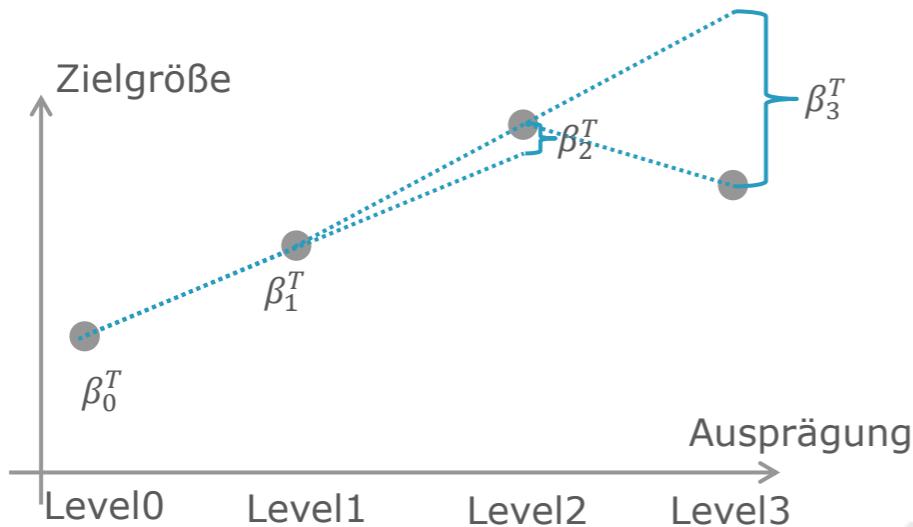
# Methodik und Ergebnisse

## Trend filtering für ordinale Merkmale

### Trend filtering

### Methodik

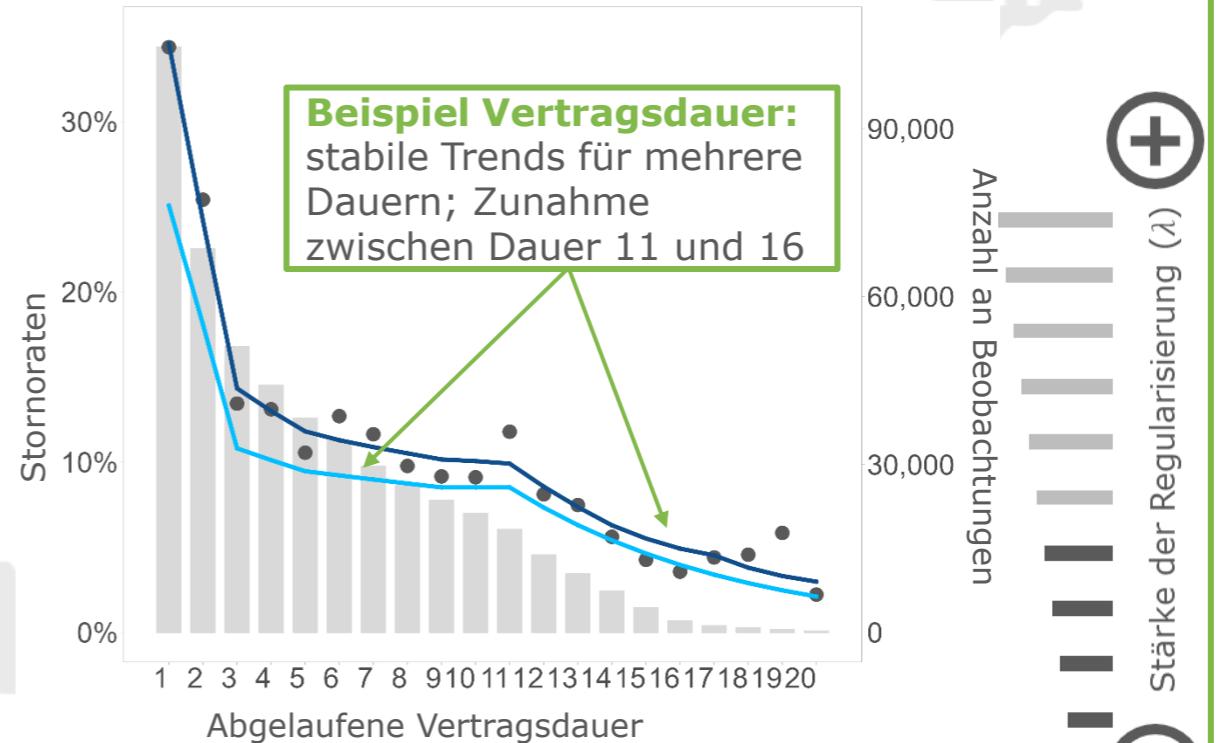
$$g_T(\beta_j) = |\beta_{j,1}| + |\beta_{j,2} - 2\beta_{j,1}| + \sum_{i=3}^{p_j} |\beta_{j,i} - 2\beta_{j,i-1} + \beta_{j,i-2}| =: \sum_{i=1}^{p_j} |\beta_{j,i}^T|$$



### Mit dieser Ausgestaltung beantwortbare Fragen:

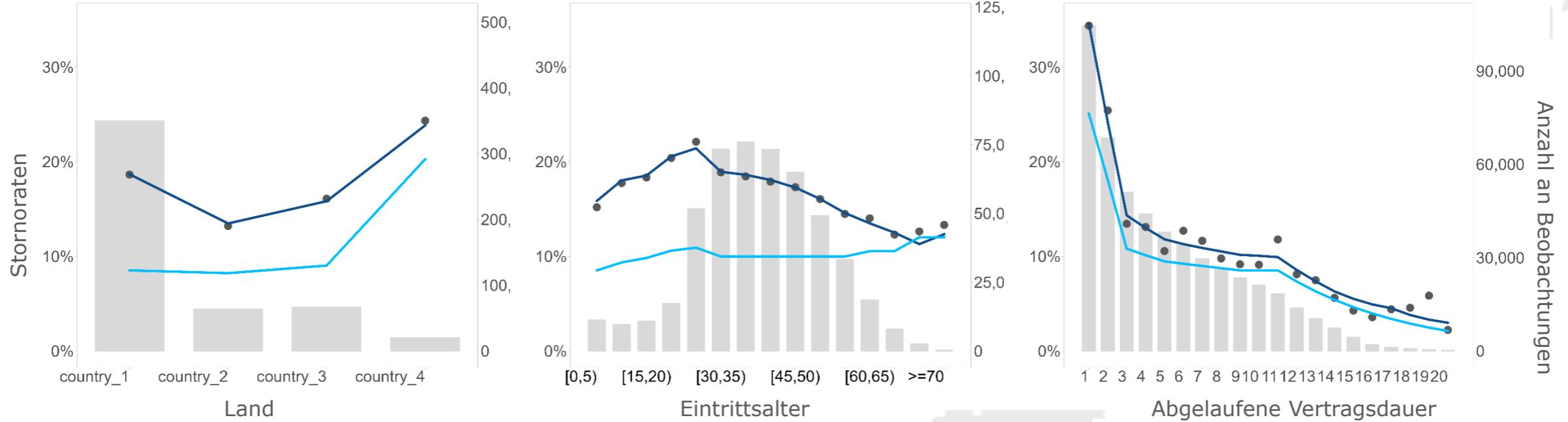
### Ergebnisse

- Welche benachbarten Ausprägungen können über einen gemeinsamen Trend modelliert werden?



# Methodik und Ergebnisse

## EIN Hyperparameter zur Steuerung der Anpassung



— Stärke der Regularisierung ( $\lambda$ ) +

Quelle: Reck, L., Schupp, J., & Reuß, A. (2023). Identifying the determinants of lapse rates in life insurance: an automated Lasso approach. *European Actuarial Journal*, 13(2), 541-569.

# Methodik und Ergebnisse

## weitere Designentscheidungen und Fallstricke

### Wahl der Kovariablen

- Welche Informationen liegen in welcher Qualität vor?
- Welche weiteren (ggf. externen) Informationen können/sollen ergänzt werden?

### Ausreißer beim Randeffect

- große Trendänderungen
- Sprünge

### Binning

- Welche Ausprägungen können/sollen vorab zusammengefasst werden?
- Wie wählt man sinnvolle Bins?

### Bestrafungsterm für Interaktionen

- Kovariable nun 2-dimensional
- Was zählt als „benachbarte“ Ausprägung?

### Wahl des Hyperparameters $\lambda$

- statistisch motiviert mit Kreuzvalidierung
- anwendungsorientiert
- „Screening vs. Selection“

Startposition beim trend filtering,

...

Quelle: Reck, L. (2024). The Automation of Core Actuarial Modelling Tasks - an Analysis and Evaluation of the Lasso. Working Paper, Ulm University.



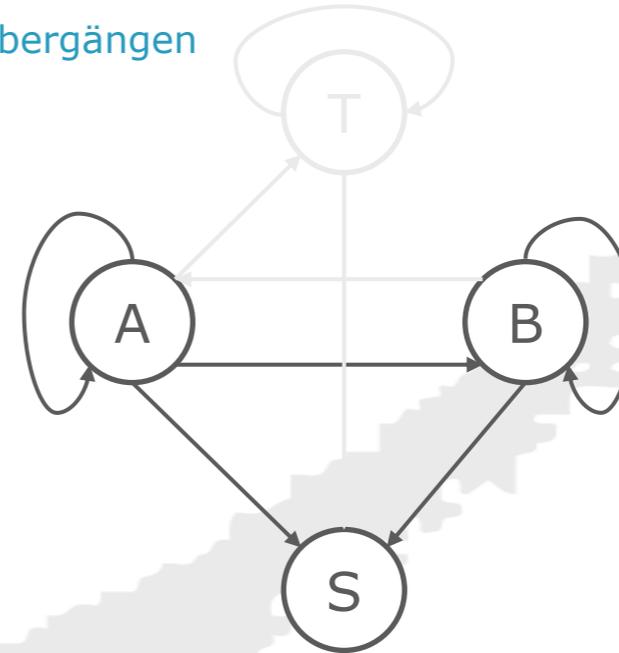
**Das Aufsetzen eines verlässlichen Lasso-Modells geht nicht auf „Knopfdruck“ – um das volle Potenzial von Lasso zu nutzen muss man sowohl die Details des Modellierungsansatzes als auch den Anwendungskontext genau kennen.**

# Mehrere VN-Optionen

## Motivation

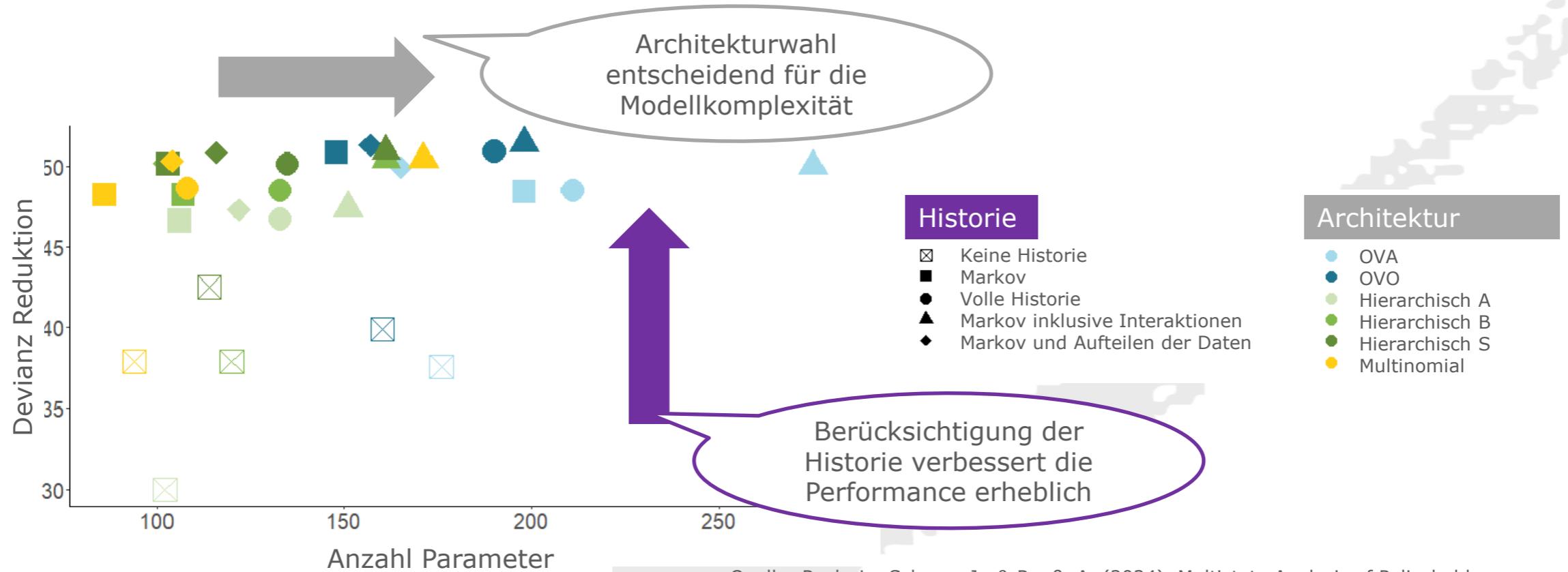
In der Praxis: Ausübung **mehrerer VN-Optionen** mit entsprechenden **Zustandsübergängen**

- Welche Zustände und welche Übergänge sind relevant (materiell)?
- Beispielsweise neben „Storno“ auch Beitragsfreistellung
  - potentiell aber auch weitere Zustände wie Teilrückkauf
  - oder weitere Übergänge wie Wiederinkraftsetzung
- hier müssen weitere Designentscheidungen berücksichtigt werden
  - Varianten der Modellarchitektur
    - Zerlegung in mehrere binäre Modelle
      - Aggregation nicht trivial!
    - Multinomial
      - Implementierungen sind hier nur noch begrenzt vorhanden
      - Glättung über Merkmale hinweg geht; Glättung über Ausprägungen der Zielgröße geht bisher nicht
  - Berücksichtigung der Historie
    - keine Historie / aktueller Zustand / volle Historie



# Mehrere VN-Optionen

## Modellierung und Ergebnisse



Quelle: Reck, L., Schupp, J., & Reuß, A. (2024). Multistate Analysis of Policyholder Behaviour in Life Insurance-Lasso based Modelling Approaches. Working Paper, Ulm University.

# Fazit

## Das Lasso zur Automatisierung aktueller Analysen



multivariater, performanter, automatisierter Ansatz  
vielfältige, **interpretierbare** Strukturen innerhalb der Kovariablen identifizierbar  
große Bandbreite an Verteilungen der Zielgröße möglich (z.B. multinomial)



Designentscheidungen („Customizing“) heben letztlich das Potenzial des Modells  
Wie bei jedem Modell ist die Kenntnis der Stärken UND Schwächen des Modells wichtig!



Monitoring biometrischer Annahmen (Tod, Invalidisierung, Reaktivierung, ...) → Identifikation auffälliger Segmente  
Prognosemodelle im Leistungsbearbeitungsprozess (z.B. Anerkennungswahrscheinlichkeiten in der BU)  
VN-Optionen in anderen Sparten (z.B. Stornotreppen in KFZ)



Bei vergleichbaren Ausgangslagen ist Lasso immer eine Methode, die man berücksichtigen sollte, d.h.  
- bei mehreren Kovariablen, mit unterschiedlichen (oder unbekannt) Strukturen  
- wenn die Interpretierbarkeit des Modells wichtig ist

# Kontakt

**Dr. Johannes Schupp**

+49 731 20644-241

[j.schupp@ifa-ulm.de](mailto:j.schupp@ifa-ulm.de)



# Beratungsangebot

