

Modellierung des Stornos nach Beitragsanpassung in der PKV

02.07.2013

Alexander Küpper – Central Krankenversicherung AG

Inhaltsverzeichnis

Einführung

Gesetzlicher Rahmen

Stornomodell

Anwendung

Ausblick und Weiterentwicklung

Einführung – Kennzahlen der Privaten Krankenversicherung in Deutschland

Gesetzlicher Rahmen

Stornomodell

Anwendung

Ausblick und Weiterentwicklung

PKV ist wichtige Stütze des sozialen Gerüsts in Deutschland

Kennzahlen der privaten Krankenversicherung

Private Krankenversicherung¹

Versicherungsunternehmen (Mitglieder des PKV-Verbands)	43	(davon 24 AG, 19 VVaG)
Vollversicherte Personen	8,98 Mio.	(zusätzlich 22,5 Mio. Zusatzversicherungen)
Beitragseinnahmen	34,67 Mrd. Euro	
Ausgezahlte Leistungen	22,77 Mrd. Euro	
Alterungsrückstellung	169,43 Mrd. Euro	

Central Krankenversicherung AG²

Vollversicherte Personen	418 Tsd.	(zusätzlich 1,4 Mio. Zusatzversicherungen)
Beitragseinnahmen	2,2 Mrd. Euro	

1 Quelle: PKV-Verband, Geschäftsjahr 2011

2 Quelle: Geschäftsbericht 2012 Central Krankenversicherung AG

Einführung

Gesetzlicher Rahmen – Wann und wie darf/muss der Beitrag angepasst werden?
– Welche Reaktionsmöglichkeiten hat der VN?

Stornomodell

Anwendung

Ausblick und Weiterentwicklung

Spielregeln zur Beitragsanpassung sind stark durch gesetzliche Vorschriften reguliert

Gesetzliche Rahmenbedingungen zur Beitragsanpassung

§12b (2) VAG

Das Versicherungsunternehmen hat für jeden nach Art der Lebensversicherung kalkulierten Tarif zumindest jährlich die erforderlichen mit den kalkulierten Versicherungsleistungen zu vergleichen. Ergibt die der Aufsichtsbehörde und dem Treuhänder vorzulegende Gegenüberstellung für einen Tarif eine Abweichung von mehr als 10 vom Hundert, sofern nicht in den allgemeinen Versicherungsbedingungen ein geringerer Vomhundertsatz vorgesehen ist, hat das Unternehmen alle Prämien dieses Tarifs zu überprüfen und, wenn die Abweichung als nicht nur vorübergehend anzusehen ist, mit Zustimmung des Treuhänders anzupassen.

Prämie des Vertrages erhöht sich

§205 (4) VVG

Erhöht der Versicherer auf Grund einer Anpassungsklausel die Prämie oder vermindert er die Leistung, kann der Versicherungsnehmer hinsichtlich der betroffenen versicherten Person innerhalb von zwei Monaten nach Zugang der Änderungsmitteilung mit Wirkung für den Zeitpunkt kündigen, zu dem die Prämienhöhung oder die Leistungsminderung wirksam werden soll.

Mehrere Optionen für Kunden

Keine Reaktion

Storno

Tarifwechsel

Inhaltsverzeichnis

Einführung

Gesetzlicher Rahmen

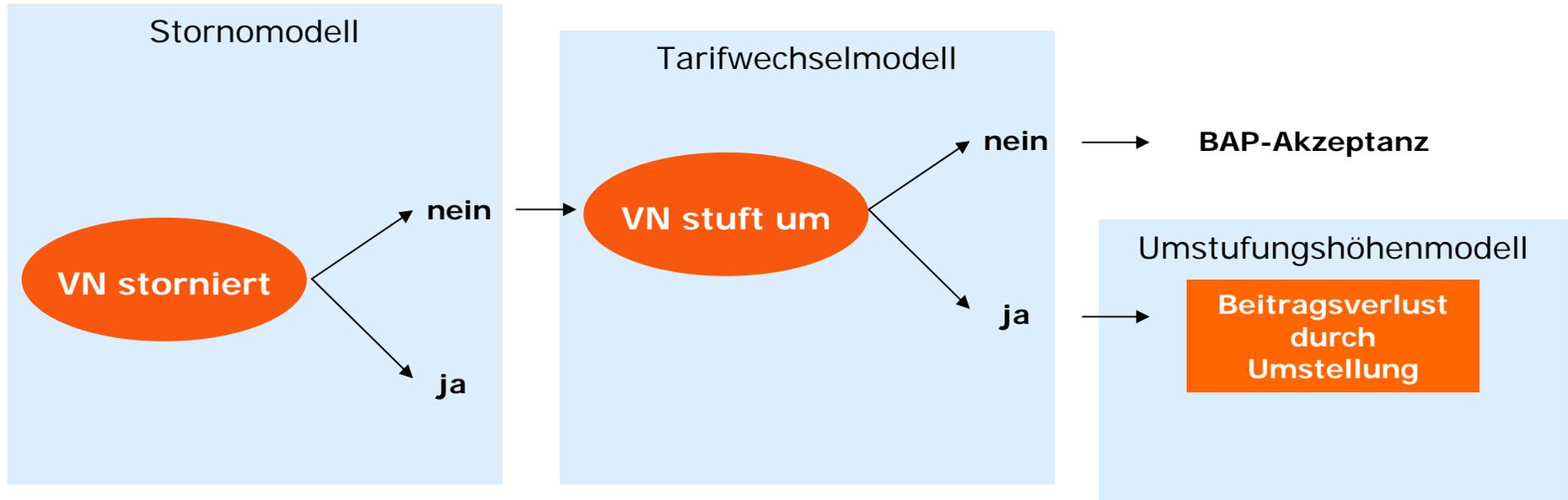
Stornomodell - Modellierung mittels eines verallgemeinerten linearen Modells

Anwendung

Ausblick und Weiterentwicklung

VN-Verhalten im Zusammenhang einer Beitragsanpassung kann als iterative Ausführung zweier 0/1-Entscheidungen modelliert werden

Modellierungsansatz



- Kundenverhalten als BAP-Reaktion wird durch iterative Ausführung zweier 0/1-Entscheidungen modelliert:
 - Im ersten Schritt entscheidet Kunde, ob er storniert oder nicht
 - Falls er nicht storniert, entscheidet Kunde im zweiten Schritt, ob er die Anpassung akzeptiert oder eine Umstufung vornimmt
 - Im Falle einer erfolgten Umstufung schätzt ein weiteres Modell den Beitragsverlust durch Umstellung

Definition Zielgrößen und Identifikation beschreibender Variablen

wesentliche Voraussetzung für Güte des Stornomodells

Datengrundlage

Bestandsdaten

Vollversicherte Erwachsene
vor den letzten drei
Beitragsanpassungen

Merkmale:

- Alter des VN
- Vertragsdauer
- Beitrag
- Geschlecht
- Vertragsumfang (KT, KHT, BE)
- mitversicherte Kinder
- SB-Höhe
- Tariflinie

Anpassungsdaten

Informationen zu den jeweils
kommenden
Beitragsanpassungen

Merkmale:

- Beitragserhöhung (absolut und relativ)
- Leistungsanpassung (SB-Erhöhung)

Weitere Daten

Informationen, die aus weiteren
Datenquellen (intern und extern)
stammen

Merkmale:

- individuelle Schadenerfahrung
- Zahlungsverhalten
- Kundenwert

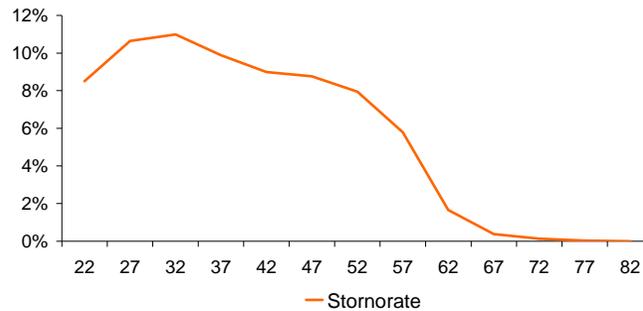
Zielgrößen

- Stornodaten nach Beitragsanpassung (0/1-Kennzeichen, 1=storniert, 0=nicht storniert)
- Bei Tarifwechsel:
 - Tarifwechselkennzeichen (0/1-Kennzeichen, 1=Tarifwechsel, 0=kein Tarifwechsel)
 - Beitragsveränderung nach Tarifwechsel

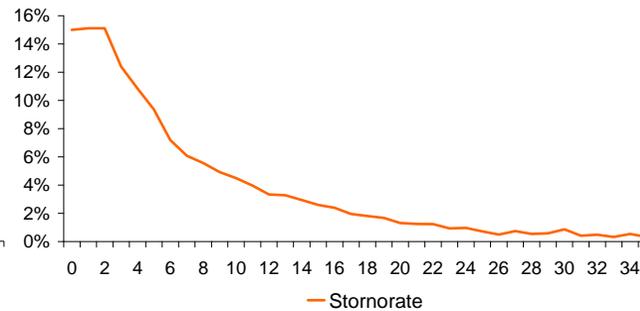
Analysen der letzten Beitragsanpassungen lassen Abhängigkeit zwischen Storno und bestimmten Merkmalen erkennen

Abhängigkeiten

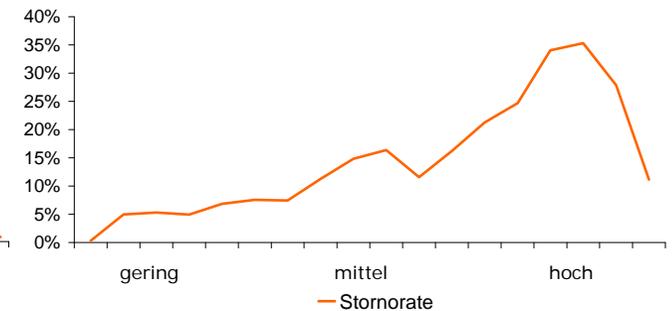
Stornorate nach Alter



Stornorate nach Dauer



Stornorate nach Anpassung (rel.)



- Univariate Analysen lassen Abhängigkeiten zwischen Bestandsmerkmalen und Stornoverhalten erkennen.
- Parameter „Alter“, „Dauer“ und „BAP-Anpassung“ zeigen größte Spreizungskraft auf.
- Problem: Wie kann man alle drei Effekte gleichzeitig (multivariat) berücksichtigen?
 - ➔ Verwendung von Verallgemeinerten Linearen Modellen (GLM)

Ausgangspunkt für eine multivariate Analyse ist die klassische lineare Regression

Theorie des linearen Modells

Klassisches lineares Modell:

$$Y_j = \sum_i \beta_i \cdot X_{ij} + \varepsilon_j$$

Zufallsgröße Y_j ist die Summe von Regressionsparameter β_i (multipliziert mit Designvektor X_{ij}) und Fehlerterm ε_j .

Wesentliche Annahmen:

- Fehlerterm ε_j ist Zufallsgröße mit $\varepsilon_j \sim N(0, \sigma^2)$
- Damit Y_1, Y_2, \dots, Y_n unabhängige normalverteilte Zufallsgrößen mit
 - $E[Y_j] = p_j = \sum_i \beta_i \cdot X_{ij}$
 - $Var(Y_j) = Var(\varepsilon_j) = \sigma^2$
- Die Schätzung der Parameter p_j erfolgt mittels kleinste-Quadrate-Methode, das entspricht unter Normalverteilungsannahme einer **Maximum-Likelihood-Schätzung**.

Probleme bei Übertragung auf Stornomodellierung:

- Die zu beschreibende Zielgröße (Storno ja/nein) ist nicht normalverteilt, sondern binomialverteilt.
- Der Schätzer für die Stornowahrscheinlichkeit, der zwischen 0 und 1 liegt, lässt sich nicht direkt durch eine lineare Gleichung beschreiben

GLM-Theorie¹ verallgemeinert Idee der linearen multivariaten Regression auf binomial verteilte Zufallsgrößen

GLM-Theorie für binomialverteilte Größen

Verallgemeinerungen der klassischen Theorie:

- **1. Verallgemeinerung:** Y_1, Y_2, \dots, Y_n unabhängige $\text{Bin}(1, p_j)$ -verteilte Zufallszahlen
- **2. Verallgemeinerung:** $E[Y_j] = p_j = g^{-1}\left(\sum_i \beta_i \cdot X_{ij}\right)$ mit Linkfunktion: $g(x) = \ln\left(\frac{x}{1-x}\right)$
- Die Schätzung der Parameter p_j erfolgt nach GLM-Theorie analog zum klassischen linearen Modell mittels **Maximum-Likelihood-Schätzung**. Die Regressionsparameter β_j lassen sich durch Statistik-Software numerisch bestimmen.

Anwendung auf das Stornomodell:

- Y_j stellen die binomialverteilten Stornoentscheidungen dar. $Y_j=1$ bedeutet, der Kunde storniert, $Y_j=0$ bedeutet, dass er bleibt. Die jeweiligen Erwartungswerte bezeichnen die zu bestimmenden Stornowahrscheinlichkeiten.
- Über die Linkfunktion wird das beliebig reelle Ergebnis des linearen Ausdrucks $\sum \beta_i \cdot X_{ij}$ auf das Intervall $[0,1]$ transformiert. Somit kann das Ergebnis als eine Wahrscheinlichkeit interpretiert werden.

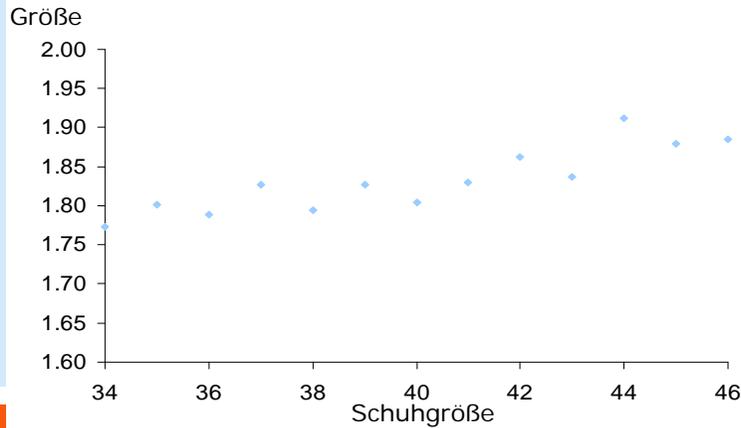
1: J. A. Nelder and R. W. M. Wedderburn: Generalized Linear Models, Journal of the Royal Statistical Society, Vol. 135, No. 3 (1972)

Wahl von Verteilungsannahme und Link-Funktion hängt von jeweiliger Anwendung ab

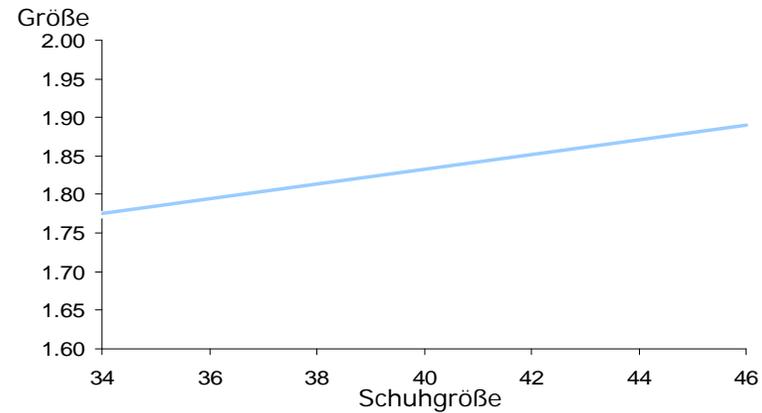
Vergleich: lineares Modell mit Binomial-GLM

lineares Modell

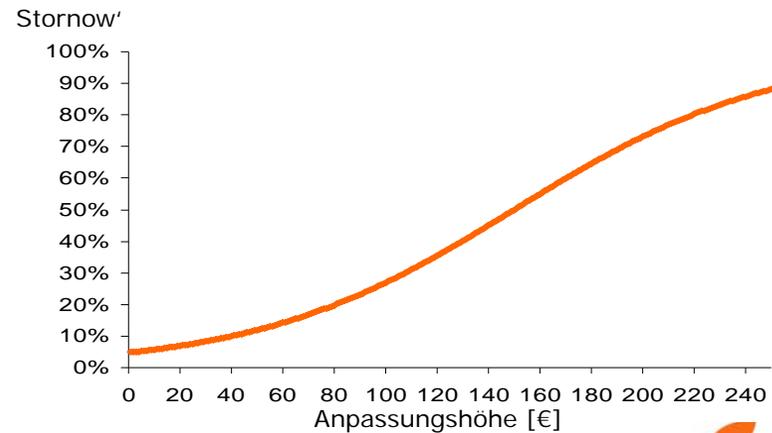
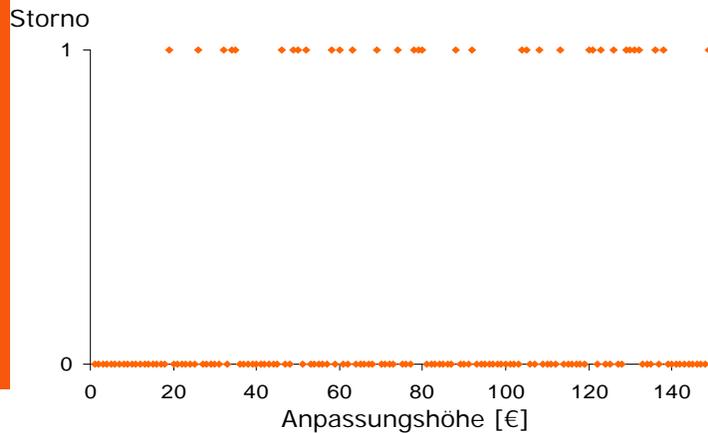
Beobachtungen



Modellschätzer



Binomial-GLM



Güte der Ergebnisse mittels Devianz und Wald-Test bewertbar

Gängige Teststatistiken zur Bewertung der Modellgüte

Test-Statistiken

- **Wald-Test**

- Der Wald-Test basiert auf der Hypothese $\beta_j = 0$ für den Regressor der j-ten unabhängigen Variablen. Es wird also angenommen, der Parameter hätte keinen signifikanten Einfluss auf das Gesamtmodell.
- Die zugehörige Teststatistik stellt den log-likelihood des Modells **ohne** den Parameter i gegen den log-likelihood des Modells **mit** Parameter i. Diese Teststatistik ist asymptotisch χ^2 -verteilt.
- Hypothese wird verworfen, falls Teststatistik zu groß ist; d.h. Parameter liefert signifikanten Einfluss, je größer Teststatistik ausfällt.

➔ **Der Wald-Test wird benutzt zur Bewertung der Signifikanz einzelner Parameter.**

- **Devianz:**

- Die Devianz misst den Unterschied zwischen log-likelihood des Modells und des saturierten Modells.
- Je geringer der Unterschied, desto signifikanter ist das Gesamtmodell

➔ **Die Devianz wird benutzt zur Bewertung der Anpassungsgüte des Gesamtmodells.**

Einführung

Gesetzlicher Rahmen

Stornomodell

Anwendung – Ein Stornomodell für die PKV

- Erkenntnisse aus dem Storno-GLM

Ausblick und Weiterentwicklung

Praktische Verwendung von beschreibenden Variablen

Dummy-Variablen, Interaktionen, Polynome

Dummy-Variablen

- werden verwendet bei kategorialen oder gruppierten Merkmalen
- jede Ausprägung bekommt einen separaten Regressor
- Beispiele:
 - Tariflinie (Top, Plus, Eco)
 - Geschlecht (männl., weibl.)
 - BE-Komponente vorhanden
 - KT oder KHT vorhanden
 - Kunde ist Nichtzahler (j/n)
 - Anzahl Kinder

Tariflinie:

Ausprägungen: Top, Plus, Eco

$$\beta^{TI} = (\beta_{Top}^{TI}, \beta_{Plus}^{TI}, \beta_{Eco}^{TI})$$

Interaktionen

- werden verwendet, wenn sich zwei oder mehrere beschreibende Variablen nicht unabhängig verhalten, sondern interagieren
- jede Ausprägung (des Kreuzproduktes) bekommt einen separaten Regressor
- Beispiele:
 - Geschlecht (m, w) x Kunde ist Nichtzahler (j/n)

Geschlecht x Nichtzahler:

Ausprägungen: (m,j), (w,j), (m,n), (w,n)

$$\beta^{PG_NiZa} = (\beta_{m,j}^{PG_NiZa}, \beta_{w,j}^{PG_NiZa}, \beta_{m,n}^{PG_NiZa}, \beta_{w,n}^{PG_NiZa})$$

Polynome

- werden verwendet bei metrischen Merkmalen
- der Grad des Polynoms gibt die Anzahl an Regressoren vor
- Beispiele:
 - Alter VN
 - Dauer
 - Beitragshöhe
 - relative Anpassung
 - SB-Höhe

Alter VN:

Ausprägungen: Alter in 5-er Schritten

$$\beta^{Alter} = (\beta_1^{Alter}, \beta_2^{Alter})$$

32-jähriger männl. Plus-Zahler: $E(Y) = g^{-1}(\beta_{Plus}^{TG} + \beta_{m,n}^{PG_NiZa} + 32\beta_1^{Alter} + 32^2\beta_2^{Alter})$

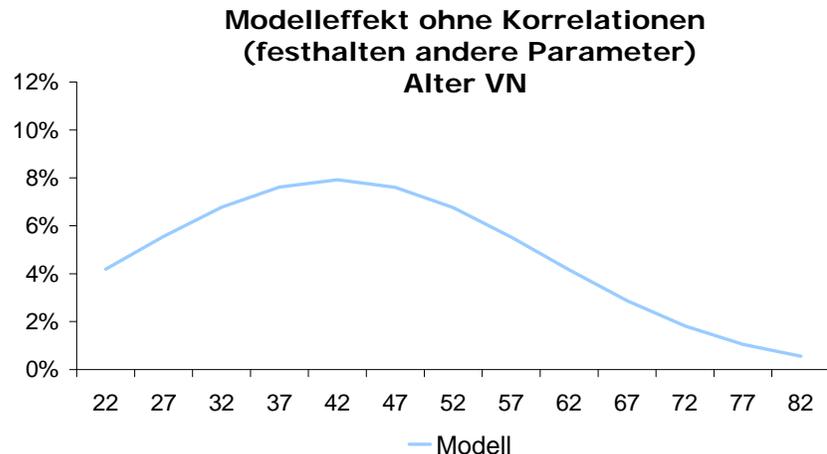
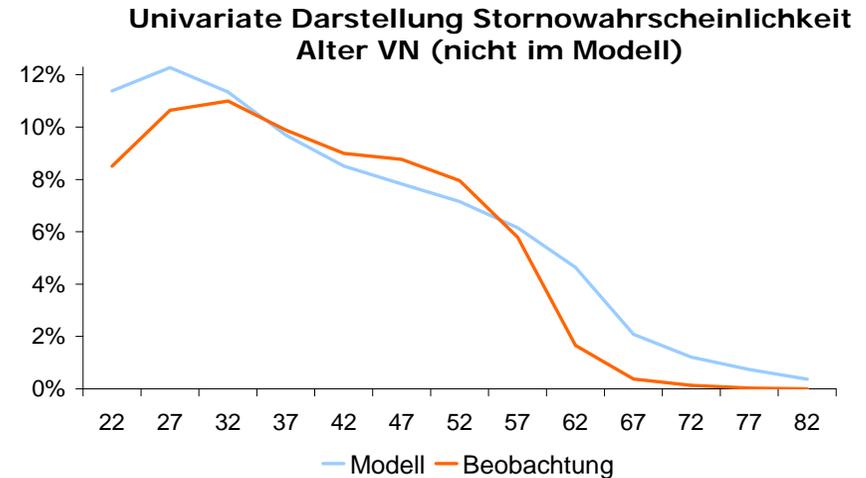
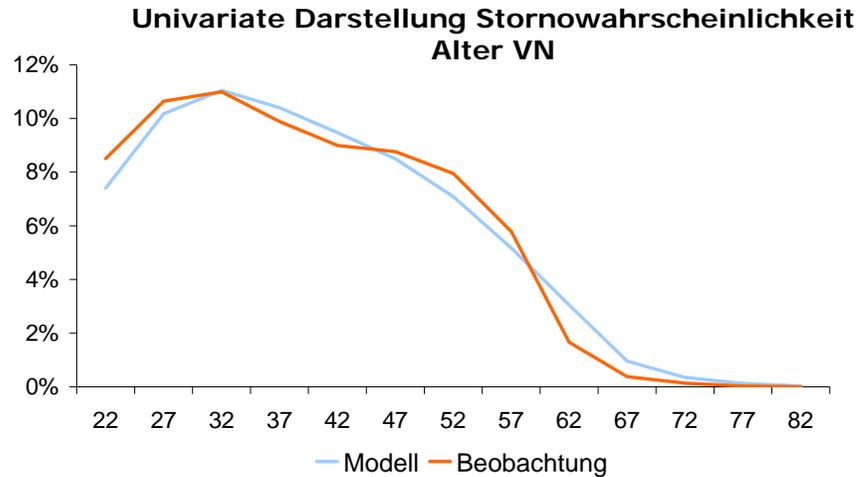
Vorhandene Merkmale zeigen signifikanten Einfluss auf Stornoverhalten

Ergebnisse aus dem GLM

Merkmalsname	Ausprägung	modelliert	Richtung	Signifikanz
Alter VN	22-82; in 5-Jahres-Schritten	Parabel	~	hoch
Dauer	0-30; in Jahres-Schritten	linear	↘	sehr hoch
Anpassung rel. x Tariflinie	in % pro Tariflinie	Interaktion	↗	sehr hoch
Leistungen im GJ	0, bis 1000€, bis 5000€, über 5000€	Dummy	↘	mittel
Monatssollbeitrag	in €	log-linear	↗	hoch
KHT x Tariflinie	ja/nein pro Tariflinie	Interaktion	↘	schwach
BE vorhanden	ja/nein	Dummy	↘	mittel
Anzahl Kinder im Vertrag	0,1,2,3	Dummy	↘	mittel
Pauschalleistungstarif	ja/nein	Dummy	↘	schwach
KT x Tariflinie	ja/nein pro Tariflinie	Interaktion	↗	schwach

Alter des VN hat signifikanten Einfluss auf Stornoverhalten

Erkenntnisse aus dem Stornomodell – Alter VN

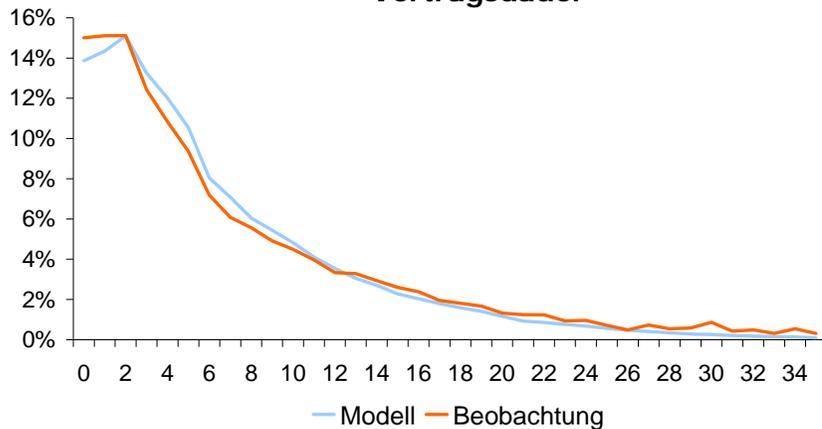


- deutliche Altersabhängigkeit beim Storno zu beobachten
- Variable als Polynom zweiten Grades in Storno-GLM
- drittstärkster Parameter in Modell
- starke Korrelation zu anderen Merkmalen (Dauer)

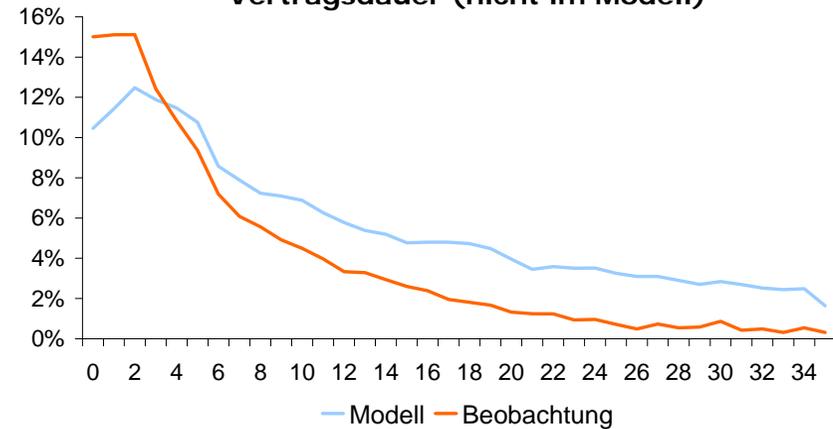
Dauer ist zweitstärkste Variable im Stornomodell

Erkenntnisse aus dem Stornomodell – Vertragsdauer

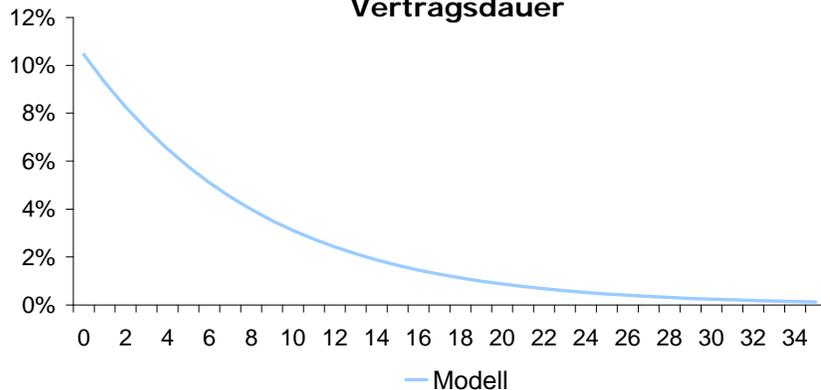
Univariate Darstellung Stornowahrscheinlichkeit
Vertragsdauer



Univariate Darstellung Stornowahrscheinlichkeit
Vertragsdauer (nicht im Modell)



Modelleffekt ohne Korrelationen
(festhalten andere Parameter)
Vertragsdauer

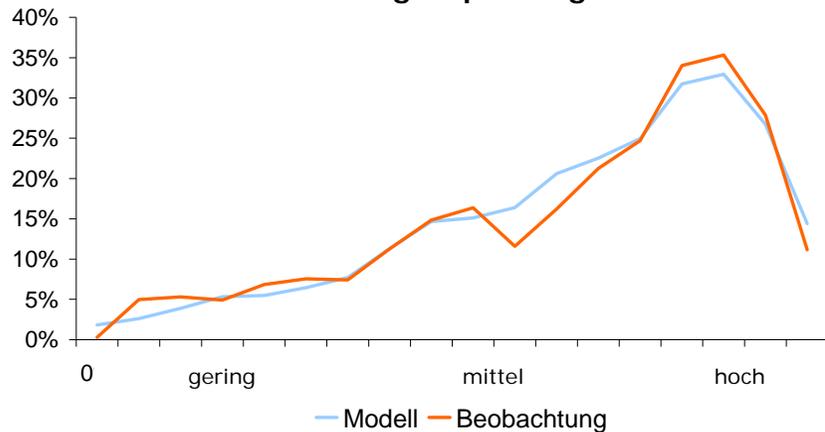


- monotone, deutliche Dauerabhängigkeit beim Storno zu beobachten
- Variable linear in Storno-GLM
- zweitstärkster Parameter in Modell
- starke Korrelation zu anderen Merkmalen (Alter)
- evtl. Korrelation zu Übertragungswerten

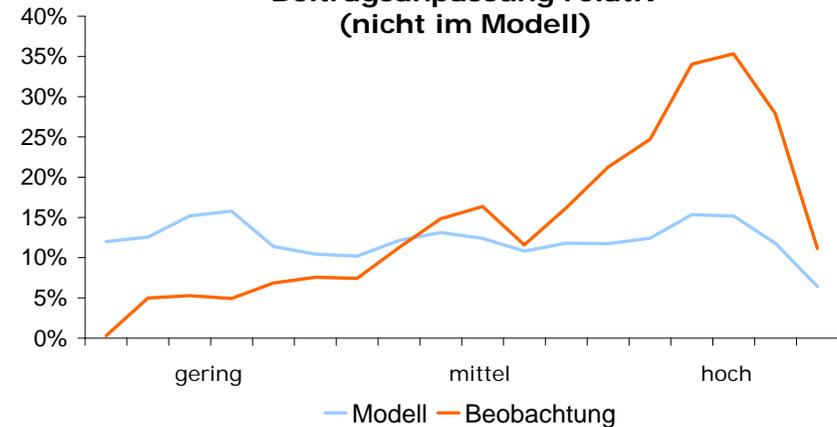
Wesentlicher Treiber für BAP-Storno ist die Beitragsanpassung

Erkenntnisse aus dem Stornomodell – Beitragsanpassung

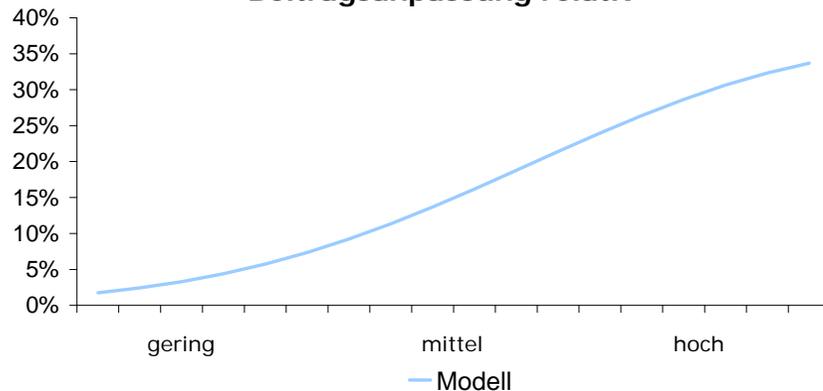
**Univariate Darstellung Stornowahrscheinlichkeit
Beitragsanpassung relativ**



**Univariate Darstellung Stornowahrscheinlichkeit
Beitragsanpassung relativ
(nicht im Modell)**



**Modelleffekt ohne Korrelationen
(festhalten andere Parameter)
Beitragsanpassung relativ**

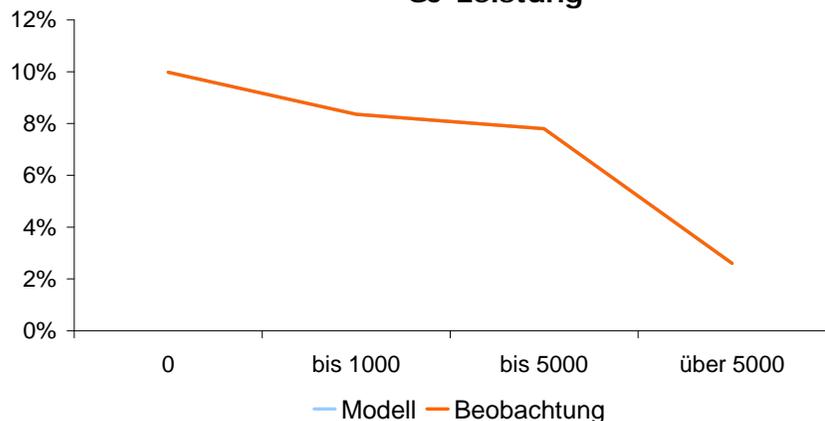


- Storno hängt wesentlich von Beitragsanpassung ab
- Variable als Polynom zweiten Grades in Storno-GLM
- mit Abstand stärkster Parameter in Modell
- wenig Korrelationen zu anderen Merkmalen
- unterschiedlicher Verlauf, je nach Tariflinie

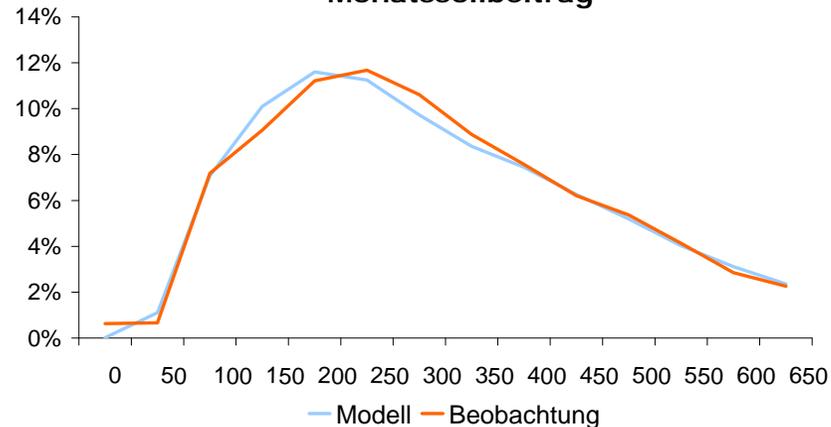
Weitere Parameter tragen zur Verbesserung der Gesamtmodellgüte signifikant bei

Erkenntnisse aus dem Stornomodell – Weitere signifikante Merkmale

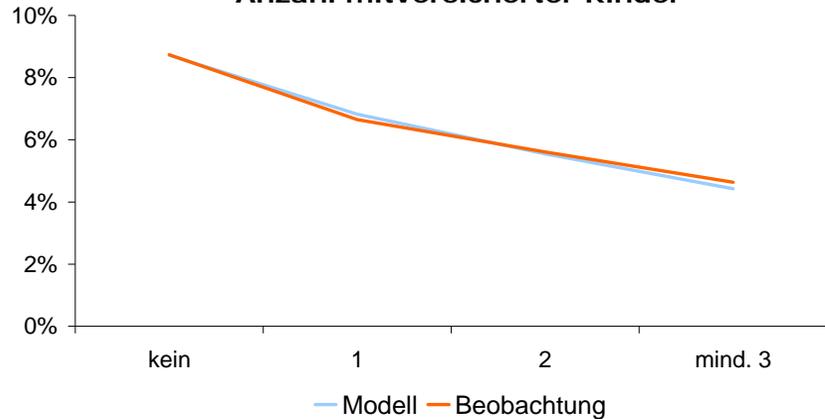
Univariate Darstellung Stornowahrscheinlichkeit GJ-Leistung



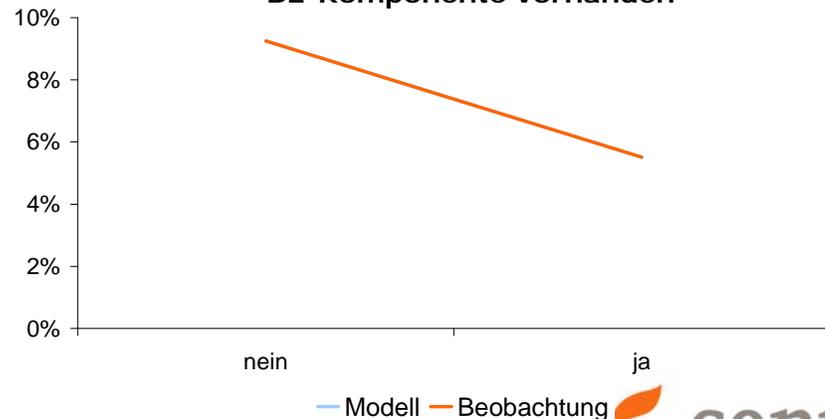
Univariate Darstellung Stornowahrscheinlichkeit Monatssollbeitrag



Univariate Darstellung Stornowahrscheinlichkeit Anzahl mitversicherter Kinder



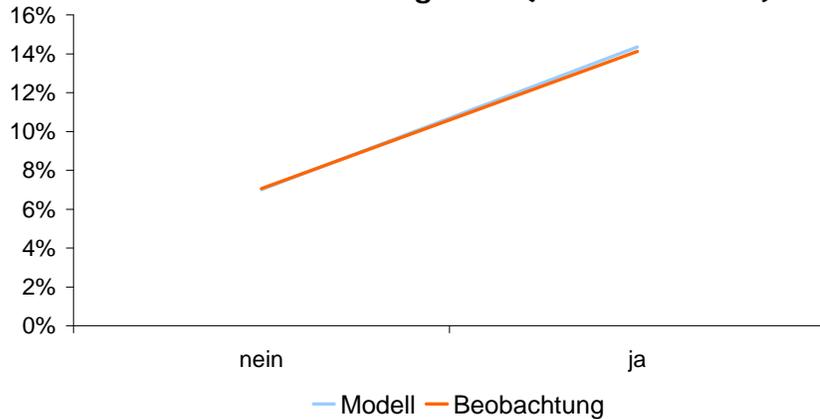
Univariate Darstellung Stornowahrscheinlichkeit BE-Komponente vorhanden



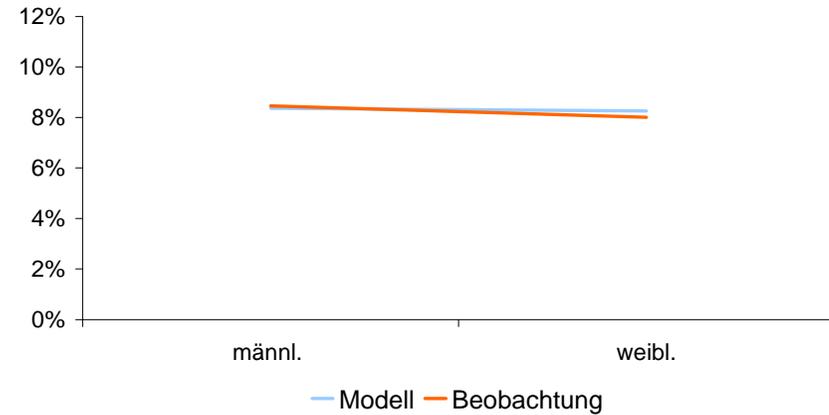
Einige Parameter sind univariat noch trennend, haben im multivariaten Ansatz aber keine Signifikanz mehr

Erkenntnisse aus dem Stornomodell – Nicht signifikante Merkmale

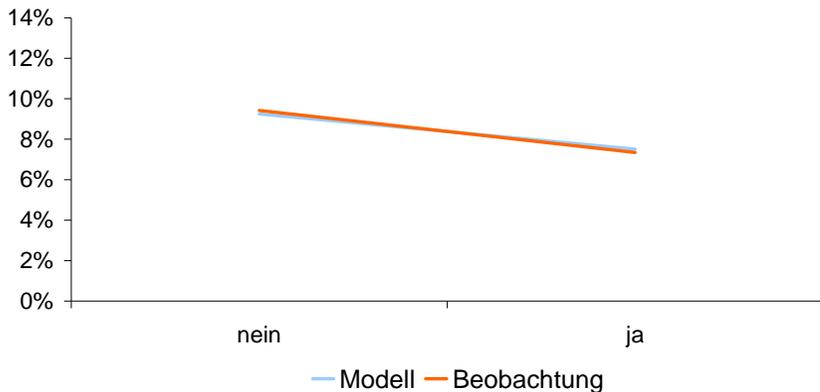
Univariate Darstellung Stornowahrscheinlichkeit Pauschalleistungstarif (nicht im Modell)



Univariate Darstellung Stornowahrscheinlichkeit Geschlecht (nicht im Modell)

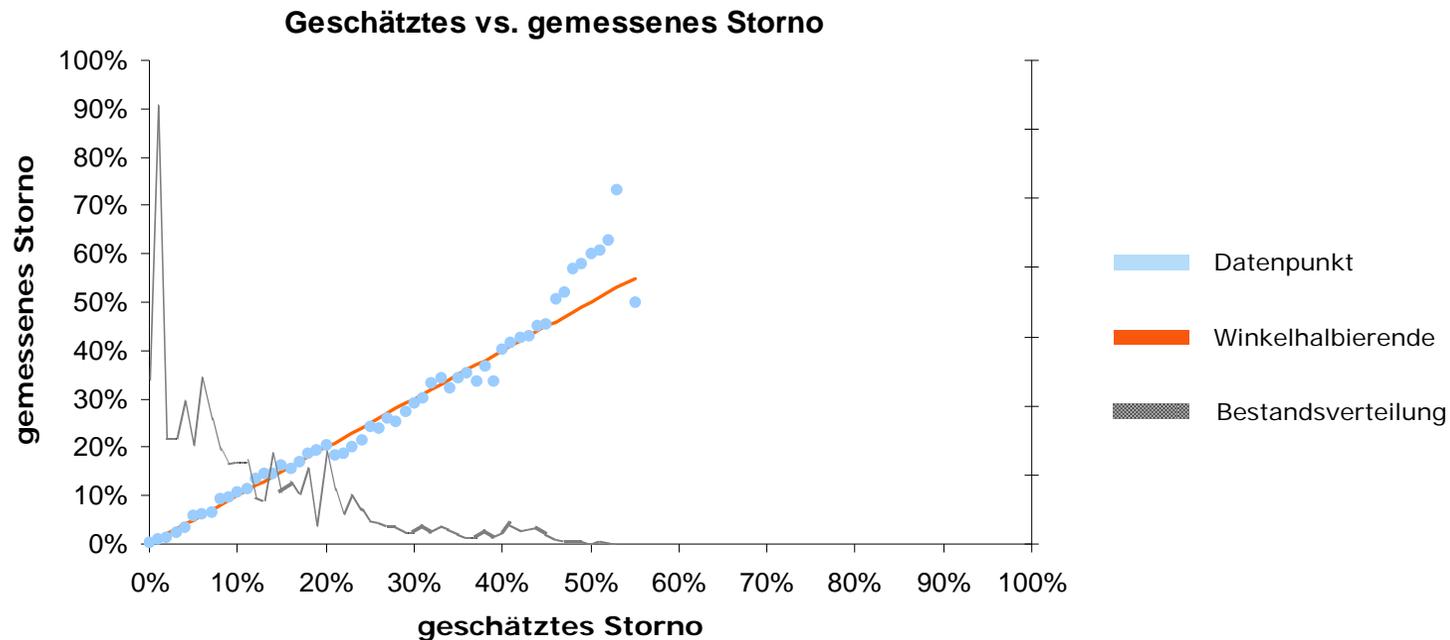


Univariate Darstellung Stornowahrscheinlichkeit KT mitversichert (nicht im Modell)



Verteilung der geschätzten Storni passt zu gemessenen Storni

Erkenntnisse aus dem Stornomodell – Anpassungsgüte



- Portfolio wird in Klassen der geschätzten Stornowahrscheinlichkeiten unterteilt
- Innerhalb der Klassen wird durchschnittliches beobachtetes Storno ermittelt
- Verlauf nahe der Winkelhalbierenden spricht für Gesamtanpassungsgüte des Modells

Erkenntnisse aus Stornomodell für verschiedene Zwecke nutzbar

Beispiele für Nutzen des Stornomodells

- **Bestands- und Leistungsprognosen**

- Wichtiger Baustein zur Prognose von Beiträgen, Anzahl Verträge und Leistungen ist die Stornoentwicklung. Eine genauere Identifizierung der Stornoaffinität der Kunden hilft bei der Bewertung der Bestandsstruktur.

- **Planung von AK-Kapazitäten im Jahreswechsel**

- Eine Stornoprognose ermöglicht die Planung von Kapazitäten in operativen Einheiten des Betriebs.

- **Baustein des VN-Verhaltens für internes Modell**

- Das Stornoverhalten, insbesondere als BAP-Reaktion, ist ein Baustein des VN-Verhaltens im internen Solvency II-Modell.

Einführung

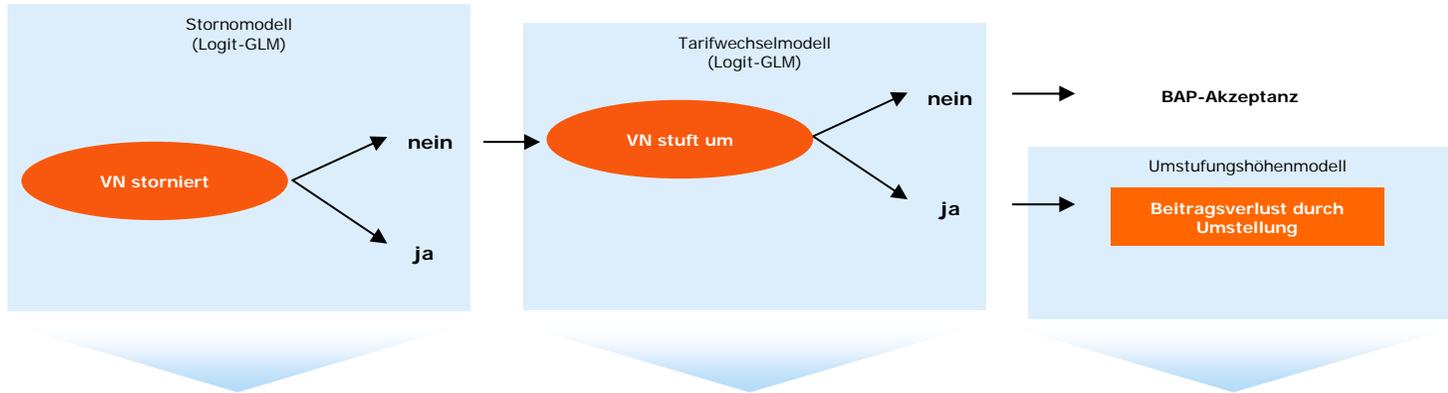
Gesetzlicher Rahmen

Stornomodell

Anwendung

Ausblick und Weiterentwicklung - Verbesserung des Modells (Identifizierung weiterer signifikanter Merkmale, Aktualisierung Daten)
- Übertragung des Modells auf Tarifwechselverhalten

Ausblick und Weiterentwicklung VN-Verhalten



- hier vorgestellt
- Weitere signifikante Variablen (Data Mining)
- auch für ZV, TG, etc.?
- Personen- vs. Tarifsicht?

- analog zu Storno, ein Logit-GLM
- andere signifikante Variablen?
- Problem der Identifikation

- Deterministisch oder stochastisch modelliert?
- Wechselmatrix
- Selektionsrisiken?

Weiterentwicklung des Modells

- mögliche zu untersuchende Parameter:

Kundenwert	Übertragungswert	Anpassungshistorie	vertriebl. Informationen	Nichtzahlerinformationen
detailliertere Leistungsdaten (Diagnosen)	externe Daten	Effekt von Deckungsrückstellungen	...	

- Validierung der Prognosekraft des Modells durch aktualisierte Daten im nächsten Jahreswechsel

Vielen Dank für Ihre Aufmerksamkeit!