

TECHNISCHE  
UNIVERSITÄT  
WIEN

Vienna University of Technology

# DIPLOMARBEIT

## Mortality Swaps und verwandte Instrumente zur Sekuritarisierung von Langlebigkeitsrisiko

Ausgeführt am Institut für  
Stochastik und Wirtschaftsmathematik  
der Technischen Universität Wien

unter der Anleitung von  
Univ.Prof. Dipl.-Math. Dr.rer.nat. Thorsten Rheinländer

durch  
Ricarda Stellner BSc BSc

Teybergasse 4/5  
1140 Wien

Wien, 27.08.2018



# Kurzfassung

In dieser Arbeit behandeln wir den  $q$ -Forward, ein grundlegendes Finanzmarktinstrument zur Absicherung des Langlebigkeitsrisikos, und dessen Bepreisung. Hierbei folgen wir der Vorgehensweise von Barrieu & Veraart (2016) in deren Paper [4]. Wir präsentieren drei aktuarielle Methoden zum Berechnen des fairen Preises, nämlich das Nettoprämienprinzip, das Standardabweichungsprinzip und das Nullnutzenprinzip. Um die Mortalität modellieren zu können, beschäftigen wir uns mit dem Lee-Carter-Modell und dem Cairns-Blake-Dowd-Modell. Nun führen wir eine empirische Studie anhand von Sterbedaten aus Österreich beziehungsweise England und Wales durch und erläutern deren abweichende Ergebnisse. Die konkrete Durchführung dieser Studie geschieht durch eine Simulation mithilfe der Programmiersprache R, wobei der verwendete Code im Anhang ersichtlich ist.

## Abstract

In this paper we discuss the  $q$ -forward, a basic financial instrument for securitization of longevity risk and its pricing. Doing this, we follow the approach of Barrieu & Veraart (2016) in their paper [4]. We review the following three actuarial methods to determine the fair value: The net premium principle, the standard deviation principle and the principle of zero utility. Furthermore, we focus on the Lee-Carter model and the Cairns-Blake-Dowd model, which we use for modelling mortality. We conduct an empirical study on mortality data from Austria as well as England and Wales and explain the different results. Specifically, we run a simulation written in the programming language R, where the used code can be found in the appendix.

# Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass die vorliegende Diplomarbeit nach den anerkannten Grundsätzen für wissenschaftliche Abhandlungen von mir selbstständig erstellt wurde. Alle verwendeten Hilfsmittel, insbesondere die zugrunde gelegte Literatur, sind in dieser Arbeit genannt und aufgelistet.

Das Thema dieser Diplomarbeit wurde von mir bisher weder im In- noch Ausland einer Beurteilerin/einem Beurteiler zur Begutachtung in irgendeiner Form als Prüfungsarbeit vorgelegt. Diese Diplomarbeit stimmt mit der von den Begutachterinnen/Begutachtern beurteilten Diplomarbeit überein.

Wien, 27.08.2018

# Danksagung

Ich danke zunächst Thorsten Rheinländer für die gute Betreuung dieser Diplomarbeit und allen weiteren Lehrenden der TU Wien und der Universität Wien, die mir in den letzten sechs Jahren ihr Wissen vermittelt haben.

Außerdem möchte ich mich bei Elfriede Paleta bedanken, die mich mit ihrem Mathematik-Olympiade-Kurs als Erste für dieses Fach begeistern konnte.

Weiters ein großes Dankeschön an alle Studienkollegen und Freunde, die meine Studienzeit so besonders gemacht haben. Danke dir Melanie, dass wir unser Studium vom ersten Tag an gemeinsam durchgezogen und uns immer gegenseitig motiviert haben.

Mein größter Dank gilt aber meiner Familie, die mich immer unterstützt hat und mir eine sorgenfreie Studienzeit ermöglicht hat. Danke, dass ihr immer für mich da wart!

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>9</b>
<b>2</b>	<b>Produktbeschreibung des q-Forwards</b>	<b>10</b>
<b>3</b>	<b>Einfache Modelle zur Preisfindung von q-Forwards</b>	<b>14</b>
3.1	Notation und grundlegende Ideen . . . . .	14
3.2	Einige Bepreisungsmethoden für q-Forwards . . . . .	15
3.2.1	Das Nettoprämienprinzip . . . . .	16
3.2.2	Das Standardabweichungsprinzip . . . . .	16
3.2.3	Das Nullnutzenprinzip . . . . .	17
3.2.4	Diskussion der Risikoprämie . . . . .	18
3.3	Modellierung der Mortalität . . . . .	19
3.3.1	Das Lee-Carter-Modell . . . . .	19
3.3.2	Das Cairns-Blake-Dowd-Modell . . . . .	27
<b>4</b>	<b>Empirischer Vergleich der verschiedenen Preisfindungsmodelle</b>	<b>30</b>
4.1	Aufbau der empirischen Analyse . . . . .	31
4.2	Fitten der Mortalitätsmodelle an die Daten . . . . .	32
4.3	Simulation und Ergebnisse . . . . .	39
<b>5</b>	<b>Zusammenfassung und Ausblick</b>	<b>51</b>
<b>A</b>	<b>Verwendeter R Code</b>	<b>52</b>

# Abbildungsverzeichnis

2.1	Zahlungsströme eines q-Forwards zur Maturität . . . . .	10
4.1	Beobachtete Sterbewahrscheinlichkeiten $q(t, x)$ für Männer mit Alter $x \in \{60, \dots, 89\}$ in den Jahren $t \in \{1961, \dots, 2014\}$ in Österreich . . . . .	31
4.2	Verlauf der Parameterschätzungen für das Lee-Carter-Modell für die männliche Bevölkerung Österreichs. . . . .	33
4.3	Verlauf der Parameterschätzungen für das Lee-Carter-Modell für die männliche Bevölkerung England und Wales. . . . .	35
4.4	Vergleich der Forecasts für ARIMA(0,1,0) und ARIMA(2,1,0) . . . . .	37
4.5	Verlauf der Parameterschätzungen für das Cairns-Blake-Dowd-Modell für die männliche Bevölkerung Österreichs für die Jahre 1961-2014. . . . .	38
4.6	Verlauf der Parameterschätzungen für das Cairns-Blake-Dowd-Modell für die männliche Bevölkerung England und Wales für die Jahre 1961-2014. . . . .	39
4.7	Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter $x = 60$ und Daten aus den Jahren 1961-2009 für Österreich . . . . .	43
4.8	Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter $x = 60$ und Daten aus den Jahren 1961-2009 für England und Wales . . . . .	44
4.9	Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter $x = 70$ und Daten aus den Jahren 1961-2009 für Österreich . . . . .	45
4.10	Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter $x = 70$ und Daten aus den Jahren 1961-2009 für England und Wales . . . . .	46
4.11	Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter $x = 60$ und Daten aus den Jahren 1966-2014 für Österreich . . . . .	47
4.12	Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter $x = 60$ und Daten aus den Jahren 1966-2014 für England und Wales . . . . .	48
4.13	Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter $x = 70$ und Daten aus den Jahren 1966-2014 für Österreich . . . . .	49
4.14	Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter $x = 70$ und Daten aus den Jahren 1966-2014 für England und Wales . . . . .	50

# Tabellenverzeichnis

2.1	Beispielhaftes Term Sheet eines q-Forwards . . . . .	12
2.2	Zahlungsströme zur Maturität des q-Forwards . . . . .	12
4.1	Schätzer der Random-Walk-Parameter $\mu_{LC}$ und $\sigma_{LC}^2$ im Lee-Carter-Modell .	36
4.2	Schätzer der Random-Walk-Parameter $\mu_{CBD}$ und $\sigma_{CBD}\sigma_{CBD}^\top$ im Cairns-Blake-Dowd-Modell . . . . .	40



# Kapitel 1

## Einleitung

Seit Beginn des 20. Jahrhunderts ist die durchschnittliche Lebensdauer der Menschen stark angestiegen und auch in Zukunft ist ein positiver Trend zu erwarten. Diese an sich sehr erfreuliche Entwicklung bringt jedoch nicht nur soziale Problemstellungen mit sich, wie etwa die Pflegedebatte, sondern stellt auch die Versicherungsbranche vor neue Herausforderungen. Besonders betroffen sind hiervon Pensionskassen und Anbieter von Rentenprodukten. Diese müssen sich - auch unter Solvency - gegen das Langlebighkeitsrisiko absichern. Eine mögliche Lösung neben der Rückversicherung ist die Übertragung des Risikos auf den Finanzmarkt, welche mithilfe von Mortality Swaps abgewickelt werden kann. Dabei handelt es sich um Derivate mit denen eine festgelegte Mortalitätsrate gegen die in der Realität eingetretene Rate getauscht werden kann. Der grundlegende Baustein dieser Finanzinstrumente ist der q-Forward, auf welchen der Fokus dieser Arbeit liegt.

Die grundlegende Funktionsweise des q-Forwards wird in Kapitel 2 beschrieben und mithilfe eines Beispiels verständlich gemacht. Weiters wird die (noch geringe) Bedeutung von q-Forwards für Versicherungsunternehmen und den Finanzmarkt aufgezeigt.

In Kapitel 3 werden die behandelten Bepreisungsmethoden vorgestellt. Weiters wird die Modellierung der Mortalität behandelt und erklärt, wie die Modelle an die Daten gefittet werden können.

Anschließend wird in Kapitel 4 die empirische Studie mit Daten aus England & Wales sowie aus Österreich beschrieben. Hierbei werden Preise mit dazugehörigen Konfidenzintervallen für die verschiedenen Bepreisungsmethoden und Mortalitätsmodelle berechnet.

Abschließend werden in Kapitel 5 die Erkenntnisse der Studie zusammengefasst und ein Ausblick auf weitere sinnvolle Modifizierungen des behandelten Verfahrens zur Bepreisung von q-Forwards gegeben.

# Kapitel 2

## Produktbeschreibung des q-Forwards

Durch die steigende Langlebigkeit stehen Anbieter von Rentenprodukten vor folgendem Problem: Je länger deren Kunden leben, desto länger müssen diesen Zahlungen gewährt werden. Somit muss durch die steigende Lebensdauer insgesamt mehr Kapital bereitgestellt werden. Das einfachste Instrument um sich gegen dieses Risiko am Finanzmarkt abzusichern ist der q-Forward, welcher wie folgt definiert ist.

Ein q-Forward ist eine Vereinbarung zwischen zwei Vertragspartnern zu einem bestimmten im Vorhinein fixierten Zeitpunkt (der Maturität des Vertrags) einen Betrag entsprechend der realisierten Mortalität der betrachteten Population gegen einen Betrag entsprechend einer schon zu Vertragsabschluss fixierten Mortalität auszutauschen.

Dies entspricht also einem Zero-Coupon-Swap zwischen der fixierten Mortalität und der realisierten Mortalität. Zum Zeitpunkt des Vertragsabschlusses fließen somit noch keine Zahlungsströme, es wird lediglich die fixe Mortalität festgesetzt. Am Maturitätszeitpunkt wird ein Differenzbetrag der in Abbildung 2.1 dargestellten Zahlungsströme geleistet.

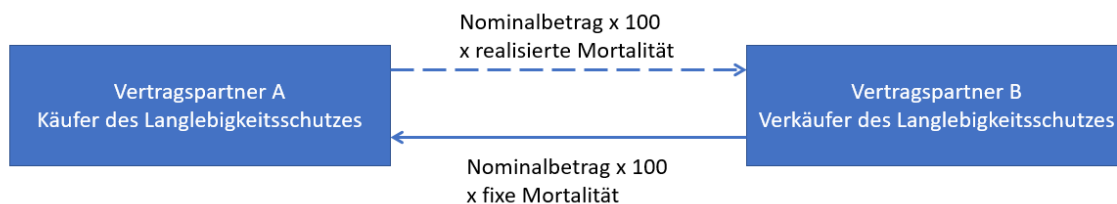


Abbildung 2.1: Zahlungsströme eines q-Forwards zur Maturität

Die zur Maturität fälligen Zahlungen setzen sich aus dem Nominalbetrag, der fixierten und der realisierten Mortalitätsrate zusammen. Der Nominalbetrag ist ein Wert, der das

Volumen des Vertrages beschreibt und zu Vertragsbeginn festgelegt wird. Auch die fixe Mortalität wird zu Beginn von den beiden Vertragspartnern festgesetzt. Die Hauptaufgabe dieser Diplomarbeit wird es sein, verschiedene Methoden zur Bestimmung eines fairen Wertes für die fixe Mortalität zu betrachten. Da zu Vertragsbeginn keine Kaufzahlungen für den  $q$ -Forward fließen, wird die Bestimmung dieses Wertes auch oft als Bepreisung des  $q$ -Forwards bezeichnet.

Die realisierte Mortalität ist durch einen Index gegeben. Hierfür gibt es zwei mögliche Arten, nämlich kundenspezifische Indizes und standardisierte Indizes. Ein kundenspezifischer Index bezieht sich genau auf das Portfolio an Versicherungsnehmern für welches eine Absicherung gewünscht ist. Im Gegensatz dazu berechnet sich ein standardisierter Index aus einer größeren Population, meist aus der Gesamtbevölkerung eines bestimmten Landes. Bei bisher abgeschlossenen Verträgen wurden standardisierte Indizes verwendet. Der erste solche Index war der LifeMetrics Index von JPMorgan. Mittlerweile werden standardisierte Indizes von der Life & Longevity Markets Association (LLMA), der auch JPMorgan angehört, herausgegeben. Anfang 2018 waren Indizes für die USA, England & Wales, Deutschland und die Niederlande verfügbar.

Nun wollen wir die Funktionsweise des  $q$ -Forwards anhand eines Beispiels betrachten. In Tabelle 2.1 ist ein beispielhafter Auszug aus einem Term Sheet für einen  $q$ -Forward gegeben. Die XYZ Pensionsvorsorge will sich durch diesen Vertrag gegen das Langlebighkeitsrisiko absichern und ist Empfänger der fixen Rate, während die Großbank JPMorgan die fixe Rate zahlt. Das Referenzjahr entspricht nicht dem Jahr der Mortalität des Vertrages, da die Indizes für die realisierte Mortalitätsrate erst mit etwas Verzögerung berechnet werden können. Die beiden entgegengesetzten Zahlungsströme die zur Maturität fällig sind, werden zu einem Nettozahlungsbetrag zusammengefasst. Dieser kann positiv oder negativ sein. Falls der Nettozahlungsbetrag positiv ist, ergibt sich eine Zahlung von der XYZ Pensionsvorsorge an JPMorgan. Ist diese Differenz jedoch negativ, so muss JPMorgan den Absolutbetrag an die XYZ Pensionsvorsorge zahlen. Diese Vorgehensweise ist in Tabelle 2.2 anhand einigen möglichen Werten für die Referenzrate dargestellt.

Warum stellt diese Vorgehensweise nun eine Absicherung für die XYZ Pensionsvorsorge gegen das Langlebighkeitsrisiko dar? Eine niedrigere Referenzrate für die Mortalität bedeutet, dass im Vergleich zur fixen Mortalitätsrate in der betrachteten Population weniger Menschen gestorben sind. Dies entspricht einer höheren Langlebighkeit gegen welche eine Absicherung gewünscht ist. Nun erhält die XYZ Pensionsvorsorge jedoch wie in Tabelle 2.2 erkennbar ist den Absolutbetrag des Nettozahlungsbetrags. Diesen kann sie als Kompensation für die höheren Zahlungen an deren Kunden verwenden.

Ist die Referenzrate jedoch höher als die fixe Rate, so entspricht dies einer höheren Mortalität und somit einer niedrigeren Langlebighkeit. Nun muss die XYZ Pensionsvorsorge eine entsprechende Zahlung an JPMorgan leisten. Durch die niedrigere Langlebighkeit fielen je-

Nominalbetrag	EUR 100.000.000
Vertragspartner A	XYZ Pensionsvorsorge
Vertragspartner B	JPMorgan
Handelsdatum	31.12.2016
Maturitätsdatum	31.12.2026
Referenzjahr	2025
fixe Rate	1,3 %
Referenzrate	von der LLMA berechnete Mortalitätsrate für 64-jährige Männer aus Deutschland im Referenzjahr
Zahlung von A zur Maturität	Nominalbetrag x 100 x Referenzrate
Zahlung von B zur Maturität	Nominalbetrag x 100 x fixe Rate
Nettozahlungsbetrag	Nominalbetrag x 100 x (Referenzrate - fixe Rate )

Tabelle 2.1: Beispielhaftes Term Sheet eines q-Forwards

Referenzrate	fixe Rate	Nettozahlungsbetrag	Empfänger des Absolutbetrages
1,1 %	1,3 %	-20.000.000	XYZ Pensionsvorsorge
1,2 %	1,3 %	-10.000.000	XYZ Pensionsvorsorge
1,3 %	1,3 %	0	-
1,4 %	1,3 %	10.000.000	JPMorgan

Tabelle 2.2: Zahlungsströme zur Maturität des q-Forwards

doch die Zahlungen an deren Rentenkunden niedriger aus.

Problematisch ist jedoch, dass sich die Mortalität des Portfolios der XYZ Pensionsvorsorge nicht zwingend so verhalten muss wie die der Gesamtbevölkerung, welche zur Berechnung der Referenzrate herangezogen wird. Dieses sogenannte Basisrisiko kann jedoch durch aufmerksames Design der Absicherung minimiert werden.

Durch den Verkauf von  $q$ -Forwards können sich also Versicherungsgesellschaften mit Rentenprodukten im Portfolio gegen das Langlebigkeitsrisiko absichern. Umgekehrt ist auch eine Absicherung gegen das Mortalitätsrisiko durch den Kauf von  $q$ -Forwards möglich. Dies kann für Versicherungsgesellschaften mit einem Ablebensportfolio interessant sein. Ist die Mortalität höher als ursprünglich gedacht, muss diese Gesellschaft nun höhere Zahlungen leisten. Um sich dagegen abzusichern wird die Versicherungsgesellschaft in einem  $q$ -Forward zum Zahler der fixen Rate und der andere Vertragspartner zum Zahler der realisierten Rate.

Der  $q$ -Forward ist ein relativ neues Produkt, welches im Jahr 2007 erstmals von der Großbank JPMorgan gehandelt wurde. Bis heute ist der Markt für Derivate zur Absicherung von Langlebigkeitsrisiko nicht sehr entwickelt und noch nicht liquide. Der Name  $q$ -Forward leitet sich von der aktuariellen Notation her, in welcher der Buchstabe  $q$  für die Sterbewahrscheinlichkeit verwendet wird. Prinzipiell gibt es sowohl natürliche Käufer als auch Verkäufer für  $q$ -Forwards. Allerdings ist das Volumen der Verbindlichkeiten, das die Pensionskassen hedgen wollen deutlich größer als jenes der Lebensversicherer mit Ablebensportfolio. Somit gibt es zu wenige Käufer für  $q$ -Forwards, was sich auf die Marktpreise auswirken wird.

# Kapitel 3

## Einfache Modelle zur Preisfindung von q-Forwards

### 3.1 Notation und grundlegende Ideen

Im Folgenden verwenden wir die Notation, welche von Cairns et al.(2009) in deren Paper [6] eingeführt wurde.

- Zunächst definieren wir ein Kalenderjahr als die Zeitspanne von Zeitpunkt  $t$  bis zum Zeitpunkt  $t + 1$ .
- Die Sterberate für das Alter  $x$  im Kalenderjahr  $t$  bezeichnen wir mit  $m(t, x)$ . Dies ist gegeben durch

$$m(t, x) := \frac{\text{Anzahl der Toten mit Alter } x \text{ beim letzten Geburtstag im Kalenderjahr } t}{\text{Durchschnittliche Population mit Alter } x \text{ beim letzten Geburtstag im Kalenderjahr } t} .$$

Die durchschnittliche Population wird meist durch eine Schätzung der Population mit Alter  $x$  beim letzten Geburtstag in der Mitte des Jahres approximiert.

- Weiters bezeichnen wir mit  $q(t, x)$  die Sterbewahrscheinlichkeit. Dies ist die Wahrscheinlichkeit für eine Person mit Alter  $x$  zum Zeitpunkt  $t$  im Zeitraum zwischen  $t$  und  $t + 1$  zu sterben. Somit entspricht  $q(t, x)$  nicht der in Österreich gebräuchlichen Notation  ${}_tq_x$ , welche die Wahrscheinlichkeit bezeichnet, dass ein  $x$ -jähriger innerhalb von  $t$  Jahren sterben wird.
- Die Sterblichkeitsintensität bezeichnen wir mit  $\mu(t, x)$ . Für kleine  $dt$  kann mit  $\mu(t, x)dt$  die Wahrscheinlichkeit, dass eine Person mit Alter  $x$  zum Zeitpunkt  $t$  zwischen  $t$  und  $t + dt$  stirbt, approximiert werden.

Eine grundlegende Formel der Lebensversicherungsmathematik stellt folgenden Zusammenhang zwischen  $q(t, x)$  und  $\mu(t, x)$  her:

$$q(t, x) = 1 - \exp\left(-\int_t^{t+1} \mu(s, x) ds\right) . \quad (3.1)$$

Um nun auch  $m(t, x)$  und  $q(t, x)$  in Verbindung bringen zu können, treffen wir folgende Annahme:

Annahme 1: Für natürliche Zahlen  $t$  und  $x$  und alle  $0 \leq s, u < 1$  gelte

$$\mu(t + s, x + u) = \mu(t, x) .$$

Anders formuliert bedeutet dies, dass die Sterblichkeitsintensität im Verlauf eines Kalenderjahres konstant bleibt. Daraus folgt nun sofort

$$m(t, x) = \mu(t, x) , \quad (3.2)$$

$$q(t, x) = 1 - \exp(-\mu(t, x)) = 1 - \exp(-m(t, x)) . \quad (3.3)$$

Annahme 1 ist im Allgemeinen nicht erfüllt, allerdings wird (3.3) als gute Approximation für die Beziehung zwischen  $q(t, x)$  und  $m(t, x)$  gesehen.

Nun wollen wir eine Notation für die wichtigsten Eckdaten eines q-Forwards einführen. Wir bezeichnen im Folgenden den Maturitätszeitpunkt mit  $T$ , den Nominalbetrag mit  $z$  und die fixierte Mortalität mit  $K$ . Weiters bezeichnen wir mit  $q(T)$  die realisierte Mortalität der Referenzpopulation zur Zeit  $T$ . Somit ergibt sich der Nettozahlungsbetrag (Net Payoff Amount) durch

$$\text{NPA}(T) = z(q(T) - K) , \quad (3.4)$$

wobei wir den Skalierungsfaktor 100 aus Kapitel 2 im Weiteren vernachlässigen.

Da zu Vertragsbeginn keine Zahlungen fließen, muss der Wert des q-Forwards, der im Weiteren mit  $\text{Value}_0(\cdot)$  bezeichnet wird, zu diesem Zeitpunkt gleich 0 sein. Somit ergibt sich im Allgemeinen die Beziehung

$$\text{Value}_0(z(q(T) - K)) = 0 , \quad (3.5)$$

woraus sich für bestimmte Bepreisungsmethoden der faire Wert für die fixierte Mortalität  $K$  berechnen lässt.

## 3.2 Einige Bepreisungsmethoden für q-Forwards

Im Folgenden wollen wir einige grundlegende Bepreisungsmethoden vorstellen und anhand dieser den fairen Wert für  $K$  berechnen. Wir betrachten nun immer einen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$ , wobei  $\mathbb{P}$  das verwendete Wahrscheinlichkeitsmaß bezeichnet. In unserer

Betrachtungsweise ist  $\mathbb{P}$  das statistische bzw. historische Wahrscheinlichkeitsmaß.

Da die betrachteten Verträge meist sehr lange Laufzeiten haben, ist es notwendig den Endwert mit einem entsprechenden Zinssatz abzuzinsen. Eine gute Prognose dieses Zinssatzes ist ein schwieriges Unterfangen, das jedoch für komplexere Langlebigkeitsderivate sehr wichtig ist. Bei den betrachteten q-Forwards ist der Einfluss des Zinssatzes auf Grund der Struktur der Zahlungsströme limitiert. Wir nehmen im Folgenden an, dass ein risikoloser Titel am Finanzmarkt verfügbar ist, dessen Zeitwert zum Zeitpunkt  $t$  sich als  $e^{rt}$  ergibt und somit den fixen Zinssatz  $r$  liefert. Somit wird (3.5) zu

$$\text{Value}_0(\exp(-rT)z(q(T) - K)) = 0 \quad (3.6)$$

erweitert.

### 3.2.1 Das Nettoprämienprinzip

Das Nettoprämienprinzip ist das grundlegendste aktuarielle Prämienkalkulationsprinzip und auch als Äquivalenzprinzip bekannt. Die Nettoprämie ergibt sich als Erwartungswert unter dem statistischen Wahrscheinlichkeitsmaß  $\mathbb{P}$ :

$$\text{Value}_0(\text{NPA}(T)) = \mathbb{E}_{\mathbb{P}}[e^{-rT}z(q(T) - K)] = 0 . \quad (3.7)$$

Daraus folgt im Weiteren

$$e^{-rT}z(\mathbb{E}_{\mathbb{P}}[q(T)] - K) = 0$$

und wir erhalten den fairen Wert für  $K$  als

$$K = \mathbb{E}_{\mathbb{P}}[q(T)] \quad (3.8)$$

Die Nettoprämie bildet die Grundlage für weitere Bepreisungsmethoden bei welchen ein Risikoabschlag bzw. -zuschlag hinzugefügt wird.

### 3.2.2 Das Standardabweichungsprinzip

Eine Möglichkeit für diese Vorgehensweise bietet das Standardabweichungsprinzip. Die Bepreisung ergibt sich hier als der Erwartungswert unter dem Wahrscheinlichkeitsmaß  $\mathbb{P}$  zu dem ein Vielfaches der Standardabweichung addiert wird. Somit erhalten wir

$$\text{Value}_0(\text{NPA}(T)) = \mathbb{E}_{\mathbb{P}}[e^{-rT}z(q(T) - K)] + \lambda\sqrt{\text{Var}_{\mathbb{P}}(e^{-rT}z(q(T) - K))} = 0 . \quad (3.9)$$

Daraus ergibt sich

$$\begin{aligned} e^{-rT}z(\mathbb{E}_{\mathbb{P}}[q(T)] - K) + \lambda e^{-rT}z\sqrt{\text{Var}_{\mathbb{P}}(q(T) - K)} &= 0 , \\ \mathbb{E}_{\mathbb{P}}[q(T)] - K + \lambda\sqrt{\text{Var}_{\mathbb{P}}(q(T))} &= 0 , \end{aligned}$$



und wir erhalten für den fairen Wert für die fixe Mortalität

$$K = \mathbb{E}_{\mathbb{P}}[q(T)] + \lambda \sqrt{\text{Var}_{\mathbb{P}}(q(T))} . \quad (3.10)$$

Die Wahl der Parameters  $\lambda$  hat offensichtlich einen großen Einfluss auf die Berechnung der fixierten Mortalität  $K$ . Diese Entscheidung kann von vielen Faktoren abhängen. Zum einen gibt es externe Faktoren wie Vorschriften der Regulierungsbehörden, zum anderen gibt es interne Faktoren, wie zum Beispiel wie stark das Basisrisiko mit der Wahl des Langlebigkeitsindex assoziiert ist. Wir werden im folgenden Abschnitt sehen, wie sich  $\lambda$  durch die Interpretation des additiven Ausdrucks als Risikoprämie auch als mit dem Sharpe-Quotienten verwandten Parameter ergibt.

### 3.2.3 Das Nullnutzenprinzip

Eine weitere Vorgehensweise zur Preisfindung ist das Nullnutzenprinzip. Hierbei wählt man eine wachsende und konkave Funktion  $U$  als Nutzenfunktion. Diese repräsentiert den Nutzen, den der Besitzer eines gewissen Assets aus diesem zieht. Da ein höherer Vermögenswert auch einen höheren Nutzen liefern sollte, wird die Funktion  $U$  als wachsende Funktion gewählt. Eine weitere ökonomische Überlegung ist die Tatsache, dass Investoren bei gleichem erwarteten Payoff das Investment mit dem geringeren Risiko wählen. Diese Risikoaversion wird durch die konkave Form von  $U$  repräsentiert.

Um nun den fairen Preis für die fixe Mortalität  $K$  zu bestimmen, ist es notwendig den erwarteten Payoff des q-Forwards mit jenem eines alternativen Investments zu vergleichen. Der Investor kann sein Anfangsvermögen  $W_0$  also nur in den risikolosen Titel anlegen oder zusätzlich mithilfe eines q-Forwards. Wurde für  $K$  der faire Wert gewählt, so sollte der erwartete Nutzen beider Möglichkeiten gleich sein. Es ergibt sich nun

$$\mathbb{E}_{\mathbb{P}}[U(W_0 e^{rT} + z(q(T) - K))] = \mathbb{E}_{\mathbb{P}}[U(W_0 e^{rT})] . \quad (3.11)$$

Im Weiteren verwenden wir die exponentielle Nutzenfunktion  $U(y) = -\exp(-\gamma y)$ , wobei  $\gamma$  den konstanten Risikoaversionsparameter bezeichnet. Somit erhalten wir

$$\mathbb{E}_{\mathbb{P}}[-\exp(-\gamma W_0 e^{rT} - \gamma z(q(T) - K))] = \mathbb{E}_{\mathbb{P}}[-\exp(-\gamma W_0 e^{rT})] . \quad (3.12)$$

Dies führt zu

$$\begin{aligned} -\exp(-\gamma W_0 e^{rT}) \mathbb{E}_{\mathbb{P}}[\exp(-\gamma z(q(T) - K))] &= -\exp(-\gamma W_0 e^{rT}) , \\ \mathbb{E}_{\mathbb{P}}[\exp(-\gamma z(q(T) - K))] &= 1 , \\ \mathbb{E}_{\mathbb{P}}[\exp(-\gamma z q(T))] \exp(\gamma z K) &= 1 , \\ \mathbb{E}_{\mathbb{P}}[\exp(-\gamma z q(T))] &= \exp(-\gamma z K) , \\ \log(\mathbb{E}_{\mathbb{P}}[\exp(-\gamma z q(T))]) &= -\gamma z K \end{aligned}$$

und somit erhalten wir für den fairen Wert der fixen Mortalität

$$K = -\frac{1}{\gamma z} \log(\mathbb{E}_{\mathbb{P}}[\exp(-\gamma z q(T))]) . \quad (3.13)$$

Ein großer Unterschied zu den anderen beiden Bepreisungsmethoden ist die Abhängigkeit des fairen Werts für die fixierte Mortalität  $K$  vom Nominalbetrag  $z$ . Während beim Nettoprämienprinzip und beim Standardabweichungsprinzip das Nominale  $z$  keinen Einfluss hat, hängt der faire Wert nun nicht-linear von  $z$  ab. Weiters treten das Nominale  $z$  und der Risikoaversionsparameter  $\gamma$  immer gemeinsam als Produkt  $\gamma z$  in der Formel für  $K$  auf. Betrachten wir nun zwei Marktteilnehmer mit Risikoaversionsparametern  $\gamma_1 < \gamma_2$ . Falls diese ihre Nominalbeträge so wählen, dass  $\gamma_1 z_1 = \gamma_2 z_2$  gilt, so erhalten beide den gleichen fairen Wert für die fixierte Mortalität  $K$ . Da somit  $z_1 > z_2$  gelten muss, ergibt sich, dass der Teilnehmer mit der niedrigeren Risikoaversion  $\gamma_1$  ein größeres Volumen  $z_1$  kaufen kann.

### 3.2.4 Diskussion der Risikoprämie

Bei der Betrachtung des Marktes für q-Forwards haben wir festgestellt, dass es deutlich mehr Marktteilnehmer gibt, die q-Forwards verkaufen wollen, als solche, die an einem Kauf interessiert sind. Somit werden die Marktpreise für q-Forwards einen Risikoabschlag beinhalten um Investoren anzulocken. Konkret lässt sich also sagen, dass q-Forwards nicht mit dem Nettoprämienpreis gehandelt werden, sondern noch mit einer negativen Risikoprämie  $R$  versehen werden. Somit ergibt sich

$$K = \mathbb{E}_{\mathbb{P}}[q(T)] + R , \quad (3.14)$$

wobei  $R < 0$  ist.

Der faire Wert im Standardabweichungsprinzip weist genau diese Struktur auf. Das Vorzeichen des additiven Ausdrucks hängt nur von  $\lambda$  ab. Somit können wir  $\lambda < 0$  annehmen. Damit Investoren an einem Kauf interessiert sind, muss ihnen ein ausreichender Return im Verhältnis zum Risiko geboten werden. Eine klassische Kennzahl dafür ist der Sharpe-Quotient, welcher als

$$S = \frac{\mathbb{E}[R - R_f]}{\sqrt{\text{Var}(R)}}$$

definiert ist. Hier bezeichnet  $R$  den Return des betrachteten Assets und  $R_f$  den konstanten Return einer risikolosen Veranlagung. Übertragen auf die Situation beim q-Forward ergibt sich der Sharpe-Quotient nun als

$$S = \frac{\mathbb{E}_{\mathbb{P}}[q(T)] - K}{\sqrt{\text{Var}_{\mathbb{P}}(q(T))}} . \quad (3.15)$$

Wir können nun einen minimalen Sharpe-Quotienten  $S$  festlegen, bei dem Investoren noch am Kauf interessiert wären und anschließend (3.15) nach  $K$  lösen. Dann ergibt sich

$$K = \mathbb{E}_{\mathbb{P}}[q(T)] - S\sqrt{\text{Var}_{\mathbb{P}}(q(T))} \quad (3.16)$$

und wir können  $\lambda = -S$  setzen. Somit haben wir eine marktnahe Möglichkeit zur Wahl des Parameters  $\lambda$  beim Standardabweichungsprinzip gefunden.

### 3.3 Modellierung der Mortalität

Im Weiteren folgen wir dem Vorgehen von Barrieu & Veraart (2016) in deren Paper [4]. Dort werden drei verschiedene Modelle verglichen, nämlich das Lee-Carter-Modell (1992) mit Random Walk, das Lee-Carter-Modell mit ARIMA und das Cairns-Blake-Dowd-Modell (2009). Alle drei haben gemeinsam, dass es sich um extrapolative Modelle handelt, die nur von den gegebenen Daten wie der Größe der Population und der dazugehörigen Anzahl an Todesfällen für bestimmte Zeitintervalle und Altersklassen abhängen.

#### 3.3.1 Das Lee-Carter-Modell

Im Lee-Carter-Modell (1992) beschreiben wir die logarithmierte Sterberate als

$$\log(m(t, x)) = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)} . \quad (3.17)$$

Hier modellieren die beiden Funktionen  $\beta_x^{(1)}$  und  $\beta_x^{(2)}$  den Einfluss des Alters, während  $\kappa_t^{(2)}$  die Veränderungen durch den Zeitablauf widerspiegelt. Dieses Modell kann auch als  $\log(m(t, x)) = \beta_x^{(1)} \kappa_t^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}$ , wobei  $\kappa_t^{(1)} = 1$  für alle Zeitpunkte  $t$ , aufgefasst werden.

Bei der Schätzung der Parameter aus den Daten stößt man auf das Problem der Identifizierbarkeit. Betrachten wir kurz das Modell

$$\log(m(t, x)) = \tilde{\beta}_x^{(1)} + \tilde{\beta}_x^{(2)} \tilde{\kappa}_t^{(2)} , \quad (3.18)$$

wobei

$$\begin{aligned} \tilde{\beta}_x^{(1)} &= \beta_x^{(1)} + b\beta_x^{(2)} , \\ \tilde{\beta}_x^{(2)} &= \frac{\beta_x^{(2)}}{a} , \\ \tilde{\kappa}_t^{(2)} &= a(\kappa_t^{(2)} - b) . \end{aligned}$$

Durch Einsetzen erkennt man, dass die Modelle in (3.17) und (3.18) übereinstimmen, jedoch die Parameter anders gewählt sind. Somit müssen wir zusätzliche Nebenbedingungen einführen um die Eindeutigkeit der Schätzung zu gewährleisten. Dafür gibt es verschiedene Möglichkeiten. Im Folgenden verwenden wir die von Cairns et al.(2009) vorgeschlagenen Einschränkungen

$$\sum_t \kappa_t^{(2)} = 0 , \quad (3.19)$$

$$\sum_x \beta_x^{(2)} = 1 . \quad (3.20)$$

Die erste Einschränkung (3.19) ist natürlich gegeben und impliziert, dass der Schätzer für  $\beta_x^{(1)}$  dem Mittel über alle  $t$  der logarithmierten Sterberate  $\log(m(t, x))$  entspricht. Um beide Parameter  $a$  und  $b$  eindeutig fixieren zu können, benötigen wir noch eine zweite Einschränkung. Dafür gibt es keine natürlich gegebene Nebenbedingung und in der Literatur finden sich verschieden Ansätze. Allerdings hat die Wahl der zweiten Nebenbedingung keinen Einfluss auf die Qualität des Fits des Modells als auch auf die Prognosen der Mortalität.

Anhand der Daten zur Sterblichkeit verschiedener Industrienationen können bereits einige grundlegende Aussagen zum Verhalten der Funktionen  $\beta_x^{(1)}$ ,  $\beta_x^{(2)}$  und  $\kappa_t^{(2)}$  getroffen werden. So ist  $\kappa_t^{(2)}$  fallend, woraus man schließen kann, dass die Sterberaten im Allgemeinen im Laufe der Zeit kleiner werden. Weiters kann man annehmen, dass  $\beta_x^{(1)}$  wachsend ist. Dies bedeutet, dass die Sterberaten größer werden je älter die betrachtete Person ist. Für  $\beta_x^{(2)}$  lässt sich eine fallende Tendenz beobachten, was bedeutet, dass die Verbesserungen bei der Langlebigkeit auf Personen mit niedrigerem Alter stärkeren Einfluss hat.

Um das Lee-Carter-Modell verwenden zu können, muss es an die vorhandenen Daten angepasst werden. Wir wollen also Schätzer für  $\beta_x^{(1)}$ ,  $\beta_x^{(2)}$  und  $\kappa_t^{(2)}$  finden. Angenommen wir haben Daten für  $n + 1 \in \mathbb{N}$  Jahre, welche wir mit  $t_0, \dots, t_n$  bezeichnen und für welche  $t_{i+1} = t_i + 1$  gelten soll. Diese Daten seien für  $m + 1$  Altersklassen gegeben, welche mit  $x_0, x_1, \dots, x_m$  bezeichnet werden. Aus den Daten kann direkt die Sterberate  $m(t, x)$  berechnet werden. Nun sollen  $\beta_x^{(1)}$ ,  $\beta_x^{(2)}$  und  $\kappa_t^{(2)}$  für  $x = x_0, \dots, x_m$  und  $t = t_0, \dots, t_n$  anhand der Modelldefinition (3.17) und den Nebenbedingungen (3.19) und (3.20) geschätzt werden. Summieren wir die Gleichungen, die durch (3.17) gegeben sind für ein festes Alter  $x$  über alle Jahre  $t$  auf, so erhalten wir

$$\sum_{t=t_0}^{t_n} \log(m(t, x)) = (n + 1)\beta_x^{(1)} + \beta_x^{(2)} \sum_{t=t_0}^{t_n} \kappa_t^{(2)} .$$

Durch die Nebenbedingung (3.19) fällt der zweite Summand weg und es ergibt sich der Schätzer für  $\beta_x^{(1)}$ ,  $x = x_0, \dots, x_m$  als

$$\hat{\beta}_x^{(1)} = \frac{1}{n + 1} \sum_{t=t_0}^{t_n} \log(m(t, x)) . \quad (3.21)$$

Wie von Lee & Carter (1992) vorgeschlagen, führen wir nun eine Singulärwertzerlegung durch, um Schätzer für  $\beta_x^{(2)}$  und  $\kappa_t^{(2)}$  zu finden. Dafür benötigen wir die Matrix der logarithmierten Sterberaten abzüglich der schon berechneten Schätzer für  $\beta_x^{(1)}$ . Somit definieren wir

$$M = \begin{pmatrix} \log(m(t_0, x_0)) - \hat{\beta}_{x_0}^{(1)} & \dots & \log(m(t_0, x_m)) - \hat{\beta}_{x_m}^{(1)} \\ \vdots & \ddots & \vdots \\ \log(m(t_n, x_0)) - \hat{\beta}_{x_0}^{(1)} & \dots & \log(m(t_n, x_m)) - \hat{\beta}_{x_m}^{(1)} \end{pmatrix}$$

$M$  ist eine reelle  $(n + 1) \times (m + 1)$  Matrix. Dessen Singulärwertzerlegung ergibt die Darstellung  $M = U\Sigma V^\top$ , wobei

- $U$  eine orthogonale  $(n + 1) \times (n + 1)$ -Matrix ist,
- $V^\top$  die Transponierte einer orthogonalen  $(m + 1) \times (m + 1)$  Matrix und
- $\Sigma$  eine reelle  $(n + 1) \times (m + 1)$  Matrix der Form

$$\Sigma = \left( \begin{array}{ccc|ccc} \sigma_1 & & & \vdots & & \\ & \ddots & & \dots & 0 & \dots \\ & & \sigma_r & \vdots & & \\ \hline & \vdots & & \vdots & & \\ \dots & 0 & \dots & \dots & 0 & \dots \\ & \vdots & & \vdots & & \end{array} \right)$$

mit  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$  ist und die Anzahl  $r$  der Diagonalelemente dem Rang der Matrix  $M$  entspricht.

Man kann für jede Matrix eine Singulärwertzerlegung finden, wobei  $\Sigma$  immer eindeutig bestimmt ist. Eine weitere Darstellungsmöglichkeit der Singulärwertzerlegung ergibt sich mithilfe der Spaltenvektoren  $U_j, V_j$  der Matrizen  $U$  und  $V$  auf folgende Weise

$$M = \sigma_1 U_{.1} V_{.1}^\top + \sigma_2 U_{.2} V_{.2}^\top + \dots + \sigma_r U_{.r} V_{.r}^\top \quad (3.22)$$

Durch die Darstellung der gesuchten Schätzer  $\hat{\beta}_x^{(2)}$  für  $x = x_0, \dots, x_m$  und  $\hat{\kappa}_t^{(2)}$  für  $t = t_0, \dots, t_n$  als Vektoren  $\hat{\beta}^{(2)} = (\hat{\beta}_{x_0}^{(2)}, \dots, \hat{\beta}_{x_m}^{(2)})^\top$  und  $\hat{\kappa}^{(2)} = (\hat{\kappa}_{t_0}^{(2)}, \dots, \hat{\kappa}_{t_n}^{(2)})^\top$  ergibt sich die aus dem Lee-Carter-Modell stammende Matrix  $M$  als  $M = \hat{\kappa}^{(2)} (\hat{\beta}^{(2)})^\top$ .

Wir approximieren nun  $\hat{\kappa}^{(2)} (\hat{\beta}^{(2)})^\top$  durch den ersten Summand der Darstellung (3.22) der Singulärwertzerlegung. Um die Einschränkung (3.20) zu erfüllen, wählen wir  $\hat{\beta}^{(2)} = \frac{1}{w} V_{.1}$ , wobei das Gewicht  $w = \sum_{i=1}^{m+1} V_{i1}$  der Summe der Einträge des Vektors  $V_{.1}$  entspricht. Somit ergibt sich der Schätzer  $\hat{\kappa}^{(2)}$  als  $\hat{\kappa}^{(2)} = w \sigma_1 U_{.1}$ .

## Das Lee-Carter-Modell mit Random Walk

Barrieu & Veraart (2016) schlagen nun vor, zunächst ein Lee-Carter-Modell (1992) mit Random Walk zu betrachten. Hierbei wird die zeitabhängige Funktion  $\hat{\kappa}_t^{(2)}$  durch

$$\hat{\kappa}_t^{(2)} = \hat{\kappa}_{t-1}^{(2)} + \mu_{LC} + \sigma_{LC} Z_t, \quad (3.23)$$

modelliert, wobei  $Z_t$  unabhängig identisch verteilte Zufallsvariablen mit einer Standardnormalverteilung und  $\mu_{LC} \in \mathbb{R}, \sigma_{LC} > 0$  sind. Diese beiden Parameter müssen nun anhand

der Daten geschätzt werden. Hierfür benötigen wir die eben hergeleiteten Schätzer  $\hat{\kappa}_t^{(2)}$  für  $t = t_0, \dots, t_n$ .

Betrachten wir nun wieder Daten aus den  $n + 1 \in \mathbb{N}$  Jahren  $t_0, t_1, \dots, t_n$ . Nun wollen wir die Maximum-Likelihood-Schätzer für  $\mu_{LC}$  und  $\sigma_{LC}^2$  konstruieren. Da die Zufallsfehler  $Z_t$  standardnormalverteilt sind, ergibt sich die bedingte Verteilung von  $\hat{\kappa}_t^{(2)}$  gegeben  $\hat{\kappa}_{t-1}^{(2)}$  als

$$\hat{\kappa}_t^{(2)} | \hat{\kappa}_{t-1}^{(2)} \sim \mathcal{N}(\hat{\kappa}_{t-1}^{(2)} + \mu_{LC}, \sigma_{LC}^2)$$

für alle  $t = t_0, \dots, t_n$ . Dessen Dichte  $f(\hat{\kappa}_t^{(2)} | \hat{\kappa}_{t-1}^{(2)})$  hat somit die Form

$$f(\hat{\kappa}_t^{(2)} | \hat{\kappa}_{t-1}^{(2)}) = \frac{1}{\sqrt{2\pi\sigma_{LC}^2}} \exp\left(-\frac{(\hat{\kappa}_t^{(2)} - \hat{\kappa}_{t-1}^{(2)} - \mu_{LC})^2}{2\sigma_{LC}^2}\right).$$

Die Likelihoodfunktion ergibt sich somit als

$$\begin{aligned} L(\mu_{LC}, \sigma_{LC}^2; \hat{\kappa}^{(2)}) &= f(\hat{\kappa}_{t_0}^{(2)}, \dots, \hat{\kappa}_{t_n}^{(2)}), \\ &= f(\hat{\kappa}_{t_0}^{(2)}) f(\hat{\kappa}_{t_1}^{(2)} | \hat{\kappa}_{t_0}^{(2)}) \cdots f(\hat{\kappa}_{t_n}^{(2)} | \hat{\kappa}_{t_{n-1}}^{(2)}), \\ &= f(\hat{\kappa}_{t_0}^{(2)}) \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_{LC}^2}} \exp\left(-\frac{(\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC})^2}{2\sigma_{LC}^2}\right), \\ &= f(\hat{\kappa}_{t_0}^{(2)}) (2\pi\sigma_{LC}^2)^{-\frac{n}{2}} \prod_{i=1}^n \exp\left(-\frac{(\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC})^2}{2\sigma_{LC}^2}\right), \end{aligned}$$

wobei die Dichte von  $\hat{\kappa}_{t_0}^{(2)}$  nicht von  $\mu_{LC}$  und  $\sigma_{LC}^2$  abhängt. Wir erhalten nun die log-Likelihood als

$$\begin{aligned} l(\mu_{LC}, \sigma_{LC}^2; \hat{\kappa}^{(2)}) &= \log(f(\hat{\kappa}_{t_0}^{(2)})) - \frac{n}{2} \log(2\pi\sigma_{LC}^2) - \sum_{i=1}^n \frac{(\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC})^2}{2\sigma_{LC}^2}, \\ &= \log(f(\hat{\kappa}_{t_0}^{(2)})) - \frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma_{LC}^2) - \frac{1}{2\sigma_{LC}^2} \sum_{i=1}^n (\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC})^2. \end{aligned}$$

Um die log-Likelihood zu maximieren bilden wir die Ableitungen nach  $\mu_{LC}$  und  $\sigma_{LC}^2$  und setzen diese gleich 0. Wir erhalten also als Ableitung nach  $\mu_{LC}$

$$\frac{\partial l(\mu_{LC}, \sigma_{LC}^2; \hat{\kappa}^{(2)})}{\partial \mu_{LC}} = -\frac{1}{2\sigma_{LC}^2} \sum_{i=1}^n 2(\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC})(-1).$$

Nun ergibt sich

$$\begin{aligned}
-\frac{1}{2\sigma_{LC}^2} \sum_{i=1}^n 2(\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC})(-1) &= 0, \\
\sum_{i=1}^n (\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC}) &= 0, \\
\sum_{i=1}^n (\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)}) - n\mu_{LC} &= 0.
\end{aligned}$$

Wir erhalten nun den Maximum-Likelihood-Schätzer für  $\mu_{LC}$  als

$$\hat{\mu}_{LC} = \frac{1}{n} \sum_{i=1}^n (\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)}) = \frac{\hat{\kappa}_{t_n}^{(2)} - \hat{\kappa}_{t_0}^{(2)}}{n}. \quad (3.24)$$

Betrachten wir nun die Ableitung nach  $\sigma_{LC}^2$  welche folgende Form hat

$$\frac{\partial l(\mu_{LC}, \sigma_{LC}^2; \hat{\kappa}^{(2)})}{\partial \sigma_{LC}^2} = -\frac{n}{2\sigma_{LC}^2} - \frac{1}{2\sigma_{LC}^4} (-1) \sum_{i=1}^n (\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC})^2.$$

Durch Nullsetzen ergibt sich

$$\begin{aligned}
-\frac{n}{2\sigma_{LC}^2} + \frac{1}{2\sigma_{LC}^4} \sum_{i=1}^n (\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC})^2 &= 0, \\
-n\sigma_{LC}^2 + \sum_{i=1}^n (\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \mu_{LC})^2 &= 0.
\end{aligned}$$

Somit erhalten wir als Maximum-Likelihood-Schätzer

$$\hat{\sigma}_{LC}^2 = \frac{1}{n} \sum_{i=1}^n (\hat{\kappa}_{t_i}^{(2)} - \hat{\kappa}_{t_{i-1}}^{(2)} - \hat{\mu}_{LC})^2. \quad (3.25)$$

Man erkennt also, dass der Schätzer für  $\hat{\mu}_{LC}$  nur von der ersten und der letzten Beobachtung abhängt. Somit wird offensichtlich das gewählte Zeitfenster der Beobachtungen großen Einfluss auf den Wert dieses Schätzers haben.

Wir wollen nun die Sterbewahrscheinlichkeit  $q(T, x)$  für die Maturität  $T$  und ein bestimmtes Alter  $x$  schätzen. Diese ist für fixes  $T$  und  $x$  wieder eine Zufallsvariable. Wir nehmen an, dass  $T = t_n + \Delta t$ , wobei  $\Delta t > 0$  und  $t_n$  das derzeitige Jahr bezeichnet. Dann ergibt sich die Verteilung von  $\hat{\kappa}_T^{(2)}$  gegeben  $\hat{\kappa}_{t_n}^{(2)}$  als

$$\hat{\kappa}_T^{(2)} | \hat{\kappa}_{t_n}^{(2)} \sim \mathcal{N}(\hat{\kappa}_{t_n}^{(2)} + \mu_{LC} \Delta t, \sigma_{LC}^2 \Delta t). \quad (3.26)$$

Dies folgt aus

$$\hat{\kappa}_T^{(2)} = \hat{\kappa}_{t_n}^{(2)} + \mu_{LC}\Delta t + \sigma_{LC} \sum_{i=t_n+1}^T Z_i$$

und der Tatsache, dass sich die Summe von unabhängigen Normalverteilungen als

$$\sum_{i=t_n+1}^T Z_i \sim \mathcal{N}(0, \Delta t)$$

ergibt.

Wir können nun Stichproben für die Verteilung von  $q(T, x)$  generieren, indem wir  $Y_i$  aus einer  $\mathcal{N}(\hat{\kappa}_{t_n}^{(2)} + \hat{\mu}_{LC}\Delta t, \hat{\sigma}_{LC}^2\Delta t)$  Verteilung ziehen und

$$m_i = \exp(\hat{\beta}_x^{(1)} + \hat{\beta}_x^{(2)}Y_i) , \quad (3.27)$$

$$q_i = 1 - \exp(-m_i) \quad (3.28)$$

setzen. Dies wird zur Berechnung der in Kapitel 3.2 vorgestellten Bepreisungsmethoden benötigt.

## Das Lee-Carter-Modell mit ARIMA

Eine weitere Möglichkeit die zeitabhängige Funktion  $\hat{\kappa}_t^{(2)}$  im Lee-Carter-Modell (1992) zu beschreiben, ist durch ein ARIMA-Zeitreihen-Modell. ARIMA steht hierbei für *autoregressive integrated moving average*.

Allgemein ist ein ARIMA( $p, d, q$ )-Modell für eine Zeitreihe  $(X_t)$ ,  $t \in \mathbb{Z}$  folgenderweise definiert

$$\left(1 - \sum_{i=1}^p \phi_i B^i\right)(1 - B)^d X_t = c + \left(1 + \sum_{i=1}^q \theta_i B^i\right)\epsilon_t , \quad (3.29)$$

wobei die Modellparameter  $p, d, q$  aus den natürlichen Zahlen gewählt werden. Weiters bezeichnet  $B$  den Backshift-Operator. Für diesen gilt  $B^i X_t = X_{t-i}$  für alle  $i \in \mathbb{N}$  und  $t \in \mathbb{Z}$ . Die Parameter  $\phi_i$ ,  $i = 1, \dots, p$  beschreiben den autoregressiven Teil des Modells, während die Parameter  $\theta_i$ ,  $i = 1, \dots, q$  die moving average modellieren. Die Konstanten  $\theta_i$  bestimmen also die Gewichtung des gleitenden Mittelwerts (moving average) der Rauschterme  $\epsilon_t$ . Diese Störterme  $\epsilon_t$  werden hierbei durch ein weißes Rauschen mit Mittelwert 0 und Varianz  $\sigma^2$  beschrieben. Weiters sei  $c$  eine reelle Konstante. Außerdem nehmen wir an, dass die Polynome  $\phi(z) = 1 - \sum_{i=1}^p \phi_i z^i$  und  $\theta(z) = 1 + \sum_{i=1}^q \theta_i z^i$  keine Nullstellen  $z$  mit  $|z| < 1$  besitzen.

Der vorhin betrachtete Random Walk mit Drift ist ein Spezialfall des durch die mögliche Wahl der Parameter viel allgemeineren ARIMA-Modells. Konkret handelt es sich um ein



ARIMA(0,1,0)-Modell. Dieses ergibt sich als

$$(1 - B)\hat{\kappa}_t^{(2)} = c + \epsilon_t$$

und somit gilt

$$\hat{\kappa}_t^{(2)} - \hat{\kappa}_{t-1}^{(2)} = c + \epsilon_t .$$

Es entspricht also die Konstante  $c$  dem früheren Drift-Parameter  $\mu_{LC}$  und der Rauschterm  $\epsilon_t$  der skalierten Zufallsvariable  $\sigma_{LC}Z_t$  welche nun  $\mathcal{N}(0, \sigma_{LC}^2)$  verteilt ist.

Um das ARIMA-Modell an die Daten zu fitten, folgen wir der Vorgehensweise von Hyndman & Khandakar (2008). Das Hauptproblem liegt hierbei in der Bestimmung der Modellparameter  $p$ ,  $d$  und  $q$ .

Falls  $d$  bekannt ist, können wir  $Y_t = (1 - B)^d X_t$  setzen und das Modell

$$(1 - \sum_{i=1}^p \phi_i B^i) Y_t = c + (1 + \sum_{i=1}^q \theta_i B^i) \epsilon_t \quad (3.30)$$

betrachten. Hier entspricht  $Y_t$  der  $d$ -fachen Differenzenbildung der Zeitreihe  $X_t$  und das neue Modell (3.30) einem ARMA( $p, q$ )-Modell. Um einem ARMA-Modell entsprechen zu können, muss  $Y_t$  jedoch schwach stationär sein.

Eine Zeitreihe ( $X_t$ ) heißt schwach stationär, falls

- $\mathbb{E}[X_t] = \mu$  konstant für alle  $t$  ist und
- die Kovarianzfunktion  $\gamma(t, t+h) = \text{Cov}(X_t, X_{t+h})$  unabhängig von  $t$  ist, d.h  $\gamma(t, t+h) = \gamma(s, s+h)$  gilt.

Der Parameter  $d$  entspricht somit der Anzahl der zu bildenden Differenzen um  $Y_t$  stationär werden zu lassen. Um  $d$  zu bestimmen, führen wir einen Unit-Root-Test durch. Wir wählen hier nun den Kwiatkowski-Phillips-Schmidt-Shin-Test, der auch als KPSS-Test bekannt ist. Die Nullhypothese dieses Tests ist die Stationarität der betrachteten Zeitreihe. Wir suchen also nach Evidenz um die Nullhypothese zu verwerfen. Falls der KPSS-Test ergibt, dass die Nullhypothese verworfen wird und wir somit keine stationäre Zeitreihe vorliegen haben, bilden wir die Differenzen  $X_t - X_{t-1}$  und führen nun für die neue Zeitreihe erneut einen KPSS-Test durch. Dieses Vorgehen wiederholen wir solange, bis die durch Differenzenbildung erhaltene Zeitreihe stationär wird. Der Parameter  $d$  wird nun als die Anzahl der verworfenen Nullhypothesen der wiederholten KPSS-Tests gewählt. In der Praxis ist es so gut wie nie notwendig öfter als zweimal Differenzen zu bilden um die Stationarität zu erreichen.

Zur Bestimmung der Parameter  $p$  und  $q$  bei gegebenen  $d$  für den autoregressive beziehungsweise moving average Teil des Modells betrachten wir das Akaike Informationskriterium (AIC)

$$\text{AIC} = -2 \log(L) - 2(p + q + k + 1) , \quad (3.31)$$

wobei  $k = 1$  falls  $c \neq 0$  und  $k = 0$  sonst. Weiters ist  $L$  die maximierte Likelihood, des an die durch  $d$ -fache Differenzenbildung erhaltenen Daten  $Y_t = (1 - B)^d X_t$  gefittete ARMA( $p, q$ )-Modell. Der Ausdruck in Klammern des zweiten Summanden entspricht der Anzahl der betrachteten Parameter im Modell, wobei auch die Varianz  $\sigma^2$  des weißen Rauschens mitgezählt wird.

Wir suchen nun Parameter  $p$  und  $q$  die das AIC minimieren. Da der Rechenaufwand explodieren würde, wenn man alle möglichen Kombinationen ausprobiert, wählen wir hier einen schrittweisen Ansatz zur Bestimmung der Parameter  $p, d$  und  $q$ .

*Der Hyndman-Khandakar-Algorithmus (2008)*

1. Wir bestimmen die Anzahl der notwendigen Differenzenbildungen  $d$  mithilfe eines wiederholten KPSS-Tests. Wir beschränken uns hierbei jedoch auf  $0 \leq d \leq 2$ .
2. Nun bestimmen wir die Parameter  $p$  und  $q$  durch Minimierung des durch (3.31) gegebenen AIC. Wir beschränken uns bei der Wahl von  $p$  und  $q$  im Laufe des gesamten Algorithmus auf  $0 \leq p, q \leq 5$ . Anstelle alle möglichen Kombinationen für  $p$  und  $q$  durchzuprobieren, gehen wir nun schrittweise vor.

(a) Zunächst fitten wir folgende vier initiale Modelle

- ARIMA(0, $d$ ,0)
- ARIMA(2, $d$ ,2)
- ARIMA(1, $d$ ,0)
- ARIMA(0, $d$ ,1)

Falls  $d \leq 1$  gilt, betrachten wir diese vier Modelle jeweils mit einer Konstanten  $c \neq 0$ . Zusätzlich fitten wir in diesem Fall auch ein ARIMA(0, $d$ ,0)-Modell ohne Konstante.

(b) Wir wählen aus den in Schritt 2.(a) betrachteten Modellen jenes mit kleinsten AIC aus.

(c) Nun variieren wir das in Schritt 2.(b) gewählte Modell, welches wir als das "derzeitige Modell" bezeichnen, auf folgende Weise:

- Wir ändern entweder  $p$  oder  $q$  um  $\pm 1$ .
- Wir ändern  $p$  und  $q$  um  $\pm 1$ .
- Wir inkludieren eine Konstante  $c \neq 0$ , falls keine vorhanden ist oder entfernen die Konstante  $c$ , falls sie im derzeitigen Modell vorkommt.

Falls eine dieser Variationen ein niedrigeres AIC liefert, wird diese als das neue "derzeitige Modell" gewählt.

(d) Wir wiederholen den Schritt 2.(c) solange, bis wir keine weiteren Verbesserungen des AIC finden können.

Dieser Algorithmus findet nicht mit Sicherheit das beste Modell, allerdings findet er effizient ein gut für weitere Prognosen geeignetes Modell. Nachdem wir nun die Ordnung des ARIMA-Modells gefunden haben, müssen wir nun die Parameter  $c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$  schätzen. Dafür führen wir eine Maximum-Likelihood-Schätzung durch. Somit haben wir eine Möglichkeit gefunden ein ARIMA-Modell an die gegebene Zeitreihe  $\hat{\kappa}_t^{(2)}$  zu fitten.

### 3.3.2 Das Cairns-Blake-Dowd-Modell

Ein moderneres Modell für die Mortalität ist das Cairns-Blake-Dowd-Modell, welches ab dem Jahr 2006 entwickelt wurde. Anders als im älteren Lee-Carter-Modell (1992) wird hier nicht  $m(t, x)$  betrachtet, sondern der Logit-Wert von  $q(t, x)$  direkt modelliert. Dieser ist gegeben als

$$\log \left( \frac{q(t, x)}{1 - q(t, x)} \right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) , \quad (3.32)$$

wobei  $\bar{x}$  das Durchschnittsalter im betrachteten Sample bezeichnet. Zur besseren Vergleichbarkeit mit dem Lee-Carter-Modell formulieren wir das Cairns-Blake-Dowd-Modell (2009) nun als

$$\log \left( \frac{q(t, x)}{1 - q(t, x)} \right) = \beta_x^{(1)} \kappa_t^{(1)} + \beta_x^{(2)} \kappa_t^{(2)} , \quad (3.33)$$

wobei  $\beta_x^{(1)} \equiv 1$  und  $\beta_x^{(2)} = x - \bar{x}$ . Hier tritt kein Identifikationsproblem auf und somit sind keine weiteren Nebenbedingungen notwendig. Anders als im Lee-Carter-Modell haben wir nun zwei zeitabhängige Funktionen  $\kappa_t^{(1)}$  und  $\kappa_t^{(2)}$ .

Wie schon im vorigen Abschnitt benötigen wir nun Schätzer für  $\kappa_t^{(1)}$  und  $\kappa_t^{(2)}$ . Seien die Daten wie im Lee-Carter-Modell gegeben. Im Folgenden verwenden wir die Notation

$$C(t, x) = \log \left( \frac{q(t, x)}{1 - q(t, x)} \right) .$$

Dann ergeben sich für ein festes Jahr  $t$  aus der Modelldefinition (3.32) die  $m + 1$  Gleichungen der Form

$$C(t, x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x})$$

für  $x = x_0, \dots, x_m$ .

Da wir in der Praxis von  $m \geq 1$  ausgehen können, bietet sich hier der Kleinste-Quadrate-Schätzer des einfachen linearen Regressionsmodells  $y_i = a + bx_i$  an, wobei  $\kappa_t^{(1)}$  und  $\kappa_t^{(2)}$  für jedes  $t = t_0, \dots, t_n$  die Rolle der Koeffizienten  $a$  und  $b$  übernehmen. Somit erhalten wir

zunächst für festes  $t$

$$\begin{aligned}\hat{\kappa}_t^{(2)} &= \frac{\sum_{i=0}^m ((x_i - \bar{x}) - \overline{(x - \bar{x})})(C(t, x_i) - \overline{C(t, x)})}{\sum_{i=0}^m ((x_i - \bar{x}) - \overline{(x - \bar{x})})^2} \\ &= \frac{\sum_{i=0}^m ((x_i - \bar{x}) - \overline{(x - \bar{x})})C(t, x_i)}{\sum_{i=0}^m ((x_i - \bar{x}) - \overline{(x - \bar{x})})^2}.\end{aligned}$$

Da  $\overline{(x - \bar{x})} = 0$  gilt, ergibt sich der KQ-Schätzer für  $\kappa_t^{(2)}$  als

$$\hat{\kappa}_t^{(2)} = \frac{\sum_{i=0}^m (x_i - \bar{x})C(t, x_i)}{\sum_{i=0}^m (x_i - \bar{x})^2} \quad (3.34)$$

Weiters finden wir für  $t = t_0, \dots, t_n$

$$\begin{aligned}\hat{\kappa}_t^{(1)} &= \overline{C(t, x)} - \hat{\kappa}_t^{(2)} \overline{(x - \bar{x})} \\ &= \overline{C(t, x)}\end{aligned}$$

und der KQ-Schätzer für  $\kappa_t^{(1)}$  hat somit die Form

$$\hat{\kappa}_t^{(1)} = \frac{1}{m+1} \sum_{i=0}^m C(t, x_i). \quad (3.35)$$

Nun folgen wir dem Ansatz von Cairns et al. (2006) und modellieren die Entwicklung der  $\hat{\kappa}_t^{(1)}$  und  $\hat{\kappa}_t^{(2)}$  mit einem zweidimensionalen Random-Walk-Modell mit Drift. Somit erhalten wir

$$\begin{pmatrix} \hat{\kappa}_t^{(1)} \\ \hat{\kappa}_t^{(2)} \end{pmatrix} = \begin{pmatrix} \hat{\kappa}_{t-1}^{(1)} \\ \hat{\kappa}_{t-1}^{(2)} \end{pmatrix} + \mu_{CBD} + \sigma_{CBD} Z_t, \quad (3.36)$$

wobei  $\mu_{CBD} \in \mathbb{R}^2$ ,  $\sigma_{CBD} \in \mathbb{R}^{2 \times 2}$  und  $Z_t$  unabhängige und identisch verteilte Zufallsvektoren mit einer zweidimensionalen Standardnormalverteilung sind. Im Folgenden bezeichnen wir den Vektor  $\begin{pmatrix} \hat{\kappa}_t^{(1)} \\ \hat{\kappa}_t^{(2)} \end{pmatrix}$  mit  $\hat{\kappa}_t$  um die Notation zu erleichtern.

Um  $\mu_{CBD}$  und  $\sigma_{CBD}$  zu schätzen, verwenden wir nun die zweidimensionale Version der Maximum-Likelihood-Schätzer aus Kapitel 3.3.1. Diese ergeben sich als

$$\hat{\mu}_{CBD} = \frac{1}{n} (\hat{\kappa}_{t_n} - \hat{\kappa}_{t_0}), \quad (3.37)$$

$$\hat{\sigma}_{CBD} \hat{\sigma}_{CBD}^\top = \frac{1}{n} \sum_{i=1}^n (\hat{\kappa}_{t_i} - \hat{\kappa}_{t_{i-1}} - \hat{\mu}_{CBD})(\hat{\kappa}_{t_i} - \hat{\kappa}_{t_{i-1}} - \hat{\mu}_{CBD})^\top. \quad (3.38)$$

Unser Ziel ist es wieder  $q(T, x)$  zu schätzen. Dafür ist es notwendig  $\hat{\kappa}_T$  zu bestimmen. Wieder ergibt sich die Verteilung von  $\hat{\kappa}_T$  gegeben  $\hat{\kappa}_{t_n}$  als Normalverteilung und zwar konkret als

$$\hat{\kappa}_T | \hat{\kappa}_{t_n} \sim \mathcal{N}_2(\hat{\kappa}_{t_n} + \mu_{CBD} \Delta t, \sigma_{CBD} \sigma_{CBD}^\top \Delta t). \quad (3.39)$$

Dies ergibt sich aus der Tatsache, dass  $\hat{\kappa}_T$  als

$$\hat{\kappa}_T = \hat{\kappa}_{t_n} + \mu_{CBD}\Delta t + \sigma_{CBD} \sum_{i=t_n+1}^T Z_i$$

geschrieben werden kann und die Summe der unabhängigen Normalverteilungen wieder normalverteilt ist. Konkret ergibt sich

$$\sum_{i=t_n+1}^T Z_i \sim \mathcal{N}_2(0, \Delta t \mathbf{I}_2) .$$

Nun können wir zweidimensionale Stichproben  $Y_1, Y_2, \dots$  aus einer  $\mathcal{N}_2(\hat{\kappa}_{t_n} + \hat{\mu}_{CBD}\Delta t, \hat{\sigma}_{CBD}\hat{\sigma}_{CBD}^\top \Delta t)$  Verteilung ziehen und anschließend

$$\log\left(\frac{q_i}{1 - q_i}\right) = Y_i^{(1)} + Y_i^{(2)}(x - \bar{x})$$

berechnen. Durch Umformen ergibt sich

$$q_i = \frac{\exp(Y_i^{(1)} + Y_i^{(2)}(x - \bar{x}))}{1 - \exp(Y_i^{(1)} + Y_i^{(2)}(x - \bar{x}))} . \quad (3.40)$$

Somit haben wir wieder eine Möglichkeit gefunden, um Stichproben für  $q(T, x)$  zu generieren.

# Kapitel 4

## Empirischer Vergleich der verschiedenen Preisfindungsmodelle

Um die in Kapitel 3.2 vorgestellten Bepreisungsmethoden vergleichen zu können, führen wir eine empirische Studie durch. Dafür verwenden wir zunächst demographische Daten aus Österreich, die über die Human Mortality Database (HMD) verfügbar sind. Die HMD bezieht diese Daten von der Statistik Austria und bezeichnet die Datenqualität der letzten 50 Jahre als sehr gut. Die benötigten Sterberaten sind für Österreich für die Jahre 1947 bis 2014 verfügbar. Im weiteren Verlauf folgen wir dem konzeptionellen Aufbau der Studie von Barrieu & Veraart (2016) und verwenden nur Daten für die männliche Bevölkerung. Weiters beschränken wir uns daher auf die Jahre ab 1961 und die Altersklasse der 60- bis 89-jährigen. Für den internationalen Vergleich führen wir außerdem dieselben Berechnungen auch für Daten aus England & Wales durch. Auch diese Daten sind über die HMD verfügbar und weisen eine gute Qualität auf. Auf diesen Datensatz wenden wir dieselben Einschränkungen bezüglich betrachtetem Geschlecht, Jahren und Altersklassen an wie auf den Datensatz aus Österreich.

In beiden Datensätzen haben wir die rohen Sterberaten für jedes Alter und Jahr gegeben. Aus diesen können wir mithilfe der Formel (3.3) die Sterbewahrscheinlichkeiten berechnen. In Abbildung 4.1 können wir deren Verlauf für die betrachteten Jahre und Altersklassen für Österreich sehen, wobei sich für die Daten aus England und Wales ein ähnliches Bild ergibt. Wir erkennen eine Abnahme der Sterbewahrscheinlichkeit über die Jahre und dass diese für höhere Alter stärker ausgeprägt ist. Weiters ist wie erwartet die Sterbewahrscheinlichkeit für höhere Alter höher. Neben dem beschriebenen allgemeinen Trend gibt es auch eine stochastische Komponente, welche am nicht glatten Verlauf erkennbar ist.

Für die praktische Durchführung der empirischen Studie verwenden wir das Programm R und insbesondere die Pakete `demography` und `forecast`. Der vollständige Code kann im Anhang eingesehen werden.

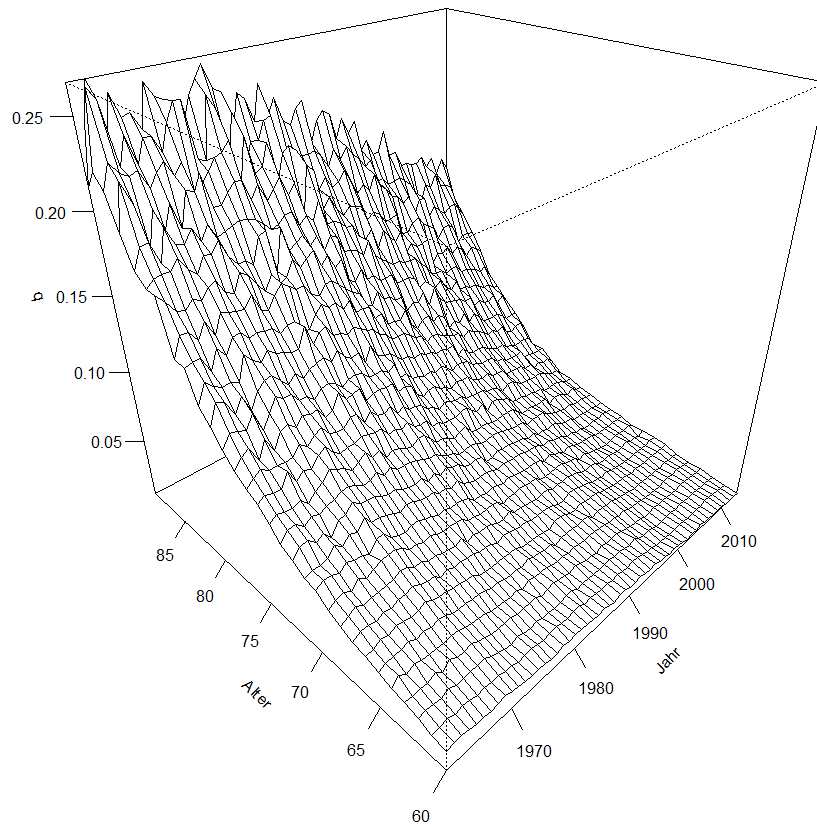


Abbildung 4.1: Beobachtete Sterbewahrscheinlichkeiten  $q(t, x)$  für Männer mit Alter  $x \in \{60, \dots, 89\}$  in den Jahren  $t \in \{1961, \dots, 2014\}$  in Österreich

## 4.1 Aufbau der empirischen Analyse

Das Ziel der empirischen Analyse ist also der Vergleich der verschiedenen Bepreisungsmethoden. Dafür bestimmen wir die Preise der q-Forwards und deren Konfidenzintervalle unter Berücksichtigung verschiedener Mortalitätsmodelle und Datenausschnitte. Im Folgenden verwenden wir:

- Daten aus zwei verschiedenen Ländern:
  - Österreich
  - England und Wales
- vier verschiedene Bepreisungsmethoden:

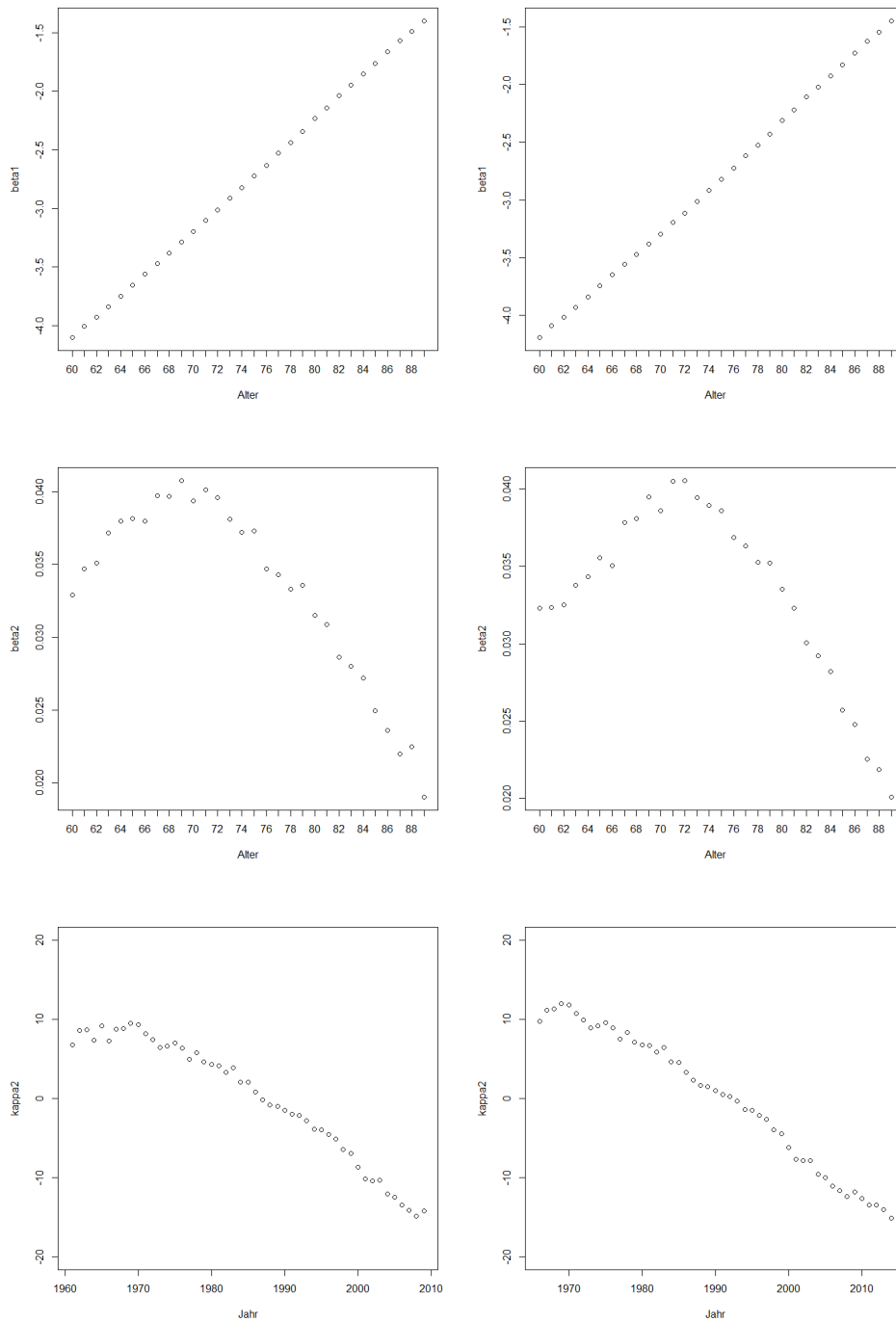
- das Nettoprämienprinzip (Net)
- das Standardabweichungsprinzip mit Parameter  $\lambda = -0,1$  (StD)
- das Nullnutzenprinzip mit  $\gamma z = 1$  (ZU1)
- das Nullnutzenprinzip mit  $\gamma z = 10000$  (ZU2)
- drei verschiedene Mortalitätsmodelle:
  - das Lee-Carter-Modell mit Random Walk (LCRW)
  - das Lee-Carter-Modell mit ARIMA (LCAR)
  - das Cairns-Blake-Dowd-Modell mit Random Walk (CBD)
- zwei verschiedene Zeiträume für die Parameterschätzung der Mortalitätsmodelle:
  - die Jahre 1961-2009
  - die Jahre 1966-2014
- jeweils zwei verschiedene Zeitfenster für die Schätzung der Zeitreihenmodelle:
  - 6 Jahre (2004-2009 bzw. 2009-2014)
  - 21 Jahre (1989-2009 bzw. 1994-2014)
- zwei verschiedene Maturitäten  $T$ :
  - nach 10 Jahren
  - nach 30 Jahren
- zwei verschiedene zugrundeliegende Alter  $x$  für die q-Forwards:
  - 60 Jahre
  - 70 Jahre

## 4.2 Fitten der Mortalitätsmodelle an die Daten

Zunächst wollen wir das Lee-Carter-Modell (1992) an die österreichischen Daten fitten. Wir verwenden dafür zwei Zeitfenster von jeweils 49 Jahren, nämlich die Jahre 1961-2009 und 1966-2014. Nun schätzen wir  $\beta_x^{(1)}$ ,  $\beta_x^{(2)}$  und  $\kappa_t^{(2)}$  mithilfe der im Kapitel 3.3.1 beschriebenen Methode für beide Zeitfenster. Deren Verlauf ist in der Abbildung 4.2 graphisch dargestellt.

Hier erkennen wir, dass  $\hat{\beta}_x^{(1)}$  mit dem Alter  $x$  wächst, was einer höheren Sterblichkeit für höhere Alter entspricht. In beiden Fällen weist  $\hat{\beta}_x^{(1)}$  beinahe eine lineare Struktur auf. Weiters fällt  $\hat{\kappa}_t^{(2)}$  im Laufe der Zeit  $t$ , was darauf schließen lässt, dass es eine zeitabhängige Verbesserung der Mortalität gibt. Betrachten wir nun die Entwicklung von  $\hat{\beta}_x^{(2)}$ , so erkennen





(a)  $\hat{\beta}_x^{(1)}$ ,  $\hat{\beta}_x^{(2)}$  und  $\hat{\kappa}_t^{(2)}$  für die Jahre 1961-2009

(b)  $\hat{\beta}_x^{(1)}$ ,  $\hat{\beta}_x^{(2)}$  und  $\hat{\kappa}_t^{(2)}$  für die Jahre 1966-2014

Abbildung 4.2: Verlauf der Parameterschätzungen für das Lee-Carter-Modell für die männliche Bevölkerung Österreichs.

wir, dass diese in beiden Fällen bis etwa zum Alter 70 eine steigende Tendenz aufweist und anschließend abfällt. Die Verbesserung der Sterblichkeit über die Zeit ist also für die Altersklassen um 70 am stärksten spürbar und nimmt anschließend mit dem Alter ab.

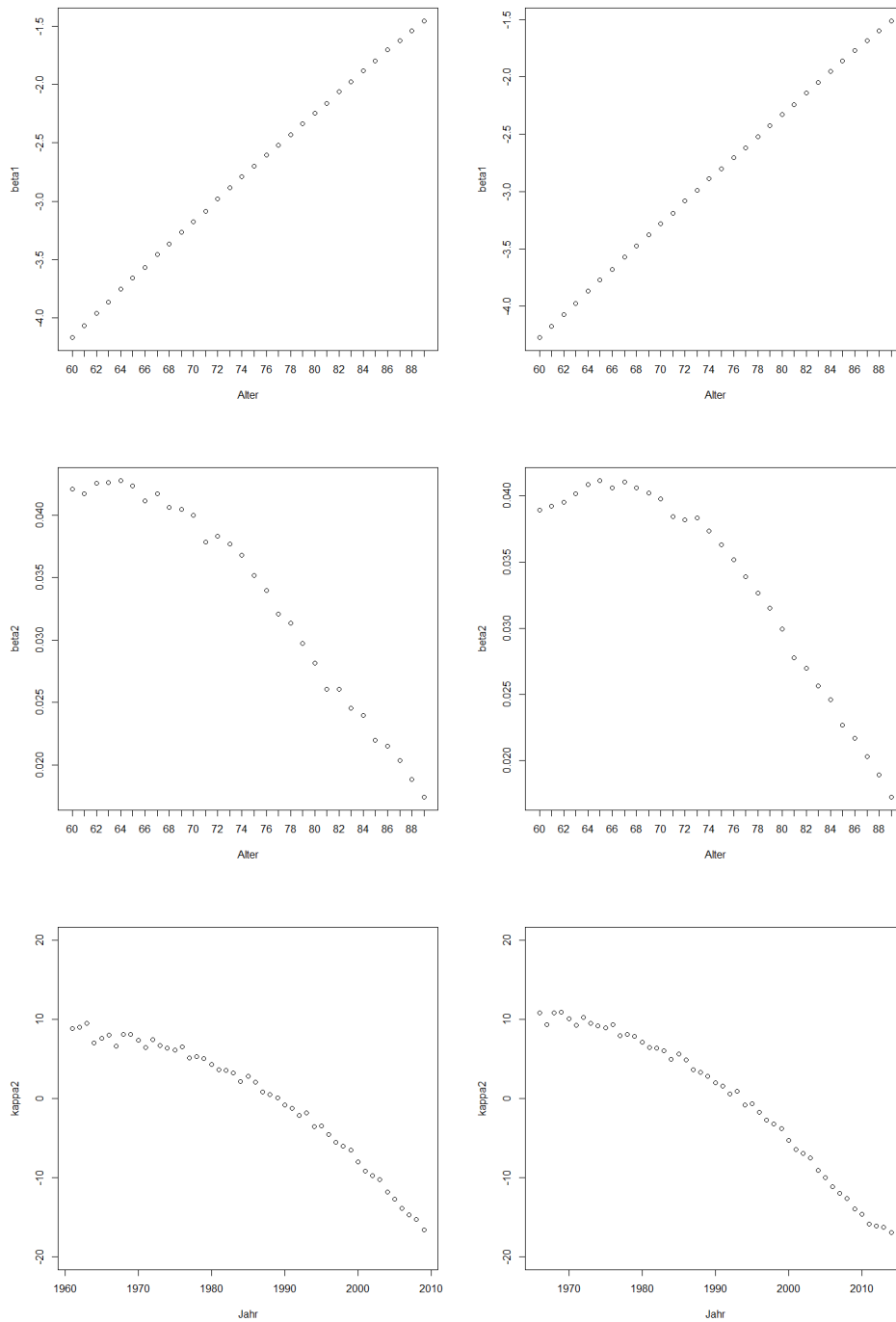
Weiters betrachten wir die Fits der Daten aus England und Wales an das Lee-Carter-Modell. Die Verläufe der Parameterschätzer  $\hat{\beta}_x^{(1)}$ ,  $\hat{\beta}_x^{(2)}$  und  $\hat{\kappa}_t^{(2)}$  sind in Abbildung 4.3 graphisch dargestellt. Die Struktur der Verläufe von  $\hat{\beta}_x^{(1)}$  und  $\hat{\kappa}_t^{(2)}$  ähneln den Verläufen der aus den österreichischen Daten erhaltenen Schätzer. Für  $\hat{\beta}_x^{(2)}$  erkennen wir jedoch, dass der bei den österreichischen Daten stark ausgeprägte Buckel um das Alter 70 fehlt. Für den Beobachtungszeitraum 1961-2009 ist der Verlauf von  $\hat{\beta}_x^{(2)}$  anfangs flach und fällt ab dem Alter von 65 Jahren ab. Betrachten wir Daten aus den Jahren 1966-2014, so erkennen wir Anzeichen einer Entwicklung wie bei den österreichischen Daten, welche allerdings deutlich schwächer ausgeprägt ist.

Nun betrachten wir die Schätzungen für  $\kappa_t^{(2)}$  als Zeitreihe. Zunächst wollen wir diese als Random Walk mit Drift  $\mu_{LC}$  und Standardabweichung  $\sigma_{LC}$  modellieren. Betrachten wir die Formeln (3.24) und (3.25) für die Schätzer  $\hat{\mu}_{LC}$  und  $\hat{\sigma}_{LC}^2$  so erkennen wir, dass die Wahl des Schätzfensters eine entscheidende Rolle spielen wird. Die konkreten Werte für beide Länder, Beobachtungszeiträume und Schätzfenster sind in Tabelle 4.1 dargestellt. Der Drift-Parameter  $\hat{\mu}_{LC}$  kann je nach Wahl des Schätzfensters deutliche Unterschiede aufweisen. Dies ist dann der Fall, wenn die Entwicklung von  $\hat{\kappa}_t^{(2)}$  in den letzten 6 Jahren deutlich steiler oder flacher war als in den letzten 21 Jahren. Betrachten wir etwa die österreichischen Daten für den Beobachtungszeitraum 1961-2009, so erkennen wir, dass der Verlauf der  $\hat{\kappa}_t^{(2)}$  in den Jahren 2004-2009 abflacht. Im Gegensatz dazu ist der Verlauf der  $\hat{\kappa}_t^{(2)}$  für die Daten aus England und Wales für denselben Beobachtungszeitraum in den letzten 6 Jahren steiler als in den letzten 21 Jahren. Für den Beobachtungszeitraum 1966-2014 ergeben sich für die österreichischen Daten keine nennenswerten Unterschiede in der Drift bei der Betrachtung von 6 oder 21 Jahren. Für denselben Beobachtungszeitraum erkennen wir bei den Daten aus England und Wales deutliche Unterschied für die beiden Schätzer  $\hat{\mu}_{LC}$  für die Drift. Diese sind durch eine deutlich flachere Entwicklung der  $\hat{\kappa}_t^{(2)}$  ab 2009 im Vergleich zum Verlauf der letzten 21 Jahre erklärbar.

Im nächsten Schritt modellieren wir die Zeitreihe  $\hat{\kappa}_t^{(2)}$  mithilfe eines ARIMA-Modells. Wir verwenden zunächst den Hyndman-Khandakar-Algorithmus (2008) um schrittweise ein geeignetes ARIMA( $p, d, q$ )-Modell zu bestimmen. Für die automatische Schätzung in R muss jedoch beachtet werden, dass hier eine andere Parametrisierung als in (3.29) vorliegt. So verwendet R die Notation

$$(1 - \sum_{i=1}^p \phi_i B^i)(1 - B)^d (X_t - \mu t^d / d!) = (1 + \sum_{i=1}^q \theta_i B^i) \epsilon_t .$$

Der Parameter  $c$  im ursprünglichen Modell (3.29) ergibt sich als  $c = \mu(1 - \sum_{i=1}^p \phi_i)$ , wobei  $\mu$  das Mittel von  $(1 - B)^d X_t$  ist.



(a)  $\hat{\beta}_x^{(1)}$ ,  $\hat{\beta}_x^{(2)}$  und  $\hat{\kappa}_t^{(2)}$  für die Jahre 1961-2009

(b)  $\hat{\beta}_x^{(1)}$ ,  $\hat{\beta}_x^{(2)}$  und  $\hat{\kappa}_t^{(2)}$  für die Jahre 1966-2014

Abbildung 4.3: Verlauf der Parameterschätzungen für das Lee-Carter-Modell für die männliche Bevölkerung England und Wales.

Land	Beobachtungszeitraum	Schätzfenster	$\hat{\mu}_{LC}$	$\hat{\sigma}_{LC}^2$
AUT	1961-2009	1989-2009	-0,6592	0,3426
AUT	1961-2009	2004-2009	-0,4270	0,3137
AUT	1966-2014	1994-2014	-0,6836	0,3508
AUT	1966-2014	2009-2014	-0,6621	0,1367
ENG	1961-2009	1989-2009	-0,8346	0,2547
ENG	1961-2009	2004-2009	-0,9581	0,0697
ENG	1966-2014	1994-2014	-0,8081	0,1997
ENG	1966-2014	2009-2014	-0,5994	0,1509

Tabelle 4.1: Schätzer der Random-Walk-Parameter  $\mu_{LC}$  und  $\sigma_{LC}^2$  im Lee-Carter-Modell

Für die österreichischen Daten ergibt sich für alle Beobachtungszeiträume und Schätzfenster ein ARIMA(0,1,0)-Modell. Dies entspricht dem vorhin betrachteten Random Walk mit Drift. Bei der Betrachtung der Daten aus den Jahren 1961-2009 und dem Schätzfenster 1989-2009 erhalten wir  $\hat{c} = -0,6592$  und  $\hat{\sigma}^2 = 0,3606$ . Beim kürzeren Schätzfenster 2004-2009 erhalten wir  $\hat{c} = -0,4270$  und  $\hat{\sigma}^2 = 0,3922$ . Verwenden wir Daten aus den Jahren 1966-2014 und das Schätzfenster 1994-2014, so ergibt sich  $\hat{c} = -0,6836$  und  $\hat{\sigma}^2 = 0,3692$ . Die Parameter für das ARIMA-Modell beim Schätzfenster 2009-2014 sind  $\hat{c} = -0,6621$  und  $\hat{\sigma}^2 = 0,1709$ . Der Parameter  $c$  entspricht also jeweils der Drift  $\hat{\mu}_{LC}$  im Random-Walk-Modell. Die Unterschiede in der Standardabweichung ergeben sich durch die verschiedenen Schätzer für die Stichprobenvarianz. So wird bei der Schätzung der ARIMA-Modelle die korrigierte Stichprobenvarianz  $\widehat{\text{Var}}(X) = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$  als Schätzer für  $\hat{\sigma}^2$  verwendet. Bei der Schätzung der Varianz im Random-Walk-Modell wird jedoch die nicht korrigierte Version  $\widehat{\text{Var}}(X) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$  verwendet. Wir erkennen somit, dass sich der Schätzer für die Varianz in der Betrachtung als ARIMA(0,1,0)-Modells als  $\hat{\sigma}^2 = \frac{n}{n-1} \hat{\sigma}_{LC}^2$  ergibt.

Für die Daten aus England und Wales erhalten wir für den Betrachtungszeitraum 1961-2009 und das Schätzfenster 1989-2009 ein ARIMA(1,1,0)-Modell mit Drift mit Parametern  $\hat{\phi}_1 = -0,4859$ ,  $\hat{c} = -1,2266$  und  $\hat{\sigma}^2 = 0,2133$ . Für alle anderen drei betrachteten Zeiträume ergibt sich wieder ein ARIMA(0,1,0)-Modell. Für Daten aus den Jahren 1961-2009 und dem Schätzfenster 2004-2009 für die Zeitreihe, erhalten wir die Parameter  $\hat{c} = -0,9581$  und  $\hat{\sigma}^2 = 0,0872$ . Betrachten wir Daten aus den Jahren 1966-2014 und dem Schätzfenster 1994-2014 ergeben sich die Parameter des ARIMA(0,1,0)-Modells als  $\hat{c} = -0,8081$  und

$\hat{\sigma}^2 = 0,2102$ . Für das kürzere Schätzfenster 2009-2014 erhalten wir  $\hat{c} = -0,5994$  und  $\hat{\sigma}^2 = 0,1866$ .

Nun wollen wir überprüfen, ob der schrittweise Ansatz des Hyndman-Khandakar-Algorithmus gute Ergebnisse für die Modellierung der ARIMA-Modelle liefert. Dafür suchen wir nun das optimale ARIMA-Modell nicht mit dem effizienteren schrittweisen Algorithmus, sondern aus allen möglichen ARIMA-Modellen.

Für die Daten aus Österreich findet der Hyndman-Khandakar-Algorithmus hier immer das ARIMA-Modell mit dem kleinsten AIC. Somit ist das zuvor gewählte Random-Walk-Modell trotz seiner verhältnismäßig einfachen Struktur die beste Wahl unter allen betrachteten ARIMA-Modellen. Für die Daten aus England und Wales ergibt sich jedoch eine andere Situation. Hier findet der schrittweise Hyndman-Khandakar-Algorithmus für beide kurzen Schätzfenster von 6 Jahren nicht das optimale ARIMA-Modell. Anstelle des gefundenen ARIMA(0,1,0)-Modells sind es in beiden Fällen ARIMA(2,1,0)-Modelle, die das AIC minimieren. Für das Schätzfenster 2004-2009 unter Verwendung der Daten von 1961-2009 erhalten wir die Parameter  $\hat{\phi}_1 = -0,8171$ ,  $\hat{\phi}_2 = -0,8931$ ,  $\hat{c} = -2,4784$  und  $\hat{\sigma}^2 = 0,0277$ . Betrachten wir Daten aus den Jahren 1966-2014 und das Schätzfenster 2009-2014 so ergeben sich  $\hat{\phi}_1 = -0,2187$ ,  $\hat{\phi}_2 = -0,8928$ ,  $\hat{c} = -1,1527$  und  $\hat{\sigma}^2 = 0,0842$ .

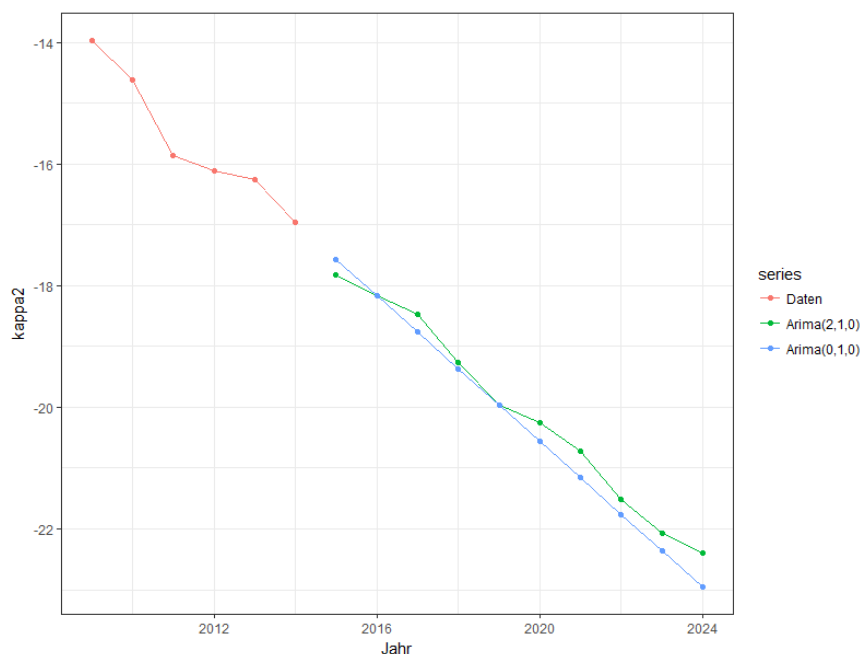


Abbildung 4.4: Vergleich der Forecasts für ARIMA(0,1,0) und ARIMA(2,1,0)

Wir vergleichen nun das ARIMA(0,1,0)-Modell mit dem ARIMA(2,1,0)-Modell für das Schätzfenster 2009-2014. Wir betrachten Forecasts für die nächsten 10 Jahre für beide Model-

le. Beide Modelle weisen einen ähnlichen Trend auf, wie man in Abbildung 4.4 erkennen kann. Auch für das zweite Schätzintervall 2004-2009, für welches der schrittweise Algorithmus nicht das optimale ARIMA-Modell findet, erhalten wir eine ähnliche Struktur. Auch wenn der schrittweise Hyndman-Khandakar-Algorithmus nicht immer das optimale  $ARIMA(p, d, q)$ -Modell findet, verwenden wir diesen im Bootstrap-Algorithmus, der im nächsten Kapitel beschrieben wird. Da dieser mit sehr vielen Iterationen ausgeführt wird, ist hier die Effizienz des Algorithmus vorrangig.

Nachdem wir nun beide Varianten des Lee-Carter-Modells erfolgreich an die Daten gefittet haben, wenden wir uns nun dem CBD-Modell zu. Zunächst müssen wir hier also die Verläufe der beiden zeitabhängigen Funktionen  $\kappa_t^{(1)}$  und  $\kappa_t^{(2)}$  schätzen. Diese sind für Daten aus Österreich in Abbildung 4.5 und für Daten aus England und Wales in Abbildung 4.6 zu finden.

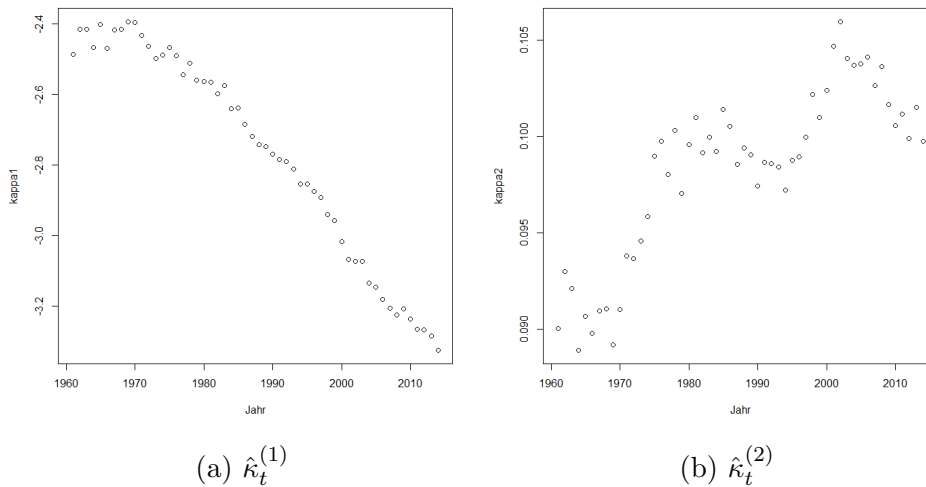


Abbildung 4.5: Verlauf der Parameterschätzungen für das Cairns-Blake-Dowd-Modell für die männliche Bevölkerung Österreichs für die Jahre 1961-2014.

Nachdem die Parameter  $\kappa_t^{(1)}$  und  $\kappa_t^{(2)}$  für jedes Jahr ohne Einfluss des betrachteten Zeitraums geschätzt werden, verändert das Hinzufügen von Jahren am Ende der Periode beziehungsweise das Weglassen von Jahren am Anfang der Periode nichts an den Schätzern für die Jahre dazwischen. Daher genügt es - anders als beim Lee-Carter-Modell - die beiden Zeitfenster 1961-2009 und 1966-2014 im Folgenden als das größere Zeitfenster 1961-2014 zu betrachten.

Wir erkennen nun, dass  $\hat{\kappa}_t^{(1)}$  für beide betrachteten Länder einen fallenden Verlauf aufweist. Anders verhält sich jedoch  $\hat{\kappa}_t^{(2)}$ . Für die englischen Daten ergibt sich eindeutig ein wachsender Verlauf. Für die österreichischen Daten erhalten wir zwar auch insgesamt einen wachsenden Trend, jedoch nicht mit einem monotonen sondern einem wellenförmigen Verlauf.

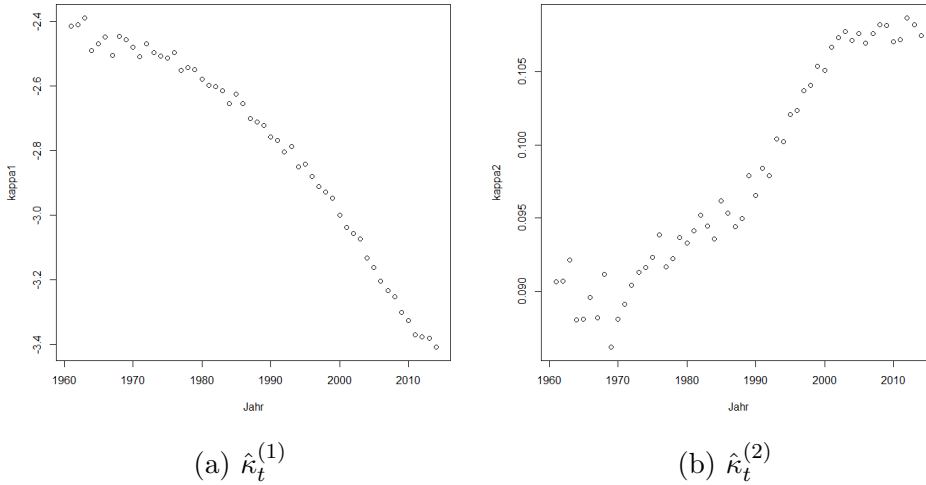


Abbildung 4.6: Verlauf der Parameterschätzungen für das Cairns-Blake-Dowd-Modell für die männliche Bevölkerung England und Wales für die Jahre 1961-2014.

Nun wollen wir die beiden zeitabhängigen Funktionen  $\hat{\kappa}_t^{(1)}$  und  $\hat{\kappa}_t^{(2)}$  als zweidimensionalen Random Walk mit Drift modellieren. Dazu schätzen wir die beiden Parameter  $\mu_{CBD}$  und  $\sigma_{CBD}\sigma_{CBD}^\top$ . Diese Schätzer können in Tabelle 4.2 für alle betrachteten Schätzfenster und beide Länder abgelesen werden. Wir erkennen die fallende Struktur von  $\hat{\kappa}_t^{(1)}$  am durchgehend negativen Vorzeichen des Driftparameters  $\hat{\mu}_{CBD}^{(1)}$  für beide Datensätze. Die wellenförmige Struktur von  $\hat{\kappa}_t^{(2)}$  für die österreichischen Daten spiegelt sich im wechselnden Vorzeichen von  $\hat{\mu}_{CBD}^{(2)}$  wider. Zu beachten ist auch, dass für die Daten aus England und Wales das Wachstum von  $\hat{\kappa}_t^{(2)}$  in den letzten beobachteten Jahren abflacht bzw. sich umkehrt, was das negative Vorzeichen des letzten Schätzwerts für  $\hat{\mu}_{CBD}^{(2)}$  erklärt.

### 4.3 Simulation und Ergebnisse

Im Folgenden wollen wir für alle im Kapitel 3.2 vorgestellten Bepreisungsmethoden konkrete Werte berechnen. Dazu führen wir für alle drei betrachteten Mortalitätsmodelle eine Simulation durch um Werte für die Sterbewahrscheinlichkeit  $q(t, x)$  zu erhalten. Konkret betrachten wir nun jeweils ein festgelegtes Alter  $x \in \{60, 70\}$  und eine fixe Maturität  $T \in \{10, 30\}$ . Mit Hilfe der im vorigen Kapitel gefitteten Modelle simulieren wir nun 10000 Pfade für  $q(t_n + T, x)$ . Dies geschieht im Fall der beiden Modelle die einen Random Walk beinhalten durch zufällig erzeugte Werte einer Normalverteilung und umformen gemäß der Formel (3.28) und (3.40). Für das ARIMA-Zeitreihenmodell ist für jede Parameterbelegung eine automatische Simulation im R-Package `forecast` implementiert, wobei auch hier die zufällige Komponente aus einer Normalverteilung stammt. Aus den 10000 simulierten Werten können wir nun Prei-

Land	Schätzfenster	$\hat{\mu}_{CBD}$	$\hat{\sigma}_{CBD}\hat{\sigma}_{CBD}^T$
AUT	1989-2009	$\begin{pmatrix} -0,0229 \\ 0,0001 \end{pmatrix}$	$\begin{pmatrix} 0,0004 & -0,00001 \\ -0,00001 & 0,000002 \end{pmatrix}$
AUT	2004-2009	$\begin{pmatrix} -0,0147 \\ -0,0004 \end{pmatrix}$	$\begin{pmatrix} 0,0003 & -0,00001 \\ -0,00001 & 0,000001 \end{pmatrix}$
AUT	1994-2014	$\begin{pmatrix} -0,0236 \\ 0,0001 \end{pmatrix}$	$\begin{pmatrix} 0,0004 & -0,00001 \\ -0,00001 & 0,000002 \end{pmatrix}$
AUT	2009-2014	$\begin{pmatrix} -0,0237 \\ -0,0004 \end{pmatrix}$	$\begin{pmatrix} 0,0002 & 0,000003 \\ 0,000003 & 0,000001 \end{pmatrix}$
ENG	1989-2009	$\begin{pmatrix} -0,0290 \\ 0,0005 \end{pmatrix}$	$\begin{pmatrix} 0,0004 & 0,00001 \\ 0,00001 & 0,000009 \end{pmatrix}$
ENG	2004-2009	$\begin{pmatrix} -0,0229 \\ 0,0001 \end{pmatrix}$	$\begin{pmatrix} 0,0001 & 0,000004 \\ 0,000004 & 0,000002 \end{pmatrix}$
ENG	1994-2014	$\begin{pmatrix} -0,0280 \\ 0,0004 \end{pmatrix}$	$\begin{pmatrix} 0,0003 & 0,000007 \\ 0,000007 & 0,000007 \end{pmatrix}$
ENG	2009-2014	$\begin{pmatrix} -0,0214 \\ -0,0001 \end{pmatrix}$	$\begin{pmatrix} 0,0002 & 0,000004 \\ 0,000004 & 0,000009 \end{pmatrix}$

Tabelle 4.2: Schätzer der Random-Walk-Parameter  $\mu_{CBD}$  und  $\sigma_{CBD}\sigma_{CBD}^T$  im Cairns-Blake-Dowd-Modell

se entsprechend der jeweiligen Formeln für das Netto-, das Standardabweichungs- und das Nullnutzenprinzip berechnen.

Wir sind jedoch nicht nur an konkreten Werten für die Preise der q-Forwards interessiert sondern auch an Konfidenzintervallen. Um die symmetrischen 95%-Konfidenzintervalle zu berechnen, führen wir ein Bootstrappingverfahren durch. Dabei wollen wir eine empirische Verteilung von  $q(t_n + T, x)$  erzeugen um anschließend das 2,5%-Quantil und das 97,5%-Quantil ablesen zu können. Dazu ist es notwendig in jedem Durchlauf einen neuen Datensatz zu generieren um einen weiteren Wert für  $q(t_n + t, x)$  bestimmen zu können. Wir bezeichnen im Folgenden die Anzahl der Todesfälle mit Alter  $x$  im Jahr  $t$  mit  $D(t, x)$  und das Exposure-to-Risk zum Alter  $x$  im Jahr  $t$  mit  $E(t, x)$ . Hier gilt nun die Beziehung  $D(t, x) = m(t, x)E(t, x)$ . Nun nehmen wir an, dass die Anzahl der Todesfälle  $D(t, x)$  einer Poisson-Verteilung mit



Parameter  $\lambda(t, x)$  folgt, also

$$\mathbb{P}(D(t, x) = d) = \frac{\lambda(t, x)^d e^{-\lambda(t, x)}}{d!}$$

gilt, wobei der Poisson-Parameter als  $\lambda(t, x) = m(t, x)E(t, x)$  definiert wird.

Wir können nun für jedes betrachtete Jahr  $t$  und Alter  $x$  den entsprechenden Poissonparameter  $\lambda(t, x)$  direkt berechnen, da wir in unseren demographischen Daten sowohl die Sterberate  $m(t, x)$  also auch das Exposure-to-Risk  $E(t, x)$  gegeben haben. In jeder Wiederholung unseres Bootstrapping-Verfahrens erzeugen wir nun einen neuen Datensatz von Anzahlen an Todesfällen indem wir für jedes Alter und Jahr einen zufälligen Wert einer Poisson-Verteilung mit jeweils passenden Parameter  $\lambda(t, x)$  ziehen. Für diesen neuen Datensatz können wir nun die Sterberate  $m(t, x)$  berechnen und damit wie oben beschrieben die Preise berechnen. Wir wiederholen dieses Verfahren 1000 mal und erhalten ebenso viele Werte für jede Bepreisungsmethode als Preise den q-Forward.

Die Ergebnisse dieses Verfahrens für Daten aus England und Wales bzw. Österreich und die beiden Beobachtungszeiträume 1961-2009 bzw. 1966-2014 und Alter 60 bzw. 70 sind in den Abbildungen 4.7 bis 4.14 graphisch dargestellt. Mit einem Punkt bzw. Dreieck ist hierbei der konkrete Schätzwert für den ursprünglichen Datensatz dargestellt und die umgebenden Linien zeigen die 95%-Konfidenzintervalle.

Vergleichen wir nun die Ergebnisse der jeweiligen Länder und Beobachtungszeiträume für Alter  $x = 60$  und  $x = 70$ , so erkennen wir sofort, dass die Preise für  $x = 60$  in allen Fällen deutlich niedriger sind als jene für  $x = 70$ . Die grafische Darstellung hat daher auch jeweils eine andere Achsenskalierung. Dieses Resultat ist allerdings nicht überraschend, da die Sterberate für das Alter 60 in allen Modellen niedriger vorhergesagt wird als jene für das Alter 70, was klar den demografischen Verhältnissen entspricht.

In jeder Grafik sind die Werte für beide Maturitäten  $T = 10$  und  $T = 30$  dargestellt. Vergleicht man die Werte für jeweils gleiche Schätzfenster und Bepreisungsmethoden, so erkennt man, dass die Preise für  $T = 10$  in allen Fällen höher sind als jene für  $T = 30$ . Dies lässt sich durch die negative Trendkomponente, die in allen Zeitreihenmodellen vorliegt, erklären. Zu beachten ist jedoch, dass je nach konkreter Struktur des gewählten Datensatzes, die Konfidenzintervalle nicht disjunkt sein müssen, wie wir etwa in Abbildung 4.7 und 4.9 erkennen.

Vergleichen wir nun in den jeweiligen Grafiken die Preise der verschiedenen Mortalitätsmodelle, so fällt auf, dass die beiden Varianten des Lee-Carter-Modells oft sehr ähnliche Modelle liefern. Das optimale ARIMA( $p, d, q$ )-Modell entspricht häufig einem ARIMA(0,1,0)-Modell, also genau dem Random Walk mit Drift. Dennoch auftretende Unterschiede lassen sich dadurch erklären, dass das optimale ARIMA( $p, d, q$ )-Modell in jedem Schritt des Bootstrap völlig neu geschätzt wird und sich dabei auch die Parameter  $p$ ,  $d$  und  $q$  ändern können. Das CBD-Modell weist häufig andere Werte als die beiden Varianten des Lee-Carter-Modells

auf.

Betrachten wir nun Abbildung 4.12 so fällt die Gruppierung der Werte in vier Ebenen auf. Neben der bereits erklärten Trennung durch die beiden Maturitäten  $T = 30$  und  $T = 10$  ist der große Einfluss des gewählten Schätzfensters offensichtlich. So ist etwa für beide Maturitäten der Preis hauptsächlich davon abhängig, ob wir für die Zeitreihenanalyse die Werte der letzten sechs Jahre (2009-2014) oder der letzten 21 Jahre (1994-2014) verwenden. Die Wahl des Mortalitätsmodells hat hierbei einen geringeren Einfluss und die Wahl der Bepreisungsmethode erscheint nebensächlich. Auch bei den anderen Grafiken ist der Einfluss des Schätzfensters erkennbar, wenngleich auch nicht so offensichtlich.

Betrachten wir nun die Ergebnisse für die österreichischen Daten so erkennen wir, dass diese im Vergleich zu den Daten aus England und Wales deutlich größere Konfidenzintervalle aufweisen. Dadurch überlappen sogar teilweise die Konfidenzintervalle für die beiden Maturitäten. Die Verläufe der Parameterschätzer  $\hat{\kappa}_t^{(1)}$  bzw.  $\hat{\kappa}_t^{(2)}$  passen sowohl im Lee-Carter-Modell als auch im Cairns-Blake-Dowd-Modell nicht immer gut zum vorgegebenen linearen Trend im Random-Walk-Modell mit Drift. Da das  $ARIMA(p, d, q)$ -Modell auch meist die Gestalt eines Random-Walk-Modells annimmt, ist das auch hier ein Problem. Die Zeitreihenmodelle reagieren somit sensibel auf geringe Änderungen in den Daten, welche bei der Projektion in die Zukunft zu größeren Unterschieden führen. Somit ergeben sich verhältnismäßig große Konfidenzintervalle. Für den Datensatz aus England und Wales ist dieses Verhalten nicht so stark ausgeprägt.

Weitere allgemeine Erkenntnisse, insbesondere ein Vergleich der verschiedenen Preisfindungsmodelle, lassen sich nur schwer finden, da die Wahl des Beobachtungszeitraumes, des Schätzfensters und des Mortalitätsmodells einen übermächtigen Einfluss auf den Preis des q-Forwards haben. Vor allem durch die unterschiedliche Länge des Schätzfensters können völlig voneinander abweichende Ergebnisse erzielt werden.

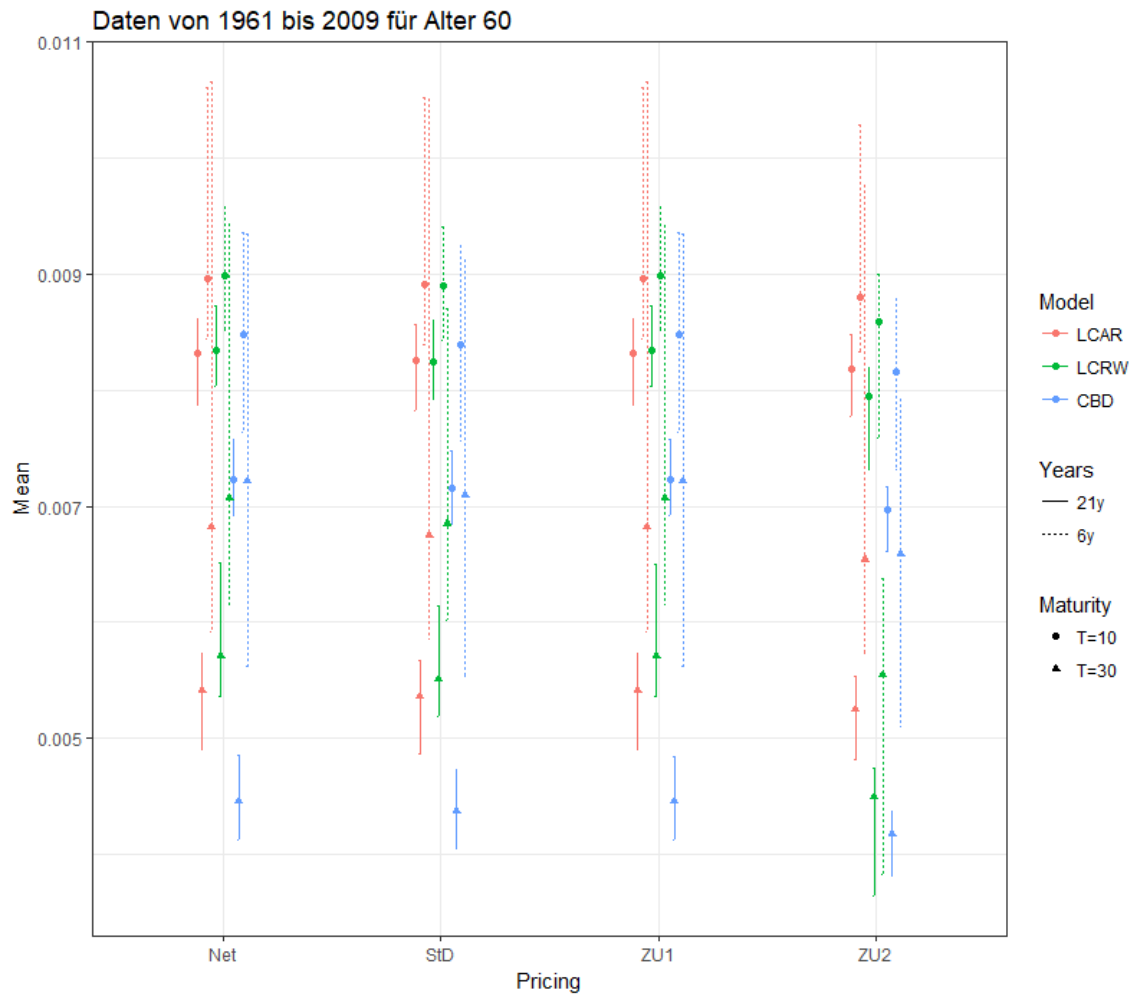


Abbildung 4.7: Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter  $x = 60$  und Daten aus den Jahren 1961-2009 für Österreich

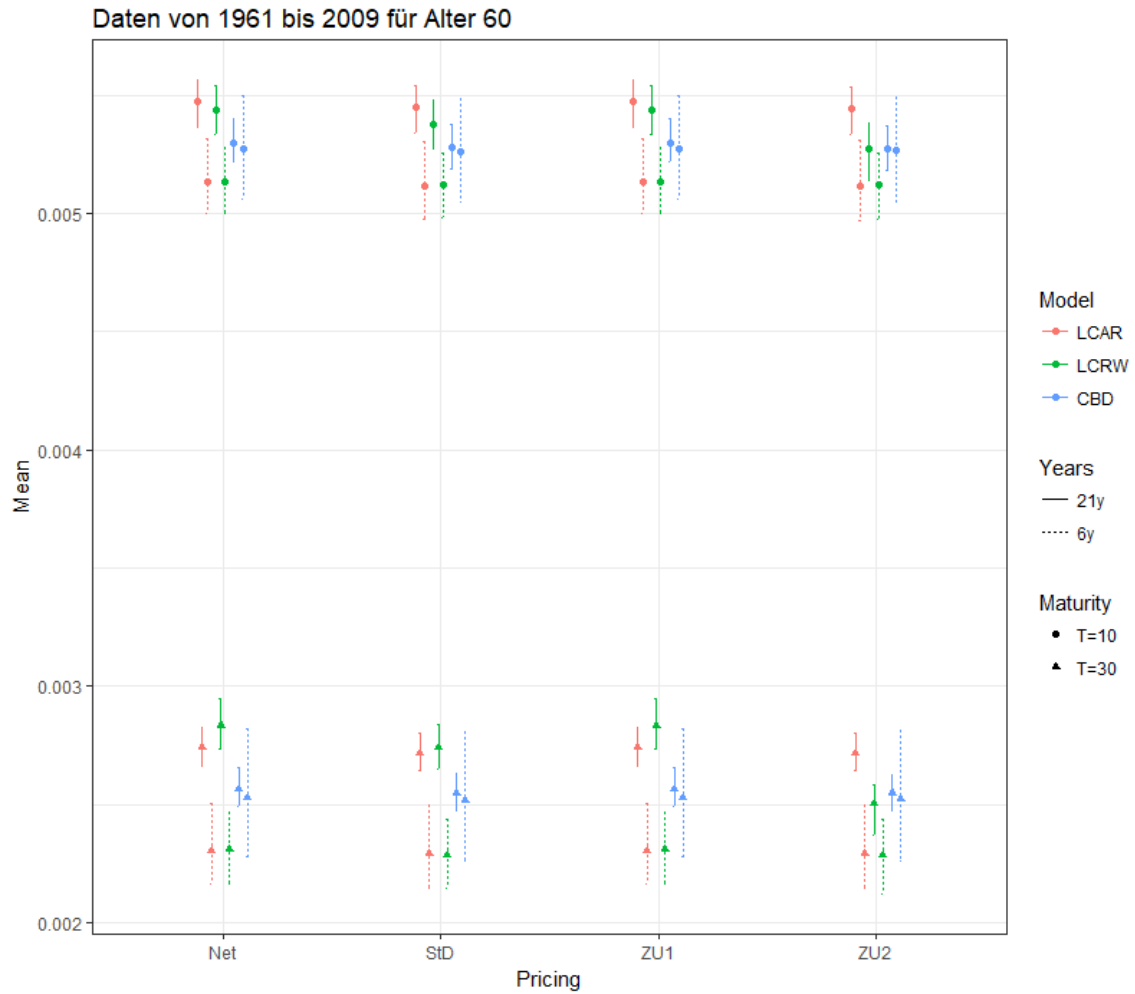


Abbildung 4.8: Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter  $x = 60$  und Daten aus den Jahren 1961-2009 für England und Wales

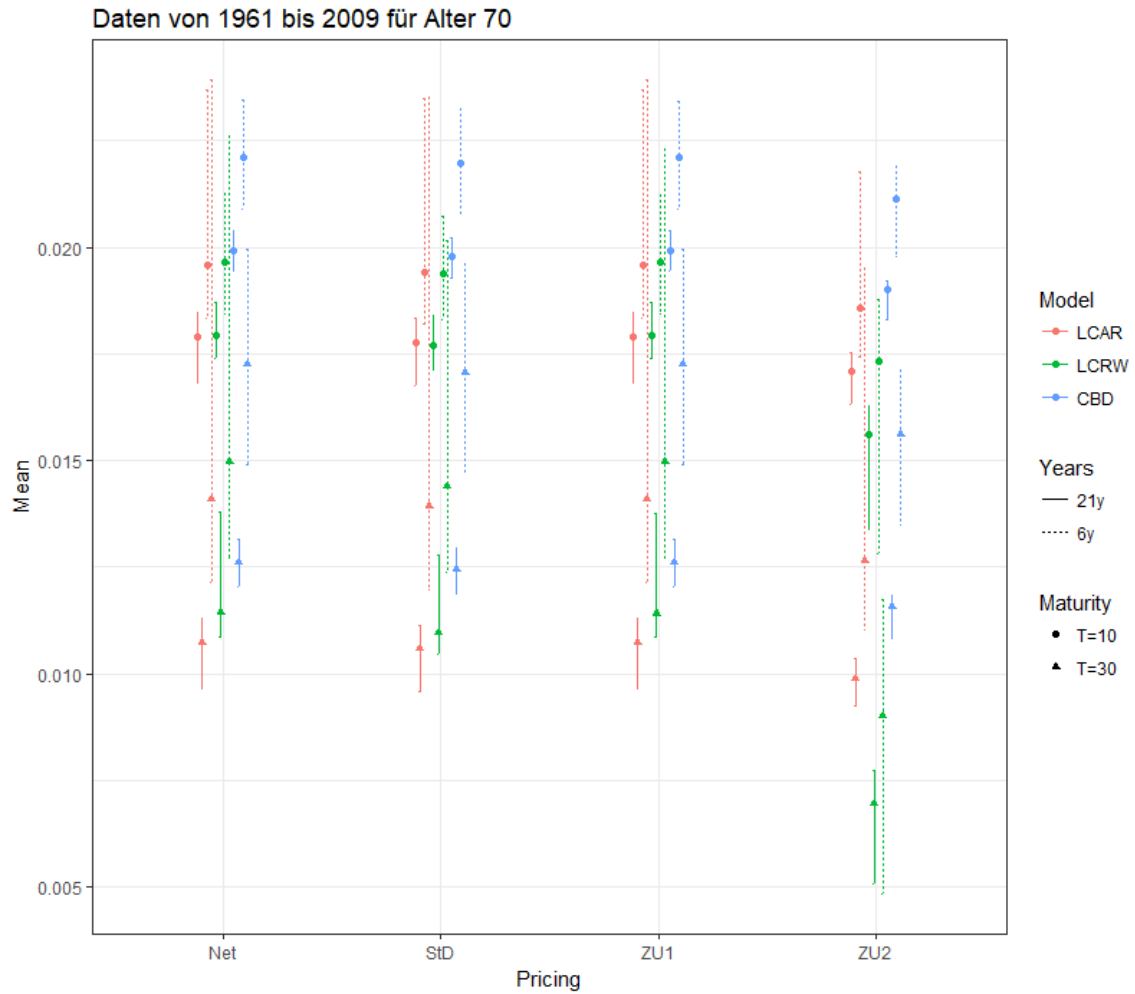


Abbildung 4.9: Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter  $x = 70$  und Daten aus den Jahren 1961-2009 für Österreich



Abbildung 4.10: Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter  $x = 70$  und Daten aus den Jahren 1961-2009 für England und Wales

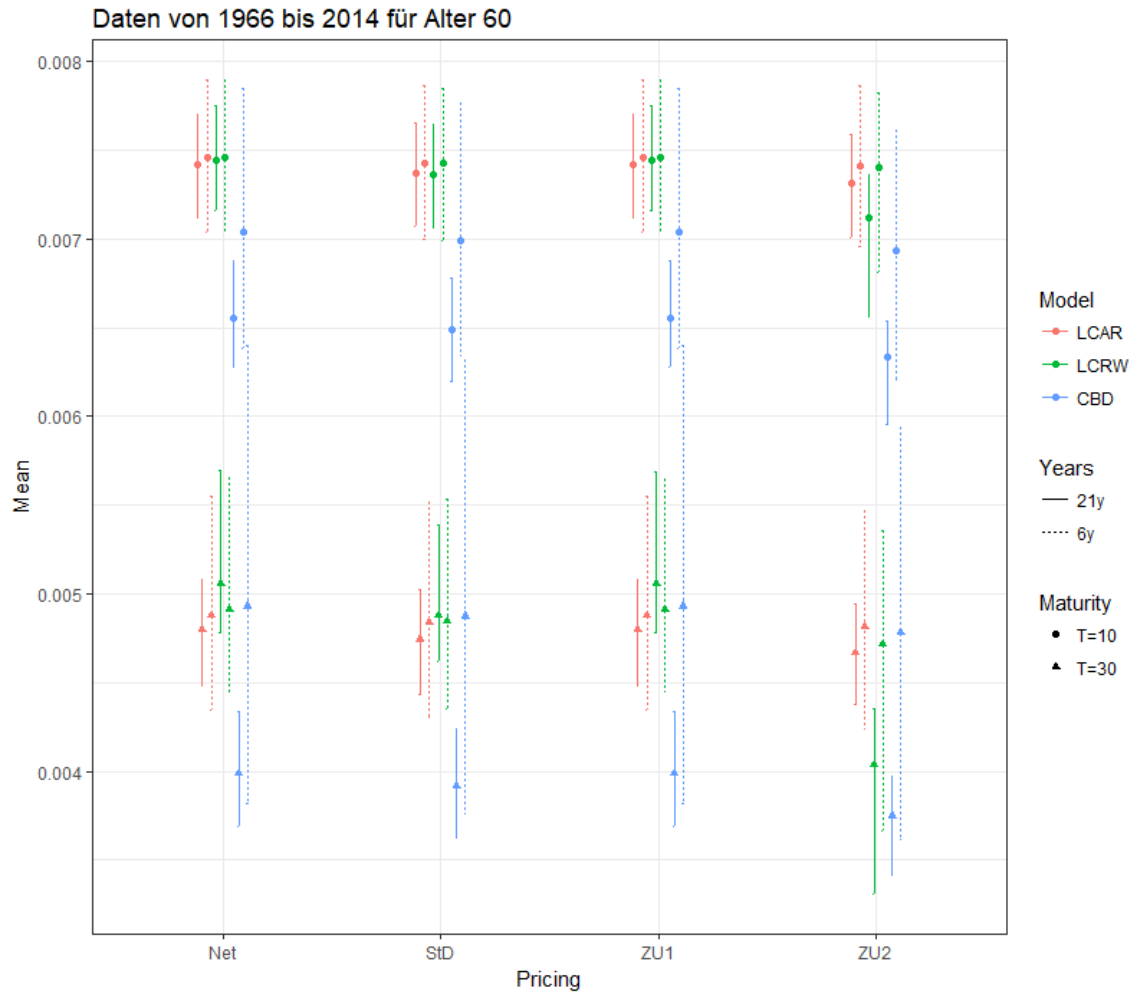


Abbildung 4.11: Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter  $x = 60$  und Daten aus den Jahren 1966-2014 für Österreich

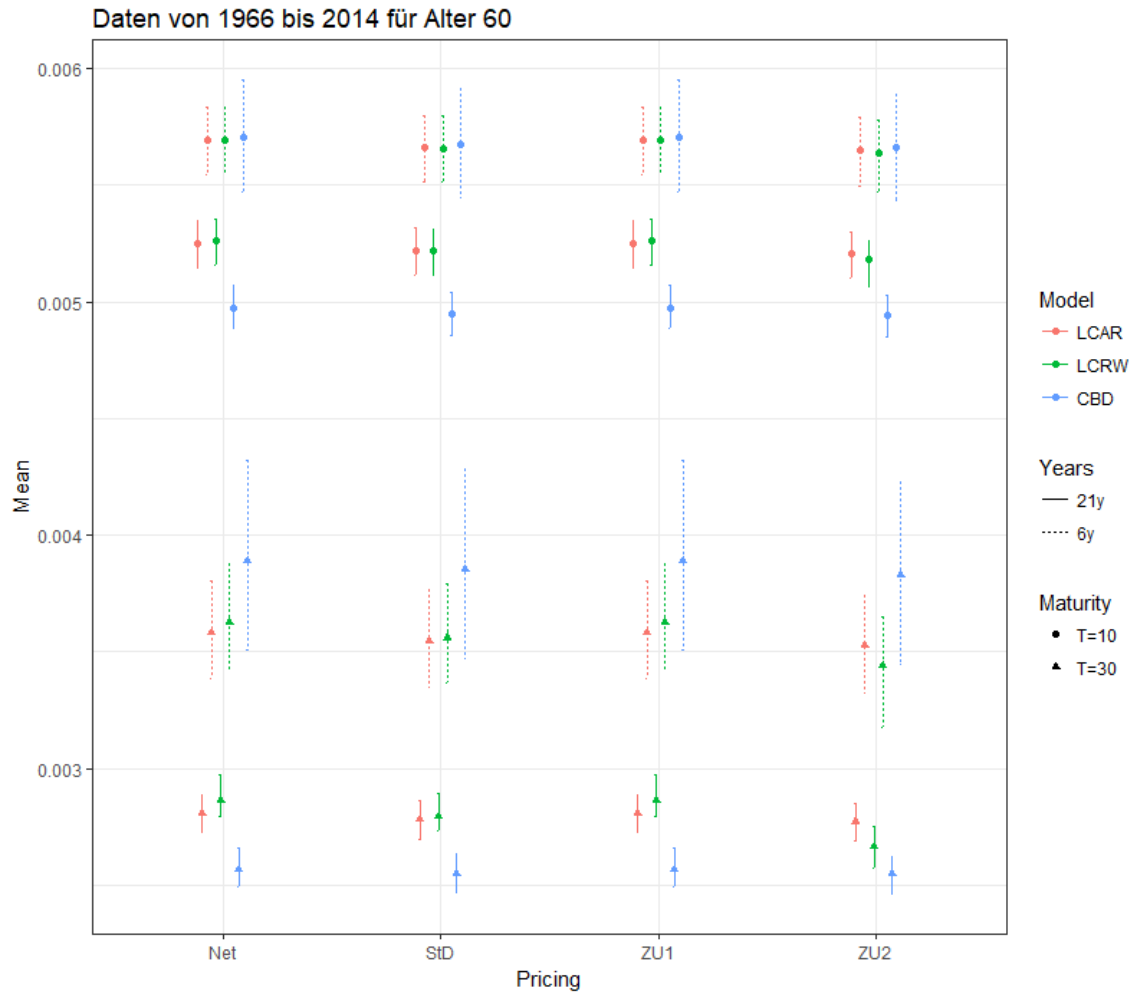


Abbildung 4.12: Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter  $x = 60$  und Daten aus den Jahren 1966-2014 für England und Wales



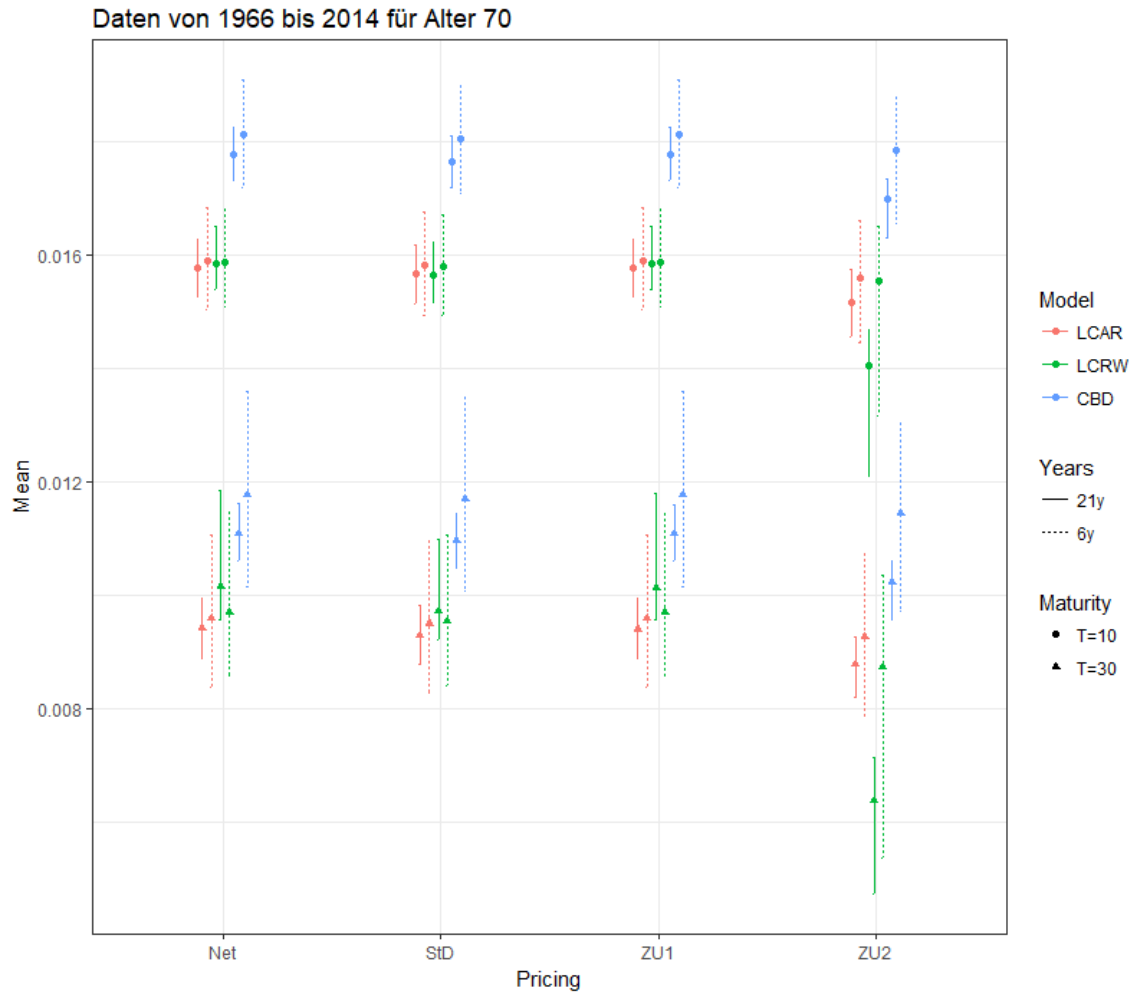


Abbildung 4.13: Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter  $x = 70$  und Daten aus den Jahren 1966-2014 für Österreich

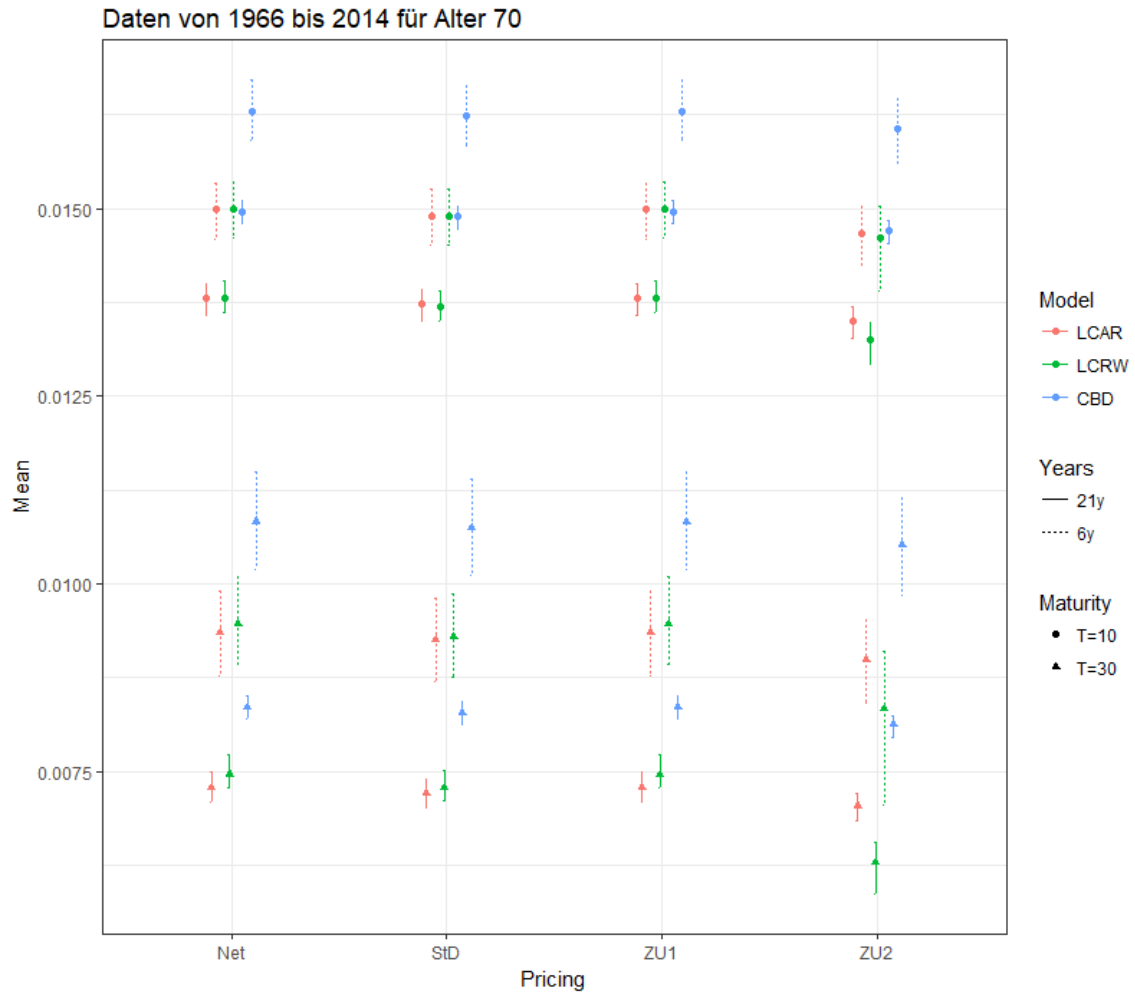


Abbildung 4.14: Preise der q-Forwards mit 95%-Konfidenzintervallen für Alter  $x = 70$  und Daten aus den Jahren 1966-2014 für England und Wales

# Kapitel 5

## Zusammenfassung und Ausblick

In dieser Arbeit wurden drei verschiedene relativ elementare Bepreisungsmethoden für  $q$ -Forwards vorgestellt. Die im Zuge der empirischen Studie damit berechneten Preise sind jedoch stark abhängig von der Art der Modellierung der Mortalität. Insbesondere die dabei verwendeten Zeitfenster haben einen überaus großen Einfluss auf den erhaltenen Preis. Weiters reagieren die Modelle mitunter relativ sensibel auf kleinere Änderungen im Datensatz, was sich in großen Konfidenzintervallen widerspiegelt. Vor allem für die österreichischen Daten tritt dieses Verhalten gehäuft auf.

Bei weiterführender Beschäftigung mit der Bepreisung von  $q$ -Forwards sollte der Fokus zunächst auf die Modellierung der Mortalität gelegt werden. Der häufig vorkommende lineare Trend in den verwendeten Zeitreihenmodellen spielt hierbei eine große Rolle. Problematisch ist jedoch, dass dieser nicht genau genug geschätzt wird oder sich im Laufe der Zeit ändert. Es ist notwendig, dieses Problem durch die Wahl anderer Zeitreihenmodelle oder überhaupt anderer Mortalitätsmodelle in den Griff zu bekommen. Erst wenn es gelingt, dass die Vorhersagen für die Sterberaten nicht mehr so stark von den gewählten Zeitparametern abhängen, ist es sinnvoll komplexere Bepreisungsmethoden einzuführen. Da es zur Zeit noch keinen liquiden Markt für  $q$ -Forwards gibt, besteht auch noch keine Möglichkeit zum Vergleich mit Marktpreisen. Die Entwicklung des Handels mit  $q$ -Forwards sollte jedoch in Zukunft im Auge behalten werden.

# Anhang A

## Verwendeter R Code

```
library(demography)
library(forecast)
library(lifecontingencies)
library(mvtnorm)
library(MASS)
library(plot3D)
library(ggplot2)
library(dplyr)

##### VERWENDETE FUNKTIONEN #####

#Plotten der Sterblichkeitsdaten in 3D
plot_dem <-function(daten, par){
  v <- data.frame(daten$rate[par$serie])
  colnames(v) <- min(daten$year):max(daten$year)
  x<- par$zt0:par$zt1
  y<-par$za0:par$za1
  cols <- as.character(x)
  rows <- as.character(y)
  q<-t(data.matrix(1-exp(-v[rows,cols] )))
  persp(x,y,q, phi = 30, theta = -45, xlab="Jahr", ylab="Alter",
        zlab="q", axes=T, ticktype="detailed")
}

#Schätzen der Parameter kappa2, beta1, beta2 im Lee-Carter-Modell
fittingLC<-function(daten, par){
  lca<-lca(daten, series=par$serie, age=par$za0:par$za1,
           years=par$zt0:par$zt1, adjust="none")
  beta1 <- lca$ax
```

```

beta2 <-lca$bx
kappa2 <- lca$kt
result <-list(beta1, beta2, kappa2)
names(result) <- c("beta1", "beta2","kappa2")
return(result)
}

#Darstellung von kappa2, beta1, beta2 im Lee-Carter-Modell als Plots
plottingLC <-function(daten, par){
  lc <- fittingLC(daten, par)
  plot(lc$kappa2,type="p", ylab="kappa2", xlab="Jahr", ylim=c(-20,20))
  plot(lc$beta1,xaxt="n", ylab="beta1", xlab="Alter")
  axis(1, at=1:(par$za1-par$za0+1),labels=names(lc$beta1))
  plot(lc$beta2, xaxt="n", ylab="beta2", xlab="Alter")
  axis(1, at=1:(par$za1-par$za0+1),labels=names(lc$beta2))
}

# Schätzen von Random Walk Parameter im Lee-Carter-Modell
fittingLCRW<- function(daten,par){
  lc<-fittingLC(daten, par)
  n <- par$tn-par$t0
  h<-tail(lc$kappa2, n+1)
  dif <-diff(h)
  myLC <-mean(dif)
  v<- (dif-myLC)^2
  sigma2LC <- mean(v)
  result <-list(lc$beta1, lc$beta2, lc$kappa2, myLC,sigma2LC)
  names(result) <- c("beta1", "beta2","kappa2", "myLC", "sigma2LC")
  return(result)
}

#Erzeugen von simulierten Werten für Mortalitätsrate q im
# Lee-Carter-Modell mit Random walk
samplingLCRW <-function(daten, par){
  a <-fittingLCRW(daten, par)
  Y<- rnorm(par$M,
            mean=a$kappa2[time(a$kappa2)==par$tn]+a$myLC*par$deltaT,
            sd = a$sigma2LC*par$deltaT)
  q<-probab(a$beta1[par$x],a$beta2[par$x], Y )
  return(q)
}
probab <- function(beta1x, beta2x, Y){
  m <-exp(beta1x+beta2x*Y)
  q <- 1- exp(-m)
  return(q)
}

```

```

}

#Schätzen des optimalen ARIMA-Modells im Lee-Carter-Modell
fittingLCAR<- function(daten, par){
  lc<-fittingLC(daten, par)
  n <- par$tn-par$t0
  v <-tail(lc$kappa2,n+1)
  w <- auto.arima(v, seasonal=FALSE, ic="aic" )
  #stepwise=FALSE würde bessere Ergebnisse liefern,
  #ist aber ineffizient bei Bootstrap
  result <-list(lc$beta1,lc$beta2, lc$kappa2, w)
  names(result) <- c("beta1", "beta2","kappa2", "w")
  return(result)
}

```

```

#Erzeugen von simulierten Werten für Mortalitätsrate q
#im Lee-Carter-Modell mit ARIMA
samplingLCAR<-function(daten, par){
  a<-fittingLCAR(daten, par)
  v<-rep(0, par$M)
  for(i in 1:par$M){
    z<-simulate(a$w,future=TRUE, nsim=par$deltaT)
    Y <- z[time(z)==(par$tn+par$deltaT)]
    q<-probab(a$beta1[par$x],a$beta2[par$x], Y )
    v[i]<- q}

  return(v)
}

```

```

# Schätzen der Parameter kappa1, kappa2 und
#Random Walk Parameter im CBD-Modell
fittingCBD<- function(daten,par){
  n <- par$tn-par$t0
  v <- data.frame(daten$rate[par$serie])
  colnames(v) <- min(daten$year):max(daten$year)
  cols <- as.character(par$zt0:par$zt1)
  rows <- as.character(par$za0:par$za1)
  u<-v[rows,cols]
  q<-1-exp(-u)
  c <- log(q/(1-q))
  x <-par$za0:par$za1- mean(par$za0:par$za1)
  kappa1 <- rep(0,(par$zt1-par$zt0+1))
  names(kappa1) <- colnames(c)
}

```

```

kappa2 <- rep(0,(par$zt1-par$zt0+1))
names(kappa2) <- colnames(c)

for(i in cols){
  y<- c[,i]
  fm <- lm(y ~ x)   #lineare Regression
  kappa1[i]<-fm$coefficients["(Intercept)"]
  kappa2[i]<-fm$coefficients["x"]
}

kappa1<- ts(kappa1, start= as.numeric(names(kappa1[1])))
kappa2 <- ts(kappa2, start= as.numeric(names(kappa2[1])))
end1 <- tail(kappa1, n+1)
end2 <-tail(kappa2, n+1)
diff1 <- diff(end1)
diff2 <- diff(end2)
my1<-mean(diff1)
my2 <-mean(diff2)
my <- unname(c(my1, my2))
summe <- matrix(0, nrow=2, ncol=2)
for(i in 1:n){
  v<-c(diff1[i]-my1, diff2[i]-my2)
  summe <- summe + v %*% t(v)
}
sigma2CBD <- unname(summe/n)
result<- list(kappa1, kappa2,x, my, sigma2CBD)
names(result)<- c("kappa1", "kappa2", "beta2", "myCBD", "sigma2CBD")
return(result)
}

#Darstellung von kappa1, kappa2 und beta2 im CBD-Modell
plottingCBD <-function(daten,par){
  v <-fittingCBD(daten,par)
  plot(names(v$kappa1), v$kappa1, type="p", xlab="Jahr", ylab="kappa1")
  plot(names(v$kappa2), v$kappa2, type="p", xlab="Jahr", ylab="kappa2")
  plot(v$beta2, xaxt="n", xlab="Alter", ylab="beta2")
  axis(1, at=1:(par$za1-par$za0+1),labels=par$za0:par$za1)
}

##Erzeugen von simulierten Werten für Mortalitätsrate q
#im CBD-Modell mit Random walk
samplingCBD <-function(daten, par){
  a<-fittingCBD(daten,par)
  tn <- as.character(par$tn)
}

```

```

Y2<-mvrnorm(n=par$M,
            c(a$kappa1[tn],a$kappa2[tn])+a$myCBD*par$deltaT,
            a$sigma2CBD*par$deltaT)
q2<-exp(Y2[,1]+Y2[,2]*(par$x1-mean(par$za0:par$za1)))/
      (1+exp(Y2[,1]+Y2[,2]*(par$x1-mean(par$za0:par$za1))))
return(q2)
}

#Berechnung der Preise für die verschiedenen Bepreisungsmethoden
pricing <-function(art, daten, par){
  q<-switch(art,
           LCRW=samplingLCRW(daten,par),
           LCAR=samplingLCAR(daten,par),
           CBD=samplingCBD(daten, par))
  PriceNet <- mean(q)
  PriceStd <- mean(q) + par$lambda*sd(q)
  PriceZU1 <- -1/par$gz1*log(mean(exp(-par$gz1*q)))
  PriceZU2 <- -1/par$gz2*log(mean(exp(-par$gz2*q)))

  result <- list (PriceNet, PriceStd, PriceZU1, PriceZU2)
  names(result) <-c("PriceNet", "PriceStd", "PriceZU1", "PriceZU2")
  return(result)
}

#Führe für alle 3 Mortalitätsmodelle Bootstrapping durch
bootstrap <- function(daten,par ){
  v_LCAR<-v_LCRW<-v_CBD <-matrix(0,nrow=par$bstr,ncol=4)

  for(i in 1:par$bstr){
    anzj <- max(daten$year)-min(daten$year)+1
    anza <- max(daten$age)- min(daten$age)+1
    la <-matrix(unlist(unname(daten$pop[par$serie])), nrow=anza)*
      matrix(unlist(unname(daten$rate[par$serie])), nrow=111)
    D <- matrix (rpois(anza*anzj, la), nrow=anza)
    daten2 <- demogdata(
      D/matrix(unlist(unname(daten$pop[par$serie])), nrow=anza),
      matrix(unlist(unname(daten$pop[par$serie])), nrow=anza),
      ages=daten$age, years=daten$year, type="mortality",
      name=par$serie, label="Austria")
    price <-pricing("LCAR", daten2, par)
    v_LCAR[i,1]<-price$PriceNet
  }
}

```



```

v_LCAR[i,2]<-price$PriceStd
v_LCAR[i,3]<-price$PriceZU1
v_LCAR[i,4]<-price$PriceZU2

price <-pricing("LCRW", daten2, par)
v_LCRW[i,1]<-price$PriceNet
v_LCRW[i,2]<-price$PriceStd
v_LCRW[i,3]<-price$PriceZU1
v_LCRW[i,4]<-price$PriceZU2

price <-pricing("CBD", daten2, par)
v_CBD[i,1]<-price$PriceNet
v_CBD[i,2]<-price$PriceStd
v_CBD[i,3]<-price$PriceZU1
v_CBD[i,4]<-price$PriceZU2
}

result <-list(v_LCAR,v_LCRW,v_CBD)
names(result)<-c("LCAR","LCRW","CBD")
return(result)
}

#Berechne 2,5% und 97,5% Quantile
quantiles <- function(bootstr){
v<-matrix(0,nrow=4,ncol=2)
for(i in 1:4){
v[i,1]<- quantile(bootstr[,i], 0.025, na.rm=TRUE)
v[i,2]<- quantile(bootstr[,i], 0.975, na.rm=TRUE)
}
return(v)
}

#Führe Bepreisung samt Bootstrapping für die Konfidenzintervalle durch
#und stelle die Ergebnisse als dataframe dar
todataframe <-function(daten, par){
#speichere direkt als DF um factors zu vermeiden

w<-data.frame(Model=character(), Pricing=character(),
Years=character(),Maturity=character(),
mean=double(), qLow= double(),qUp =double())
for(Mat in par$deltaT_vector ){ #c(10,30)
par$deltaT<-Mat
for(t0 in par$t0_vector){
par$t0 <-t0

```

```

qs<-bootstrap(daten,par)

for( art in c("LCAR", "LCRW", "CBD")){
  print(paste(Mat,t0,art,sep="□"))
  pr <- pricing(art, daten, par)
  quant <- quantiles(matrix(unlist(qs[art]),ncol=4))
  Model<- rep(art,4)
  Pricing <- c("Net","StD","ZU1", "ZU2")
  Years <- rep(paste(as.character(par$tn-par$t0+1),"y",sep=""),4)
  Maturity <- rep(paste("T=",as.character(par$deltaT), sep=""),4)
  Mean <-c(pr$PriceNet,pr$PriceStd,pr$PriceZU1,pr$PriceZU2 )
  qLow <- quant[,1]
  qUp <-quant[,2]

  v1 <-data.frame(Model,Pricing,Years,Maturity, Mean, qUp,qLow)
  w <- rbind(w,v1)
}
}
}
return(w)
}

##### PARAMETER UND SIMULATION #####

#Daten bis 2009, Alter 60
par1 <- list(t0_vector =c(1989,2004),
            #Beginn der estimation windows für die Zeitreihenanalyse
            tn =2009,                #Ende davon
            za0 = 60,                #Beginn der beobachteten Alter
            za1 = 89,                #Ende davon
            zt0 = 1961,              #Begin der beobachteten Jahre
            zt1 = 2009,              #Ende davon
            serie = "male",          #Betrachtete Serie (male, female, total)
            deltaT_vector = c(10,30), #Maturität in Jahren
            x1 =60,                  #betrachtetes Alter für qForward
            x=as.character(par1$x1),
            M = 10000,                #Anzahl der Monte Carlo Simulationen
            lambda = -0.1,            #Parameter im Standardabweichungsprinzip
            gz1 = 1,                  #gamma*z im Zero Utility Prinzip-Variante 1
            gz2 = 1000,              #gamma*z im Zero Utility Prinzip-Variante 2
            bstr=1000                 #Anzahl der Durchläufe beim Bootstrapping
)

#Daten bis 2009, Alter 70

```

```

par2 <- list( t0_vector =c(1989,2004),
#Beginn des estimation windows für die Zeitreihenanalyse
tn =2009,           #Ende davon
za0 = 60,          #Beginn der beobachteten Alter
za1 = 89,          #Ende davon
zt0 = 1961,        #Begin der beobachteten Jahre
zt1 = 2009,        #Ende davon
serie = "male",    #Betrachtete Serie (male, female, total)
deltaT_vector = c(10,30), #Maturität in Jahren
x1 =70,            #betrachtetes Alter für qForward
x=as.character(par2$x1),
M = 10000,         #Anzahl der Monte Carlo Simulationen
lambda = -0.1,     #Parameter im Standardabweichungsprinzip
gz1 = 1,           #gamma*z im Zero Utility Prinzip-Variante 1
gz2 = 1000,        #gamma*z im Zero Utility Prinzip-Variante 2
bstr=1000          #Anzahl der Durchläufe beim Bootstrapping
)

#Daten bis 2014, Alter 60
par3 <- list( t0_vector =c(1994,2009),
#Beginn des estimation windows für die Zeitreihenanalyse
tn =2014,           #Ende davon
za0 = 60,          #Beginn der beobachteten Alter
za1 = 89,          #Ende davon
zt0 = 1966,        #Begin der beobachteten Jahre
zt1 = 2014,        #Ende davon
serie = "male",    #Betrachtete Serie (male, female, total)
deltaT_vector = c(10,30), #Maturität in Jahren
x1 =60,            #betrachtetes Alter für qForward
x=as.character(par3$x1),
M = 10000,         #Anzahl der Monte Carlo Simulationen
lambda = -0.1,     #Parameter im Standardabweichungsprinzip
gz1 = 1,           #gamma*z im Zero Utility Prinzip-Variante 1
gz2 = 1000,        #gamma*z im Zero Utility Prinzip-Variante 2
bstr=1000          #Anzahl der Durchläufe beim Bootstrapping
)

#Daten bis 2014, Alter 70
par4 <- list( t0_vector =c(1994,2009),
#Beginn des estimation windows für die Zeitreihenanalyse
tn =2014,           #Ende davon
za0 = 60,          #Beginn der beobachteten Alter
za1 = 89,          #Ende davon
zt0 = 1966,        #Begin der beobachteten Jahre
zt1 = 2014,        #Ende davon

```

```

    serie = "male",      #Betrachtete Serie (male, female, total)
    deltaT_vector = c(10,30), #Maturität in Jahren
    x1 =70,              #betrachtetes Alter für qForward
    x=as.character(par4$x1),
    M = 10000,          #Anzahl der Monte Carlo Simulationen
    lambda = -0.1,      #Parameter im Standardabweichungsprinzip
    gz1 = 1,            #gamma*z im Zero Utility Prinzip-Variante 1
    gz2 = 1000,         #gamma*z im Zero Utility Prinzip-Variante 2
    bstr=1000           #Anzahl der Durchläufe beim Bootstrapping
)

#fürs Plotten von CBD
par <- list( t0 =2009,
             #Beginn des estimation windows für die Zeitreihenanalyse
             tn =2014,      #Ende davon
             za0 = 60,      #Beginn der beobachteten Alter
             za1 = 89,      #Ende davon
             zt0 = 1961,    #Begin der beobachteten Jahre
             zt1 = 2014,    #Ende davon
             serie = "male" #Betrachtete Serie (male, female, total)
)

#Lade Daten für Österreich direkt aus der Human Mortality Database
daten <-hmd.mx("AUT", "ricarda1@live.at", "gouryella", "Austria")

#für England/Wales leider nicht so einfach möglich.
#Lade Daten aus gespeicherten Dateien
daten_eng <-read.demogdata("C:/Users/Ricarda/Documents/Masterarbeit/
empirischer_teil/engwal_deathrates.txt",
"C:/Users/Ricarda/Documents/Masterarbeit/
empirischer_teil/engwal_exposure.txt",
type="mortality", label="EnglandWales")

#3D-Plot der österreichischen Daten
plot_dem(daten, par)

#Plots der Zeitreihen
plottingLC(daten, par1)
plottingLC(daten_eng, par1)
plottingLC(daten, par3)
plottingLC(daten_eng, par3)
plottingCBD(daten,par)

```

```

plottingCBD(daten_eng, par)

#erzeuge Dataframe mit allen wichtigen Variablen für LCRW
#und Listen für CBD und LCAR
LCRW <-data.frame(Jahre=character(),Land=character(),
                  my=double(),sigma2=double())

LCAR<-list()
CBD<-list()
for(la in c("AUT","ENG")){
data<-switch(la,
              AUT=daten,
              ENG=daten_eng)
for(y in c("par1","par3")){
par<-switch(y,
            par1=par1,
            par3=par3)
for(x in par$t0_vector){
par$t0<-x
#LCRW
fit<-fittingLCRW(data, par)
my<-fit$myLC
sigma2<-fit$sigma2LC
Jahre <-paste(as.character(par$t0),"-",as.character(par$tn))
Land <-la
v1 <-data.frame(Jahre,Land,my,sigma2)
LCRW <- rbind(LCRW,v1)
#LCAR
fit1<-fittingLCAR(data, par)
coefs<-fit1$w$coef
sigma2<-fit1$w$sigma2
name <-paste(as.character(par$t0),"-",as.character(par$tn),la)
liste1<-list(name,coefs,sigma2)
names(liste1)<-c("Jahr+Land","Koeffs","Sigma2")
LCAR<-c(LCAR,liste1)
#CBD
fit2<-fittingCBD(data,par)
my<-fit2$myCBD
sigma2<-fit2$sigma2CBD
liste2<-list(name,my,sigma2)
names(liste2)<-c("Jahr+Land","My","Sigma2")
CBD<-c(CBD,liste2)
}}}}

#Plotte Vergleich der beiden Arima Forecasts

```

```

par3$t0<-2009
lc<-fittingLC(daten_eng, par3)
n <- par3$tn-par3$t0
v <-tail(lc$kappa2,n+1)
w_total <- auto.arima(v,stepwise=FALSE, seasonal=FALSE, ic="aic" )
w_stepwise <- auto.arima(v, seasonal=FALSE, ic="aic" )
f1<-forecast(w_total, 10)
f2<-forecast(w_stepwise, 10)
f<-cbind(f1=f1$mean,f2=f2$mean )
df<-cbind(v,f)
colnames(df)<- c("Daten", "Arima(2,1,0)", "Arima(0,1,0)")
autoplot(df)+geom_point()+xlab("Jahr")+ylab("kappa2")+theme_bw()

```

*#Führe gesamte Simulation + Bootstrapping durch*

```

aut_60_2009_neu<-todataframe(daten, par1)
aut_70_2009_neu<-todataframe(daten, par2)
aut_60_2014_neu<-todataframe(daten,par3)
aut_70_2014_neu <-todataframe(daten,par4)

```

```

eng_60_2009 <-todataframe(daten_eng, par1)
eng_70_2009 <-todataframe(daten_eng, par2)
eng_60_2014 <-todataframe(daten_eng, par3)
eng_70_2014 <-todataframe(daten_eng, par4)

```

*#Stelle erhaltene Werte in Plots dar*

```

pd <- position_dodge(0.25)
finalplot <-function(ergebnis, jahr, alter){
  ggplot(ergebnis, aes(x=Pricing, y=Mean, colour=Model,
                      linetype=Years, shape=Maturity))+
    geom_errorbar(aes(ymin=qLow,ymax=qUp), width=.1, position=pd)+
    geom_point(position=pd)+ theme_bw()
  +ggtitle(paste("Daten von",jahr-48, "bis",jahr, "für Alter", alter))
}

```

```

finalplot(aut_60_2009_neu, 2009,60)
finalplot(aut_70_2009_neu, 2009,70)
finalplot(aut_60_2014_neu,2014,60)
finalplot(aut_70_2014_neu, 2014, 70)
finalplot(eng_60_2009,2009,60)
finalplot(eng_70_2009,2009,70)
finalplot(eng_60_2014,2014,60)
finalplot(eng_70_2014,2014,70)

```

# Literatur

- [1] The Life & Longevity Markets Association. *Longevity Pricing Framework*. 2010. URL: <https://llma.org/publications/>.
- [2] The Life & Longevity Markets Association. *Sample Term Sheet – q-forward*. 2010. URL: <https://llma.org/publications/>.
- [3] The Life & Longevity Markets Association. *Technical Note – The q-forward*. 2010. URL: <https://llma.org/publications/>.
- [4] Pauline M. Barrieu und Luitgard A.M. Veraart. »Pricing q-forward contracts: an evaluation of estimation window and pricing method under different mortality models«. In: *Scandinavian Actuarial Journal* 2016.2 (2016), S. 146–166.
- [5] Peter J. Brockwell und Richard A. Davis. *Time series theory and methods*. 2nd edition. New York: Springer, 1991.
- [6] Andrew J. G. Cairns, David Blake, Kevin Dowd, Guy D. Coughlan, David Epstein, Alen Ong und Igor Balevich. »A Quantitative Comparison of Stochastic Mortality Models Using Data From England and Wales and the United States«. In: *North American Actuarial Journal* 13.1 (2009), S. 1–35.
- [7] Guy Coughlan, David Epstein, Amit Sinha und Paul Honig. *q-Forwards: Derivatives for Transferring Longevity and Mortality Risks*. 2007.
- [8] Rob Hyndman und Yeasmin Khandakar. »Automatic Time Series Forecasting: The forecast Package for R«. In: *Journal of Statistical Software, Articles* 27.3 (2008), S. 1–22.
- [9] Rob J Hyndman und George Athanasopoulos. *Forecasting: principles and practice*. 2nd edition. 2018. URL: <https://www.otexts.org/fpp2/>.
- [10] Ronald D. Lee und Lawrence R. Carter. »Modeling and Forecasting U.S. Mortality«. In: *Journal of the American Statistical Association* 87.419 (1992), S. 659–671.
- [11] Robert H. Shumway und David S. Stoffer. *Time series analysis and its applications : with R examples*. 2nd edition. New York: Springer, 2006.
- [12] Giorgio Spedicato. *Mortality projection with demography and lifecontingencies packages*. URL: <https://CRAN.R-project.org/package=lifecontingencies>.