



Analyse der Sterblichkeit in Osteuropa

Bachelorarbeit

im Studiengang Wirtschaftsmathematik
Hochschule Koblenz, RheinAhrCampus Remagen

vorgelegt von
Takam Tchedjou
geb. am 22.02.1998 in Jaunde/Kamerun
Matrikelnr: 535481

Erstgutachter : Herr Prof. Dr. Claus Neidhardt
Zweitgutachter : Frau Prof. Dr. Martina Brück

19. März 2020

Inhaltsverzeichnis

Abbildungsverzeichnis	iv
Tabellenverzeichnis	vii
1 Einleitung	1
2 Datenbeschreibung und Deskriptive Analyse	3
2.1 Datenbeschreibung und -vorbereitung	4
2.2 Deskriptive Analyse der Sterblichkeitsverbesserung	4
2.3 Visualisierung der logarithmischen Sterblichkeitsrate	13
3 Lee-Carter Modell	19
3.1 Ansatz	19
3.2 Parameterschätzung	20
3.2.1 Singulärwertzerlegung (SWZ)	21
3.3 Random Walk with Drift (RWD)	23
4 Sterblichkeitsprognosen für kleine Populationen	24
4.1 Idee	24
4.2 Replikation von der Sterblichkeitsveränderung	25
4.2.1 Methode der kleinsten Quadrate	25
5 Auswertung	27
5.1 Lee-Cartell Modell	28
5.1.1 Verlauf von $\hat{\alpha}_x$	28
5.1.2 Verlauf von $\hat{\beta}_x$	29
5.1.3 Verlauf von $\hat{\kappa}_t$	33
5.2 Ergebnisse der Vorhersage der Sterblichkeitsveränderung ($\hat{\kappa}_t^{[AVE]}$) in Estland, Lettland und Slowenien	36
5.3 Überprüfung der Stabilität der Gewichte	39
5.4 Backtesting	42

6	Fazit und Ausblick	46
7	Anhang	47
	Literaturverzeichnis	58

Abbildungsverzeichnis

2.1	Farbverlauf der Heatmap	5
2.2	Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Lettland	6
2.3	Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Estland	6
2.4	Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Slo- wenien	7
2.5	Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Slowakei	7
2.6	Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Un- garn	8
2.7	Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Polen	8
2.8	Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Tsche- chien	9
2.9	Verlauf vom Logarithmus der Sterberate ($\mu_{[30:100],t}$) für Männer zwischen 30-100 Jahre alt in 1995 und 2015	14
2.10	Verlauf vom Logarithmus der Sterberate ($\mu_{[30:100],t}$) für Frauen zwischen 30-100 Jahre alt in 1995 und 2015	15
2.11	Verlauf vom Logarithmus der Sterberate ($\mu_{40,[1990:2016]}$) für 40-jährige Männer	16
2.12	Verlauf vom Logarithmus der Sterberate ($\mu_{80,[1990:2016]}$) für 80-jährige Männer	16
2.13	Verlauf vom Logarithmus der Sterberate ($\mu_{40,[1990:2016]}$) für 40-jährige Frauen	17
2.14	Verlauf vom Logarithmus der Sterberate ($\mu_{80,[1990:2016]}$) für 80-jährige Frauen	17
5.1	Verläufe von $\hat{\alpha}_x$ und $\hat{\alpha}_y$ bei Männern und Frauen von 30-100 J.a. in 1990-2016	28
5.2	Verlauf von $\hat{\beta}_x$ bei Männern und Frauen von 30-90 J.a. in 1990-2016 . .	31
5.3	Verlauf von $\hat{\kappa}_t$ bei Männern und Frauen von 30-100 J.a. in 1990-2016 .	35
5.4	Vergleich zwischen $\hat{\kappa}_t$ und $\hat{\kappa}_t^{[AVE]}$ in Estland, Lettland und Slowenien .	39

5.5	Vergleich zwischen $\hat{\kappa}_t$, $\hat{\kappa}_t^{RWD}$ und $\hat{\kappa}_t^{[AVE]/RWD}$	45
7.1	medizinische Fortschritte 1870 bis 2015. (Quelle: siehe [Pha])	47
7.2	Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Estland	48
7.3	Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Lettland	49
7.4	Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Slowenien	49
7.5	Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Polen	50
7.6	Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Slowakei	51
7.7	Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Tschechien	51
7.8	Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Ungarn	52
7.9	Vergleich zwischen der logarithmieren Sterberate $\mu_{x,t}$ und der logarithmieren Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Estland	53
7.10	Vergleich zwischen der logarithmieren Sterberate $\mu_{x,t}$ und der logarithmieren Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Lettland	54
7.11	Vergleich zwischen der logarithmieren Sterberate $\mu_{x,t}$ und der logarithmieren Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Slowenien	54
7.12	Vergleich zwischen der logarithmieren Sterberate $\mu_{x,t}$ und der logarithmieren Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Polen	55
7.13	Vergleich zwischen der logarithmieren Sterberate $\mu_{x,t}$ und der logarithmieren Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Slowakei	56
7.14	Vergleich zwischen der logarithmieren Sterberate $\mu_{x,t}$ und der logarithmieren Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Tschechien	56

7.15 Vergleich zwischen der logarithmieren Sterberate $\mu_{x,t}$ und der logarithmieren Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Ungarn	57
--	----

Tabellenverzeichnis

2.1	Bevölkerungszahl in den Jahren 2017, 2018 und 2019	3
5.1	Gewichtsdarstellung ($\hat{\omega}_i$) in Estland	36
5.2	Gewichtsdarstellung ($\hat{\omega}_i$) in Lettland	36
5.3	Gewichtsdarstellung ($\hat{\omega}_i$) in Slowenien	37
5.4	Gewichtsdarstellung ($\hat{\omega}_i$) in Estland	40
5.5	Gewichtsdarstellung ($\hat{\omega}_i$) in Lettland	40
5.6	Gewichtsdarstellung ($\hat{\omega}_i$) in Slowenien	41
5.7	Drift ($\hat{\theta}$) berechnet aus $\hat{\kappa}_t$ (Lee-Carter) zwischen 1990 und 2005	42
5.8	Drift ($\hat{\theta}$) berechnet aus $\hat{\kappa}_t^{[AVE]}$ (Replikation) zwischen 1990 und 2005	42

Kapitel 1

Einleitung

In der biometrischen Mathematik wird grundsätzlich zwischen dem Sterberisiko und dem Langlebigkeitsrisiko unterschieden. Beide beinhalten Risikofaktoren, die z.B. je nach Ausgestaltung einer Versicherung besonders berücksichtigt werden müssen. Das Sterberisiko ist das Risiko, eine höhere Sterblichkeitsrate als erwartet zu haben, was eine höhere Sterbewahrscheinlichkeit bedeutet. Die Faktoren, die die Sterblichkeit beeinflussen, sind Kriege, Epidemien oder Pandemien, die zu höheren Sterblichkeitsraten führen ([GBRF12]). Das Langlebigkeitsrisiko ist das Risiko, dass Menschen länger als erwartet überleben oder die Sterblichkeitsrate niedriger als erwartet ist. Fortschritte in der Medizin, technologische Verbesserungen und Änderungen im Lebensstil führen tendenziell zu einer Verringerung der Zahl der Todesfälle ([GBRF12]).

Aus wirtschaftlicher Sicht hat der Rückgang der Sterblichkeit erhebliche negative Auswirkungen auf die Pensionspläne und Rentenversicherer. Eine deutliche Überschätzung der Sterblichkeitsrate impliziert ein hohes Risikoprofil für Pensionsfonds und Rentenversicherer ([GBRF12]). Auf dieser Weise spielt die Sterblichkeitsmodellierung von Ländern eine wichtige Rolle in der Versicherungsbranche. Das Lee-Carter Modell (1992) scheint das erste Modell zu sein, das die gestiegenen Lebenserwartungstrends bei den Sterblichkeitsraten berücksichtigt.

Bei solchen Aktivitäten stehen Aktuarien oft vor dem Problem, die Sterblichkeit kleiner Bevölkerungen zu modellieren: z.B. kleine Länder, eine bestimmte Region eines Landes. In den kleinen Populationen sind die Sterblichkeitsraten stärker von zufälligen Fluktuationen beeinflusst als in den größeren ([MMS19]).

Meine Aufgabe ist es, die Sterblichkeit in Osteuropa zu untersuchen, indem ich mit Mortalitätsdaten von Estland, Litauen, Slowenien, Polen, Tschechien, Slowakei und Ungarn arbeite.

Mein Projekt wird in vier Hauptteile gegliedert sein. Der erste Teil umfasst zunächst die Beschreibung und Vorbereitung der Daten, die für die Durchführung meines Pro-

jekts verwendet wurden, dann eine deskriptive Analyse der Sterblichkeitsverbesserung mit unterstützenden Grafiken und schließlich eine Visualisierung der logarithmischen Sterblichkeitsrate in den sieben oben genannten Ländern. Der zweite und dritte Teil sind die theoretischen Teile der Arbeit. Sie werden sowohl zeigen, wie das Lee-Carter Modell verwendet werden kann, als auch, wie die Mortalitätsdaten von Ländern mit kleinen Bevölkerungen mit Hilfe größerer Länder repliziert werden können. Und schließlich fasst der letzte Teil die Interpretation der im Rahmen unserer Aufgabe erzielten Ergebnisse zusammen.

Kapitel 2

Datenbeschreibung und Deskriptive Analyse

Laut der Online-Datenbank *Countrymeters* ([Cou20]) zeigt die folgende Tabelle zunächst die Bevölkerungszahl (in Millionen von Einwohnern) jedes der sieben Länder im Zeitraum von 2017 bis 2019, damit wir kleine von größeren Ländern leicht unterscheiden können.

Land	2017 (Mio.)	2018 (Mio.)	2019 (Mio.)
Estland	1,304	1,300	1,296
Lettland	1,933	1,910	1,887
Polen	38,63	38,64	38,64
Slowakei	5,432	5,436	5,440
Slowenien	2,073	2,076	2,079
Tschechien	10,557	10,565	10,572
Ungarn	9,809	9,777	9,746

Tabelle 2.1: Bevölkerungszahl in den Jahren 2017, 2018 und 2019

Wir können daher den Schluss ziehen, dass Lettland, Estland und Slowenien als kleine Länder betrachtet werden können, weil ihre Bevölkerung weniger als 2,5 Millionen Einwohner hat.

2.1 Datenbeschreibung und -vorbereitung

Die für unser Projekt verwendeten Daten stammen aus der Online-Datenbank *Human Mortality Database* (siehe [Mor]). Diese Datenbank enthält detaillierte Bevölkerungs- und Mortalitätsdaten für Männer und Frauen in 41 verschiedenen Ländern. Der Zeitraum unserer Untersuchung von Sterblichkeitsdaten für Männer und Frauen im Alter von 0-100 Jahren ist von 1990 bis 2016. Die in unserem Fall verwendeten Daten sind die Anzahl der Todesfälle und Anzahl der lebenden Personen in Polen, Tschechien, Ungarn, Slowakei, Slowenien, Lettland und Estland. Wenn die Anzahl der Todesfälle von x-jährigen Personen in einem Land gleich Null ist, ist eine Weiterverarbeitung dieser Daten unmöglich. Daher wird die Anzahl der Todesfälle in diesem Fall durch 1 ersetzt (Dies war der Fall in Estland, Lettland, Slowenien und Ungarn).

2.2 Deskriptive Analyse der Sterblichkeitsverbesserung

In diesem Abschnitt analysieren wir die ein-Jahres-Verbesserung der männlichen und weiblichen Sterblichkeit in der Altersgruppe zwischen 40 und 90 Jahren von 1992 bis 2013, die mithilfe der statistischen Software R berechnet wurde, in den sieben betroffenen Ländern anhand einer Heatmap.

Die Sterblichkeitsverbesserung $\Delta_{x,t}$ vom Jahr t bis zum Jahr $t+1$ im Alter x lässt sich wie folgt ermitteln:

$$\Delta_{x,t} = \frac{q_{x,t} - q_{x,t+1}}{q_{x,t}} = 1 - \frac{q_{x,t+1}}{q_{x,t}}, \quad (2.1)$$

wobei $q_{x,t}$ die Sterbewahrscheinlichkeit im Alter x und im Kalenderjahr t bezeichnet. Somit ist $\frac{q_{x,t+1}}{q_{x,t}}$ das Verhältnis zwischen der Sterblichkeit im Jahr t und im Jahr $t+1$ im Alter x .

Um die allgemeine Tendenz der Sterblichkeit zu verdeutlichen, wurde die Sterbewahrscheinlichkeit $q_{x,t}$ im Laufe der Zeit mithilfe der gleitenden Durchschnitte 5. Ordnung geglättet ([Eck10]). Daher wurde die Sterbewahrscheinlichkeit dann wie folgt bestimmt:

$$q_{x,t}^{gl} = \frac{1}{5}(q_{x,t-2} + q_{x,t-1} + q_{x,t} + q_{x,t+1} + q_{x,t+2}). \quad (2.2)$$

Eine Heatmap ist eine Darstellung von Daten in Form einer Karte oder eines Diagramms, in der die Datenwerte als Farben dargestellt werden. Das Heat Map-Verfahren zeigt die Verteilung einer quantitativen Variable über alle Kombinationen von zwei kategorialen Faktoren. Wenn einer der beiden Faktoren die Zeit repräsentiert, kann die Entwicklung der Variable mit Hilfe der Heatmap leicht eingesehen werden. Zur Darstellung der Werte der quantitativen Variable wird eine Farbverlaufsskala verwendet ([Hea17]). Viele verschiedene Farbschemata können zur Veranschaulichung der Heatmap verwendet werden, mit den jeweiligen Vor- und Nachteilen für die Wahrnehmung. Regenbogenfarbkarten werden oft verwendet, weil Menschen mehr Farbschattierungen als Grau wahrnehmen können, was die Menge der im Bild wahrnehmbaren Details erhöhen würde.

Um in unserem Fall eine Heatmap zu erstellen, benötigen wir einerseits die berechneten Sterblichkeitsverbesserungen, denn die Sterblichkeitsverbesserungen werden durch zwei Faktoren beeinflusst: Alter und Zeit. Andererseits ist das R-Paket *"gplots"* (siehe [WBB⁺19]) notwendig.

Die folgende Abbildung (Abbildung 2.1) zeigt zunächst den Farbverlauf unserer Heatmaps, d.h. eine fortlaufende Änderung von Dunkelblau nach Violett, basierend auf einzelnen Datenwerten. Zunächst einmal zeigt die Farbe Blau an, dass die Mortalität relativ schnell ansteigt (d.h. $\Delta_{x,t} < -5\%$). Dann weist die grüne Farbe darauf hin, dass die Sterblichkeit leicht ansteigt (d.h. $\Delta_{x,t} \in [-5\%, 0\%[$), während die orange Farbe eine leichte Abnahme anzeigt (d.h. $\Delta_{x,t} \in [0\%, 5\%[$). Und schließlich deuten Farben wie Rot und Violett auf starke Verbesserungen hin (d.h. die Sterblichkeitsrate sinkt und $\Delta_{x,t} > 5\%$).

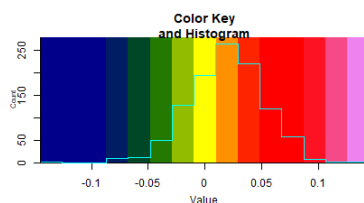
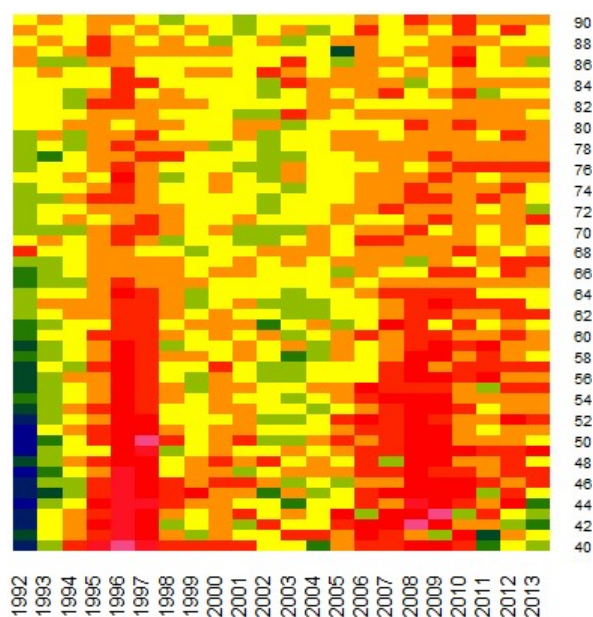
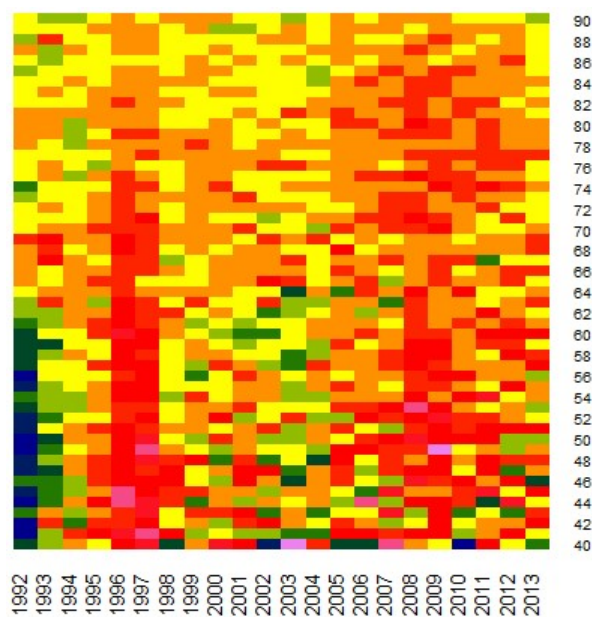


Abbildung 2.1: Farbverlauf der Heatmap

Die folgenden Abbildungen beschreiben die Sterblichkeitsverbesserung in den sieben Ländern durch die Heatmaps.

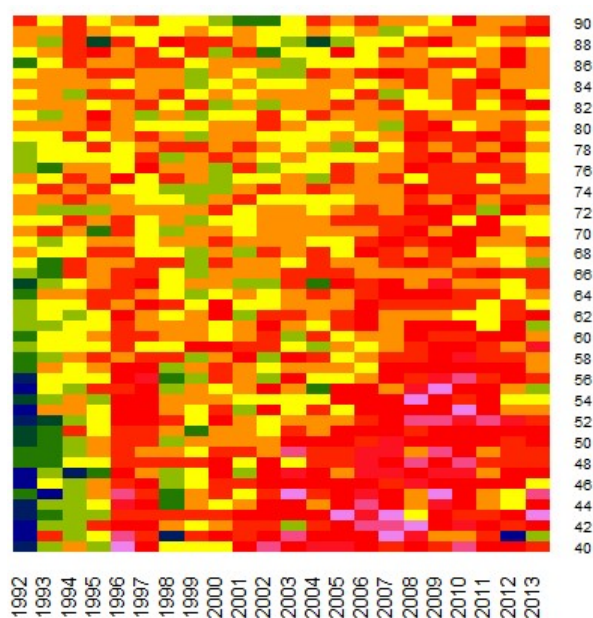


a) Männer

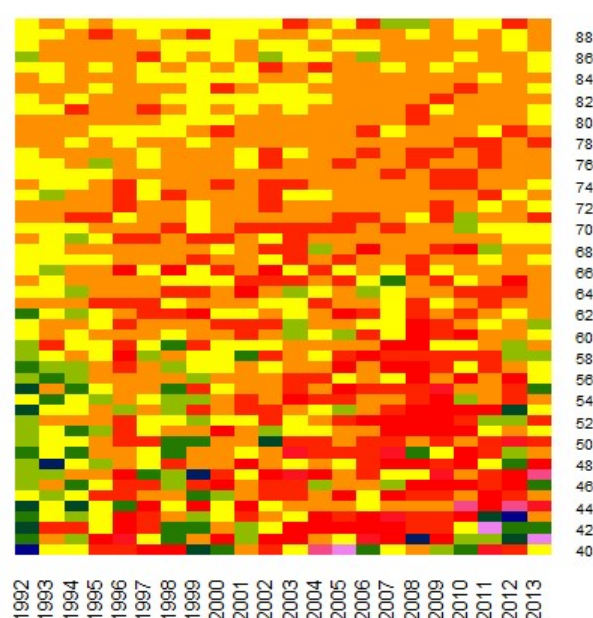


b) Frauen

Abbildung 2.2: Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Lettland

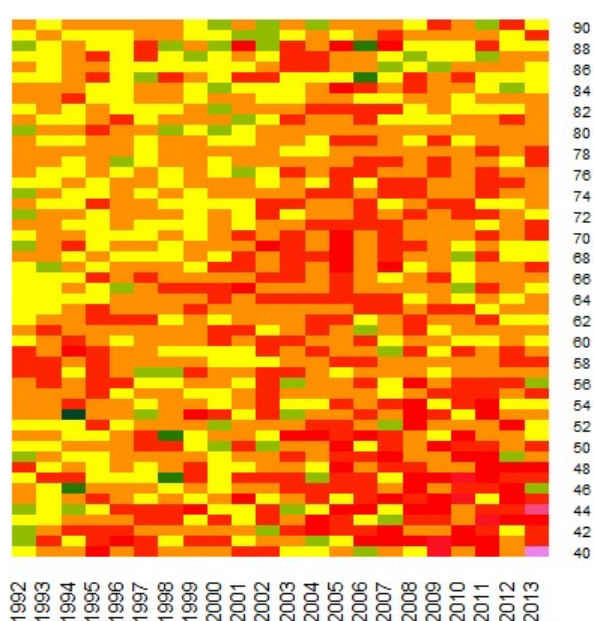


a) Männer

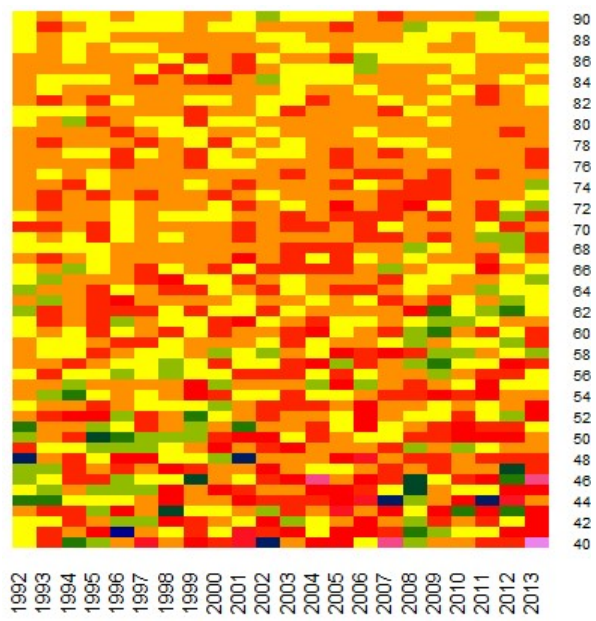


b) Frauen

Abbildung 2.3: Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Estland

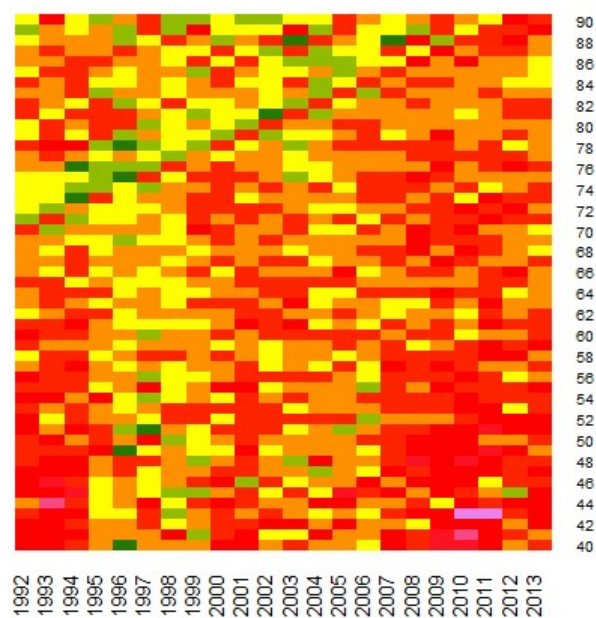


a) Männer

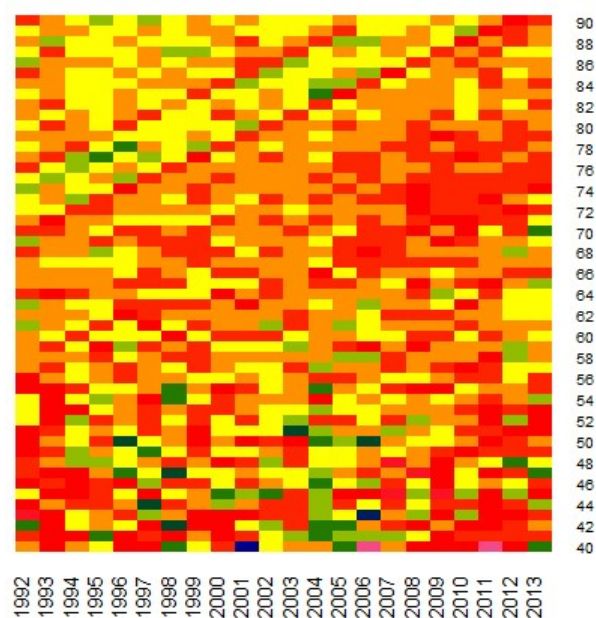


b) Frauen

Abbildung 2.4: Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Slowenien

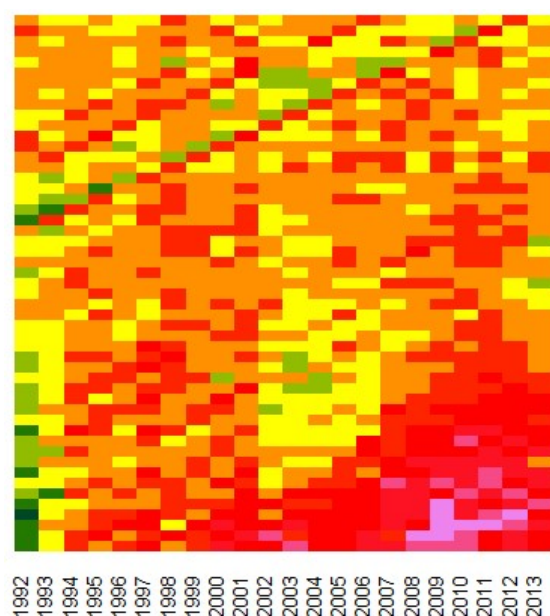


a) Männer

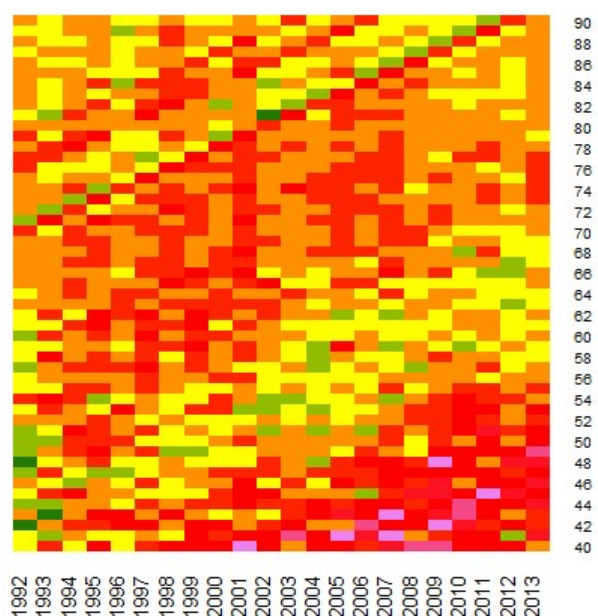


b) Frauen

Abbildung 2.5: Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Slowakei

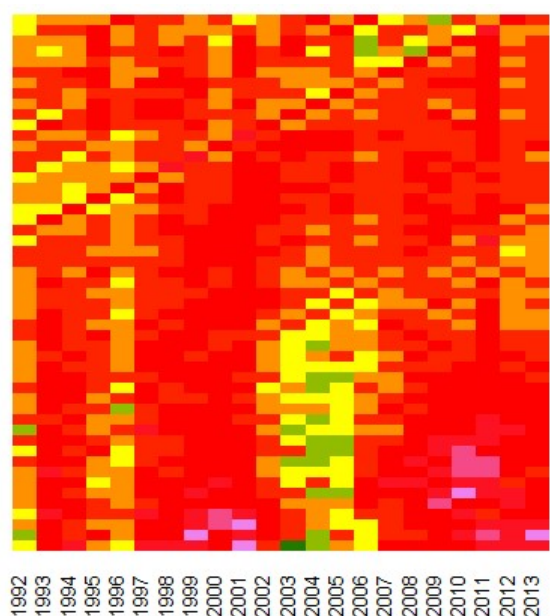


a) Männer

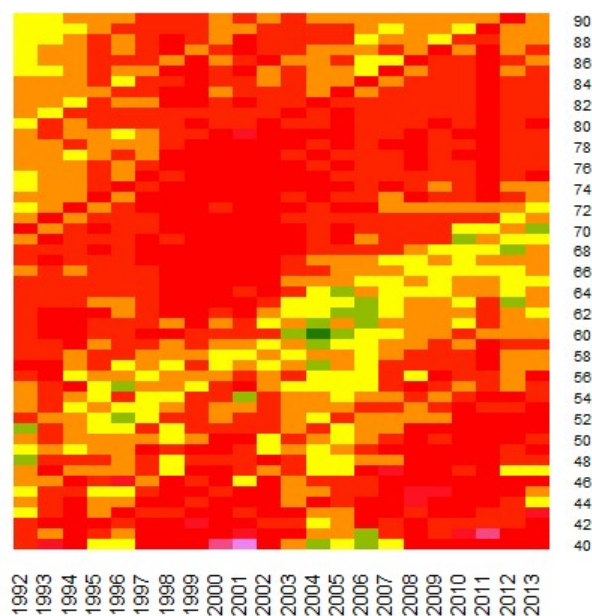


b) Frauen

Abbildung 2.6: Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Ungarn



a) Männer



b) Frauen

Abbildung 2.7: Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Polen

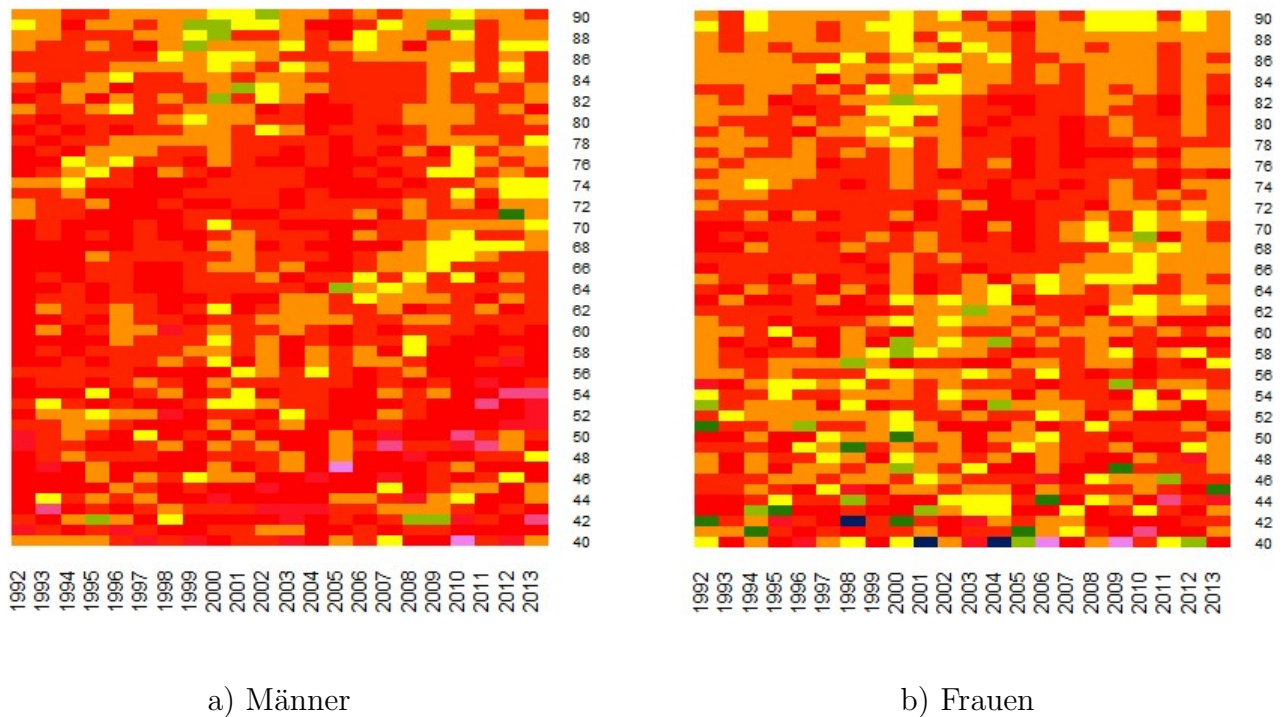


Abbildung 2.8: Sterblichkeitsverbesserung für Männer und Frauen (40-90 J.a.) in Tschechien

Durch diese Heatmaps können *Kohorteneffekte*, *Alterseffekte* und *Periodeneffekte* beobachtet werden.

Eine Kohorte ist eine Gruppe von Menschen mit dem gleichen demographischen Merkmal. Kohorteneffekte treten dann auf, wenn eine Kohorte durch ein bestimmtes Verhalten gekennzeichnet ist. Diese Menschengruppe wird oft mit dem Geburtsjahr oder auch dem Geburtsjahrzehnt in Verbindung gebracht, d.h. Geburtsjahreseffekte werden einigermaßen als Kohorteneffekte betrachtet ([Koh]). In einer Heatmap werden Kohorteneffekte durch farbige Diagonalstrukturen entlang des Bildes von einem Kalenderjahr bis zu mehreren Beobachtungsjahren wahrgenommen. Dies bezieht sich auf Sterblichkeitsverbesserungen von Menschen im selben Geburtsjahr (bzw. in derselben Generation).

Periodeneffekte sind Sterblichkeitsveränderungen, die in einer bestimmten Zeitspanne unabhängig von der Generation (oder vom Alter) auftreten ([Koh]). Das Auftreten einer Epidemie kann durchaus die Ursache für die Beobachtung von Periodeneffekten sein. Sie sind in einer Heatmap durch farbige vertikale Strukturen zu erkennen.

Unter Alterseffekten versteht man die Auswirkungen des Alters auf die Entwicklung

der Sterblichkeit. Sie werden in einer Heatmap dank ihrer horizontalen Struktur identifiziert.

Lettland ist das einzige Land, in dem die Kohorteneffekte nicht signifikant sind (siehe Abbildung 2.2). Im Gegenteil, es werden sowohl bei Männern als auch bei Frauen Kalenderjahreffeekte festgestellt. Zwischen den Jahren 1992-1993 sind vertikale Strukturen von blauen und grünen Farben bei Männern im Alter von 40-78 Jahren und Frauen im Alter von 40-64 Jahren sichtbar. Das Vorhandensein dieser Farben weist darauf hin, dass die Sterblichkeit in dieser Zeit besonders hoch war. Danach wird in den folgenden Jahren ein Rückgang der Mortalität beobachtet. Zwischen den Jahren 1995-1998 sind vertikale Strukturen von orangen und rötlichen Farben sichtbar, wodurch eine starke Verbesserung der Mortalität (durch die Farbe Rot) bei Männern zwischen 40 und 66 Jahren und bei Frauen zwischen 40 und 78 Jahren zu erkennen ist. Zwischen den Jahren 1998-2004 sind sowohl bei Männern als auch bei Frauen keine Verbesserungen zu verzeichnen (d.h. $\Delta_{x,t} \simeq 0\%$). Und schließlich sind ab dem Jahr 2006 Verbesserungen durch vertikale Strukturen in orange und rötlichen Farben genauso wie in den Jahren 1995-1998 zu bemerken.

Estland weist im Vergleich zum Nachbarland Kalenderjahres- und Kohorteneffekte auf (siehe Abbildung 2.3). Erstens ist zu beachten, dass, wie in Lettland, auch in Estland die Sterblichkeit in den Jahren 1992-1993 bei Männern im Alter von 40-66 Jahren und Frauen im Alter von 40-60 Jahren hoch war. In den Jahren 1996-1997 wurden dann bei Männern vertikale Strukturen von orangefarbenen und rötlichen Farben beobachtet, bei Frauen jedoch erst 1996. Immer noch bei den Männern beobachten wir einerseits diagonale Strukturen der roten Farbe ab dem Jahr 2002, ausgehend in den Altersgruppen zwischen 40 und 50 Jahren. Dies zeigt deutlich das Vorhandensein von Kohorteneffekten. Andererseits sind bei Männern zwischen 2007-2012 immer noch Kalenderjahreffeekte durch rote und orangefarbene vertikale Strukturen zu beobachten. Man kann daher sagen, dass die Verbesserung der männlichen Sterblichkeit ab 2002 hoch war (mit $\Delta_{x,t} > 3\%$). Bei den Frauen hingegen ist ab 1996 auch ein Kohorteneffekt ab der Altersgruppe der 60- bis 80-Jährigen durch das Vorhandensein der orangefarbenen Diagonalstrukturen sichtbar. Darüber hinaus werden ab 2002 auch rote Diagonalstrukturen ausgehend bei Frauen der 40-54-jährigen Generation beobachtet, wobei diese roten Diagonalstrukturen jedoch durch das Auftreten bestimmter gelber, grüner und blauer Farben beeinflusst werden, die entweder keine Verbesserung oder einen Anstieg der Mortalität anzeigen. Die Sterblichkeit bei Frauen hat sich sicherlich

verbessert, ist aber immer noch niedriger als bei Männern.

In Abbildung 2.4 zeigt Slowenien auch Kohorteneffekte für Männer und Frauen. Ein Kohorteneffekt wird sowohl bei Männern als auch bei Frauen über die orangefarbenen Diagonalstrukturen ab der Altersgruppe 54-62 Jahre in den Jahren 1992-2013 wahrgenommen. Bei den Männern ist das Vorkommen blauer Farben, die einen starken Anstieg der Mortalität anzeigen, fast nicht vorhanden (nur das Vorkommen von Blau bei 53-jährigen Männern im Jahr 1994). Stattdessen bemerken wir an einigen Stellen das Erscheinen von grünen Farben. Darüber hinaus beobachten wir eine Fülle von orangefarbenen und gelben Farben. Das bedeutet, dass die Sterblichkeit in den meisten Fällen entweder konstant geblieben ist oder sich im Laufe der Zeit verbessert hat. Bei den Frauen hingegen können wir beobachten, dass die Farbe Blau an einigen kleinen Orten recht präsent ist, wie z.B. bei den 48-jährigen Frauen im Jahr 2001 oder bei den 40-jährigen Frauen im Jahr 2002.

Abbildung 2.5 zeigt auch Kohorteneffekte in der Slowakei sowohl für Männer als auch für Frauen. Bei Männern werden zunächst in der Altersgruppe 40-68 Jahre in den Jahren 1992-1995 vertikale Strukturen der Rotfärbung beobachtet. Weiterhin wird ab der Altersgruppe der Männer zwischen 56 und 78 Jahren ab dem Jahr 1999 ein Kohorteneffekt, bestehend aus den Farben Orange und Rot, beobachtet, was bedeutet, dass sich die Mortalität dieser Generation weiter verbessert. Und schließlich werden auch bei Männern zwischen den Jahren 2007 und 2013 Kalenderjahreffekte mit vertikalen roten und orangefarbenen Strukturen beobachtet, die auf einen Rückgang der Mortalität hinweisen. Bei Frauen gibt es jedoch keine Kalenderjahreffekte wie bei Männern, sondern nur einen großen Kohorteneffekt durch die orange und rötliche Diagonalstruktur beginnend in einer Altersgruppe von 56 bis 72 Jahren in den Jahren 1992-2013. Die blaue Farbe ist bei Männern nicht vorhanden und bei Frauen an einigen wenigen Stellen sichtbar.

In Ungarn verhält sich die Verbesserung der Mortalität bei Männern und Frauen fast identisch, außer dass bei Männern in den Jahren 1994-2001 orangefarbene Kalenderjahreffekte auftreten (siehe Abbildung 2.6). Das Jahr 1993 ist auch durch eine leicht erhöhte Mortalität bei Männern zwischen 40 und 58 Jahren und bei Frauen zwischen 42 und 50 Jahren gekennzeichnet (aufgrund der grünen Farbe auf den Heatmaps). Diagonale Strukturen von orangefarbener Farbe sind dann einerseits bei Männern ab der Altersgruppe von 54 bis 72 Jahren ab dem Jahr 1994 und bei Frauen ab der Alters-

gruppe von 48 bis 72 Jahren ab dem Jahr 1993 zu beobachten. Andererseits wurden ab 1994 rote und violette Kohorteneffekte ausgehend bei Männern in einer Altersgruppe bis 42 Jahre und ab 1997 bei Frauen in einer Altersgruppe bis 40 Jahre beobachtet, was auf eine starke Verbesserung der Mortalität dieser Generation hinweist.

Einer der bemerkenswertesten Fälle in dieser Analyse ist der von Polen (siehe Abbildung 2.7). In Polen gibt es nicht nur eine reichliche Rotfärbung und klare Beweise für Kohorteneffekte bei Männern und Frauen, sondern es gibt auch keine Stellen, die auf einen Anstieg der Mortalität durch Anwesenheit der Farbe Blau hinweisen würden. Bei Männern wurden vertikale rote und orangefarbene Strukturen zuerst in den Jahren 1993-1994 für eine Altersgruppe von 40 bis 70 Jahren beobachtet, dann in den Jahren 1997-2002 und 2006-2013 für alle Altersgruppen (d.h. von 40 bis 90 Jahren). Dies würde bedeuten, dass die Sterblichkeit in diesem Land mit der Zeit stark abnimmt. Bei den Männern werden dann auch die roten und orangenen Diagonalstrukturen ab der Altersgruppe von 50 bis 70 Jahren von 1992 bis zum Jahr 2013 beobachtet. Bei den Frauen lassen sich ab 1992 rote und orangene Kohorteneffekte beginnend in den Altersgruppen 50 bis 84 Jahre und ab 1992 in einer Altersgruppe bis 42 Jahre beobachten. Man kann also sagen, dass die Sterblichkeit in Polen immer besser wird.

Der Fall der Tschechischen Republik ist auch deshalb außergewöhnlich, weil die vorherrschende Farbe auf den Heatmaps rot ist (siehe Abbildung 2.8). Bei den Männern kann man in der Tat sagen, dass die durch das Vorhandensein der roten Farbe gekennzeichneten Kohorten- und Kalenderjahreseffekte für fast alle Altersgruppen (d.h. von 40 bis 90 Jahren) festzustellen sind. Bei Frauen ist dies auch der Fall, allerdings wird an einigen Stellen das Vorhandensein von blauen und grünen Farben beobachtet, die bei Männern nicht zu finden sind, wie z.B. das Vorhandensein der Farbe Blau 1998 bei 40-jährigen Frauen oder das Auftreten von Dunkelgrün noch 1998, aber bei 49-jährigen Frauen. Abgesehen von diesen wenigen kleinen Bereichen, die auf einen Anstieg der Sterblichkeit hinweisen, können wir auch feststellen, dass sich die Sterblichkeit in der Tschechischen Republik im Laufe der Zeit enorm verbessert hat.

Am Ende dieser Analyse können wir feststellen, dass in vielen Ländern ab dem Jahr 1992 Kohorteneffekte für die 50-70-jährigen Generationen deutlich sichtbar waren, z.B. in Slowenien, Ungarn, Polen und der Tschechischen Republik. Obwohl alle diese sieben Länder seit 2004 Mitglieder der Europäischen Union sind, wurde nicht in allen von ihnen eine enorme Verbesserung der Sterblichkeit beobachtet, wie z.B. in Lettland.

Allerdings ist eine Verbesserung der Mortalität zu verzeichnen, die vor allem auf die Entwicklung der Medizin zurückzuführen ist. Schon vor den 90er Jahren hat die Medizin riesige Fortschritte gemacht, zum Beispiel die Entwicklung des Antibiotikums Penicillin 1944 zur Bekämpfung bakterieller Infektionskrankheiten, oder 1968 die Anwendung der Knochenmarktransplantation, die es ermöglicht, Blutkrebs zu bekämpfen. Auch in und nach den 90er Jahren hat sich die Medizin weiter entwickelt, wie zum Beispiel die Gentherapie im Jahr 1990, die darin besteht, Gene in die Zellen eines Individuums eindringen zu lassen, um eine Krankheit zu bekämpfen. Darüber hinaus ermöglicht die ab 2011 durchgeführte Immuntherapie die Bekämpfung von Krebs mit Hilfe des körpereigenen Immunsystems (weitere Informationen zu medizinischen Entwicklungen lassen sich auf [\[Pha\]](#) oder [\[Lan\]](#) finden.)

Es ist zu beachten, dass es die älteren Menschen sind, die diese medizinische Behandlung mehr benötigen als die jungen.

2.3 Visualisierung der logarithmischen Sterblichkeitsrate

In diesem Abschnitt werden zunächst Grafiken betrachtet, die zum einen den Verlauf der logarithmischen Sterblichkeitsrate von Männern und Frauen in einer Altersgruppe von 30 bis 100 Jahren im Jahr 1995 und im Jahr 2015 in den sieben Ländern und zum anderen die Entwicklung von der logarithmischen Sterblichkeitsrate bei Männern und Frauen im Alter von 40 Jahren und im Alter von 80 Jahren in den Jahren 1990 bis 2016.

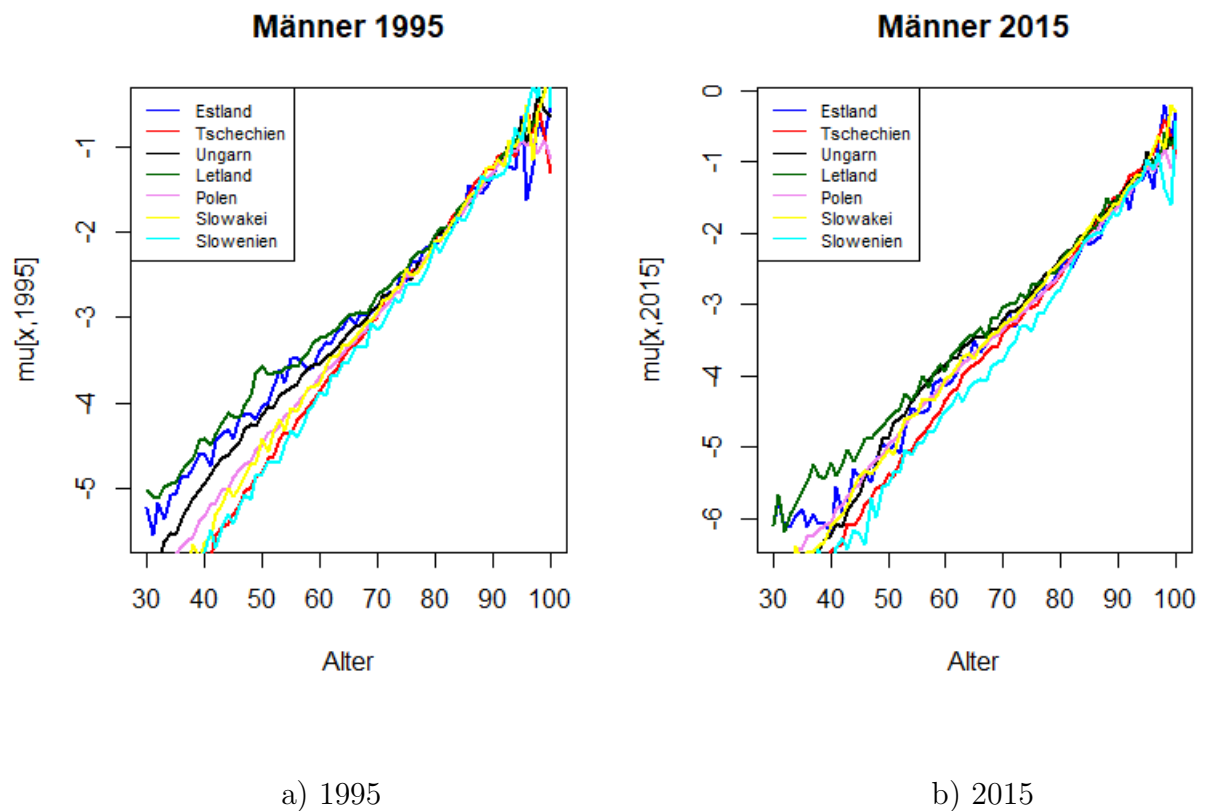


Abbildung 2.9: Verlauf vom Logarithmus der Sterberate ($\mu_{[30:100],t}$) für Männer zwischen 30-100 Jahre alt in 1995 und 2015

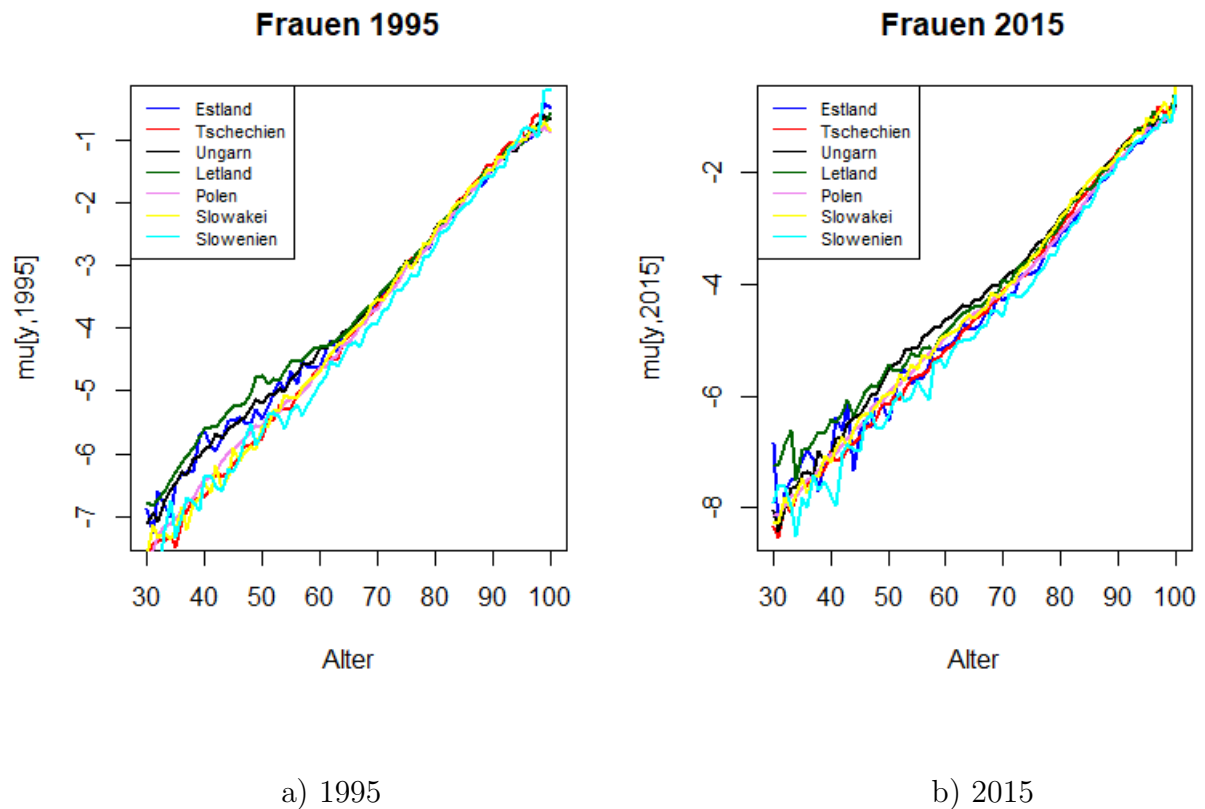


Abbildung 2.10: Verlauf vom Logarithmus der Sterberate ($\mu_{[30:100],t}$) für Frauen zwischen 30-100 Jahre alt in 1995 und 2015

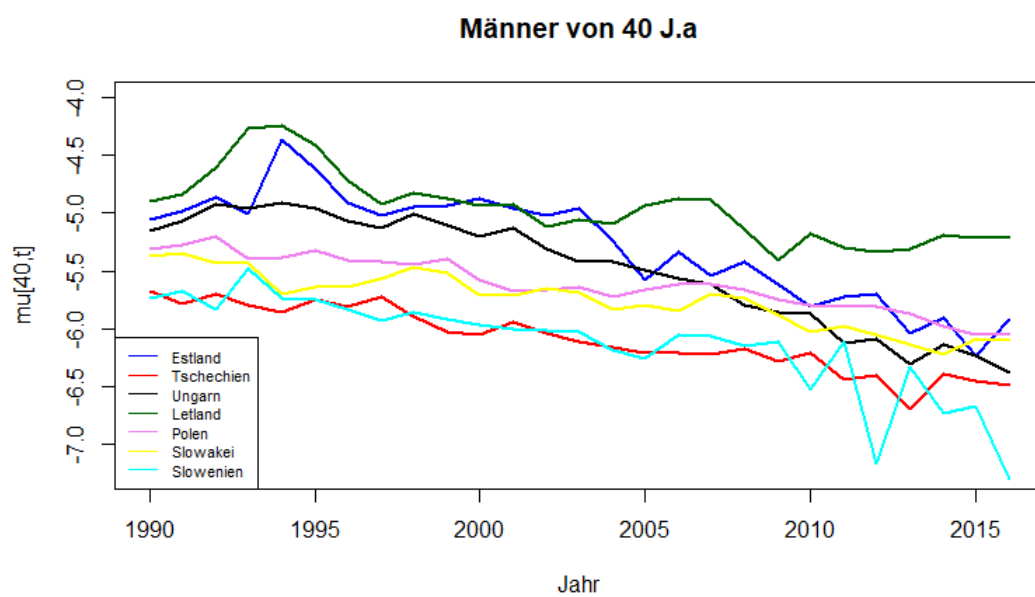


Abbildung 2.11: Verlauf vom Logarithmus der Sterberate ($\mu_{40,[1990:2016]}$) für 40-jährige Männer

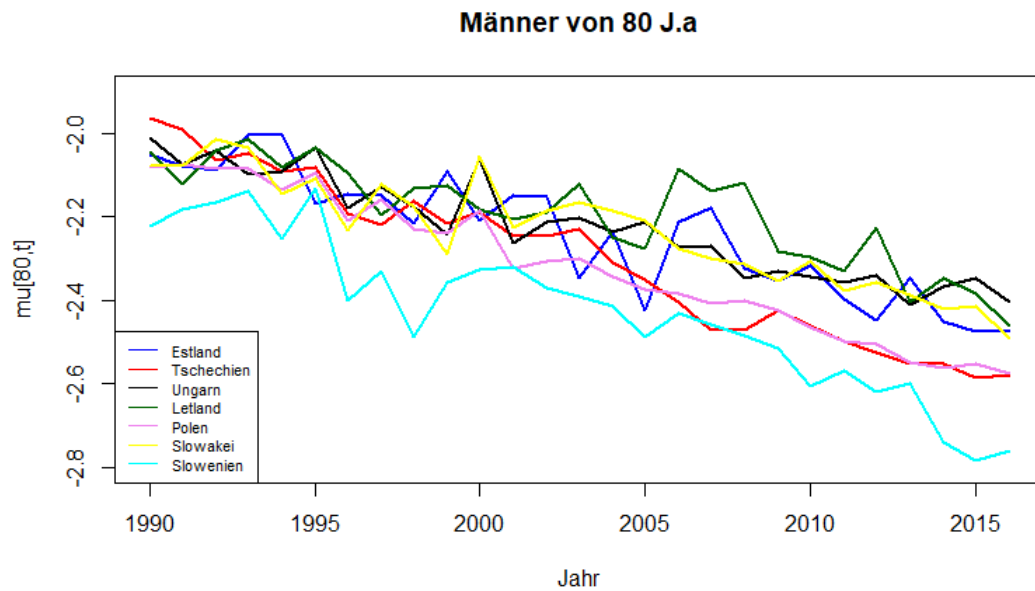


Abbildung 2.12: Verlauf vom Logarithmus der Sterberate ($\mu_{80,[1990:2016]}$) für 80-jährige Männer

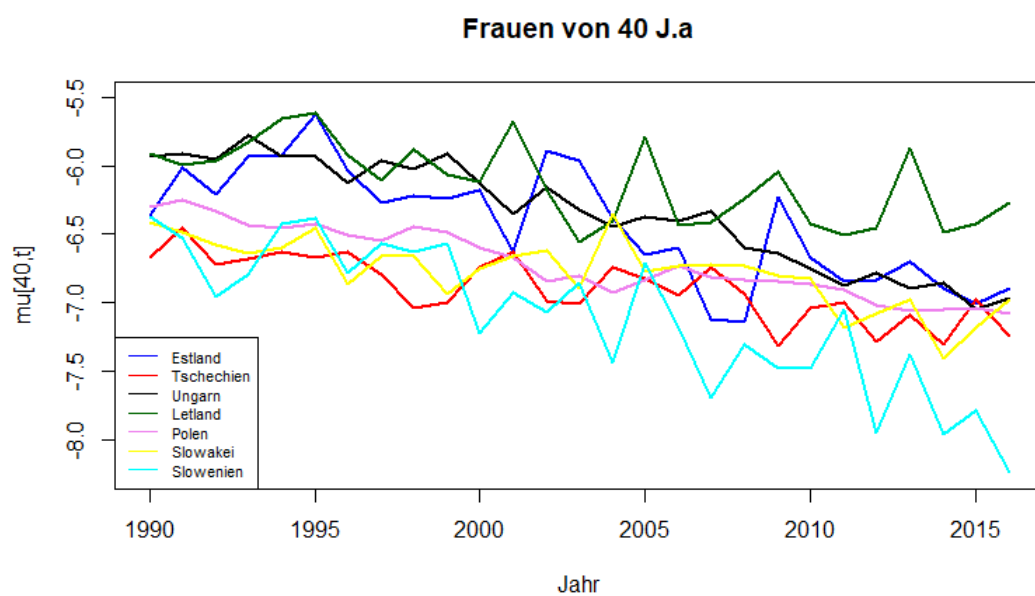


Abbildung 2.13: Verlauf vom Logarithmus der Sterberate ($\mu_{40,[1990:2016]}$) für 40-jährige Frauen

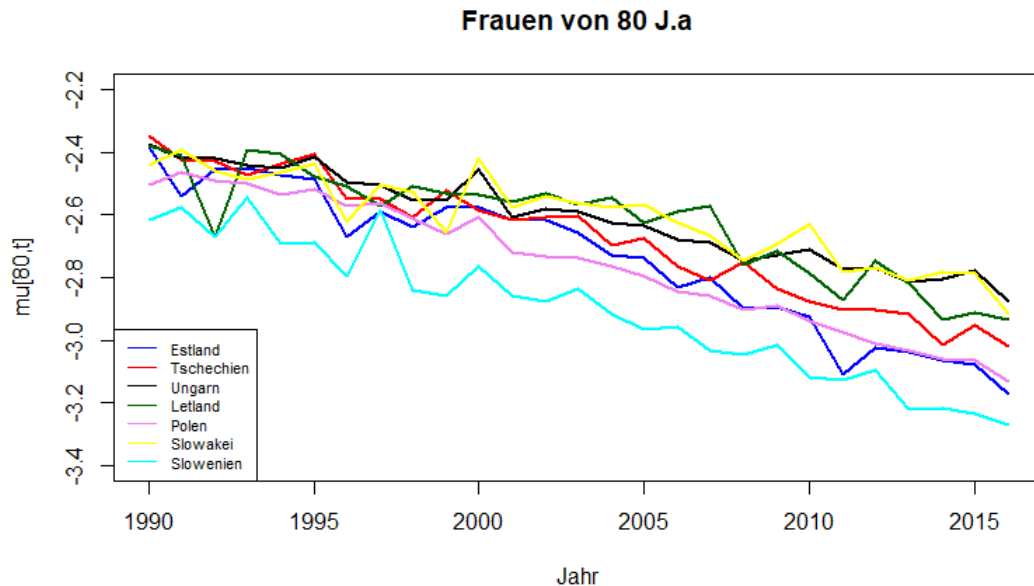


Abbildung 2.14: Verlauf vom Logarithmus der Sterberate ($\mu_{80,[1990:2016]}$) für 80-jährige Frauen

Die Abbildungen 2.9 und 2.10 zeigen uns zunächst die logarithmische Entwicklung der Sterblichkeitsrate von Männern und Frauen im Alter von 30 bis 100 Jahren im Jahr 1995 und im Jahr 2015. Einerseits können wir aus der Abbildung 2.9 erkennen, dass die Sterblichkeit von Männern im Alter von 30-60 Jahren im Jahr 1995 in Estland und Lettland deutlich höher ist als in anderen Ländern. Andererseits ist die Sterblichkeit in Tschechien und Slowenien niedriger als in den anderen Ländern. 2015 hat sich die Sterblichkeit in allen Ländern im Vergleich zu 1995 verbessert. Darüber hinaus ist die Sterblichkeit in Estland in 2015 zurückgegangen, um näher an die anderer Länder heranzukommen, während die Weiterentwicklung in Lettland weniger offensichtlich ist. Im Gegensatz dazu liegt die Sterblichkeit bei Frauen in Estland und Lettland (siehe Abbildung 2.10) bereits nahe an den anderen Ländern. Daher gibt es nicht viel zu sagen, außer dass sich die Sterblichkeit in allen Ländern zwischen 1995 und 2015 ver-

bessert hat.

Zweitens stellen die Abbildungen 2.11 - 2.14 das Verlaufsprofil der logarithmischen Sterblichkeitsrate von Männern und Frauen im Alter von 40 und 80 Jahren von 1990 bis 2016 dar. Diese Abbildungen bestätigen erneut, dass Slowenien die niedrigste und Lettland die höchste Sterblichkeitsrate aufweist. Es zeigt sich auch, dass die Sterblichkeit unabhängig vom Alter mit der Zeit abnimmt. Darüber lässt sich auch beobachten, dass die Abstände zwischen den verschiedenen Verläufe von $\mu_{x,[1990:2016]}$ und $\mu_{y,[1990:2016]}$ der sieben Länder mit zunehmendem Alter abnehmen. Bei 40-jährigen Männern beispielsweise liegen die Verlaufsprofile der logarithmischen Sterberaten all dieser Länder im Wertebereich von $[-7.5, -4]$, während das Intervall bei den 80-jährigen Männern $[-3, -1.75]$ beträgt.

Kapitel 3

Lee-Carter Modell

Das vorliegende Kapitel basiert auf [GK07], [GBRF12] und [MMS19] .

Lee und Carter entwickelten ihren Ansatz speziell für US-Sterblichkeitsdaten zwischen 1933 und 1987, aber die Methode wird heute auf Mortalitätsdaten aus vielen Ländern und Zeiträumen angewandt, weit über die Anwendung hinaus, für die sie konzipiert wurde. Die Methode von Lee-Carter kombiniert ein demographisches Modell mit einem statistischen Zeitreihenmodell, um die Sterblichkeitsraten vorherzusagen.

3.1 Ansatz

Sei $m_{x,t}$ die Sterblichkeitsrate in der Altersgruppe x ($x = 0, \dots, A$) und der Zeit t ($t = t_0, \dots, t_n$) für ein Land. Dann lässt sich $m_{x,t}$ darstellen als:

$$m_{x,t} = \frac{d_{x,t}}{E_{x,t}} \quad ,$$

wobei $d_{x,t}$ und $E_{x,t}$ bzw. die Anzahl von Todesfällen und die Anzahl von Personen, die dem Todesrisiko ausgesetzt sind.

Der erste Schritt des Lee-Carter Modells ist die Modellierung des Logarithmus dieser Sterblichkeitsraten ($\mu_{x,t}$) als:

$$\mu_{x,t} := \ln(m_{x,t}) = \alpha_x + \beta_x \kappa_t + \epsilon_{x,t} \quad (3.1)$$

wobei α_x, β_x altersspezifische Parameter sind, κ_t ein zeitvariabler Parameter ist, und $\epsilon_{x,t} \sim N(0, \sigma^2)$ normalverteilte Fehler mit Erwartungswert 0 und Varianz σ^2 sind.

Genauer gesagt, beschreibt α_x der durchschnittliche Logarithmus von $m_{x,t}$ im betrachteten Kalenderjahrbereich $[t_{min}, t_{max}]$ für das Alter x , β_x teilt mit, wie stark das Alter x auf der Verbesserung der Sterblichkeit reagiert, und κ_t ist die Veränderung der durchschnittliche Sterblichkeit (oder auch Sterblichkeitsverbesserung) im Laufe der Kalenderjahre.

Die Parametrisierung in (3.1) ist nicht eindeutig, da sie in Bezug auf die Transformationen invariant ist:

$$\begin{aligned} (\beta_x, \kappa_t) &\rightarrow (c\beta_x, \frac{1}{c}\kappa_t) \quad \forall c \in \mathbb{R}, c \neq 0 \\ (\alpha_x, \kappa_t) &\rightarrow (\alpha_x - c\beta_x, \kappa_t + c) \quad \forall c \in \mathbb{R}. \end{aligned}$$

Die Parameter β_x und κ_t müssen dann die folgenden Bedingungen zur Identifizierbarkeit des Modells erfüllen:

$$\sum_{x=0}^A \beta_x = 1 \quad \text{und} \quad \sum_{t=t_0}^{t_n} \kappa_t = 0. \quad (3.2)$$

Mithilfe dieser Bedingungen können die Parameter des Lee-Carter Modells bestimmt werden.

3.2 Parameterschätzung

Aus (3.1) lässt sich die folgende Gleichung darstellen:

$$\sum_{t=t_0}^{t_n} \mu_{x,t} = (t_n - t_0 + 1)\alpha_x + \beta_x \sum_{t=t_0}^{t_n} \kappa_t + \sum_{t=t_0}^{t_n} \epsilon_{x,t}. \quad (3.3)$$

Die Bedingung $\sum_{t=t_0}^{t_n} \kappa_t = 0$ ermöglicht uns, die geschätzte α_x (also $\hat{\alpha}_x$) darzustellen als:

$$\hat{\alpha}_x = \frac{1}{t_n - t_0 + 1} \sum_{t=t_0}^{t_n} \mu_{x,t} \quad \text{mit} \quad \sum_{t=t_0}^{t_n} \hat{\epsilon}_{x,t} = 0. \quad (3.4)$$

Angenommen, dass $\mu_{x,t} - \hat{\alpha}_x = \beta_x \kappa_t + \epsilon_{x,t} \sim N(\beta_x \kappa_t, \sigma^2)$ normalverteilt mit Erwartungswert $\beta_x \kappa_t$ und Varianz σ^2 sind, dann können die Parameter β_x und κ_t mithilfe von der Maximum-Likelihood Schätzung geschätzt werden. Aber da diese Methode zur Bestimmung von Parametern nicht ganz überzeugend war, fanden Lee und Carter heraus, dass der optimale Weg, β_x und κ_t zu bestimmen, darin besteht, eine Singulärwertzerlegung

von der Matrix $z_{x,t} = \mu_{x,t} - \hat{\alpha}_x$ durchzuführen.

3.2.1 Singulärwertzerlegung (SWZ)

Gegeben sei die Matrix $\mathbf{Z} \in \mathbb{R}^{(A+1) \times (t_n - t_0 + 1)}$ mit

$$\mathbf{Z} = \begin{pmatrix} z_{0,t_0} & z_{0,t_1} & \cdots & z_{0,t_{n-1}} & z_{0,t_n} \\ z_{1,t_0} & z_{1,t_1} & \cdots & z_{1,t_{n-1}} & z_{1,t_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z_{A-1,t_0} & z_{A-1,t_1} & \cdots & z_{A-1,t_{n-1}} & z_{A-1,t_n} \\ z_{A,t_0} & z_{A,t_1} & \cdots & z_{A,t_{n-1}} & z_{A,t_n} \end{pmatrix} \quad (3.5)$$

$$= \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \cdots \\ \beta_{A-1} \\ \beta_A \end{pmatrix}}_{\boldsymbol{\beta}} \cdot \underbrace{\begin{pmatrix} \kappa_{t_0}, \kappa_{t_1}, \cdots, \kappa_{t_{n-1}}, \kappa_{t_n} \end{pmatrix}}_{\boldsymbol{\kappa}^T} + \underbrace{\begin{pmatrix} \epsilon_{0,t_0} & \epsilon_{0,t_1} & \cdots & \epsilon_{0,t_{n-1}} & \epsilon_{0,t_n} \\ \epsilon_{1,t_0} & \epsilon_{1,t_1} & \cdots & \epsilon_{1,t_{n-1}} & \epsilon_{1,t_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \epsilon_{A-1,t_0} & \epsilon_{A-1,t_1} & \cdots & \epsilon_{A-1,t_{n-1}} & \epsilon_{A-1,t_n} \\ \epsilon_{A,t_0} & \epsilon_{A,t_1} & \cdots & \epsilon_{A,t_{n-1}} & \epsilon_{A,t_n} \end{pmatrix}}_{\boldsymbol{\epsilon}} \quad (3.6)$$

$$= \boldsymbol{\beta} \cdot \boldsymbol{\kappa}^T + \boldsymbol{\epsilon}. \quad (3.7)$$

Es existiert dann orthogonale Matrizen $\mathbf{U} \in \mathbb{R}^{(A+1) \times (A+1)}$ und $\mathbf{V} \in \mathbb{R}^{(t_n - t_0 + 1) \times (t_n - t_0 + 1)}$ sowie eine Matrix $\boldsymbol{\Sigma} = (s_{ij}) \in \mathbb{R}^{(A+1) \times (t_n - t_0 + 1)}$ mit $s_{ij} = 0$ für alle $i \neq j$ und nichtnegativen Diagonaleinträgen $s_{11} \geq s_{22} \geq \cdots \geq s_{kk} > 0$, für die

$$\mathbf{Z} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T \quad (3.8)$$

gilt.

Die Darstellung (3.8) heißt Singulärwertzerlegung von \mathbf{Z} . Gemäß [Fur95] kann eine SWZ wie folgt durchgeführt werden:

Es wird nur der Fall $m := A + 1 \geq t_n - t_0 + 1 =: n$ betrachtet. Man bildet zuerst $\mathbf{A} := \mathbf{Z}^T \mathbf{Z}$, damit man eine quadratische $n \times n$ Matrix bekommt. Danach können die Eigenvektoren (λ_i) von \mathbf{A} berechnet werden. Diese sind nicht negativ und werden in der Reihenfolge $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_k > \lambda_{k+1} = \cdots = \lambda_n = 0$ gegeben, wobei k der Rang von \mathbf{A} ist. Dann kann eine orthonormalbasis $\vec{v}_1, \cdots, \vec{v}_n$ des \mathbb{R}^n gebildet werden, wobei \vec{v}_i der Eigenvektor zum Eigenwert λ_i ist. Die Matrix \mathbf{V} ergibt sich dann als $[\vec{v}_1, \cdots, \vec{v}_n]$

und ist somit orthogonal (d.h. $\mathbf{V}^T = \mathbf{V}^{-1}$). Die Singulärwerte von \mathbf{Z} sind als $s_i = \sqrt{\lambda_i}$ definiert. Die Matrix $\mathbf{\Sigma} = (s_{ij})$ ist diagonal, d.h. für $i \neq j$ ist $s_{ij} = 0$. Die Diagonalelemente sind durch die Singulärwerte gegeben: $s_{ii} = s_i$. Der nächste Schritt besteht darin, die Vektoren \vec{u}_i zu bestimmen. Für $i \leq k$ gilt: $\vec{u}_i = \frac{1}{\sqrt{\lambda_i}} \mathbf{Z} \vec{v}_i$. Diese Vektoren bilden auch ein Orthonormalsystem. Danach ergänzt man diese Vektoren zu einer ON-Basis $\vec{u}_1, \dots, \vec{u}_m$ des \mathbb{R}^m und fasst man sie in der Matrix $\mathbf{U} = [\vec{u}_1, \dots, \vec{u}_m]$ zusammen. Anschließend lässt sich \mathbf{Z} wie in (3.7) darstellen.

Darüber hinaus ist es auch möglich, \mathbf{Z} in dem folgenden Form zu schreiben:

$$\mathbf{Z} = s_1 \cdot \vec{u}_1 \cdot \vec{v}_1^T + \dots + s_k \cdot \vec{u}_k \cdot \vec{v}_k^T \quad (3.9)$$

$$= \sum_{i=1}^k s_i \cdot \vec{u}_i \cdot \vec{v}_i^T, \quad (3.10)$$

wobei $s_1 \cdot \vec{u}_1 \cdot \vec{v}_1^T$ die beste Approximation von \mathbf{Z} durch den Rang 1 ist.

Auf dieser Weise können $\hat{\beta}$ und $\hat{\kappa}$ mithilfe von \vec{u}_1 , \vec{v}_1^T und s_1 ermittelt werden.

Aus

$$s_1 \cdot \vec{u}_1 \cdot \vec{v}_1^T = \hat{\beta} \cdot \hat{\kappa}^T, \quad (3.11)$$

folgt

$$s_1 \cdot u_{1,i} \cdot \vec{v}_1^T = \hat{\beta}_i \cdot \hat{\kappa}^T, \quad (3.12)$$

wobei $u_{1,i}$ und $\hat{\beta}_i$ bzw. die i-te Komponente von \vec{u}_1 und $\hat{\beta}$ sind. Daher kann die folgende Gleichung formuliert werden:

$$s_1 \cdot \sum_{i=1}^m u_{1,i} \cdot \vec{v}_1^T = \sum_{i=1}^m \hat{\beta}_i \cdot \hat{\kappa}^T \quad \text{mit } m = A + 1. \quad (3.13)$$

Mithilfe der Bedingung aus (3.2) kann zuerst $\hat{\kappa}$ ermittelt werden:

$$\hat{\kappa} = (\kappa_{t_0}, \dots, \kappa_{t_n})^T = s_1 \left(\sum_{i=1}^{A+1} u_{1,i} \right) \vec{v}_1. \quad (3.14)$$

Ersetzt man dann $\hat{\kappa}$ aus (3.14) in der Gleichung (3.11), erhält man:

$$\hat{\beta} = (\hat{\beta}_0, \dots, \hat{\beta}_A)^T = \frac{\vec{u}_1}{\sum_{i=1}^{A+1} u_{1,i}}. \quad (3.15)$$

3.3 Random Walk with Drift (RWD)

Um die zukünftige Sterblichkeitsrate zu prognostizieren, gehen Lee und Carter davon aus, dass α_x und β_x über die Zeit konstant bleiben und der Zeitfaktor κ_t an sich als stochastischer Prozess betrachtet wird. Sie schlagen vor, das folgende Random Walk mit Drift Modell (RWD) für κ_t zu verwenden:

$$\hat{\kappa}_t = \hat{\kappa}_{t-1} + \theta + \xi_t \quad (3.16)$$

wobei $\xi_t \sim N(0, \sigma_{rw}^2)$ normalverteilte Fehler und θ der sogenannte *Drift* sind. Dann lässt sich der geschätzte Drift $\hat{\theta}$ wie folgt ermitteln:

Aus

$$\begin{aligned} \hat{\kappa}_{t_1} &= \hat{\kappa}_{t_0} + \theta + \xi_{t_1} \\ \hat{\kappa}_{t_2} &= \hat{\kappa}_{t_1} + \theta + \xi_{t_2} \\ \vdots &= \vdots \\ \hat{\kappa}_{t_{n-1}} &= \hat{\kappa}_{t_{n-2}} + \theta + \xi_{t_{n-1}} \\ \hat{\kappa}_{t_n} &= \hat{\kappa}_{t_{n-1}} + \theta + \xi_{t_n} \quad \text{mit } t_{i+1} = t_i + 1, \end{aligned}$$

folgt,

$$\hat{\kappa}_{t_n} + \sum_{i=1}^{n-1} \hat{\kappa}_{t_i} = \hat{\kappa}_{t_0} + \sum_{i=1}^{n-1} \hat{\kappa}_{t_i} + (t_n - t_0) \cdot \theta + \sum_{i=1}^n \xi_{t_i}.$$

Dann ergibt sich $\hat{\theta}$ als:

$$\hat{\theta} = \frac{\hat{\kappa}_{t_n} - \hat{\kappa}_{t_0}}{t_n - t_0} \quad \text{mit} \quad \sum_{i=1}^n \hat{\xi}_{t_i} = 0. \quad (3.17)$$

Weiterhin ist die geschätzte Varianz der Fehler σ_{rw}^2 gegeben durch:

$$\hat{\sigma}_{rw} = \frac{1}{t_n - t_0} \sum_{i=0}^{n-1} (\hat{\kappa}_{t_{i+1}} - \hat{\kappa}_{t_i} - \hat{\theta})^2 \quad (3.18)$$

Kapitel 4

Sterblichkeitsprognosen für kleine Populationen

In kleinen Bevölkerungen sind die Sterblichkeitsraten durch eine große Volatilität gekennzeichnet. Außerdem sind die Datensätze oft nur für einige Jahre verfügbar und leiden unter fehlenden Daten. Daher können Standard-Sterblichkeitsmodelle (wie z.B. das Lee-Carter Modell) unsichere und biologisch unwahrscheinliche Projektionen liefern. Daher wurde ein weiterer Ansatz zur Modellierung der Sterblichkeit von kleinen Ländern entwickelt.

Das Ziel dieser Methode ist es, die Sterblichkeit von Ländern mit kleiner Bevölkerung durch eine geeignete Mischung von Sterblichkeitsdaten aus den Nachbarländern zu replizieren (siehe [AMOP14]).

4.1 Idee

Laut [AMOP14] ist das Verfahren wie folgt durchzuführen:

Es werden Sterblichkeitsdaten aus $n + 1$ Ländern verwendet, in denen es ein Land 0 gibt, das eine kleine Bevölkerung mit möglichen schweren zufälligen Schwankungen hat. Dann gibt es die anderen Länder i ($i=1,2,\dots,n$) mit großer/mittlerer/ oder kleiner Bevölkerung. vier mögliche Ansätze zur Replikation von Sterblichkeitsdaten werden vorgeschlagen, und nämlich:

- (1) Wahrscheinlichkeitsverteilung der Lebenszeit
- (2) Sterblichkeitsprofil
- (3) Relative Sterblichkeitsverbesserung

(4) Absolute Sterblichkeitsverbesserung.

In [AMOP14] wird erklärt, wie die Wahrscheinlichkeitsverteilung der Lebenszeit von kleiner Ländern repliziert werden kann. Aber wir werden in dem nächsten Abschnitt anders vorgehen, indem wir mit der geschätzten Sterblichkeitsveränderung $\hat{\kappa}_t$ vom Lee-Carter Modell arbeiten werden.

4.2 Replikation von der Sterblichkeitsveränderung

Sei $\hat{\kappa}_t^{(0)}$ die Sterblichkeitsveränderung von dem Land 0 und $\hat{\kappa}_t^{(i)}$ für die Länder $i = 1, \dots, n$. Die Replikation der Sterblichkeitsveränderung für das Land 0 ist gegeben durch:

$$\hat{\kappa}_t^{[AVE]} = \sum_{i=1}^n \omega_i \cdot \hat{\kappa}_t^{(i)}, \quad (4.1)$$

wobei $\hat{\kappa}_t^{[AVE]}$ die durchschnittliche Sterblichkeitsveränderung für die Länder $1, \dots, n$ ist und ω_i das Gewicht für das entsprechende Land i ist.

Zwei Bedingungen in Bezug auf die Gewichte ω_i sollen erfüllt sein:

$$\omega_i \geq 0 \text{ für } i = 1, \dots, n \quad \text{und} \quad \sum_{i=1}^n \omega_i = 1.$$

Der nächste Schritt wäre dann die Gewichte ω_i so zu bestimmen, dass der Abstand zwischen $\hat{\kappa}_t^{(0)}$ und $\hat{\kappa}_t^{[AVE]}$ minimal wird, damit $\hat{\kappa}_t^{[AVE]}$ $\hat{\kappa}_t^{(0)}$ ersetzen kann. Der bekannteste Ansatz dafür ist die sogenannte *Methode der kleinsten Quadrate*.

4.2.1 Methode der kleinsten Quadrate

Mithilfe von [FKL09] besteht die Methode der kleinsten Quadrate in unserem Fall darin $\omega_1, \dots, \omega_n$ zu bestimmen, sodass:

$$KQ(\omega_1, \dots, \omega_n) = \sum_{t=t_0}^{t_n} (\hat{\kappa}_t^{(0)} - \hat{\kappa}_t^{[AVE]})^2 = \sum_{t=t_0}^{t_n} (\hat{\kappa}_t^{(0)} - \sum_{i=1}^n \omega_i \cdot \hat{\kappa}_t^{(i)})^2 \quad (4.2)$$

$$= \sum_{t=t_0}^{t_n} (\hat{\kappa}_t^{(0)} - \omega_1 \cdot \hat{\kappa}_t^{(1)} - \dots - \omega_n \cdot \hat{\kappa}_t^{(n)})^2 \quad (4.3)$$

minimal wird. Anhand der Vektor- und Matrixdarstellung lässt sich ω_i einfach bestimmen.

Die Gleichung

$$\hat{\kappa}_t^{[AVE]} = \omega_1 \cdot \hat{\kappa}_t^{(1)} + \dots + \omega_n \cdot \hat{\kappa}_t^{(n)}, \quad t = t_0, \dots, t_n$$

lässt sich auch in folgender Form darstellen:

$$\hat{\kappa}^{[AVE]} = \begin{pmatrix} \hat{\kappa}_{t_0}^{[AVE]} \\ \hat{\kappa}_{t_1}^{[AVE]} \\ \dots \\ \dots \\ \hat{\kappa}_{t_{n-1}}^{[AVE]} \\ \hat{\kappa}_{t_n}^{[AVE]} \end{pmatrix} = \underbrace{\begin{pmatrix} \hat{\kappa}_{t_0}^{(1)} & \hat{\kappa}_{t_0}^{(2)} & \dots & \hat{\kappa}_{t_0}^{(n-1)} & \hat{\kappa}_{t_0}^{(n)} \\ \hat{\kappa}_{t_1}^{(1)} & \hat{\kappa}_{t_1}^{(2)} & \dots & \hat{\kappa}_{t_1}^{(n-1)} & \hat{\kappa}_{t_1}^{(n)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\kappa}_{t_{n-1}}^{(1)} & \hat{\kappa}_{t_{n-1}}^{(2)} & \dots & \hat{\kappa}_{t_{n-1}}^{(n-1)} & \hat{\kappa}_{t_{n-1}}^{(n)} \\ \hat{\kappa}_{t_n}^{(1)} & \hat{\kappa}_{t_n}^{(2)} & \dots & \hat{\kappa}_{t_n}^{(n-1)} & \hat{\kappa}_{t_n}^{(n)} \end{pmatrix}}_{\mathbf{K}} \cdot \underbrace{\begin{pmatrix} \omega_1 \\ \omega_2 \\ \dots \\ \dots \\ \omega_{n-1} \\ \omega_n \end{pmatrix}}_{\boldsymbol{\omega}} = \mathbf{K}\boldsymbol{\omega}. \quad (4.4)$$

Dann lässt sich die Gleichung (4.3) auf der folgenden Weise umformen:

$$KQ(\boldsymbol{\omega}) = \sum_{t=t_0}^{t_n} (\hat{\kappa}_t^{(0)} - \omega_1 \cdot \hat{\kappa}_t^{(1)} - \dots - \omega_n \cdot \hat{\kappa}_t^{(n)})^2 \quad (4.5)$$

$$= \|\hat{\kappa}^{(0)} - \mathbf{K}\boldsymbol{\omega}\|^2 \quad (4.6)$$

$$= \|\hat{\kappa}^{(0)} - \hat{\kappa}^{[AVE]}\|^2. \quad (4.7)$$

Wird im allgemeinen Fall der Gradient von KQ bezüglich der ω_i gebildet und null gesetzt, dann folgt:

$$\nabla KQ = \nabla \|\hat{\kappa}^{(0)} - \mathbf{K}\boldsymbol{\omega}\|^2 = \nabla \left((\hat{\kappa}^{(0)} - \mathbf{K}\boldsymbol{\omega})^T (\hat{\kappa}^{(0)} - \mathbf{K}\boldsymbol{\omega}) \right) = -2\mathbf{K}^T (\hat{\kappa}^{(0)} - \mathbf{K}\boldsymbol{\omega}) = \mathbf{0},$$

also

$$\mathbf{K}^T \hat{\kappa}^{(0)} = (\mathbf{K}^T \mathbf{K}) \boldsymbol{\omega}.$$

Falls $\mathbf{K}^T \mathbf{K}$ invertierbar ist, lässt sich der Vektor der geschätzten Gewichte $\hat{\boldsymbol{\omega}}$ darstellen als:

$$\hat{\boldsymbol{\omega}} = (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{K}^T \hat{\kappa}^{(0)}. \quad (4.8)$$

Kapitel 5

Auswertung

Die Ergebnisse in diesem Kapitel wurden mit Hilfe der statistischen Software R berechnet.

Dieses Kapitel wird in zwei Hauptteile gegliedert. Der erste Teil umfasst zunächst die Interpretation und den Vergleich der Ergebnisse des Lee-Carter-Modells in den sieben Ländern zwischen 1990 und 2016. Der zweite Teil besteht darin, die durchschnittliche Sterblichkeitsveränderung $\hat{\kappa}_t^{[AVE]}$ in Lettland, Estland und Slowenien anhand von Daten aus Polen, der Slowakei, Ungarn und Tschechien zu replizieren, allerdings in einem Zeitraum von 1990 bis 2005. Nachdem die neuen Sterblichkeitsveränderungen $\hat{\kappa}_t^{[AVE]}$ in diesen drei Ländern repliziert werden, müssen wir natürlich testen, ob dieses Replikationsmodell in unserem Fall gut oder schlecht ist. Wir werden dafür zwei Methoden gebrauchen:

Die erste Methode wird darin bestehen, zu testen, ob die bei der Bestimmung der durchschnittlichen Sterblichkeitsveränderung $\hat{\kappa}_t^{[AVE]}$ ermittelten Gewichte $\hat{\omega}_i$ aus Polen, Ungarn, Slowakei und Tschechien stabil sind oder nicht. Die zweite Methode besteht darin, ein Backtesting durchzuführen, d.h wir werden uns zunächst auf den Zeitraum 1990 bis 2005 beschränken, dann werden wir die Sterblichkeitsveränderung $\hat{\kappa}_t$ dieser drei kleinen Länder anhand von dem Random Walk with Drift im Zeitraum von 2006 bis 2016 zweimal vorhersagen: In der ersten Vorhersage werden wir die zukünftige Sterblichkeitsveränderung (d.h. $\hat{\kappa}_{[2006:2016]}^{RWD}$) anhand der im Lee-Carter-Modell berechneten Sterblichkeitsveränderung $\hat{\kappa}_t$ (oder auch $\hat{\kappa}_t^{LC}$) bestimmen, jedoch mit Abstraktion der nach dem Jahr 2005 kommenden Werte (d.h. $\hat{\kappa}_{[1990:2005]}^{LC}$). Dann werden wir die zweite Vorhersage von $\hat{\kappa}_{[2006:2016]}^{RWD}$ mithilfe der replizierten durchschnittlichen Sterblichkeitsveränderung $\hat{\kappa}_{[1990:2005]}^{[AVE]}$ anstelle von $\hat{\kappa}_{[1990:2005]}^{LC}$ durchführen. Schließlich werden wir die Ergebnisse dieser beiden Vorhersagen interpretieren, indem wir sie mit der aktuellen Sterblichkeitsveränderung der Jahre 2006-2016 des Lee-Carter Modells $\hat{\kappa}_{[1990:2016]}^{LC}$ verglichen werden.

5.1 Lee-Cartell Modell

Mithilfe der berechneten Sterberaten $m_{x,t}$ wurden die Parameter $\hat{\alpha}_x$, $\hat{\beta}_x$ und $\hat{\kappa}_t$ des Lee-Carter Modells bestimmt. Die folgenden Grafiken zeigen den Verlauf der einzelnen Parameter des Lee-Carter Modells.

5.1.1 Verlauf von $\hat{\alpha}_x$

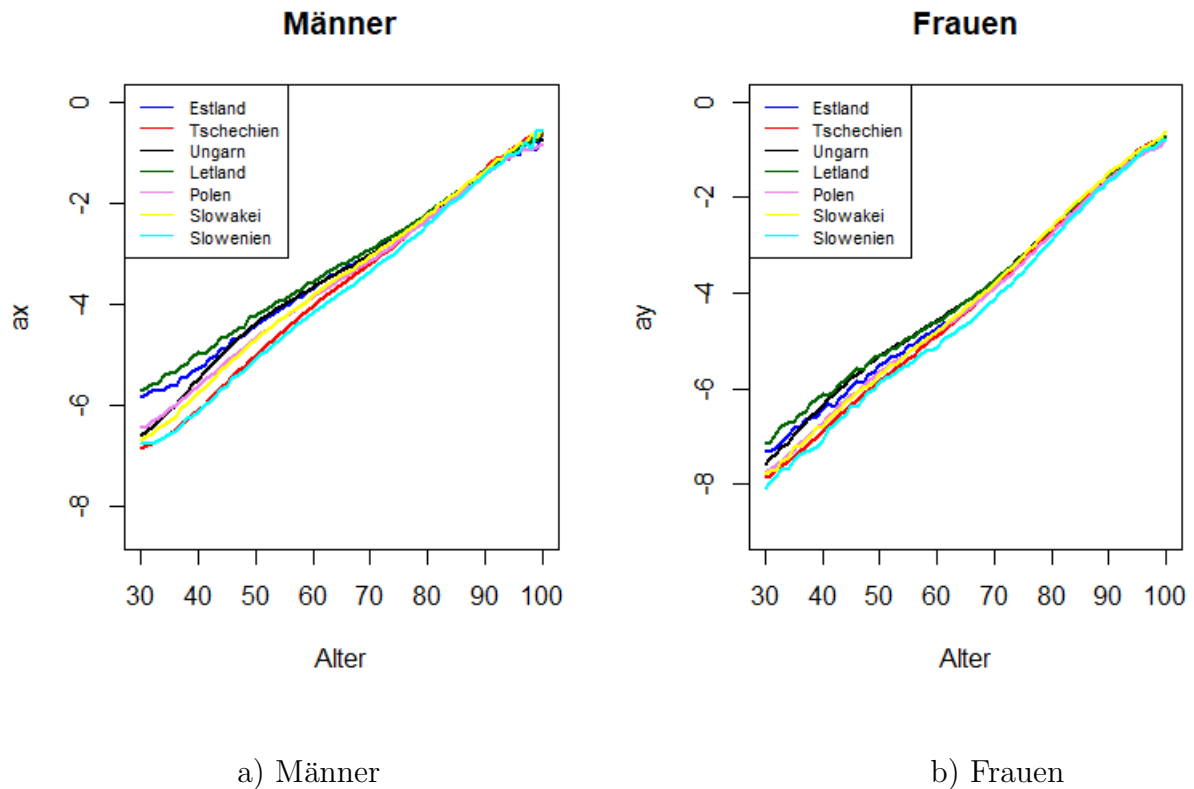
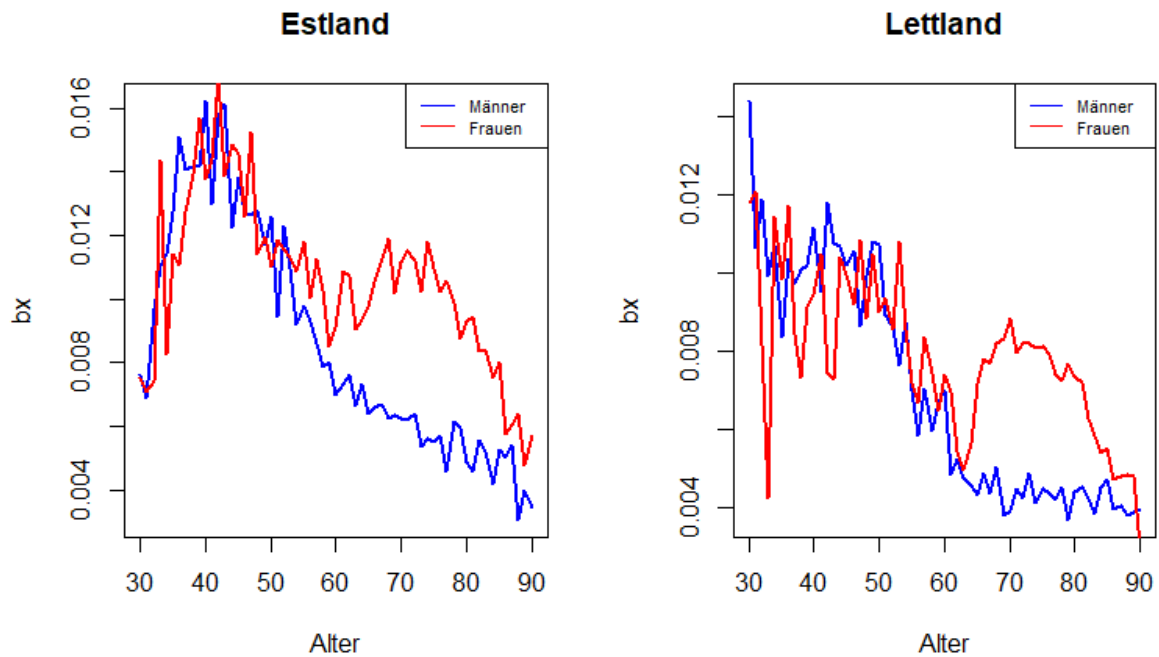


Abbildung 5.1: Verläufe von $\hat{\alpha}_x$ und $\hat{\alpha}_y$ bei Männern und Frauen von 30-100 J.a. in 1990-2016

In der Abbildung 5.1 sind die Verläufe von $\hat{\alpha}_x$ und $\hat{\alpha}_y$ für Männer und Frauen im Alter von 30 bis 100 Jahren im Zeitraum 1990-2016 für alle sieben Länder ersichtlich.

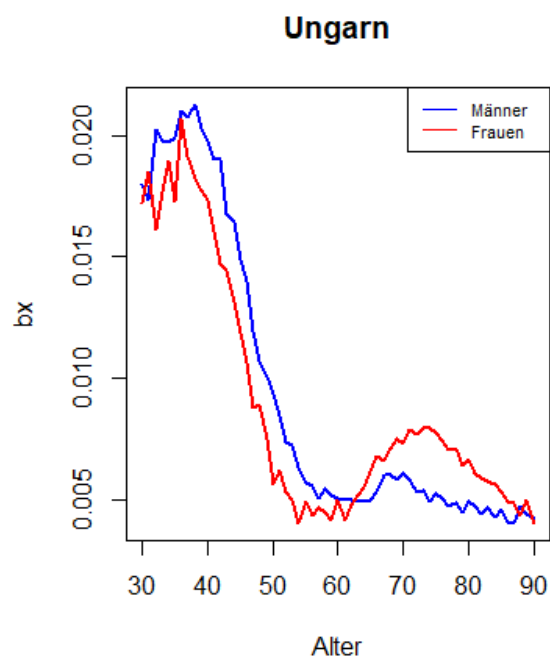
Dieser Verlauf ähnelt dem Verlauf von μ_x und μ_y aus den Abbildungen 5.2 und 5.3, da α_x den durchschnittlichen μ_x für das Alter x im Zeitintervall $[t_{min}, t_{max}]$ beschreibt. Wir können in diesen obigen Grafiken (Abbildung 5.1) sehen, dass die Sterblichkeit in Slowenien und Tschechien am niedrigsten sind, während sie in Lettland am höchsten ist. Außerdem beobachten wir für alle Länder, dass die Sterblichkeitsrate praktisch proportional zum Alter steigt, da es keine Stelle gibt, wo sie drastisch ansteigt. Und schließlich bestätigen diese beiden Abbildungen, dass die Sterblichkeit bei Frauen niedriger ist als bei Männern.

5.1.2 Verlauf von $\hat{\beta}_x$

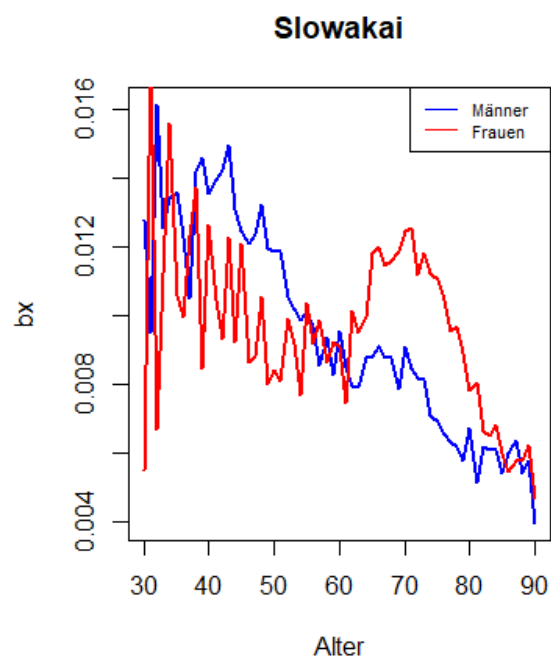


a) Estland

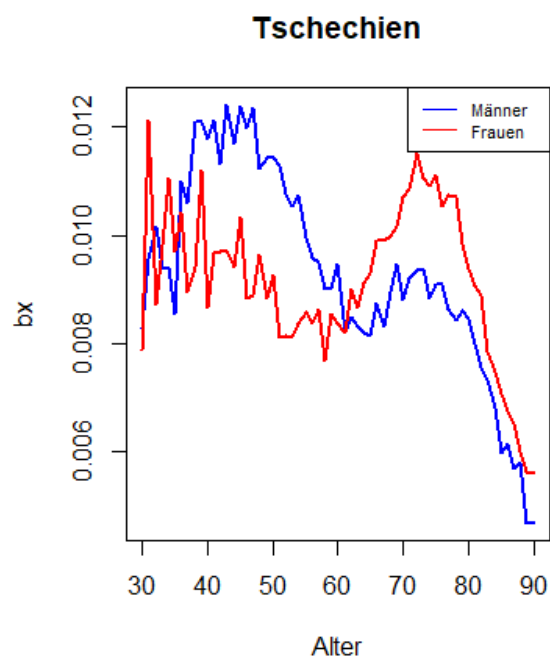
b) Lettland



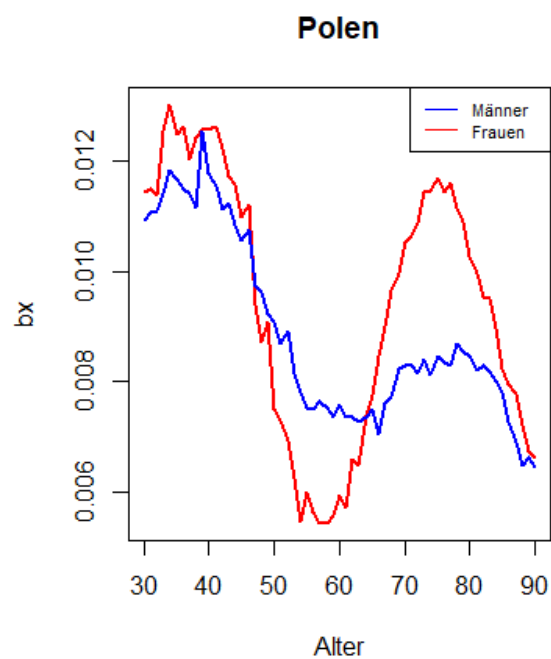
c) Ungarn



d) Slowakei



e) Tschechien



f) Polen

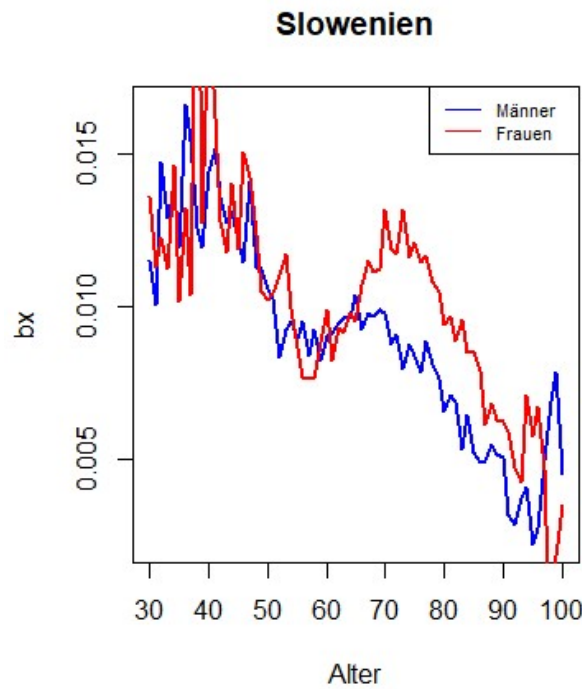


Abbildung 5.2: Verlauf von $\hat{\beta}_x$ bei Männern und Frauen von 30-90 J.a. in 1990-2016

Abbildung 5.2 zeigt den Verlauf von $\hat{\beta}_x$ in jedem Land sowohl für Männer als auch für Frauen. Aufgrund fehlender Daten über die Anzahl der Todesfälle in einigen Ländern, insbesondere bei Frauen, weisen einige Verläufe von $\hat{\beta}_x$ in diesen Ländern Schwankungen auf, wie z.B. in Estland, Lettland, der Slowakei usw.

Wie bereits erwähnt, beschreibt β_x , wie stark das Alter x auf die Veränderung der Sterblichkeit κ_t reagiert. Es lassen sich verschiedene Trends erkennen. Einige Altersgruppen zeigen größere Fortschritte als andere. Allerdings sind die Fortschritte in den einzelnen Ländern unterschiedlich.

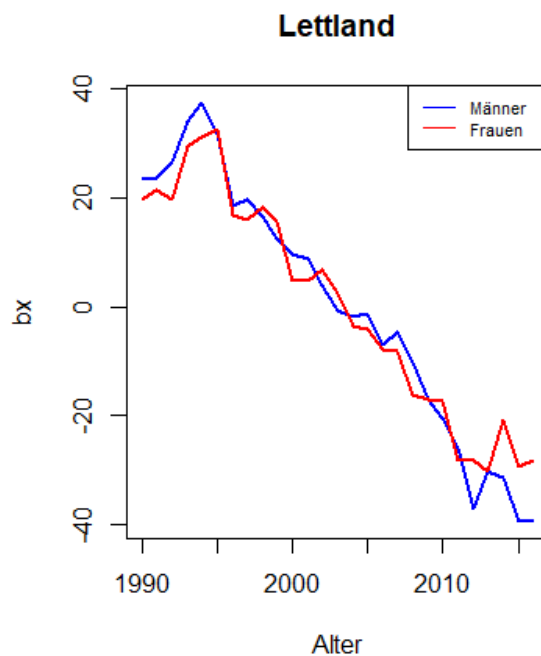
Zunächst fällt auf, dass in fast allen Ländern die Verbesserungen bei den Frauen ab dem 60. Lebensjahr echt höher sind als bei den Männern. Dies könnte erklären, warum Frauen eine bessere Lebenserwartung haben als Männer. In Estland sind die Verbesserungen in der Altersgruppe 30-50 Jahre für Männer und Frauen relativ hoch und ähnlich. Ab dem Alter von 50 Jahren nehmen die Verbesserungen bei Männern jedoch viel schneller ab als bei Frauen (bei Frauen sind die Verbesserungen immer noch ziemlich hoch). Die Verbesserungen in Lettland verhalten sich wie in Estland, aber in der

Altersgruppe 30 bis 60 Jahre, d.h. von 30 auf 60 Jahre, sind die Verbesserungen bei Männern und Frauen fast identisch aber obwohl sie hoch sind, sind sie immer noch geringer als in Estland. Ab dem Alter von 60 Jahren fallen die Verbesserungen bei Männern dann schneller als bei Frauen. In Ungarn sind die größten Verbesserungen im Alter zwischen 30 und 45 Jahren sowohl bei Männern als auch bei Frauen zu beobachten. Einerseits sind die Verbesserungen bei Männern höher als bei Frauen zwischen 30 und 60 Jahren, andererseits sind die Verbesserungen ab dem Alter von 60 Jahren für beide Geschlechter ziemlich gering, aber der Fall bei Männern ist immer noch geringer als der bei Frauen. Die Verbesserungen in der Slowakei sind ähnlich wie in Ungarn. Zwischen dem 30. und 50. Lebensjahr sind die Fortschritte bei den Männern größer als bei den Frauen. Nach dem 50. Lebensjahr nehmen die Fortschritte bei den Männern jedoch mit dem Alter ab, während bei den Frauen noch gute Verbesserungen zu verzeichnen sind. In Tschechien, zwischen 30 und 60 Jahren, sind die Fortschritte bei Männern weitaus größer als bei Frauen, nach 60 Jahren ist das Verhältnis genau umgekehrt. In Polen sind zwei Punkte auffällig: Zum einen gibt es zwischen 50 und 60 Jahren fast keine Verbesserung bei den Frauen, zum anderen übertreffen die Verbesserungen bei Frauen ab 70 Jahren die bei Männern deutlich. Der Verlauf von $\hat{\beta}_x$ in Slowenien, ist ähnlich wie der Verlauf in Estland oder in Lettland.

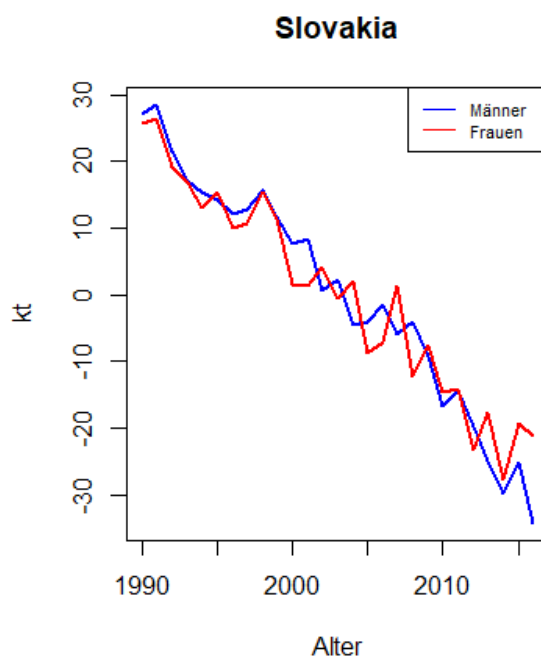
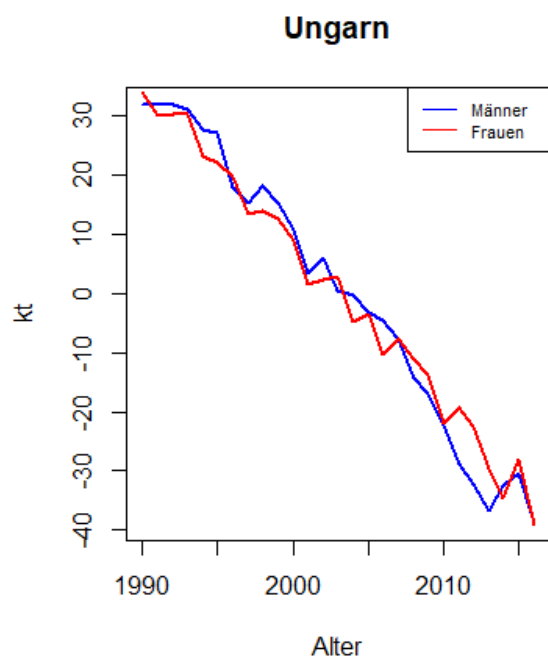
5.1.3 Verlauf von $\hat{\kappa}_t$



a) Estland

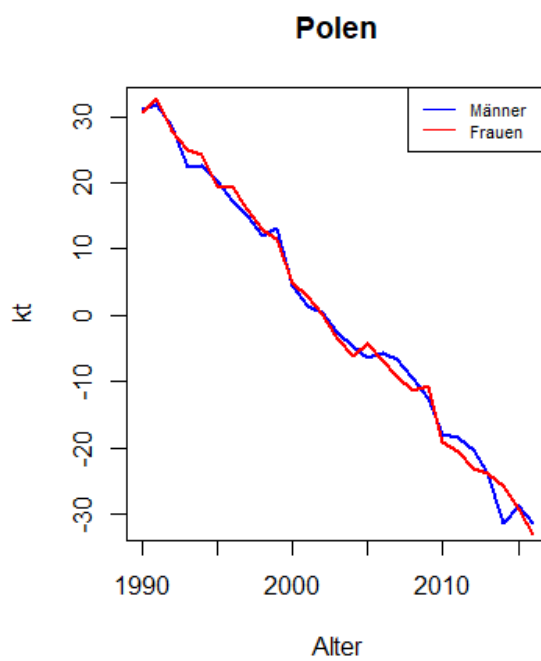
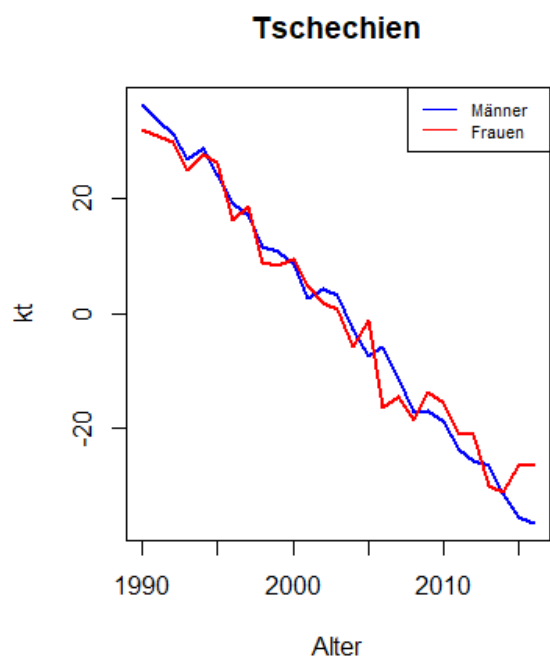


b) Lettland



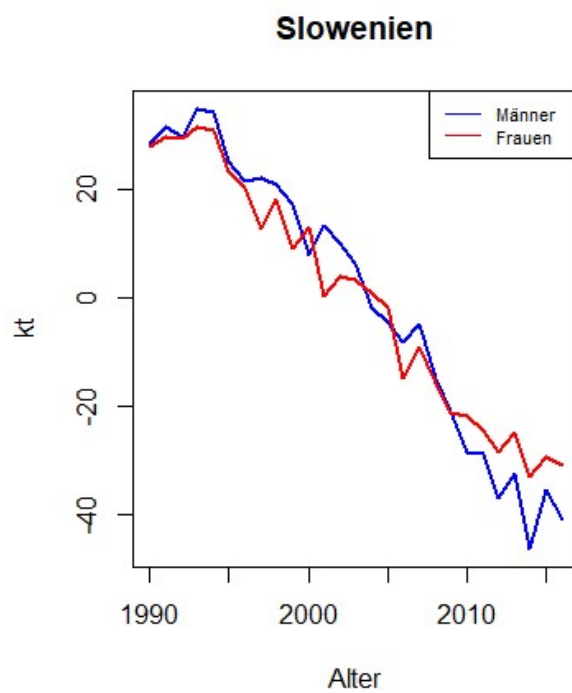
c) Ungarn

d) Slowakei



e) Tschechien

f) Polen



g) Slowenien

Abbildung 5.3: Verlauf von $\hat{\kappa}_t$ bei Männern und Frauen von 30-100 J.a. in 1990-2016

In den obigen Grafiken können wir die Entwicklung von $\hat{\kappa}_t$ für Männer und Frauen in jedem Land beobachten. In all diesen Ländern stellen wir fest, dass die Verläufe von $\hat{\kappa}_t$ für Männer und Frauen recht ähnlich sind, insbesondere in Polen und der Tschechischen Republik. In den anderen Ländern wird manchmal beobachtet, dass die Sterblichkeit bei Männern in einem bestimmten Zeitintervall niedriger ist als bei Frauen, manchmal ist es auch das Gegenteil. In Slowenien beispielsweise ist zu bemerken, dass die Sterblichkeit bei Frauen zwischen 1990 und 2000 einerseits niedriger ist als bei Männern und andererseits ab 2009 höher ist.

Außerdem können wir noch feststellen, dass die Sterblichkeitsrate in allen Ländern im Laufe der Zeit sinkt. Dies ist insbesondere auf den Fortschritt der Medizin zurückzuführen (einige Fortschritte wurden schon in Kapitel 2 erwähnt). Es ist jedoch zu beachten, dass in Estland und Lettland die Sterblichkeit zwischen 1990 und 1995 noch zunimmt, bevor sie in den folgenden Jahren zurückgeht. Dies wurde auch in Kapitel 2 festgestellt, als die Verbesserung der Sterblichkeit in diesen beiden Ländern untersucht wurde. Dann ist in Ländern wie Lettland, Estland und Slowakei (nur für Frauen) der Verlauf von $\hat{\kappa}_t$ ein wenig instabil, denn obwohl $\hat{\kappa}_t$ im Laufe der Zeit sinkt, steigt er manchmal in den folgenden Jahren wieder an, wie z.B. in Estland bei Männern, wo die Sterblichkeit 1996 niedriger ist als 1997, oder in der Slowakei bei Frauen, wo die Sterblichkeit 2005 viel kleiner ist als 2008. Hingegen sind die Verläufe in Polen und Tschechien (bei Männer) sehr stabil, da sie nur im Laufe der Zeit abnehmen.

Mit den Parametern des Modells von Lee-Carter wurde die Sterblichkeit in diesen Ländern gut untersucht. Jeder Parameter spielt eine wichtige Rolle, d.h. der Parameter $\hat{\alpha}_x$ gibt uns einen Überblick über die Sterblichkeitsrate in jeder Altersgruppe (in unserem Fall von 30 bis 100 Jahren). Der Parameter $\hat{\beta}_x$ erlaubte uns dann, Verbesserungen und Fortschritte in verschiedenen Altersintervallen zu sehen, und schließlich gab uns der Zeitfaktor $\hat{\kappa}_t$ einen guten Gesamtüberblick über die Entwicklung der Sterblichkeit im Zeitraum 1990-2016.

5.2 Ergebnisse der Vorhersage der Sterblichkeitsveränderung $(\hat{\kappa}_t^{[AVE]})$ in Estland, Lettland und Slowenien

Hier werden wir $\hat{\kappa}_t^{[AVE]}$ in Lettland, Estland und Slowenien mit Hilfe der Sterblichkeitsveränderung $\hat{\kappa}_t$ in Polen, der Slowakei, Ungarn und Tschechien im Zeitraum 1990-2005 replizieren. Für diese Aufgabe werden wir zunächst die Gewichte genau wie in Kapitel 4 berechnen.

Die folgenden Tabellen enthalten die verschiedenen Gewichte ($\hat{\omega}_i$) bei Männern und Frauen in diesen drei Ländern.

Land	Männer ($\hat{\omega}_i$)	Frauen ($\hat{\omega}_i$)
Polen	-1.18	-0.49
Slowakei	-0.2	-0.21
Tschechien	0.72	0.91
Ungarn	1.53	0.72
Gesamt	0.88	0.92

Tabelle 5.1: Gewichtsdarstellung ($\hat{\omega}_i$) in Estland

Land	Männer ($\hat{\omega}_i$)	Frauen ($\hat{\omega}_i$)
Polen	-0.36	-0.41
Slowakei	-0.46	0.24
Tschechien	0.21	0.70
Ungarn	1.39	0.43
Gesamt	0.78	0.95

Tabelle 5.2: Gewichtsdarstellung ($\hat{\omega}_i$) in Lettland

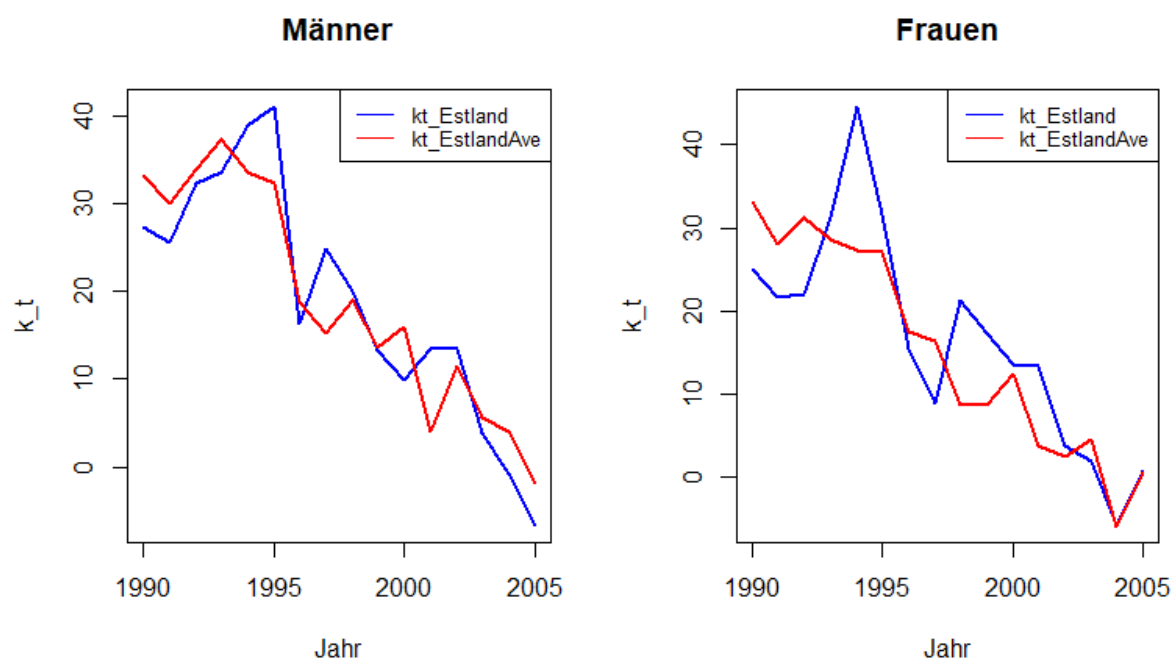
Land	Männer ($\hat{\omega}_i$)	Frauen ($\hat{\omega}_i$)
Polen	-0.85	-0.26
Slowakei	0.44	-0.03
Tschechien	0.43	0.29
Ungarn	1.05	0.99
Gesamt	1.07	0.99

Tabelle 5.3: Gewichtsdarstellung ($\hat{\omega}_i$) in Slowenien

Das erste an diesen auffälligen Ergebnissen ist, dass es Gewichte gibt, die negative Werte haben. Aber das ist in unserem Fall nicht sehr überraschend, denn die Entwicklung der Sterblichkeit verhält sich in jedem Land anders. Dies würde bedeuten, dass für ein Land wie Polen, das eine sehr große Verbesserung der Sterblichkeit hatte, es einigermassen normal wäre, dass es sich nicht für die Bestimmung von $\hat{\kappa}_t^{[AVE]}$ in Estland und oder Lettland, die eine normale Verbesserung hatten, geeignet wäre. Zweitens können wir auch merken, dass die Summe der Gewichte der Männer in Estland und Lettland im Vergleich zu den Frauen nicht sehr nahe bei 1 liegt.

In Estland hat Tschechien die besten Ergebnisse für die Gewichten bei Männern und Frauen, und kann daher als Referenzpopulation gewählt werden. Andererseits sind die Gewichte in der Slowakei und in Polen nicht geeignet, weil sie negativ sind. In Lettland hat Ungarn bei den Männern ein ziemlich hohes Gewicht, ist aber immer noch das geeignetste Land bei der Replikation von der durchschnittliche Sterblichkeit $\kappa_t^{[AVE]}$. Allerdings hat Tschechien bei den Frauen in Lettland das bessere Gewicht. In Slowenien gibt es nichts zu sagen, außer dass Ungarn hervorragend geeignet ist, um die Referenzpopulation zu sein.

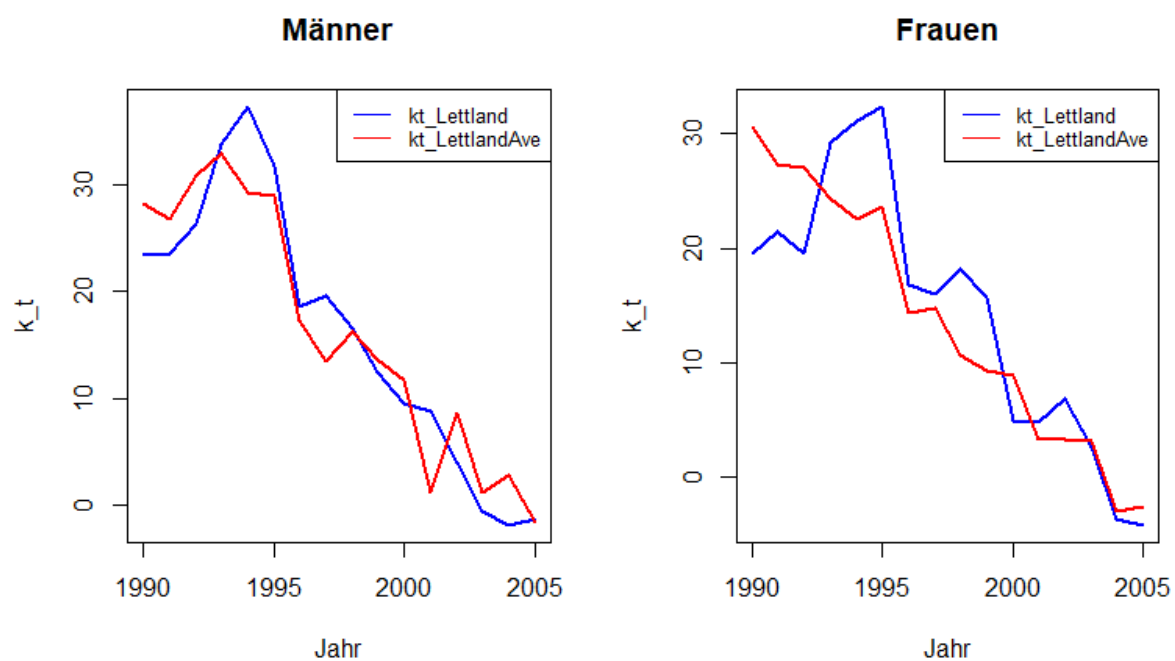
Nach der Bestimmung der Gewichte können die replizierten Sterblichkeitsveränderungen $\hat{\kappa}_{[1990:2005]}^{[AVE]}$ für diese 3 Länder bestimmt werden. Die folgenden Abbildungen veranschaulichen den Vergleich zwischen $\hat{\kappa}_t$ und $\hat{\kappa}_t^{[AVE]}$ in den einzelnen Ländern.



Männer

Frauen

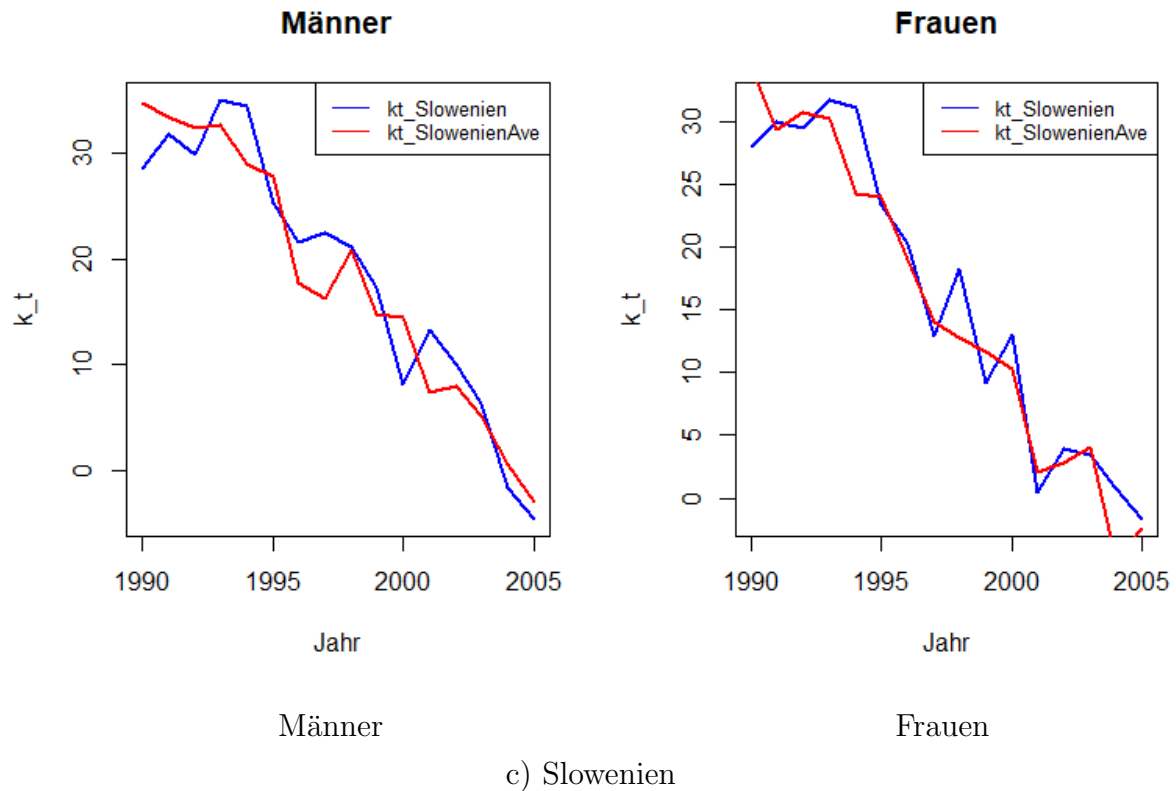
a) Estland



Männer

Frauen

b) Lettland



c) Slowenien

Abbildung 5.4: Vergleich zwischen $\hat{\kappa}_t$ und $\hat{\kappa}_t^{[AVE]}$ in Estland, Lettland und Slowenien

Zwei Methoden können nun getestet werden, um die Qualität von $\hat{\kappa}_t^{[AVE]}$ zu untersuchen: *Die Überprüfung der Stabilität der Gewichte* und *das Backtesting*.

5.3 Überprüfung der Stabilität der Gewichte

Da die Vorhersage der Sterblichkeitsveränderung ($\hat{\kappa}_t^{[AVE]}$) erfolgreich bestimmt wird, müssen wir jetzt prüfen, ob das Replikationsmodell in unserem Fall gut oder schlecht ist. Auf dieser Weise müssen wir prüfen, ob die oben ermittelten Gewichte stabil sind. Um dies erfolgreich zu verifizieren, müssen wir lediglich die Gewichte wie oben berechnen, allerdings im Zeitintervall 2006-2016. Wenn die neu erzielten Ergebnisse im Wesentlichen mit denen des Zeitintervalls 1990-2005 übereinstimmen, dann werden die Gewichte als stabil, ansonsten aber als instabil bezeichnet.

Die folgenden Tabellen vergleichen die berechneten Gewichte zwischen den Zeiträumen 1990-2005 und 2006-2016 in Estland, Lettland und Slowenien.

Land	Männer $\hat{\omega}_{[1990:2005]}$	Männer $\hat{\omega}_{[2006:2016]}$	Frauen $\hat{\omega}_{[1990:2005]}$	Frauen $\hat{\omega}_{[2006:2016]}$
Polen	-1.18	0.33	-0.49	0.96
Slowakei	-0.2	0.02	-0.21	0.48
Tschechien	0.72	0.15	0.91	0.26
Ungarn	1.53	0.76	0.72	-0.34
Gesamt	0.88	1.26	0.92	1.36

Tabelle 5.4: Gewichtsdarstellung ($\hat{\omega}_i$) in Estland

Land	Männer $\hat{\omega}_{[1990:2005]}$	Männer $\hat{\omega}_{[2006:2016]}$	Frauen $\hat{\omega}_{[1990:2005]}$	Frauen $\hat{\omega}_{[2006:2016]}$
Polen	-0.36	0.05	-0.41	1.36
Slowakei	-0.46	0.16	0.24	0.17
Tschechien	0.21	0.39	0.70	0.42
Ungarn	1.39	0.49	0.43	-0.78
Gesamt	0.78	1.09	0.95	1.17

Tabelle 5.5: Gewichtsdarstellung ($\hat{\omega}_i$) in Lettland

Land	Männer $\hat{\omega}_{[1900:2005]}$	Männer $\hat{\omega}_{[2006:2016]}$	Frauen $\hat{\omega}_{[1900:2005]}$	Frauen $\hat{\omega}_{[2006:2016]}$
Polen	-0.85	2.25	-0.26	0.39
Slowakei	0.44	-0.78	-0.02	0.41
Tschechien	0.43	-0.78	0.29	0.47
Ungarn	1.05	0.66	0.98	-0.04
Gesamt	1.07	1.35	0.99	1.22

Tabelle 5.6: Gewichtsdarstellung ($\hat{\omega}_i$) in Slowenien

Der Unterschied zwischen den Gewichten der beiden Zeiträumen ist eindeutig. Zunächst einmal ist festzustellen, dass Polen im Zeitraum 2006-2016 positive Gewichte bei Männern und Frauen in Estland, Lettland und Slowenien hat, während seine Gewichte im Zeitraum 1990-2005 negativ waren. Darüber hinaus hat Polen in einigen Fällen einen starken Einfluss auf das Replikationsmodell zwischen 2006 und 2016, wie z.B. in Estland mit $\hat{\omega}_{[2006:2016]}^{EST/POL} = 0,96$ bei Frauen oder in Lettland mit $\hat{\omega}_{[2006:2016]}^{LET/POL} = 1,35$ bei Frauen. Bei den Gewichten für die Slowakei gibt es ebenfalls Unterschiede in den beiden Zeiträumen, abgesehen von dem Fall in Lettland bei den Frauen, bei dem $\hat{\omega}_{[1990:2005]}^{LET/SLK} = 0,24$ und $\hat{\omega}_{[2006:2016]}^{LET/SLK} = 0,17$. Außerdem kann es behauptet werden, dass in Tschechien und Ungarn die Gewichte der beiden Zeitintervalle ebenfalls unterschiedlich sind.

Nach Prüfung der Stabilität der Gewichte kann daher festgestellt werden, dass die Gewichte nicht stabil sind, und daher ist unser Replikationsmodell gemäß dieser 1. Methode nicht angemessen.

Jetzt müssen wir also prüfen, ob das Backtesting das gleiche Ergebnis liefert wie die Überprüfung der Stabilität der Gewichte.

5.4 Backtesting

In der zweiten Methode werden wir die zukünftigen Sterblichkeitsveränderungen $\hat{\kappa}_t^{RWD}$ im Zeitraum 2006-2016 in Estland, Lettland und Slowenien mit Hilfe der in Kapitel 3 beschriebenen Random Walk with Drift bestimmen und dabei zwei Fälle unterscheiden: Der erste Fall ist die Berechnung von $\hat{\kappa}_{[2006:2016]}^{RWD}$ auf Basis der Sterblichkeitsveränderungen $\hat{\kappa}_{[1990:2005]}^{LC}$ von dem Lee-Carter-Modell (also $\hat{\kappa}_{[2006:2016]}^{LC/RWD}$). Der zweite Fall wird darin bestehen, $\hat{\kappa}_{[2006:2016]}^{RWD}$ auf Basis der replizierten Veränderungen der Sterblichkeit $\hat{\kappa}_{[1990:2005]}^{[AVE]}$ zu bestimmen, nämlich $\hat{\kappa}_{[2006:2016]}^{[AVE]/RWD}$.

Zunächst muss in jedem der beiden Fälle der Drift Parameter $\hat{\theta}$ bestimmt werden, der es uns erlaubt, zukünftige Veränderungen der Sterblichkeit $\hat{\kappa}_{[2006:2016]}^{RWD}$ zu visualisieren. Die folgenden Tabellen stellen die ermittelten Drift-Parameter $\hat{\theta}$ dar, die in jedem der beiden Fälle berechnet wurden.

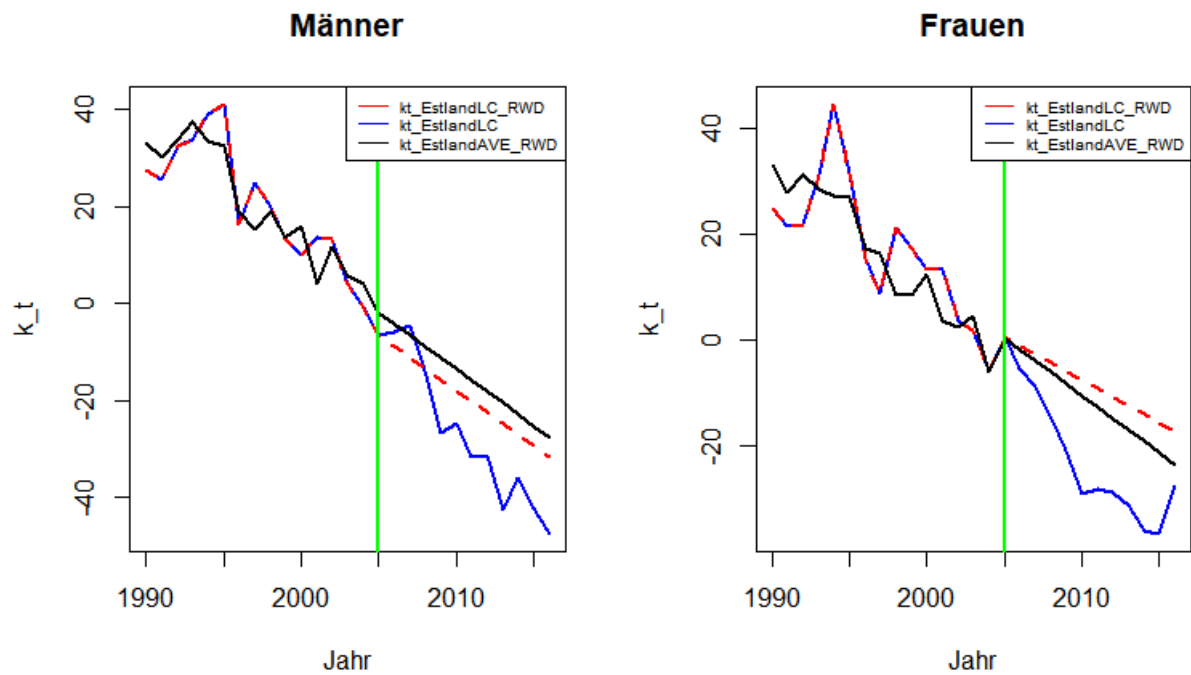
Land	Männer ($\hat{\theta}$)	Frauen ($\hat{\theta}$)
Estland	-2.27	-1.62
Lettland	-1.65	-1.58
Slowenien	-2.21	-1.97

Tabelle 5.7: Drift ($\hat{\theta}$) berechnet aus $\hat{\kappa}_t$ (Lee-Carter) zwischen 1990 und 2005

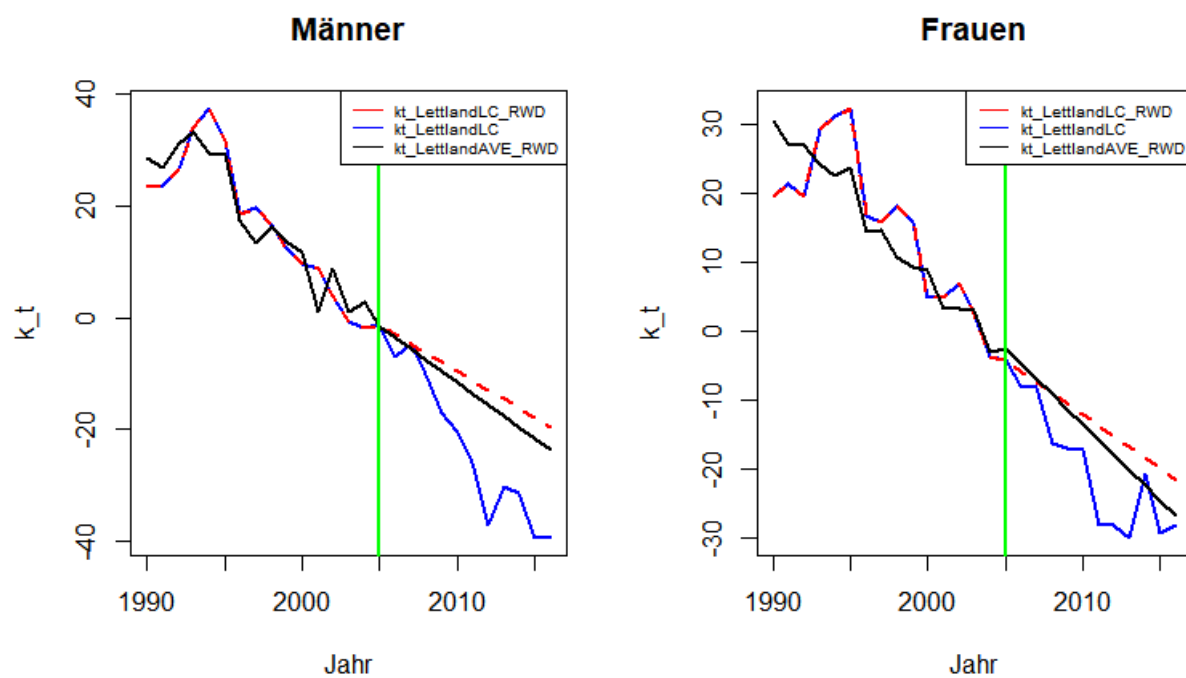
Land	Männer ($\hat{\theta}$)	Frauen ($\hat{\theta}$)
Estland	-2.34	-2.17
Lettland	-2	-2.21
Slowenien	-2.51	-2.45

Tabelle 5.8: Drift ($\hat{\theta}$) berechnet aus $\hat{\kappa}_t^{[AVE]}$ (Replikation) zwischen 1990 und 2005

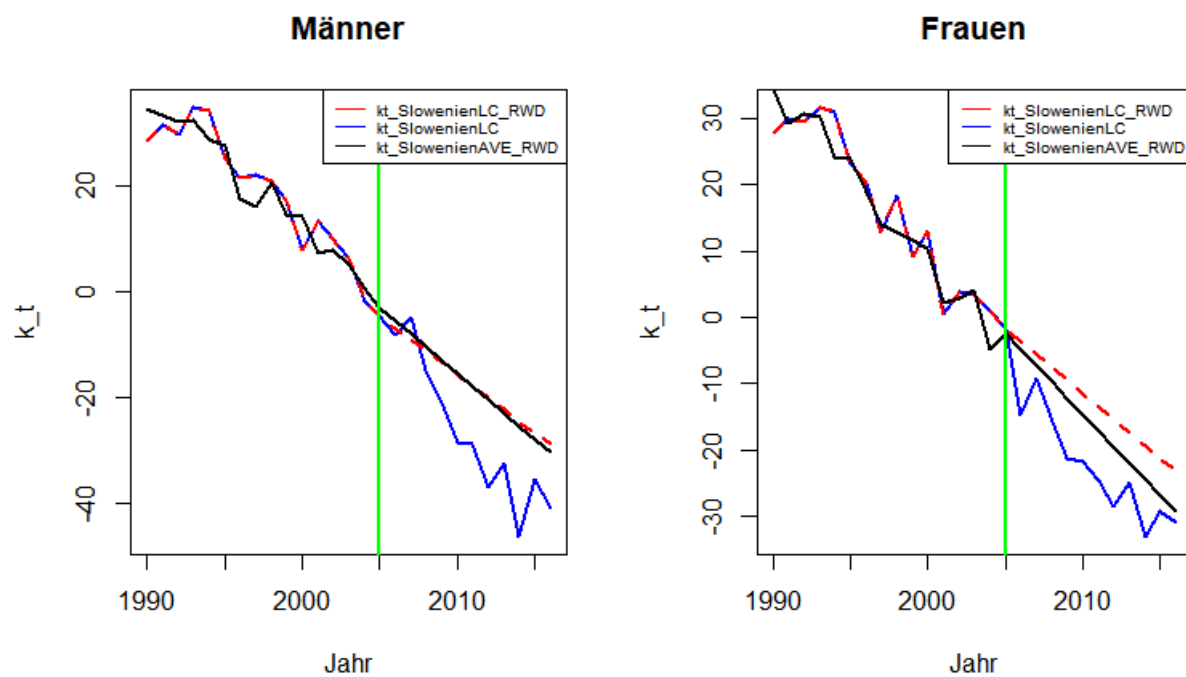
Der nächste Schritt ist jetzt die Bestimmung zukünftiger Sterblichkeitsveränderungen ($\hat{\kappa}_t^{RWD}$). Abbildung 5.5 verdeutlicht den Vergleich zwischen der aktuellen Sterblichkeitsveränderung aus dem Lee-Carter Modell ($\hat{\kappa}_{[1990:2016]}$ bzw. $\hat{\kappa}_{[1990:2016]}^{LC}$), der zukünftigen Sterblichkeitsveränderung ($\hat{\kappa}_t^{LC/RWD}$) auf Basis von $\hat{\kappa}_{[1990:2005]}^{LC}$ (Lee-Carter Modell in 1990-2005) und der zukünftigen Sterblichkeitsveränderung ($\hat{\kappa}_t^{[AVE]/RWD}$) auf Basis von $\hat{\kappa}_{[1990:2005]}^{[AVE]}$.



a) Vergleich in Estland.



b) Vergleich in Lettland.



c) Vergleich in Slowenien.

Abbildung 5.5: Vergleich zwischen $\hat{\kappa}_t$, $\hat{\kappa}_t^{RWD}$ und $\hat{\kappa}_t^{[AVE]/RWD}$

Die blaue Linie bezeichnet den Verlauf von κ_t^{LC} , die rote Linie den Verlauf von $\kappa_t^{LC/RWD}$ und die schwarze Linie den Verlauf von $\kappa_t^{[AVE]/RWD}$. Obwohl $\hat{\kappa}_{[2006:2016]}^{LC/RWD}$ die Vorhersage von $\hat{\kappa}_{[1990:2005]}^{LC}$ ist, scheint erstaunlicherweise in fast allen Fällen, $\hat{\kappa}_{[2006:2016]}^{[AVE]/RWD}$ näher an $\hat{\kappa}_{[2006:2016]}^{LC}$ zu liegen als $\hat{\kappa}_{[2006:2016]}^{LC/RWD}$. Beide Vorhersagen sind jedoch noch recht weit von den aktuellen Werten $\hat{\kappa}_{[1990:2016]}^{LC}$ entfernt. Das Problem liegt darin, dass sich die Veränderung der Sterblichkeit vom Lee-Carter-Modell $\hat{\kappa}_t^{LC}$ in diesen drei Ländern nach 2005 viel besser entwickelt hat, und da der Drift $\hat{\theta}$ aus den Daten von 1990 bis 2005 bestimmt wird, sind die Kurven von $\hat{\kappa}_{[2006:2016]}^{LC/RWD}$ und $\hat{\kappa}_{[2006:2016]}^{[AVE]/RWD}$ tendenziell höher als die Kurve von $\hat{\kappa}_{[2006:2016]}^{LC}$. Dieses Problem wurde erstmals beim Vergleich der Gewichte bemerkt, da Länder wie Polen, die zwischen 1990 und 2005 negative Gewichte hatten, zwischen 2006 und 2016 positive Gewichte hatten. Mit anderen Worten: Man kann keine unvorhersehbaren Ereignisse vorhersagen.

Auf dieser Weise wäre es uns auch unmöglich zu sagen, dass beide Vorhersagen in unserem Fall geeignet sind. Daher können wir sagen, dass das Replikationsmodell zur Ermittlung von $\hat{\kappa}_t^{[AVE]}$ nicht gut ist.

Kapitel 6

Fazit und Ausblick

Die vorliegende Arbeit beschäftigte sich mit der Analyse der Sterblichkeit in Osteuropa.

Einerseits bestand der erste Schritt unserer Aufgabe darin, die Verbesserung der Sterblichkeit in Estland, Lettland, Slowenien, Polen, der Slowakei, Ungarn und Tschechien mit Hilfe einer Heatmap zu untersuchen. Einige Länder wie Lettland zeigten nur Kalenderjahreffekte, andere wie Ungarn oder Slowenien nur Kohorteneffekte, während Länder wie Polen beide Effekte aufwiesen. Andererseits sollte auch die logarithmische Sterblichkeitsrate untersucht werden, um herauszufinden, wie die Sterblichkeit von Männern und Frauen bisher voranschreitet.

In einem zweiten Schritt mussten wir dann die Sterblichkeit mit dem Lee-Carter-Modell untersuchen, indem wir die Parameter $\hat{\alpha}_x$, $\hat{\beta}_x$ und $\hat{\kappa}_t$ bestimmten. Hinsichtlich der Ergebnisse von $\hat{\alpha}_x$ stellten wir fest, dass die Sterblichkeit in Slowenien etwas niedriger war als in anderen Ländern und dass die Sterblichkeit in Lettland am höchsten war. Im Fall von $\hat{\beta}_x$ waren die Fortschritte bei Frauen in fast allen Ländern größer als bei Männern ab 50 Jahren. Für $\hat{\kappa}_t$ stellten wir fest, dass in einigen Ländern wie Lettland der Verlauf von $\hat{\kappa}_t$ ein wenig instabil war und in anderen wie Polen sehr gut stabilisiert wurde.

Da die Mortalitätsdaten in Ländern mit kleinen Bevölkerungszahlen im Allgemeinen durch Fluktuationen beeinflusst werden, wurden im dritten Schritt die Sterblichkeitsveränderungen $\hat{\kappa}_{[1990:2015]}$ in Estland, Lettland und Slowenien anhand der Sterblichkeitsveränderungen $\hat{\kappa}_{[1990:2015]}$ in Polen, Tschechien, Ungarn und der Slowakei repliziert. Dann wurde dieses Replikationsmodell mit zwei Methoden getestet, um herauszufinden, ob es akzeptabel ist oder nicht: Überprüfung der Gewichtsstabilität und Backtesting. Wir stellten fest, dass einerseits die Gewichte nicht stabil waren und andererseits das Backtesting nicht sehr positiv ausfiel. Daher wurde der Schluss gezogen, dass das Replikationsmodell für $\kappa_t^{[AVE]}$ nicht gut ist.

Kapitel 7

Anhang

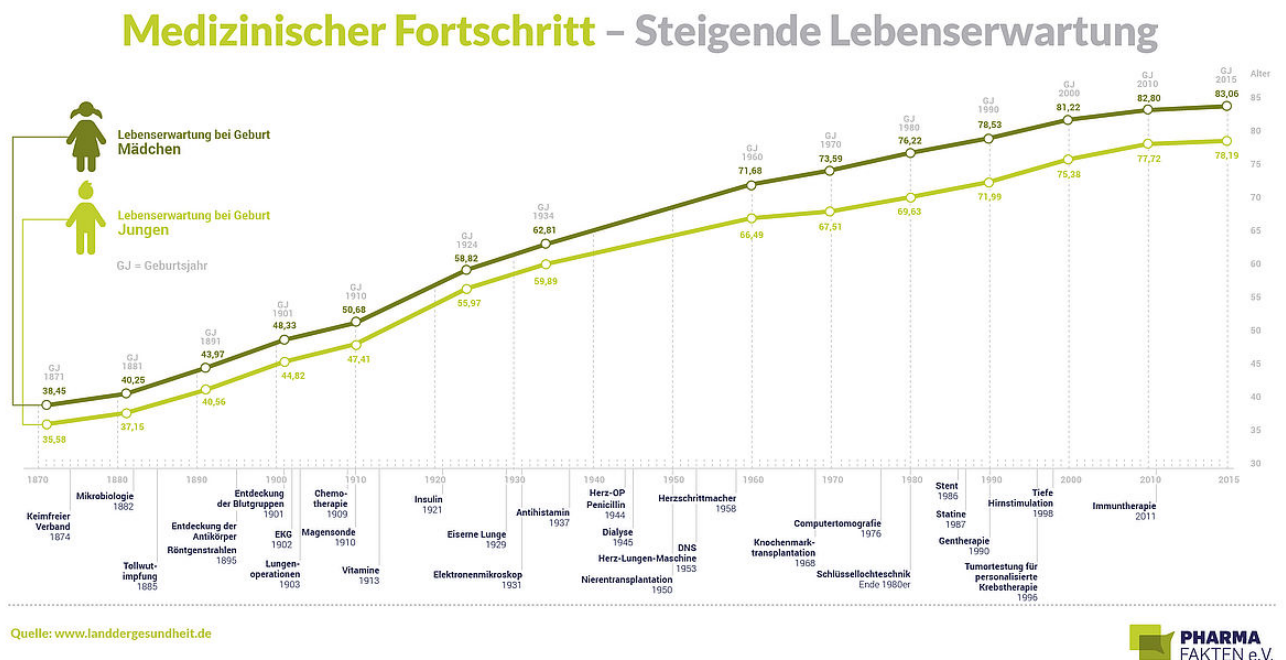


Abbildung 7.1: medizinische Fortschritte 1870 bis 2015. (Quelle: siehe [Pha])

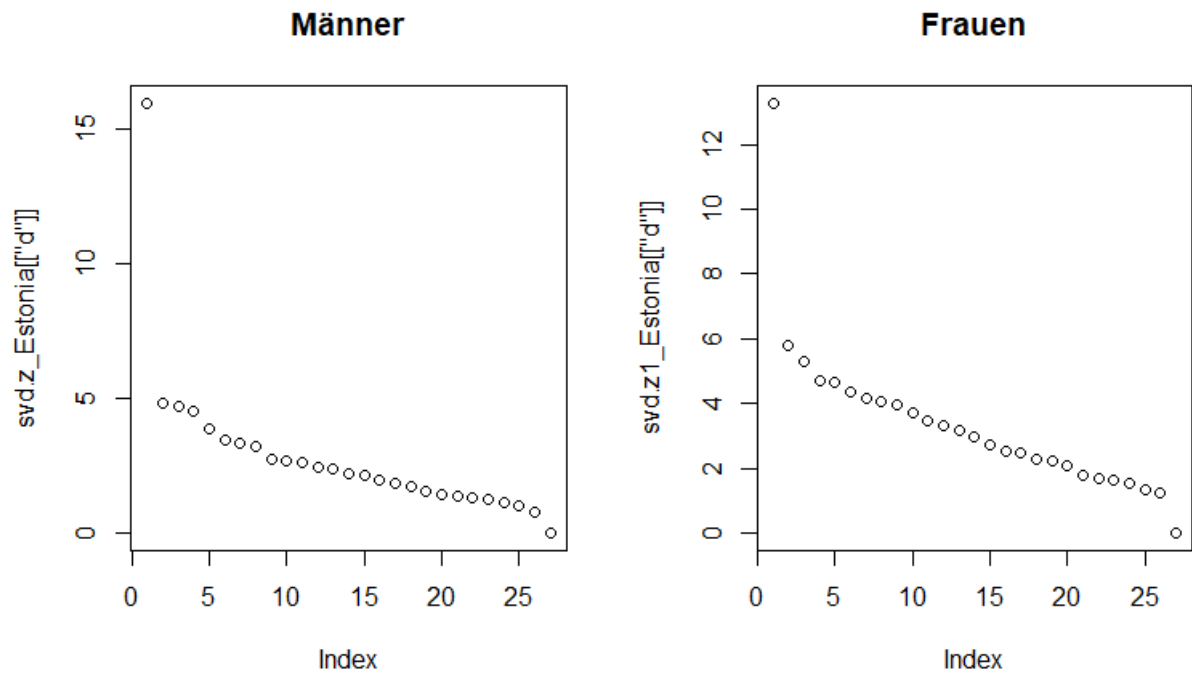


Abbildung 7.2: Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Estland

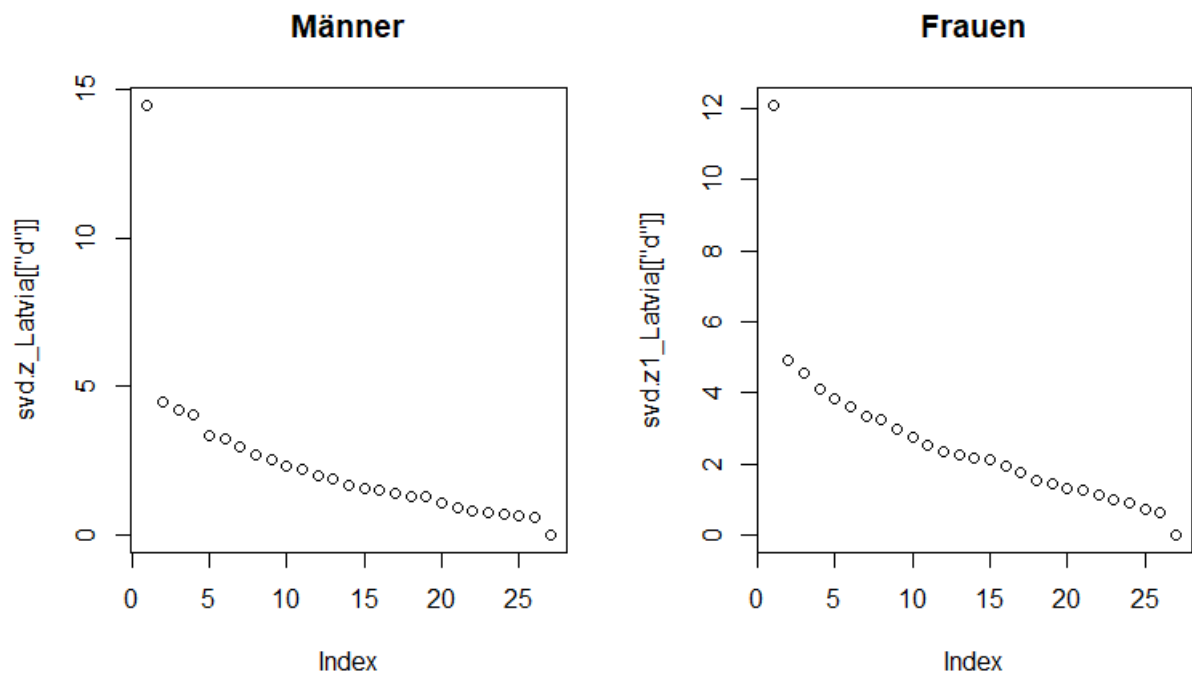


Abbildung 7.3: Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Lettland

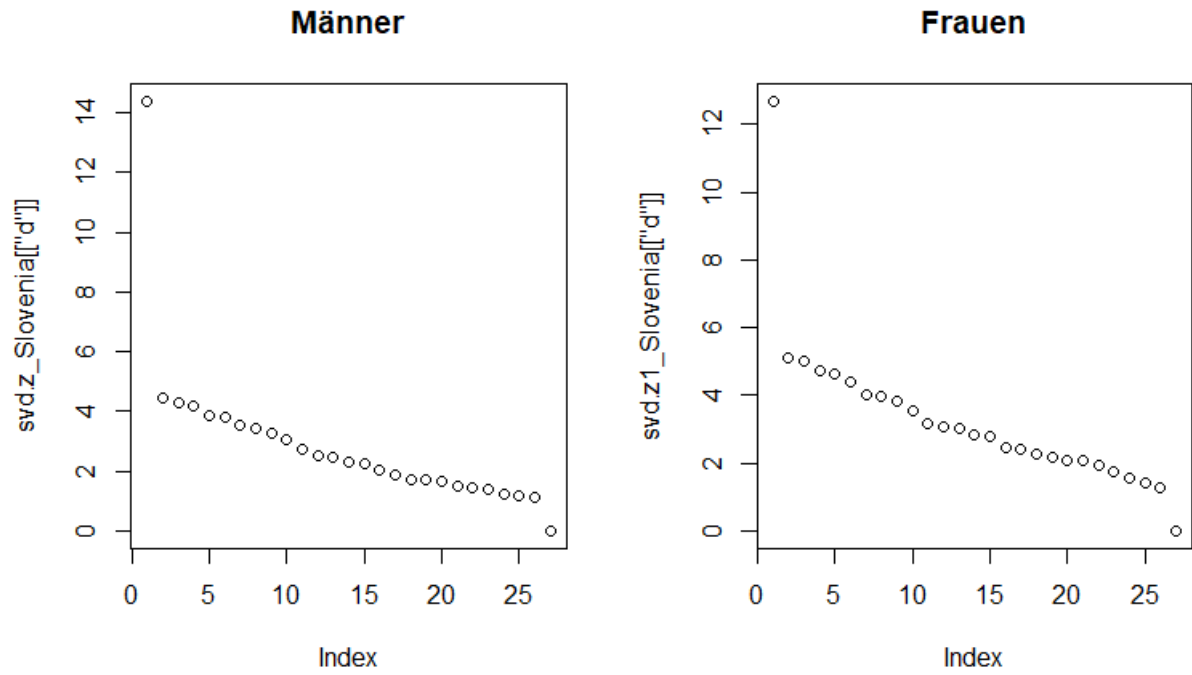


Abbildung 7.4: Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Slowenien

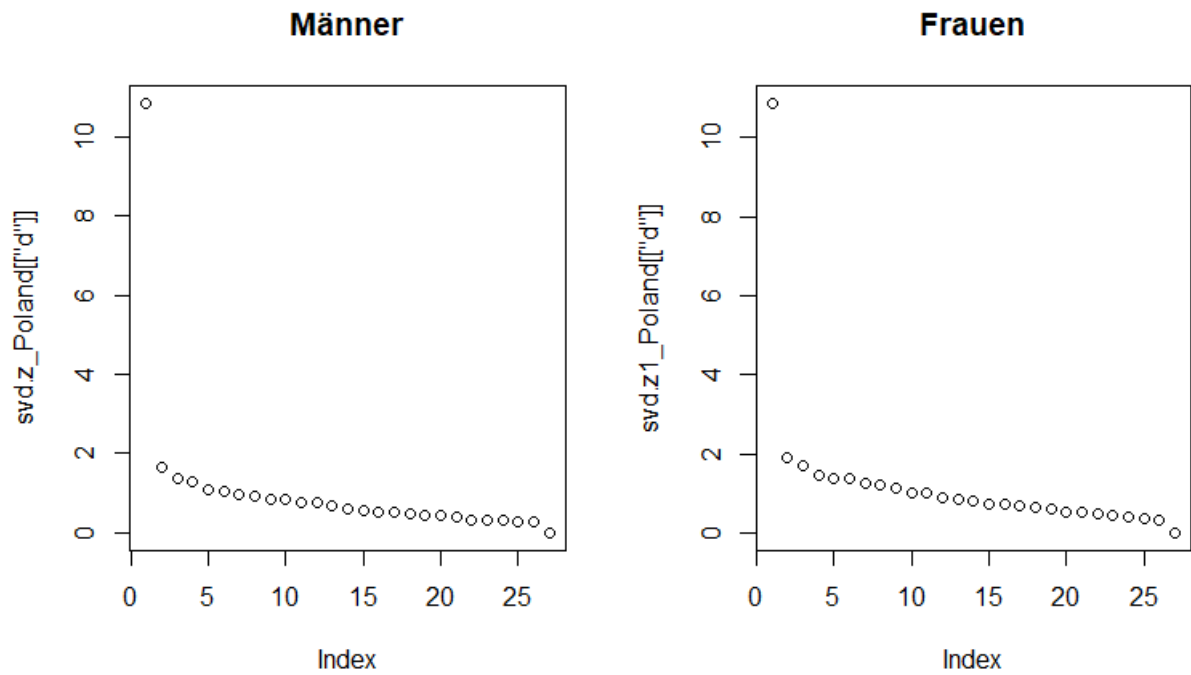


Abbildung 7.5: Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Polen

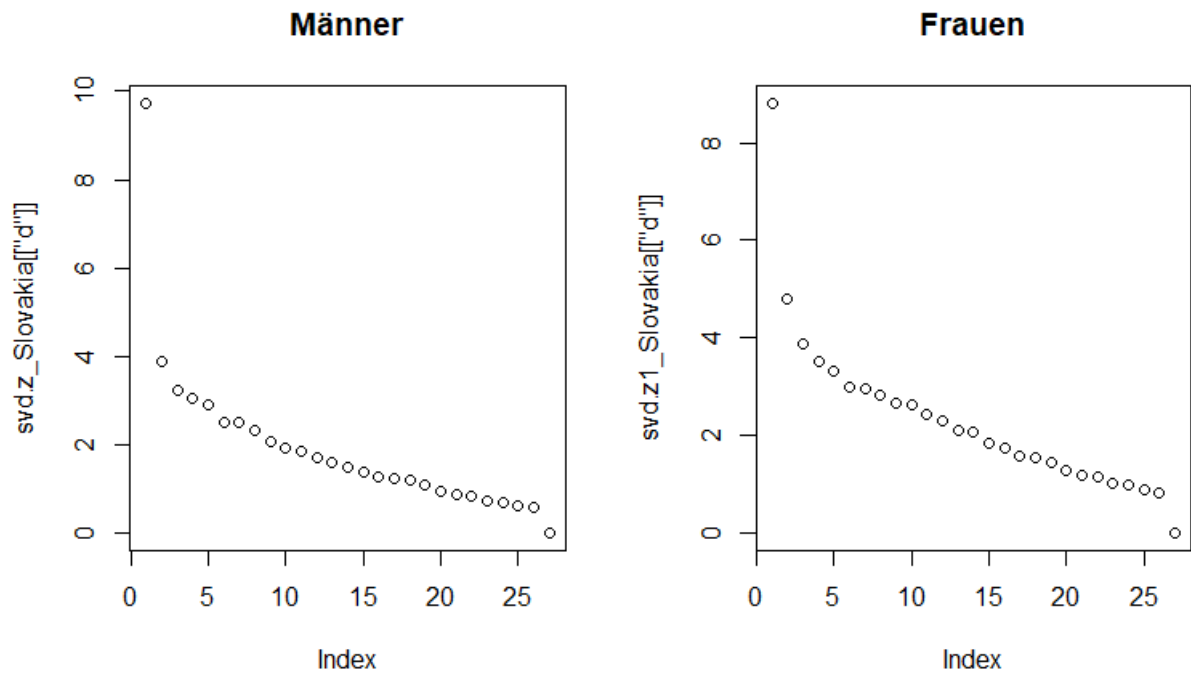


Abbildung 7.6: Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Slowakei

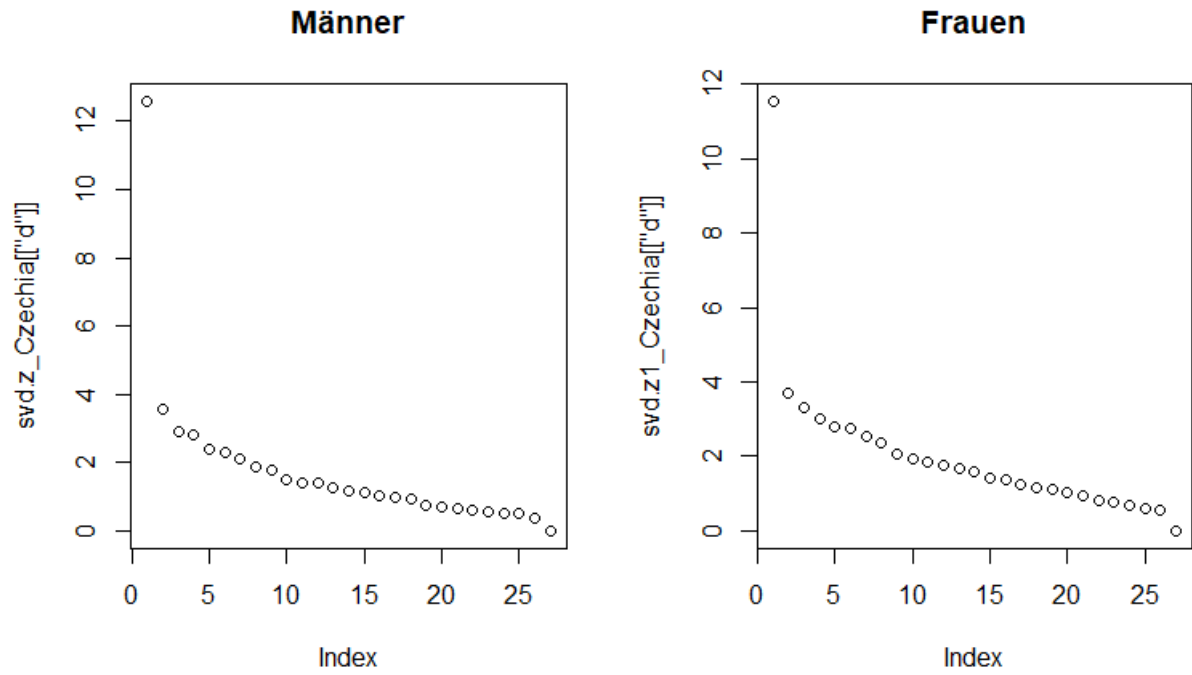


Abbildung 7.7: Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Tschechien

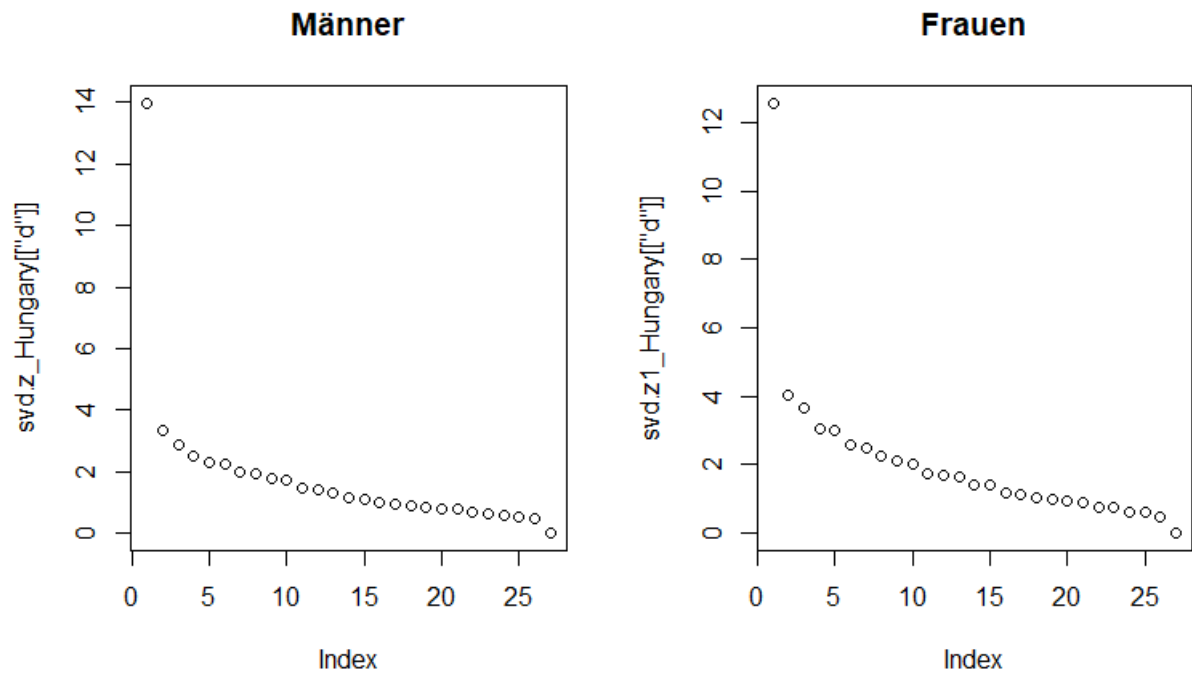


Abbildung 7.8: Darstellung der Singulärwerten s_i der Matrix Z aus dem Lee-Carter Modell in Ungarn

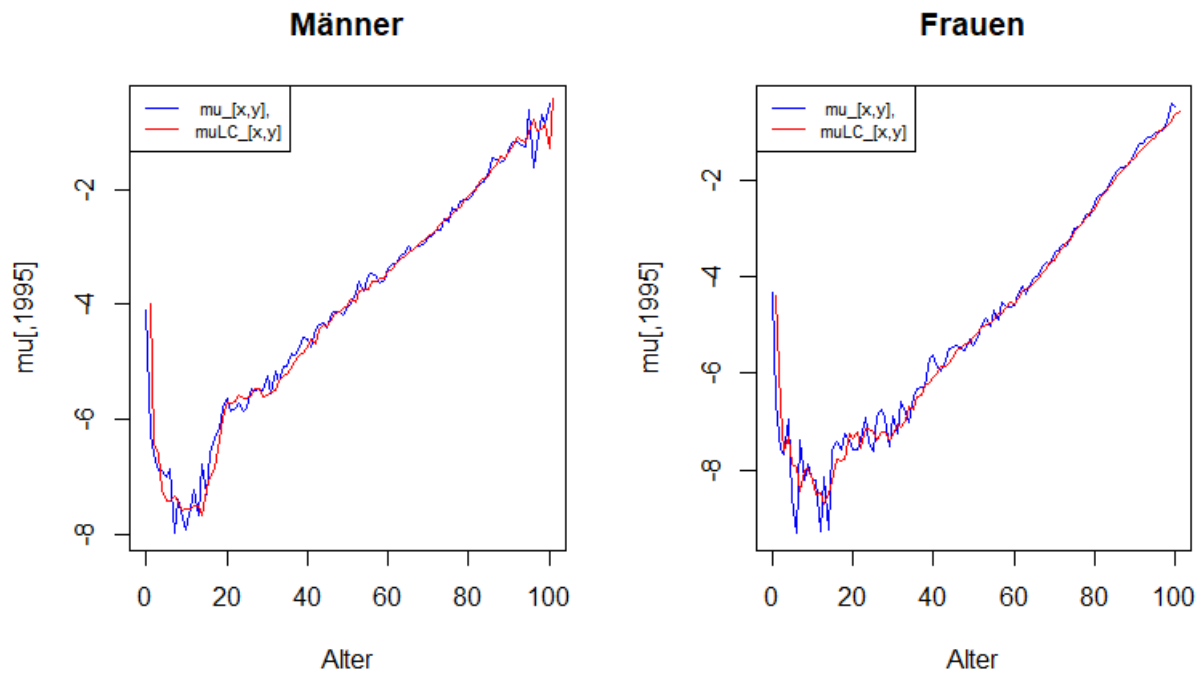


Abbildung 7.9: Vergleich zwischen der logarithmierten Sterberate $\mu_{x,t}$ und der logarithmierten Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Estland

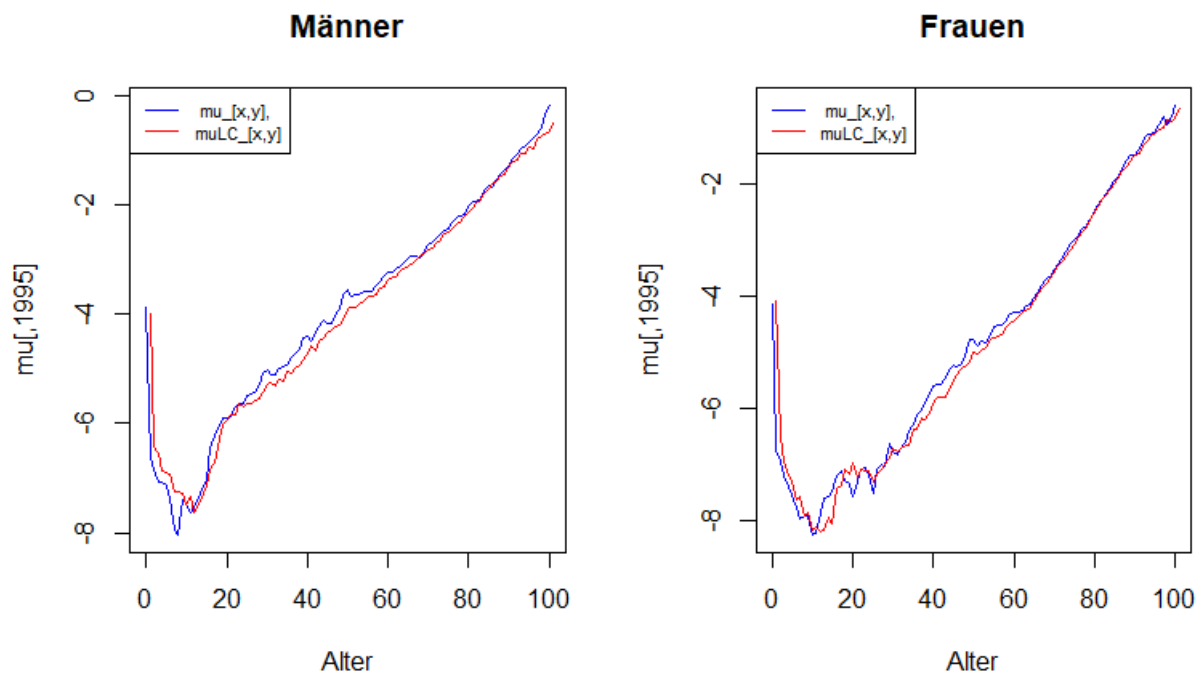


Abbildung 7.10: Vergleich zwischen der logarithmierten Sterberate $\mu_{x,t}$ und der logarithmierten Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Lettland

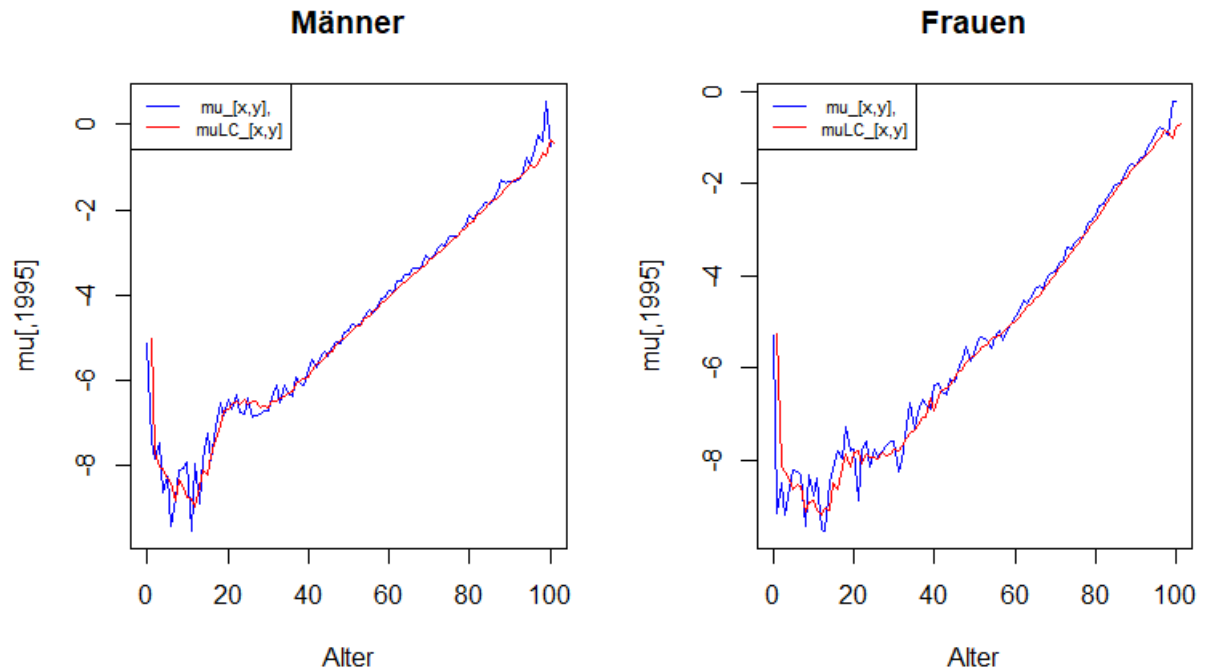


Abbildung 7.11: Vergleich zwischen der logarithmierten Sterberate $\mu_{x,t}$ und der logarithmierten Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Slowenien

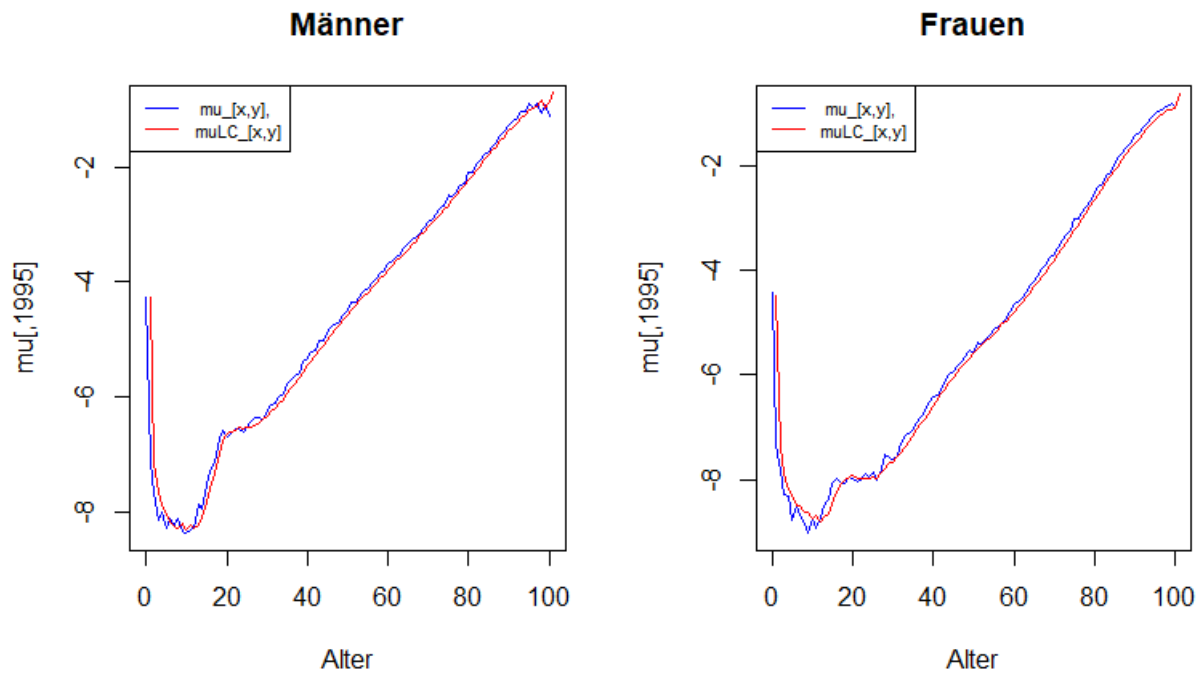


Abbildung 7.12: Vergleich zwischen der logarithmierten Sterberate $\mu_{x,t}$ und der logarithmierten Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Polen

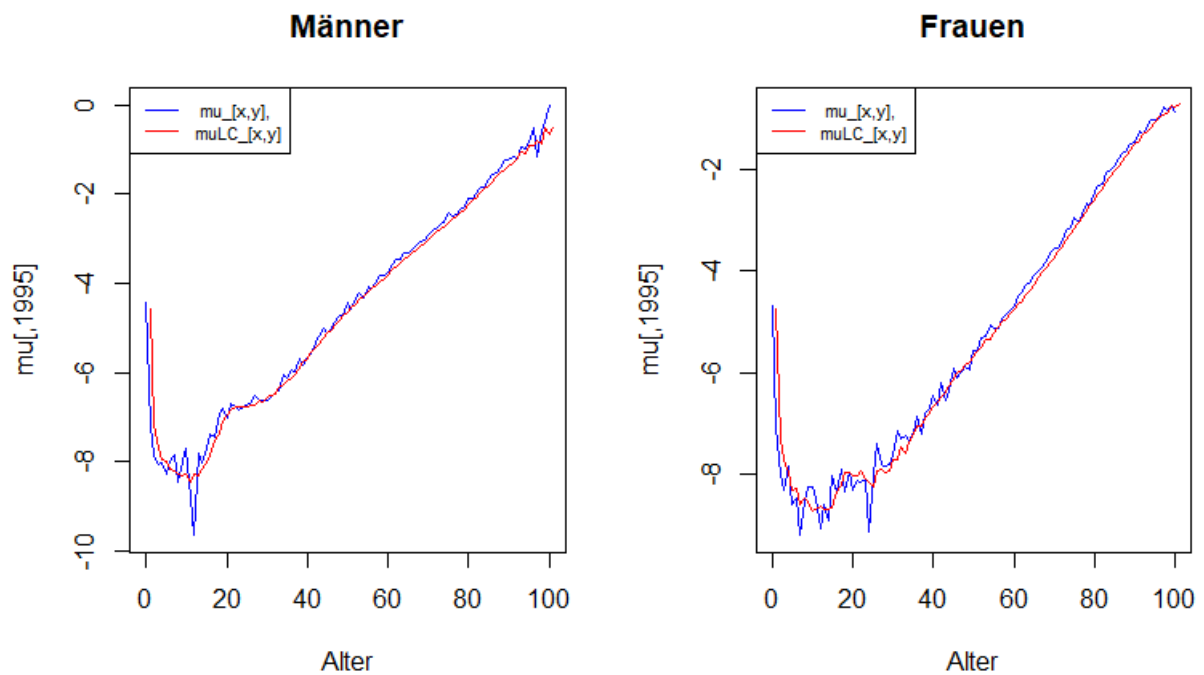


Abbildung 7.13: Vergleich zwischen der logarithmierten Sterberate $\mu_{x,t}$ und der logarithmierten Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Slowakei

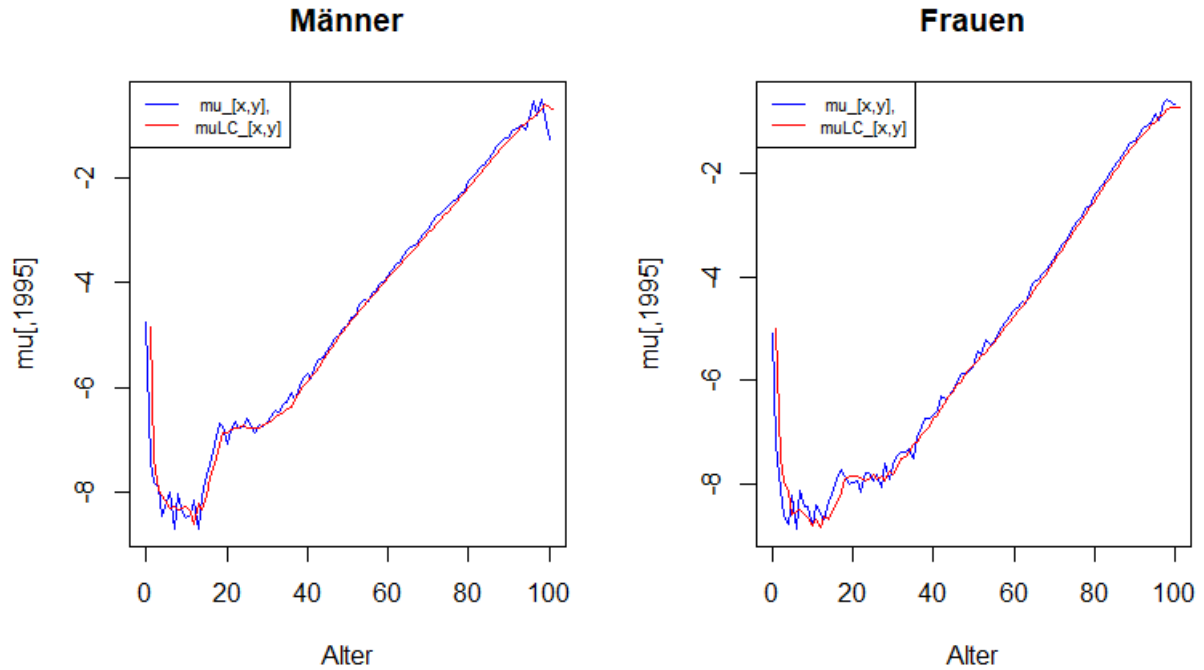


Abbildung 7.14: Vergleich zwischen der logarithmierten Sterberate $\mu_{x,t}$ und der logarithmierten Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Tschechien

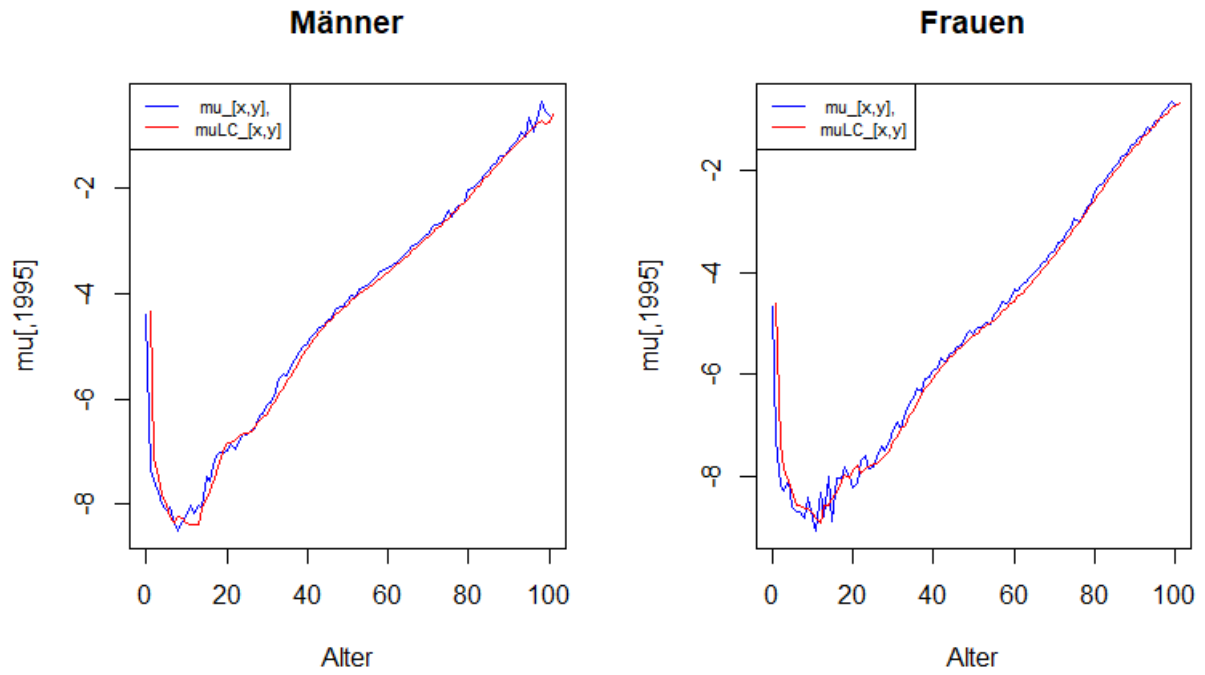


Abbildung 7.15: Vergleich zwischen der logarithmierten Sterberate $\mu_{x,t}$ und der logarithmierten Sterberate mit dem Lee-Carter Modell $\hat{\mu}_{x,t}$ (oder $\mu_{x,t}^{LC}$) in 1995 in Ungarn

Literaturverzeichnis

- [AMOP14] AHCAN, Ales ; MEDVED, Darko ; OLIVIERI, Annamaria ; PITACCO, Eman-
no: *Forecasting mortality for small populations by mixing mortality da-
ta*. [http://www.actuaries.org/CTTEES_TFM/Documents/Hague_Item20_
Projecting_Mortality_Small_Populations.pdf](http://www.actuaries.org/CTTEES_TFM/Documents/Hague_Item20_Projecting_Mortality_Small_Populations.pdf). Version: 01 2014. –
Insurance: Mathematics and Economics 54: 12-27
- [Cou20] *Population of the world and countries*. <http://countrymeters.info/en>.
Version: 2020. – Countrymeters.info
- [Eck10] ECKSTEIN, Peter P.: *Statistik für Wirtschaftswissenschaftler: Eine real-
datenbasierte Einführung mit SPSS*. Wiesbaden : Gabler, 2010. – 341–
380 S. http://dx.doi.org/10.1007/978-3-8349-8933-8_10. http://dx.doi.org/10.1007/978-3-8349-8933-8_10. – ISBN 978–3–8349–
8933–8
- [FKL09] FAHRMEIR, Ludwig ; KNEIB, Thomas ; LANG, Stefan: *Regression: Modelle,
Methoden und Anwendungen*. Springer, 2009. – 90–93 S.
- [Fur95] FURLAN, Peter: *Das gelbe Rechenbuch. Für Ingenieure, Naturwissen-
schaftler und Mathematiker*. Verlag Martina Furlan, 1995. [http://www.
das-gelbe-rechenbuch.de/download/Swz.pdf](http://www.das-gelbe-rechenbuch.de/download/Swz.pdf)
- [GBRF12] GIACOMETTI, Rosella ; BERTOCCHI, Marida ; RACHEV, Svetlozar ; FA-
BOZZI, Frank: *A comparison of the Lee-Carter model and AR-ARCH mo-
del for forecasting mortality rates*. [https://statistik.econ.kit.edu/
download/alpine.DEB.pdf](https://statistik.econ.kit.edu/download/alpine.DEB.pdf). Version: 01 2012. – Insurance: Mathematics
and Economics: 85-93
- [GK07] GIROSI, Federico ; KING, Gary: *Understanding the Lee-Carter Mortality
Forecasting Method*. [https://gking.harvard.edu/files/gking/files/
lc.pdf](https://gking.harvard.edu/files/gking/files/lc.pdf). Version: 09 2007

- [Hea17] *Heat Map*. http://cdn2.hubspot.net/hubfs/402067/PDFs/Heat_Map.pdf. Version: 2017. – Statgraphics Centurion 18
- [Koh] *Kohortenanalyse*. <https://de.ryte.com/wiki/Kohortenanalyse#Kohorte>. – Ryte Wiki
- [Lan] *Meilensteine der Medizin*. <https://www.landdergesundheits.de/fortschritt/meilensteine-medicin>. – Landdergesundheits
- [MMS19] MENZIETTI, Massimiliano ; MORABITO, Maria F. ; STRANGES, Manuela: *Mortality Projections for Small Populations: An Application to the Maltese Elderly*. <https://www.mdpi.com/2227-9091/7/2/35#>. Version: 03 2019. – Risks, MDPI, Open Access Journal 7(2): 1-25
- [Mor] *Human Mortality Database*. www.mortality.org. – University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany)
- [Pha] *Medizinischer Fortschritt - Steigende Lebenserwartung*. <https://www.pharma-fakten.de/grafiken/detail/701-medizinischer-fortschritt-steigende-lebenserwartung/#&gid=lightbox-group-803&pid=0>. – Pharma-Fakten
- [WBB⁺19] WARNES, Gregory R. ; BOLKER, Ben ; BONEBAKKER, Lodewijk ; GENTLEMAN, Robert ; LIAW, Wolfgang Huber A. ; LUMLEY, Thomas ; MAECHLER, Martin ; MAGNUSSON, Arni ; MOELLER, Steffen ; SCHWARTZ, Marc ; VENABLES, Bill: *gplots: Various R Programming Tools for Plotting Data*, 2019. <https://CRAN.R-project.org/package=gplots>. – R package version 3.0.1.1

Danksagung

Zum Schluss möchte ich mich bei all denjenigen bedanken, die mich während der Anfertigung dieser Bachelorarbeit unterstützt und motiviert haben.

Zuerst möchte ich mich bei Herrn Prof. Dr. Claus Neidhardt, der meine Bachelorarbeit betreut und begutachtet hat. Für die hilfreiche Anregungen und die konsequente Kritik bei der Erstellung dieser Arbeit möchte ich mich herzlich bedanken.

Außerdem möchte ich Herrn René Schneidmann für das Korrekturlesen meiner Bachelorarbeit danken.

Abschließend möchte ich mich bei meinen Eltern bedanken, die mir mein Studium durch ihre Unterstützung ermöglicht haben.

Remagen, 19. März 2020

Takam Tchedjou

Ehrenwörtliche Erklärung

Hiermit versichere ich, die vorliegende Bachelorarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben. Alle Stellen, die aus den Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Remagen, 19. März 2020

Takam Tchedjou