

**Stochastische  
Bevölkerungsvorausberechnung  
für Deutschland und ihre Bedeutung  
für ein zukünftiges Rentenmodell**

ABSCHLUSSARBEIT

zur Erlangung des akademischen Grades

Master of Science (M.Sc.)

im Masterstudiengang Statistik

an der Wirtschaftswissenschaftlichen Fakultät

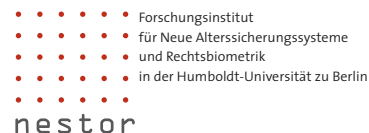
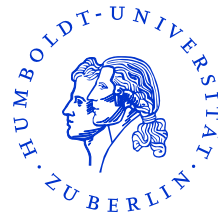
Humboldt-Universität zu Berlin

von

Alena Myšičková

geboren am 19.03.1981 in Čáslav

Matrikelnr.: 184568



Gutachter:

Prof. Dr. Wolfgang Härdle

Prof. Dr. Hans-Peter Schwintowski

Berlin, den 14. November 2007

# **Selbstständigkeitserklärung**

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel und Quellen angefertigt habe.

Die Arbeit hat keiner anderen Prüfungsbehörde vorgelegen.

Berlin, den 14. November 2007

Alena Myšíčková

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>8</b>
<b>2</b>	<b>Mortalität</b>	<b>10</b>
2.1	Datensatz . . . . .	10
2.2	Modellierung der Mortalität . . . . .	13
2.3	Prognose der Mortalität . . . . .	19
<b>3</b>	<b>Fertilität</b>	<b>26</b>
3.1	Datensatz . . . . .	27
3.2	Modellierung der Fertilität . . . . .	28
3.3	Prognose der Fertilität mit dem transformierten Fertilitätsindex . . . . .	32
<b>4</b>	<b>Migration</b>	<b>38</b>
<b>5</b>	<b>Bevölkerungsvorausberechnung</b>	<b>41</b>
5.1	Kohorten-Komponenten-Methode . . . . .	41
5.2	Ergebnisse der Modellrechnung . . . . .	44
5.3	Vergleich mit der Prognose des Statistischen Bundesamtes . . . . .	49
<b>6</b>	<b>Auswirkung auf ein neues Rentensystem</b>	<b>53</b>
6.1	Konzept des Rentenmodells mit einer Basisrente . . . . .	54
6.2	Modellrechnung zur Finanzierung des Rentensystems . . . . .	57
<b>7</b>	<b>Fazit</b>	<b>65</b>

# Tabellenverzeichnis

2.1	Geschätzte Werte (SVD) der Parameter $a_x$ und $b_x$ getrennt für Männer und Frauen. . . . .	18
2.2	Geschätzte Lebenserwartung eines Neugeborenen getrennt für Männer und Frauen zusammen mit 5%- und 95%-Quantilen nach 5000 Simulationen. . . . .	24
3.1	Geschätzte Werte (SVD) der Parameter $a_x$ und $b_x$ . . . . .	31
5.1	Mittelwerte, 2.5%- und 97.5%-Quantile der Bevölkerungsprognose in Mio.	45
6.1	Geschätzte Kosten beiden Rentensystemen in Mrd. EUR. . . . .	61
6.2	Aufsummierte Kosten beiden Rentensystemen in Mrd. EUR. . . . .	62
6.3	Geschätzte Belastung ausgewählter Personen im neuen Rentensystem mit einem Rentenbetrag von 730 EUR und 1000 EUR (mit *) monatlich. . .	63
6.4	Geschätzte Belastung ausgewählter Personen im derzeitigen Rentensystem mit einem Rentenbetrag von 730 EUR. . . . .	63
6.5	Vergleich der Belastung ausgewählter Personen in Prozent des Gesamteinkommens. . . . .	64

# Abbildungsverzeichnis

2.1	Lebenserwartung bei Geburt für Frauen (rote Linie) und Männer (blaue Linie) im früheren Bundesgebiet. . . . .	11
2.2	Altersspezifische Sterbeziffer für Männer und Frauen im Ostdeutschland (rote Linie) und Westdeutschland (dunkelblaue Linie). . . . .	12
2.3	Logarithmierte altersspezifische Sterbeziffer für Männer (oben) und für Frauen (unten). . . . .	15
2.4	Geschätzter Mortalitätsindex $k_t$ für Männer (blaue Linie) und Frauen (rote Linie). . . . .	17
2.5	Tatsächliche (rote Punkte) und geschätzte (blaue Oberfläche) Sterbeziffern für Frauen auf logarithmischer Skala. . . . .	20
2.6	Der Mortalitätsindex mit seiner Prognose von 2005 bis 2074 mit 95%-Konfidenzintervall nach 5000 Simulationen. Blaue Linien stehen für Werte der männlichen Mortalität, rote Linien für Werte der weiblichen Mortalität. . . . .	22
2.7	Einige Simulationen der Lebenserwartung bis zum Jahr 2074 für Männer (blau) und für Frauen (schwarz). . . . .	24
2.8	Histogramm der männlichen (oben) und weiblichen (unten) Lebenserwartung bei Geburt im Jahr 2070. . . . .	25
3.1	Total Fertility Rate pro Frau in Westdeutschland. . . . .	27
3.2	Altersspezifische Geburtenziffern für Ostdeutschland (rot) und Westdeutschland (blau) in den Jahren 1990 und 2004. . . . .	28

3.3	Altersspezifische Geburtenziffern der 15- bis 44-jährigen Mütter von 1950 bis 2005. . . . .	29
3.4	Geschätzter Fertilitätsindex $f_t$ . . . . .	32
3.5	Total Fertility Rate (schwarz) mit der Prognose und ausgewählten Simulationen von 2006 bis 2074. Die rote gestrichelte Linien bezeichnen den Median und den 95%-Konfidenzintervall. . . . .	36
3.6	Histogramme der TFR im Jahr 2030 (oben) und im Jahr 2070 (unten). . . . .	37
4.1	Verlauf des Wanderungssaldo vom Jahr 1990 bis 2004. . . . .	39
4.2	Altersstruktur der weiblichen (rot) und männlichen (blau) Zu- und Fortzüge, Mittelwerte aus dem Zeitraum 1990 bis 2004. . . . .	39
5.1	Mittelwert (durchgezogene Linie) und 95%-Intervall (gestrichelte Linie) für den Bevölkerungsumfang bis zum Jahr 2074 mit Schätzungen des Statistischen Bundesamtes (rote Dreiecke). . . . .	45
5.2	Verteilung der Bevölkerungszahl in den Jahren 2010, 2030, 2050 und 2070 in Mio. . . . .	46
5.3	Bevölkerungspyramiden in den Jahren 2010, 2030, 2050 und 2070 zusammen mit der Bevölkerungspyramide des Jahres 2007 (schwarz). Das rote Feld stellt die weibliche, das blaue Feld die männliche Bevölkerung in 1000 dar. . . . .	47
5.4	Jugendquotient (durchgezogene Linie) mit dem 95%-Konfidenzintervall (gestrichelte Linie). . . . .	49
5.5	Altenquotient mit Altersgrenze 65 (schwarze Linie) und mit Altersgrenze 67 (rote Linie) zusammen mit dazugehörigen 95%-Konfidenzintervallen. . . . .	50
5.6	Geschätzte Bevölkerungszahlen in den Jahren 2030 und 2050 (schwarze Kreuze) mit dem Mittelwert (grüner Punkt) und 95%-Konfidenzintervalle (gestrichelte Linien). Als rotes Dreieck bzw. blaues Viereck sind die Varianten des Statistischen Bundesamtes dargestellt („mittlere“, „relativ junge“ und „relativ alte“ Bevölkerung). . . . .	52

6.1	Der Mindestbeitragssatz in dem neuen Rentenmodell (schwarze durchgezogene Linie) mit dem 95%-Konfidenzintervall (schwarze gestrichelte Linien). . . . .	58
6.2	Der Histogramm für die variable Beitragssatz im Jahr 2070. . . . .	59
6.3	Die geschätzte Kostenentwicklung im derzeitigen (blaue Linie) und im neuen Rentenmodell (schwarze Linie), die sich aus dem Bundeszuschuss (rote Linie) und den Beitragserträgen (schwarze gestrichelte Linie) ergibt. . . . .	60

# 1 Einleitung

Bevölkerungsvorausberechnungen erläutern die Folgen der bereits angelegter Strukturen und erkennbarer Veränderungen in der Bevölkerung auf die zukünftige Population. Sie sind entscheidend für zahlreiche politische, gesellschaftliche und wirtschaftliche Entscheidungsprozesse, wie z.B. die Finanzierung der sozialen Sicherungssysteme, die Arbeitsnachfrage und -angebot aber auch die Finanzierung öffentlicher Infrastrukturen.

In der vorliegenden Arbeit wird eine stochastische Bevölkerungsvorausberechnung durchgeführt, die im Gegensatz zu deterministischen Modellrechnungen neben einer Punktschätzung auch Angaben von Intervallschätzungen und von Wahrscheinlichkeitsverteilungen liefert. Dies gilt nicht nur für die eigene Prognose der Bevölkerungsgröße sondern auch für andere Kenngrößen wie z.B. Jugend- oder Altenquotient, die uns eine Information über die Bevölkerungsstruktur geben. Zuletzt werden in der hier vorliegenden stochastischen Modellen die zeitliche Korrelationen bzw. Autokorrelationen der demografischen Variablen berücksichtigt, siehe Babel (2007).

Die Bevölkerungsvorausberechnung basiert auf der Modellierung der demografischen Variablen Mortalität, Fertilität und Migration, welche in den folgenden drei Kapiteln beschrieben werden. Das Kapitel 2 beginnt mit der Beschreibung des zur Modellierung der Mortalität verwendeten Datensatzes. Im Abschnitt 2.2 wird das Lee-Carter-Modell für die altersspezifischen Sterbeziffern vorgestellt und im letzten Abschnitt des zweiten Kapitels wird die Prognose der Mortalität und der Lebenserwartung aufgezeigt. Das Kapitel 3 widmet sich analog der Analyse der Fertilitätsdaten (Abschnitt 3.1), der Modellbeschreibung für die Fertilität (Abschnitt 3.2) und zuletzt auch die Prognose der



Geburtenhäufigkeit. Kapitel 4 schließt sich mit der Analyse und Festlegung der Annahmen an die Migration an. Kapitel 5 beschreibt zuerst im Abschnitt 5.1 die demografischen Methoden zur Bevölkerungsentwicklung, im Weiteren werden die Ergebnisse der stochastischen Modellrechnung präsentiert. Im letzten Abschnitt 5.3 werden die Ergebnisse der stochastischen Vorausberechnung mit den Ergebnissen der 11. koordinierten Bevölkerungsvorausberechnung des Statistischen Bundesamtes aus dem Jahre 2006 verglichen, siehe Stat.Bundesamt (2006a). Schließlich werden in Kapitel 6 Auswirkungen der demografischen Entwicklung in Deutschland auf ein neues Rentensystem analysiert. Eine Zusammenfassung der erworbenen Ergebnissen bietet zum Schluss das Kapitel 7.

## 2 Mortalität

Aufgrund des medizinischen Fortschritts und der besseren Lebensbedingungen ist die Lebenserwartung in Deutschland im Laufe des 20. Jahrhunderts gestiegen. Den Trend der Lebenserwartung im früheren Bundesgebiet ist in der Abbildung 2.1 abgebildet. So ist beispielsweise die Lebenserwartung eines neugeborenen Mädchen (rote Linie) von 70.9 Jahren im Jahr 1956 auf 81.8 Jahre im Jahr 2004 gestiegen. Bei einem neugeborenen Jungen betrug die Lebenserwartung (blaue Linie) im Jahr 1956 65.9 Jahre, im Jahr 2004 waren es bereits 76.5 Jahre. Die Mortalität ändert sich jedoch nicht nur im Laufe der Zeit, sie unterscheidet sich auch in einzelnen Altersgruppen, was auch auf die Bevölkerungsentwicklung Auswirkungen hat.

Der folgende Abschnitt 2.1 widmet sich der Beschreibung des in der Modellrechnung verwendeten Datensatzes. Im Abschnitt 2.2 wird das Lee-Carter-Modell für die Mortalität dargestellt, der letzte Abschnitt dieses Kapitels zeigt die Punktschätzung zusammen mit Konfidenzbändern der Mortalität und der Lebenserwartung.

### 2.1 Datensatz

Für die Mortalitätsanalyse wurden geschlechts- und altersspezifische Sterbeziffern verwendet. Die geschlechts- und altersspezifische Sterbeziffer wird definiert als “die Zahl der männlichen oder weiblichen Gestorbenen eines Alters bezogen auf 1000 Lebende des entsprechenden Alters und Geschlechts“<sup>1</sup>. Die altersspezifischen Sterbeziffern dienen als

---

<sup>1</sup>Quelle: [lexikon.meyers.de](http://lexikon.meyers.de)

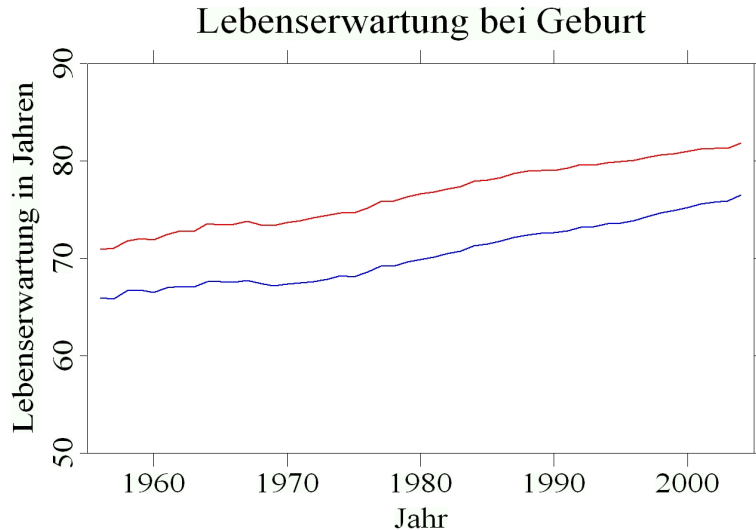


Abbildung 2.1: Lebenserwartung bei Geburt für Frauen (rote Linie) und Männer (blaue Linie) im früheren Bundesgebiet.

das einfachste Maß der Sterblichkeitsintensität und sind auch für den internationalen Vergleich geeignet, da sie von der Altersstruktur der Bevölkerung im jeweiligen Land unabhängig sind, siehe Schmerbach (2001).

Der für die Analyse verwendete Datensatz stammt aus der Online-Datenbank "The Human Mortality Database" ([www.mortality.org](http://www.mortality.org)). Es wurden die geschlechtsspezifischen Sterbeziffern für einjährige Altersgruppen von unter 1 Jahr bis 110 und mehr Jahre im Zeitraum vom Jahr 1956 bis 2004 verwendet. Insbesondere in den ältesten Altersgruppen kommen aufgrund der kleinen Fallzahlen viele fehlende oder inkonsistente Werte (größer als 1) vor. Um diese zu beseitigen werden die Werte ab dem ersten inkonsistenten bzw. fehlenden Wert bis zum Jahr 110 linear interpoliert. Für die letzte Altersgruppe (110-jährige und ältere) wird eine maximale Sterberate angenommen. Damit können die fehlenden Werte in den ältesten Altersgruppen mit einer einheitlichen Methode plausibel fortgesetzt werden, wobei die Ergebnisse der Bevölkerungsvorausberechnung nur unerheblich beeinflusst werden.

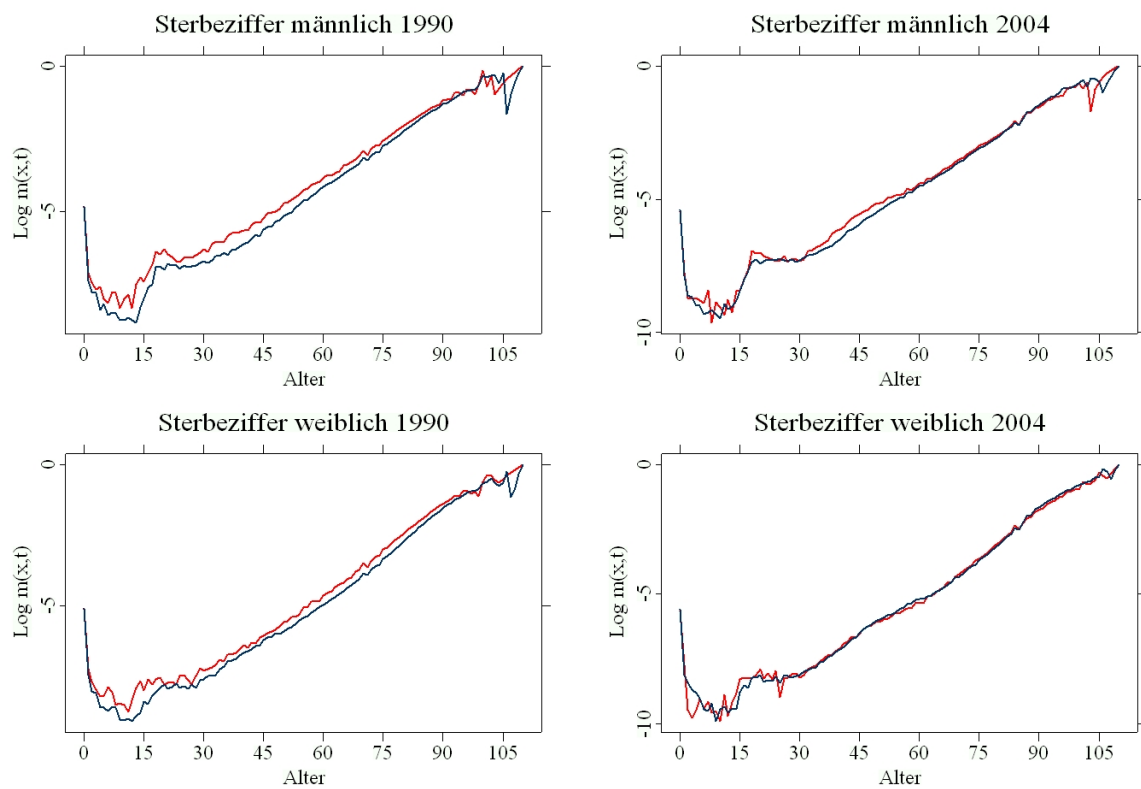


Abbildung 2.2: Altersspezifische Sterbeziffer für Männer und Frauen im Ostdeutschland (rote Linie) und Westdeutschland (dunkelblaue Linie).

Für den Zeitraum 1956 bis 1990 liegen nur getrennte Daten für die alten Bundesländer und die ehemalige DDR vor. Seit der Wiedervereinigung Deutschlands liegen Daten sowohl für das gesamte Bundesgebiet als auch separat für das ehemalige West- und Ostdeutschland vor. Wie man in der Abbildung 2.2 beobachten kann, haben sich die Mortalitätsverhältnisse der Männer und Frauen im Ostdeutschland an die Verhältnisse in Westdeutschland in den letzten 15 Jahren angepasst. Während im Jahr 1990 die Sterbeziffern der ostdeutschen Männer bzw. Frauen deutlich über diejenige der Westdeutschen liegen, ist der Verlauf der Sterbeziffern der ost- und westdeutschen Frauen im Jahr 2004 beinahe identisch. Kleine Unterschiede sind nur bei den jüngeren Frauen zwischen 2 und 20 Jahren zu beobachten. Die Anpassung der Mortalitätsverhältnisse der ostdeutschen an die der westdeutschen Männer erfolgt ebenfalls, wenn auch nicht so schnell wie bei den Frauen. Geringe Unterschiede sind sowohl bei den jungen Männern zwischen 2 und 15 Jahren als auch bei den Männern im produktiven Alter zwischen 30 und 60 Jahren zu erkennen.

Trotz der nur noch geringen Unterschiede zwischen den Sterberaten im Ost- und Westdeutschland ist anzunehmen, dass sich die Sterbeziffern der Menschen in Ostdeutschland noch weiter anpassen werden. Im Weiteren erfolgt daher nur eine Analyse der Sterberaten für das ehemalige Bundesgebiet, die prognostizierte Werte für Westdeutschland werden auch für ostdeutsche Männer und Frauen angenommen. Einen weiterführenden Vergleich der Mortalitätsentwicklung in Ost- und Westdeutschland bietet Lipps und Betz (2005); Babel (2007).

## 2.2 Modellierung der Mortalität

Wie bereits erwähnt, ändern sich die altersspezifischen Sterbeziffern nicht nur im Laufe der Zeit sondern auch in einzelnen Altersgruppen. Eine dreidimensionale Darstellung der männlichen und weiblichen Sterbeziffern zeigt die Abbildung 2.3, wobei an der  $x$ -Achse die Jahreszahl, an der  $y$ -Achse die einzelne Altersgruppen und an der  $z$ -Achse

die dazugehörige Sterbeziffer abgebildet werden. Die Farben der Punkte werden mit steigender Sterblichkeit dunkler. In beiden Grafiken kann man in der ersten Linie eine kontinuierlich abnehmende Säuglingssterblichkeit erkennen. Deutlich ist auch die höhere Sterblichkeit der jungen Männer bis zum 20. Lebensjahr im Vergleich zu den Frauen, die beinahe ausschließlich Verkehrsunfällen zuzuschreiben ist, siehe Dinkel und Luy (1999). Ein rapider Anstieg der Sterblichkeit ist bei Frauen ab dem 70. Lebensjahr, bei Männern ab dem 60. Lebensjahr zu beobachten.

Um die Mortalitätseinflüsse in die altersspezifischen und zeitabhängigen Komponenten zu trennen und anschließend die Prognose in der Zukunft modellieren zu können wird auf die Datenmatrix das Lee-Carter Modell beschrieben von Lee und Carter (1992) angewendet. Für die altersspezifische Sterbeziffer  $\mathbf{m}_{x,t}$  der Altersgruppe  $x$  im Jahr  $t$  gilt dann:

$$\log(\mathbf{m}_{x,t}) = \mathbf{a}_x + \mathbf{b}_x \mathbf{k}_t + \varepsilon_{x,t} , \quad (2.1)$$

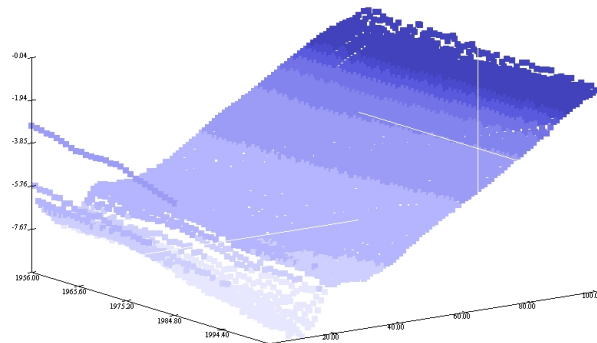
mit den altersspezifischen Konstanten  $\mathbf{a}_x$  und  $\mathbf{b}_x$ , dem zeitabhängigen Mortalitätsindex  $\mathbf{k}_t$  und dem Störterm  $\varepsilon_{x,t}$ . Die Funktion  $\exp(\mathbf{a}_x)$  stellt die generelle Form der Mortalitätsfunktion dar. Aus der Ableitung  $\frac{d}{dt} \log(\mathbf{m}_{x,t}) = \mathbf{b}_x \frac{dk}{dt}$  folgt die Interpretation der altersspezifischen Konstanten  $\mathbf{b}_x$ ; sie beschreiben, welche Sterberaten schneller bzw. langsamer hinsichtlich der Veränderungen in  $\mathbf{k}_t$  sinken. Der Störterm  $\varepsilon_{x,t}$  mit Erwartungswert 0 und Varianz  $\sigma_\varepsilon^2$  spiegelt die zufälligen, nicht durch das Modell abgebildeten Einflüsse. Das logarithmische Modell wird verwendet um negative Sterbeziffern in der Vorhersage zu verhindern.

Um die Parameter des Modells für die vorhandenen Daten der deutschen Sterbeziffer zu schätzen, sucht man nach einer Lösung der Gleichung (2.1) mit Hilfe der Methode der kleinsten Quadrate. Jedoch ist das Modell unterbestimmt, daher werden Einschränkungen an die Komponenten festgelegt:

$$\sum_{x=0}^{\omega} \mathbf{b}_x = 1 \quad \sum_{t=1}^T \mathbf{k}_t = 0 , \quad (2.2)$$

wobei  $\omega$  die längste Lebensdauer bezeichnet (für die verwendeten Daten gilt:  $\omega = 110$ ).

## Sterbeziffern männlich



## Sterbeziffern weiblich

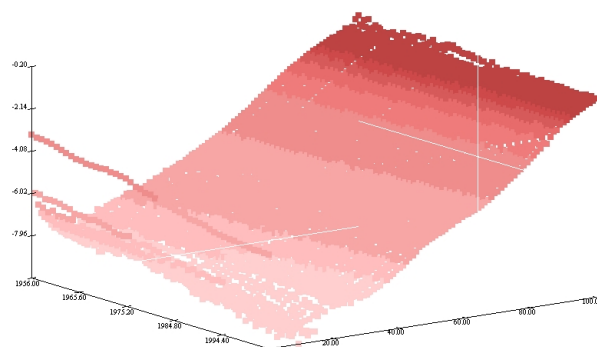


Abbildung 2.3: Logarithmierte altersspezifische Sterbeziffer für Männer (oben) und für Frauen (unten).

Für das Ausgangsjahr des hier benutzten Datensatzes 1956 ist  $t = 1$ , für den Zeithorizont  $T$  gilt dann mit dem Endjahr des Datensatzes 2004:  $T = 2004 - 1956 + 1 = 49$ .

Aus den Gleichungen (2.1) und (2.2) folgt für  $\mathbf{a}_x$  folgendes Verhältnis:

$$\mathbf{a}_x = \frac{1}{T} \sum_{t=1}^T \log(\mathbf{m}_{x,t}) . \quad (2.3)$$

Auf das Modell (2.1) können jedoch keine Standardregressionsverfahren angewendet werden, weil an der rechten Seite keine Regressoren, sondern nur zu schätzende Parameter zusammen mit dem unbekanntem Index  $\mathbf{k}_t$  stehen. Deshalb wird für die Schätzung des Modells die ‘‘Singular Value Decomposition‘‘ (SVD, in deutscher Literatur auch Singulärwertzerlegung) verwendet. Es wird die Matrix  $\mathbf{M} = \log(\mathbf{m}_{x,t}) - \mathbf{a}_x$  zerlegt, so dass gilt:

$$\mathbf{M} = [\log(\mathbf{m}_{x,t}) - \mathbf{a}_x] = \mathbf{\Gamma} \mathbf{\Lambda} \mathbf{\Delta}^\top , \quad (2.4)$$

wobei  $\mathbf{\Gamma}(\omega \times r)$  und  $\mathbf{\Delta}(T \times r)$  orthonormale Matrizen mit  $r = \text{rank}(\mathbf{M})$  und  $\mathbf{\Lambda} = \text{diag}(\lambda_1^{1/2}, \dots, \lambda_r^{1/2})$  mit den Eigenwerten  $\lambda_i$ ,  $i = 1, \dots, r$  von  $\mathbf{M}^\top \mathbf{M}$  repräsentieren, siehe Härdle und Simar (2007). Der erste Spaltenvektor der Matrix  $\mathbf{\Gamma}$  nach Normierung (2.2) entspricht dem Vektor der altersspezifischen Parameter  $\mathbf{b}_x$ . Der erste Spaltenvektor der Matrix  $\mathbf{\Delta}$  multipliziert mit dem größten Eigenwert  $\lambda_1$  liefert die Schätzwerte des zeitabhängigen Index  $\mathbf{k}_t$ .

Die geschätzten Werte der altersspezifischen Konstanten  $\mathbf{a}_x$  und  $\mathbf{b}_x$  zeigt die Tabelle 2.1 getrennt für Männer und Frauen. Das generelle Niveau der Mortalität repräsentiert mit  $\mathbf{a}_x$  ist bei Frauen in allen Altersgruppen (außer jener der 106-jährigen) niedriger als bei Männern. Die Werte  $\mathbf{a}_x$  der Neugeborenen sind deutlich höher als bei älteren Kindern, sie entsprechen etwa den Werten der 58-jährigen Männern bzw. der 63-jährigen Frauen, was Ausdruck der erhöhten Säuglingssterblichkeit ist. Wie bereits erwähnt, beschreiben die Konstanten  $\mathbf{b}_x$  die Veränderungen hinsichtlich der Änderungen in  $\mathbf{k}_t$ . Die größten Werte sowohl bei Männern als auch bei Frauen sind in den jüngsten Altersgruppen zu finden, was dem Rückgang der Säuglings- und Kindersterblichkeit in der 2. Hälfte des 20. Jahrhunderts zu zuschreiben ist. Noch zu erwähnen sind die etwas höheren Werte der



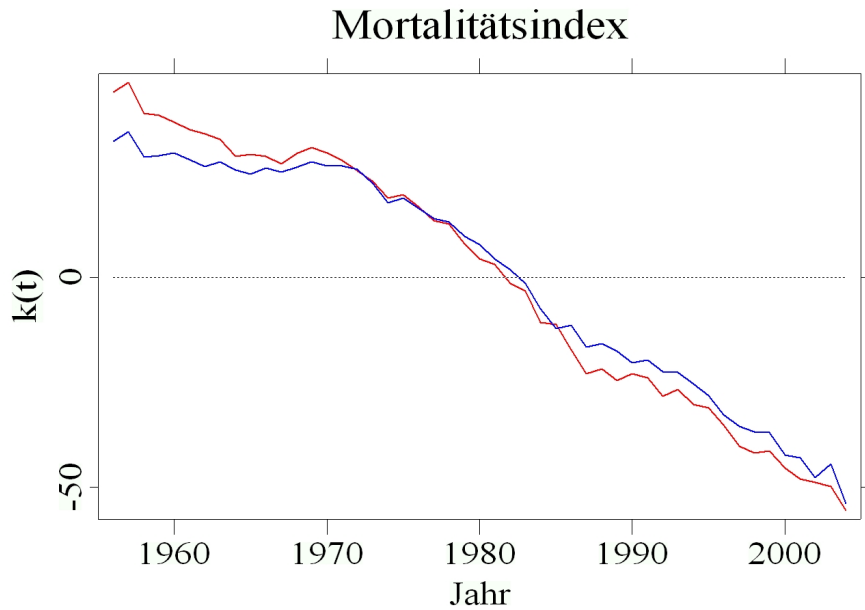


Abbildung 2.4: Geschätzter Mortalitätsindex  $\mathbf{k}_t$  für Männer (blaue Linie) und Frauen (rote Linie).

Konstante  $\mathbf{b}_x$  der 23- bis 40-jährigen Frauen, die den Rückgang der Müttersterblichkeit und der mit Schwangerschaft in Verbindung stehenden Todesursachen abbilden, siehe Dinkel und Luy (1999).

Der Verlauf der Indizes  $\mathbf{k}_t$  der männlichen (blaue Linie) und weiblichen (rote Linie) Sterblichkeit in der Zeitspanne von 1956 bis 2004 zeigt die Abbildung 2.4. Beide Indizes sinken annähernd linear, wobei der weibliche Mortalitätsindex schneller als der männliche abfällt. Von Vorteil für die weitere Vorhersage ist auch die kleine Varianz der Schwankungen.

Die Güte der Anpassung zeigen die dreidimensionalen Abbildungen 2.5. Die roten Punkte stellen hierbei die tatsächlichen logarithmierten Sterbeziffern dar, die blaue Oberfläche zeigt die durch das Lee-Carter Modell geschätzten Werte. Auf der  $x$ -Achse sind die Beobachtungsjahre abgebildet, die  $y$ -Achse zeigt die einzelne Altersgruppen.

Alter	Männer		Frauen		Alter	Männer		Frauen	
	$a_x$	$b_x \cdot 10^3$	$a_x$	$b_x \cdot 10^3$		$a_x$	$b_x \cdot 10^3$	$a_x$	$b_x \cdot 10^3$
0	-4.27	27.24	-4.52	23.58	56	-4.45	7.64	-5.16	7.13
1	-6.90	22.56	-7.05	19.45	57	-4.35	7.84	-5.08	7.29
2	-7.37	22.35	-7.60	19.06	58	-4.26	8.04	-4.98	7.44
3	-7.58	22.06	-7.83	18.58	59	-4.16	8.03	-4.90	7.69
4	-7.78	24.27	-8.04	19.35	60	-4.06	8.26	-4.80	7.94
5	-7.87	24.33	-8.18	19.42	61	-3.96	8.32	-4.71	8.13
6	-7.95	24.13	-8.27	18.40	62	-3.86	8.39	-4.61	8.40
7	-8.04	24.34	-8.37	17.85	63	-3.77	8.42	-4.51	8.54
8	-8.07	21.56	-8.50	16.30	64	-3.68	8.40	-4.40	8.90
9	-8.20	22.48	-8.58	16.66	65	-3.58	8.53	-4.30	9.06
10	-8.27	21.91	-8.67	14.91	66	-3.48	8.57	-4.19	9.21
11	-8.22	17.34	-8.63	13.53	67	-3.39	8.30	-4.08	9.29
12	-8.18	17.57	-8.62	13.68	68	-3.30	8.37	-3.97	9.36
13	-8.11	17.21	-8.57	13.04	69	-3.21	8.27	-3.86	9.47
14	-7.92	16.83	-8.38	13.49	70	-3.11	8.21	-3.74	9.75
15	-7.65	15.58	-8.13	11.96	71	-3.02	8.04	-3.63	9.77
16	-7.19	15.16	-7.93	11.65	72	-2.92	7.91	-3.51	9.74
17	-7.00	14.14	-7.85	11.00	73	-2.83	7.85	-3.39	9.96
18	-6.65	10.86	-7.66	7.99	74	-2.74	7.72	-3.27	9.98
19	-6.60	11.64	-7.66	8.14	75	-2.65	7.53	-3.14	9.97
20	-6.61	11.75	-7.67	8.38	76	-2.55	7.64	-3.03	9.85
21	-6.61	11.73	-7.72	9.92	77	-2.46	7.32	-2.90	9.85
22	-6.65	11.24	-7.71	10.13	78	-2.37	7.23	-2.79	9.56
23	-6.68	10.70	-7.70	10.80	79	-2.28	6.92	-2.67	9.37
24	-6.71	10.54	-7.66	10.47	80	-2.18	6.31	-2.54	8.70
25	-6.73	10.30	-7.64	11.23	81	-2.08	6.03	-2.42	8.24
26	-6.74	10.58	-7.61	11.38	82	-1.99	5.77	-2.31	8.02
27	-6.73	10.17	-7.55	11.52	83	-1.90	5.41	-2.19	7.62
28	-6.71	10.01	-7.50	11.51	84	-1.81	5.20	-2.08	7.21
29	-6.69	10.15	-7.44	11.86	85	-1.72	5.03	-1.96	6.96
30	-6.66	9.97	-7.38	11.65	86	-1.62	4.45	-1.86	6.33
31	-6.61	9.59	-7.31	11.51	87	-1.53	4.16	-1.75	5.87
32	-6.57	9.62	-7.23	11.45	88	-1.45	4.03	-1.65	5.57
33	-6.52	9.43	-7.16	10.91	89	-1.37	3.65	-1.55	5.13
34	-6.46	9.03	-7.08	10.24	90	-1.28	3.44	-1.46	4.76
35	-6.40	8.86	-6.98	10.05	91	-1.19	3.28	-1.36	4.42
36	-6.33	8.64	-6.90	9.64	92	-1.12	3.25	-1.28	4.17
37	-6.24	8.57	-6.80	9.21	93	-1.04	3.01	-1.20	3.85
38	-6.16	8.14	-6.71	9.19	94	-0.97	2.47	-1.11	3.71
39	-6.06	8.18	-6.63	8.87	95	-0.90	2.63	-1.03	3.48
40	-5.98	7.30	-6.54	8.42	96	-0.84	2.01	-0.96	3.32
41	-5.90	7.08	-6.44	8.32	97	-0.80	1.95	-0.89	3.10
42	-5.80	6.72	-6.35	7.96	98	-0.77	1.19	-0.83	2.75
43	-5.71	6.55	-6.25	7.49	99	-0.53	5.78	-0.67	4.67
44	-5.62	6.60	-6.17	7.38	100	-0.43	2.19	-0.59	4.37
45	-5.52	6.23	-6.07	7.20	101	-0.52	1.13	-0.62	1.61
46	-5.42	6.24	-5.98	7.09	102	-0.47	2.17	-0.54	3.09
47	-5.33	6.17	-5.90	7.11	103	-0.49	2.14	-0.55	1.94
48	-5.23	6.34	-5.81	6.99	104	-0.51	0.63	-0.54	2.62
49	-5.14	6.30	-5.73	7.04	105	-0.42	1.96	-0.51	2.84
50	-5.03	6.28	-5.64	6.91	106	-0.46	7.98	-0.39	3.95
51	-4.94	6.87	-5.56	6.95	107	-0.35	7.53	-0.41	7.58
52	-4.84	6.70	-5.48	7.06	108	-0.21	4.31	-0.28	6.04
53	-4.75	6.84	-5.40	7.03	109	-0.09	1.72	-0.20	5.20
54	-4.65	7.10	-5.32	7.02	110	0.00	0.00	-0.04	2.10
55	-4.55	7.38	-5.25	7.17					

Tabelle 2.1: Geschätzte Werte (SVD) der Parameter  $a_x$  und  $b_x$  getrennt für Männer und Frauen.

Die aufgeklärte Varianz (explained variance) der einzelnen Altersgruppen bei Frauen liegt überwiegend über 90%, abgesehen von einigen Altersgruppen zwischen 13 und 21 Jahren und den ältesten Altersgruppen ab dem 93. Lebensjahr. Bei Männern liegt die aufgeklärte Varianz der 16- bis 18-jährigen, der 40-jährigen und der 81-jährigen und älteren unter der Neunzigprozentgrenze. Als ein gemeinsames Maß kann man die so genannte aufgeklärte Gesamtvarianz (total explained variance) verwenden, die man durch das Verhältnis der Summe der ungeklärten Varianzen aller Altersgruppen und der Summe der Gesamtvarianzen aller Altersgruppen abgezogen von 1 bekommt. Das Modell für Frauen erklärt somit 87.6% der Gesamtvarianz, das Modell für Männer 87.8% der Gesamtvarianz, was eine relativ gute Anpassung des Modells zeigt.

## 2.3 Prognose der Mortalität

Zum Erstellen der Mortalitätsprognose ist es erforderlich ein adäquates Zeitreihenmodell für den Mortalitätsindex  $\mathbf{k}_t$  zu finden. Nach standardmäßigen Modellidentifikationsprozeduren, siehe z.B. Hamilton (1994), zeigte sich der Random Walk mit Drift (auch Zufallsbewegung oder Irrfahrt mit Drift) für die beiden Mortalitätsindizes (weiblich und männlich) als geeignet.

Der Random Walk mit Drift ist ein ARIMA (Autoregressive Integrated Moving Average) Prozess, der nur ein integriertes Teil ersten Grades I(1) enthält, was einer Stationarität der ersten Differenzen entspricht, siehe Hamilton (1994). Für den Mortalitätsindex  $\mathbf{k}_t$  lautet der Random Walk mit Drift folgendermaßen:

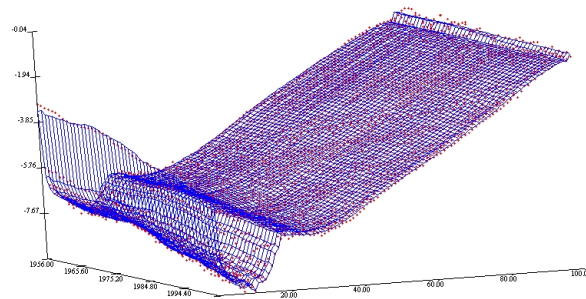
$$\mathbf{k}_t = \delta + \mathbf{k}_{t-1} + u_t, \quad (2.5)$$

wobei  $\delta$  den stochastischen Trend und  $u_t$  das weiße Rauschen<sup>2</sup> mit Nullerwartungswert und konstanter Varianz  $\sigma_u^2$  für alle  $t$  darstellt.

---

<sup>2</sup>Das weiße Rauschen (auch White Noise) ist eine Sequenz unkorrelierter Variablen mit einem Nullerwartungswert und einer konstanten endlichen Varianz  $\sigma_u^2$ , bezeichnet als  $u_t \sim \text{WN}(0, \sigma_u^2)$

## Sterbeziffern männlich (Log-Skala)



## Sterbeziffern weiblich (Log-Skala)

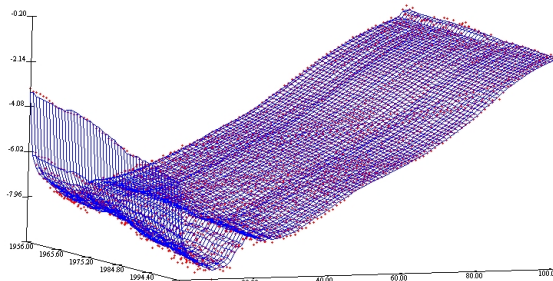


Abbildung 2.5: Tatsächliche (rote Punkte) und geschätzte (blaue Oberfläche) Sterbeziffern für Frauen auf logarithmischer Skala.

Mit Hilfe der statistischen Software XploRe werden die Parameter des Modells getrennt für den männlichen und den weiblichen Mortalitätsindex geschätzt. Folgende zwei Gleichungen zeigen die geschätzten Modelle mit den Standardabweichungen:

$$\begin{aligned} \mathbf{k}_t &= -1.84 + \mathbf{k}_{t-1} + u_t \quad \text{mit } \hat{\sigma}_u = 6.33 \quad \text{für Männer,} \\ \mathbf{k}_t &= -2.15 + \mathbf{k}_{t-1} + u_t \quad \text{mit } \hat{\sigma}_u = 5.11 \quad \text{für Frauen.} \end{aligned}$$

Die Konstante  $\delta = -1.84(-2.15)$  bezeichnet die durchschnittliche jährliche Änderung in  $\mathbf{k}$ . Wie in der Abbildung 2.4 zu sehen ist, sinkt der Mortalitätsindex bei Frauen schneller als bei Männern. Das bestätigt auch die geschätzte Konstante  $\delta$  im Zeitreihenmodell, die bei Frauen einen größeren absoluten Wert als bei Männern aufweist. Die geschätzte Standardabweichung  $\hat{\sigma}_u$  bezeichnet die Unsicherheit für die einjährige Prognose des Prozesses  $\mathbf{k}_t$ , die Unsicherheit steigt mit der Quadratwurzel des vorhersagten Zeithorizonts.

Die Vorhersagefunktion für den Random Walk mit Drift bekommt man als Lösung folgender Differentialgleichung, die man aus der Definition (2.5) nach Vernachlässigung des weißen Rauschens bekommt, für Details siehe z.B. Moral und González (2003):

$$k_t - k_{t-1} = \delta . \tag{2.6}$$

Für den Vorhersagewert  $k_T(l)$  mit Ausgangszeitpunkt  $T$  im Zeitpunkt  $T + l$  gilt dann:

$$k_T(l) = c_T + l \cdot \delta \quad l = 1, 2, \dots . \tag{2.7}$$

Die Vorhersagefunktion ist damit eine lineare Funktion, die nur durch die Konstante  $c_T$  vom Ausgangspunkt  $T$  abhängt. Für  $l = 0$  gilt dann  $c_T = k_T$ . Die Abbildung (2.6) zeigt die Vorhersagefunktion  $\mathbf{k}_T(l)$  bis zum Jahr 2070 zusammen mit dem 95%-Konfidenzintervall, der nach 5000 Simulationen des Prozesses mit einem normalverteilten weißen Rauschen geschätzt wurde.

Für eine bessere Darstellung der Mortalität wird aus den geschätzten altersspezifischen Sterbeziffern die Lebenserwartung berechnet. Die (fernere) Lebenserwartung im Alter  $x$  im Zeitpunkt  $t$  lässt sich nach demographischen Standardverfahren als Funktion der

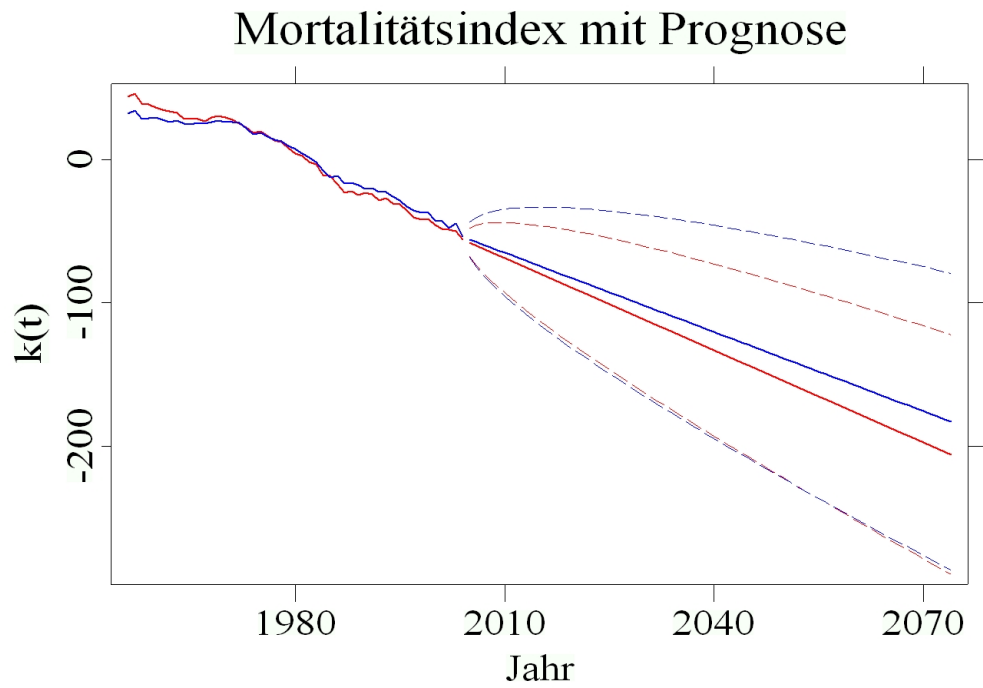


Abbildung 2.6: Der Mortalitätsindex mit seiner Prognose von 2005 bis 2074 mit 95%-Konfidenzintervall nach 5000 Simulationen. Blaue Linien stehen für Werte der männlichen Mortalität, rote Linien für Werte der weiblichen Mortalität.

einjährigen Sterbewahrscheinlichkeiten ausdrücken, siehe Babel (2007):

$$e_{x,t} = \frac{1}{2} + \sum_{y=x}^{\omega-1} \prod_{i=x}^y (1 - q_{i,t}) . \quad (2.8)$$

Hierbei bezeichnet  $q_{x,t}$  die „Wahrscheinlichkeit für eine genau  $x$ -jährige Person, vor Vollendung des nächsten Lebensjahres zu sterben“, nach Schmerbach (2001). Die Sterbewahrscheinlichkeit berechnet man mit Hilfe der Methode nach Farr aus den altersspezifischen Sterbeziffern wie folgt:

$$q_{x,t} = \frac{2m_{x,t}}{2 + m_{x,t}} . \quad (2.9)$$

Die Abbildung (2.7) zeigt den Zeitverlauf und einige der 5000 Simulation der Lebenserwartung eines neugeborenen Mädchens (schwarze Linien) und eines Jungens (blaue Linien). Auch in Zukunft liegt damit die Lebenserwartung der Männer deutlich unter der Lebenserwartung der Frauen, bei beiden Geschlechtern ist einen weiteren Anstieg der Lebenserwartung in der Zukunft zu erwarten.

Die Verteilung der Lebenserwartung eines Neugeborenen im Jahr 2070 zeigt die Abbildung 2.8 in Form eines Histogramms. Der Median der Lebenserwartung beträgt 90.5 Jahren bei Frauen bzw. 84.3 Jahren bei Männern. Der 90%-Konfidenzintervall liegt zwischen 86.8 und 93.5 Jahren bei Frauen bzw. 78.8 und 88.4 Jahren bei Männern. Eine Übersicht der vorausgesagten Werte mit 5%- und 95%-Quantilen in ausgewählten Jahren bietet die Tabelle 2.2. Bei den Frauen ist einen durchschnittlichen Anstieg 1.3 Jahre pro Jahrzehnt bis zum Jahr 2050 und danach 1.1 Jahre pro Jahrzehnt zu beobachten. Bei Männern zeigt sich der höheren Anstieg nur bis zum Jahr 2040, danach erhöht sich die Lebenserwartung um 1.1 bzw. nur 1 Jahr pro Dekade.

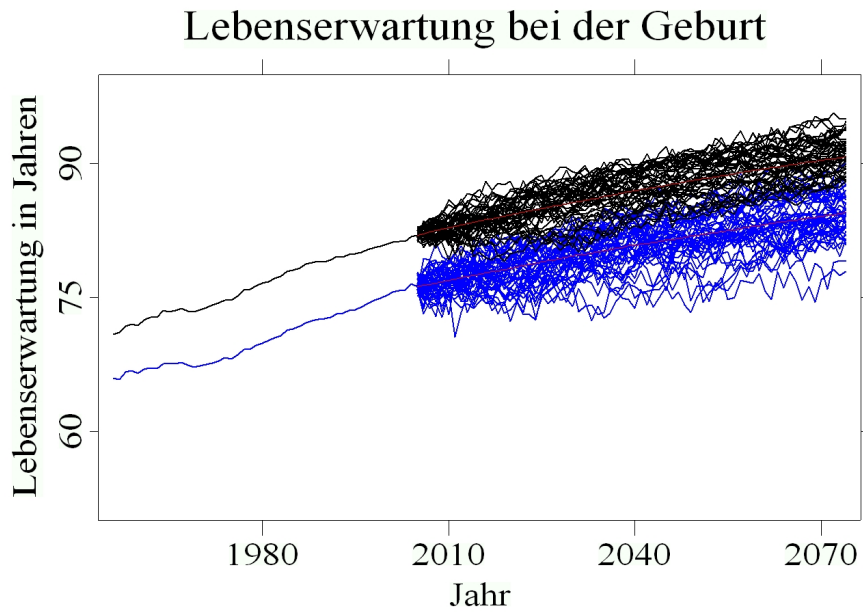


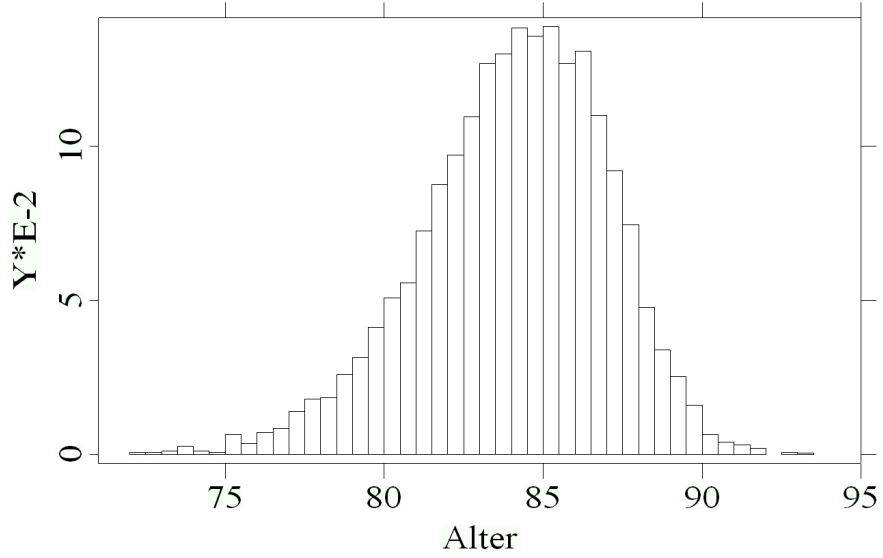
Abbildung 2.7: Einige Simulationen der Lebenserwartung bis zum Jahr 2074 für Männer (blau) und für Frauen (schwarz).

Jahr	Männer			Frauen		
	5%-Qntl.	Median	95%-Qntl.	5%-Qntl.	Median	95%-Qntl.
2010	74.3	77.0	79.3	81.0	82.8	84.6
2020	74.8	78.4	81.6	81.8	84.4	86.7
2030	75.5	79.7	83.4	82.6	85.8	88.5
2040	76.2	81.0	84.9	83.7	87.1	89.9
2050	77.0	82.2	86.2	84.8	88.3	91.3
2060	77.9	83.3	87.4	85.7	89.4	92.5
2070	78.8	84.3	88.4	86.8	90.5	93.5

Tabelle 2.2: Geschätzte Lebenserwartung eines Neugeborenen getrennt für Männer und Frauen zusammen mit 5%- und 95%-Quantilen nach 5000 Simulationen.



### Lebenserwartung bei Männern im Jahr 2070



### Lebenserwartung bei Frauen im Jahr 2070

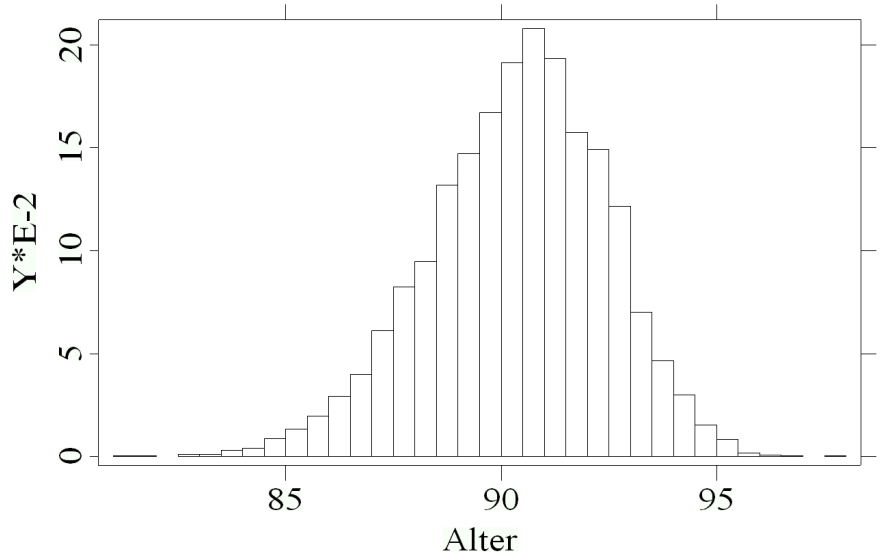


Abbildung 2.8: Histogramm der männlichen (oben) und weiblichen (unten) Lebenserwartung bei Geburt im Jahr 2070.

## 3 Fertilität

Fertilität oder auch Geburtenhäufigkeit ist nach der Mortalität eine weitere demographische Komponente, die die Bevölkerungsstruktur und -größe beeinflusst.

Die Geburtenhäufigkeit wird unter anderem mittels altersspezifischer Geburtenziffern gemessen, die als „Anzahl der Lebendgeborenen der Mütter bestimmten Alters je 1000 Frauen gleichen Alters“ definiert ist, siehe Stat.Bundesamt (2006b). Die Summe der altersspezifischen Geburtenziffern gibt die Total Fertility Rate (TFR), auch zusammengefasste Geburtenziffer, die sich als „Zahl der Kinder, die 1000 Frauen im Laufe ihres Lebens gebären, und zwar unter der Annahme, dass sich die altersspezifischen Geburtenziffern künftig nicht ändern“ ergibt, siehe Stat.Bundesamt (2006b). Die Total Fertility Rate wird von Änderungen im Altersaufbau nicht beeinflusst und ist damit für den internationalen Vergleich geeignet<sup>1</sup>. Die Abbildung 3.1 zeigt den Verlauf der TFR bezogen auf eine Frau in Westdeutschland zwischen den Jahren 1950 und 2005. Das Maximum erreichte die TFR im Jahr 1964 mit 2.54 Kindern pro Frau. Nach der sog. „Babyboom“-Phase Ende der 50er und Anfang der 60er Jahren kam nach der Erfindung der Antibabypille Ende der 60er Jahren ein starker Rückgang der Fertilität bis zum Tiefpunkt im Jahr 1985 mit 1.28 Kinder pro Frau. Ab diesem Zeitpunkt schwankt die Total Fertility Rate um den Wert 1.4 Kinder pro Frau.

In diesem Kapitel wird die stochastische Modellierung der altersspezifischen Fertilität vorgestellt: der nächste Abschnitt beschreibt den dafür verwendeten Datensatz, der Abschnitt 3.2 zeigt das verwendete Modell und der letzte Abschnitt 3.3 fasst die gewonnenen Ergebnisse zusammen.

---

<sup>1</sup>In den internationalen Statistiken wird oft die Total Fertility Rate bezogen auf 1 Frau verwendet.

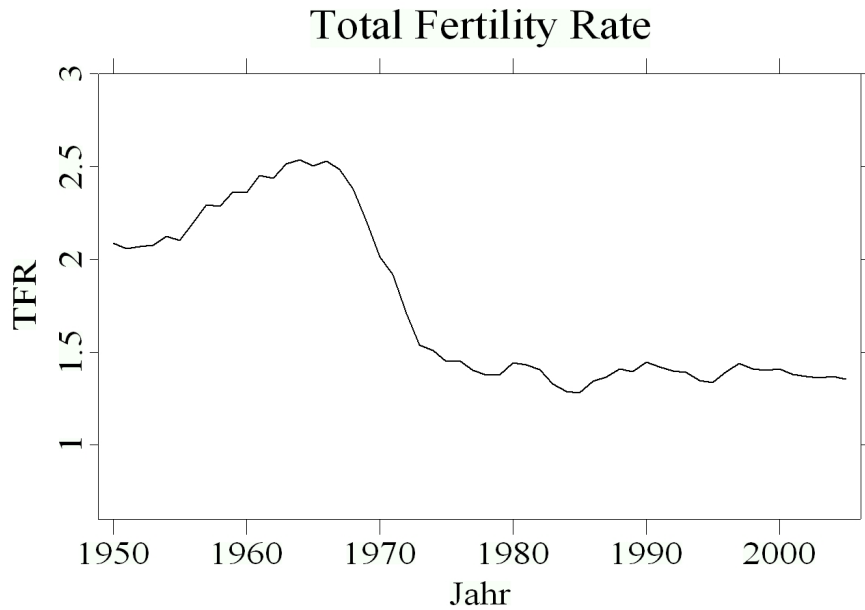


Abbildung 3.1: Total Fertility Rate pro Frau in Westdeutschland.

### 3.1 Datensatz

Für die Fertilitätsanalyse werden die altersspezifischen Geburtenziffer der 15- bis 44-jährigen Mütter verwendet. Vom Statistischen Bundesamt stehen die altersspezifischen Geburtenziffern für Westdeutschland vom Jahr 1950 bis 2005, seit der Wiedervereinigung auch für Ostdeutschland<sup>2</sup> zur Verfügung. Analog zu dem Mortalitätsverfahren ist anzunehmen, dass sich die Fertilitätsverhältnisse der ostdeutschen Frauen an diejenigen der westdeutschen Frauen anpassen werden, was auch in Abbildung 3.2 zu erkennen ist. Die blaue Linie stellt die altersspezifische Geburtenziffer für Westdeutschland, die rote Linie diejenige für Ostdeutschland dar. Im Jahr 1990 sind insgesamt höhere Kinderzahlen bei jüngeren Müttern im Ostdeutschland zu beobachten. Das durchschnittliche Alter der Mutter bei Geburt eines Kindes lag 1990 im Ostdeutschland bei 25.1 Jahren, im Westdeutschland waren es dagegen 28.2 Jahre. Die Situation hat sich in 15 Jahren geändert: die Frauen im Ostdeutschland bekommen ihre Kinder immer später, im Durchschnitt

<sup>2</sup>©Statistisches Bundesamt, Wiesbaden, 2006

mit 28.5 Jahren, was bereits nah an dem Durchschnitt der westdeutschen Frauen (29.6 Jahre) liegt.

Deshalb folgt im Weiteren nur eine Analyse der Geburtenziffern für Westdeutschland, die geschätzten Werte zusammen mit der Zukunftsprognose werden für das gesamte Bundesgebiet übernommen.

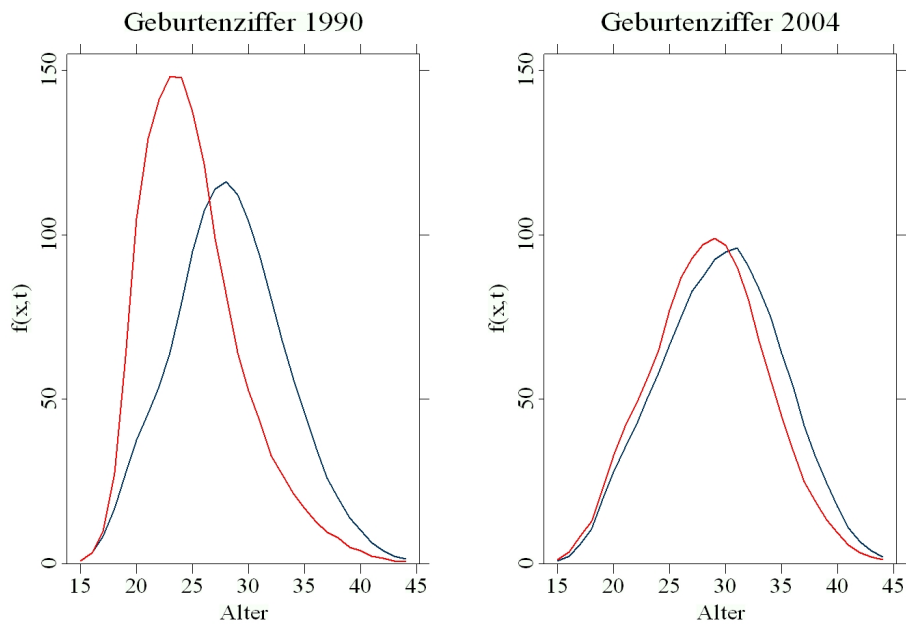


Abbildung 3.2: Altersspezifische Geburtenziffern für Ostdeutschland (rot) und Westdeutschland (blau) in den Jahren 1990 und 2004.

## 3.2 Modellierung der Fertilität

Analog zu den altersspezifischen Sterbeziffern ändern sich auch die altersspezifischen Geburtenziffer in der Zeit und gleichzeitig in den einzelnen Altersgruppen der Mütter. Eine dreidimensionale Darstellung der Geburtenziffern bietet die Abbildung 3.3, wo an der  $x$ -Achse die Zeitspanne, an der  $y$ -Achse die einzelne Altersgruppen der Mütter und an der  $z$ -Achse die dazugehörigen Geburtenziffern abgebildet sind. Die Grafik zeigt sowohl

# Geburtenziffern

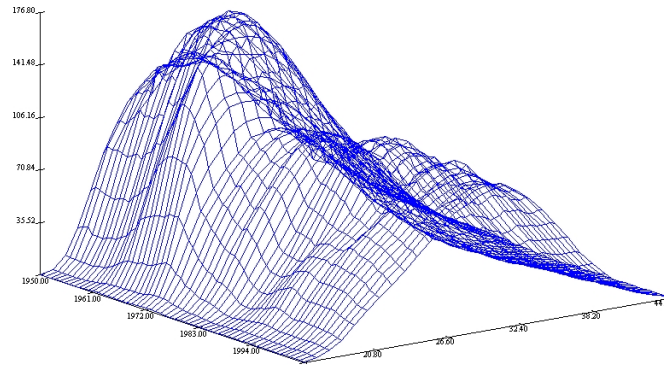


Abbildung 3.3: Altersspezifische Geburtenziffern der 15- bis 44-jährigen Mütter von 1950 bis 2005.

die abnehmende Anzahl der geborenen Kindern ab Ende der 60er Jahren als auch das steigende Alter der Mütter beim Geburt eines Kindes.

Für die weitere Analyse werden die zeit- und altersabhängige Komponente der Geburtenhäufigkeit getrennt. Analog zu Modell (2.1) wird auf die Geburtenziffer  $\mathbf{f}_{x,t}$  der  $x$ -jährigen Mutter im Jahr  $t$  das Lee-Carter-Modell angewendet:

$$\mathbf{f}_{x,t} = \mathbf{a}_x + \mathbf{b}_x \mathbf{f}_t + \varepsilon_{x,t} . \quad (3.1)$$

$\mathbf{a}_x$  und  $\mathbf{b}_x$  repräsentieren ebenfalls die altersabhängigen Parameter;  $\mathbf{f}_t$  den zeitabhängigen Fertilitätsindex und  $\varepsilon_{x,t}$  den Störterm.

Das Lee-Carter-Modell für die Fertilität (3.1) wurde als lineares Modell angewendet, da man hiermit bessere Schätzwerte als mit einem logarithmischen Modell, wie z.B. bei der Mortalitätsmodellierung verwendet, erhält, siehe Lee (1993). Im Fall der Geburtenziffern

handelt es sich außerdem um relativ große Werte im Bereich zwischen 1 und 180 Kindern. Aus diesem Grund besteht kein Risiko der Schätzung einer negativen Geburtenziffern bei der späteren Prognoseschätzung, weshalb die Anwendung eines logarithmischen Modells nicht zwingend notwendig ist.

Da die Gleichung (3.1) keine eindeutige Lösung hat, müssen zunächst an die Komponenten, analog zu (2.2), folgende Beschränkungen gesetzt werden:

$$\sum_{x=15}^{44} \mathbf{b}_x = 1 \quad \sum_{t=1}^T \mathbf{k}_t = 0 . \quad (3.2)$$

Der Ausgangsjahr mit  $t = 1$  ist das Jahr 1950, für das Endjahr 2005 gilt dann:

$$T = 2005 - 1950 + 1 = 56.$$

Aus den Beschränkungen (3.2) und der Gleichung (3.1) folgt für die altersspezifischen Konstanten:

$$\mathbf{a}_x = \frac{1}{T} \sum_{t=1}^T \mathbf{f}_{x,t} . \quad (3.3)$$

Die altersspezifischen Konstanten  $\mathbf{b}_x$  und den Fertilitätsindex  $\mathbf{f}_t$  bekommt man als Hauptvektoren der orthonormalen Matrizen, die man mit Hilfe der Singulärwertzerlegung der Matrix  $[\mathbf{f}_{x,t} - \mathbf{a}_x]$  ableitet, analog zu (2.4).

Die Summe der Konstanten  $\mathbf{a}_x$  über das Alter  $x$ , bezeichnet als  $A$ , ist als der Durchschnittswert der TFR über die ganze Zeitspanne zu interpretieren. Summiert man die beiden Seiten der Gleichung (3.1) über  $x$ , bekommt man für den Zeitverlauf der TFR:

$$\text{TFR}_t = A + \mathbf{f}_t + \mathbf{E}_t , \quad (3.4)$$

wobei  $\mathbf{E}_t$  die Summe der  $\varepsilon_{x,t}$ s über das Alter  $x$  bezeichnet. So kann man den Fertilitätsindex  $\mathbf{f}_t$  als die Abweichung der TFR in der Periode  $t$  von ihrem langzeitigen Durchschnittswert erklären.

Die geschätzten Werte der altersspezifischen Konstanten  $\mathbf{a}_x$  und  $\mathbf{b}_x$  zeigt die Tabelle 3.1. Die größten Werte über 100 des Parameters  $\mathbf{a}_x$  liegen in den Altersgruppen zwischen 23 und 30 Jahren, in welchen die Mütter die meisten Kinder bekommen. Die größte Veränderung in der Zeitperiode geschah in den Altersgruppen der 21- bis 25-jährigen Müttern, wo die Konstante  $\mathbf{b}_x$  ihre größte Werte über  $7 \cdot 10^{-2}$  annimmt.

Alter	$a_x$	$b_x \cdot 10^3$	Alter	$a_x$	$b_x \cdot 10^3$
15	0.77	-0.02	30	101.59	30.30
16	3.65	1.17	31	92.03	26.68
17	12.24	8.59	32	81.65	24.97
18	27.17	24.17	33	71.18	24.29
19	46.25	42.39	34	60.93	23.62
20	64.03	59.06	35	51.83	23.31
21	78.73	70.77	36	43.39	23.85
22	91.96	79.47	37	35.29	22.86
23	101.71	80.41	38	28.15	21.12
24	109.47	77.57	39	21.92	18.47
25	115.45	70.74	40	16.67	15.67
26	118.47	62.32	41	12.13	12.95
27	118.02	53.56	42	8.35	9.73
28	114.79	44.47	43	5.42	6.78
29	109.05	36.45	44	3.21	4.30

Tabelle 3.1: Geschätzte Werte (SVD) der Parameter  $a_x$  und  $b_x$ .

In der Abbildung 3.4 ist erkennbar, dass der Verlauf des Fertilitätsindex  $f_t$  beinahe den Verlauf der TFR (Abb. 3.1) kopiert, was auch eine große Varianz des Index  $f_t$  verursacht ( $\sigma_f^2 = 2 \cdot 10^5$ ).

Trotz der großen Schwankungen liegt die aufgeklärte Varianz des geschätzten Modells (3.1) für Mütter zwischen 20 und 29 und zwischen 40 und 44 Jahren über 80%. Alle anderen Altersgruppen weisen eine kleinere aufgeklärte Varianz auf, wobei bei den jüngsten Müttern zwischen 15 und 19 Jahren die kleinsten Werte zwischen 10 und 54% erreicht werden. Berechnet man die aufgeklärte Varianz über alle Altersgruppen, erklärt somit das Modell 80.5% der Gesamtvarianz.

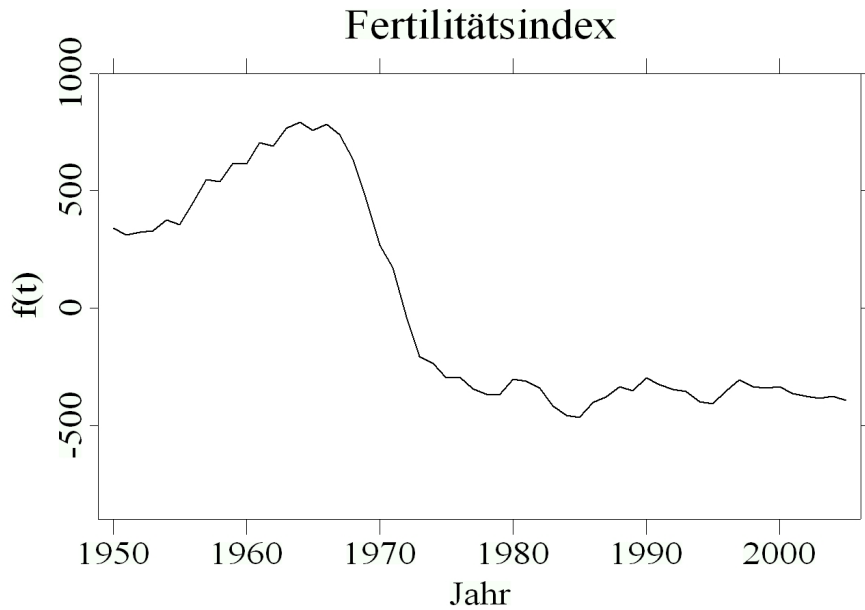


Abbildung 3.4: Geschätzter Fertilitätsindex  $f_t$ .

### 3.3 Prognose der Fertilität mit dem transformierten Fertilitätsindex

Nach der Modellschätzung im Abschnitt 3.2 folgt nun die Prognose der Fertilität mit dem Modell für den angepassten Wert der TFR  $F_t$ :

$$F_t = f_t + A, \quad (3.5)$$

$f_t$  beschreibt die Abweichungen von dem langzeitigen Durchschnittswert  $A$ . Aufgrund der hohen Varianz von  $f_t$ , die mit dem Prognosejahr noch steigt, kann eine direkte Prognose für  $F_t$  nicht demographisch plausible Ergebnisse liefern, siehe Lee (1993).

Um negative oder zu hohe Fertilitätsraten in der Prognose zu verhindern ist es hilfreich, vorherdefinierte obere und untere Grenzen in das Modell zu integrieren. Beide Grenzen, bezeichnet als  $L$  (untere) und  $U$  (obere), können großzügig wie z.B. 0 und 10 Geburten pro Frau oder auch mehr realistisch mit z.B. 1 und 3 Kindern pro Frau gewählt werden.



Modelliert wird dann ein transformierter Fertilitätsindex  $g_t$ , definiert als:

$$g_t = \log \left( \frac{F_t - L}{U - F_t} \right), \quad (3.6)$$

sodass für  $F_t$  gilt:

$$F_t = \frac{U \cdot \exp(g_t) + L}{1 + \exp(g_t)}. \quad (3.7)$$

Für die Konvergenzen ergibt sich damit:

$$\begin{aligned} g_t \rightarrow \infty &\implies F_t \rightarrow U \\ g_t \rightarrow -\infty &\implies F_t \rightarrow L. \end{aligned}$$

Hieraus folgt, dass die Prognose der TFR und auch der Konfidenzintervall zwischen den beiden Werten  $L$  und  $U$  liegen.

Selbst in dem oben beschriebenen Modell mit einer unteren und oberen Grenze liefert die Prognose keine demographisch konsistenten und stabilen Werte. Nach dem klassischen Box-Jenkins-Verfahren, siehe Hamilton (1994), wurde als adäquat ein ARMA(1,1)<sup>3</sup> Modell geschätzt, welches als Punktschätzung für das Jahr 2050 1.3 Geburten pro Frau generiert. Ein beinahe identisches Modell liefert jedoch eine Schätzung für dasselbe Jahr von 1.7 Geburten pro Frau. Solche Unterschiede in der Punktschätzung der Fertilität haben einen starken Einfluss auf die Prognose der Altersverteilung und die Gesamtgröße der Bevölkerung und könnten damit erhebliche Schwankungen in den Simulationen verursachen.

Deshalb wird ein Modell konstruiert, in dem die Fertilität auf einem prädefinierten Niveau (auch „ultimativer“ oder „endgültiger Level“) balanciert, siehe Lee (1993). Gegen dieses Niveau konvergiert die Prognose langfristig.

Folgenderweise wird der Level in das Zeitreihenmodell eingebaut: Man betrachte einen allgemeinen ARMA ( $p, q$ ) Prozess für den Fertilitätsprozess  $F_t$ :

$$F_t = c + \phi_1 F_{t-1} + \phi_2 F_{t-2} + \dots + \phi_p F_{t-p} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q},$$

---

<sup>3</sup>„Autoregressive Moving Average Process“ mit einem autoregressiven Parameter und einem Parameter der gleitenden Mittelwerte.

wobei  $\phi_1, \phi_2, \dots, \phi_p$  die Konstanten des autoregressiven Teils und  $\theta_1, \theta_2, \dots, \theta_p$  die Konstanten des Teils der gleitenden Mittelwerte darstellen.  $u_t$  bezeichnet ein weißes Rauschen mit einem Nullerwartungswert und der konstanten Varianz  $\sigma_u^2$ . Ein endgültiges Niveau  $F^*$  bekommt man dann als

$$F^* = \frac{c}{1 - \sum_{i=1}^p \phi_i},$$

unter der Voraussetzung, dass  $\sum_{i=1}^p \phi_i \neq 1$ . Danach kann man den Konstantterm in dem ARMA Modell schreiben als  $c' = F^* \cdot (1 - \sum_{i=1}^p \phi_i)$ . Jetzt kann man den Prozess  $y_t = F_t - F^*$  definieren, der sich mit einem ARMA Prozess mit einer Nullkonstante beschreiben lässt. Es werden die Parameter und die Prognose für den Prozess  $y_t$  geschätzt, die Prognosewerte für  $F_t$  bekommt man als die Addition des Levels  $F^*$  zu den Prognosewerten für  $y_t$ . Kombiniert man diesen Ansatz mit der logarithmischen Transformation (3.7), ist der Prozess  $[g_t - G^*]$  mit  $G^* = \log\left(\frac{F^* - L}{U - F^*}\right)$  abzuschätzen.

An dieser Stelle kann man die Frage stellen, warum überhaupt ein stochastischer Ansatz verwendet wird, wenn das Modell anschließend so viel eingeschränkt wird. Das stochastische Modell liefert jedoch zwei wichtige Vorteile: Erstens liefert es nicht nur eine Punktschätzung der prognostizierten Variable sondern auch eine Risikoschätzung der Prognose. Und zweitens kann man mit dem stochastischen Ansatz die (Auto-)Korrelationen der Fertilitätsprognose abschätzen, die eine wichtige Rolle bei der Bevölkerungsvorausberechnung spielen, siehe Lee (1993); Babel (2007).

Die standardmäßigen Modellidentifikationsprozeduren (siehe Hamilton (1994)) angewandt an den Prozess  $[g_t - G^*]$  führen zu dem ARMA Prozess erster Ordnung in beiden Teilen, bezeichnet auch als ARMA(1,1):

$$g_t - G^* = \phi_1 \cdot (g_{t-1} - G^*) + u_t + \theta_1 \cdot u_{t-1}, \quad (3.8)$$

mit den Konstanten  $\phi_1$  und  $\theta_1$  und dem weißen Rauschen  $u_t$  mit Varianz  $\sigma_u^2$ .

Anhand der Entwicklung der TFR in der 2. Hälfte des 20. Jahrhunderts wurden die Grenzen an das Modell 0 und 4 Geburten pro Frau gewählt, als das endgültige Niveau

wurde der Durchschnittswert der TFR der letzten 20 Jahren genommen:  $F^* = 1.395$ . Daraus ergibt sich für  $G^* = \log\left(\frac{1.395 - 0}{4 - 1.395}\right) = -0.63$ . Die Parameter des Modells (3.8) mit den erwähnten Beschränkungen wurden mit Hilfe des statistischen Programm XploRe geschätzt. Das angepasste Modell lautet:

$$g_t + 0.63 = 0.98 \cdot (g_{t-1} + 0.63) + u_t + 0.25 \cdot u_{t-1} , \quad (3.9)$$

mit  $\hat{\sigma}_u = 0.09$ . Das geschätzte ARMA(1,1) Modell ist damit stationär ( $|\phi_1| = 0.98 < 1$ ) und invertierbar ( $|\theta| = 0.25 < 1$ ), es kann daher eine direkte Vorhersage für den Prozess  $g_t$  hergeleitet werden, siehe Moral und González (2003). Für den Vorhersagewert  $g_T(l)$  mit Ausgangspunkt  $T$  im Zeitpunkt  $T + l$  gilt dann:

$$g_T(l) = G^* + c_T \cdot \theta^l \quad l = 1, 2, \dots , \quad (3.10)$$

wobei  $c_T$  eine vom Ausgangspunkt  $T$  abhängige Konstante darstellt. Mit  $l = 0$  gilt dann  $c_T = g_T - G^*$ . Unter Anwendung der Gleichung (3.7) berechnet man die Prognosewerte für  $F_t$  und anschließend  $f_t = F_t - A$ . Die Werte der altersspezifischen Geburtenziffern bekommt man aus der Gleichung (3.1).

Die Abbildung (3.5) zeigt die Vorhersagefunktion für  $\mathbf{F}_t$  bis zum Jahr 2070 zusammen mit dem 95%-Konfidenzintervall, der nach 5000 Simulationen des Prozesses mit einem normalverteilten weißen Rauschen geschätzt wurde. Die rote gestrichelte Linie zeigt den Median der Prognose, der schnell auf das vordefinierte Niveau steigt. Die Verteilung der totalen Fertilitätsrate in den Jahren 2030 und 2070 zeigt die Abbildung 3.6. Im Jahr 2030 liegen die Werte der TFR mit 95%er Wahrscheinlichkeit zwischen 1.24 und 1.49 Geburten pro Frau, im Jahr 2070 zwischen 1.22 und 1.54 Geburten pro Frau, was aus heutiger Sicht realistisch erscheint.

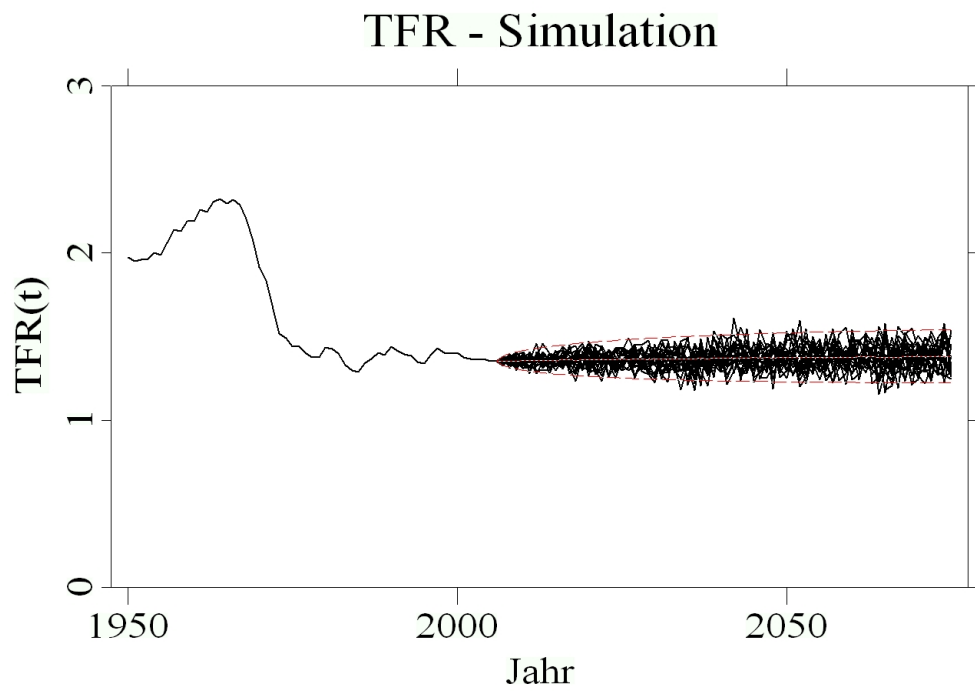


Abbildung 3.5: Total Fertility Rate (schwarz) mit der Prognose und ausgewählten Simulationen von 2006 bis 2074. Die rote gestrichelte Linien bezeichnen den Median und den 95%-Konfidenzintervall.

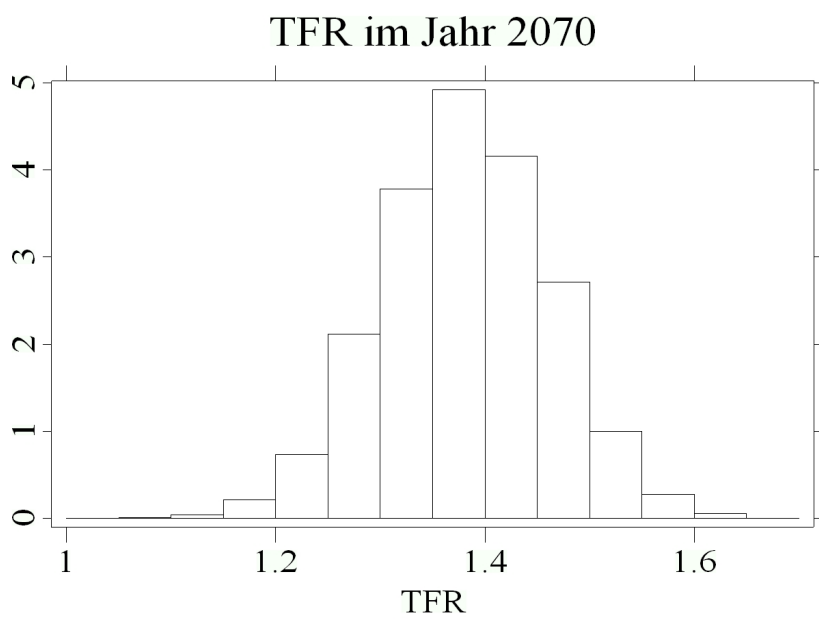
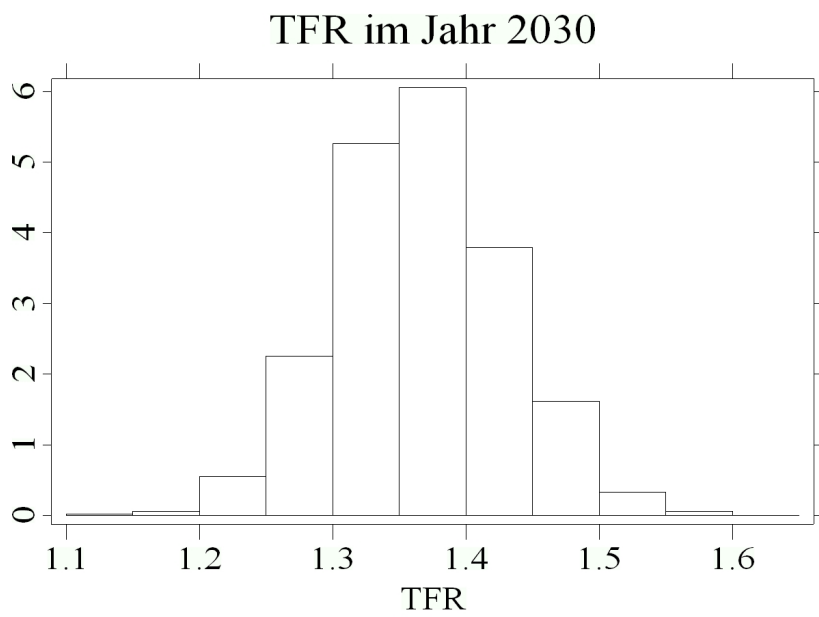


Abbildung 3.6: Histogramme der TFR im Jahr 2030 (oben) und im Jahr 2070 (unten).

## 4 Migration

Die Zu- und Auswanderung in Deutschland unterlag in den letzten Dekaden großen Schwankungen. Im Unterschied zur Mortalität beziehungsweise zur Fertilität lässt sich in der Entwicklung der Wanderung kein Trend erkennen. Für die zukünftige Bevölkerungszahl und die Altersstruktur ist dabei die Differenz zwischen Zu- und Fortzügen, das heißt der „Wanderungssaldo“, ausschlaggebend. Die Migration hängt von zahlreichen Faktoren ab. Auf der einen Seite wird die Zuwanderung von politischer, ökonomischer und demographischer Entwicklung der Herkunftsländer (z.B. Polen und Russland) beeinflusst, auf der anderen Seite kann die Migrationspolitik und die Anziehungskraft Deutschlands als Zielland die Wanderung steuern.

Die Abbildung 4.1 zeigt den Verlauf des Wanderungssaldos in den Jahren 1990 bis 2004. Sein Maximum mit über 1 Million Personen erreichte er im Jahr 1990, was Folge der Zuzügen der deutschstämmigen Bevölkerung (so genannte Spätaussiedler) aus Ost- und Südosteuropa war. Im Jahr 1997 sank der Saldo auf seinen Tiefpunkt mit einem Überschuss von 22 Tausend Fortzügen. Hierin spiegelten sich die Auswirkungen gesetzlicher Zuzugsbeschränkungen für Spätaussiedler wie auch der Aufstieg der Auswanderung von Deutschen wieder.

Betrachtet man sich die Entwicklung des Wanderungssaldos zusammen mit der Altersverteilung in Abbildung 4.2, zeigt sich ein deutlicher Überschuss der zugezogenen Personen im Alter zwischen 18 und 28 Jahren. Dies führt zu einem „Verjüngungseffekt“ der Deutschen Bevölkerung, da die zuziehenden Personen im Durchschnitt jünger sind als die fortziehende, siehe Stat.Bundesamt (2006b).

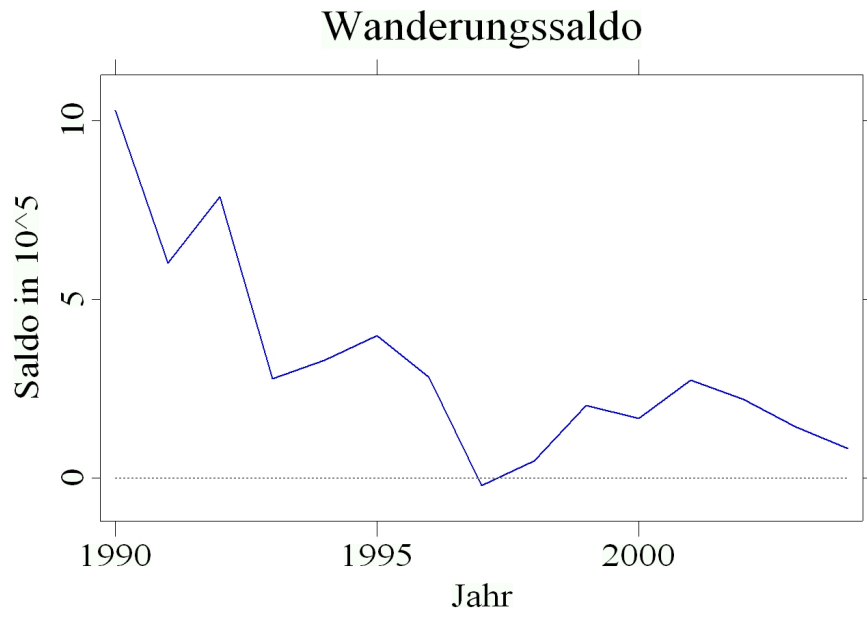


Abbildung 4.1: Verlauf des Wanderungssaldo vom Jahr 1990 bis 2004.

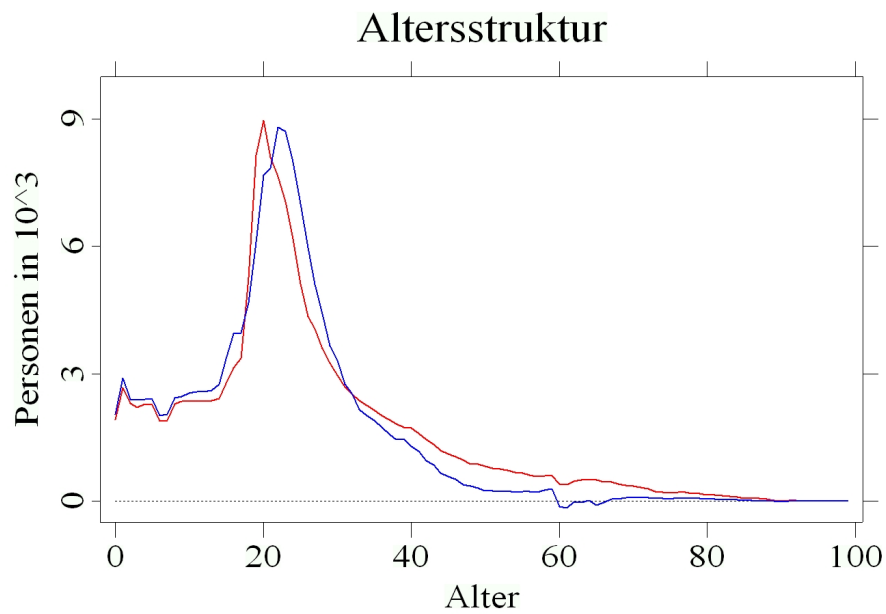


Abbildung 4.2: Altersstruktur der weiblichen (rot) und männlichen (blau) Zu- und Fortzüge, Mittelwerte aus dem Zeitraum 1990 bis 2004.

Wegen der großen Anzahl der Einflussfaktoren ist eine adäquate stochastische Modellierung der Migration problematisch. Für die Bevölkerungsvorausberechnung wird ein konstanter jährlicher Wanderungssaldo von 155 000 Personen angenommen, was dem Durchschnittswert der letzten 8 Jahren entspricht. Die Altersverteilung der zu- und fortgezogenen Personen wird dementsprechend als Durchschnitt der jeweiligen Altersgruppen gewählt. Diese Annahme weicht nicht von den Annahmen des Statistischen Bundesamtes ab, welches mit zwei Szenarien mit jeweils 100 000 bzw. 200 000 zugezogenen Personen rechnet, siehe Stat.Bundesamt (2006b).



# 5 Bevölkerungsvorausberechnung

Nachdem in den vorherigen Kapiteln die Modellrechnung der Vitalprozesse beschrieben wurde, kann die Bevölkerungsprognose hergestellt werden. Die Prognose wird mit Hilfe der Kohorten-Komponenten-Methode berechnet, die im folgenden Abschnitt vorgestellt wird. Im dann nachfolgenden Abschnitt werden die Ergebnisse der Modellrechnung präsentiert und zuletzt wird die stochastische Bevölkerungsvorausberechnung mit den Ergebnissen der 11. koordinierten Bevölkerungsvorausberechnung des Statistischen Bundesamtes verglichen.

## 5.1 Kohorten-Komponenten-Methode

Die Kohorten-Komponenten-Methode verbindet die vorgestellten Ansätze zur Modellierung von Mortalität, Fertilität und Migration und ermöglicht Modellrechnungen vom Umfang und der Altersstruktur der Bevölkerung durchzuführen.

In der mathematischen Notation der Kohorten-Komponenten-Methode werden folgende demographische Variablen verwendet:

- $N_{x,t}^M$  bzw.  $N_{x,t}^F$  als die männliche bzw. weibliche Bevölkerung der  $x$ -jährigen Altersgruppe am Ende des Jahres  $t$ ;
- $F_{x,t}$  als die altersspezifische Geburtenrate der  $x$ -jährigen Mutter im Jahr  $t$ , was der altersspezifischen Geburtenziffer bezogen auf eine Frau entspricht,

- $p_{x,t}^M$  bzw.  $p_{x,t}^F$  als die Wahrscheinlichkeit eines  $x$ -jährigen Mannes bzw. Frau das nächste Lebensjahr zu erreichen;
- $M_{x,t}^M$  bzw.  $M_{x,t}^F$  als die Summe der männlichen bzw. weiblichen Migranten der Altersgruppe  $x$  im Jahr  $t$ ;
- $s$  als Geschlechtsverhältnis von weiblichen zu männlichen Neugeborenen (In der Modellrechnung wird ein Verhältnis 100:106 angenommen.);
- $x = 0, 1, 2, \dots, k-1, k$ , wobei  $k$  die älteste Altersgruppe der 110-jährigen und älteren bezeichnet.

Dann ergibt sich für die weibliche Bevölkerung am Ende des Jahres  $t+1$ :

$$\begin{pmatrix} N_{0,t+1}^F \\ N_{1,t+1}^F \\ N_{2,t+1}^F \\ \vdots \\ \vdots \\ N_{k,t+1}^F \end{pmatrix} = \begin{pmatrix} 0 & \dots & sF_{15,t}p_{-1,t} & \dots & sF_{44,t}p_{-1,t} & 0 & \dots & 0 \\ p_{0,t}^F & 0 & 0 & \dots & \dots & 0 & \dots & 0 \\ 0 & p_{1,t}^F & 0 & \dots & \dots & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & & & \vdots & \\ \vdots & \vdots & & & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \dots & 0 & p_{k-1,t}^F & 0 \end{pmatrix} \cdot \begin{pmatrix} N_{0,t}^F \\ N_{1,t}^F \\ N_{2,t}^F \\ \vdots \\ \vdots \\ N_{k,t}^F \end{pmatrix} + \begin{pmatrix} M_{0,t}^F \\ M_{1,t}^F \\ M_{2,t}^F \\ \vdots \\ \vdots \\ M_{k,t}^F \end{pmatrix},$$

und für die männliche Bevölkerung (ohne Neugeborene) am Ende des Jahres  $t+1$ :

$$\begin{pmatrix} N_{0,t+1}^M \\ N_{1,t+1}^M \\ N_{2,t+1}^M \\ \vdots \\ \vdots \\ N_{k,t+1}^M \end{pmatrix} = \begin{pmatrix} 0 & \dots & 0 & \dots & \dots & 0 & \dots & 0 \\ p_{0,t}^M & 0 & 0 & \dots & \dots & 0 & \dots & 0 \\ 0 & p_{1,t}^M & 0 & \dots & \dots & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & & & \vdots & \\ \vdots & \vdots & & & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \dots & 0 & p_{k-1,t}^M & 0 \end{pmatrix} \cdot \begin{pmatrix} N_{0,t}^M \\ N_{1,t}^M \\ N_{2,t}^M \\ \vdots \\ \vdots \\ N_{k,t}^M \end{pmatrix} + \begin{pmatrix} M_{0,t}^M \\ M_{1,t}^M \\ M_{2,t}^M \\ \vdots \\ \vdots \\ M_{k,t}^M \end{pmatrix}.$$

Das bedeutet, dass man die Bevölkerung in dem folgenden Jahr  $t+1$  als die jeweilige Bevölkerung des Vorjahres  $t$  multipliziert mit der zutreffenden Überlebenswahrscheinlichkeit und erhöht bzw. reduziert um den Wanderungssaldo im Jahr  $t$  erhält. Einfachheitshalber wird angenommen, dass alle Geburten in der Mitte eines Jahres stattfinden. Damit wird auch in der Jahresmitte ein neues Lebensalter erreicht. Dies ist auch der

Ausgangspunkt für die Berechnung der Überlebenswahrscheinlichkeit einer  $x$ -jährigen Person am Ende des Jahres  $t$ , die sich als eine bedingte Wahrscheinlichkeit aus der einjährigen Sterbewahrscheinlichkeiten<sup>1</sup> ergibt, siehe Diekmann et al. (2000):

$$p_{x,t} = \frac{(1 - q_{x,t}) \cdot (1 - \frac{q_{x+1,t+1}}{2})}{1 - \frac{q_{x,t}}{2}} . \quad (5.1)$$

Die Formel enthält im Nenner die Wahrscheinlichkeiten einer  $x$ -jährigen Person das nächste Lebensjahr zu erleben und einer  $x + 1$ -jährigen Person die Hälfte des nächsten Jahres zu erleben, dividiert durch die Wahrscheinlichkeit, dass eine  $x$ -jährige Person die Hälfte des Jahres  $t$  erreicht, da die  $x$ -jährige Person bis zum Jahresende des Jahres  $t$  bereits die Hälfte ihres  $x$ -ten Lebensjahres überlebt hat<sup>2</sup>. Diese Berechnung erfolgt für alle Altersgruppen vom 1. bis zum 109. Lebensjahr, die Überlebenswahrscheinlichkeit für die 110-jährige ist damit im Modell gleich Null.

Die Lebendgeborenen werden als die Summe der weiblichen Bevölkerung im gebärfähigen Alter multipliziert mit der altersspezifischen Geburtenziffern berechnet und anschließend nach dem Geschlechtsverhältnis in männliche bzw. weibliche Neugeborene aufgeteilt. Außerdem werden die Anzahl der Neugeborenen multipliziert mit der Wahrscheinlichkeit, die erste Hälfte des ersten Lebensjahres zu überleben (bezeichnet als  $p_{-1,t}$ ), die sich aus der Hälfte der einjährigen Sterbewahrscheinlichkeit für eine 0-jährige Person ergibt:

$$p_{-1,t} = 1 - \frac{q_{0,t}}{2} .$$

Für die männlichen und weiblichen Neugeborenen am Ende des Jahres  $t + 1$  gilt dann:

$$N_{0,t+1}^M = p_{-1,t}^M \cdot (1 - s) \cdot \sum_{x=15}^{44} F_{x,t} \cdot N_{x,t}^F$$

$$N_{0,t+1}^W = p_{-1,t}^F \cdot s \cdot \sum_{x=15}^{44} F_{x,t} \cdot N_{x,t}^F .$$

Für die Migration wird vereinfacht angenommen, dass die Fort- und Zuzüge am Ende des Jahres stattfinden und dass sich die zugezogenen Personen an die Mortalitäts- und Fertilitätsverhältnisse in Deutschland umgehend anpassen.

<sup>1</sup>Für eine einfache Notation wird hier auf eine Geschlechtsunterscheidung verzichtet.

<sup>2</sup>Unter der Annahme, dass die unterjährigen Sterbewahrscheinlichkeiten gleichverteilt sind.

## 5.2 Ergebnisse der Modellrechnung

Für die Erstellung der stochastischen Bevölkerungsprognose werden nach den bereits beschriebenen Modellen die Zufallsvariablen Mortalitäts- und Fertilitätsindex in jedem Prognosejahr generiert, daraus die altersspezifischen Mortalität und Fertilität bestimmt und zuletzt nach der Kohorten-Komponenten-Methode die Bevölkerungszahl in jeder Altersgruppe berechnet. Dieser Prozess wird vielfach wiederholt, um eine große Zahl der Stichprobenpfade für Bevölkerungs- und Vitalratenprozesse zu bilden und hierraus werden wiederum die Mittelwerte und die Wahrscheinlichkeitsverteilungen kalkuliert.

Als Ausgangsbevölkerung wird die Bevölkerungsstruktur vom 31.12.2004 nach Angaben des Statistischen Bundesamtes verwendet, wobei zur Bevölkerung alle Personen, die ihr Wohnsitz in Deutschland haben und auch gemeldete Ausländer und Ausländerinnen (einschließlich der Staatenlosen) zählen, siehe Stat.Bundesamt (2006b).

Alle Modellrechnungen und Grafiken wurden mit Hilfe der statistischen Software XploRe erzeugt. Bei der Bevölkerungsprognose wurden 5000 Simulationen generiert.

Die Tabelle 5.1 gibt die Mittelwerte und ausgewählten Quantile für die Prognose des Bevölkerungsumfangs in 5-Jahren-Intervallen an. Die Entwicklung der Bevölkerungszahl zwischen den Jahren 2004 und 2074 zeigt die Abbildung 5.1. Die durchgezogene Linie stellt den Mittelwert, die gestrichelte Linie den 95%-Konfidenzintervall dar und die roten Dreiecke bezeichnen die Prognosewerte der 11. koordinierten Bevölkerungsvorausberechnung des Statistischen Bundesamtes (Ober- und Untergrenze der mittleren Bevölkerung), siehe Stat.Bundesamt (2006a).

Die Bevölkerungszahl sinkt im Mittelwert von 82.4 Millionen im Jahr 2004 bis 2030 um 2.4 Millionen, bis 2050 um 10.5 Millionen und bis 2070 um fast 19 Millionen Menschen. Der Rückgang ist ab den 40er Jahren intensiver, da die starke Generation der „Babyboom“-Kinder (siehe Abbildung 3.1) ausstirbt. Der 95%-Intervall hat im Jahr 2050 eine Spannweite von 4.4 Millionen, im Jahr 2070 von 6.1 Millionen, was auch die Abbildung 5.2 mit Histogrammen der Bevölkerungsgröße in Jahren 2010, 2030, 2050 und 2070 zeigt.

Jahr	2.5%-Quantil	Mittelwert	97.5%-Quantil
2005	82.24	82.43	82.61
2010	82.01	82.53	83.04
2015	81.47	82.20	82.91
2020	80.71	81.63	82.56
2025	79.94	81.05	82.17
2030	78.71	80.00	81.26
2035	76.91	78.40	79.87
2040	74.81	76.55	78.26
2045	72.34	74.33	76.30
2050	69.71	71.92	74.15
2055	67.10	69.51	71.95
2060	64.78	67.39	70.02
2065	62.75	65.58	68.39
2070	60.81	63.83	66.90

Tabelle 5.1: Mittelwerte, 2.5%- und 97.5%-Quantile der Bevölkerungsprognose in Mio.

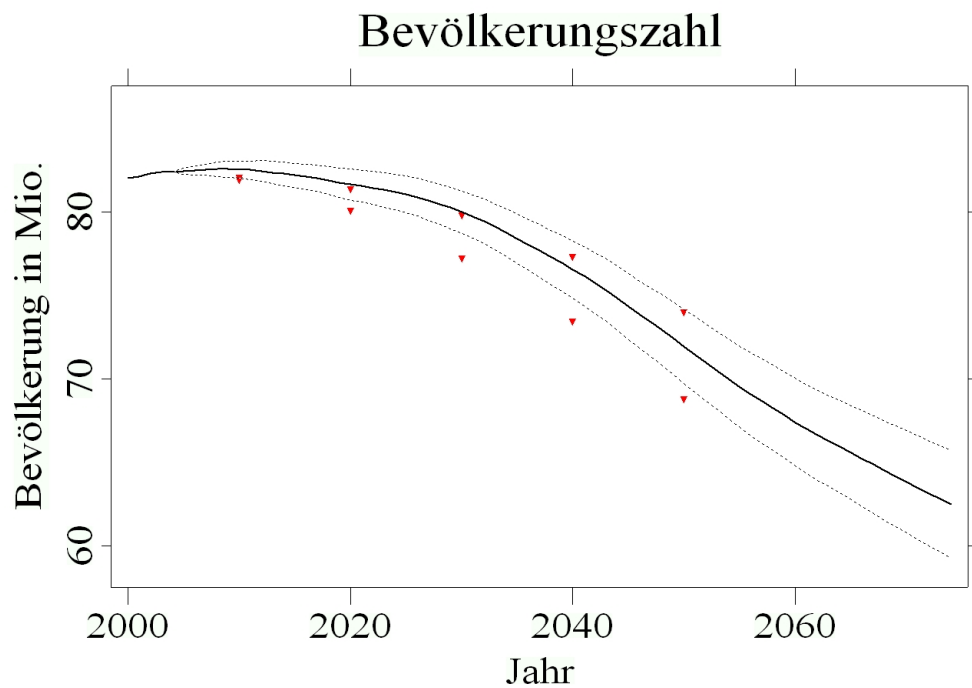


Abbildung 5.1: Mittelwert (durchgezogene Linie) und 95%-Intervall (gestrichelte Linie) für den Bevölkerungsumfang bis zum Jahr 2074 mit Schätzungen des Statistischen Bundesamtes (rote Dreiecke).

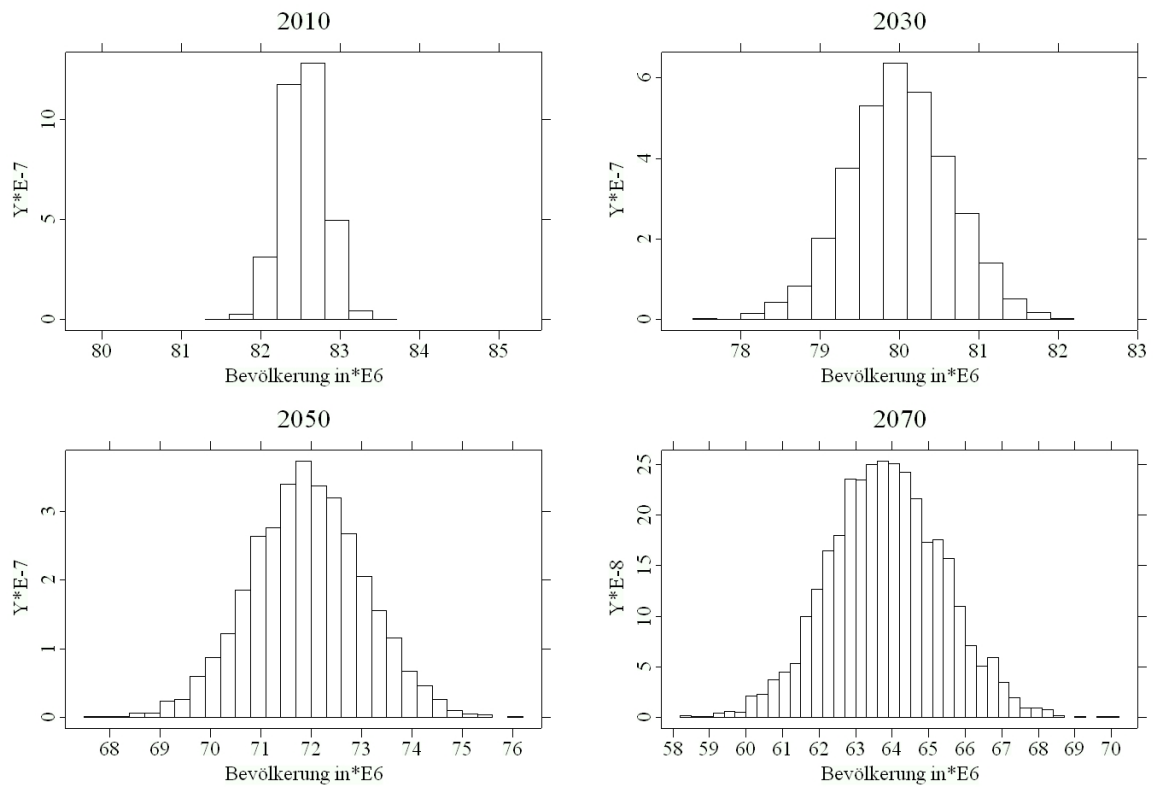


Abbildung 5.2: Verteilung der Bevölkerungszahl in den Jahren 2010, 2030, 2050 und 2070 in Mio.

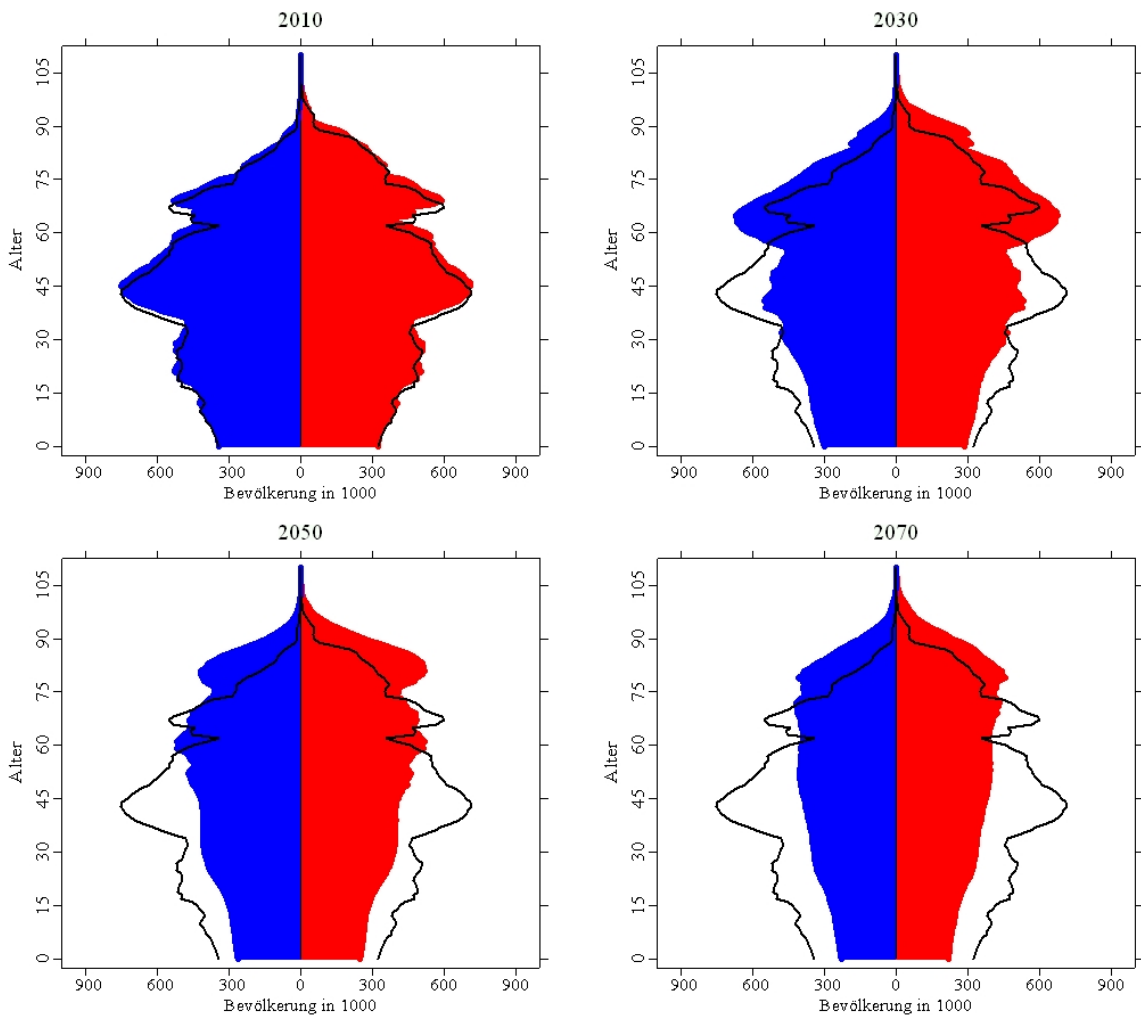


Abbildung 5.3: Bevölkerungspyramiden in den Jahren 2010, 2030, 2050 und 2070 zusammen mit der Bevölkerungspyramide des Jahres 2007 (schwarz). Das rote Feld stellt die weibliche, das blaue Feld die männliche Bevölkerung in 1000 dar.

Die Abbildung 5.3 zeigt die sogenannten Bevölkerungspyramiden in den Jahren 2010, 2025, 2050 und 2070. An der  $x$ -Achse ist die Bevölkerungsgröße in 1000, an der  $y$ -Achse die einzelnen Altersgruppen abgebildet. Das rote Feld stellt die weibliche, das blaue Feld die männliche Bevölkerung dar. Zum Vergleich ist mit der schwarzen Linie die Bevölkerungspyramide des Jahres 2007 abgebildet. Es ist deutlich zu erkennen, dass in Deutschland immer weniger Kinder geboren werden. Das liegt nicht nur an der niedrigen Geburtenraten, sondern auch an der immer sinkenden Zahl der Frauen, die Kinder gebären können. Bereits im Jahr 2030 ist die Anzahl der jungen Menschen unter 20 Jahren mit 13.2 Mio. um 3.7 Mio. gegenüber dem Jahr 2007 rückläufig. Bis zum Jahr 2050 sinkt die Zahl um weitere 2 Mio. bis auf 9.9 Mio. im Jahr 2070, was ein Unterschied zum Jahr 2007 von 7 Mio. Menschen darstellt.

Ein gegenläufiger Trend ist bei der älteren Bevölkerung zu beobachten, was die Folge der steigenden Lebenserwartung und der alternden starken Jahrgängen ist. Während die Anzahl der über 65-jährigen im Jahr 2007 bei 16.3 Mio. liegt, steigt diese bis zum Jahr 2030 um mehr als 6 Mio. auf 22.6 Mio. Diese Zahl bleibt bis zum Vorhersagehorizont relativ stabil, sie liegt im Jahr 2050 bei 22.7 Mio. und sinkt auf 21.1 Mio im Jahr 2070. Die Zahl der Personen im Erwerbsalter sinkt von 50 Mio. im Jahr 2007 um 5.7 Mio. im Jahr 2030, und um weitere 6.3 Mio. bis zum Jahr 2050. Bis zum Jahr 2070 sinkt die Zahl der Personen im Erwerbsalter um weitere 5.2 Mio. auf 32.2 Mio., was einen Unterschied gegenüber 2007 über 17 Mio. darstellt.

Die Veränderung in der Bevölkerungsstruktur zeigen auch der Jugend- und Altenquotient in den Abbildungen 5.4 bzw. 5.5. Der Jugendquotient zeigt das prozentuale Verhältnis der unter 20-jährigen zu den Personen im Erwerbsalter zwischen 20 bis unter 65 Jahren. Er sinkt von heutigen 33% auf sein Minimum im Jahr 2023 auf 28.4%, ab dem Jahr 2028 schwankt er zwischen 29 und 30%. Die Abbildung 5.5 zeigt den Altenquotient mit der Altersgrenze 65 Jahre (schwarz) und den Altenquotient mit der Altersgrenze 67 Jahre (rot), d.h. das Verhältnis der Rentnergeneration ab 65 bzw. 67 Jahren zu der Bevölkerung im Erwerbsalter zwischen 20 bis unter 65 bzw. 67 Jahre. In beiden Fällen erfolgt ab Mitte der 20er Jahren ein starker Anstieg des Altenquotientens, da in diesen



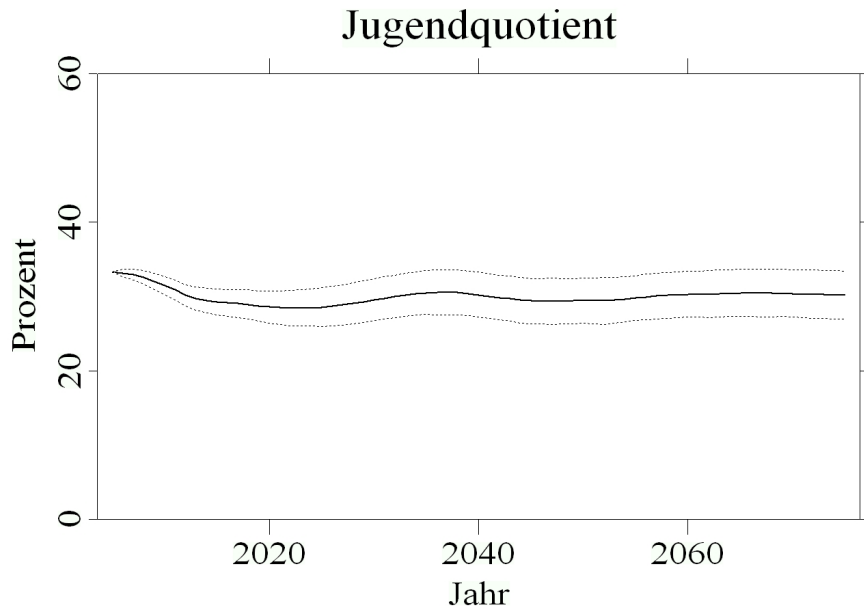


Abbildung 5.4: Jugendquotient (durchgezogene Linie) mit dem 95%-Konfidenzintervall (gestrichelte Linie).

Jahren die starke „Babyboom“-Generation in das Rentenalter kommt. Der Altenquotient mit Altersgrenze 65 verdoppelt sich von heutigen 30% auf 60% im Jahr 2040 und steigt dann langsamer bis zum Jahr 2074 auf 64%. Der Altenquotient mit Altersgrenze 67 erweist einen analogen Verlauf – er verdoppelt sich ebenso von heutigen 26% auf 53% im Jahr 2040 und steigt allmählich bis auf 57% im Jahr 2074.

## 5.3 Vergleich mit der Prognose des Statistischen

### Bundesamtes

Die hier präsentierte Ergebnisse werden mit den Ergebnissen der 11. koordinierten Bevölkerungsvorausberechnung des Statistischen Bundesamtes verglichen, für Details siehe Stat.Bundesamt (2006a). Das Statistische Bundesamt verwendet eine deterministische Methode der Vorausberechnung mit unterschiedlichen Annahmen an Lebenser-

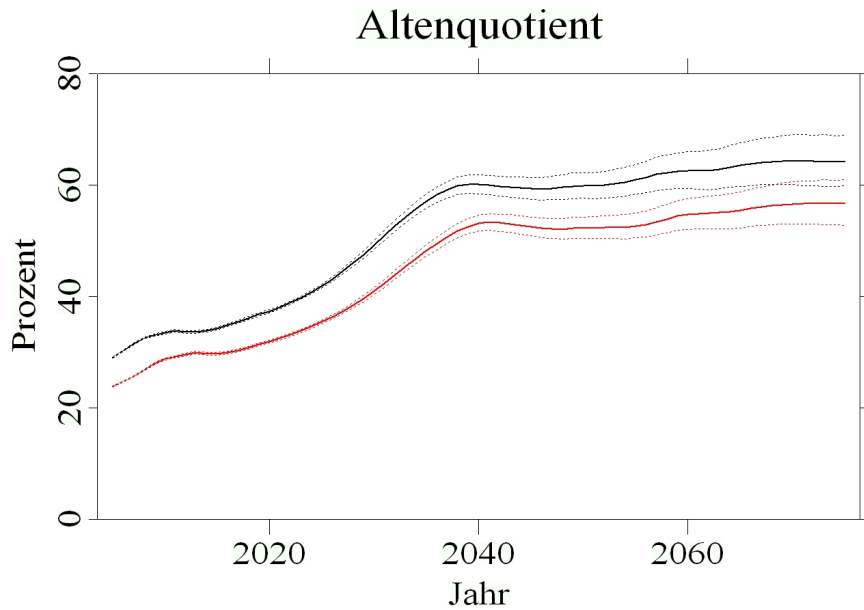


Abbildung 5.5: Altenquotient mit Altersgrenze 65 (schwarze Linie) und mit Altersgrenze 67 (rote Linie) zusammen mit dazugehörigen 95%-Konfidenzintervallen.

wartung, Fertilität und Migration.

Die Basisannahme an die Lebenserwartung der Männer bei Geburt im Jahr 2050 ist 83.5 Jahre, was ein Jahr über dem geschätzten Median aus dem Kapitel 2 liegt. Die Basisannahme und auch die Annahme mit hohem Anstieg (85.4 Jahre) fallen in das 90%-Konfidenzintervall (siehe Tabelle 2.2). Die Basisannahme an die Lebenserwartung der Frauen bei Geburt ist 88 Jahre, was nur 0.3 Jahre unter dem geschätzten Median aus dem Kapitel 2 liegt. Ebenfalls beide Annahmen (Annahme mit hohem Anstieg mit 89.8 Jahren bei Geburt) fallen unter die 5%- und 95% Quantile der geschätzten Lebenserwartung.

Das Statistische Bundesamt setzt 3 Varianten für die Entwicklung der Geburtenhäufigkeit voraus: eine leicht fallende TFR auf 1.2 Kinder pro Frau im Jahr 2050, eine annähernd konstante TFR bei 1.4 Kinder pro Frau und eine leicht steigende ab 2025 auf 1.6 Kinder pro Frau. Die beiden äußeren Werte liegen damit leicht außerhalb des 95%-Intervalls, der im Jahr 2030 Grenzwerte 1.24 und 1.49 hat. Die konstante Annahme ist beinahe identisch

mit dem vordefinierten „endgültigen Niveau“, was 1.395 Kinder pro Frau beträgt.

Für die Wanderungen setzt das Statistische Bundesamt 2 Varianten voraus-ein Wanderungssaldo von 100 000 bzw. von 200 000 Personen jährlich. Damit liegt die im Kapitel 4 beschriebene Annahme von 150 000 Migranten jährlich zwischen den beiden Varianten des statistischen Bundesamtes.

Die 11. koordinierte Bevölkerungsvorausberechnung liefert Ergebnisse für insgesamt 12 Kombinationen aller Varianten bis zum Jahr 2050. Die Abbildung 5.1 zeigt die Werte der „mittleren Bevölkerung“ (Varianten 1-W1 und 1-W2), die nach der Basisannahme an die Lebenserwartung und mit der konstant bleibenden Geburtenhäufigkeit berechnet wurden. Die Untergrenze wird durch das niedrige Wanderungssaldo, die Obergrenze durch das höhere Wanderungssaldo bestimmt. Die Werte der oberen Schätzung liegen bis zum Jahr 2030 knapp unter dem Mittelwert der stochastischen Schätzung, ab dem Jahr 2040 sind sie oberhalb des Mittelwertes im Bereich des Konfidenzintervalls. Die unteren Werte liegen alle außerhalb des 95%-Konfidenzintervalls.

Einen Vergleich der Ergebnisse bietet auch die Abbildung 5.6, die die geschätzte Bevölkerungszahlen im Jahr 2030 (abgebildet an der  $x$ -Achse) und im Jahr 2050 (abgebildet an der  $y$ -Achse) zeigt. Die schwarzen Kreuze stellen die simulierten Werte der stochastischen Modellrechnung mit ihrem Mittelwert (grüner Punkt) und dem 95%-Konfidenzintervall (gestrichelte Linien) dar. Die Varianten „mittlere“ Bevölkerung (1-W1 und 1-W2) sind mit den roten Dreiecken gekennzeichnet, die blauen Vierecke demonstrieren die Variante der „relativ jungen“ Bevölkerung und der „relativ alten“ Bevölkerung (Basisannahme der Lebenserwartung, leicht steigende TFR und höherer Wanderungssaldo bzw. Annahme des hohen Anstiegs der Lebenserwartung, leicht fallende Geburtenhäufigkeit und niedrigerer Wanderungssaldo). Nur die Werte der Obergrenze der Variante der „mittleren“ Bevölkerung (1-W2) fallen in die beiden Konfidenzintervalle der Jahre 2030 und 2050. Die Werte der Untergrenze der „mittleren“ Bevölkerung (1-W1) und auch die Werte der Variante der „relativ alten“ Bevölkerung liegen sowie im Jahr 2030 als auch im Jahr 2050 unter dem 2.5%-Quantil der stochastischen

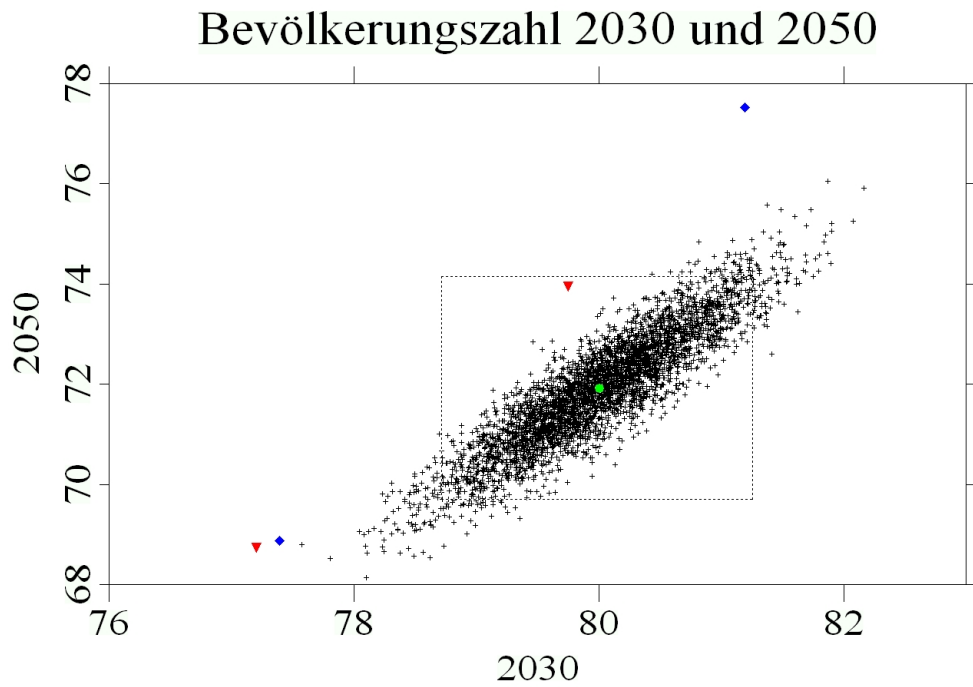


Abbildung 5.6: Geschätzte Bevölkerungszahlen in den Jahren 2030 und 2050 (schwarze Kreuze) mit dem Mittelwert (grüner Punkt) und 95%-Konfidenzintervalle (gestrichelte Linien). Als rotes Dreieck bzw. blaues Viereck sind die Varianten des Statistischen Bundesamtes dargestellt („mittlere“, „relativ junge“ und „relativ alte“ Bevölkerung).

Prognose. Die Variante der „relativ jungen“ Bevölkerung liegt im Jahr 2030 noch im 95%-Konfidenzintervall, im Jahr 2050 ist der Wert höher als der 97.5%-Quantil.

## 6 Auswirkung auf ein neues Rentensystem

Die Bevölkerungszahl in Deutschland nimmt ab und die Lebenserwartung steigt, dieser Trend wird sich auch in der nächsten Jahrzehnten fortsetzen, wie bereits die Ergebnisse der stochastischen Modellrechnung im Abschnitt 5.2 gezeigt haben. Diese Veränderung hat große Auswirkungen auf die Versorgung der Menschen im Rentenalter – denn es gibt immer weniger Personen im Erwerbsalter, die die Rentner versorgen können. Das heutige gesetzliche Rentensystem wird im Wege des Umlageverfahrens finanziert, d.h. dass die laufende Rentenzahlungen aus den aktuellen Beiträgen der Rentenversicherungsträger finanziert werden. Die Versicherten erhalten als Gegenleistung für ihre Beitragszahlung einen Anspruch auf Bezug einer Altersrente, die dann von der nächsten Beitragszahler-Generation finanziert wird<sup>1</sup>.

Wird die Anzahl der Beitragszahler kleiner und wächst damit sogar die Anzahl der Rentner, müssen entweder die Beiträge erhöht werden um das gleiche Rentenniveau zu halten oder bei gleich bleibendem Beitragssatz die Rentenleistungen gesenkt werden. Im Jahr 2006 lag die Höhe der durchschnittlichen gesetzlichen Rente bei 720 EUR im Monat, der Beitragssatz betrug dabei 19.5% des Bruttogehalts<sup>2</sup>.

Zusätzlich können die Defizite zwischen der Beitragseinnahmen und der Leistungen aus dem Bundessteueraufkommen finanziert werden. In Deutschland machten die staatlichen

---

<sup>1</sup>Siehe §153 I SGB VI und [www.deutsche-rentenversicherung-bund.de](http://www.deutsche-rentenversicherung-bund.de)

<sup>2</sup>Ab dem Jahr 2007 erfolgt eine Beitragssatzerhöhung auf 19.9%.

Zuschüsse im Jahr 2006 mit 55 Mrd. EUR etwa 24% der Einnahmen der gesetzlichen Rentenversicherung, siehe Deutsche Rentenversicherung Bund (2006a).

Wie bereits die Abbildung 5.5 zeigte, wird sich das Verhältnis der Rentner zu den Menschen im erwerbsfähigen Alter noch drastisch erhöhen. Um den zukünftigen Rentnern ein würdiges Alter zu gewähren und gleichzeitig die jüngeren Generationen nicht allzu belasten, muss ein neues Alterssicherungssystem konzipiert werden, welches eine Grundsicherung für alle Menschen im Ruhestand gewährleistet.

Ein Konzept eines neuen Rentensystems, das in dem Forschungsinstitut für Neue Alterssicherungssysteme und Rechtsbiometrik der Humboldt Universität in Berlin entwickelt wurde, wird in dem folgenden Abschnitt 6.1 vorgestellt. Der darauf folgende Abschnitt zeigt die geschätzte Belastung der einzelnen Personengruppen sowie der Staatskasse und vergleicht das neue System mit dem heutigen Umlagesystem der gesetzlichen Rentenversicherung.

## **6.1 Konzept des Rentenmodells mit einer Basisrente**

Nach dem hier vorgeschlagenen Konzept eines neuen Alterssicherungssystems soll allen Rentnern ab dem 67. Lebensjahr eine Grundsicherung in Höhe von 1000 EUR gewährleistet werden. Diese Alterssicherung wird zur Hälfte weiterhin aus der Umlage finanziert, die zweite Hälfte wird im Wege der Kapitaldeckung angespart. Jedoch werden alle Erwerbstätigen in beiden Teilen zwangsversichert: Auf der einen Seite werden sie zur Beitragszahlungen im Umlagesystem gezwungen, auf der anderen Seite werden sie, vergleichbar der Kfz-Haftpflichtversicherung<sup>3</sup>, zu einer kapitalgedeckten Altersvorsorge verpflichtet. In das Pflichtsystem werden alle erwerbstätigen Personen (inklusive z.B. Beamte und Selbstständige) einbezogen.

Zum Zwecke der Kapitalersparnis wird für jede Person von Geburt an ein Altersvorsorgekonto eingerichtet, auf welches nicht nur der Vorsorgende selbst, sondern auch seine

---

<sup>3</sup>Siehe §1 PflVG.

Familienmitglieder für ihn, z.B. solange er noch ein Kind ist und über kein eigenes Einkommen verfügt, Geld einzahlen können. Durch die längere Ansparphase kann dank des Zinseszins effekts mehr Kapital erzielt werden. Das Altersvorsorgekonto soll sowohl zum Aufbau als auch zur Entnahme von Kapital dienen (z.B. für Ausbildungs- oder Studienzwecke). Es sollte sich um ein einfaches, transparentes Produkt handeln, dass sich durch niedrige Abschluss- und Verwaltungskosten auszeichnet. Denkbar sind grundsätzlich Lebensversicherungsverträge (ohne Risikoabsicherung) und Banksparpläne.

Zur Einführung des neuen Rentensystems ist es erforderlich, eine Übergangslösung zu konzipieren, um den Vorsorgenden zu ermöglichen, ausreichend Kapital anzusparen und zugleich die Versorgung derjenigen zu gewährleisten, die noch aus dem alten System Leistungen erhalten. Der Umstieg auf das neue Alterssicherungssystem sollte möglichst zeitnah geschehen - um die heutige, noch relativ ausgeglichene demographische Situation zwischen arbeitender und Rentnergeneration zu nutzen. Die Übergangsphase wirkt sich jedoch auf mehrere Generationen aus.

Bei einer Umstellung werden die Beiträge in das Umlagesystem sofort abgesenkt, auf den Betrag, der erforderlich ist, das notwendige Niveau, einer Basisrente für alle Rentner von 500 EUR, zu gewährleisten. Die ersparten Beiträge kommen den Rentenversicherten zu Gute.

Da die heutige Rentnergeneration oder Menschen kurz vor dem Renteneintritt nicht mehr in der Lage sind den zweiten Teil der Rente als Kapital anzusparen, behalten sie den Anspruch auf ihre Altersrente in gleicher Höhe wie im derzeitigen System. Die Differenz zwischen den niedrigeren Beitragseinnahmen und den Rentenausgaben muss durch das Bundessteueraufkommen ausgeglichen werden. Dies wird zwar zunächst zu einer Steuererhöhung führen, es wird dadurch aber eine sozial gerechtere Ausfinanzierung der Rentnergeneration erzielt. Denn es werden alle Bürger über die Steuern und nicht nur die Beitragszahler an der Finanzierung der derzeitigen Rentnergeneration beteiligt.

Um die Staatskasse zu entlasten, werden die Versicherten verpflichtet, einen Teil ihrer zukünftigen Rente selber zu sparen. Mit späterem Geburtsjahr (d.h. für die jüngeren

Menschen) wird sich der angesparte Anteil folgendermaßen stufenweise erhöhen: Menschen, die heute zwischen 50 und 55 Jahren alt sind, bekommen aus dem Umlagesystem eine um 8% gekürzte Rente - dieser Betrag wird durch das angesparte Kapital ersetzt. Menschen zwischen 45 und 50 Jahren werden in Zukunft eine um 16% gekürzte Rente aus dem „Umlagepotf“ erhalten, dieser Teil ihrer Altersrente wird durch eigene Ersparnisse ersetzt. Menschen zwischen 40 und 45 Jahren können noch einen höheren Prozentsatz ihrer zukünftigen Rente durch Kapitalersparnisse ersetzen, etwa 24%. 40-jährige und spätere Geburtsjahrgänge können dank längerer Anspardauer bereits den ganzen zweiten Teil ihrer Rente aus dem eingesparten Kapital finanzieren<sup>4</sup>.

Zusammengefasst, werden nach der Umstellung auf das neue Rentensystem ab dem 1.1.2008 folgende Schritte erfolgen:

- Einbeziehung aller Erwerbspersonen in das Rentensystem.
- Beitragssenkung auf ein Niveau, das zur Finanzierung der 500 EUR Rente aus dem Umlagesystem ausreicht.
- Pflicht zur kapitalgedeckten Vorsorge.
- Steuererhöhung zur Defizitfinanzierung (in der Übergangszeit)
- Stufenweise Entlastung des Haushalts durch die Teilfinanzierung aus der Kapitaldeckung

Folgender Abschnitt präsentiert die Modellrechnung für die Finanzierung des neuen Rentenmodells und stellt die Belastung der einzelnen Altersgruppen in dem heutigen und in dem neuen System gegenüber.

---

<sup>4</sup>Unter der Voraussetzung eines Zinssatzes von 2% und der Abschlusskosten von 3% kann jeder Versicherte mit einem maximalen Betrag von 95 EUR monatlich das benötigte Kapital für den restlichen Teil seiner Rente auf dem heutigen Niveau ansparen.



## 6.2 Modellrechnung zur Finanzierung des Rentensystems

Für die Modellrechnung wird zur Vereinfachung angenommen, dass alle Beitragszahler und alle Leistungsempfänger identisch sind. D.h. dass alle Beitragszahler dasselbe Einkommen haben und alle Rentner dieselbe Rente beziehen. Dann gilt für den Beitragssatz  $s$  eine einfache Formel:

$$s_t = \frac{B \cdot R_t}{L_t \cdot V_t} \quad (6.1)$$

wobei  $B$  die Rentenhöhe,  $R_t$  die Zahl der Rentenempfänger,  $L_t$  den Arbeitslohn und  $V_t$  die Zahl der Beitragszahler in Periode  $t$  darstellt. Somit kann man den benötigten Beitragssatz für jedes Prognosejahr abschätzen.

Der Berechnung wird ein durchschnittliches Bruttojahresentgelt<sup>5</sup> von 30 000 EUR unterstellt, wobei für das Jahr das durchschnittliche Bruttojahresentgelt der gesetzlich Versicherten für Jahr 2007 29 488 EUR beträgt, siehe Deutsche Rentenversicherung Bund (2006a). Außerdem wird angenommen, dass 70% aller Personen im Erwerbsalter erwerbstätig sind, was dem Durchschnittswert der letzten Jahre nach den Statistiken des Statistischen Bundesamtes und der Deutschen Rentenversicherung Bund entspricht, siehe Stat.Bundesamt (2006b); Deutsche Rentenversicherung Bund (2006b).

Da bereits in der gesetzlichen Rentenversicherung mit einer Erhöhung des Renteneintrittsalter auf 67 Jahre gerechnet wird, ist diese auch in der Berechnung für das neue Rentensystem einbezogen – zur Vereinfachung wird sie linear aufgehoben, sodass ab dem Jahr 2012 alle Menschen mit 67 in Rente gehen. Außerdem wird eine Pauschale von 3% für die Verwaltungskosten gerechnet, die Verwaltungskosten der gesetzlichen Rentenversicherung liegen zwar bei 1.5%, allerdings soll sicherheitshalber mit einer Reserve für eventuelle sonstige Kosten gerechnet werden.

Die Abbildung 6.1 zeigt den geschätzten Mindestbeitragssatz mit dem 95%-Konfidenzintervall, der für die Finanzierung der 500 EUR-Rente benötigt wird. Der Beitragssatz

---

<sup>5</sup>Alle Werte auch in der weiteren Berechnungen sind inflationsbereinigt.

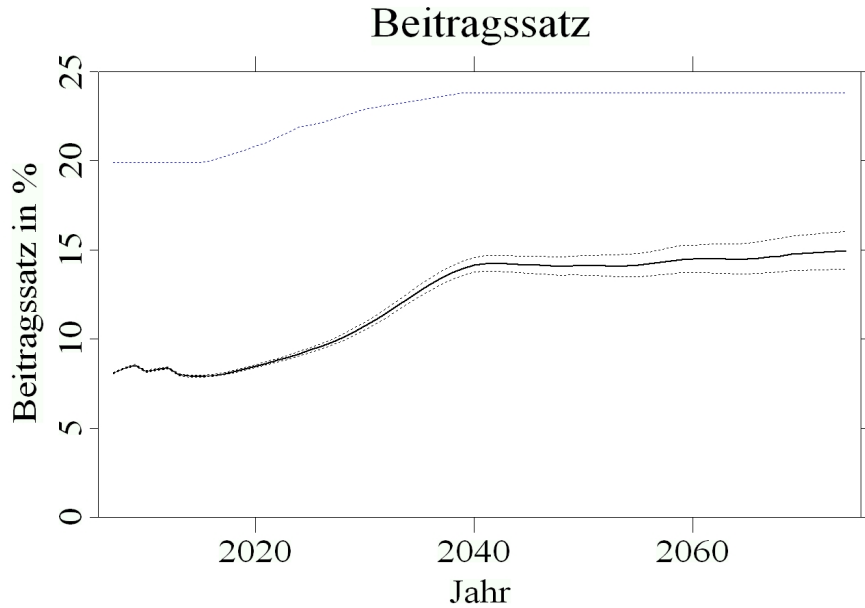


Abbildung 6.1: Der Mindestbeitragssatz in dem neuen Rentenmodell (schwarze durchgezogene Linie) mit dem 95%-Konfidenzintervall (schwarze gestrichelte Linien).

steigt analog zu dem Altenquotient (Abb. 5.5) von benötigten 8.5% im Jahr 2008 auf 14.5% im Jahr 2041. Er balanciert um den Wert 14.5% bis zum Ende der 60er Jahre und nähert sich der 15%-Grenze am Ende der 70er Jahre. Die blaue gestrichelte Linie zeigt die Höhe des Beitragssatzes in der gesetzlichen Rentenversicherung nach Rürup-Kommission (2003), der im Jahr 2007 bei 19.9% liegt. Er steigt auf 23.8% ab dem Jahr 2040. Der Histogramm in der Abbildung 6.2 zeigt die Verteilung des Beitragssatzes im Jahr 2070. Mit einer Wahrscheinlichkeit von 95% liegt er zwischen 13.9 und 15.8%.

Die Entwicklung des Beitragssatzes und damit auch der Kosten in dem umlagefinanzierten Teil des Rentenmodells hängt von dem Verhältnis der Rentenempfänger zu den Beitragszahlern ab. Die Höhe der Bundeszuschüsse lässt sich von der Zahl der noch nach heutigem Niveau beziehenden Rentenempfänger und der heutigen Rentenhöhe ableiten. Der durchschnittliche Rentenzahlbetrag für Altersrenten lag zum 31.12.2005 nach Angaben der Deutsche Rentenversicherung Bund (2006a) für Männer bei 991 EUR, für Frauen

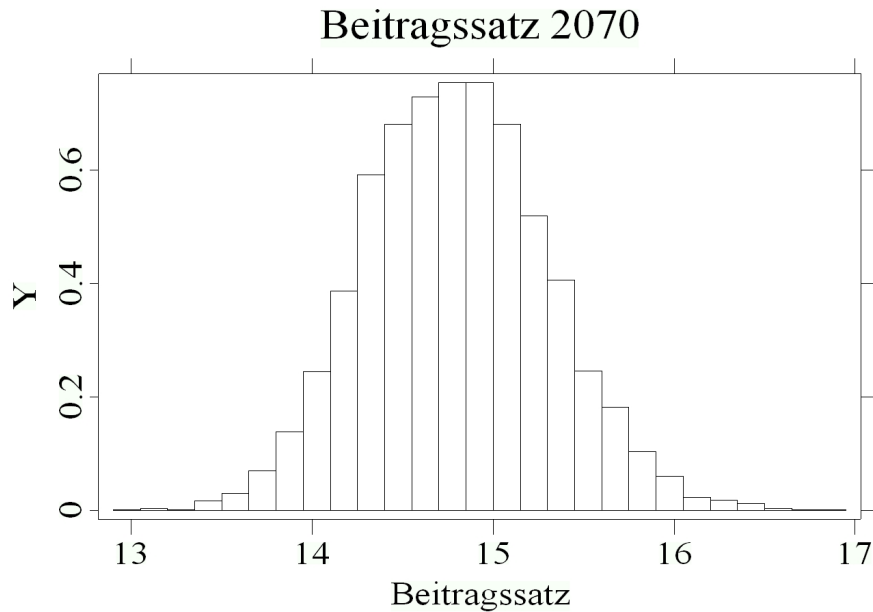


Abbildung 6.2: Der Histogramm für die variable Beitragssatz im Jahr 2070.

bei 509 EUR monatlich. Da zum 1.7.2007 eine Rentenerhöhung von 0.54% stattgefunden hat, ist auch mit einer Erhöhung des durchschnittlichen Rentenzahlbetrags zu rechnen. Im folgender Berechnung wurde mit einem durchschnittlichen Rentenbetrag von 1000 EUR für Männer bzw. von 515 EUR für Frauen kalkuliert.

Die Tabelle 6.1 gibt ein Überblick über die geschätzte Kosten für die umlagefinanzierten Renten (Spalte „Umlage“) und über die notwendigen Bundeszuschüssen (Spalte „Steuer“) an. Zum Vergleich werden auch die geschätzten Kosten für die Altersrenten nach dem heutigen Modell gezeigt, die aus der Kosten für Altersrenten im Jahr 2006 in der Höhe von 151 Mrd. EUR hergeleitet wurden<sup>6</sup>.

Eine grafische Darstellung der zukünftigen geschätzten Kosten bietet die Abbildung 6.3. Die blaue Linie zeigt die geschätzten Kosten des derzeitigen Modells, die schwarze Linie stellt die geschätzten Kosten des neuen Modells dar, die sich aus dem umlagefinanzierten (gestrichelte Linie) und aus den Bundeszuschüssen (rote Linie) zusammensetzen.

<sup>6</sup>Quelle: Forschungsportal der Deutschen Rentenversicherung mit unterjährlich aktualisierten Daten ([forschung.deutsche-rentenversicherung.de](http://forschung.deutsche-rentenversicherung.de))

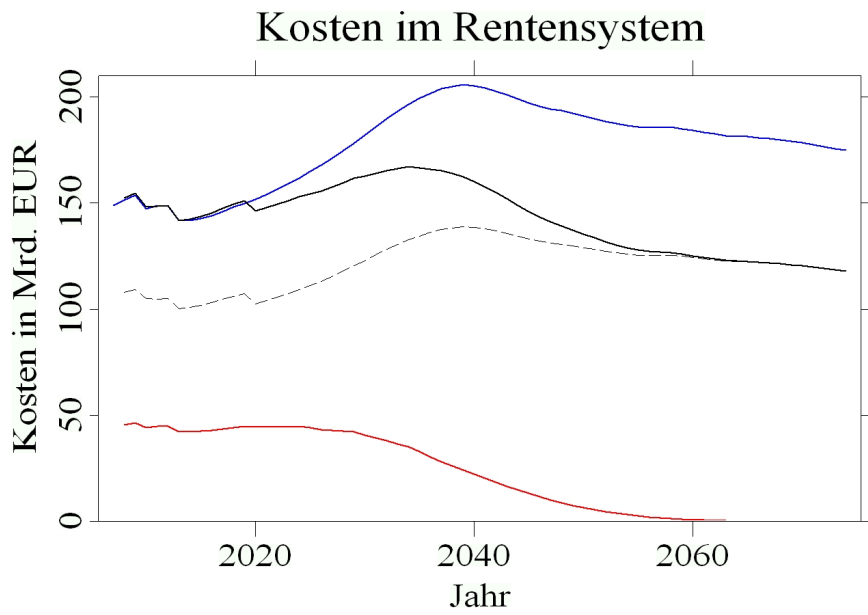


Abbildung 6.3: Die geschätzte Kostenentwicklung im derzeitigen (blaue Linie) und im neuen Rentenmodell (schwarze Linie), die sich aus dem Bundeszuschuss (rote Linie) und den Beitragserträgen (schwarze gestrichelte Linie) ergibt.

Jahr	Neues System			Heutiges System
	Umlage	Steuer	Gesamt	Ausgaben
2008	100.51	48.33	148.84	148.84
2010	102.10	49.13	151.23	151.08
2015	95.55	45.75	141.30	140.28
2020	101.09	46.55	148.65	149.41
2025	109.23	46.86	156.09	162.13
2030	120.30	43.30	163.60	179.05
2035	132.57	33.64	166.22	196.00
2040	138.81	22.60	161.41	201.92
2045	134.36	13.55	147.92	194.01
2050	129.69	6.71	136.40	187.91
2055	125.80	2.65	128.45	182.75
2060	124.84	0.77	125.61	181.10
2065	122.47	0.15	122.62	178.37
2070	120.89	0.02	120.91	175.58

Tabelle 6.1: Geschätzte Kosten beiden Rentensystemen in Mrd. EUR.

Aufgrund der Erhöhung des Renteneintrittsalters sinken bis zum Jahr 2015 kurzfristig die Rentenausgaben in beiden Modellen. Danach folgt ein rapider Anstieg der Kosten, der sich aber in dem neuen Modell durch den Einsatz der kapitalgedeckten Ersparnisse mildern lässt. Demgegenüber steigen die geschätzten Kosten des heutigen Systems bis auf 200 Mrd. EUR im Jahr 2040. Die Belastung der Staatskasse in dem neuen System sinkt ab dem Jahr 2020 und liegt ab 2060 unter 1 Mrd. EUR.

Rechnet man die geschätzten Kosten in beiden Systemen zusammen, bekommt man einen Vergleich des benötigten Gesamtkapitals, wie die Tabelle 6.2 zeigt. In dem neuen System wurden die Ausgaben aus dem „Umlagetopf“, die Bundeszuschüsse und das angesparte Kapital, aus welchem sich die kapitalgedeckte Rente von 500 EUR finanzieren lässt<sup>7</sup>, addiert. Für das heutige System wurden die geschätzten zukünftigen Kosten hochgerechnet, sodass der durchschnittliche Rentenbetrag bei 1 000 EUR monatlich liegt. Summiert man alle Kosten ab dem Jahr 2008, bekommt man bereits im Jahr 2020 einen

<sup>7</sup>Eine 50-jährige Person kann mit einem Betrag von 285 EUR monatlich ihre zukünftige Rente in Höhe von 500 EUR finanzieren. Eine 40-jährige Person braucht nur 145 EUR monatlich und eine 30-jährige Person nur 120 EUR monatlich zu sparen (gerechnet mit nominellen Zins von 2.5% p.a.).

Unterschied der Gesamtkosten von 200 Mrd. EUR zu Gunsten des neuen Systems. Bis zum Jahr 2060 ist das neue Rentensystem um 3 150 Mrd. EUR billiger, was ganze 25% der Kosten ab dem Jahr 2008 darstellt. Dieser Trend zeigt sich auch in den weiteren Jahren, bis zum Jahr 2070 wird das neue Rentensystem insgesamt um mehr als 4 000 Mrd. EUR und damit um 27% der Gesamtkosten entlastet.

Bis zum Jahr	Neues System	Heutiges System	Entlastung
2010	560	620	-60
2020	2 420	2 620	-200
2030	4 380	5 000	-620
2040	6 380	7 730	-1 350
2050	8 160	10 430	-2 270
2060	9 730	12 880	-3 150
2070	11 230	15 360	-4 130

Tabelle 6.2: Aufsummierte Kosten beiden Rentensystemen in Mrd. EUR.

Die Tabellen 6.3 und 6.4 vergleichen die Belastung einzelnen Personen zwischen 35 und 55 Jahren<sup>8</sup> im neuen (Tabelle 6.3) und im derzeitigen (Tabelle 6.4) Rentensystem. Die Spalte „Bezahlt“ gibt die Information über die bereits eingezahlten Beiträge in die gesetzliche Rentenversicherung, wobei der Berechnung die jeweiligen durchschnittlichen Bruttojahresentgelte und die jeweiligen Beitragssätze nach Angaben der Deutsche Rentenversicherung Bund (2006b) zugrunde gelegt wurden. In der Spalte „Neue Beiträge“ sind die geschätzten Beitragszahlungen in das Umlagesystem des neuen Rentenmodells eingegeben. Zusätzlich wird unter „Kapitaldeckung“ das benötigte Kapital berechnet, was man insgesamt aufwenden muss, um eine Rente in Höhe von 230 EUR bzw. 500 EUR (Spalten bezeichnet mit \*) kapitalgedeckt zu finanzieren<sup>9</sup>. Die Spalte „Gesamt“ gibt den Gesamtaufwand in EUR an.

In der Tabelle 6.4 wird analog die geschätzte Belastung für das derzeitige System angege-

<sup>8</sup>Alter im Jahr 2008

<sup>9</sup>Der Berechnung wurde ein konstantes Jahresentgelt, ein nomineller Zins von 2% bzw. 2.5% bei den langfristigen Sparverträgen und Abschlusskosten von 3% der Vertragssumme unterstellt.

ben, falls es in der Zukunft ohne Veränderungen bleibt. In der Spalte „Noch zu zahlen“ sind die geschätzten Beitragszahlungen in die gesetzliche Rentenversicherung bis zum Renteneintritt unter Annahme der Beitragsentwicklung nach der Annahmen der Rürup-Kommission (2003) geschildert.

Die Tabelle 6.5 vergleicht den Gesamtaufwand der Personen in dem neuen System mit der Rentenleistung von 730 bzw. 1 000 EUR und in dem heutigen System in Prozentzahl der Gesamteinkommen der jeweiligen Personen.

Alter	Bezahlt	Umlage	Kapital- deckung	Kapital- deckung*	Gesamt	Gesamt*
55	110.7	30.0	9.4	46.8	150.1	187.5
50	92.5	43.6	18.4	55.1	154.5	191.2
45	71.1	59.0	23.8	60.2	153.9	190.3
40	44.9	77.0	30.8	66.4	152.7	188.3
35	17.3	97.6	28.8	62.2	143.7	177.1

Tabelle 6.3: Geschätzte Belastung ausgewählter Personen im neuen Rentensystem mit einem Rentenbetrag von 730 EUR und 1000 EUR (mit \*) monatlich.

Alter	Noch zu		
	Bezahlt	zahlen	Gesamt
55	110.7	72.1	182.8
50	92.5	104.1	196.6
45	71.1	137.6	208.7
40	44.9	172.2	217.1
35	17.3	207.6	224.9

Tabelle 6.4: Geschätzte Belastung ausgewählter Personen im derzeitigen Rentensystem mit einem Rentenbetrag von 730 EUR.

Die Belastung aller Personen in dem neuen Rentensystem ist mit gleichbleibendem Rentenniveau deutlich niedriger. Selbst mit einer höheren Rentenleistung von insgesamt 1 000 EUR ist das neue System für die meisten Menschen günstiger: Eine 55-jährige Person würde für die gleiche Rentenleistung insgesamt um 30 000 EUR bzw. 3.5% des Gesamteinkommens mehr bezahlen, falls sich bei dem derzeitigen Rentensystem nichts ändert. Eine 45-jährige Person wird sich mit einem Aufwand von über 190 000 EUR in

Alter	Neues System mit 730	Neues System mit 1000	Heutiges System mit 730
55	16.0	20.0	19.5
50	15.6	19.3	19.8
45	15.0	18.6	20.4
40	14.7	18.1	20.9
35	13.7	16.9	21.5

Tabelle 6.5: Vergleich der Belastung ausgewählter Personen in Prozent des Gesamteinkommens.

dem neuen Rentensystem eine Gesamrente von 1 000 EUR leisten können, gegenüber einer Rente von 730 EUR im derzeitigen System mit vergleichbarem Aufwand. Bei der jüngeren Menschen wird der Unterschied dank der längeren Ansparphase für die kapitalgedeckte Vorsorge noch größer. Der Aufwand bei gleichem Rentenniveau bei einer 40-jährige Person wird um 60 000 EUR bzw. um 6% des Gesamteinkommens kleiner sein. Für eine 35-jährige Person unterscheiden sich die beiden Systeme mit gleicher Rentenleistung um 80 000 EUR bzw. 7.5 Prozentpunkte des gesamten Bruttoeinkommens. Das neue System ist auch hier mit einer höheren Rentenleistung um 45 000 bzw. 4.5 Prozentpunkte günstiger.

Damit zeigt die an dem nestor-Forschungsinstitut durchgeführte Modellrechnung, dass die Umstellung auf ein hälftig Umlage- und hälftig kapitalgedecktes Rentensystem:

- das Rentensystem bezahlbar macht
- den Generationenvertrag stabil hält
- die Leistungen für Frauen und Männer erhöht (1 000 EUR pro Monat)
- die Leistungen für Frauen und Männer vorhersehbar macht
- den Staat entlastet um den Bürger zu belasten.

Dieses neue System der Sicherung für alle Menschen, welches durch den Zinseszinsseffekt, die Einbeziehung aller Bürger, dank seiner Transparenz, Einfachheit und der Entlastung der „schwächeren“ Generationen billiger ist, ist für alle Generationen bezahlbar.



## 7 Fazit

In der vorliegenden Arbeit wurde die Bevölkerungsentwicklung in den nächsten 70 Jahren für Deutschland abgeschätzt. Hierzu wurden verschiedene stochastische und demographische Verfahren herangezogen, mit denen die demographischen Prozesse Mortalität, Fertilität und Migration analysiert wurden.

Die Modelrechnungen lieferten folgende ausgewählte Ergebnisse:

- Die Lebenserwartung bei Geburt für Frauen (bzw. Männer), abgeleitet aus den altersspezifischen Sterbeziffern, erhöht sich weiterhin von 81.8 (76.5) Jahren im Jahr 2005 bis auf 90.5 (84.3) Jahre im Jahr 2070, der 90%-Intervall hat eine Breite von 6.7 (9.6) Jahren.
- Die Bevölkerungszahl sinkt von 82.4 Mio. im Jahr 2004 auf 71.9 Mio. im Jahr 2050 bis auf 63.8 Mio. im Jahr 2070. Der 95%-Intervall weist eine Breite von 4.4 Mio. im Jahr 2050 bzw. 6.1 Mio. im Jahr 2070.
- Es erfolgt eine Veränderung in der Bevölkerungsstruktur: der Altenquotient mit Altersgrenze 65 steigt rapid von heutigen 30% auf mindestens 60% (2.5%-Quantil) und maximalen 69% (97.5%-Quantil) im Jahr 2070. Gleichzeitig sinkt gering der Jugendquotient von heutigen 33% auf 30% im Jahr 2070.
- Die Schätzwerte der Bevölkerungszahl sind zum Teil vergleichbar mit den Ergebnissen der 11. koordinierten Bevölkerungsvorausberechnung, wobei die Obergrenze der Variante „mittlere Bevölkerung“ am nächsten liegt, siehe Stat.Bundesamt (2006a).

Die demographischen Komponenten Mortalität, Fertilität und Migration erfordern eine entsprechende stochastische Modellierung, die jedoch demographisch plausible Ergebnisse liefert. In dieser Arbeit wurden folgende Ansätze verwendet: ein auf dem Lee-Carter-Modell basierender Ansatz zur stochastischen Modellierung der altersspezifischen Sterbeziffer, ein auf dem endgültigen Niveau und den oberen und unteren Einschränkungen beruhender Ansatz zur stochastischen Modellierung der altersspezifischen Geburtenziffern und eine konstante Annahme zur deterministischen Modellierung der Migration. Die Modelle für die demographischen Prozesse bilden eine Basis für die eigene Bevölkerungsvorausberechnung, die mit Hilfe der Kohorten-Komponenten-Methode durchgeführt wurde.

Der zweite Teil dieser Arbeit beschäftigt sich mit der Auswirkung der hier berechneten Bevölkerungsentwicklung auf das gesetzliche Rentensystem in Deutschland und schlägt ein neues, zur Hälfte umlagefinanziertes, zur Hälfte kapitalgedecktes Rentensystem mit einer gleichen Grundleistung für alle Rentner vor. Die hier durchgeführte Modellrechnung zeigt, dass das neue System für alle Altersgruppen und vor allem für die jungen Generationen günstiger ist als das derzeitige System, obwohl es eine höhere durchschnittliche Rentenleistung gewährleistet.

Alle in dieser Arbeit vorgenommenen Berechnungen basieren auf der demographischen Entwicklung bzw. Entwicklung der Renten in den letzten Jahrzehnten und bilden damit die Strukturen und erkennbare Veränderungen aus der Vergangenheit auf die Zukunft ab. Hiermit handelt es sich um eine möglichst genaue Schätzung der Variablen in den nächsten sieben Jahrzehnten, jedoch um keine Prognose, welche die Zukunft vorhersagt.

# Literaturverzeichnis

- Babel, B.: *Bevölkerungsvorausberechnungen unter besonderer Berücksichtigung von Unsicherheiten*. EUL Verlag, Lohmar - Köln, 2007.
- Börsch-Supan, A., Reil-Held, A. und Wilke, Ch.B.: *Der Nachhaltigkeitsfaktor und andere Formelmodifikationen zur langfristigen Stabilisierung des Beitragssatzes zur GRV*. MEA – Mannheimer Forschungsinsitut Ökonomie und Demografischer Wandel, 2003.
- Deutsche Rentenversicherung Bund, DRV: *Aktuelle Daten 2007*. Deutsche Rentenversicherung Bund, 2006a.
- Deutsche Rentenversicherung Bund, DRV: *Rentenversicherung in Zeitreihen*. Deutsche Rentenversicherung Bund, 2006b.
- Diekmann, A., Mueller, U. und Nauck, B.: *Handbuch der Demographie*. Springer, Heidelberg, 2000.
- Dinkel, R.H. und Luy, M.: Natur oder Verhalten? Ein Beitrag zur Erklärung der männlichen Übersterblichkeit durch einen Vergleich von Kloster- und Allgemeinbevölkerung. In: *Zeitschrift für Bevölkerungswissenschaft*, Band 24,2:S. 105–132, 1999.
- Hamilton, J.D.: *Time Series Analysis*. Princeton University Press, New Jersey, 1994.
- Härdle, W. und Simar, L.: *Applied Multivariate Statistical Analysis*. Springer Berlin, 2007.

- Hyndman, R.J. und Booth, H.: *Stochastic population forecasts using functional data models for mortality, fertility and migration*. ANU - The Australian National University, Working Papers in Demography, No.99, 2006.
- Hyndman, R.J. und Ullah, Md.S.: *Robust forecasting of mortality and fertility rates: a functional data approach*. MONASH University, Working Paper No 2/05, 2005.
- Lee, R.D.: Modeling and forecasting the time series of US fertility: Age distribution, range, and ultimate level. In: *International Journal of Forecasting*, Band 9:S. 187–202, 1993.
- Lee, R.D. und Carter, L.R.: Modeling and Forecasting U.S. Mortality. In: *Journal of the American Statistical Association*, Band 87:S. 659–671, 1992.
- Lee, R.D. und Tuljapurkar, S.: *Population Forecasting for Fiscal Planning: Issues and Innovations*. Center for the Economics and Demography of Aging, CEDA Papers: Paper 2000-0004CL, 2000.
- Lipps, O. und Betz, F.: Stochastische Bevölkerungsprojektion für West- und Ostdeutschland. In: *Zeitschrift für Bevölkerungswissenschaft*, Band 30,1:S. 3–42, 2005.
- Moral, P. und González, P.: *Univariate Time Series Modelling*. in XploRe Tutorials MD\*TECH Method and Data Technologies, 2003.
- Pflaumer, P.: Confidence intervals for population projections based on Monte Carlo methods. In: *International Journal of Forecasting*, Band 4:S. 135–142, 1988.
- Pflaumer, P.: Forecasting US population totals with the Box-Jenkins approach. In: *International Journal of Forecasting*, Band 8:S. 329–338, 1992.
- Rentenversicherungsbericht: *Bericht der Bundesregierung über die gesetzliche Rentenversicherung, insbesondere über die Entwicklung der Einnahmen und Ausgaben, der Nachhaltigkeitsrücklage sowie des jeweils erforderlichen Beitragssatzes in den künftigen 15 Kalenderjahren gemäß §154 SGB VI*. Bundesministerium für Arbeit und Soziales, 2005.

- Rürup-Kommission: *Abschlussbericht der Kommission für die Nachhaltigkeit in der Finanzierung der Sozialen Sicherungssysteme unter Vorsitz von B. Rürup*. Bundesministerium für Gesundheit und Soziale Sicherung, Berlin, 2003.
- Schmerbach, S.: *Wirtschafts- und Bevölkerungsstatistik I*. Vorlesungskript, HU Berlin, 2001.
- Stat.Bundesamt: *Bevölkerung Deutschlands bis 2050, Ergebnisse der 11. koordinierten Bevölkerungsvorausberechnung*. Statistisches Bundesamt, Wiesbaden, 2006a.
- Stat.Bundesamt: *Statistisches Jahrbuch 2006, Für die Bundesrepublik Deutschland*. Statistisches Bundesamt Wiesbaden, 2006b.
- Tuljapurkar, S.: *Population Forecasts, Fiscal Policy, and Risk*. The Levy Economics Institute of Bard College, Working Paper No.471, 2006.
- Wilmoth, J.R., Andreev, K., Jdanov, D., Gleijer, D.A., Boe, C., Bubenheim, M., Philipov, D., Shkolnikov, V. und Vachon, P.: *Methods Protocol for the Human Mortality Database*, version 5 Auflage, 2007.
- Wolfsdorf, K.: *Versicherungsmathematik, Teil 1 Personenversicherung*. B.G. Teubner Stuttgart, 1997.