



BEITRÄGE ZUR  
BEVÖLKERUNGSWISSENSCHAFT

Herausgegeben durch das  
Bundesinstitut für Bevölkerungsforschung (BiB)

*Band 45*

Christian Dudel

# Vorausberechnung von Verwandtschaft

Wie sich die gemeinsame Lebenszeit von Kindern,  
Eltern und Großeltern zukünftig entwickelt

Verlag Barbara Budrich  
Opladen • Berlin • Toronto 2014

Bibliografische Information der Deutschen Nationalbibliothek  
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;  
detaillierte bibliografische Daten sind im Internet über  
<http://dnb.d-nb.de> abrufbar.

Zgl. Dissertation an der Ruhr-Universität Bochum

Gedruckt auf säurefreiem und alterungsbeständigem Papier.

Alle Rechte vorbehalten.

© 2014 Verlag Barbara Budrich, Opladen, Berlin & Toronto  
[www.budrich-verlag.de](http://www.budrich-verlag.de)

ISBN 978-3-8474-0182-7 (Paperback)

**eISBN 978-3-8474-0287-9 (eBook)**

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Titelbildnachweis: Sybille Steinmetz, Bundesinstitut für Bevölkerungsforschung (BiB)

Umschlaggestaltung: Bettina Lehfeldt, Kleinmachnow – [www.lehfeldtgraphic.de](http://www.lehfeldtgraphic.de) und BiB

Schriftleitung: Dr. Martin Bujard, BiB

Satz: Sybille Steinmetz, BiB

## Vorwort

Der demografische Wandel gilt als Megathema in Gesellschaft und Politik. Neue Daten zur Geburtenrate oder zur Bevölkerungsentwicklung lösen regelmäßig ein breites mediales Echo aus. Die Bundesregierung hat Demografie zur Chefsache gemacht und in der letzten Legislaturperiode einen Demografiebericht vorgelegt und eine Demografiestrategie entworfen. Es gibt mittlerweile unzählige Presseartikel, Webseiten und populärwissenschaftliche Publikationen zum demografischen Wandel – nicht alle vermitteln ein exaktes Bild.

Ein derart zentrales Thema bedarf einer soliden wissenschaftlichen Verankerung. In der Wissenschaftslandschaft Deutschlands ist dieses zunehmend in den Fokus rückende Thema v.a. durch zwei Fachzeitschriften, *Comparative Population Studies* und *Demographic Research*, und durch die begutachtete Schriftenreihe „Beiträge zur Bevölkerungswissenschaft“ institutionalisiert. Mit dieser Reihe verfolgt das Bundesinstitut für Bevölkerungsforschung das Ziel, der Befundlage zum weiten Feld des demografischen Wandels eine wissenschaftliche Plattform für grundlegende Monografien und Sammelbände zu geben. Dafür wird diese seit 1975 erscheinende Reihe stetig weiterentwickelt. Ab dem hier vorliegenden Band 45 wird sie vom Verlag Barbara Budrich betreut. Alle Bände erscheinen ab jetzt auch als E-Book und werden fünf Jahre nach Erscheinen im Internet dauerhaft frei zugänglich gemacht.

Dieser Band widmet sich der Vorausberechnung von Verwandtschaft – eine bisher nur begrenzt erforschte Thematik mit wichtigen gesellschaftlichen und sozialpolitischen Implikationen. Der Autor Christian Dudel hat hierzu ein nichtparametrisch-stochastisches Verfahren entwickelt. Darauf basierend zeigt er mehrere Vorausberechnungen bis zum Jahr 2060 für die erwartete Zahl lebender Kinder und Enkel in bestimmten Altersgruppen sowie die Wahrscheinlichkeit lebender Mütter, Väter und Großeltern, jeweils differenziert nach Geschlecht und West- und Ostdeutschland. Mit diesen Befunden liegt nun eine Grundlage für mehrere weitere familiensoziologische und -ökonomische Fragestellungen vor.

Zwei Anwendungen sind speziell hervorzuheben: (1) Dudels Analysen ermöglichen Daten zur gemeinsamen Lebenszeit von Enkeln und Großeltern, die durch die steigende Lebenserwartung verstärkt und den Aufschub der Geburten reduziert wird. (2) Auch wird eine Abschätzung des Pflegepotenzials von Verwandten ermöglicht, das sich auf der individuellen Ebene weitaus anders verändert als die für die Pflegeversicherung relevante Verschiebung der Altersstruktur.

Wir danken Christian Dudel für die Erstellung des vorliegenden Buches und die Bereitschaft, sein Werk im Rahmen des Begutachtungsverfahrens zu überarbeiten. Ein großer Dank geht an Sybille Steinmetz für eine tatkräftige Betreuung des Layouts. Wir wünschen allen Leserinnen und Lesern eine anregende Lektüre.

Wiesbaden, Januar 2014  
Bundesinstitut für Bevölkerungsforschung



## Danksagung

Der vorliegende Text ist eine überarbeitete Fassung meiner Doktorarbeit, die an der Fakultät für Sozialwissenschaft der Ruhr-Universität Bochum als Dissertation angenommen wurde.

Meinem Doktorvater Prof. Dr. Götz Rohwer, an dessen Lehrstuhl die Arbeit geschrieben wurde, bin ich für das entgegengebrachte Vertrauen, hilfreiche Anmerkungen und die Möglichkeit, diese Arbeit zu erstellen, zu großem Dank verpflichtet. Für weitere hilfreiche Anmerkungen und Ratschläge danke ich der Zweitgutachterin Prof. Dr. Notburga Ott.

Die Anregung zum Thema dieser Arbeit verdanke ich Prof. Dr. Jörg Althammer. Ebenso zu Dank verpflichtet bin ich Prof. Dr. Martin Werding, Prof. Dr. Cornelia Weins und Prof. Dr. Jörg-Peter Schräpler, von denen ebenfalls etliche Anregungen kamen.

Für zahlreiche Gespräche und Unterstützung danke ich meinen Kolleginnen und Kollegen (in alphabetischer Reihenfolge) Hanna Beneker, Dr. Thilo Ernst, Jan Marvin Garbuszus, Mine Hancioglu, Bastian Hartmann, Katharina Knüttel, Dr. Ulrich Pötter und Dr. Karolina Simonic. Besonderer Dank gilt Sebastian Jeworutzki und Katharina Lerch für ihr sorgfältiges Korrekturlesen.

Dem Bundesinstitut für Bevölkerungsforschung möchte ich meinen Dank aussprechen für die Möglichkeit, meine Dissertation im Rahmen der Schriftenreihe „Beiträge zur Bevölkerungswissenschaft“ zu veröffentlichen. Insbesondere danke ich Sybille Steinmetz für ihre Hilfe beim Textsatz, Dr. Evelyn Grünheid für ihr sorgfältiges Korrekturlesen und Dr. Martin Bujard für hilfreiche Anmerkungen und Ratschläge.

Christian Dudel



# Inhaltsverzeichnis

Tabellenverzeichnis	11
Abbildungsverzeichnis	13
Symbolverzeichnis	15
1 Einleitung	17
2 Verwandtschaft und Anwendungsbeispiele	21
2.1 Definition und Abgrenzung des Begriffs Verwandtschaft	21
2.2 Zwei Beispiele für die Bedeutung von Verwandtschaft	25
3 Demographische Grundlagen	31
3.1 Demographische Grundbegriffe	31
3.2 Demographische Entwicklung in Deutschland	39
3.3 Demographischer Wandel und Verwandtschaft in der wissenschaftlichen Diskussion	46
4 Ansätze zur Vorausberechnung von Verwandtschaft	49
4.1 Allgemeine Probleme demographischer Vorausberechnungen	49
4.2 Varianten der Vorausberechnung von Verwandtschaft	53
4.3 Verzweigungsprozesse	54
4.4 Bevölkerungsmathematische Ansätze	61
4.5 Multi-State Ansätze	71
4.6 Mikrosimulationen	81
4.7 Zwischenfazit	92
4.8 Eingrenzung der Fragestellung und Ausblick	93
5 Ein Ansatz zur Vorausberechnung linearer Verwandtschaft	95
5.1 Probabilistische Vorausberechnungen	95
5.2 Diskrete Multityp-Galton-Watson-Prozesse	100
5.3 Lineare Verwandtschaft im diskreten Multityp-Galton-Watson-Prozess	109
5.4 Zusammenfassung und Diskussion	125
6 Fertilität von Männern	127
6.1 $\phi(y x)$ : Datengrundlage und Berechnung	127
6.2 Geburtenraten für westdeutsche Männer ab 1951	133
6.3 Vergleiche mit Ergebnissen aus anderen Quellen	142
7 Validierung des Ansatzes	145
7.1 Vorbemerkungen	145
7.2 Das Sozio-oekonomische Panel	146
7.3 Vergleich der Ergebnisse der Modellrechnung mit Daten des SOEP	152

7.4	Zusammenfassung und Diskussion . . . . .	162
8	Vorausberechnung der demographischen Komponenten und der Bevölkerung	163
8.1	Vorbemerkungen . . . . .	163
8.2	Nichtparametrisch-stochastische Vorausberechnung demographischer Komponenten . . . . .	163
8.3	Nichtparametrische Vorausberechnung von Sterberaten . . . . .	167
8.4	Nichtparametrische Vorausberechnung von Geburtenraten . . . . .	171
8.5	Nichtparametrische Vorausberechnung von Ein- und Auswanderung . . . . .	176
8.6	Fortschreibung der Bevölkerung . . . . .	180
8.7	Zusammenfassung . . . . .	187
9	Ausgewählte Ergebnisse der Vorausberechnung linearer Verwandtschaft	189
9.1	Vorbemerkungen . . . . .	189
9.2	Ergebnisse zur Zahl der Kinder . . . . .	190
9.3	Ergebnisse zur Zahl der Enkelkinder . . . . .	195
9.4	Ergebnisse zur Zahl lebender Eltern . . . . .	201
9.5	Ergebnisse zur Zahl lebender Großeltern . . . . .	206
9.6	Sensitivitätsanalysen . . . . .	210
9.7	Zusammenfassung . . . . .	216
10	Zwei Anwendungsbeispiele	219
10.1	Mehr gemeinsame Lebenszeit von Generationen? . . . . .	219
10.2	Pflege und Verwandtschaft: Entwicklung verwandtschaftlicher Unterstützungspotenziale am Beispiel von Kindern . . . . .	223
10.3	Bedingungen für die Gültigkeit der Ergebnisse . . . . .	228
11	Zusammenfassung, Diskussion und Ausblick	231
11.1	Ziel und Motivation . . . . .	231
11.2	Methodisches Vorgehen . . . . .	231
11.3	Zentrale Ergebnisse . . . . .	232
11.4	Diskussion . . . . .	236
11.5	Ausblick . . . . .	238
	Literaturverzeichnis	241

Der Anhang ist online zugänglich auf der Seite zum Buch beim Budrich Verlag: <http://www.budrich-verlag.de>. DOI: 10.3224/84740183A

## Tabellenverzeichnis

2.1	Verwandschaftsschema nach Atkins . . . . .	23
4.1	Ergebnisse des Ansatzes nach Goodman et al. (1974) sowie Ergebnisse für durchschnittliche Anzahl Verwandter im SOEP . . . . .	69
4.2	Zusammengefasste Ergebnisse für den Ansatz nach Goodman et al. (1974) für 2006 und 2050 . . . . .	70
4.3	Anzahl der Geschwister von Kindern, die mit einer (Stief-)Mutter zusammen im Haushalt leben . . . . .	78
6.1	Fallzahlen der Statistik der Geburten . . . . .	128
7.1	Wahrscheinlichkeiten für lebende Großeltern mütterlicherseits 2009 . . . . .	161
7.2	Wahrscheinlichkeiten für lebende Großeltern väterlicherseits 2009 . . . . .	161
8.1	Ergebnisse für die stochastische Vorausberechnung der Mortalität . . . . .	169
8.2	Ergebnisse für die stochastische Vorausberechnung der Fertilität . . . . .	174
8.3	Ausgewählte Ergebnisse für die stochastische Vorausberechnung der Migration . . . . .	179
8.4	Ergebnisse für die stochastische Bevölkerungsvorausberechnung . . . . .	183
8.5	Ergebnisse zur Bevölkerungsgröße und zum Altenquotienten im Jahr 2050 aus der Literatur . . . . .	186
9.1	Übersicht der Varianten der Sensitivitätsanalyse . . . . .	212
9.2	Ergebnisse zu Kindern und Enkelkindern westdeutscher Frauen 2060 bei den sechs Varianten der Sensitivitätsanalyse . . . . .	213
9.3	Ergebnisse zur Mutter und Großmutter mütterlicherseits in Westdeutschland für 2060 bei den sechs Varianten der Sensitivitätsanalyse . . . . .	215
10.1	Geteilte Lebenszeit mit Großeltern von Altersklasse 0 bis Altersklasse 20 . . . . .	222
11.1	Zentrale Ergebnisse der stochastischen Bevölkerungsvorausberechnung . . . . .	233



## Abbildungsverzeichnis

3.1	Bevölkerungsentwicklung in der BRD . . . . .	40
3.2	Entwicklung der zusammengefassten Geburtenziffer in West- und Ostdeutschland . . . . .	40
3.3	Entwicklung der Lebenserwartung von Männern und Frauen in West- und Ostdeutschland . . . . .	44
3.4	Einwanderung, Auswanderung und Nettowanderung in der BRD . . . . .	44
3.5	Entwicklung der Zahl der Eheschließungen und -scheidungen in Deutschland . . . . .	46
5.1	Vereinfachte Darstellung eines Multityp-Galton-Watson-Prozesses . . . . .	103
6.1	Dissimilaritätsindex für $\phi(y x)$ , berechnet zwischen 2008 und 2007, 2008 und 2006 sowie 2007 und 2006 . . . . .	130
6.2	Dissimilaritätsindex für $\phi(y x)$ , berechnet zwischen 2006/2008 und 1995 sowie 2000 . . . . .	130
6.3	$\phi(y 25)$ für 1995, 2000, 2006, 2007 und 2008 . . . . .	132
6.4	$\phi(y 35)$ für 1995, 2000, 2006, 2007 und 2008 . . . . .	132
6.5	Verteilungen $\phi(y x)$ für $x = 15, 20, 25, 30, 35, 40, 45$ . . . . .	134
6.6	Verteilungen $\hat{\phi}(y x)$ für $x = 15, 20, 25, 30, 35, 40, 45$ . . . . .	134
6.7	Zusammengefasste Geburtenziffer für Männer und Frauen von 1951 bis 2008 . . . . .	135
6.8	Ausgewählte altersspezifische Geburtenraten für Männer von 1951 bis 2008 . . . . .	137
6.9	Ausgewählte altersspezifische Geburtenraten für Frauen von 1951 bis 2008 . . . . .	137
6.10	Altersspezifische Geburtenraten für Männer für die Jahre 1951, 1960, 1970, 1980 und 1990 . . . . .	138
6.11	Altersspezifische Geburtenraten für Männer und Frauen von 2003 bis 2008 . . . . .	138
6.12	Altersstruktur der Bevölkerung in der reproduktiven Phase 1951 . . . . .	140
6.13	Verhältniszahlen für Fertilität und Bestand von Männern und Frauen von 1951 bis 2008 . . . . .	140
6.14	Verhältnisse von zusammengefassten Geburtenziffern von Männern und Frauen für Westdeutschland, England und Wales, Dänemark, Frankreich und Norwegen . . . . .	143
7.1	Durchschnittliche Zahl jemals geborener Kinder von Frauen 2009 . . . . .	156
7.2	Durchschnittliche Zahl jemals geborener Kinder von Männern 2009 . . . . .	156
7.3	Durchschnittliche Zahl lebender Enkel von Frauen 2006 und 2009 . . . . .	157
7.4	Durchschnittliche Zahl lebender Enkel von Männern 2006 und 2009 . . . . .	157
7.5	Wahrscheinlichkeit für eine lebende Mutter 2009 . . . . .	159
7.6	Wahrscheinlichkeit für einen lebenden Vater 2009 . . . . .	159
8.1	Ziehungswahrscheinlichkeiten bei 49 Differenzen und unterschiedlichen Werten für $a$ . . . . .	167

8.2	Entwicklung des Bevölkerungsbestandes bis 2060 . . . . .	184
8.3	Entwicklung des Altenquotienten bis 2060 . . . . .	184
9.1	Erwartete Anzahl lebender Kinder westdeutscher Frauen 2060 . . . . .	191
9.2	Erwartete Anzahl lebender Kinder westdeutscher Männer 2060 . . . . .	191
9.3	Erwartete Anzahl lebender Kinder ostdeutscher Frauen 2060 . . . . .	194
9.4	Erwartete Anzahl lebender Kinder ostdeutscher Männer 2060 . . . . .	194
9.5	Erwartete Anzahl lebender Enkel westdeutscher Frauen 2060 . . . . .	196
9.6	Erwartete Anzahl lebender Enkel westdeutscher Männer 2060 . . . . .	196
9.7	Erwartete Anzahl lebender Enkel ostdeutscher Frauen 2060 . . . . .	199
9.8	Erwartete Anzahl lebender Enkel ostdeutscher Männer 2060 . . . . .	199
9.9	Wahrscheinlichkeit einer lebenden Mutter 2060, Westdeutschland . . . . .	202
9.10	Wahrscheinlichkeit einer lebenden Mutter, eines lebenden Vaters und für keine lebenden Eltern in Altersklasse 50 2009-2060, Westdeutschland . . . . .	202
9.11	Wahrscheinlichkeit einer lebenden Mutter, eines lebenden Vaters und für keine lebenden Eltern in Altersklasse 50 2009-2060, Ostdeutschland . . . . .	205
9.12	Wahrscheinlichkeiten für lebende Großeltern mütterlicherseits 2009 bis 2060, Westdeutschland . . . . .	207
9.13	Wahrscheinlichkeiten für lebende Großeltern väterlicherseits 2009 bis 2060, Westdeutschland . . . . .	207
9.14	Wahrscheinlichkeiten für lebende Großeltern mütterlicherseits 2009 bis 2060, Ostdeutschland . . . . .	209
9.15	Wahrscheinlichkeiten für lebende Großeltern väterlicherseits 2009 bis 2060, Ostdeutschland . . . . .	209
10.1	Erwartete Anzahl lebender Kinder westdeutscher Frauen in Altersklasse 50, 2009 bis 2060 . . . . .	220
10.2	Verhältnis der Zahl der Pflegebedürftigen zur Zahl der potenziellen Pflegepersonen bis 2060 . . . . .	225
10.3	Median der erwarteten Kinderzahl westdeutscher Frauen in den Altersklassen 70, 75, 80, 85 und 90 von 2009 bis 2060 . . . . .	226
10.4	Median der erwarteten Kinderzahl ostdeutscher Frauen in den Altersklassen 70, 75, 80, 85 und 90 von 2009 bis 2060 . . . . .	226
11.1	Erwartete Anzahl lebender Kinder westdeutscher Frauen 2060 . . . . .	235
11.2	Wahrscheinlichkeiten für lebende Großeltern mütterlicherseits 2009 bis 2060, Ostdeutschland . . . . .	235

## Symbolverzeichnis

Symbol(e)	Bedeutung
$\mathbf{A}$	Leslie-Matrix
$B_t$	Zahl der Geburten im Jahr $t$
$d_x$	Zahl der Sterbefälle (Sterbetafel)
$D_t$	Zahl der Sterbefälle im Jahr $t$
$E(\cdot)$	Erwartungswert
$e_0$	Durchschnittliche fernere Lebenserwartung bei der Geburt
$e_x$	Durchschnittliche fernere Lebenserwartung im exakten Alter $x$
$e_i$	Vektor mit $e_i = 1$ und $e_j = 0$ für $j \neq i$
$f_x$	Altersspezifische Geburtenrate für Alter $x$
$F_x$	Geburtenrate für Altersklasse $x$
$F_x^*$	Geburtenrate für Altersklasse $x$ (Überleben vorausgesetzt)
$J_x$	Personenjahre im Alter $x$
$l_0$	Radix einer Sterbetafel
$l_x$	Größe einer Sterbetafelkohorte
$M_t^i, M_t^o$	Ein- bzw. Auswanderung im Jahr $t$
$m_x$	Nettomigration für Altersklasse $x$
$\mathbf{m}_t$	Vektor der Nettomigration
$\mathbb{N}$	Menge der natürlichen Zahlen (einschließlich 0)
$n_t, n_{x,t}$	Bevölkerungsbestand zu $t$ (in Altersklasse $x$ )
$\mathbf{n}_t$	Bevölkerungsvektor
$p_x$	Überlebenswahrscheinlichkeit (Sterbetafel)
$P_x$	Überlebenswahrscheinlichkeit für Altersklasse $x$
$P_x^*$	Wahrscheinlichkeit, von der Geburt bis Altersklasse $x$ zu überleben
$\text{Pr}(\cdot)$	Wahrscheinlichkeit
$q_x$	Sterbewahrscheinlichkeit (Sterbetafel)
$r_x$	Sterberate (Sterbetafel)
TFR	Zusammengefasste Geburtenziffer
$W(x t)$	Anteil von Eltern Neugeborener im Alter $x$ an allen Eltern Neugeborener zu $t$
$\bar{x}$	Durchschnittliches Alter bei der Geburt im Jahr $t$
$\mathbb{Z}$	Menge der ganzen Zahlen
$\alpha$	Untere Altersgrenze der reproduktiven Phase
$\beta$	Obere Altersgrenze der reproduktiven Phase
$\rho$	Höchstes erreichbares Alter
$\phi(\cdot  \cdot)$	Bedingte Verteilung der Alter von Vätern in Abhängigkeit vom Alter der Mutter
$\#\{\cdot\}$	Anzahl der Elemente einer Menge
$ \mathbf{x} $	Summe der Einträge eines Vektors



# 1 Einleitung

Die unter dem Begriff des „demographischen Wandels“ subsumierten Tendenzen der Bevölkerungsentwicklung in der BRD nehmen in der wissenschaftlichen und öffentlichen Diskussion breiten Raum ein.<sup>1</sup> Als kennzeichnend für die aktuelle Entwicklung werden niedrige Geburtenraten und eine hohe Lebenserwartung gesehen, die zum einem zu einer Alterung der Bevölkerung und zum anderen zu einem Bevölkerungsrückgang führen (vgl. Börsch-Supan 1991; Mai 2005; Bundesministerium des Innern 2011), deren mögliche Konsequenzen auf etliche gesellschaftliche Bereiche thematisiert werden. Dabei liegt ein Schwerpunkt der Betrachtung oftmals auf der zukünftigen Entwicklung, so dass neben der offiziellen Bevölkerungsvorausberechnung des Statistischen Bundesamtes noch eine Vielzahl weiterer Vorausberechnungen zu diversen Themen existiert. Als Beispiele für solche Themen können der Arbeitsmarkt und der Bereich der sozialen Sicherung genannt werden.<sup>2</sup>

Ein weiteres Thema ist der Einfluss des demographischen Wandels auf Familie und Verwandtschaft. So findet sich in der Literatur beispielsweise häufig die Überlegung, dass ein Anstieg der Lebenserwartung dazu führt, dass die gemeinsame Lebenszeit von Enkeln und Großeltern steigt, wovon ausgehend die Bedeutsamkeit der Beziehung zwischen Enkeln und Großeltern zunehmen wird (vgl. insb. Bengtson 2001; Bengtson & Martin 2001).

Dabei sind bei einer Veränderung verwandtschaftlicher Strukturen auch Konsequenzen für andere gesellschaftliche Bereiche zu erwarten, da „[...] in Deutschland familiäre und verwandtschaftliche Netzwerke in Korrespondenz mit der Absicherung wichtiger Risiken durch die Sozialversicherungssysteme für das Funktionieren des gesellschaftlichen Zusammenhangs eine zentrale Bedeutung haben“ (Rosenbaum & Timm 2008: 33). Als konkretes Beispiel lässt sich anführen, dass die Versorgung von pflegebedürftigen Menschen momentan in nicht unerheblichem Umfang von Verwandten gewährleistet wird, so dass sich eine Veränderung des Vorhandenseins von Verwandten direkt auf den Pflegebereich auswirken sollte (vgl. Galler 1990: 63). Dieses Beispiel ist insofern von besonderer Bedeutung, da für die Zukunft von einer steigenden Zahl pflegebedürftiger Menschen ausgegangen wird (vgl. bspw. Blinkert & Klie 2004; Bomsdorf et al. 2008; Statistisches Bundesamt 2008).

Im Gegensatz zu anderen Themenbereichen liegen allerdings keine aktuellen beziehungsweise nur eingeschränkte Vorausberechnungen zur Abschätzung des Einflusses des demographischen Wandels auf Familie und Verwandtschaft vor. Aus diesem Grund sind Überlegungen zur zukünftigen Entwicklung von Verwandtschaft und zu daraus resultierenden Konsequenzen auf Spekulationen angewiesen. Erschwerend kommt hinzu, dass der Einfluss der demographischen Entwicklung auf Verwandtschaft nicht unbedingt so eindeutig ist, wie weiter oben anhand des Beispiels von Enkeln und Großeltern angedeutet. Beispielsweise kann ein Anstieg des durchschnittlichen Alters von Eltern bei der Geburt eines Kindes einen verkürzenden

---

1 Für die wissenschaftliche Diskussion s. bspw. die Zusammenstellung von Arbeiten zum demographischen Wandel in der BRD von Kolle (2006).

2 Für den Bereich des Arbeitsmarktes s. bspw. Börsch-Supan & Wilke (2007), Fuchs et al. (2011) und Schulz (2008). Mögliche Konsequenzen und Entwicklungen im Bereich der sozialen Sicherung werden unter anderem in den Beiträgen in Schmähl & Ulrich (2001) diskutiert. Konkrete Beispiele für Vorausberechnungen aus dem Gesundheits- und Pflegebereich sind Blinkert & Klie (2004), Breyer & Felder (2004), Bomsdorf et al. (2008) und Statistisches Bundesamt (2008). S. auch Abschnitt 10.2.

Effekt auf die gemeinsame Lebenszeit von Kindern und (Groß-)Eltern aufweisen, während eine Zunahme der Lebenserwartung zu einem Anstieg dieser gemeinsamen Lebenszeit führen kann. Der Einfluss dieser entgegengesetzten Effekte auf die gemeinsame Lebenszeit insgesamt ist dabei nicht ohne weiteres bestimmbar (vgl. bspw. Uhlenberg 2005, 2006).

Hiervon ausgehend ergibt sich das Ziel der vorliegenden Arbeit: die Vorausberechnung von Verwandtschaft. Zur beispielhaften Klärung dieses Ziels sei wieder auf Großeltern zurückgegriffen. Ausgehend von diesem Verwandtschaftsgrad könnte man beispielsweise fragen, wie hoch die Wahrscheinlichkeit ist, dass im Jahr 2050 alle vier Großeltern einer 30-jährigen Person noch leben und in welchem Alter sie jeweils sind. Über eine Vorausberechnung von Verwandtschaft können diese und ähnliche Fragen beantwortet werden.

Ergebnisse einer solchen Vorausberechnung bieten etliche Anwendungsmöglichkeiten. Unter anderem kann untersucht werden, wie sich die geteilte Lebenszeit von Generationen in Zukunft entwickeln wird und es kann konkret abgeschätzt werden, welcher der bereits genannten Effekte überwiegt: der Effekt eines Anstiegs der Lebenserwartung oder der Effekt von Veränderungen im durchschnittlichen Alter bei der Geburt. Entsprechende Ergebnisse sind nicht nur in Hinblick auf die in der Literatur zu findenden Spekulationen von Interesse, sondern auch bezogen auf die Bedeutung intergenerationaler Beziehungen.

Eine weitere Anwendungsmöglichkeit ist die Betrachtung der Entwicklung des verwandtschaftlichen Unterstützungspotentials, das pflegebedürftigen Menschen zur Verfügung steht. So kann beispielsweise betrachtet werden, wie sich die Zahl an Kindern entwickeln wird, die ältere Pflegebedürftige potentiell betreuen können. Dies erlaubt eine Einschätzung darüber, ob Pflege in Zukunft weiterhin zu einem großen Teil durch Angehörige übernommen werden kann, oder ob verwandtschaftliche Unterstützungspotentiale abnehmen und von einer steigenden Nachfrage nach professioneller Pflege auszugehen ist.

Während in den vorausgegangenen Absätzen der Begriff Verwandtschaft teils zusammen mit dem Begriff Familie verwendet wurde, wird im weiteren nur von der „Vorausberechnung von Verwandtschaft“ die Rede sein und eine Trennung zwischen den beiden Begriffen vorgenommen, die ihrer Bedeutung in der wissenschaftlichen Diskussion entspricht. Inwiefern sich diese beiden Begriffe unterscheiden und was im Detail unter „Vorausberechnung von Verwandtschaft“ verstanden werden kann, wird im weiteren Verlauf der Arbeit geklärt.

Der Ablauf der Arbeit gliedert sich wie folgt: Die Kapitel 2 bis 4 behandeln theoretische und empirische Grundlagen. In Kapitel 2 wird eine Definition des Begriffs Verwandtschaft gegeben, diese gegen den Begriff Familie abgegrenzt sowie die gesellschaftliche Bedeutung von Verwandtschaft anhand der bereits angedeuteten Beispiele (intergenerationale Beziehungen; Pflege) erläutert. Eine Klärung demographischer Grundbegriffe und eine Beschreibung der demographischen Entwicklung in der BRD findet sich in Kapitel 3, welches mit einer kurzen Diskussion des Einflusses des demographischen Wandels auf Verwandtschaft schließt. Kapitel 4 beginnt mit einigen Ausführungen zu demographischen Vorausberechnungen im Allgemeinen und zur Vorausberechnung von Verwandtschaft im Speziellen. Ferner werden Verfahren beschrieben und diskutiert, die bisher in der Literatur zur Vorausberechnung von Verwandtschaft verwendet wurden.

Die Kapitel 5 bis 8 bilden den methodischen und empirischen Schwerpunkt der Arbeit. Das fünfte Kapitel ist theoretisch-methodisch orientiert und enthält eine Beschreibung und Diskussion des in dieser Arbeit verwendeten Ansatzes zur Vorausberechnung von Verwandt-

schaft. Zur Umsetzung dieses Ansatzes sind Informationen notwendig, die durch eine eigene Auswertung der Statistik der Geburten ermittelt wurden. Die Daten der Statistik der Geburten, deren Auswertung sowie abgeleitete Ergebnisse sind in Kapitel 6 beschrieben. In Kapitel 7 findet sich eine Validierung des verwendeten Vorausberechnungsansatzes, bei der überprüft wird, als wie realistisch Ergebnisse angesehen werden können, die mit diesem erzielt wurden. Schließlich wird in Kapitel 8 diskutiert, wie die zukünftige Entwicklung der Sterblichkeit, der Geburten sowie der Ein- und Auswanderung modelliert werden kann und es werden entsprechende Ergebnisse präsentiert.

Die eigentlichen Ergebnisse der Arbeit finden sich in den Kapiteln 9 und 10. Die Ergebnisse der Vorausberechnung von Verwandtschaft werden in Kapitel 9 beschrieben. Ergänzende Ergebnisse zur geteilten Lebenszeit von Generationen und zu verwandtschaftlichen Unterstützungspotentialen werden in Kapitel 10 vorgestellt.

Kapitel 11 enthält eine Zusammenfassung, offene Fragen und Probleme werden diskutiert und es wird ein Ausblick gegeben. Der Anhang der Arbeit findet sich online auf der Website des Verlages.<sup>3</sup> Auf die Inhalte des Anhangs wird im weiteren Verlauf der Arbeit eingegangen.<sup>4</sup>

---

3 Siehe <http://www.budrich-verlag.de>. DOI: 10.3224/84740183A

4 Daneben sei noch auf einige Formalia hingewiesen:

- Als Dezimaltrennzeichen wird wie im Englischen üblich ein Punkt anstelle eines Kommas verwendet; also beispielsweise 1.5 anstelle von 1, 5.
- Der Text gliedert sich in vier durchnummerierte Ebenen mit folgenden Bezeichnungen: Kapitel (bspw. Kapitel 3), Abschnitte (bspw. Abschnitt 3.1), Unterabschnitte (bspw. Unterabschnitt 3.1.2), sowie eine vierte, als Paragraph bezeichnete Gliederungsebene (bspw. Paragraph 3.1.2.4).
- Kürzere wörtliche Zitate werden im Fließtext in Anführungszeichen gesetzt. Längere wörtliche Zitate werden vom Fließtext abgesetzt und eingerückt wiedergegeben, wobei auf Anführungszeichen verzichtet wird.
- Einen Überblick über die verwendete Notation gibt das Symbolverzeichnis auf Seite 15. In dieses wurden nur die am häufigsten verwendeten Symbole aufgenommen. Ferner werden manche Symbole auch mit mehreren Bedeutungen benutzt. Grundsätzlich werden alle Symbole an der Stelle ihrer ersten Nennung erläutert. Bei mehrfach verwendeten Symbolen wird zudem immer darauf hingewiesen, wie diese an der entsprechenden Stelle zu interpretieren sind.



## 2 Verwandtschaft und Anwendungsbeispiele

### 2.1 Definition und Abgrenzung des Begriffs Verwandtschaft

#### 2.1.1 Definition von Verwandtschaft und damit verbundener Begriffe

Unter Verwandtschaft wird allgemein eine Verbindung von Personen entweder über direkte beziehungsweise gemeinsame Abstammung oder Heirat verstanden (vgl. Hill & Kopp 2006: 18; Huinink & Konietzka 2007: 27; Jakoby 2008b: 20).<sup>5</sup> Im ersten Fall spricht man von konsanguiner Verwandtschaft, im zweiten Fall von Affinialverwandtschaft (vgl. Jakoby 2008b: 20).<sup>6,7</sup> Beispiele für konsanguine Verwandtschaft sind die Großeltern und Geschwister einer Person. Affinialverwandtschaft umfasst Ehepartner sowie „angeheiratete“ Verwandtschaft, wie beispielsweise Schwiegereltern.

Ausgehend von dieser Definition lassen sich diverse weitergehende Unterteilungen von Verwandtschaft vornehmen. Beispielsweise ist eine Untergliederung konsanguiner Verwandtschaft in „lineare“ und „kollaterale“ Verwandtschaft möglich (vgl. Jakoby 2008b: 21).<sup>8</sup> Unter linearer Verwandtschaft werden direkte Vor- oder Nachfahren zusammengefasst (bspw. Eltern, Enkel) und unter kollateraler Verwandtschaft versteht man Personen, die zwar gemeinsame Vorfahren haben, allerdings nicht direkt voneinander abstammen (bspw. Geschwister) (vgl. ebd.: 21).

Verwandtschaft kann weiter nach Verwandtschaftsgraden differenziert werden. Der Grad der Verwandtschaft ergibt sich über die Zahl der vermittelnden Geburten (vgl. Jakoby 2008b: 25).<sup>9</sup> Je nach Grad der Verwandtschaft und je nach vermittelnden Geburten existieren unterschiedliche Bezeichnungen für Verwandte, die teils schon verwendet wurden, wie beispielsweise „Großeltern“. Eine einfache schematische Darstellung konsanguiner Verwandtschaft, die eine genauere Bestimmung der Bezeichnungen und der bisher eingeführten Begriffe erlaubt, wurde von Atkins (1974) entwickelt.<sup>10</sup> Diese findet sich in Tabelle 2.1, in der Bezeichnungen für Ver-

5 Eine weitere Möglichkeit ist Adoption. Diese wird aufgrund ihrer quantitativ eher geringen Bedeutung im weiteren nicht beachtet. Für eine Diskussion s. Wild (1998).

6 Diese Unterteilung von Verwandtschaft entspricht im Wesentlichen der rechtlichen Definition von Verwandtschaft in der BRD, wie sie im Bürgerlichen Gesetzbuch festgehalten ist (s. insb. §1589 und §1590). Eine Diskussion dieser rechtlichen Situation findet man bei Lucke (1998) und Ecarius (2007: 224ff.).

7 In der Literatur finden sich teils noch weitere Bezeichnungen, die den Verwandtschaftsbegriff erweitern. Ein Beispiel ist „fiktive Verwandtschaft“, womit in aller Regel Freunde einer Person gemeint sind, die wie enge Verwandte behandelt werden (vgl. Day 1989: 187). Einen Überblick über weitere Bezeichnungen findet man bei Day (1989: 187f.). Diese Erweiterungen des Verwandtschaftsbegriffs werden hier im weiteren nicht berücksichtigt, da durch sie beliebige Personen als Verwandtschaft aufgefasst werden können, womit das Verwandtschaftskonzept sehr unscharf werden würde (vgl. ebd.: 188).

8 Zuweilen werden anstelle der Begriffe „linear“ und „kollateral“ auch die Begriffe „vertikal“ und „lateral“ verwendet (s. bspw. Rosenbaum & Timm 2008: 41).

9 Eine andere Möglichkeit, den Grad der Verwandtschaft zu ermitteln, besteht darin, die genetische Ähnlichkeit von Personen zu betrachten (vgl. Voland & Paul 1998: 38). Dies wird hier im weiteren nicht berücksichtigt.

10 Bezeichnungen für andere Verwandtschaftsbeziehungen lassen sich teilweise durch das anfügen eines Präfix an eine der in der Tabelle genannten Bezeichnungen erzeugen (bspw. „Stief-“, „Schwieger-“ oder „Halb-“). Dabei ist aber zu beachten, dass teils spezielle und teils keine Bezeichnungen existieren. Beispielsweise wird entferntere affinale Verwandtschaft zuweilen unter dem Begriff „Schwippsschwägerschaft“ subsumiert (vgl. Ecarius 2007: 220).

wandte, sowie Pfeile, die vermittelnde Geburten repräsentieren, eingezeichnet sind (vgl. ebd.: 9). Die Bezeichnungen sind direkte Übersetzungen der bei Atkins verwendeten englischen Begriffe, die dem deutschen Sprachgebrauch entsprechen (vgl. Narrol 1958: 753f.).<sup>11</sup>

Ausgangspunkt der Betrachtung ist eine Person, die als „Ego“ bezeichnet wird – eine aus der Ethnologie stammende Benennung (vgl. bspw. Goodman et al. 1974: 20). Ausgehend von Ego lässt sich nun beliebigen Verwandten ein Zahlenpaar  $(i, j)$  zuordnen, wobei Ego selbst das Paar  $(0, 0)$  zugeordnet wird (vgl. Atkins 1974: 3).  $j$  gibt die Zahl der vermittelnden Geburten bezogen auf Vorfahren von Ego an (vgl. ebd.: 3). Beispielsweise wird die verwandtschaftliche Beziehung zwischen Ego und ihren Eltern durch eine Geburt vermittelt (die Geburt von Ego selbst), so dass für die Eltern  $j = 1$  ist. Zwischen Ego und ihren Großeltern steht zusätzlich noch die Geburt eines Elternteils, so dass für die Großeltern  $j = 2$  gesetzt wird.

Die erste Zahl  $i$  des Zahlenpaars  $(i, j)$  gibt die Zahl der vermittelnden Geburten bezogen auf Nachfahren an, allerdings immer konditional auf  $j$  (vgl. Atkins 1974: 3). Dies bedeutet, dass zunächst entsprechend des Wertes von  $j$  zu Vorfahren von Ego gezählt wird.  $i$  gibt dann die Zahl der Geburten an Nachfahren ausgehend von den über  $j$  bestimmten Verwandten an. So weisen beispielsweise alle Geschwister von Ego das Zahlenpaar  $(1, 1)$  auf. Denn es gibt  $j = 1$  vermittelnde Geburt in Richtung der Vorfahren von Ego und von diesen ausgehend  $i = 1$  Geburt von Nachfahren.

Der Grad der Verwandtschaft kann ausgehend von einem Zahlenpaar  $(i, j)$  über  $i + j$  bestimmt werden. Für lineare Verwandtschaft ist anhand der bereits gegebenen verbalen Definition leicht ersichtlich, dass entweder  $i = 0$  oder  $j = 0$  vorliegen muss. Bei kollateraler Verwandtschaft gilt sowohl  $i \neq 0$  als auch  $j \neq 0$ . Ferner kann durch vertauschen von  $i$  und  $j$  bestimmt werden, in welcher Beziehung Ego zu einem Verwandten  $(i, j)$  steht (vgl. Atkins 1974: 7). Beispielsweise wird Onkel und Tanten das Zahlenpaar  $(1, 2)$  zugeordnet. Durch vertauschen erhält man  $(2, 1)$  – Ego ist eine Nichte oder ein Neffe der Onkel und Tanten.

Anhand dieses Schemas lässt sich eine weitergehende Definition von Verwandtschaft, die von Schütze & Wagner (1998) gegeben wurde, klären. Sie definieren Verwandtschaft als eine „[...] Form der sozialen Beziehung zwischen Menschen, die mindestens nach den Kriterien Geschlecht, Generationenzugehörigkeit, Bluts- versus Schwiegerverwandtschaft variiert“ (ebd.: 7). Die Unterscheidung zwischen Blutsverwandtschaft (konsanguiner Verwandtschaft) und Schwiegerverwandtschaft (affinialer Verwandtschaft) wurde bereits weiter oben erläutert.

Die Differenzierung nach Geschlecht meint zwei verschiedene Unterscheidungen. Zum einen gibt es für bestimmte Verwandte geschlechtsspezifische Bezeichnungen, die sich bei Fehlen eines geschlechtsneutralen Oberbegriffs wie „Eltern“ bereits in Tabelle 2.1 finden. Ansonsten kann entsprechend des üblichen Sprachgebrauchs differenziert werden, wie beispielsweise zwischen Mutter und Vater (vgl. Narrol 1958: 752). Zum anderen kann bei über mehrere Geburten vermittelte Verwandten nach dem Geschlecht der vermittelnden Geburten differenziert werden.<sup>12</sup> So lässt sich beispielsweise von Großeltern mütterlicherseits sprechen.

11 Teilweise existieren neben den hier angegebenen Bezeichnungen noch weitere, allerdings oftmals nicht mehr gebräuchliche Begriffe. Für einen Überblick s. Narrol (1958: 753f.).

12 Dies kann insofern von besonderer Bedeutung sein, als Verwandtschaft in manchen kulturellen Kontexten entweder nur über männliche oder nur über weibliche Vorfahren konstruiert wird (vgl. Huinink & Konietzka 2007: 27). Ersteres bezeichnet man als patrilineare Deszendenz, zweiteres als matrilineare Deszendenz (vgl. Hill & Kopp 2006: 18). Wird wie in Deutschland Verwandtschaft im gleichen Maße sowohl über männliche als auch weibliche Vorfahren gebildet, spricht man von bilinearer Deszendenz (vgl. Huinink & Konietzka 2007: 27).

Tabelle 2.1: Verwandtschaftsschema nach Atkins

(0,0)	←		←		
Ego		(0,1)		(0,2)	(0,3)
↓		Eltern		Großeltern	Urgroßeltern
(1,0)		↓		↓	↓
Kinder		(1,1)		(1,2)	(1,3)
↓		Geschwister		Onkel & Tanten	Großonkel u. -tanten
(2,0)		↓		↓	↓
Enkelkinder		(2,1)		(2,2)	(2,3)
↓		Neffen u. Nichten		Cousins & Cousinen	Großcousins & -cousinen
(3,0)		↓		↓	↓
Urenkel		(3,1)		(3,2)	(3,3)
		Großneffen u. -nichten		Cousins & Cousinen 2. Grades	Großcousins & -cousinen 2. Grades

Quelle: Atkins (1974: 9)

Die Differenzierung nach Generation erfolgt über die Bildung der Differenz  $i - j$  (vgl. Atkins 1974: 8). Alle Verwandten, bei denen sich der selbe Wert  $x = i - j$  ergibt, gehören dabei zu einer Generation. Beispielsweise ergibt sich der Wert  $-1$  bei Eltern, Onkel und Tanten sowie Großcousins und -cousinen, die somit einer Generation angehören. Allgemein gehören alle die Verwandten zu einer Generation, die auf einer der Diagonalen der Tabelle 2.1 liegen.<sup>13</sup>

Neben den hier vorgestellten Definitionen und Unterscheidungen existieren zwar noch etliche weitere, allerdings sind diese für die vorliegende Arbeit nicht von Bedeutung.<sup>14</sup> Im weiteren Verlauf wird die sich aus diesem Unterabschnitt ergebende Bestimmung von Verwandtschaft verwendet, die sich im Wesentlichen mit dem alltäglichen Sprachgebrauch decken sollte.

### *2.1.2 Weitere Definitionen und Abgrenzung von Verwandtschaft gegen andere Begriffe*

Die Begriffe „Verwandtschaft“ und „Familie“ werden im alltäglichen Sprachgebrauch oft synonym verwendet (vgl. Jakoby 2008a: 210). Allerdings weichen in der wissenschaftlichen Literatur vorzufindende Definitionen des Begriffs Familie teils deutlich von der im vorausgegangenen Unterabschnitt gegebenen Definition von Verwandtschaft ab.

Hier wird im weiteren eine Definition des Begriffs Familie verwendet, die aus der amtlichen Statistik stammt (Statistisches Bundesamt 2011c: 32):

Die Familie im „statistischen Sinn“ umfasst [...] alle Eltern-Kind-Gemeinschaften, d. h. Ehepaare, nichteheliche (gemischtgeschlechtliche) und gleichgeschlechtliche Lebensgemeinschaften sowie alleinerziehende Mütter und Väter mit ledigen Kindern im Haushalt. Einbezogen sind in diesen Familienbegriff – neben leiblichen Kindern – auch Stief-, Pflege- und Adoptivkinder ohne Altersbegrenzung. Damit besteht eine Familie immer aus zwei Generationen (Zwei-Generationen-Regel): Eltern/-teile und im Haushalt lebende ledige Kinder.

Unter dem zur Definition von Familie verwendeten Begriff „Haushalt“ versteht man (vgl. ebd.: 32)

[...] jede zusammenwohnende und eine wirtschaftliche Einheit bildende Personengemeinschaft (Mehrpersonenhaushalte) sowie Personen, die allein wohnen und wirtschaften (Einpersonenhaushalte, z. B. Einzeluntermieter). Zum Haushalt können verwandte und familienfremde Personen gehören (z. B. Hauspersonal).

Diese Haushaltsdefinition basiert im Wesentlichen also auf zwei Ideen: der Idee der Wohneinheit und der Idee der Wirtschaftseinheit (vgl. Keilman 1995: 113f.). Genauere Erörterungen und Diskussionen des Haushaltsbegriffs findet man bei Burch (1995a) und Keilman (1995). Vor- und Nachteile der Beschränkung von Familie auf den Haushaltskontext werden von Bien & Marbach (1991) diskutiert.

Eine Definition von Familie, die explizit auf den Verwandtschaftsbegriff zurückgreift, findet sich bei Burch (1979: 174):

13 Der Begriff „Generation“ wird in der sozialwissenschaftlichen Literatur teilweise auch mit anderen Bedeutungen belegt. Für eine Diskussion s. Kertzer (1983). Die hier verwendete Definition wird von Kertzer (1983: 129) unter anderem als „genealogisch“ bezeichnet.

14 Für eine Übersicht s. insb. Jakoby (2008b: 19ff.).

'Family' refers only to those kin with whom one co-resides. Persons living in the same dwelling unit comprise a 'household', whether they are related or not.

Entsprechend dieser Definitionen ist Familie eine „Untergruppe“ (Ecarius 2007: 221) von Verwandtschaft.<sup>15</sup> Einerseits ist Verwandtschaft unabhängig von Koresidenz, während wiederum Koresidenz, die als konstituierendes Merkmal eines Haushalts aufgefasst wird, unabhängig von Verwandtschaft ist (vgl. auch Imhoff & Keilmann 1991: 9). Insofern kann Familie dann vereinfacht als Schnittmenge aus Verwandtschaft und Haushalt aufgefasst werden.<sup>16</sup>

## 2.2 Zwei Beispiele für die Bedeutung von Verwandtschaft

### 2.2.1 Ein Beispiel für die Bedeutung bestimmter Verwandtschaftsbeziehungen: Intergenerationale Beziehungen

Verwandtschaft gilt als eher vernachlässigtes Thema in der sozialwissenschaftlichen Forschung (s. bspw. Hill & Kopp 2006: 258f., Schütze & Wagner 1998: 9). Betrachtet man den Forschungsstand zur Bedeutung von Verwandtschaft zeigen sich allerdings deutliche Unterschiede, je nachdem welche Verwandten man betrachtet. Zu Stiefverwandtschaft und entfernteren Verwandten wie Onkel und Tanten existieren beispielsweise kaum Untersuchungen (vgl. Jakoby 2008a: 209; Steinbach 2008: 154f.). Ähnlich verhält es sich bei Geschwistern, insbesondere bei Beziehungen zwischen Geschwistern im Erwachsenenalter (vgl. Hill & Kopp 2006: 258, Nave-Herz & Feldhaus 2005: 111). Die wenigen vorliegenden Studien sind oftmals nur konzeptuell oder methodisch qualitativ orientiert, betrachten lediglich die Kontakthäufigkeiten zwischen Verwandten oder stammen aus anderen wissenschaftlichen Kontexten wie der Psychologie (bspw. zu Onkeln und Tanten Jakoby 2008a: 216ff.).<sup>17</sup> Zu affinaler Verwandtschaft über den Ehepartner hinaus, wie Schwiegereltern, lassen sich kaum Informationen finden.<sup>18 19</sup> Gleiches gilt für „Halbverwandtschaft“, wie beispielsweise Halbgeschwister.

15 Dies ist bei Verwendung der Definition des Statistischen Bundesamtes nicht gänzlich korrekt, zumal auch nicht-eheliche Partnerschaften berücksichtigt werden, wobei entsprechend der im vorausgegangenen Unterabschnitt gegebenen Definition von Verwandtschaft nicht verheiratete Partner nicht verwandt sind. Somit überlappen sich Familie und Verwandtschaft lediglich.

16 Diese Definition von Familie ist nur eine von etlichen möglichen. Für eine in der Familiensoziologie relativ weit verbreitete Definition, die nicht explizit auf Koresidenz abhebt, s. Nave-Herz (1994: 5f.). Eine weitergehende Diskussion des Familienbegriffs findet man u.a. bei Marbach (2008).

17 Mögliche Ursachen dafür, dass Verwandtschaft zumindest in Teilen ein vernachlässigtes Thema ist, werden bei Jakoby (2008a: 209,214) und Schütze & Wagner (1998: 9) diskutiert.

18 Dabei ist allerdings zu bedenken, dass zwar kaum Informationen zu Schwiegereltern vorliegen, allerdings sehr wohl zu Eltern, wie im weiteren Verlauf dieses Unterabschnitts noch erläutert wird. Wenn nun beispielsweise bekannt ist, dass die Eltern eines Mannes die Betreuung dessen Kinder übernehmen und dieser Mann verheiratet ist, ist zugleich auch bekannt, dass die Schwiegereltern der Frau die Kinder betreuen. Insofern lassen sich wahrscheinlich einige der Befunde zur Beziehung zwischen Eltern und erwachsenen Kindern auch teilweise auf Schwiegereltern übertragen. Ein Gegenbeispiel das zeigt, dass dies nicht immer der Fall ist, findet sich bei Diewald (1989: 455). Im Rahmen des Wohlfahrtssurvey 1984 gaben 57% der Befragten mit lebenden Eltern an, diese mindestens einmal in der Woche zu treffen. Für Schwiegereltern gaben dies hingegen 45% an. Insofern weisen Eltern und Schwiegereltern also nicht die selbe Kontakthäufigkeit auf.

19 Eine Ausnahme ist die Betreuung von pflegebedürftigen Menschen durch ihre Schwiegerkinder, s. Unterabschnitt 2.2.2.

Anders verhält es sich mit der Beziehung zwischen Eltern und Kindern, welche mit Hinblick auf die kindliche Sozialisation ein zentrales Thema sozialwissenschaftlicher Forschung zu Familie ist. Darüber hinaus ist seit den 1980er Jahren verstärkt die Triade aus Großeltern, Eltern und Kindern beziehungsweise Enkelkindern in den Fokus der Forschung gerückt (vgl. Nave-Herz 2001: 22).<sup>20</sup> Dass dieses Thema erst vergleichsweise spät aufgegriffen wurde, wird zum einen damit begründet, dass die geteilte Lebenszeit von Großeltern und Enkeln historisch betrachtet erst seit Beginn des 20. Jahrhunderts und durch die in diesem stattgefundenen demographischen Veränderungen nennenswerte Ausmaße angenommen hat (vgl. Lange & Lauterbach 1998: 228). Zum anderen wird darauf verwiesen, dass die Thematisierung vor allem durch „wissenschaftsexterne Faktoren“ ausgelöst wurde, insbesondere durch sozialpolitische Debatten (vgl. Nave-Herz 2001: 22).

Seit es dieses Thema in den wissenschaftlichen Blickpunkt geschafft hat, gibt es eine wachsende Zahl an Studien, die sich mit diesem unter Schlagworten wie „intergenerationale Beziehungen“, „intergenerative Beziehungen“, „Generationenbeziehungen“, „Drei-Generationen-Familie“ und allgemeiner „Mehrgenerationenfamilie“ befassen (Huinink & Konietzka 2007: 209; Nave-Herz 2001: 21).<sup>21</sup> Die Verwendung des Generationenbegriffs entspricht der im vorausgegangenen Abschnitt gegebenen Definition eingegrenzt auf lineare Verwandte. Allerdings setzen die Bezeichnungen „Drei-Generationen-Familie“ und „Mehrgenerationenfamilie“ im Gegensatz zu der in der vorliegenden Arbeit verwendeten Definition von Familie keinen gemeinsamen Haushalt voraus (vgl. Nave-Herz 2001: 21) und werden in der Fachliteratur mehr oder weniger synonym mit den anderen Begriffen verwendet. Im weiteren wird der Klarheit halber nur die Bezeichnung „intergenerationale Beziehungen“ verwendet.

Bei der Betrachtung intergenerationaler Beziehungen wird in der Literatur häufig auf das Konzept der „intergenerationalen Solidarität“ zurückgegriffen (vgl. u.a. Bengtson & Roberts 1991; Bengtson 2001). Bei diesem Konzept werden intergenerationale Beziehungen anhand von sechs Dimensionen betrachtet (Bengtson & Roberts 1991: 857, Huinink & Konietzka 2007: 209):

- associational solidarity: Art und Häufigkeit von Interaktionen (bspw. Besuche, Telefonate, etc.), Kontaktintensität
- affectual solidarity: Art und Grad positiver Gefühle, emotionale Verbundenheit
- consensual solidarity: Ausmaß der Übereinstimmung an Werten und Einstellungen
- functional solidarity: Ausmaß an Unterstützung und Austausch von Ressourcen
- normative solidarity: Ausmaß der Erfüllung von familiären Verpflichtungen
- structural solidarity: räumliche Nähe, „Opportunitätsstruktur“

Dabei wird unter „Solidarität“ allgemein eine „positive Interdependenz“ verstanden (Höpflinger nach Brandt 2009: 19). Dieses Konzept legt somit einen Schwerpunkt auf die Potenziale von intergenerationalen Beziehungen (vgl. Huinink & Konietzka 2007: 210). In der Forschung

---

20 Formal werden hierbei drei aufeinander folgende Generationen linearer Verwandtschaft betrachtet. Ausgehend von Ego umfasst dies in der Notation nach Atkins (1974) drei mögliche Kombinationen: erstens (0, 2), (0, 1) und (0, 0); zweitens (0, 1), (0, 0) und (1, 0); drittens (0, 0), (1, 0) und (2, 0). Unabhängig von der konkreten Kombination treten immer vier Verwandtschaftsbeziehungen auf: (2, 0), (1, 0), (0, 1) und (0, 2). Betrachtet wird also eine Teilmenge linearer Verwandtschaft.

21 Eine Literaturübersicht zu auf Deutschland bezogene Forschung findet sich bei Ross-Strajhar (2005).

werden allerdings auch mögliche negative Aspekte und Konflikte thematisiert (Bengtson & Martin 2001: 212; Lowenstein 2005: 406f.) und die Verwendung des Begriffs „Ambivalenz“ anstelle von Solidarität vorgeschlagen, um diese negativen Aspekte mit zu berücksichtigen (vgl. Lüscher & Pillemer 1998). Im weiteren soll dieses Konzept wie in der Literatur nicht unüblich lediglich zu einer einfachen, heuristischen Systematisierung der Betrachtung von intergenerationalen Beziehungen dienen (vgl. Huinink & Konietzka 2007: 210), wobei die Diskussion um den Solidaritätsbegriff vernachlässigt wird.<sup>22</sup>

Beispiele für konkrete, auf Deutschland bezogene Studien, die sich mit der räumlichen Nähe und Kontakthäufigkeit und -arten auseinander setzen, sind die Arbeiten von Kohli et al. (2000: 185ff.), Kreyenfeld et al. (2009), Lange & Lauterbach (1998) sowie Murphy (2008). Wesentliche Ergebnisse sind, dass es in der BRD zwar sehr wenige „Mehrgenerationenhaushalte“, also Haushalte, in denen mehr als zwei Generationen leben, gibt. Allerdings ist die räumliche Entfernung der Wohnorte von Eltern und erwachsenen Kindern beziehungsweise Großeltern und Enkeln in der Regel nicht so hoch, dass die Möglichkeit für direkte Begegnungen durch sie eingeschränkt wäre. Zudem gibt es oft regelmäßige Kontakte zwischen diesen Personen. Bei einem Vergleich von Daten aus den Jahren 1986 und 2001 kommen Kalmijn & De Vries (2009) zu dem Schluss, dass sich die Kontakthäufigkeit zumindest zwischen den beiden genannten Jahren praktisch nicht verändert hat.

Studien zu emotionaler Nähe stammen beispielsweise von Bengtson & Martin (2001), Nauck & Suckow (2006), Neyer & Lang (2003) und Kohli et al. (2000: 188f.). Bei diesem Aspekt lässt sich verkürzt festhalten, dass emotionale Verbundenheit von den Mitgliedern einzelner Generationen allgemein zwar als relativ hoch angesehen wird, allerdings das Ausmaß der wahrgenommenen Verbundenheit je nach Generation, zu der eine Person gehört, schwankt (vgl. bspw. Nauck & Suckow 2006: 1169f. u. 1182).

Arbeiten zu *consensual solidarity* sind vergleichsweise selten (vgl. Bengtson & Martin 2001: 215). Zwar gibt es Untersuchungen zur intergenerationalen „Vererbung“ von Einstellungen, allerdings beschränken sich diese in der Regel auf die Untersuchung des Einflusses der elterlichen Einstellungen auf die Einstellungen von Kindern und Jugendlichen (vgl. bspw. Baier & Hadjar 2004).

Ein relativ großer Teil der Literatur befasst sich mit intergenerationalen Transfers, die den Austausch von Unterstützung und Ressourcen umfassen. Beispiele sind die Arbeiten von Brandt et al. (2008), Hoff (2007) und Kohli et al. (2000: 191ff.). Es lässt sich zusammenfassen, dass es rege Austauschbeziehungen zwischen den Generationen gibt, wobei erwachsene Kinder und seltener Enkelkinder ihren Eltern beziehungsweise Großeltern eher instrumentelle Hilfe geben, während Großeltern ihren erwachsenen Kindern und Enkelkindern eher finanzielle Hilfe zukommen lassen. Dabei ist aber auch instrumentelle Hilfe durch Großeltern nicht unüblich, wie beispielsweise die Betreuung von Enkelkindern (vgl. Hank & Buber 2009). Die Austauschbeziehung ist oftmals nicht reziprok und es finden sich Hinweise darauf, dass sich die Richtung von Transfers im Lebensverlauf ändert: Zunächst fließen Transfers von Eltern an ihre erwachsenen Kinder und Enkelkinder. Mit steigendem Alter der Eltern sind diese selbst

---

22 Eine Diskussion des Konzepts der intergenerationalen Solidarität und von alternativen Vorschlägen wie intergenerationaler Ambivalenz findet sich bei Brandt (2009: Kap. 2). Daneben spielt in der Literatur auch das theoretische Konzept des Sozialkapitals eine Rolle, auf welches hier nicht weiter eingegangen werden soll. Eine Diskussion findet sich bei Astone et al. (1999).

zunehmend auf Unterstützung angewiesen und erhalten Transfers von ihren erwachsenen Kindern (vgl. Brandt et al. 2008: 379ff.). Auch nach dem Tod einer Person sind Transfers durch diese in Form von Erbschaften möglich (vgl. Szydlík 2011).

Das Thema der *normative solidarity* wird kaum in der Literatur behandelt. Es existieren lediglich Arbeiten, die versuchen, das Ausmaß von familiären Verpflichtungen festzustellen. Ein Beispiel stammt von Daatland & Herlofson (2003), die anhand von Umfragedaten international vergleichend untersuchten, inwieweit die Unterstützung von pflegebedürftigen Eltern durch ihre erwachsenen Kinder als Verpflichtung angesehen wird. Für die BRD zeigt sich, dass circa 66% der Befragten dies als Verpflichtung wahrnehmen – ein im internationalen Vergleich relativ niedriger Anteil (vgl. ebd.: 546ff.).

## 2.2.2 *Ein Beispiel für die Bedeutung von Verwandtschaft in einem bestimmten Themenbereich: Betreuung Pflegebedürftiger durch Verwandte*

Ein Beispiel für die Bedeutung von Verwandtschaft in einem bestimmten Themenbereich ist die Betreuung von pflegebedürftigen Menschen. Laut amtlicher Pflegestatistik gab es in der BRD zum Ende des Jahres 2009 circa 2.34 Millionen pflegebedürftige Menschen (vgl. Statistisches Bundesamt 2011b: 6). Von diesen wurden ungefähr 1.62 Millionen Personen (ca. 69%) in Privathaushalten versorgt, wobei die Pflege oftmals ausschließlich von Angehörigen oder von Angehörigen und Pflegediensten zusammen geleistet wurde (vgl. ebd.: 6).

In der Pflegestatistik des Statistischen Bundesamtes werden allerdings nur Personen als pflegebedürftig erfasst, die Leistungen nach dem SGB XI – also aus der Pflegeversicherung – beziehen (vgl. Statistisches Bundesamt 2011b: 29).<sup>23 24</sup> Zudem umfasst der Begriff „Angehörige“ sowohl Verwandte als auch Freunde, Nachbarn und Bekannte.<sup>25</sup> Detailliertere Informationen zu pflegenden Personen finden sich in diversen Erhebungen, wie beispielsweise dem „Survey of Health, Ageing and Retirement in Europe“ (SHARE), bei dem Personen im Alter von 50 und mehr Lebensjahren in elf europäischen Ländern befragt wurden, darunter auch Deutschland (vgl. bspw. Attias-Donfut et al. 2005). Pflegebedürftigkeit wird dabei über Angaben der Befragten zu Einschränkungen bei täglichen Verrichtungen erfasst (vgl. Blome et al. 2008: 197f.). Durch diese von der amtlichen Statistik abweichende Erfassung wird auch dann ein Pflegebedarf festgestellt, wenn kein Anspruch auf Leistungen der Pflegeversicherung besteht beziehungsweise diese (noch) nicht beantragt wurden (vgl. Brandt et al. 2008:

23 In §14 Abs. 1 SGB XI findet sich folgende Definition von Pflegebedürftigkeit: „Pflegebedürftig im Sinne dieses Buches sind Personen, die wegen einer körperlichen, geistigen oder seelischen Krankheit oder Behinderung für die gewöhnlichen und regelmäßig wiederkehrenden Verrichtungen im Ablauf des täglichen Lebens auf Dauer, voraussichtlich für mindestens sechs Monate, in erheblichem oder höherem Maße (§ 15) der Hilfe bedürfen.“ Weitergehende Erläuterungen finden sich in §14 Abs. 2-3 SGB XI. Eine Abgrenzung unterschiedlicher Stufen der Pflegebedürftigkeit wird in §15 SGB XI gegeben.

24 Ausführlichere Diskussionen der deutschen Pflegeversicherung findet man bei Blome et al. (2008: 171ff.) und Haberkern (2009: 63ff.).

25 Genauer entspricht in der Pflegestatistik die Zahl der Pflegebedürftigen, die ausschließlich durch Angehörige betreut werden, der Zahl der Empfänger von Pflegegeld nach §37 SGB XI, die keine zusätzlichen Sachleistungen aus der Pflegeversicherung erhalten (vgl. Statistisches Bundesamt 2011b: Tabelle 1.1, Fußnote 2). Dabei ist Pflegegeld nach §37 SGB XI Geld für „selbst beschaffte Pflegehilfen“, die anstelle von professioneller Pflege in Anspruch genommen werden können. Was für Personen aber „Pflegehilfen“ sind, wird in der amtlichen Statistik nicht erfasst. Daneben besteht zudem wie bereits angedeutet auch die Möglichkeit, sogenannte Kombinationsleistungen aus Pflegegeld und Sachleistungen zu beziehen (s. §38 SGB XI).

301), wodurch die mit den Daten des SHARE ermittelte Pflegequote höher als die amtlich ausgewiesene Pflegequote ist (vgl. Blome et al. 2008: 201).

Bei den im SHARE erfassten Pflegebedürftigen wurde ein Großteil der Pflege durch Privatpersonen geleistet (vgl. Haberkern & Szydlik 2008: 90). Dabei handelte es sich oft um die (Ehe-)Partner beziehungsweise Kinder einer pflegebedürftigen Person (vgl. Blome et al. 2008: 207). Genauer pflegten im Jahr 2003 Partner in 55.9% und Kinder in 43.1% der Fälle, bei denen die Betreuung durch Angehörige stattfand (vgl. ebd.: 207). Bei den Angaben, auf denen diese Werte basieren, waren Mehrfachnennungen möglich, so dass beispielsweise sowohl Partner als auch Kinder genannt werden konnten. Hierdurch weisen die gerade genannten Anteilswerte nicht notwendigerweise auf eine alleinige Pflege durch entsprechende Verwandte hin. Im Vergleich zu Partnern und eigenen Kindern wesentlich seltener pflegten Schwiegerkinder und andere Verwandte sowie sonstige, nicht verwandte Personen. Letztere traten in 15.3% der Fälle auf, bei denen Pflege durch Privatpersonen erfolgte (vgl. ebd.: 207). Schneekloth (2006) kam mit Daten der Erhebung „Möglichkeiten und Grenzen selbstständiger Lebensführung“ (MuG III) aus dem Jahr 2002 zu qualitativ ähnlichen Ergebnissen.<sup>26</sup>

Ein nicht unerheblicher Teil der Pflege wird in Deutschland also durch (enge) Verwandte erbracht. Schneekloth merkt hierzu an: „Der ‚größte Pflegedienst in Deutschland‘ ist und bleibt die Familie.“ (Schneekloth 2006: 408). Dabei ist aber auch zu bedenken, dass einerseits für Leistungen der Pflegeversicherung bestimmte Bezugsvoraussetzungen erfüllt sein müssen und andererseits diese Leistungen in aller Regel nicht bedarfsdeckend sind, so dass eine alleinige Versorgung durch professionelle Dienste ohne instrumentelle oder finanzielle Transfers von Verwandten kaum möglich ist (vgl. Haberkern 2009: 65). Zudem sind Kinder in begrenztem Umfang dazu verpflichtet, die Pflegekosten für ihre Eltern mitzutragen, wenn diese nicht selbst für ihre Pflege aufkommen können (vgl. ebd.: 65). Umgekehrt besteht die Möglichkeit, dass pflegenden Angehörigen das Pflegegeld durch die Betreuten weitergereicht wird (s. Fußnote 25; vgl. ebd.: 65). Insofern ist also eine zentrale Stellung von Verwandtschaft bei der Pflege zumindest in gewissen Grenzen rechtlich verankert (vgl. ebd.: 65).<sup>27</sup> Des Weiteren besteht auch eine hohe normative Verpflichtung zur Pflege der eigenen Eltern (vgl. Haberkern 2009: 66), die sich wie im vorausgegangenen Unterabschnitt bereits erwähnt auch in Bevölkerungsumfragen widerspiegelt (vgl. Daatland & Herlofson 2003: 546ff.). Zwar zeigt sich in Befragungen auch, dass sich eine knappe Mehrheit für eine Finanzierung von Pflege durch den Staat findet (vgl. Blome et al. 2008: 329). Umgekehrt präferieren je nach Altersgruppe allerdings 60% bis 70% die familiäre Pflege ihrer eigenen Eltern gegenüber der Pflege durch professionelle Dienste (vgl. ebd.: 327f.; Daatland & Herlofson 2003: 549). Ähnliche Tendenzen zeigen sich bei der Frage, ob man selbst im Pflegefall durch Verwandte oder Pflegedienste betreut werden möchte (vgl. Daatland & Herlofson 2003: 550).

---

26 Dabei ist zu beachten, dass pflegebedürftige Personen auf unterschiedlich große Netzwerke zurückgreifen können, je nachdem, welche Verwandten vorhanden sind (Haberkern vgl. 2009: 136). So ist bei kinderlosen keine Pflege durch eigene Kinder möglich. Zudem gibt es auch deutliche Unterschiede zwischen den Geschlechtern. Beispielsweise pflegen Töchter tendenziell eher als Söhne (vgl. Blome et al. 2008: 208), wobei aber Schneekloth (2006: 408) einen steigenden Anteil von männlichen Pflegepersonen feststellt.

27 Dies spiegelt sich auch in § 3 SGB XI wider, der den „Vorrang der häuslichen Pflege“ regelt: „Die Pflegeversicherung soll mit ihren Leistungen vorrangig die häusliche Pflege und die Pflegebereitschaft der Angehörigen und Nachbarn unterstützen, damit die Pflegebedürftigen möglichst lange in ihrer häuslichen Umgebung bleiben können. [...]“.

Die in diesem und dem vorausgegangenen Unterabschnitt vorgestellten Beispiele zeigen die große gesellschaftliche Bedeutung von Verwandtschaft auf. Eine starke Veränderung des Vorhandenseins von Verwandten könnte somit deutliche Konsequenzen haben, beispielsweise in Bezug auf die Nachfrage nach professioneller Pflege und die Finanzierung der Pflege. Bevor einige beispielhafte Überlegungen vorgestellt werden, inwiefern aufgrund des demographischen Wandels mit Veränderungen des Vorhandenseins von Verwandten zu rechnen ist, werden im folgenden Kapitel zunächst einige demographische Grundbegriffe erläutert und die historische demographische Entwicklung in der Bundesrepublik vorgestellt.

## 3 Demographische Grundlagen

### 3.1 Demographische Grundbegriffe

#### 3.1.1 Der Bevölkerungsbegriff

Unter dem Begriff Bevölkerung soll allgemein eine Menge von Personen verstanden werden (vgl. Preston et al. 2001: 1). Dabei lassen sich Bevölkerungen in zeitlicher, räumlicher und sachlicher Hinsicht abgrenzen (vgl. Rinne 1994: 42). Eine zeitliche Abgrenzung erfolgt durch die Angabe eines genauen Zeitpunktes, wie beispielsweise einem bestimmten Stichtag, zu dem die Bevölkerung betrachtet werden soll (vgl. ebd.: 42). Eine räumliche Abgrenzung ist durch die Angabe eines bestimmten Gebietes, wie beispielsweise dem Staatsgebiet der BRD, möglich (vgl. ebd.: 42). Durch die sachliche Abgrenzung wird die Bevölkerung weitergehend eingegrenzt, wobei zwischen drei Arten der sachlichen Eingrenzung unterschieden wird: einer technischen, einer staatsrechtlichen sowie einer administrativen Eingrenzung (vgl. ebd.: 42).<sup>28</sup> Bei ersterer Zählen alle die Personen zu einer Bevölkerung, die sich zum festgelegten Zeitpunkt im betrachteten Gebiet aufhalten (vgl. ebd.: 42). Die staatsrechtliche Abgrenzung zählt alle Personen zur Bevölkerung, die Staatsbürger eines betrachteten Staates sind (vgl. ebd.: 42). Bei der administrativen Eingrenzung wird die tatsächliche Einwohnerschaft eines Gebietes als dessen Bevölkerung aufgefasst (vgl. ebd.: 42).

In der amtlichen Statistik der BRD wird das administrative Bevölkerungskonzept verwendet und die „Bevölkerung am Ort der alleinigen bzw. der Hauptwohnung“ betrachtet (Statistisches Bundesamt 2011c: 30).<sup>29</sup> Dabei zählen genauer alle Personen zur Bevölkerung, deren Wohnung beziehungsweise bei Personen mit mehreren Wohnungen deren Hauptwohnung im Staatsgebiet der BRD liegt (vgl. ebd.: 32). Unter der „Hauptwohnung“ wird die „[...] vorwiegend benutzte Wohnung [...]“ verstanden (ebd.: 32).

Bevölkerungen lassen sich zum einen bezüglich ihrer „Struktur“ und zum anderen bezüglich ihrer „Dynamik“ beschreiben (Mueller 1993: 2; Münz & Ulrich 2001: 478). Unter der Struktur einer Bevölkerung versteht man die absolute und relative Verteilung von interessierenden Merkmalen bei den die Bevölkerung konstituierenden Personen (Mueller 1993: 2). Zwei üblicherweise betrachtete Merkmale sind Alter und Geschlecht (vgl. Poston 2006: 19). Die Dynamik einer Bevölkerung drückt sich in Veränderungen der Struktur zwischen zwei Zeitpunkten aus (vgl. Mueller 1993: 2). Beispielsweise weisen Personen zu zwei Zeitpunkten verschiedene Alter auf und es können durch Geburten oder Einwanderung neue Personen zur Bevölkerung hinzukommen (vgl. ebd.: 2).

Wenn eine Person eine bestimmte Merkmalsausprägung beziehungsweise eine bestimmte Kombination von Merkmalsausprägungen aufweist, wird auch davon gesprochen, dass sich

---

28 Shryock & Siegel (1976: 49) unterscheiden zwischen einer „de facto“ und einer „de jure“ Abgrenzung. Erstere entspricht der technischen Abgrenzung, letztere umfasst die staatsrechtliche wie auch die administrative Abgrenzung (vgl. ebd.: 49).

29 Bis 1983 wurde in der amtlichen Statistik das Konzept der „Wohnbevölkerung“ verwendet, welches vom Konzept der „Bevölkerung am Ort der alleinigen bzw. der Hauptwohnung“ abweicht (vgl. Statistisches Bundesamt 2011c: 30). S. auch Fußnote 142 auf Seite 127.

diese Person in einem bestimmten „Zustand“ befindet (vgl. Mueller 1993: 2). Die Veränderung von Zuständen sowie das Hinzukommen in eine Bevölkerung beziehungsweise das Verlassen einer Bevölkerung bezeichnet man allgemein als „Ereignisse“ (vgl. ebd.: 2).

### 3.1.2 Bevölkerungsfortschreibung

Die Größe einer Bevölkerung kann lediglich durch Geburten und Einwanderungen zunehmen, während umgekehrt eine Abnahme nur durch Sterbefälle und Auswanderung möglich ist (vgl. Preston et al. 2001: 2). Geburten und Sterbefälle werden als „natürliche Bevölkerungsbewegungen“ zusammengefasst, wobei teils noch Eheschließungen und -scheidungen hinzugenommen werden (vgl. Statistisches Bundesamt 2011c: 30). Ein- und Auswanderung werden unter dem Begriff „räumliche Bevölkerungsbewegungen“ subsumiert (vgl. ebd.: 30). Natürliche Bevölkerungsbewegungen werden in der BRD von den Standesämtern registriert, während räumliche Bevölkerungsbewegungen durch das Meldewesen erfasst werden (vgl. ebd.: 30). Ferner werden Fertilität, Mortalität und Migration auch unter dem Begriff der „demographische Komponenten“ zusammengefasst (s. bspw. Statistisches Bundesamt 2010c: 9).

Sind die natürlichen und räumlichen Bevölkerungsbewegungen zwischen zwei Zeitpunkten bekannt und kennt man die Bevölkerungsgröße zum ersten Zeitpunkt, lässt sich die Bevölkerungsgröße zum zweiten Zeitpunkt über die sogenannte Buchführungsgleichung berechnen. Die Buchführungsgleichung ist wie folgt definiert (s. bspw. Preston et al. 2001: 2; Rinne 1994: 77; Rohwer & Pötter 2003: 29):

$$n_{t+1} = n_t + B_{[t,t+1)} - D_{[t,t+1)} + M_{[t,t+1)}^i - M_{[t,t+1)}^o \quad (3.1)$$

$n_t$  steht für die Größe einer Bevölkerung zu einem bestimmten Zeitpunkt  $t$ , wobei im weiteren allgemein davon ausgegangen wird, dass es sich um den 1. Januar eines Jahres  $t$  handelt.  $n_{t+1}$  ist dann entsprechend die Größe der Bevölkerung am 1. Januar des folgenden Jahres.  $B_{[t,t+1)}$  erfasst die Zahl der Geburten, die im Intervall  $[t, t + 1)$  stattfinden.<sup>30</sup>  $D_{[t,t+1)}$  steht für die Sterbefälle und  $M_{[t,t+1)}^i$  und  $M_{[t,t+1)}^o$  erfassen die Zahl der Einwanderungen beziehungsweise Auswanderungen im selben Zeitraum.

Bei  $n_t$  und  $n_{t+1}$  handelt es sich um sogenannte „Bestandsgrößen“, die auf einen bestimmten Zeitpunkt bezogen Auskunft über eine Größe geben (vgl. Rohwer & Pötter 2003: 31). In diesem Sinne wird im weiteren der Begriff „Bevölkerungsbestand“ analog zum Begriff Bevölkerungsgröße verwendet. Bei den anderen Größen in Gleichung (3.1) handelt es sich um sogenannte „Stromgrößen“, die die Häufigkeit eines bestimmten Ereignisses in einem bestimmten Zeitraum wiedergeben (vgl. ebd.: 31).<sup>31</sup> Im weiteren wird bei diesen für die Angabe des Zeitraums eine verkürzte Notation verwendet und  $t$  anstelle von  $[t, t + 1)$  angegeben – also beispielsweise  $D_t$  anstelle von  $D_{[t,t+1)}$ . Somit wird davon ausgegangen, dass sich

30 Im weiteren Verlauf werden geschlossene Intervalle mit eckigen Klammern und offene Intervalle mit runden Klammern gekennzeichnet. Halboffene Intervalle werden entsprechend mit eckigen und runden Klammern gesetzt. Beispielsweise steht  $[a, b)$  für die Menge  $\{x \in \mathbb{R} | a \leq x < b\}$ .

31 Im weiteren wird versucht, der Konvention zu folgen, dass Stromgrößen mit Großbuchstaben und Bestandsgrößen mit Kleinbuchstaben bezeichnet werden. Für diese und die im weiteren noch einzuführenden Konventionen der Symbolverwendung ist zu beachten, dass teils Abweichungen vorkommen, beispielsweise wenn mehrere Konventionen einander widersprechen (s. z.B. Fußnote 43 auf S. 38).

Stromgrößen in der Regel auf einen Zeitraum von einem Jahr erstrecken und es entsprechend ausreicht, lediglich den Beginn des Zeitraums anzugeben. Allgemein wird Zeit im folgenden soweit nicht anders angegeben in Kalenderjahren gemessen.

### 3.1.3 Grundlegende demographische Kennwerte

Zur Beschreibung der Dynamik einer Bevölkerung und der dieser Dynamik zugrunde liegenden Bevölkerungsbewegungen existiert eine Vielzahl an speziellen Kennwerten.<sup>32</sup> Ein zentrales Konzept ist das der „demographischen Rate“ (vgl. Preston et al. 2001: 3). Bei demographischen Raten werden Ereignisse in einem bestimmten Zeitraum auf die Personen in einer Bevölkerung bezogen, die diese „erzeugen“ (vgl. ebd.: 3). Genauer werden sogenannte *occurrence-exposure rates* verwendet, die wie folgt definiert sind (vgl. ebd.:3):

$$\text{Rate} = \frac{\text{Häufigkeit des interessierenden Ereignisses}}{\text{Personenjahre mit Risiko des Auftretens}} \quad (3.2)$$

Im Zähler steht somit eine Stromgröße, während im Nenner erfasst wird, wie lange die interessierenden Personen dem Risiko des Auftretens des Ereignisses im betrachteten Zeitraum ausgesetzt sind. Wird beispielsweise ein Zeitraum von einem Jahr betrachtet, geht für jede Person maximal ein Personenjahr in den Nenner ein (vgl. ebd.: 5). Dies ist für eine Person dann der Fall, wenn sie während des gesamten Zeitraums dem Risiko des Auftretens des Ereignisses ausgesetzt ist. Wenn eine Person aber beispielsweise zur Mitte des Jahres auswandert und somit nicht mehr zur Bevölkerung gehört, wird lediglich ein halbes Personenjahr berücksichtigt (vgl. ebd.: 5). Durch dieses Vorgehen werden zum einen unterschiedliche große Bevölkerungen, in denen in aller Regel mehr Ereignisse zu erwarten sind als in kleinen, und zum anderen auch unterschiedlich lange Betrachtungszeiträume miteinander vergleichbar (vgl. ebd.: 5 u. 8). Zudem wird berücksichtigt, dass sich die Zahl der Personen während des Betrachtungszeitraums ändern kann.<sup>33</sup>

Ein Beispiel für demographische Raten sind sogenannte altersspezifische Geburtenraten. Diese sind definiert als (vgl. Preston et al. 2001: 94)

$$f_{x,t} = \frac{B_{x,t}}{J_{x,w,t}} \quad (3.3)$$

$B_{x,t}$  erfasst die Zahl von Geburten, die im Zeitraum  $[t, t + 1)$  Frauen im Altersintervall  $[x, x + 1)$  zugeordnet werden können.  $J_{x,w,t}$  erfasst die von Frauen im Zeitraum  $[t, t + 1)$  im Altersintervall  $[x, x + 1)$  gelebten Personenjahre.<sup>34 35</sup>

32 Eine Übersicht findet man beispielsweise in den ersten Kapiteln des Buches von Preston et al. (2001) sowie bei Mueller (2000).

33 Eine abweichende Definition demographischer Raten findet man bei Rohwer & Pötter (2003: 31). Ferner ist zu beachten, dass in der Regel lediglich Strom- und Bestandsgrößen beobachtet werden und nicht exakte Informationen zu Personenjahren, die geschätzt werden müssen (vgl. bspw. Preston et al. 2001: 15).

34 In der amtlichen Statistik wird anstelle von altersspezifischen Geburtenraten auch von „altersspezifischen Geburtenziffern“ gesprochen (vgl. Rinne 1994: 129).

35 Altersspezifische Geburtenraten, die sich auf Männer beziehen, können ganz analog berechnet werden. Allerdings liegen sowohl für Deutschland als auch international in den wenigsten Fällen die benötigten Daten vor (vgl. bspw. Poston et al. 2006: 871f.).

Beim Alter wird davon ausgegangen, dass es prinzipiell beliebig genau messbar ist, so dass ein Alter  $x$  das exakte Alter einer Person ist. Werden Altersintervalle in der Form  $[x, x + 1)$  angegeben, wird davon ausgegangen, dass  $x \in \mathbb{N}$  gilt. Wenn sich also eine Person im Altersintervall  $[x, x + 1)$  befindet, bedeutet dies, dass sie  $x$  Lebensjahre vollendet hat oder umgangssprachlich „ $x$  Jahre alt“ ist. Im weiteren wird von einigen Ausnahmen größtenteils einfach  $x$  als Abkürzung für  $[x, x + 1)$  angegeben.

Ein aus altersspezifischen Geburtenraten abgeleitetes Maß ist die zusammengefasste Geburtenziffer, die definiert ist als (vgl. Preston et al. 2001: 95)

$$\text{TFR}_t = \sum_{x=\alpha}^{\beta} f_{x,t} \quad (3.4)$$

Die Abkürzung TFR ergibt sich aus der englischen Bezeichnung für die zusammengefasste Geburtenziffer, *total fertility rate*. Sie ergibt sich als Summe aller altersspezifischen Geburtenraten eines Jahres  $t$ , wobei die Summenbildung über die Alter von  $\alpha$  bis einschließlich  $\beta$  läuft.  $\alpha$  steht dabei für das erste und  $\beta$  für das letzte Alter der reproduktiven Phase, die alle gebärfähigen Alter umfasst. Die reproduktive Phase wird unterschiedlich abgegrenzt, beispielsweise von  $\alpha = 14$  bis  $\beta = 45$  (vgl. Statistisches Bundesamt 2011c: 57).<sup>36</sup> Die zusammengefasste Geburtenziffer kann interpretiert werden als die durchschnittliche Zahl an Kindern, die Frauen bekommen würden, die bis zum Ende der reproduktiven Phase überleben und entsprechend der altersspezifischen Geburtenraten des Jahres  $t$  Kinder bekommen. Dabei ist diese Betrachtung rein hypothetisch und bezieht sich nicht auf eine bestimmte, real existierende Gruppe von Frauen, wie beispielsweise einen bestimmten Geburtsjahrgang (vgl. Preston et al. 2001: 95).<sup>37</sup>

Ausgehend von der zusammengefassten Geburtenziffer lässt sich das durchschnittliche Alter bei der Geburt berechnen als

$$\bar{x}_t = \sum_{x=\alpha}^{\beta} \frac{f_{x,t} \cdot x}{\text{TFR}_t} \quad (3.5)$$

Hierbei handelt es sich wie bei der zusammengefassten Geburtenziffer um ein hypothetisches Maß. Es gibt das durchschnittliche Alter bei der Geburt wieder, dass sich für Frauen ergeben würde, die bis zum Ende der reproduktiven Phase überleben und entsprechend der altersspezifischen Geburtenraten des Jahres  $t$  Kinder bekommen. Eine zusammengefasste Geburtenziffer von 1.5 würde bedeuten, dass die hypothetische Gruppe von Frauen durchschnittlich 1.5

36 Bei  $\alpha$  und  $\beta$  handelt es sich zumeist eigentlich um nach unten beziehungsweise oben offene Altersklassen, so dass beispielsweise  $\beta = 45$  für Geburten im Alter von 45 oder mehr Lebensjahren steht (vgl. Statistisches Bundesamt 2011c: 57).

37 In der Literatur findet sich eine umfangreiche Diskussion darüber, ob die zusammengefasste Geburtenziffer ein geeigneter Indikator für die Höhe der Fertilität ist. Diskussionsbeiträge stammen unter anderem von Akers (1965), Bongaarts & Feeney (1998), Ni Bhrolchain (1992) und Whelpton (1946). Ein wesentliches Problem ist, dass die zusammengefasste Geburtenziffer sich nicht nur aufgrund von Änderungen der Höhe der Fertilität verändern kann, sondern auch aufgrund von Änderungen im *timing* von Geburten bei ansonsten gleicher Geburtenhäufigkeit (für ein Beispiel s. Dorbritz 2004). Das *timing* von Geburten gibt an, wann Frauen im Lebenslauf Kinder bekommen (vgl. bspw. Preston et al. 2001: 102f.). Trotz dieser Diskussion ist die zusammengefasste Geburtenziffer der am weitesten verbreitete Kennwert zur Beschreibung von Fertilität, weshalb sie auch in dieser Arbeit verwendet werden soll.

Kinder bekommt und  $\bar{x} = 30$  würde bedeuten, dass die Frauen diese Kinder durchschnittlich im Alter von 30 Jahren geboren haben.<sup>38</sup>

Raten und aus diesen abgeleitete Kennwerte werden auch zur Beschreibung der Sterblichkeit verwendet, wobei die Betrachtung im Rahmen von sogenannten Sterbetafeln erfolgt.<sup>39</sup> Bei einer Sterbetafel handelt es sich um „[...] ein mathematisches Modell der Sterblichkeitsverhältnisse einer Bevölkerung während eines bestimmten Beobachtungszeitraums [...]“ (Statistisches Bundesamt 2011c: 33). Genauer wird die Sterblichkeit einer tatsächlichen oder hypothetischen Kohorte betrachtet (vgl. Rinne 1994: 142). Unter einer „Kohorte“ versteht man allgemein eine Menge von Personen, für die ein bestimmtes Ereignis zu einem gleichen Zeitpunkt stattgefunden hat (vgl. Rohwer & Pötter 2003: 36). Bei Sterbetafeln ist dies in aller Regel die Geburt der Mitglieder der Kohorte. Wird eine tatsächliche Geburtskohorte betrachtet, spricht man von einer Kohortensterbetafel (vgl. Rinne 1994: 142). Wird hingegen ähnlich wie bei der zusammengefassten Geburtenziffer von einer hypothetischen Kohorte ausgegangen, spricht man von einer Periodensterbetafel (vgl. ebd.: 142). Unter einer „Periode“ wird hierbei allgemein ein bestimmter, festgelegter Zeitraum, wie beispielsweise ein Jahr, verstanden (vgl. Wilmoth et al. 2007: 6). Anmerkungen zur Interpretation von Periodensterbetafeln folgen weiter unten.

Ausgangspunkt der Betrachtung ist sowohl bei Kohorten- als auch bei Periodensterbetafeln die Ausgangsgröße der real existierenden oder hypothetischen Kohorte im Alter 0,  $l_0$ , die auch als „Radix“ der Sterbetafel bezeichnet wird (vgl. Keyfitz & Caswell 2005: 29). Allgemein wird die Größe der Kohorte in einem Alter  $x$  mit  $l_x$  bezeichnet (vgl. Preston et al. 2001: 40). Zu beachten ist, dass  $x$  hierbei als exaktes Alter aufgefasst wird und nicht als Abkürzung für ein Altersintervall (vgl. ebd.: 40).  $l_0$  kann bei der Betrachtung einer hypothetischen Kohorte prinzipiell auf einen beliebigen Wert festgelegt werden, üblich sind aber die Werte 100000 oder 1 (vgl. Keyfitz & Caswell 2005: 29).

Gegeben  $l_x$  und  $l_{x+1}$  lässt sich die Zahl der Sterbefälle im Altersintervall  $[x, x + 1)$  berechnen als (vgl. Keyfitz & Caswell 2005: 30)

$$d_x = l_{x+1} - l_x \quad (3.6)$$

Die Wahrscheinlichkeit, im Altersintervall  $[x, x + 1)$  zu sterben, kann dann berechnet werden über (vgl. ebd.: 30)

$$q_x = \frac{d_x}{l_x} \quad (3.7)$$

Die Wahrscheinlichkeit, vom exakten Alter  $x$  bis zum exakten Alter  $x + 1$  zu überleben, erhält man wiederum durch (vgl. Rinne 1994: 144)

$$p_x = 1 - q_x = \frac{l_{x+1}}{l_x} \quad (3.8)$$

38 Eine Alternative besteht in der Berechnung über  $\sum_x B_{x,t}x / (\sum_y B_{y,t})$ . Dies entspricht zwar dem „tatsächlichen“ durchschnittlichen Alter bei der Geburt für ein Jahr  $t$ , führt allerdings zu einem inhaltlich nicht mit der zusammengefassten Geburtenziffer konsistenten Ergebnis, da keine hypothetische Frauengruppe betrachtet wird. Das tatsächliche und das hypothetische durchschnittliche Alter bei der Geburt sind nur identisch, wenn  $\forall x \in \{\alpha, \dots, \beta\} : J_{x,w,t} = c_t$  mit  $c_t > 0$  gilt, da  $B_{x,t} = f_{x,t}J_{x,w,t}$ .

39 Umfassende Darstellungen zu Sterbetafeln finden sich bei Chiang (1984) und Namboodiri & Suchindran (1987).

Für den Fall, dass  $l_0 = 1$  gesetzt wird, also die Ausgangsgröße der betrachteten Kohorte gleich 1 ist, folgt somit  $p_x = l_x$  (vgl. Keyfitz & Caswell 2005: 29f.).

Hiervon ausgehend lässt sich die „durchschnittliche fernere Lebenserwartung“ für das Alter  $x$  berechnen als (vgl. Rinne 1994: 147)

$$e_x = \frac{1}{2} + \sum_{y=x+1}^{\rho} \frac{l_y}{l_x} \quad (3.9)$$

Hierbei ist  $\rho$  das höchste erreichbare Alter, so dass  $p_\rho = 0$  gilt. Wenn beispielsweise  $e_{50} = 34.2$  berechnet wurde, bedeutet dies, dass Personen, die zur betrachteten Kohorte gehören, im exakten Alter von 50 Jahren durchschnittlich noch 34.2 Jahre leben. Ausgehend von der durchschnittlichen ferneren Lebenserwartung für das Alter  $x$  kann die durchschnittliche fernere Lebenserwartung bei der Geburt definiert werden als

$$e_0 = \frac{1}{2} + \sum_{y=1}^{\rho} \frac{l_y}{l_0} \quad (3.10)$$

Dies vereinfacht sich bei  $l_0 = 1$  zu

$$e_0 = \frac{1}{2} + \sum_{y=1}^{\rho} p_y \quad (3.11)$$

Im weiteren werden die Bezeichnungen „durchschnittliche fernere Lebenserwartung bei der Geburt“, „Lebenserwartung bei der Geburt“ sowie „Lebenserwartung“ synonym verwendet.

Demographische Raten sind nun für die Konstruktion von Sterbetafeln insofern von großer Bedeutung, als dass Größen wie beispielsweise  $l_x$  nicht direkt beobachtet werden können. Stattdessen werden zunächst sogenannte Sterberaten berechnet – im Falle von Periodensterbetafeln, die sich auf ein Jahr  $t$  beziehen, geschieht dies über (vgl. Chiang 1984: 72)

$$r_{x,t} = \frac{D_{x,t}}{J_{x,t}} \quad (3.12)$$

Hierbei ist  $D_{x,t}$  die Zahl der Sterbefälle im Beobachtungszeitraum und  $J_{x,t}$  die Zahl der Personenjahre. Auch so definierte Sterberaten sind zumeist nicht beobachtbar (s. Fußnote 33 auf S. 33), weshalb es verschiedene Vorschläge für ihre Schätzung gibt, wie beispielsweise (vgl. ebd.: 72)

$$r_{x,t} = \frac{D_{x,t}}{(n_{x,t} - D_{x,t}) + a_{x,t}D_{x,t}} \quad (3.13)$$

$a_{x,t}$  ist dabei die durchschnittlich im Intervall von  $t$  bis  $t + 1$  gelebte Zeit von Personen, die im Alter von  $x$  bis  $x + 1$  sterben. Setzt man beispielsweise  $a_{x,t} = 1/2$ , geht in den Nenner für jede überlebende Person ein volles Personenjahr ein und für jede sterbende Person nur ein halbes. Gegeben  $r_{x,t}$  lassen sich  $q_{x,t}$  und  $p_{x,t}$  herleiten und von diesen ausgehend abgeleitete Kennwerte wie  $e_{x,t}$  (vgl. Preston et al. 2001: 42ff.).<sup>40</sup>

<sup>40</sup> Hierauf wird in Anhang A zurückgekommen.

Eine Periodensterbetafel, die ausgehend von den Daten eines Jahres  $t$  berechnet wurde, bezieht sich wie bereits erwähnt auf eine fiktive Kohorte, ähnlich wie bei der zusammengefassten Geburtenziffer. Beispielsweise kann die aus einer solchen Periodensterbetafel berechnete Lebenserwartung verstanden werden als die Lebenserwartung, die sich für eine Kohorte ergeben würde, die den Mortalitätsverhältnissen der betrachteten Periode unterliegt (vgl. Preston et al. 2001: 42). Im weiteren Verlauf der Arbeit werden lediglich Periodensterbetafeln verwendet. Diese werden aufgrund deutlicher Unterschiede in der Sterblichkeit zwischen Männern und Frauen getrennt nach Geschlechtern betrachtet (s. bspw. Abb. 3.3 auf S. 44).

### 3.1.4 Projektionen, Prognosen und Vorausberechnungen

Kennt man die auf der rechten Seite der Buchführungsgleichung (3.1) stehenden Größen, kann ausgehend von einem Bevölkerungsbestand zu  $t$  die Bevölkerungsgröße zu  $t + 1$  berechnet werden. Ist man bereit, Annahmen über die zukünftige Zahl an Geburten, Sterbefällen und Wanderungen zu treffen, kann auch die zukünftige Entwicklung einer Bevölkerung hergeleitet werden. Bei der Berechnung der zukünftigen Entwicklung unterscheidet man zwischen „Projektionen“ und „Prognosen“ (vgl. Hummel 2000: 281). Als Prognosen werden Berechnungen verstanden, die den Anspruch haben, die „tatsächliche“ zukünftige Entwicklung aufzuzeigen (vgl. ebd.: 281). Projektionen hingegen weisen einen deutlich geringeren Geltungsanspruch auf. Die einer Projektion zugrunde liegenden Annahmen über die zukünftige Entwicklung werden als rein hypothetisch aufgefasst, so dass die Ergebnisse der Projektion lediglich Entwicklungen aufzeigen, die sich bei zutreffen der Annahmen ergeben (vgl. ebd.: 281). Es handelt sich somit um „Wenn-Dann“-Aussagen, die zeigen, wie sich eine Bevölkerung entwickeln könnte (vgl. Keyfitz & Caswell 2005: 63). Damit verbunden ist in aller Regel kein Anspruch auf die Richtigkeit der Annahmen über die zukünftige Entwicklung.

Allerdings ist die Unterscheidung zwischen diesen beiden Varianten insbesondere bezogen auf die Interpretation von Ergebnissen zumeist nicht sehr trennscharf, da beispielsweise die hypothetischen Annahmen, die Projektionen zugrunde liegen, oftmals mit dem Anspruch verbunden sind, „wahrscheinliche“ oder zumindest denkbare Entwicklungen aufzuzeigen (s. auch S. 49 in Paragraph 4.1.1; vgl. Sebald 2008: 24). Im weiteren Verlauf dieser Arbeit wird anstelle dieser beiden Begriffe deshalb zumeist der Begriff „Vorausberechnung“ verwendet, der als neutraler Oberbegriff aufgefasst werden soll, unabhängig vom Geltungsanspruch.<sup>41</sup>

### 3.1.5 Die Kohorten-Komponenten-Methode

Wie demographische Vorausberechnungen allgemein durchgeführt werden können, wurde bereits im vorausgegangenen Unterabschnitt angedeutet. Ein formaler Rahmen, der eine detaillierte Beschreibung erlaubt, ist die sogenannte Kohorten-Komponenten-Methode, die auf

---

41 Laut amtlicher Statistik liefert „[e]ine Bevölkerungsvorausberechnung [...] Erkenntnisse darüber, wie sich die Bevölkerungsanzahl und deren alters- und geschlechtsspezifische Struktur bei Eintreffen der Annahmen zur künftigen Entwicklung der Geburtenhäufigkeit, der Lebenserwartung sowie der Wanderungen entwickeln würden.“ (Statistisches Bundesamt 2011c: 32). Dies lässt sich zwar im Sinne einer Projektion auffassen. Zugleich wird aber der Gültigkeitsanspruch der Annahmen nicht festgelegt. Insofern ist unklar, ob diese Definition mit der Verwendung des Begriffs „Vorausberechnung“ in dieser Arbeit übereinstimmt.

Cannan (1895), Bowley (1924) und Whelpton (1936) zurückgeht. Eine kompakte Formulierung dieses Ansatzes in Matrixnotation stammt von Leslie (1945), der von einer rein weiblichen Bevölkerung ausging (vgl. ebd.: 183).<sup>42</sup> Diese Formulierung sieht wie folgt aus (vgl. ebd.: 183f.; Caswell 2001: 10):<sup>43</sup>

$$\mathbf{n}_{t+1} = \mathbf{A}_t \mathbf{n}_t \quad (3.14)$$

mit

$$\mathbf{n}_t = (n_{0,t}, n_{1,t}, \dots, n_{\rho,t})' \quad (3.15)$$

und

$$\mathbf{A}_t = \begin{pmatrix} F_{0,t} & F_{1,t} & F_{2,t} & \dots & F_{\rho-1,t} & F_{\rho,t} \\ P_{0,t} & 0 & 0 & \dots & 0 & 0 \\ 0 & P_{1,t} & 0 & \dots & 0 & 0 \\ 0 & 0 & P_{2,t} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & P_{\rho-1,t} & 0 \end{pmatrix} \quad (3.16)$$

$\mathbf{n}_t$  ist ein  $(\rho+1 \times 1)$  Spaltenvektor, dessen Einträge  $n_{x,t}$  den Bevölkerungsbestand an Frauen im Altersintervall  $[x, x+1)$  zu  $t$  wiedergeben.<sup>44</sup> Die als „Leslie-Matrix“ bezeichnete  $(\rho+1 \times \rho+1)$  Matrix  $\mathbf{A}$  enthält altersspezifische Geburtenraten  $F_{x,t}$  sowie Überlebenswahrscheinlichkeiten  $P_{x,t}$  (vgl. Caswell 2001: 9f.). Genauer gibt  $P_{x,t}$  die Wahrscheinlichkeit an, dass eine Frau, die zu  $t$  im Altersintervall  $[x, x+1)$  ist, bis zu  $t+1$  überlebt und sich dann im Altersintervall  $[x+1, x+2)$  befindet (vgl. ebd.: 9).  $F_{x,t}$  gibt die Geburtenrate weiblicher Nachfahren für den Zeitraum von  $t$  bis  $t+1$  für Frauen an, die sich zu  $t$  im Altersintervall  $[x, x+1)$  befinden (vgl. ebd.: 10), wobei diese Geburtenraten für manche Alter auch gleich Null sein können.<sup>45</sup>

Die Bevölkerung zu  $t+1$  ergibt sich also aus dem Teil der Bevölkerung zu  $t$ , der bis zu  $t+1$  überlebt, sowie aus durch Geburten in die Bevölkerung hinzukommende Frauen. Dies entspricht im Wesentlichen der Logik der Buchführungsgleichung (3.1). Zwar ist Migration

42 Möglichkeiten für eine Erweiterung auf zwei Geschlechter werden in Abschnitt 4.5 diskutiert.

43 Im weiteren werden fett gesetzte Großbuchstaben für Matrizen und fett gesetzte Kleinbuchstaben für Vektoren verwendet. Einträge von Matrizen werden mit dem selben klein gesetzten Buchstaben wie die entsprechende Matrix bezeichnet. Für Einträge von Vektoren gilt entsprechendes. Eine Abweichung von dieser Konvention findet sich direkt in Formel (3.16). Dort werden unterschiedliche Großbuchstaben für die Einträge verwendet, um diese entsprechend ihrer unterschiedlichen inhaltlichen Interpretation unterscheiden zu können. Die Verwendung von Groß- anstelle von Kleinbuchstaben entspricht der in Fußnote 45 festgehaltenen Konvention.

44 Für die Transponierte eines Vektors  $\mathbf{x}$  bzw. einer Matrix  $\mathbf{X}$  wird im weiteren wie in (3.15) ein Häkchen verwendet, also  $\mathbf{x}'$  bzw.  $\mathbf{X}'$ .

45 Somit unterscheiden sich also  $P_{x,t}$  und  $F_{x,t}$  von  $p_{x,t}$  und  $f_{x,t}$ , die in Unterabschnitt 3.1.3 eingeführt wurden. Zwar beziehen sich alle diese Größen auf einen Zeitraum von einem Jahr, allerdings gelten  $p_{x,t}$  und  $f_{x,t}$  ab dem exakten Alter  $x$  bis ausschließlich zum Alter  $x+1$ , während  $P_{x,t}$  und  $F_{x,t}$  für beliebige einjährige Altersintervalle mit Anfang in  $[x, x+1)$  gelten sollen. Wie  $P_{x,t}$  und  $F_{x,t}$  aus  $p_{x,t}$  und  $f_{x,t}$  berechnet werden können, wird in 5.2.4 und 5.3.1 erklärt. Wie durch die verwendete Notation bereits gezeigt, werden für Wahrscheinlichkeiten und Raten, die sich auf beliebige Altersintervalle mit Anfang in  $[x, x+1)$  beziehen, Großbuchstaben verwendet und Wahrscheinlichkeiten und Raten, die sich auf exakte Alter  $x$  bzw. auf das Intervall  $[x, x+1)$  beziehen, werden mit Kleinbuchstaben bezeichnet.

noch unberücksichtigt, allerdings lässt sich diese leicht ergänzen (vgl. bspw. Alho & Spencer 2005: 186):

$$\mathbf{n}_{t+1} = \mathbf{A}_t \mathbf{n}_t + \mathbf{m}_t \quad (3.17)$$

mit

$$\mathbf{m}_t = (m_{0,t}, m_{1,t}, \dots, m_{\rho,t})' \quad (3.18)$$

$\mathbf{m}_t$  ist ein  $(\rho + 1, 1)$ -Spaltenvektor, dessen Einträge  $m_{x,t}$  den Nettoeffekt der Migration aus dem Zeitraum  $[t, t + 1)$  auf die Bevölkerung im Alter  $[x, x + 1)$  zu  $t + 1$  erfassen. Befindet sich zum Beispiel eine Person zu  $t$  im Altersintervall  $[x, x + 1)$  und wandert zwischen  $t$  und  $t + 1$  aus, „fehlt“ diese Person zu  $t + 1$  im Altersintervall  $[x + 1, x + 2)$ .<sup>46</sup>

Um eine Bevölkerung nun  $y$  Jahre vorauszuberechnen, benötigt man einen Ausgangsbestand der Bevölkerung  $\mathbf{n}_t$  sowie Leslie-Matrizen  $\mathbf{A}_t, \mathbf{A}_{t+1}, \dots, \mathbf{A}_{t+y-1}$  und Vektoren  $\mathbf{m}_t, \mathbf{m}_{t+1}, \dots, \mathbf{m}_{t+y-1}$ . Hierüber lässt sich durch eine wiederholte Anwendung von (3.14)  $\mathbf{n}_{t+y}$  berechnen (vgl. Caswell 2001: 11).

Die in diesem und den vorausgegangenen Unterabschnitten dargestellten Begriffe, Kennwerte und Konzepte stellen die Grundlage der weiteren Betrachtung dar. Andere demographische Begriffe und Verfahren werden im weiteren Verlauf an der Stelle ihrer ersten Nennung formal eingeführt, zumeist in Fußnoten.

## 3.2 Demographische Entwicklung in Deutschland

In diesem Abschnitt werden die Bevölkerungsentwicklung, die Entwicklung der Fertilität, der Mortalität, der Migration und der Eheschließungen und -scheidungen (Nuptialität) in der BRD seit den fünfziger Jahren überblicksartig skizziert. Eine Beschreibung der zugrunde liegenden Daten und deren Besonderheiten findet man in Anhang A.<sup>47</sup> Soweit nichts anderes angegeben wird, basieren alle im weiteren Verlauf der Arbeit vorgestellten Werte und Kennwerte auf den in diesem Anhang beschriebenen Daten.

### 3.2.1 Bevölkerungsentwicklung

In Abbildung 3.1 ist die Entwicklung des Bevölkerungsbestandes der BRD für die Jahre von 1956 bis 2009 dargestellt. Für den Zeitraum ab 1991 beziehen sich die Werte auf den seit dem 3.10.1990 geltenden Gebietsstand. Für die Jahre davor wurden die Bevölkerungsbestände der BRD und der DDR zusammengefasst. Alle Werte beziehen sich auf den 1. Januar eines Jahres.

<sup>46</sup> Wie  $m_{x,t}$  ermittelt werden kann, wird in Unterabschnitt 8.6.1 diskutiert.

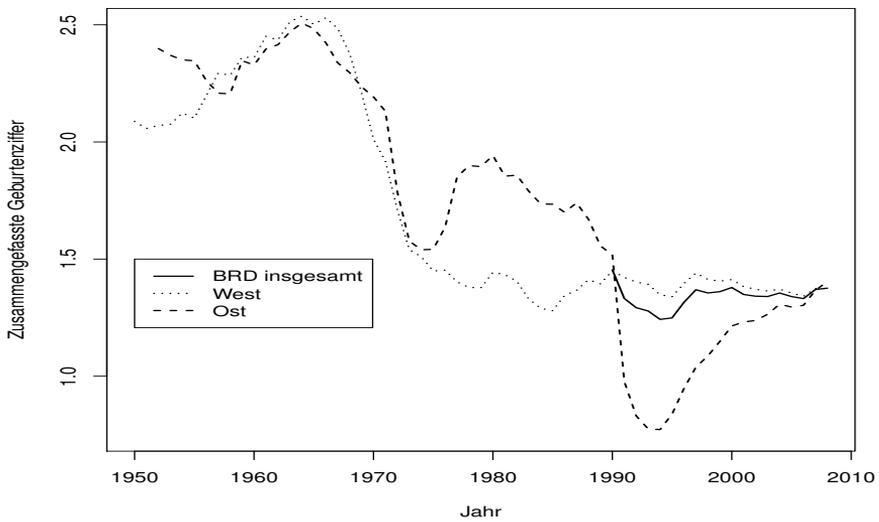
<sup>47</sup> Die Daten zur Nuptialität stellen eine Ausnahme dar und werden im Anhang nicht besprochen, da sie lediglich in diesem Unterabschnitt beschrieben werden, ansonsten aber keine große Rolle in dieser Arbeit spielen. Die Datenreihen wurden aus der GENESIS-Online-Datenbank des Statistischen Bundesamtes bezogen: <https://www-genesis.destatis.de/>.

Abbildung 3.1: Bevölkerungsentwicklung in der BRD (vor 1990 inkl. DDR)



Quelle: Human Mortality Database; eigene Darstellung

Abbildung 3.2: Entwicklung der zusammengefassten Geburtenziffer in West- und Ostdeutschland



Quelle: Statistisches Bundesamt; eigene Darstellung

1956 lag der Bevölkerungsbestand bei etwa 71.4 Millionen Personen. Bis 1971 fand ein relativ stetiger Anstieg auf 79.1 Millionen Personen statt. Anschließend ging die Bevölkerung bis 1986 mit leichten Schwankungen auf circa 77.7 Millionen Personen zurück. 1987 und 1988 nahm die Bevölkerung um jeweils etwa 120 Tausend Personen zu. Ab 1989 und etwa bis zur Mitte der neunziger Jahre stieg der Bevölkerungsbestand stark an. Beispielsweise war die Bevölkerung 1990 um etwa 722 Tausend Personen größer als 1989. Ab Mitte der neunziger Jahre verlangsamte sich das Bevölkerungswachstum deutlich. Seit dem Jahr 2004 sind die Bevölkerungsbestände verglichen mit den jeweiligen Vorjahresbeständen kontinuierlich zurück gegangen. 2009 lag der Bevölkerungsbestand bei knapp 82 Millionen Personen, wobei dies einer Abnahme um etwa 220 Tausend Personen im Vergleich zum Vorjahr entspricht. Dabei sind die historischen Zu- und Abnahmen der Bevölkerung jeweils auf Unterschiedliche Faktoren zurückzuführen und insbesondere durch die Entwicklung der Fertilität und der Migration gekennzeichnet.

### 3.2.2 Entwicklung der Fertilität

Die Entwicklung der zusammengefassten Geburtenziffer von 1950 bis 2008 ist in Abbildung 3.2 zu sehen.<sup>48 49</sup> Eingezeichnet ist die Entwicklung für Westdeutschland, die DDR beziehungsweise Ostdeutschland sowie ab 1990 für das wiedervereinigte Deutschland. Die Datenreihe für die DDR beginnt erst mit dem Jahr 1952. Für den ersten Teil der fünfziger Jahre zeigen sich zwar relativ deutliche Unterschiede zwischen Ost und West, allerdings verläuft die Entwicklung etwa ab dem Jahr 1957 bis etwa 1975 ungefähr parallel. Für diesen Zeitraum zeigt sich zunächst eine relativ deutliche Zunahme der zusammengefassten Geburtenziffer, wobei jeweils im Jahr 1964 der höchste Wert erreicht wird. Für Westdeutschland war dies eine zusammengefasste Geburtenziffer von etwa 2.54 und für die DDR von circa 2.51.<sup>50</sup> Hierauf folgte eine starke Abnahme dieser Kennziffer.<sup>51</sup> 1975 lag sie bei etwa 1.45 (West) beziehungsweise 1.54 (Ost). Dabei wird allgemein davon ausgegangen, dass ein Wert von mindestens 2.1 „bestandserhaltend“ ist (vgl. bspw. Espenshade et al. 2003). Seit 1975 lag die zusammengefasste Geburtenziffer in Westdeutschland mit Schwankungen konstant unter

48 Entsprechend der Ausführungen im vorausgegangenen Abschnitt beziehen sich die den zusammengefassten Geburtenziffern zugrunde liegenden Geburtenraten auf Frauen. Für Geburtenraten von Männern s. Abschnitt 6.2. Eine zusammenfassende Beschreibung der Entwicklung der Fertilität in West- und Ostdeutschland findet man unter anderem bei Dorbritz (2008) und bei El-Khorazaty (1997).

49 Die Entwicklung der absoluten Zahl der Geburten weist einen verglichen mit der zusammengefassten Geburtenziffer grob ähnlichen Verlauf auf, weshalb sie hier nicht gesondert dargestellt wird. Zu beachten ist allerdings, dass keine direkte Korrespondenz zwischen diesen beiden Größen besteht, da für die Zahl der Geburten auch die Zahl an Frauen im gebärfähigen Alter mit entscheidend ist, während dieser Faktor bei der zusammengefassten Geburtenziffer keine direkte Rolle spielt.

50 Zumindes im Wert für Westdeutschland zeigt sich allerdings die in Fußnote 37 auf Seite 34 genannte Schwäche der zusammengefassten Geburtenziffer: In der Literatur werden die hohen Werte der fünfziger und sechziger Jahre nicht nur auf ein Ansteigen der Geburtenhäufigkeit, sondern auch auf Veränderungen im *timing* von Geburten zurückgeführt (vgl. bspw. Rohwer & Pötter 2003: 223). Zur Entwicklung des durchschnittlichen Alters bei der Geburt, in dem sich solche Veränderung des *timing* teilweise widerspiegeln, siehe weiter unten im Text.

51 Mögliche Ursachen der starken Abnahme der Fertilität werden unter anderem von Dorbritz (2008), Klein & Lengsfeld (1985), Klein (1989) und Goldstein et al. (2003) diskutiert. Diese Abnahme ist dabei ein internationales Phänomen, welches auch in etlichen anderen Ländern beobachtet werden kann und in der Literatur als „zweiter demographischer Übergang“ (*second demographic transition*) diskutiert wird (Van de Kaa 1987; Lesthaeghe 1992, 2010). Eine umfangreiche, international vergleichende Diskussion findet sich bei Bujard (2011).

einem Wert von 1.5. Der niedrigste Wert wurde im Jahr 1985 erreicht und betrug circa 1.28. Der letzte in der Abbildung eingetragene Wert für Westdeutschland im Jahr 2008 war ungefähr gleich 1.37.

Die Entwicklung in der DDR verlief seit 1975 deutlich anders als im Westen. Zunächst zeigte sich bis 1980 ein Anstieg der zusammengefassten Geburtenziffer auf einen Wert von etwa 1.94. Anschließend folgte eine kontinuierliche Abnahme und im Jahr 1990 wurde ein Wert von 1.52 erreicht. 1991 betrug die zusammengefasste Geburtenziffer lediglich 0.98 und nahm in den folgenden Jahren nochmals ab, wobei im Jahr 1994 mit einem Wert von 0.77 ein extremer Tiefstand erreicht wurde.<sup>52</sup> Ab 1995 zeigte sich dann eine kontinuierliche Zunahme, wobei die zusammengefasste 2008 mit einem Wert von etwa 1.4 leicht über dem westdeutschen Niveau lag.<sup>53</sup> Die zusammengefasste Geburtenziffer für die gesamte, wiedervereinigte BRD verschleierte die drastischen Unterschiede zwischen West und Ost größtenteils und ist vor allem durch die Entwicklung im Westen geprägt.

Das durchschnittliche Alter bei der Geburt lag für westdeutsche Frauen 1950 bei 28.5 Jahren, für ostdeutsche Frauen 1952 bei 26.1 Jahren. Bis zur Mitte der siebziger Jahre sank das durchschnittliche Alter für westdeutsche Frauen bis auf Werte knapp unter 27 Jahren, gefolgt von einem kontinuierlichen Anstieg bis zum Jahr 2008 und einem durchschnittlichen Alter von 30.2 Jahren. Die Entwicklung in Ostdeutschland zeigt ebenfalls erst eine Ab- und dann eine Zunahme, wobei aber grundsätzlich niedrigere Werte realisiert wurden als im Westen. 2008 lag das durchschnittliche Alter bei der Geburt bei 29.1 Jahren, also ungefähr 1.1 Jahre niedriger als für westdeutsche Frauen.

Wie in Unterabschnitt 3.1.3 diskutiert, wird bei den bisher zur Fertilität dargestellten Kennwerten von einer hypothetischen Betrachtungsweise ausgegangen und hypothetische Kohorten als Grundlage der Auswertungen verwendet. Berechnet man die Kennziffern hingegen für Frauen einzelner Geburtsjahrgänge und damit „echter“ Kohorten, zeigt sich teils ein anderes Bild. So ergibt sich bei Berechnung der zusammengefassten Geburtenziffer für einzelne Geburtsjahrgänge insbesondere für Westdeutschland ein relativ kontinuierlicher Rückgang. Beispielsweise erreichten westdeutsche Frauen des Jahrgangs 1935 eine durchschnittliche Kinderzahl von etwa 2.17 Kindern, Frauen des Jahrgangs 1945 eine durchschnittliche Kinderzahl von 1.78 und Frauen des Jahrgangs 1955 eine durchschnittliche Kinderzahl von etwa 1.62. Beim durchschnittlichen Alter bei der Geburt zeigt sich wiederum zunächst eine Abnahme, gefolgt von einem Anstieg. Für den Jahrgang 1935 lag das durchschnittliche Alter bei 27.18 Jahren, für den Jahrgang 1945 bei 25.67 Jahren und für den Jahrgang 1955 bei 27.10 Jahren.<sup>54</sup>

52 Eine kritische Diskussion dieser Entwicklung findet man bei Kreyenfeld & Konietzka (2004: 4ff.). Folgt man diesen Autoren, ist die starke Veränderung der zusammengefassten Geburtenziffer zu Beginn der neunziger Jahre in nicht unerheblichem Maße auf Veränderungen im *timing* von Geburten zurückzuführen.

53 Inwieweit sich diese Entwicklung als Angleichung der ost- an die westdeutschen Verhältnisse interpretieren lässt, wird von Goldstein & Kreyenfeld (2011) diskutiert.

54 Die Berechnung der zusammengefassten Geburtenziffer für einen Geburtsjahrgang  $C$  erfolgte über

$$\text{TFR}_C = \sum_{x=\alpha}^{\beta} f_{x,C+x}$$

### 3.2.3 *Entwicklung der Mortalität*

Einen Überblick über die Entwicklung der Lebenserwartung von 1956 bis 2008 liefert Abbildung 3.3.<sup>55</sup> Dabei ist die Lebenserwartung unterteilt nach Geschlecht und West- und Ostdeutschland eingezeichnet. Für westdeutsche Frauen lag die Lebenserwartung im Jahr 1956 bei 70.9 Jahren. Bis 2008 zeigt sich ein kontinuierlicher Anstieg auf einen Wert von 82.4 Jahren. Dieser ist insbesondere auf eine Abnahme der Säuglingssterblichkeit zurückzuführen (vgl. Luy 2004: 10). Bei westdeutschen Männern zeigt sich praktisch die selbe Entwicklung, allerdings ausgehend von einem deutlich niedrigeren Wert als bei westdeutschen Frauen. 1956 lag die Lebenserwartung bei 65.9 Jahren und 2008 bei 77.6 Jahren. Dabei sind dies die beiden Jahre mit den niedrigsten Differenzen der Lebenserwartungen von Frauen und Männern mit Unterschieden von etwa 5 (1956) beziehungsweise 4.9 (2008) Jahren. Für die anderen Jahre fallen die Differenzen teils nicht unwesentlich höher aus. Beispielsweise wurde die höchste Differenz im Jahr 1980 erreicht und lag bei 6.8 Jahren.

Für ostdeutsche Frauen betrug die Lebenserwartung im Jahr 1956 wie für westdeutsche Frauen 70.9 Jahre. Bis etwa zur Mitte der siebziger Jahre verlief die Entwicklung der Lebenserwartung relativ parallel zur Entwicklung im Westen mit etwas niedrigeren Werten. Ab Mitte der siebziger Jahre aber nahm die Lebenserwartung in geringerem Umfang zu als im Westen. Dies führte dazu, dass ostdeutsche Frauen 1990 eine Lebenserwartung von 76.3 aufwiesen, während westdeutsche Frauen eine Lebenserwartung von 79 Jahren hatten. Seit den neunziger Jahren fand allerdings eine aufholende Angleichung der Lebenserwartung ostdeutscher Frauen an das westdeutsche Niveau statt. 2008 lag die Lebenserwartung ostdeutscher Frauen bei etwa 82.4 Jahren und die Differenz zu westdeutschen Frauen betrug nur noch knapp 0.1 Jahre. Für ostdeutsche Männer lässt sich größtenteils das selbe festhalten wie für ostdeutsche Frauen – es gab verglichen mit westdeutschen Männern erst eine ähnliche, dann eine auseinanderklaffende Entwicklung und schließlich seit der Wiedervereinigung eine Angleichung. Im Gegensatz zu Frauen ist diese Angleichung bisher allerdings nicht vollständig erfolgt. 2008 lag die Lebenserwartung ostdeutscher Männer bei etwa 76.5 Jahren, was circa 1.1 Jahre weniger sind als bei westdeutschen Männern.

### 3.2.4 *Entwicklung der Migration*

In der Abbildung 3.4 ist der Umfang der Einwanderung in und der Auswanderung aus der BRD von 1950 bis 2009 zu sehen.<sup>56</sup> Zu beachten ist, dass sich diese Daten vor dem Jahr 1990

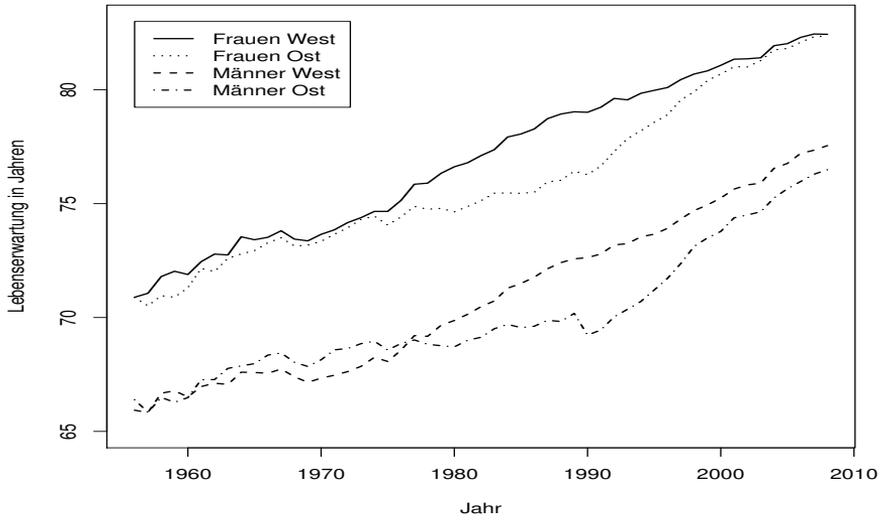
---

Dabei ist zu beachten, dass dies streng genommen nicht genau der Fertilität der Kohorte  $C$  entspricht, da sich  $f_{x,C+x}$  nicht nur über Frauen der Kohorte  $C$ , sondern auch über Frauen der Kohorte  $C - 1$  ergibt. Erstere sind zu Beginn des Jahres  $C + x$  im Alter  $x - 1$  und erreichen das Alter  $x$  erst, womit es noch Geburten im Alter  $x - 1$  geben kann, die nicht mitgezählt werden. Frauen des Jahrgangs  $C - 1$  sind zu Beginn des Jahres  $C + x$  im Alter  $x$  und werden im Laufe des Jahres  $x + 1$  Jahre alt, so dass deren Geburten im Alter  $x$  mit in  $TFR_C$  eingehen. Dinkel & Milenovic (1992) sprechen bei solch einem Vorgehen von einer Betrachtung von „Quasikohorten“.

55 Übersichtsdarstellungen zur historischen Entwicklung der Mortalität findet man beispielsweise bei Luy (2004) und Wiesner (2001). Speziell mit Kindersterblichkeit beschäftigen sich Ahmad et al. (2000). Kohortensterbetafeln findet man bei Dinkel (1992) und Statistisches Bundesamt (2006).

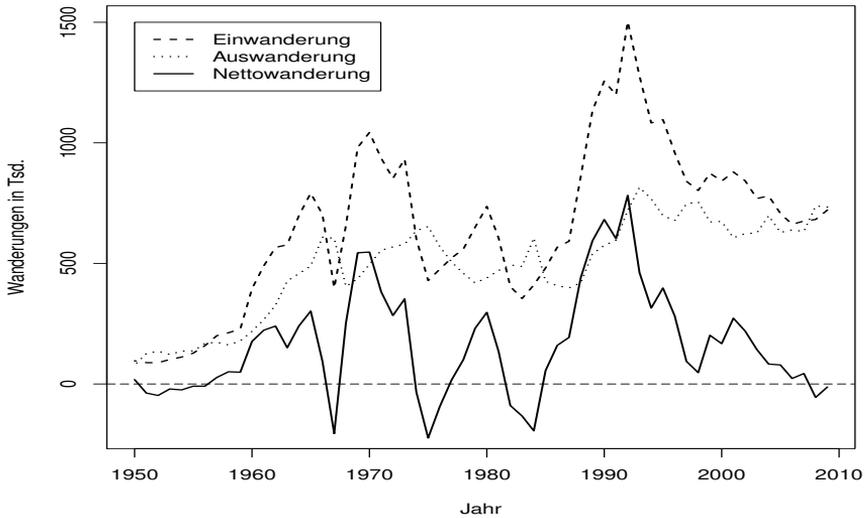
56 Eine Beschreibung der Geschichte der Einwanderung in Deutschland findet man bei Siebert (2003). Auswanderung wird unter anderem von Sauer & Ette (2007) diskutiert.

Abbildung 3.3: Entwicklung der Lebenserwartung von Männern und Frauen in West- und Ostdeutschland



Quelle: Human Mortality Database; eigene Darstellung

Abbildung 3.4: Einwanderung, Auswanderung und Nettowanderung in der BRD (vor 1990 ohne DDR)



Quelle: Statistisches Bundesamt; eigene Darstellung

ausschließlich auf Westdeutschland beziehen. Neben dem Umfang der Ein- und Auswanderung ist mit der Nettowanderung noch die Differenz aus diesen beiden Größen eingezeichnet. Zudem ist eine horizontale Linie ergänzt, die eine Nettomigration von 0 anzeigt.

Bei der Einwanderung zeigen sich insgesamt vier Einwanderungswellen: die erste ab Mitte der fünfziger Jahre bis 1966, die zweite von 1968 bis 1973, die dritte von der Mitte der siebziger bis zu Beginn der achtziger Jahre und schließlich die vierte Welle von Ende der achtziger bis zur Mitte der neunziger Jahre (vgl. Siebert 2003: 5ff.). Die Einwanderung vor der letzten Einwanderungswelle umfasste vor allem Arbeitsmigranten aus dem Mittelmeerraum und später Nachzüge ihrer Familienmitglieder und war stark durch die wirtschaftliche Entwicklung in der BRD geprägt (vgl. ebd.: 5f.). Die letzte und mit Abstand größte Einwanderungswelle speiste sich primär aus den ehemaligen Ostblockstaaten (vgl. ebd.: 7). Bei der Auswanderung zeigen sich nicht so deutliche Schwankungen wie bei der Einwanderung. Die auftretenden Veränderungen sind dabei zeitlich zu den Schwankungen der Einwanderung versetzt: auf jede der Einwanderungswellen folgten Rückwanderungen.

Insgesamt war der Nettoeffekt von Wanderungen auf die Größe der Bevölkerung zumeist deutlich positiv und weist nur für wenige Jahre ein negatives Vorzeichen auf. In den letzten Jahren des in Abbildung 3.4 dargestellten Zeitraums war die Nettomigration aber entweder vergleichsweise niedrig oder negativ. Beispielsweise betragen die Werte für die Jahre 2007, 2008 und 2009 der Reihe nach etwa +43 Tausend, -56 Tausend und -13 Tausend.

Betrachtet man die dargestellten Entwicklungen zusammen, lässt sich festhalten, dass der in Abbildung 3.1 dargestellte Anstieg der Bevölkerung bis zur Mitte der siebziger Jahre vor allem auf hohe Werte der Fertilität bei gleichzeitig relativ hoher Einwanderung zurückzuführen ist. Dass der Bevölkerungsbestand anschließend nicht aufgrund der niedrigen Fertilität gesunken, sondern um 1990 nochmals deutlich gestiegen ist, ist allein auf einen erheblichen Überschuss an Einwanderungen zurückzuführen. In Kombination mit dem deutlichen Anstieg der Lebenserwartung führte die niedrige Fertilität zudem zu einer deutlichen Alterung der Bevölkerung. 1960 betrug der Anteil von Personen im Alter von 65 oder mehr Lebensjahren etwa 11.5% der Gesamtbevölkerung. 1980 waren es 15.7%, im Jahr 2000 16.2% und schließlich 2008 etwa 20.3%.<sup>57</sup>

### 3.2.5 *Entwicklung der Nuptialität*

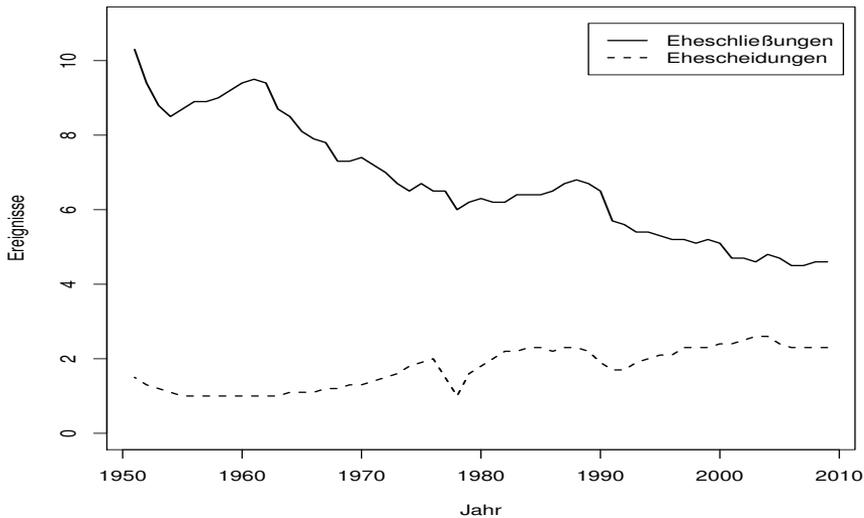
Ergänzend zu den bisherigen Ausführungen sind in Abbildung 3.5 grobe Indikatoren für die Entwicklung des Heiratsverhaltens abgebildet. Eingezeichnet ist die Zahl der Eheschließungen und Ehescheidungen jeweils pro 1000 Einwohner und für den Zeitraum von 1951 bis 2009, wobei die Daten die Entwicklungen in West- und Ostdeutschland zusammenfassen.

Bei der Zahl der Eheschließungen pro 1000 Einwohnern zeigt sich über den betrachteten Zeitraum von einzelnen Schwankungen abgesehen eine kontinuierliche Abnahme. 1951 lag der Wert bei 10.3, 2009 bei 4.6. Bei den Ehescheidungen wiederum zeigt sich eine Zunahme von 1.5 im Jahr 1951 auf 2.3 im Jahr 2009. Verbunden mit dieser Entwicklung ist eine deutliche Zunahme nicht-ehelicher Geburten: 1960 lag der Anteil nicht-ehelicher Geburten an allen Geburten bei etwa 8%, 2009 bei knapp 33% (West- und Ostdeutschland zusammengefasst).

---

<sup>57</sup> Eine ausführliche Diskussion der „demographischen Alterung“ der Bevölkerung findet man bei Mai (2005).

Abbildung 3.5: Entwicklung der Zahl der Eheschließungen und -scheidungen pro 1000 Einwohnern



Quelle: Statistisches Bundesamt; eigene Darstellung

Dabei gibt es aber erhebliche Unterschiede zwischen West- und Ostdeutschland. Im Westen lag der Anteil nicht-ehelicher Geburten 2009 bei ungefähr 26%, im Osten bei 61% (vgl. Kreyenfeld & Konietzka 2010: 8).<sup>58</sup>

Ausgehend von den dargestellten Entwicklungen der Fertilität, Mortalität, Nuptialität und Migration stellt sich die Frage nach der zukünftigen Entwicklung dieser Bevölkerungsbewegungen und der sich hieraus ergebenden Größe und Altersstruktur der Bevölkerung. Ebenso stellt sich die Frage, inwieweit sich die zukünftige Entwicklung auf das Vorhandensein von Verwandtschaft auswirkt.

### 3.3 Demographischer Wandel und Verwandtschaft in der wissenschaftlichen Diskussion

Ausgehend von der im vorausgegangenen Abschnitt beschriebenen demographischen Entwicklung finden sich in der Fachliteratur zahlreiche Spekulationen über die Auswirkungen auf das Vorhandensein von Verwandtschaft und hiervon ausgehend auch Überlegungen über die zukünftige Entwicklung von Verwandtschaft. Beispielsweise konstatiert Nave-Herz (2001: 25) bezogen auf den historischen Anstieg der Lebenserwartung:

<sup>58</sup> Zur historischen Entwicklung des Anteils nicht-ehelicher Geburten getrennt für Ost und West vgl. Kreyenfeld & Konietzka (2010) u. Rohwer & Pötter (2003: 172ff.). Zur Entwicklung der „Heiratsneigung“ s. Heilig (1985). Ferner ist zu beachten, dass der hier gewählte Indikator für die Entwicklung des Heiratsverhaltens zwei Entwicklungen umfasst: zum einen Veränderungen der grundsätzlichen Neigung, zu heiraten, und zum anderen Veränderungen im Timing von Heiraten (vgl. Schneider & Rüger 2007: 132f.). Dies ähnelt dem in Fußnote 37 auf Seite 34 beschriebenen Problem der zusammengefassten Geburtenziffer.

Diese Verlängerung der Lebenserwartung ließ die Mehrgenerationen-Familie zu einem Massenphänomen werden; denn sie ermöglicht heutzutage die Chance des Sich-Gegenseitigen-Erfahrens von mindestens drei (bei Frauen zunehmend: bis zu fünf) Familiengenerationen.

Bei Bengtson (2001: 6) findet sich die Überlegung, dass der Umfang der Verwandtschaft zwar durch geringe Fertilität abnimmt, zugleich aber durch eine geringe Sterblichkeit die geteilte Lebenszeit von Generationen zunimmt. Hierdurch sei von einer zunehmenden Bedeutung von intergenerationalen Beziehungen auszugehen (vgl. ebd.: 2). Eine ähnliche Ansicht vertreten Blome et al. (2008: 24f.):

Mit der niedrigen Kinderzahl schrumpft innerhalb der Familien die Zahl der Seitenverwandten, während die Generationen für längere Lebensphasen zusammenleben. Damit werden die Familien „länger und dünner“.

Bezogen auf die Konsequenzen einer solchen Entwicklung merkt Brandt (2009: 13) an: „Die Generationenforschung gewinnt innerhalb der Diskussion um die Zukunft des Wohlfahrtsstaats rasant an Bedeutung, kündigen sich doch im Zuge der Bevölkerungsalterung in den meisten europäischen Staaten weitreichende Veränderungen im Generationengefüge an.“

Die von den genannten Autoren beschriebene Entwicklung wird als „Vertikalisierung“ bezeichnet und in weiten Teilen der Literatur als gegeben hingenommen (vgl. Puur et al. 2011: 138).<sup>59</sup> Dabei gibt es allerdings kaum Arbeiten zur historischen Entwicklung und zur gegenwärtigen Situation (vgl. Puur et al. 2011: 138; Herlofson & Hagestad 2011). Die für die BRD vorliegenden Arbeiten zeigen zwar durchaus einen Anstieg der gemeinsamen Lebenserwartung von Generationen (Lange & Lauterbach 1998; Lauterbach 1995, 2002), allerdings sind Vier- oder gar Fünf-Generationen-Familien nichtsdestotrotz extrem selten (Farkas & Hogan 1995; Puur et al. 2011). Eine historische Studie für die Niederlande von Post et al. (1997) zeigt zudem keinen einfachen, linearen Trend hin zu einer Vertikalisierung, sondern vielmehr ein historisches „Auf und Ab“. Diese Befunde widersprechen der Vertikalisierungshypothese zwar nicht, verdeutlichen aber, dass allzu einfache Spekulationen nicht unbedingt eine treffende Beschreibung liefern.

Für die zukünftige Entwicklung liegen bis auf eine Ausnahme (Grünheid & Scharein 2011) keine aktuellen demographischen Vorausberechnungen vor, um die weiter oben genannten Überlegungen gegenzuprüfen.<sup>60</sup> Einige Schwierigkeiten lassen sich jedoch unabhängig hiervon benennen. Betrachtet man beispielsweise die geteilte Lebenszeit von Großeltern und ihren Enkelkindern, ist wie in Kapitel 1 bereits angedeutet nicht unbedingt von einem Anstieg auszugehen, auch wenn man eine zukünftige Zunahme der Lebenserwartung annimmt. Denn einerseits kann ein Anstieg der Lebenserwartung zwar die geteilte Lebenszeit verlängern, gleichzeitig kann aber von einem Anstieg des Alters bei der Geburt ein gegenläufiger Effekt ausgehen (vgl. Murphy et al. 2006: 220; Watkins et al. 1987: 346). Vorausberechnungen von Grünheid & Scharein (2011) deuten auf einen solchen Effekt hin. Die Autoren kommen dabei zu folgendem Schluss: „[...] it becomes evident that it is not sufficient to conclude from the

59 Überlegungen zu Effekten des demographischen Wandels auf laterale und entfernte Verwandtschaft findet man bei Jakoby (2008a: 215f.).

60 Neben der Studie von Grünheid & Scharein (2011) gibt es noch eine ältere Arbeit von Galler (1990) sowie eine Arbeit von Dudel (2008), die aber von sehr einschränkenden Annahmen ausgeht. Daneben gibt es zwar weitere Arbeiten neueren Datums, allerdings gehen diese in aller Regel vom Haushaltskontext aus und erfassen keine Verwandtschaft außerhalb von Haushalten. Hierauf wird in Kapitel 4 zurückgekommen.

longer life expectancy a generally longer joint lifetime of the generations [...]“ (ebd.: 72). Daneben ist festzuhalten, dass Vermutungen über die zukünftige Entwicklung oftmals nur vage und ohne genaue Angaben zum zeitlichen Verlauf sowie ohne Angaben zum genauen Umfang von Veränderungen geäußert werden (vgl. Uhlenberg 1993: 233).

Nimmt man die im vorausgegangenen Absatz angedeutete Unsicherheit bezüglich des Effekts der demographischen Entwicklung, den Mangel an aktuellen Vorausberechnungen sowie die in Kapitel 2 beschriebene Bedeutung von Verwandtschaft zusammen, zeigt sich ein Bedarf an einer aktuellen Vorausberechnung. Möglichkeiten, wie Vorausberechnungen von Verwandtschaft durchgeführt werden können, werden im nächsten Kapitel diskutiert.

## 4 Ansätze zur Vorausberechnung von Verwandtschaft

### 4.1 Allgemeine Probleme demographischer Vorausberechnungen

Bei Bevölkerungsvorausberechnungen gibt es etliche potentielle Fehlerquellen. Diese zeigen sich grundsätzlich bei demographischen Vorausberechnungen jeglicher Art, womit sie auch die Vorausberechnung von Verwandtschaft betreffen. Die in der Literatur diskutierten Fehlerquellen lassen sich grob einer der folgenden drei Ursachen von Unsicherheit zuordnen (vgl. Tuljapurkar 1992: 385):<sup>61</sup>

1. Unsicherheit bezüglich der zukünftigen Entwicklung der demographischen Komponenten und gegebenenfalls weiterer Parameter
2. Unsicherheit von Ausgangsgrößen
3. Unsicherheit bezüglich der den verwendeten Verfahren zugrunde liegenden Annahmen

Diese drei Bereiche sollen im folgenden erläutert werden. Diese Erläuterungen sollen als Grundlage für die Bewertung verschiedener Verfahren zur Vorausberechnung von Verwandtschaft dienen, die in den folgenden Abschnitten dieses Kapitels vorgestellt werden.

#### 4.1.1 Unsicherheit bezüglich der Entwicklung demographischer Komponenten

Vergleicht man Ergebnisse älterer Bevölkerungsvorausberechnungen mit der tatsächlich eingetretenen Bevölkerungsentwicklung, zeigt sich, dass oftmals große Abweichungen vorliegen, die auf falschen Annahmen zur Entwicklung der Komponenten basieren (vgl. Bretz 2002; Grummer-Strawn & Espenshade 1991; Keyfitz 1981; Stoto 1983). Beispielsweise wurde bei in den vierziger und fünfziger Jahren durchgeführten Vorausberechnungen der Bevölkerungsentwicklung in den USA der deutliche Anstieg der Fertilität ab dem Ende der fünfziger Jahre nicht vorhergesehen, obwohl es bereits zum Zeitpunkt der Erstellung der Vorausberechnungen erste Anzeichen für einen solchen Anstieg gab (Alho vgl. 1997: 73; für Deutschland s. Bretz 2002). Dies unterstreicht, dass die zukünftige Entwicklung demographischer Komponenten und aus dieser Entwicklung abgeleitete Ergebnisse mit einer großen Unsicherheit behaftet sind (vgl. Lee 1998: 156).

Dies ist mit der wichtigste Grund, warum oftmals von demographischen Projektionen und nicht von Prognosen gesprochen wird. Hierdurch wird das Problem allerdings nicht behoben, da davon ausgegangen werden kann, dass Nutzer demographischer Vorausberechnungen, wie beispielsweise politische Entscheidungsträger, Projektionen als Prognosen auffassen (vgl. Ahlburg & Lutz 1998: 2; Ryder 1990: 433). Zum anderen sind auch die Annahmen, die bei Projektionen bezüglich der Entwicklung der demographischen Komponenten getroffen werden, oftmals aufgrund von Plausibilitätsüberlegungen gewählt und stellen Entwicklungen dar, die die Ersteller der Projektion als „wahrscheinlich“ einschätzen (vgl. Ryder 1990: 433).

---

61 Eine alternative Einteilung findet man bei Lee (1998: 157f.).

Keyfitz & Caswell (2005: 271) merken dementsprechend an: „Insofar as the assumptions of a projection are realistic, it is indeed a forecast.“

Es stellt sich somit die Frage, inwiefern ein Vorausberechnungsverfahren die Unsicherheit bezüglich der zukünftigen Entwicklung der demographischen Komponenten und gegebenenfalls weiterer benötigter Parameter direkt berücksichtigt. Ist dies nicht der Fall – wie beispielsweise bei der Kohorten-Komponenten-Methode, bei der von gegebenen demographischen Raten ausgegangen wird – muss überlegt werden, ob sich ergänzende Verfahren finden lassen, die Unsicherheit berücksichtigen und mit der eigentlichen Vorausberechnung kombiniert werden können. Diese Verfahren wiederum müssen auch auf potentielle Fehlerquellen untersucht werden.

#### *4.1.2 Unsicherheit der Ausgangsgrößen*

Neben Annahmen über die zukünftige Entwicklung gehen auch bereits vorliegende Daten mit in Vorausberechnungen ein. Ein Beispiel ist der Bevölkerungsvektor für das Ausgangsjahr, der bei der Kohorten-Komponenten-Methode benötigt wird. Ist dieser Bevölkerungsvektor allerdings fehlerbehaftet, beispielsweise weil er nicht genau bekannt ist und geschätzt werden muss oder weil Probleme bei der Datenerhebung vorliegen, so ist auch nicht davon auszugehen, dass vorausberechnete Bevölkerungsvektoren die zukünftige Entwicklung treffend wiedergeben, selbst wenn die Entwicklung der demographischen Komponenten richtig vorausgeschätzt wird (vgl. Lee 1998: 158). Daneben werden eventuell auch historische Daten verwendet, die zwar nicht direkt bei der Vorausberechnung, aber für eine Abschätzung der zukünftigen Entwicklung demographischer Komponenten benötigt werden. Beispielsweise könnte die historische Entwicklung der Lebenserwartung verwendet werden, um die zukünftige Entwicklung der Lebenserwartung einzuschätzen. Auch in einem solchen Fall muss betrachtet werden, ob die Daten fehlerbehaftet sind.

Wenn Daten fehlerbehaftet sind, sind die eigentlich interessierenden Größen nicht genau bekannt und je nach Ausmaß der Fehler nur in gewissen Grenzen abschätzbar, so dass auch hier von Unsicherheit gesprochen werden kann. Bei der Bewertung verschiedener Vorausberechnungsverfahren stellt sich hier zunächst die Frage, welche Daten benötigt werden und ob diese überhaupt vorliegen. Anschließend ist zu klären, wie die „Qualität“ der verwendeten Daten eingeschätzt werden kann.

#### *4.1.3 Unsicherheit bezüglich der den verwendeten Verfahren zugrunde liegenden Annahmen*

Dieser Problembereich umfasst eine Vielzahl von möglichen Annahmen, die bei Vorausberechnungen getroffen werden, um diese zu vereinfachen, und deren Gültigkeit in Frage gestellt werden kann. Allerdings ist oft unklar, wie stark beziehungsweise häufig Verletzungen der Annahmen auftreten und wie stark ihr Einfluss auf Ergebnisse ist, so dass bezüglich der Annahmen beziehungsweise ihrer genauen Folgen Unsicherheit vorliegt.

Etlliche Annahmen werden nur für bestimmte Verfahren getroffen. Deshalb werden in den folgenden Absätzen nur solche Annahmen diskutiert, bei denen davon ausgegangen werden

kann, dass sie zumindest für die meisten Ansätze zutreffen:

- *female dominance*
- Homogenität
- Unabhängigkeit von Personen

Ein zentrales Problem demographischer Vorausberechnungen ist das sogenannte „Zwei-Geschlechter-Problem“ (vgl. Sebold 2008). Betrachtet man beispielsweise die Kohorten-Komponenten-Methode, wird wie erwähnt von einer vollständig weiblichen Bevölkerung ausgegangen. Zwar lassen sich Männer ohne weiteres im Modell berücksichtigen, allerdings ist die Frage, wie dies die Geburtenhäufigkeit beeinflussen soll. Eine einfache und oft genutzte Möglichkeit besteht darin, weiterhin lediglich die Fertilität von Frauen zu betrachten, so dass die Annahme getroffen wird, dass Männer keinen Einfluss auf die Zahl an Geburten haben. Dies wird als *female dominance* Ansatz bezeichnet (vgl. Sebold 2008: 33) und bedeutet insbesondere, dass keine Veränderungen der Fertilität erfasst werden können, die sich aufgrund von Veränderungen des quantitativen Geschlechterverhältnisses in der Bevölkerung ergeben (vgl. Schoen 1988a: 121). Sollen bei der Modellierung von Fertilität Männer mit berücksichtigt und die gerade genannte Annahme aufgehoben werden, ergibt sich zum einen das Problem, dass für Männer in aller Regel keine oder kaum Daten zur Fertilität vorliegen (vgl. bspw. Poston 2006: 871). Zum anderen muss nun darauf geachtet werden, dass die Zahl der Geburten, die Frauen zugeordnet werden, der Zahl der Geburten entsprechen muss, die Männern zugeordnet werden, was kein einfach zu lösendes Problem darstellt (vgl. bspw. Keilman 1985).

Des Weiteren wird bei Vorausberechnungen davon ausgegangen, dass sich Mitglieder der betrachteten Bevölkerung, die die selben Merkmalsausprägungen aufweisen, nicht systematisch voneinander unterscheiden. Personen im selben Zustand werden somit als homogene Gruppe aufgefasst (vgl. Heckman & Singer 1982: 567). Dies kann sich als problematisch erweisen, wenn nicht berücksichtigte Merkmale einen Einfluss auf die demographischen Komponenten haben (vgl. bspw. Vaupel et al. 1979). Werden eigentlich relevante Merkmale nicht berücksichtigt, spricht man von nicht berücksichtigter Heterogenität (vgl. bspw. Galler 1995; Heckman & Singer 1982). Allgemeiner formuliert handelt es sich um ein Problem des Aggregationsniveaus bei der Betrachtung (vgl. bspw. de Beer 2000: 4ff.), da alle Merkmalsausprägungen eines nicht berücksichtigten Merkmals gewissermaßen zusammengefasst, mithin also aggregiert betrachtet werden. Dabei wird allerdings nicht nur von Aggregation gesprochen, wenn Merkmale zusammengefasst werden, sondern auch, wenn dies mit Ereignissen geschieht. Denn auch demographische Ereignisse wie Geburten oder Sterbefälle können weiter unterteilt werden, beispielsweise nach der Nummer der Geburt (Parität) oder nach Todesursache (vgl. ebd.: 9).

Aggregation beziehungsweise nicht berücksichtigte Heterogenität kann in mehrerlei Hinsicht problematisch sein. Zunächst liefern hoch aggregierte Vorausberechnungen weniger Informationen als tiefer gegliederte. Gleiches gilt auch für die Analyse von historischen Entwicklungen, die im Extremfall irreführend sein können. Betrachtet man beispielsweise eine Bevölkerung, die sich in zwei Gruppen mit jeweils konstanter Mortalität auf unterschiedlichem Niveau unterteilen lässt, und nimmt im betrachteten Zeitraum der Anteil der Gruppe mit der höheren Mortalität ab, entsteht der Eindruck, dass die Lebenserwartung in der Bevölkerung gestiegen sei, obwohl sich an den Sterberisiken der Individuen nichts geändert hat (vgl. bspw. Keyfitz & Caswell 2005: 482f.). Hiervon ausgehend kann nicht berücksichtigte Heterogenität

in vielerlei Hinsicht verzerrend auf Ergebnisse von Vorausberechnungen wirken, beispielsweise indem die fälschlicherweise angenommene Zunahme der Lebenserwartung extrapoliert wird.<sup>62</sup>

Ein drittes Problem ist, dass bei Vorausberechnungen oft angenommen wird, dass die Personen in der betrachteten Bevölkerung als voneinander „unabhängig“ aufgefasst werden können (vgl. bspw. Ruggles 1990: 26). Dies meint, dass die Ereignisse, die einem Mitglied der Bevölkerung widerfahren, keinen Einfluss auf andere Personen haben. Ein Beispiel ist die Annahme, dass die Fertilität einer Person nicht von ihrer Geschwisterzahl und somit der Fertilität ihrer Eltern abhängt (vgl. Ruggles 1993: 162). Allgemein wird in der Literatur zuweilen auch von einer fehlenden Berücksichtigung einer „Korrelation zwischen Personen“ gesprochen (für entsprechende Formulierungen s. bspw. Pullum & Wolf 1991). Wenn es beispielsweise einen Zusammenhang zwischen Fertilität und Geschwisterzahl gibt, dieser aber nicht berücksichtigt wird, läuft dies auf eine Verletzung der Homogenitätsannahme hinaus. Denn durch die fehlende Berücksichtigung des genannten Zusammenhangs wird unterstellt, dass es keine systematischen Unterschiede zwischen Personen in der Fertilität gibt, obwohl diese eigentlich vorliegen.

Beim Vergleich verschiedener Vorausberechnungsansätze muss hiervon ausgehend betrachtet werden, welche der genannten Annahmen getroffen werden (beziehungsweise in welchem Ausmaß) und welche zusätzlichen Annahmen hinzu kommen. Sollte ein Ansatz eine der genannten Annahmen nicht aufweisen, ist zu überprüfen, inwieweit dies durch andere Annahmen erreicht wird, die an die Stelle der ursprünglichen Annahme treten.

#### 4.1.4 Weitere allgemeine Probleme

Neben den genannten Problemen gibt es allgemein auftretende „Fehlerquellen“, bei denen unklar ist, ob sie überhaupt beseitigt werden können und müssen, und die im weiteren nicht berücksichtigt werden. Hierzu zählt erstens, dass Ergebnisse von Modellen nie exakt mit der Realität übereinstimmen werden, da es sich bei diesen um Näherungen handelt (vgl. Alho & Spencer 1985: 306). Dies bezieht sich auf die Idee, dass es praktisch nicht möglich ist, in demographischen Modellen alle Faktoren zu berücksichtigen, die Einfluss auf die interessierenden Prozesse haben können (vgl. ebd.: 306). Insofern sind Abweichungen zwischen Modellrechnungen und Realität unvermeidbar. Zweitens scheint es kaum möglich, extreme Ereignisse wie beispielsweise große Katastrophen, Seuchen oder Kriege vorherzusehen und deren Einfluss auf die demographische Entwicklung zu berücksichtigen (vgl. Lee 1998: 159).<sup>63</sup> Drittens ist zu bedenken, dass sich Vorausberechnungen als „selbsterstörende Prophezeiung“ erweisen können: Weist eine Vorausberechnung beispielsweise auf eine wie auch immer gartete ungünstige Entwicklung hin, werden vielleicht Maßnahmen ergriffen, diese Entwicklung abzuwenden, wodurch sie möglicherweise verhindert wird (vgl. de Beer 2000: 25).

62 Für Matrizenmodelle wie die Kohorten-Komponenten-Methode nach Leslie findet sich eine formale Diskussion von Aggregation bei Rogers (1969, 1975).

63 Es gibt Ausnahmen, bei denen versucht wird, solche Faktoren zu berücksichtigen. Beispiele sind die Arbeiten von Lee & Carter (1992) und Keilman & Pham (2004). Allerdings laufen zum Beispiel die Annahmen von Keilman & Pham (2004: 8f.) darauf hinaus, dass ungefähr alle 50 Jahre mit einem Weltkrieg zu rechnen ist. Dies verweist auf das grundlegende Problem, Wahrscheinlichkeiten für extreme Ereignisse in irgendeiner Art und Weise zu quantifizieren.

#### 4.1.5 Ausblick

Wenn bei einem Ansatz zur Vorausberechnung von Verwandtschaft einige oder alle der genannten Probleme vorliegen, stellt sich die Frage, inwieweit dies einen verfälschenden Einfluss auf die Ergebnisse haben kann. Neben formal-mathematischen Überlegungen kann diese Frage auch durch empirische Validierungen überprüft werden (vgl. Wilson & Rees 2005: 338f.). Dabei soll unter dem Begriff „Validierung“ allgemein verstanden werden, dass eine Vorausberechnungsmethode anhand von empirischen Daten dahingehend überprüft wird, ob sie realistische Ergebnisse liefert (vgl. Wachter et al. 1997: 103).

Bei der in den nächsten Abschnitten dieses Kapitels folgenden Darstellung verschiedener Verfahren zur Vorausberechnung von Verwandtschaft soll hiervon ausgehend auf drei Dinge geachtet werden:

1. Ob und in welcher Form die in diesem Abschnitt diskutierten Probleme und Annahmen vorliegen;
2. Ob weitere, verfahrensspezifische Probleme oder Annahmen vorliegen und ob letztere problematisch sind;
3. Ob empirische Validierungen vorliegen, die angeben können, wie stark sich diese auf die Ergebnisse auswirken;

## 4.2 Varianten der Vorausberechnung von Verwandtschaft

Bevor die einzelnen, in der Literatur zu findenden Verfahren zur Vorausberechnung von Verwandtschaft diskutiert werden, soll zunächst eine kurze Systematik dieser Verfahren gegeben werden. In der Literatur werden drei Arten unterschieden (vgl. Brass 1983: 38; DeVos & Palloni 1989: 178f.; Ruggles 1990: 23; Imhoff & Post 1998: 99):

- Analytische Modelle
- Makrosimulationen
- Mikrosimulationen

Analytische Modelle geben explizite Formeln an, über die ein Zusammenhang zwischen demographischen Komponenten und Verwandtschaftsstruktur hergestellt werden kann (vgl. Ruggles 1990: 24). Zur Gruppe der Makrosimulationen werden Verfahren gezählt, die die Kohorten-Komponenten-Methode erweitern (vgl. ebd.: 24). Mikrosimulationen schließlich gehen von einzelnen Individuen aus, deren Lebensverläufe über zufällige Zuweisungen von demographischen Ereignissen fortgeschrieben werden (vgl. ebd.: 24f.).

Mit den einzelnen Arten von Ansätzen lassen sich sehr unterschiedliche Ergebnisse erzeugen, die sich zugespißt wie folgt charakterisieren lassen. Mikrosimulationen liefern Informationen auf Individualebene. Makrosimulationen geben Auskunft über absolute und relative Verteilungen von Merkmalen. Analytische Modelle schließlich heben primär auf ausgewählte Kennwerte zur Beschreibung von Verteilungen ab. Dabei ist aber auch zu beachten, dass die Trennung zwischen diesen drei Verfahrensgruppen teils eher analytischer Natur ist.

Im weiteren Verlauf dieses Kapitels werden vier Gruppen von Ansätzen dargestellt:

1. Verzweigungsprozesse
2. Bevölkerungsmathematische Ansätze
3. Multi-State Ansätze
4. Mikrosimulationen

Die beiden zuerst genannten bilden zusammen die Gruppe der analytischen Verfahren. Sie werden dennoch getrennt voneinander behandelt, da sie von unterschiedlichen Annahmen ausgehen. Multi-State Ansätze umfassen die oben genannten Makrosimulationen.

Bei den einzelnen Ansätzen werden zunächst jeweils die theoretischen Grundlagen erläutert, die sich nicht unbedingt auf die Modellierung von Verwandtschaft beziehen. Anschließend wird jeweils die konkrete Anwendung auf Verwandtschaft vorgestellt. Wenn für einen Ansatz auf die BRD bezogene Ergebnisse vorliegen, werden diese als nächstes diskutiert. Abschließend erfolgt jeweils eine Diskussion, bei der auf die Ausführungen des vorausgegangenen Abschnitts zurückgekommen wird. Nach der Vorstellung der einzelnen Ansätze folgt in den letzten beiden Abschnitten dieses Kapitels ein Fazit sowie eine Eingrenzung der Fragestellung der Arbeit.

## 4.3 Verzweigungsprozesse

### 4.3.1 Grundlagen

Die theoretischen und analytischen Grundlagen zur Untersuchung von Verzweigungsprozessen wurden von Galton und Watson ausgearbeitet (vgl. Galton & Watson 1977).<sup>64 65</sup> Ausgangspunkt war folgende Problemstellung (ebd.: 401):

A large nation, of whom we will concern ourselves with the adult males,  $N$  in number, and who each bear separate surnames, colonise a district. Their law of population is such that, in each generation,  $a_0$  per cent of the adult males have no male children who reach adult life;  $a_1$  have one such male child;  $a_2$  have two; and so on up to  $a_5$  who have five. Find (1) what proportion of the surnames will have become extinct after  $r$  generations; and (2) how many instances there will be of the same surname being held by  $m$  persons.

Das Interesse lag hierbei vor allem auf Adelsfamilien, deren Titel an männliche Nachfahren vererbt wurden (vgl. ebd.: 400).

Die Zahl der Kinder einer Person wird als Zufallsvariable aufgefasst (vgl. Haccou et al. 2005: 5), wobei diese im weiteren Verlauf mit  $K$  bezeichnet wird und Werte  $k \in \mathbb{N}$  annehmen kann. Ferner wird davon ausgegangen, dass sich aufeinander folgende Generationen nicht überlappen (vgl. Galton & Watson 1977: 402). Geht man zum Beispiel von einem Individuum in einer ersten Generation aus, welches einen Nachfahren in einer zweiten Generation hat, beträgt die Häufigkeit des Nachnamens in beiden Generationen jeweils eins. Zeit wird in

64 engl.: *branching processes*; allgemeine Einführungen in Theorie und Anwendung von Verzweigungsprozessen sind zu finden bei Haccou et al. (2005), Harris (1963) und Mode (1971).

65 Die hier angegebene Quelle ist ein Nachdruck der ursprünglichen Arbeit: Galton & Watson (1874). Die Fragestellung wurde von Galton bereits 1873 in der Zeitschrift *Educational Times* veröffentlicht (vgl. Hull 1998: 105).

diesem Modell nur implizit über die Abfolge von Generationen ausgedrückt und als diskret aufgefasst (vgl. Waugh 1981: 631).<sup>66</sup> Schließlich gelten die Wahrscheinlichkeiten  $a_1, \dots, a_5$  für die gesamte betrachtete Bevölkerung und für alle Generationen (vgl. Galton & Watson 1977: 402). Dies umfasst auch, dass sich die Mitglieder einer Generation oder aufeinander folgender Generationen nicht beeinflussen, so dass Unabhängigkeit zwischen diesen besteht (vgl. Harris 1963: 3). Der so allgemein definierte stochastische Prozess wird zumeist als Galton-Watson-Prozess bezeichnet und kann als Markov-Kette aufgefasst werden:<sup>67</sup> wenn  $Z_0, Z_1, Z_2 \dots$  die Abfolge der Größe von Generationen ist, wobei  $Z_0$  die Größe der Ausgangsgeneration ist, dann hängt die Wahrscheinlichkeitsverteilung der Größe der Generation  $Z_{j+1}$  nur von der Größe der vorausgegangenen Generation  $Z_j$  ab, so dass der Prozess die Markov-Eigenschaft aufweist (vgl. ebd.: 3). Wohlgermerkt gibt es etliche Erweiterungen der ursprünglichen Idee von Galton und Watson, die zusammen die Gruppe der Verzweigungsprozesse bilden, von denen der Galton-Watson-Prozess eine spezielle Variante ist. Eine weitere, allgemeinere Variante wird in Abschnitt 5.2 eingeführt.

Erweitert man die Zahl der möglichen Nachfahren von maximal fünf auf einen nicht näher bestimmten, maximalen Wert  $n \in \mathbb{N}$ , lassen sich die einzelnen Wahrscheinlichkeiten  $a_1, \dots, a_n$  in einer wahrscheinlichkeitserzeugenden Funktion zusammenfassen (vgl. Feller 1977: 408):

$$f(s) = \sum_{k=0}^n a_k s^k \quad (4.1)$$

Zusätzlich wird davon ausgegangen, dass für alle  $k$   $a_k \neq 1$  und dass  $a_0 + a_1 < 1$  gilt (vgl. Harris 1963: 5).<sup>68 69</sup>

Nun ist die in der Ausgangsfragestellung genannte Wahrscheinlichkeit gesucht, dass einzelne Abstammungslinien aussterben. Es lässt sich zeigen, dass bei  $f'(1) = E(K) < 1$  langfristig jede Abstammungslinie aussterben muss – wenn also der Erwartungswert der Zahl der Nachfahren kleiner 1 ist (vgl. Feller 1977). Gleiches gilt für den Fall, dass  $E(K) = 1$  ist (vgl. ebd.). Bei  $E(K) > 1$  schließlich ergibt sich eine Aussterbewahrscheinlichkeit kleiner 1, die allerdings nur für den trivialen Fall, dass  $a_0 = 0$  ist, gleich 0 ist (vgl. ebd.; Meintrup & Schäffler 2005: 265). Den Fall  $E(K) < 1$  bezeichnet man als „subkritischen Prozess“, den Fall  $E(K) > 1$  als „superkritischen Prozess“ und für  $E(K) = 1$  spricht man von einem „kritischen Prozess“ (vgl. Kimmel & Axelrod 2002: 14).

66 Wenn  $t$  einen betrachteten Zeitpunkt angibt und Werte  $t \in \mathbb{R}$  annehmen kann, spricht man von einer „stetigen“ Betrachtung; wenn  $t \in \mathbb{Z}$ , von einer „diskreten“ Betrachtung. Im ersten Fall werden Zeit und gegebenenfalls von dieser abhängige Merkmale wie das Alter als kontinuierlich aufgefasst (vgl. Rohwer & Pötter 2003: 13). Im zweiten Fall „springt“ die Betrachtung zwischen einzelnen Zeitpunkten (vgl. ebd.: 13).

67 Zuweilen wird auch die Bezeichnung Bienayme-Galton-Watson-Prozess verwendet, da sich der französische Mathematiker Bienayme bereits vor Galton und Watson mit dem selben Problem beschäftigte (vgl. Kendall 1975).

68 Die wahrscheinlichkeitserzeugende Funktion  $f(s)$  einer beliebigen diskreten Zufallsvariable  $Y$  bestimmt die Verteilung derselben eindeutig. Insbesondere gilt allgemein, dass  $E(Y) = f'(1)$  und  $\text{Var}(Y) = f''(1) + f'(1) - f'(1)^2$  ist (vgl. bspw. Klenke 2008: 79ff.).

69 Wenn im weiteren eine Notation der Art  $f(x)$  verwendet wird, handelt es sich allgemein immer um eine Funktion, wovon die Notation  $f_x$  zu unterscheiden ist, die weiterhin für Geburtenraten steht.

### 4.3.2 Verzweigungsprozesse und Verwandtschaft

#### 4.3.2.1 Jemals geborene Verwandtschaft

Einfache Betrachtungen der Anzahl von Verwandten im Rahmen des Galton-Watson-Prozesses stammen von Waugh (1981) und Pullum (1982).

Waugh (1981: 632) geht der Frage nach, wie viele Verwandte ein aus einer Bevölkerung zufällig ausgewähltes Individuum (Ego) in vorangehenden und nachfolgenden Generationen hat. Es gelten hierbei die im vorherigen Abschnitt gemachten Annahmen, wobei als einzige Abweichung hiervon nur Frauen anstatt Männer im Modell berücksichtigt werden (vgl. ebd.: 632), so dass  $K$  für die Zahl der Töchter steht. Betrachtet werden für jede Generation alle jemals geborenen Verwandten, unabhängig davon, ob diese zeitgleich mit Ego leben oder nicht (vgl. ebd.: 632).  $Z_j$  sei die Größe der  $j$ -ten Generation, wobei 5 aufeinander folgende Generationen betrachtet werden:  $Z_0$ , die Generation der Großmutter mütterlicherseits von Ego, die als Ausgangspunkt der Betrachtung dient und deren Größe auf  $Z_0 = 1$  gesetzt wird.  $Z_1$  umfasst die Generation der Mutter von Ego inklusive Egos Tanten,  $Z_2$  erfasst die Größe von Egos Generation einschließlich Schwestern und Cousinen,  $Z_3$  entspricht den Kindern der Personen aus Generation  $Z_2$  und  $Z_4$  bezieht sich auf die Enkelkinder von  $Z_2$  (vgl. ebd.: 644). Insofern werden nur Vorfahren mütterlicherseits betrachtet.

Der Erwartungswert der Anzahl der Töchter  $K$  wird als Nettoerproduktionsrate interpretiert (vgl. ebd.: 644):<sup>70</sup>

$$f'(1) = E(K) = N \quad (4.2)$$

Nun lässt sich zeigen, dass für ein zufällig ausgewähltes Individuum gilt (vgl. ebd.: 645):

$$\begin{aligned} E(Z_1) &= N + \frac{\sigma_K^2}{N} \\ E(Z_2) &= N^2 + \frac{N+1}{N} \sigma_K^2 \\ E(Z_3) &= N \cdot E(Z_2) \\ E(Z_4) &= N \cdot E(Z_3) \end{aligned} \quad (4.3)$$

Hierbei ist  $\sigma_K^2$  die Varianz der Kinderzahl und gegeben über  $f''(1) + f'(1) - f'(1)^2$ . Über die gerade genannten Formeln lässt sich lediglich die (erwartete) Größe einer Generation in einer Abstammungslinie errechnen.

70 Die Nettoerproduktionsrate für ein Jahr  $t$  ist definiert als (vgl. bspw. Preston et al. 2001: 113)

$$N_t = \sum_{x=\alpha}^{\beta} \frac{B_{x,t}^f l_{x,t}}{J_{x,w,t} l_{0,t}}$$

wobei  $B_{x,t}^f$  für die Zahl an Mädchen steht, die von Frauen im Altersintervall  $[x, x+1)$  zwischen  $t$  und  $t+1$  geboren werden. Über den ersten Bruch auf der rechten Seite wird dann erfasst, wie viele Töchter Frauen durchschnittlich bekommen würden, die bis zum Ende der reproduktiven Phase überleben und deren Fertilität des Jahres  $t$  entspricht. Durch den zweiten Bruch wird berücksichtigt, dass das Überleben bis zum Ende der reproduktiven Phase nicht sicher ist, wobei die Überlebenswahrscheinlichkeiten aus der Periodensterbetafel des Jahres  $t$  gewonnen werden. Wie bei der zusammengefassten Geburtenziffer wird also eine hypothetische Kohorte betrachtet.

Eine differenziertere Betrachtung liefert Pullum (1982), wobei er ebenfalls von der niemals geborenen Verwandtschaft ausgeht. Hierbei orientiert er sich an Atkins (1974) Verwandtschaftsschema, welches bereits in Tabelle 2.1 auf Seite 23 dargestellt wurde, und übernimmt aus diesem die Bezeichnung von Verwandten über Zahlenpaare  $(i, j)$ . Er betrachtet zwar sowohl die Zahl der weiblichen als auch der männlichen Verwandten, allerdings sei die Darstellung zunächst auf die Zahl der weiblichen Verwandtschaft beschränkt.

Unter Beibehaltung der Annahmen von Waugh lässt sich der Erwartungswert der Zahl der niemals geborenen weiblichen Verwandten  $(i, j)$ ,  $E(V_{ij}^{(w)})$ , berechnen als (vgl. Pullum 1982: 550):

$$E(V_{ij}^{(w)}) = \begin{cases} 1 & \text{für } i = 0 \\ N^i & \text{für } j = 0 \\ SN^{i-1} & \text{sonst} \end{cases} \quad (4.4)$$

Hierbei entspricht  $S$  dem Erwartungswert der Zahl der Schwestern gegeben  $f(s)$ ,  $E(V_{1,1}^{(w)})$ . Dieser ergibt sich als (vgl. ebd.: 552):

$$S = \frac{f''(1)}{N} = \frac{E(K^2) - E(K)}{E(K)} \quad (4.5)$$

Die Berechnung der Zahl der Schwestern und hierüber die Berechnung der Anzahl an weiblichen Verwandten  $(i, j)$  für  $j \neq 0$  setzt somit genaue Kenntnis der einzelnen Wahrscheinlichkeiten  $a_k$  voraus. Die Erwartungswerte der Zahl direkter Nachfahren hingegen lassen sich bereits bei Kenntnis der Nettoerproduktionsrate berechnen.

Für alle  $E(V_{0j}^{(w)})$  gilt, dass sie gleich eins sind, was der Anzahl der direkten Vorfahren von Ego und Ego selbst entspricht. Dies ergibt sich aus der Betrachtung von niemals geborenen Verwandten: das Ego eine Mutter, eine Großmutter mütterlicherseits und so fort hat, kann als gegeben vorausgesetzt werden. Die Anzahl der  $i$ -ten direkten Nachfahren von Ego ergibt sich als Potenz der Nettoerproduktionsrate. Das heißt, dass der Erwartungswert der Zahl der Töchter von Ego gleich  $N$  ist. Aufgrund der Unabhängigkeit von Personen kann dann beispielsweise für Töchter von Ego ebenfalls angenommen werden, dass die Zahl ihrer Töchter gleich  $N$  ist, so dass sich die Zahl der Enkelkinder als  $N^2$  ergibt und so fort.<sup>71</sup>

Zur Illustration der Fälle  $i \neq 0$  beziehungsweise  $j \neq 0$  sei ein ebenfalls Beispiel angeführt. Gesucht sei der Erwartungswert der Anzahl der Großnichten  $E(V_{3,1}^{(w)})$ . Hierfür wird zunächst betrachtet wie viele Schwestern Ego hat, wobei der Erwartungswert  $S$  entspricht. Der Erwartungswert der Zahl der Töchter ist für jede dieser Schwestern gleich  $N$ . Auch für diese Töchter wiederum ist der Erwartungswert der Zahl der weiblichen Nachfahren gleich  $N$ , so dass  $E(V_{3,1}^{(w)}) = SN^2$  ist. Auch für diese Herleitung ist die Annahme der Unabhängigkeit von Personen zentral.

Um das Modell auf zwei Geschlechter zu erweitern, trifft Pullum die Annahme, dass Männer und Frauen Nachfahren mit gleichen Wahrscheinlichkeiten  $a_k$  bekommen (vgl. Pullum 1982:

71 Diese Formulierung ist nicht gänzlich korrekt. Eigentlich müsste vom „Erwartungswert der Zahl der Töchter der Töchter“ gesprochen werden. Hierauf wird der Einfachheit halber verzichtet, was auch für den weiteren Text gilt.

554). Weiterhin wird von strikter Monogamie ausgegangen, wodurch zum Beispiel Halbgeschwister von Ego ausgeschlossen werden (vgl. ebd.: 554). Die Wahrscheinlichkeit, dass ein Kind weiblich ist, sei  $\gamma$ , wobei die Geschlechterverteilung bei Geburten binomialverteilt sei (vgl. ebd.: 555). Es wird angenommen, dass diese Wahrscheinlichkeit nicht gezielt beeinflusst wird und nicht von der Parität der Mutter oder des Vaters abhängt (vgl. ebd.: 554). Nun werden zunächst über (4.4) Ergebnisse wie in der Ein-Geschlechter-Variante berechnet und diese Ergebnisse anschließend mit folgenden Faktoren multipliziert (vgl. ebd.: 555):

$$k_{ij} = \begin{cases} 2^j & \text{für } i = 0 \\ (1/\gamma)^i & \text{für } j = 0 \\ 2^{j-1}(1/\gamma)^i & \text{sonst} \end{cases} \quad (4.6)$$

Für  $i = 0$ , also die direkten Vorfahren von Ego, ist das Ergebnis leicht ersichtlich: Es gibt einen Vater und eine Mutter, Großvater und Großmutter jeweils mütterlicher- und väterlicherseits und so fort. Die Zahl der direkten Nachfahren beider Geschlechter wird einfach um  $\gamma$  verändert: wenn der Erwartungswert der Zahl der Töchtern für Ego  $N$  beträgt und der Anteil von Töchtern an Geburten  $\gamma$  beträgt, ist die durchschnittliche Gesamtzahl der Kinder gleich  $N/\gamma$ . Der Erwartungswert der Anzahl weiblicher Nachfahren für eines dieser Kinder ist wieder  $N/\gamma$ , so dass die Zahl der Enkel gleich  $N^2/\gamma^2$  ist und so weiter.

Die Zahl der männlichen und weiblichen Verwandten ( $i, j$ ) für  $j \neq 0$  sei an einem Beispiel erklärt. Um den Erwartungswert der Zahl der Cousins und Cousinen  $E(V_{2,2})$  zu berechnen, muss man zunächst die Zahl der Geschwister der Eltern kennen. Die Zahl der Eltern ist  $2^{2-1}$ , wobei beide Elternteile eine Geschwisterzahl von  $S(1/\gamma)$  haben. Die Anzahl der Kinder der Geschwister der Eltern ist gleich  $N(1/\gamma)$ , so dass sich schreiben lässt:

$$E(V_{2,2})2^{2-1}(1/\gamma)^2 = SN2^{2-1}(1/\gamma)^2 = 2S(1/\gamma)N(1/\gamma)$$

#### 4.3.2.2 Korrelation von Fertilität über Generationen

Eine Erweiterung des Modells, die Korrelation von Fertilität über Generationen hinweg berücksichtigt, liefern Pullum & Wolf (1991), wobei wieder nur Frauen berücksichtigt werden. Zunächst lässt sich für das Ein-Geschlechter-Modell ohne explizite Berücksichtigung von Abhängigkeit der Fertilität zwischen Generationen zeigen, dass der Korrelationskoeffizient zwischen der Zahl der Töchter und der Zahl der Enkeltöchter gegeben ist durch (vgl. ebd.: 397):

$$R = \sqrt{\frac{N}{N+1}} \quad (4.7)$$

Wird also angenommen, dass die Fertilität von Individuen unabhängig ist, hängt die Korrelation zwischen der Zahl der Töchter und der Zahl der Enkeltöchter nur von der Nettoerproduktionsrate ab. Das Ergebnis zeigt gewissermaßen, dass Frauen, die viele Töchter haben, durchschnittlich mehr Enkeltöchter erwarten können als Frauen mit wenigen Töchtern. Wohl-gemerkt impliziert dies keine Korrelation zwischen der Zahl der Schwestern einer Frau und der Zahl ihrer Kinder, da die zugrunde liegenden Wahrscheinlichkeiten  $a_k$  für alle Individuen gleich sind (vgl. ebd.: 398).

Um die Annahme der Unabhängigkeit der Individuen aufzuheben, definieren Pullum und Wolf zunächst mehrere wahrscheinlichkeitserzeugende Funktionen, jeweils eine für Frauen mit  $z$  Schwestern (vgl. Pullum & Wolf 1991: 400):

$$f_z(s) = \sum_{k=0}^n a_{k,z} s^k, \quad (4.8)$$

wobei  $a_{k,z}$  die Wahrscheinlichkeit ist, dass eine Frau mit  $z$  jemals geborenen Schwestern  $k$  Töchter bekommt.

Nun sei

$$N^{(z)} = f'_z(1) = \sum_{k=0}^n a_{k,z} k \quad (4.9)$$

die Nettoreproduktionsrate für Frauen mit  $z$  Schwestern und ferner sei  $g_z$  die Wahrscheinlichkeit dafür, dass eine Frau  $z$  Schwestern hat (vgl. Pullum & Wolf 1991: 400). Dann ist der Erwartungswert der Anzahl an Töchtern gleich (vgl. ebd.: 400):

$$E(K) = \sum_z g_z N^{(z)} \quad (4.10)$$

Für den Erwartungswert der Anzahl der Enkeltöchter  $E(V_{2,0}^{(w)})$  gilt (vgl. ebd.: 401):

$$E(V_{2,0}^{(w)}) = \sum_z g_z \sum_k a_{k,z} k N^{(k-1)} \quad (4.11)$$

Hiervon ausgehend geben die Autoren an, wie die Korrelation zwischen der Zahl der Töchter und der Zahl der Enkeltöchter und wie die Korrelation zwischen der Zahl der Schwestern und der Zahl der Töchter berechnet werden können (vgl. Pullum & Wolf 1991: 401ff.). Allerdings läuft dies darauf hinaus, die durch  $g_z$  und  $a_{k,z}$  implizierten Korrelationen zu berechnen, was deren Kenntnis voraussetzt. Zudem sind die Korrelationen relativ komplexe nicht-lineare Funktionen von  $g_z$  und  $a_{k,z}$ , so dass sich keine allgemeinen Aussagen über sie treffen lassen und mit konkreten Werten gerechnet werden muss.

#### 4.3.2.3 Einfache Zwei-Geschlechter-Modelle mit Berücksichtigung von Stief- und Halbverwandtschaft

In der Literatur finden sich zwei Vorschläge, wie die Zwei-Geschlechter-Variante des Ansatzes um Stief- und Halbverwandtschaft erweitert werden kann. Einer dieser Vorschläge stammt von Goldstein (1996) und der andere von Suzuki (1996). Allerdings weisen beide etliche Annahmen auf, die als sehr restriktiv angesehen werden können. Aus diesem Grunde sollen die Ansätze hier nicht genauer erläutert und lediglich ihre Annahmen aufgezeigt werden, um eine Bewertung zu ermöglichen.

Suzuki (1996) schlägt eine Erweiterung des Zwei-Geschlechter-Ansatzes vor, die mehrere Heiraten der Eltern von Ego und eine Unterscheidung von Geschwistern und Halbgeschwistern

erlaubt. Hierbei werden zwei Varianten voneinander getrennt betrachtet: eine erste Variante erlaubt mehrere Heiraten der Mutter, eine zweite Variante mehrere Heiraten des Vaters (vgl. ebd.: 22ff.). In beiden Fällen wird davon ausgegangen, dass maximal zwei mal geheiratet wird und dass Geburten nur in der Ehe erfolgen (vgl. ebd.: 21).

Ein allgemeinerer Ansatz, der mehrfache Trennungen und Wiederheiraten sowie Stiefverwandtschaft berücksichtigt, wird von Goldstein (1996) vorgeschlagen. Bei diesem wird neben der Verteilung der Nachfahren noch die Verteilung der Zahl der Ehen berücksichtigt. Unter anderem ist  $m_i$  allgemein der Anteil aller Männer und Frauen, die in ihrem Leben  $i$ -mal verheiratet sein werden (vgl. ebd.: 17). Es wird bei der Zahl der Ehen also nicht zwischen den Geschlechtern unterschieden. Ferner handelt es sich bei  $m_i$  wie bei der Zahl der erwarteten Nachfahren um ein prospektives Maß – es werden die Wahrscheinlichkeiten für bestimmte Anzahlen an Ehen über ein komplettes Leben betrachtet und nicht, wann diese Ehen bestehen. Er führt zudem etliche weitere Annahmen an (vgl. ebd.: 15f.):

- Geburten finden nur in Ehen statt.
- Hieraus ergibt sich, dass Egos Eltern mindestens einmal verheiratet sein müssen.
- Die Zahl weiterer Ehen des Vaters und der Mutter Egos sind unabhängig voneinander.
- Die Zahl der Kinder und die Zahl der Ehen sind unabhängig voneinander, abgesehen davon, dass nie verheiratete Personen entsprechend der ersten Annahme keine Kinder haben können.
- Geburten sind zufällig über Ehen verteilt.
- Ferner existiert eine Nachfahrenverteilung nur für Frauen; die Zahl der Nachfahren von Männern ergibt sich über die Zahl der Nachfahren der Frauen, mit denen sie verheiratet sind.

### 4.3.3 Diskussion

Die hier dargestellten Anwendungen von Verzweigungsprozessen stellen zwar nicht die einzigen Ansätze dieser Art dar, allerdings sind die weiteren in der Literatur auffindbaren Überlegungen primär mathematischer Natur und nicht zur tatsächlichen Anwendung mit empirischen Daten konzipiert. Insbesondere gelten etliche Ergebnisse nur für superkritische Populationen oder lassen sich nur unter sehr restriktiven Annahmen herleiten, die eher an Problemen der Zellbiologie orientiert sind.<sup>72</sup>

Betrachtet man die in diesem Abschnitt diskutierten Arbeiten mit Hinblick auf die in Abschnitt 4.1 genannten Kriterien, zeigt sich zunächst, dass Unsicherheit nicht berücksichtigt wird. Es ist eher das Gegenteil der Fall, da von konstanten demographischen Bedingungen ausgegangen wird. Der sich hieraus ergebende Datenbedarf ist zwar bezogen auf manche Ausdrücke relativ gering, allerdings allgemein nicht unerheblich. Dies wird auch dadurch deutlich, dass lediglich bei den hier vorgestellten Arbeiten von Pullum (1982), Pullum & Wolf (1991) und Waugh (1981) tatsächlich auf empirischen Daten beruhende Berechnungen

---

<sup>72</sup> Hierzu zählen Arbeiten von: Samuels (1971), Bühler (1971), Bühler (1974), Jagers (1982), Joffe & Waugh (1982), Nerman & Jagers (1984), Joffe & Waugh (1985) und Joffe & Waugh (1986).

vorgestellt werden. Auch diese sind nicht unbedingt unproblematisch, da diverse Größen aus Umfragedaten geschätzt werden müssen, die Stichprobenfehler aufweisen und möglicherweise aufgrund erhebungsbedingter Probleme, wie beispielsweise eines Survivor-Bias, verzerrt sein können (vgl. hierzu ausführlicher Preston 1976). Dabei liegen aber keine Validierungsversuche vor, so dass keine Einschätzung möglich ist, inwieweit solche Probleme einen Einfluss auf Ergebnisse haben. Zwar lassen sich die Wahrscheinlichkeiten  $a_k$  unter Verwendung von altersspezifischen Geburtenraten auch ohne Surveydaten schätzen, allerdings nur unter der Annahme, dass die Wahrscheinlichkeit einer Geburt in einem bestimmten Alter nicht von der bisherigen Kinderzahl abhängt (vgl. Krishnamoorthy 1979).

Zwar wird in der Literatur versucht, den zugrunde liegenden Ansatz auf zwei Geschlechter zu erweitern, allerdings werden hierfür weitere, restriktive Annahmen nötig, wie beispielsweise lebenslange Monogamie. Wird wie bei Goldstein (1996) oder Suzuki (1996) versucht, diese Annahmen teilweise aufzuheben und zusätzlich Nuptialität explizit zu berücksichtigen, werden abermals neue Annahmen nötig, wie die bei Suzuki (1996) zu findende Annahme, dass maximal zwei mal im Lebensverlauf geheiratet wird. Zudem liegt ein hohes Maß an Aggregation vor und es wird höchstens nach Geschlecht differenziert, nicht aber nach Alter. Schließlich wird in der Literatur zwar ein formaler Rahmen gegeben, durch den auf eine Annahme von Unabhängigkeit zwischen Personen verzichtet werden kann, allerdings weist dieser nichtsdestotrotz die anderen bereits genannten Schwächen auf und geht mit einem erheblichen Datenbedarf einher.

Weiterhin scheint problematisch, dass nur jemals geborene Verwandte betrachtet werden und nicht zeitgleich lebende Verwandte. Hierfür müsste explizit das Alter von Personen berücksichtigt werden, was auch eine Trennung der Effekte von Mortalität und Fertilität sowie eine Berücksichtigung von altersspezifische Geburten- und Sterberaten erlauben würde, die in den hier vorgestellten Modellen in  $N$  zusammengefasst sind. Eine weitere Schwierigkeit ist, dass Migration nicht berücksichtigt wird und faktisch von einer „geschlossenen“ Bevölkerung ausgegangen wird – wenn der Ansatz ein realistisches Bild liefern soll, dürfen weder Ein- noch Auswanderungen stattfinden.<sup>73</sup>

## 4.4 Bevölkerungsmathematische Ansätze

### 4.4.1 Grundlagen

Ansätze aus der Bevölkerungsmathematik beruhen auf Sterbetafeln und Geburtenraten. Dabei wird davon ausgegangen, dass für die betrachtete Bevölkerung konstante Raten gelten, sich Mortalität und Fertilität also nicht ändern. Ferner wird häufig eine zeitstetige Perspektive gewählt. Die zugrunde liegende Logik dieser Ansätze geht auf Lotka (1931) zurück, welcher

---

73 Die fehlende Berücksichtigung von Migration kann beispielsweise problematisch sein, wenn in ein betrachtetes Gebiet eine große Anzahl von Personen einwandert, die andere Verwandtschaftsstrukturen aufweist, als die einheimische Bevölkerung. In diesem Fall liegen dann Wechsel von Personen zwischen Gebieten mit unterschiedlichen demographischen Regimen vor, die nicht berücksichtigt werden. Detailliertere Beispiele, die sich auf den in dieser Arbeit verwendeten Vorausberechnungsansatz beziehen, finden sich in Anhang C.5.

den Einfluss unterschiedlicher Mortalitäts- und Fertilitätsregime auf den Anteil an Waisen und Halbwaisen in bestimmten Altersgruppen einer Bevölkerung untersuchte.<sup>74</sup>

Gegeben konstante Sterbe- und Geburtenraten sei  $v$  das durchschnittliche Alter von Müttern bei der Geburt eines Kindes. Betrachtet wird nun ein Kind im Alter  $a$ . Wenn die Mutter des Kindes noch lebt, ist diese durchschnittlich  $v + a$  Jahre alt (vgl. Lotka 1931: 39). Wenn  $l_x$  die Wahrscheinlichkeit bezeichnet, dass ein neugeborenes Mädchen bis zum Alter  $x$  überlebt, ist die Wahrscheinlichkeit, von Alter  $v$  bis zum Alter  $a + v$  zu überleben, gleich  $l_{a+v}/l_v$  (vgl. ebd.: 40). Entsprechend gibt das Komplement dieses Terms die Wahrscheinlichkeit an, dass die Mutter des Kindes nicht mehr lebt, wenn dieses  $a$  Jahre alt ist (vgl. ebd.: 41):

$$\Theta_a^w = 1 - \frac{l_{a+v}}{l_v} \quad (4.12)$$

Auf analoge Weise lässt sich die Wahrscheinlichkeit  $\Theta_a^m$ , dass der Vater des Kindes nicht mehr lebt, berechnen. Allerdings muss hierbei beachtet werden, dass nicht das durchschnittliche Alter von Vätern bei der Geburt des Kindes, sondern bei dessen Zeugung verwendet wird (vgl. ebd.: 44). Hierdurch gilt insbesondere, dass  $\Theta_0^w = 0 < \Theta_0^m$  ist, wobei  $\Theta_0^m$  allerdings selbst bei hoher Mortalität nur geringfügig über 0 liegt (vgl. ebd.: 44f.).<sup>75</sup>

Diese Überlegungen lassen sich vergleichsweise einfach auf weitere Verwandte übertragen. Die maßgebliche Ausarbeitung hierzu ist ein Aufsatz von Goodman et al. (1974). Diese sowie einige Erweiterungen und Anwendungen werden in den folgenden Unterabschnitten diskutiert.

#### 4.4.2 Der Ansatz nach Goodman, Keyfitz und Pullum

##### 4.4.2.1 Grundlagen

Der Ansatz von Goodman et al. (1974) berücksichtigt nur weibliche Verwandtschaft mütterlicherseits. Dies gilt auch über mehrere Generationen hinweg: soll beispielsweise ausgehend von einer Person die Wahrscheinlichkeit einer lebenden Urgroßmutter berechnet werden, handelt es sich hierbei nur um die Mutter der Mutter der Mutter, nicht aber die Mutter der Mutter des Vaters. Es wird ferner davon ausgegangen, dass zum einen Wahrscheinlichkeiten, von der Geburt bis zum exakten Alter  $x$  zu überleben,  $l_x$ , sowie Wahrscheinlichkeiten für Geburten in exakten Altern  $x$ ,  $h_x$ , gegeben sind (vgl. ebd.: 1f.). Das gebärfähige Alter ist durch eine Untergrenze  $\alpha$  und eine Obergrenze  $\beta$  eingeschränkt, wobei  $h_x = 0$  für  $x < \alpha$  und  $x > \beta$  (vgl. ebd.: 2). Die einzelnen altersspezifischen Wahrscheinlichkeiten  $l_x$  und  $h_x$  werden als konstant angenommen (vgl. ebd.: 1). Darüber hinaus gelten diese für alle Frauen in der betrachteten Bevölkerung in gleicher Weise. Es wird also weder Heterogenität noch

74 Bevölkerungsmathematische Ansätze, die sich am Konzept des *family life cycle* orientieren (vgl. hierzu bspw. Höhn 1982), wie beispielsweise von Feichtinger (1977, 1987), Hanslůvka (1976), Krishnamoorthy (1979) und Preston (1987) vorgestellt, werden hier nicht diskutiert, da sich diese primär auf Verweildauern in bestimmten Zuständen, wie beispielsweise Ehen, beschränken und das Vorhandensein von Verwandten im weiteren Sinne nicht erfasst wird.

75 Beispielsweise ergibt sich unter den Mortalitätsbedingungen der USA im Jahre 1920 ein Wert für  $\Theta_0^m$  von etwa 0.5% (vgl. Lotka 1931: 45). Zudem hält Lotka selbst diesen Wert für eine relativ hohe Schätzung, die den tatsächlichen Wert deutlich übersteigt (für Details vgl. Lotka 1931: 45).

Korrelation von Fertilität beziehungsweise Mortalität über Generationen hinweg berücksichtigt (vgl. Goodman et al. 1975: 377f.).<sup>76</sup>

Ausgehend von den Überlegungen im letzten Abschnitt ist die Wahrscheinlichkeit, dass die Mutter einer Frau im Alter  $a$  noch lebt allgemein gleich  $l_{x+a}/l_x$ , wobei  $x$  hier das Alter der Mutter bei der Geburt der Frau bezeichnet (vgl. Goodman et al. 1974: 7).  $x$  liegt zwar zwischen  $\alpha$  und  $\beta$ , ist ansonsten aber nicht weiter eingeschränkt. Um dies zu berücksichtigen, wird das Integral über das Intervall  $\alpha$  bis  $\beta$  gebildet und der Bruch gewichtet mit der Verteilung von  $x$  zum Zeitpunkt  $t - a$ ,  $W(x|t - a)$  (vgl. ebd.: 7):

$$M_1(a) = \int_{\alpha}^{\beta} \frac{l_{a+x}}{l_x} W(x|t - a) dx \quad (4.13)$$

Die Wahrscheinlichkeit, dass die Großmutter mütterlicherseits noch lebt ist  $M_1(a + x)$ , allerdings muss auch hierbei wieder die Altersverteilung von Frauen bei der Geburt eines Kindes zu  $t - a$  berücksichtigt werden, so dass man

$$M_2(a) = \int_{\alpha}^{\beta} M_1(a + x) W(x|t - a) dx \quad (4.14)$$

erhält (vgl. ebd.: 8). Wie bereits an den Indizes in (4.13) und (4.14) ersichtlich, gehört die Mutter zur ersten (nächsten) Vorfahrgeneration, die Großmutter zur zweiten Vorfahrgeneration, die Urgroßmutter zur dritten Vorfahrgeneration und so weiter (vgl. ebd.: 8). Die Wahrscheinlichkeit für eine lebenden Vorfahrin der  $i$ -ten Vorfahrgeneration erhält man allgemein als (vgl. ebd.: 9)

$$M_i(a) = \int_{\alpha}^{\beta} M_{i-1}(a + x) W(x|t - a) dx \quad (4.15)$$

Der Erwartungswert der Zahl der bisher geborenen Töchter einer Frau im Alter  $a$  lässt sich berechnen als (vgl. Goodman et al. 1974: 2)<sup>77</sup>

$$\int_{\alpha}^a h_x dx, \quad (4.16)$$

Um zum Erwartungswert der Zahl der lebenden Töchter einer Frau im Alter  $a$  zu gelangen, wird  $l_{a-x}$  in das Integral eingefügt, was die Wahrscheinlichkeit wiedergibt, dass eine von der

76 Letzteres meint wie bereits auf Seite 61 angedeutet, dass beispielsweise die Wahrscheinlichkeit, in einem bestimmten Alter  $x$  ein Kind zu bekommen, nicht von der bis zum Alter  $x$  realisierten Anzahl von Geburten abhängt.

77 Die Verwendung der Begriffe „Wahrscheinlichkeit“ und „Erwartungswert“ durch Goodman et al. (1974) ist in diesem Kontext eigentlich formal nicht gänzlich korrekt. Zwar gehen sie von den weiter oben angegebenen Wahrscheinlichkeiten aus, führen aber an einigen Stellen die Annahme einer stabilen Bevölkerung ein (vgl. bspw. ebd.: 7), was ein deterministisches Konzept ist (s. Fußnote 81 auf S. 66). Hier wird sich im weiteren nichtsdestotrotz an den ursprünglichen Formulierungen bei Goodman et al. (1974) orientiert.

betrachteten Frau im Alter  $x$  geborene Tochter bis zum Alter  $a - x$  überlebt (vgl. ebd.: 3):

$$\int_{\alpha}^a l_{a-x} h_x dx \quad (4.17)$$

Um die Zahl lebender Enkeltochter zu berechnen, wird wie folgt vorgegangen. Die Zahl der Töchter einer Frau ergibt sich aus (4.16). Die Wahrscheinlichkeit, dass eine dieser Töchter bis zum Alter  $y$  überlebt und selbst eine Tochter in diesem Alter gebärt, ist  $\int_{\alpha}^{a-x} l_y h_y dy$ . Diese Enkeltochter müssen bis zum Alter  $a - x - y$  überleben, womit sich

$$\int_{\alpha}^a \left[ \int_{\alpha}^{a-x} l_{a-x-y} l_y h_y dy \right] h_x dx \quad (4.18)$$

ergibt (vgl. Goodman et al. 1974: 4).

Bei der Berechnung des Erwartungswertes der Zahl der lebenden Schwestern muss zunächst zwischen älteren und jüngeren Schwestern unterschieden werden. Denn bei letzteren muss im Gegensatz zu ersteren das Überleben der Mutter nach der Geburt der betrachteten Frau berücksichtigt werden (vgl. Goodman et al. 1974: 11). Wenn Ego geboren wurde, als ihre Mutter  $x$  Jahre alt war, ergibt sich die Zahl der lebenden älteren Schwestern aus der Fertilität vom Alter  $\alpha$  bis zum Alter  $x$  unter Berücksichtigung der Mortalität von der Geburt bis zum Alter  $a + x - y$  und  $W(x|t - a)$ , wobei  $y$  das Alter der Mutter bei der Geburt einer älteren Schwester ist (vgl. ebd.: 11):

$$S'(a) = \int_{\alpha}^{\beta} \left[ \int_{\alpha}^x l_{a+x-y} h_y dy \right] W(x|t - a) dx \quad (4.19)$$

Für jüngere Schwestern ergibt sich (vgl. ebd.: 11):

$$S''(a) = \int_{\alpha}^{\beta} \left[ \int_0^a \frac{l_{x+y}}{l_x} l_{a-y} h_{x+y} dy \right] W(x|t - a) dx \quad (4.20)$$

Die Zahl der lebenden Schwestern insgesamt ist dann gegeben durch  $S(a) = S'(a) + S''(a)$ . Hierbei ist zu beachten, dass dies der Zahl der Schwestern exklusive Ego entspricht. Die Zahl der Schwestern inklusive Ego ist gleich  $S(a) + 1$  (vgl. Keyfitz & Caswell 2005: 383). Lässt man Mortalität unberücksichtigt, entspricht beim hier vorgestellten Verfahren die Zahl der Schwestern der Bruttoreproduktionsrate, gegebenenfalls eingeschränkt bis zum Alter  $x < \beta$  der Mutter (vgl. ebd.: 383).<sup>78</sup>

Geht man der Einfachheit halber zusätzlich davon aus, dass die Mutter von Ego die reproduktive Phase abgeschlossen hat, lässt sich zeigen, dass der Erwartungswert der Zahl der

<sup>78</sup> Die Bruttoreproduktionsrate entspricht der Nettoproduktionsrate ohne Berücksichtigung von Mortalität (vgl. bspw. Preston et al. 2001: 113).

Schwestern (exklusive Ego) gleich

$$S = G + \frac{\sigma^2}{G} - 1 \quad (4.21)$$

ist, wobei  $G$  für die Bruttoreproduktionsrate und  $\sigma^2$  für die Varianz der Zahl der Töchter steht (vgl. Goodman et al. 1975: 378). Damit aber wie oben erwähnt  $S = G$  gilt, muss  $\sigma^2 = G$  sein, also die Varianz der Zahl der Töchter muss dem Erwartungswert der Zahl der Töchter entsprechen (vgl. ebd.: 378). Dieses Merkmal ist eine typische Eigenschaft der Poisson-Verteilung (Equidispersion), allerdings wird diese nach Ansicht der Autoren hierdurch nicht zwingend für das Modell impliziert (vgl. ebd.: 378).

Neben den hier vorgestellten Ausdrücken leiten Goodman et al. (1974) noch Formeln für die Erwartungswerte der Zahl lebender Nichten, Tanten, Cousinen und Urenkel her. Daneben finden sich bei Keyfitz & Caswell (2005) Formeln für das durchschnittliche Alter von ausgewählten Verwandten. Beispielsweise berechnet sich das durchschnittliche Alter von Töchtern von Frauen im Alter  $a$  als (vgl. Keyfitz & Caswell 2005: 386):

$$\frac{\int_{\alpha}^a (a-x)l_{a-x}h_x dx}{\int_{\alpha}^a l_{a-x}h_x dx} \quad (4.22)$$

Entsprechende Ausdrücke lassen sich nach der selben Logik auch für andere Verwandte herleiten.

#### 4.4.2.2 Berechnung mittels demographischer Raten

Die von Goodman et al. (1974) gegebenen Formeln enthalten teils etliche Integrale, die zur Berechnung evaluiert werden müssen. Allerdings liegen die benötigten Raten in aller Regel für 1- oder 5-Jahres-Intervalle vor (vgl. Burch 1995b: 162). Goodman et al. (1974: 24ff.) gehen von Daten in 5-Jahres-Intervallen aus und nutzen eine trapezoidale Approximation der Integrale, während Dudel (2008) ausgehend von Daten in 1-Jahres-Intervallen eine einfache Treppenfunktion verwendet. Eine (annähernd) stetige Schätzung, die eine Evaluation der Integrale für beliebige  $a$  erlaubt, wurde von Burch (1995b) vorgeschlagen.<sup>79</sup> Hier seien kurz zwei Beispiele von Dudel (2008) aufgegriffen.

Sei  $f_x$  wie bereits eingeführt die altersspezifische Geburtenrate für das Altersintervall von  $x$  bis  $x + 1$  (vgl. Dudel 2008: 4f.). Dann lässt sich die Zahl der lebenden Töchter einer Frau im Alter  $a$  berechnen als (vgl. ebd.: 5):

$$\sum_{\alpha}^{a-1} l_{a-x} f_x \quad (4.23)$$

<sup>79</sup> Ein weiterer, möglicherweise verwendbarer Ansatz, der eine stetige Evaluierung erlaubt, ist in der Spline-Interpolation von Tafelfunktionen zu sehen (vgl. bspw. Hsieh 1991).

Die Wahrscheinlichkeit für eine lebende Großmutter wiederum ergibt sich als (vgl. ebd.: 6):

$$\sum_{\alpha}^{\beta-1} \left[ \sum_{\alpha}^{\beta-1} \frac{l_{x+y+a}}{l_y} W(y|t-a-x) \right] W(x|t-a) \quad (4.24)$$

Neben demographischen Raten wird zur Berechnung noch allgemein  $W(x|t-z)$  für diverse  $z$  benötigt, je nachdem welche Vorfahrgenerationen berücksichtigt werden. Wird von einer stabilen Bevölkerung ausgegangen, ist  $W(x|t-z) = W(x)$ , wobei sich dies berechnen lässt als (vgl. ebd.: 11):<sup>80</sup>

$$W(x) = f_x l_x e^{-rx} \quad (4.25)$$

Hierbei ist  $r$  die intrinsische Rate des Wachstums der stabilen Bevölkerung (vgl. hierzu ausführlicher Keyfitz & Caswell 2005: Kap. 5).<sup>81</sup>

#### 4.4.2.3 Erweiterung des Ansatzes auf zwei Geschlechter

Der Ansatz nach Goodman et al. (1974) berücksichtigt lediglich weibliche Verwandte mütterlicherseits. Ein Ansatz, der nicht nur die Wahrscheinlichkeit des Vorhandenseins der Mutter, sondern auch des Vaters liefert, wurde von Le Bras (1973) vorgeschlagen.<sup>82</sup> Hierbei wird wie bei den bereits vorgestellten Überlegungen auch von konstanten Raten ausgegangen (vgl. Le Bras & Wachter 1978: 164). Sei wie bisher  $l_x$  die Wahrscheinlichkeit, von der Geburt bis zum Alter  $x$  zu überleben, wobei zusätzliche Indizes  $w$  oder  $m$  das Geschlecht spezifizieren. Ferner sei  $W(x)$  die Altersverteilung von Müttern bei der Geburt eines Kindes und  $T(x, y)$  sei die bedingte Altersverteilung von Vätern bei der Geburt eines Kindes in Abhängigkeit des Alters der Mutter, wobei sich  $x$  auf das Alter der Mutter und  $y$  auf das Alter des Vaters bezieht, so dass  $\int T(x, y) dy = 1$  gilt (vgl. ebd.: 165f.). Schließlich sei  $q_x$  die Sterberate für das exakte Alter  $x$ , wobei hier ebenfalls wieder zusätzliche Indizes  $w$  und  $m$  verwendet werden (vgl. ebd.: 166). Dann ist die Wahrscheinlichkeit, dass eine Person  $u$  Jahre alt ist wenn die Mutter stirbt,

80 Dies ist eine Schätzung eines stetigen Ergebnisses von Lotka (1931).

81 Unter einer „stabilen Bevölkerung“ versteht man eine Bevölkerung, die dauerhaft konstante demographische Raten aufweist. Für solch eine Bevölkerung lassen sich etliche Eigenschaften, wie beispielsweise die Altersstruktur, mathematisch aus den demographischen Raten herleiten. Dabei ist zu beachten, dass eine stabile Bevölkerung zwingend konstante demographische Raten impliziert, aber nicht umgekehrt. Ausführliche Diskussionen findet man bei Keyfitz & Caswell (2005: Kap. 5) und Preston et al. (2001: Kap. 7).

82 Eine weitere Verallgemeinerung des Ansatzes auf zwei Geschlechter sowie eine Kombination mit einem Mikrosimulationsansatz findet man bei Bartlema (1987, 1988, 1989) sowie Bartlema & de Jong (1985) und Bartlema & Winkelbauer (1986). Allerdings ist dieser Ansatz nicht unproblematisch und wird im weiteren nicht berücksichtigt. Beispielsweise wird bei etlichen, zu den von Goodman et al. (1974) vorgestellten Formeln äquivalenten Ausdrücken folgender multiplikativer Korrekturfaktor verwendet (s. Bartlema 1987: 64ff.):

$$\frac{\text{TFR} - 1}{\text{TFR}} \quad (4.26)$$

Je nachdem, ob sich die entsprechende Formel auf Verwandtschaft mütterlicher- oder väterlicherseits bezieht, ist mit TFR entweder die zusammengefasste Geburtenziffer der Frauen oder aber der Männer gemeint. Dabei ist aber leicht ersichtlich, dass für  $\text{TFR} < 1$  eine negative Verwandtenanzahl und für  $\text{TFR} = 1$  0 Verwandte resultieren. Und auch für  $\text{TFR} > 1$  wird die Zahl der erwarteten lebenden Verwandten eines Verwandtschaftsgrades gegebenenfalls erheblich reduziert. Für  $\text{TFR} = 1.3$  beispielsweise erhält man einen Faktor von etwa 0.23.

und  $v$  Jahre, wenn der Vater stirbt, gegeben über (vgl. Le Bras 1973: 12):<sup>83</sup>

$$B(u, v) = \int \int W(x) q_{x+u, w} \frac{l_{x+u, w}}{l_{x, w}} q_{y+v, m} \frac{l_{y+v, m}}{l_{y, m}} T(x, y) dx dy \quad (4.27)$$

$q_x$  steht für die Sterbewahrscheinlichkeit und die Indizes  $w$  und  $m$  geben an, auf welches Geschlecht sich eine Wahrscheinlichkeit bezieht. Um  $B(u, v)$  zu berechnen, werden Simulationen vorgeschlagen, bei denen für eine bestimmte Zahl an Individuen das Alter der Eltern bei der Geburt entsprechend  $W(x)$  und  $T(x, y)$  festgelegt wird und das Sterbealter über  $l_x$  und  $q_x$  bestimmt wird (vgl. Le Bras & Wachter 1978: 166).

#### 4.4.2.4 Berücksichtigung von Heterogenität

In Unterabschnitt 4.4.2.1 wurde bereits darauf hingewiesen, dass der Ansatz nach Goodman et al. (1974) bei der Berechnung der Zahl der Schwestern Gleichheit von Mittelwert und Varianz der Kinderzahl voraussetzt. Coresh & Goldman (1988) zeigten darüber hinaus, wie die fehlende Berücksichtigung von Heterogenität bei Fertilität zu falschen Schätzungen der Zahl lateraler Verwandte führen kann. Weiterhin beibehalten wird dabei die Annahme einer stabilen Bevölkerung (vgl. ebd.: 141).

Sei  $h_{x, k}$  die Geburtenrate von Frauen im Alter  $x$  und vom „Typus“  $k$  (vgl. Coresh & Goldman 1988: 141). Frauen sind entsprechend einer zunächst nicht weiter spezifizierten Verteilung  $p(k)$  auf die verschiedenen Typen verteilt (vgl. ebd.: 141). Ferner sei  $\bar{h}_x = \int_k p(k) h_{x, k} dk$  die durchschnittliche Geburtenrate für das Alter  $x$  über alle Typen (vgl. ebd.: 141). Damit die betrachtete Bevölkerung stabil ist, muss  $\bar{h}_x$  konstant sein, wofür wiederum  $p(k)$  und alle  $h_{x, k}$  konstant sein müssen. Dies wird erreicht, indem davon ausgegangen wird, dass ein neugeborenes Mädchen mit einer zeitkonstanten Wahrscheinlichkeit  $p(k)$  zum Typus  $k$  gehört, unabhängig vom Typus der Mutter (vgl. ebd.: 142).

Zunächst zeigen die Autoren, dass Heterogenität der Fertilität bei direkten Nachfahren nicht zu verzerrten Werten führt, da beispielsweise die Zahl jemals geborener Töchter einer Frau im Alter  $a$  gleich

$$\int_{\alpha}^a \bar{h}_x dx \quad (4.28)$$

ist, was der in Abschnitt 4.4.2.1 Formel vorgestellten Formel ohne Berücksichtigung von Mortalität entspricht (vgl. Coresh & Goldman 1988: 142).

Für ältere und jüngere Schwestern von Frauen im Alter  $a$  mit einer Mutter von Typus  $k$  gilt zunächst analog zum Ansatz von Goodman et al. (1974)

$$S'(a, k) = \int_{\alpha}^{\beta} \left[ \int_{\alpha}^x h_{y, k} dy \right] W(x, k) dx \quad (4.29)$$

83 Der Integrationsbereich der beiden Integrale ist in keiner der beiden zugrunde gelegten Quellen (Le Bras 1973; Le Bras & Wachter 1978) genauer spezifiziert.

sowie

$$S''(a, k) = \int_{\alpha}^{\beta} \left[ \int_{\alpha}^{x+a} h_{y,k} \frac{l_y}{l_x} dy \right] W(x, k) dx \quad (4.30)$$

wobei in diesem Fall der Einfachheit halber das Überleben der Schwestern nicht berücksichtigt wird (vgl. Coresh & Goldman 1988: 143). Nun wird weiterhin die Wahrscheinlichkeit  $P(k)$  benötigt, dass Egos Mutter vom Typus  $k$  ist. Es kann gezeigt werden, dass

$$P(k) = p(k) \int_{\alpha}^{\beta} e^{-rx} l_x h_{x,k} dx \quad (4.31)$$

gilt (vgl. ebd.: 143). Dann liegt der Erwartungswert der Anzahl der älteren Schwestern einer Frau im Alter  $a$  bei

$$S'(a) = \int_k P(k) S'(a, k) dk \quad (4.32)$$

und die erwartete Anzahl jüngerer Schwestern ist durch

$$S''(a) = \int_k P(k) S''(a, k) dk \quad (4.33)$$

gegeben (vgl. ebd.: 143). Sind  $h_{x,k}$  und  $p(k)$  bekannt, können die Gleichungen evaluiert werden und gegebenenfalls zu Resultaten führen, die von Ergebnissen, die über den Ansatz von Goodman et al. (1974) berechnet wurden, abweichen (vgl. ebd.: 144). Über die Richtung der Abweichung können allerdings nur unter weiteren Annahmen allgemeine Aussagen getroffen werden. Nimmt man beispielsweise zu den bisherigen Annahmen zusätzlich an, dass sich die betrachteten Typen lediglich bezogen auf die Höhe der Bruttoreproduktionsrate unterscheiden, ist die erwartete Zahl an Schwestern bei Berücksichtigung dieser Heterogenität immer höher als der Erwartungswert, der sich ergibt, wenn Heterogenität nicht berücksichtigt wird (vgl. Coresh & Goldman 1988: 144f.).

#### 4.4.3 Anwendungsbeispiel

Anwendungen des ursprünglichen Ansatzes nach Goodman et al. (1974) finden sich neben dem ursprünglichen Aufsatz bei Bartlema (1987), Dudel (2008), Goldman (1986), Keyfitz (1986), Keyfitz & Caswell (2005), Martin & Culter (1983) und Tu & Liang (1990). Die als zweites genannte Arbeit bezieht sich auf die BRD und enthält zudem einen Validierungsversuch, weshalb kurz auf diese Ergebnisse eingegangen sei.

Zunächst findet sich bei Dudel (2008) eine Berechnung aller von Goodman et al. (1974) gegebenen Formeln für lebende Verwandte mit der Sterbetafel 2004/2006 für Frauen des Statistischen Bundesamtes sowie altersspezifischen Geburtenraten für das Jahr 2006. Die

*Tabelle 4.1:* Ergebnisse des Ansatzes nach Goodman et al. (1974) mit Daten für die BRD sowie Ergebnisse für durchschnittliche Anzahl Verwandter im SOEP

Alter	Modellrechnung			SOEP		
	30	50	70	30	50	70
Großmutter	0.225	0.003	0.000	0.716	0.035	0.000
Mutter	0.931	0.586	0.029	0.965	0.658	0.045
Schwestern	0.627	0.652	0.541	0.884	1.048	0.800
Tanten / Nichten	0.864	0.778	0.466	2.787	2.841	1.359
Töchter	0.313	0.644	0.638	0.390	0.996	0.974
Enkel	0.000	0.055	0.384	0.000	0.447	2.162

Quelle: Dudel (2008: 14ff.)

resultierenden Ergebnisse werden mit Daten des Sozio-oekonomischen Panels (SOEP) aus dem Jahr 2006 verglichen (vgl. Dudel 2008: 15ff.).<sup>84</sup> Die Ergebnisse finden sich in Tabelle 4.1.

Zunächst sind bei diesem Vergleich etliche Einschränkungen zu machen. Für die Zahlen der im SOEP erfassten Verwandten ist es teils unklar, ob es sich um aktuell lebende, oder aber jemals vorhandene Verwandte handelt (vgl. Dudel 2008: 16). Daneben erfasst der verwendete Ansatz nur weibliche Verwandte mütterlicherseits, während beim SOEP sowohl mütterliche als auch väterliche Verwandte berücksichtigt und zudem teilweise männliche und weibliche Verwandte zusammengefasst werden (vgl. ebd.: 16f.). Beispielsweise wird im SOEP die Zahl an Tanten und Nichten zusammengefasst erhoben, wobei sowohl Tanten väterlicher- als auch mütterlicherseits sowie Nichten sowohl über Schwestern als auch Brüder erfasst werden (vgl. ebd.: 17). Ähnliches gilt für Großmütter – im SOEP wird sowohl die Großmutter väterlicher- als auch mütterlicherseits erfasst. Die einzigen eindeutigen Vergleiche sind bei der Zahl der Töchter und dem Vorhandensein der Mutter möglich (vgl. ebd.: 17). Für diese beiden Verwandtschaftskategorien kommt Dudel (2008) zu dem Ergebnis, dass für jüngere Alter zumindest eine grobe Abschätzung der Größenordnung der Anzahl der Verwandten möglich ist, während insbesondere für höhere Alter deutliche Abweichungen vorliegen (vgl. ebd.: 19).

Weiterhin berechnet Dudel (2008: 17ff.) Schätzungen für das Jahr 2050 mit einer mittels der Lee-Carter-Methode vorausgerechneten Sterbetafel unter Beibehaltung der altersspezifischen Geburtenraten von 2006.<sup>85</sup> Bei dieser vorausgerechneten Sterbetafel ergibt sich eine Lebenserwartung von etwa 89.45 Jahren, wobei der Anstieg verglichen zur Sterbetafel 2004/2006 vor allem auf eine Reduktion der Mortalität in höheren Lebensaltern zurückzuführen ist. Ein zusammenfassender Vergleich der insgesamt erwarteten Verwandtenzahl für 2006 und 2050 findet sich in Tabelle 4.2.

Zunächst fällt wenig überraschend auf, dass eine Abnahme der Mortalität zu mehr lebenden Verwandten während aller Lebensalter führt (vgl. Dudel 2008: 19). Für mittlere Lebensalter fällt der Zuwachs allerdings deutlich stärker aus als für ältere oder jüngere Alter. Dies wird

<sup>84</sup> Eine Beschreibung des SOEP und der darin enthaltenen Daten zu Verwandtschaft findet sich in Abschnitt 7.2.

<sup>85</sup> Die Lee-Carter-Methode wurde von Lee & Carter (1992) vorgeschlagen. Eine kurze Diskussion der dem Ansatz zugrunde liegenden Idee findet man zu Beginn von Unterabschnitt 8.2.1.

Tabelle 4.2: Zusammengefasste Ergebnisse für den Ansatz nach Goodman et al. (1974) für 2006 und 2050 sowie Differenzen zwischen den Ergebnissen

Alter	10	30	50	70	90
2006 (1)	3.535	3.434	3.129	2.386	1.825
2050 (2)	3.747	3.716	3.548	2.648	2.089
Differenz (2)-(1)	0.212	0.282	0.419	0.262	0.264

Quelle: Dudel (2008: 21)

darauf zurückgeführt, dass in jungen Altern die meisten Verwandten selbst nicht in hohen Altern und somit nur einem geringen Sterblichkeitsrisiko ausgesetzt sind (vgl. ebd.: 19). Für höhere Lebensalter hingegen gilt, dass zumindest Vorfahren ein extrem hohes Alter erreichen müssten, um zeitgleich mit Ego zu leben, so dass hier ebenfalls nicht vom Anstieg der Lebenserwartung profitiert werden kann (vgl. ebd.: 19). Für mittlere Alter von Ego aber befinden sich Vorfahren zumeist in den Altersbereichen, die am meisten vom Anstieg der Lebenserwartung profitieren (vgl. ebd.: 19).

#### 4.4.4 Diskussion

Betrachtet man die vorgestellten Arbeiten in Hinblick auf die in Unterabschnitt 4.1 zuerst genannte Fehlerquelle der Unsicherheit der demographischen Komponenten, zeigt sich, dass diese nicht berücksichtigt wird. Insbesondere wird in allen Arbeiten von konstanten Raten ausgegangen. Wie im vorausgegangenen Unterabschnitt deutlich wurde, ist dadurch einerseits der Datenbedarf für die Umsetzung des ursprünglichen Ansatzes von Goodman et al. (1974) relativ gering. Andererseits zeigt die im letzten Unterabschnitt diskutierte Validierung mittels Umfragedaten, dass die bei einer solchen Annahme resultierenden Ergebnisse eher nicht realistisch sind. Eine Erweiterung des Modells auf zeitveränderliche demographische Komponenten wird in der Literatur darüber hinaus als möglich erachtet (vgl. Keyfitz & Caswell 2005: 388).

Das Zwei-Geschlechter-Problem sowie das Problem unberücksichtigter Heterogenität werden zwar in der Literatur behandelt, allerdings geht mit den Vorschlägen zur Lösung dieser Probleme ein erhöhter Datenbedarf einher. Beispielsweise stellt sich beim Ansatz von Coresh & Goldman (1988) die Frage, woher „typenspezifische“ Geburtenraten bezogen werden könnten. Zudem ist fraglich, ob eine Betrachtung der lateralen Verwandtschaft bei Einführung von zwei Geschlechtern möglich ist. Zwar könnte beispielsweise die Zahl lebender Geschwister über den Vater berechnet werden. Allerdings wäre fraglich, wie man einen solchen Wert mit der erwarteten Zahl der lebenden Geschwister über die Mutter in Einklang bringen kann – wie also beispielsweise Halbgeschwister berücksichtigt werden können.

Daneben existieren etliche weitere Schwierigkeiten. Eine mögliche Abhängigkeit von Personen untereinander wird nicht berücksichtigt. Zudem ist die Annahme notwendig, dass einzelne Ereignisse voneinander unabhängig sind. Affiniale Verwandtschaft wird nicht berücksichtigt und der ursprüngliche Ansatz von Goodman et al. (1974) setzt für die Ermittlung lateraler Verwandtschaft voraus, dass die durchschnittliche Kinderzahl gleich der Varianz der Kin-

derzahl ist. Weiterhin liefern die vorgestellten Formeln immer nur Durchschnittswerte, so dass beispielsweise im Falle der Kinderzahl keine Auskunft darüber geben, wie viele Personen kinderlos sind. Schließlich wird Migration nicht berücksichtigt, womit nur geschlossene Bevölkerungen betrachtet werden können.

## 4.5 Multi-State Ansätze

### 4.5.1 Grundlagen

Multi-State-Tafeln stellen eine Verallgemeinerung der Kohorten-Komponenten-Methode dar, bei der neben dem Alter noch weitere Merkmale berücksichtigt werden (vgl. Land & Rogers 1982: 2).<sup>86</sup> Aufgrund der allgemeinen Formulierung dieses Ansatzes, welche die Modellierung von Übergängen zwischen beliebigen Zuständen erlaubt, bieten sich diverse Anwendungsmöglichkeiten, beispielsweise bei der Vorausberechnung der Humankapitalentwicklung (vgl. Lutz & Goujon 2001), der Untersuchung von Lebensformen junger Erwachsener (vgl. Schoen et al. 2007) oder der Analyse des Arbeitsangebotes (vgl. Schoen & Woodrow 1980).

Formale Grundlagen des Multi-State Ansatzes wurden in den siebziger Jahren gelegt (vgl. insb. Hoem 1977; Rogers 1975; Rogers & Ledent 1976; Schoen & Nelson 1974; Schoen 1975; Schoen & Land 1979). Grundlegend für den Multi-State Ansatz ist dabei die Theorie der Markov-Prozesse.<sup>87</sup> Ausgegangen wird von einem inhomogenen, zeitstetigen Markov-Prozess mit endlichem Zustandsraum  $I$  (vgl. Schoen & Land 1979: 763).  $I$  umfasst  $k + 1$  Zustände, wobei  $k \in \mathbb{N}$  und  $k > 1$  ist, von denen mindestens einer absorbierend ist und den Zustand „tot“ repräsentiert (vgl. ebd.: 763). Von den verbleibenden  $k$  Zuständen kommunizieren mindestens zwei miteinander (vgl. ebd.: 763). Über den Zustandsraum  $I$  wird ein stochastischer Prozess  $\{S(x) : x > 0\}$  definiert, wobei  $x$  zeitstetig ist und das exakte Alter in Jahren erfasst (vgl. ebd.: 764).

Für Analysen und Vorausberechnungen sind die bedingten Übergangswahrscheinlichkeiten zwischen Zuständen  $i$  und  $j$  von Interesse:

$$\Pr(S(x + t) = j | S(x) = i) = p_{ij}(x, t) \quad (4.34)$$

Dies gibt die Wahrscheinlichkeit wieder, dass ein Individuum, welches im exakten Alter  $x$  in Zustand  $i$  ist, sich im Alter  $x + t$  in Zustand  $j$  befindet (vgl. Willekens 2005: 94). Hier wird deutlich, dass die Wahrscheinlichkeit für einen Wechsel nach Zustand  $j$  nur vom Alter und vom aktuellen Status abhängt (vgl. Schoen 1988a: 65). Die zuvor besetzten Zustände und die bisherige Verweildauer in Zustand  $i$  werden nicht berücksichtigt, womit der Prozess die Markov-Eigenschaft aufweist (vgl. ebd.: 65). Die Bevölkerung in einem bestimmten Alter

86 engl.: *increment-decrement table, multi-state table, multistate table* oder *multidimensional table*; diese englischsprachigen Bezeichnungen und deren deutschen Entsprechungen werden hier synonym verwendet; Übersichtsdarstellungen finden sich bei Rogers (1975, 1980) und Schoen (1988a,b).

87 Eine sehr stark an der mathematischen Theorie von Markov-Prozessen ausgerichtete Beschreibung von Inkrement-Dekrement-Tafeln findet sich bei Hoem & Jensen (1982). Eine Beschreibung über Stichprobenpfade gibt Mode (1982).

und Zustand wird als homogen betrachtet (vgl. ebd.: 65). Zudem setzt diese Vorgehensweise voraus, dass die einzelnen Mitglieder einer Bevölkerung unabhängig voneinander sind – eine Annahme, die in der Literatur oftmals nicht explizit genannt wird. Wenn  $t = 1$  ist, wird für  $p_{ij}(x, t)$  abkürzend  $p_{ij}(x)$  verwendet. Die Sterbewahrscheinlichkeit von Personen im Alter  $x$  und Zustand  $i$  wird mit  $p_{i\delta}(x, t)$  beziehungsweise  $p_{i\delta}(x)$  bezeichnet.

Die altersspezifischen Übergangswahrscheinlichkeiten lassen sich in einer Matrix zusammenfassen (vgl. Schoen 1988a: 65):

$$\mathbf{P}(x) = \begin{bmatrix} p_{11}(x) & p_{12}(x) & \dots & p_{1(k-1)}(x) & p_{1k}(x) \\ p_{21}(x) & p_{22}(x) & \dots & p_{2(k-1)}(x) & p_{2k}(x) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{k1}(x) & p_{k2}(x) & \dots & p_{k(k-1)}(x) & p_{kk}(x) \end{bmatrix} \quad (4.35)$$

Für die Zeilensumme der Zeile  $i$  gilt allgemein  $\sum_j p_{ij}(x) = 1 - p_{i\delta}(x)$ . Sterbewahrscheinlichkeiten werden also nur implizit berücksichtigt. Die einzelnen Matrizen für altersspezifische Übergangswahrscheinlichkeiten lassen sich wiederum in einer weiteren Matrix kombinieren (vgl. Keyfitz & Caswell 2005: 455):

$$\mathbf{P} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{P}(0)' & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}(1)' & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{P}(\rho-1)' & \mathbf{0} \end{bmatrix} \quad (4.36)$$

wobei  $\mathbf{P}(x)'$  die transponierte Matrix von  $\mathbf{P}(x)$  und  $\mathbf{0}$  eine  $(k \times k)$  Nullmatrix ist. Kennt man ferner die Bevölkerung zum Zeitpunkt  $t$ , welche durch einen Vektor

$$\mathbf{n}_t = (\mathbf{n}(0), \mathbf{n}(1), \dots, \mathbf{n}(\rho))' \quad (4.37)$$

repräsentiert wird, wobei

$$\mathbf{n}(x) = (n_1(x), n_2(x), \dots, n_k(x))' \quad (4.38)$$

und  $n_i(x)$  die Anzahl der Personen im Alter  $x$  in Zustand  $i$  erfasst, lässt sich die Bevölkerung einfach fortschreiben über (vgl. ebd.: 455)<sup>88</sup>

$$\mathbf{n}_{t+1} = \mathbf{P}\mathbf{n}_t \quad (4.39)$$

Soll zusätzlich Fertilität berücksichtigt werden, wird eine Matrix  $\mathbf{F}$

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}(0) & \mathbf{F}(1) & \dots & \mathbf{F}(\rho-1) & \mathbf{F}(\rho) \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (4.40)$$

<sup>88</sup> Die Einträge von  $\mathbf{n}_t$  beziehen sich genauer auf Personen im Altersintervall  $[x, x + 1)$ , während die Einträge von  $\mathbf{P}$  wie weiter oben definiert für exakte Alter gelten, also eigentlich nur für Personen im exakten Alter  $x$  und somit zu Beginn des Altersintervalls  $[x, x + 1)$ . Diese Diskrepanz wird in der Literatur zumeist nicht beachtet.

mit

$$\mathbf{F}(x) = \begin{bmatrix} f_{11}(x) & f_{12}(x) & \dots & f_{1(k-1)}(x) & f_{1k}(x) \\ f_{21}(x) & f_{22}(x) & \dots & f_{2(k-1)}(x) & f_{2k}(x) \\ \vdots & \vdots & \dots & \vdots & \vdots \\ f_{k1}(x) & f_{k2}(x) & \dots & f_{k(k-1)}(x) & f_{kk}(x) \end{bmatrix} \quad (4.41)$$

zur Matrix  $\mathbf{P}$  hinzu addiert (vgl. ebd.: 455). Dabei ist  $f_{ij}(x)$  eine Geburtenrate, die sich auf die Bevölkerung im Altersintervall  $x$  bis  $x + 1$  in Zustand  $i$  und Geburten dieser Gruppe bezieht, bei denen sich die Neugeborenen in Zustand  $j$  befinden.

Zur Bevölkerungsfortschreibung wird folgende Gleichung verwendet (vgl. Keyfitz & Caswell 2005: 455):

$$\mathbf{n}_{t+1} = (\mathbf{P} + \mathbf{F})\mathbf{n}_t \quad (4.42)$$

Die Bevölkerung in Zustand  $i$  im Alter  $x + 1$  setzt sich dann zusammen aus der Bevölkerung in Zustand  $i$  im Alter  $x$  abzüglich der Personen, die in Zustand  $i$  im Altersintervall  $x$  bis  $x + 1$  sterben oder aber in einen anderen Zustand  $j$  wechseln, plus die Zahl der Personen, die im Alter  $x$  in einem der anderen  $k - 1$  Zustände sind und zu Zustand  $i$  wechseln. Die Zahl der Geburten ergibt sich aus der Bevölkerung in einem bestimmten Alter und Zustand und der entsprechenden alters- und zustandsspezifischen Geburtenrate.

Bei dem bisher vorgestellten Vorgehen wird nicht explizit zwischen den Geschlechtern unterschieden. Ein einfaches *female dominance* Modell, welches diese Unterscheidung einführt, ist gegeben durch (vgl. Alho & Spencer 2005: 182):

$$\begin{bmatrix} \mathbf{n}_{t+1,w} \\ \mathbf{n}_{t+1,m} \end{bmatrix} = \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{P}_w & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_m \end{bmatrix} \begin{bmatrix} \mathbf{n}_{t,w} \\ \mathbf{n}_{t,m} \end{bmatrix} \quad (4.43)$$

Der Index  $w$  gibt an, dass sich der entsprechende Vektor beziehungsweise die entsprechende Matrix auf Frauen bezieht, der Index  $m$  steht analog für das männliche Geschlecht. Geburten werden in diesem Modell nur Frauen zugerechnet, so dass die Zahl der Geburten lediglich von der Zahl der Frauen in den einzelnen Altern und Zuständen und den entsprechenden Fertilitätsraten abhängt, während die Zahl der Männer keinen Einfluss hat.

Geburtenraten und Übergangswahrscheinlichkeiten müssen nicht zwingend konstant sein, sondern können sich für einzelne Zeitstellen ändern. Zudem kann auch Ein- und Auswanderung berücksichtigt werden. Ein einfaches Modell, welches diese beiden Punkte umfasst, lässt sich wie folgt darstellen (vgl. Alho & Spencer 2005: 186):

$$\begin{bmatrix} \mathbf{n}_{t+1,w} \\ \mathbf{n}_{t+1,m} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_t & \mathbf{0} \\ \mathbf{P}_{t,w} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{t,m} \end{bmatrix} \begin{bmatrix} \mathbf{n}_{t,w} \\ \mathbf{n}_{t,m} \end{bmatrix} + \begin{bmatrix} \mathbf{m}_{t,w} \\ \mathbf{m}_{t,m} \end{bmatrix} \quad (4.44)$$

Die Matrizen  $\mathbf{F}$ ,  $\mathbf{P}_m$  und  $\mathbf{P}_w$  erhalten zusätzliche Indizes, die angeben, auf welche Zeitstelle sich die Matrix bezieht. Zur Berücksichtigung von Migration wird analog zum Vorgehen bei der Leslie-Matrix ein Vektor hinzu addiert, der den Netto-Effekt von Migration nach Geschlecht, Alter und weiteren Merkmalen erfasst.

#### 4.5.2 Ein Ansatz zur Vorausberechnung von Verwandtschaft im Haushaltskontext

Ein auf den im vorhergegangenen Abschnitt vorgestellten Überlegungen basierender Ansatz, welcher die Vorausschätzung von Verwandten erlaubt, wurde von Zeng Yi (1986) und Zeng Yi et al. (1997, 1998, 2003) entwickelt und erweitert und basiert auf Überlegungen von Brass (1983) und Bongaarts (1987). Anwendungen dieses Verfahrens mit deutschen Daten finden sich bei Hullen (2002, 2004, 2006) und Hullen et al. (2004) (s. Unterabschnitt 4.5.3). Ziel des Ansatzes ist die Vorausberechnung der Zahl, Verteilung und Zusammensetzung von Familienhaushalten. Hierdurch ist die Vorausschätzung von Verwandtschaft auf nahe Verwandte, die in einem gemeinsamen Haushalt leben, beschränkt.<sup>89</sup>

Trotz des Fokus auf Haushalte setzt das Verfahren auf der Ebene des Individuums an. Die Grundidee ist, dass ein Mitglied jedes Haushaltes als sogenannter *marker* fungiert (vgl. Brass 1983: 39). Dies meint, dass eine Referenzperson festgelegt und deren Lebensverlauf verfolgt wird (vgl. Zeng Yi 1986: 679). Die Zahl der Haushalte ist dann gleich der Zahl der *marker* und Haushaltscharakteristika werden über Merkmale der Referenzpersonen hergeleitet (vgl. ebd.: 679). In der Regel dient eine erwachsene Frau als Referenzperson, während Männer nur für den Fall als *marker* gewählt werden, wenn keine Frau im Haushalt lebt (vgl. Zeng Yi et al. 1998: 63). Durch bestimmte Ereignisse, wie beispielsweise den Auszug aus dem elterlichen Haushalt, kann eine Person zum *marker* werden, während wiederum das Zusammenziehen mit einem Partner für eine der beiden Personen den Wechsel zum *non-marker* Status bedeutet (vgl. Brass 1983: 39).

Im weiteren werden bei den Mitgliedern der betrachteten Bevölkerung folgende Merkmalen und Merkmalsausprägungen berücksichtigt (vgl. Zeng Yi et al. 1997: 190):

- *Alter*: 0 bis  $\rho$
- *Geschlecht*: männlich; weiblich
- *Familienstand*: Single; verheiratet; geschieden; verwitwet; nicht-eheliche Lebensgemeinschaft
- *Koresidenz mit Eltern*: mit beiden Eltern; mit einem Elternteil; ohne Eltern; Haushalt, in dem keinerlei verwandtschaftliche Beziehungen bestehen;
- *Parität*: 0 bis zu einem maximalen Wert  $n$
- *Zahl der koresidierenden Kinder*: 0 bis zu einem maximalen Wert  $\kappa$
- *Region / Wohnort*: Stadt; Land

Wird  $n = 5$  gesetzt ergeben sich ohne Berücksichtigung von Alter und Geschlecht insgesamt 630 Merkmalskombinationen, die als Zustände verwendet werden (vgl. ebd.: 192).

---

<sup>89</sup> Ein dem hier vorgestellten Modell ähnlicher Ansatz, der ebenfalls auf die Modellierung der Haushaltsentwicklung ausgerichtet ist, wird von Imhoff (1995) und Imhoff & Keilmann (1991) beschrieben. Eine Anwendung dieses Modells auf deutsche Daten zur Vorausberechnung der Verfügbarkeit von Ehepartnern im hohen Alter als mögliches Unterstützungspotential im Falle von Pflegebedürftigkeit findet sich bei Dinkel et al. (1997). Eine Vorausberechnung, die auf dem selben Modell beruht und dem im nächsten Unterabschnitt vorgestellten Ansatz ähnlich ist, präsentieren Dorbritz et al. (1997). Anwendungen des von Bongaarts (1987) vorgeschlagenen Modells werden von Menken (1985) und Watkins et al. (1987) sowie Lee & Palloni (1992) beschrieben. Ein Modell, welches sich auf Kinder und deren Verwandte im Haushaltskontext beschränkt, wurde von Hofferth (1987) entwickelt.

Ob eine Person *marker* oder *non-marker* ist, wird hierbei also nicht explizit modelliert und aus den oben angegebenen Merkmalen hergeleitet. Die Zahl der Singlehaushalte von Frauen ist beispielsweise gleich der Zahl der Frauen die (1) entweder Single, verwitwet oder geschieden sind, (2) nicht mit ihren Eltern zusammenleben, (3) eine beliebige Parität aufweisen, (4) mit keinem ihrer gegebenenfalls vorhandenen Kinder zusammenleben und (5) einen beliebigen Wohnort aufweisen (vgl. Zeng Yi et al. 1997: 191). Die Zahl der Drei-Generationen-Haushalte bestehend aus Großeltern, Eltern und Kindern als weiteres Beispiel ist gegeben durch die Zahl der Frauen, die (1) entweder verheiratet sind oder in einer nicht-ehelichen Lebensgemeinschaft leben, (2) mit mindestens einem Elternteil zusammenleben, (3) deren Parität größer Null ist, (4) mit mindestens einem Kind zusammenleben und (5) einen beliebigen Wohnort aufweisen (vgl. ebd.: 191). Hierbei wird davon ausgegangen, dass (Ehe-)Partner grundsätzlich in einem gemeinsamen Haushalt leben. Ferner wird bei der Koresidenz mit Eltern nicht differenziert, ob dies die Eltern des Mannes oder der Frau sind.

Insgesamt werden zwölf Haushaltstypen unterschieden (vgl. Zeng Yi et al. 1997: 191). Um die absolute und relative Verteilung der Haushaltstypen über die gerade beschriebene Methode berechnen zu können, müssen etliche Schwierigkeiten gelöst werden.

Da bei  $n = 5$  insgesamt 630 Zustände für beide Geschlechter und Alter 0 bis  $\rho$  betrachtet werden, ist eine große Zahl an Übergangswahrscheinlichkeiten zu schätzen (vgl. Zeng Yi et al. 1997: 194). Wird  $\rho = 100$  gesetzt, enthält die Matrix der Übergangswahrscheinlichkeiten unter Berücksichtigung beider Geschlechter über 15 Milliarden Einträge. Von diesen wird zwar ein Großteil gleich Null sein beziehungsweise vorab auf Null gesetzt werden können, da bestimmte Übergänge nicht möglich oder zumindest unplausibel sind. Nichtsdestotrotz dürfte die Zahl der zu schätzenden Wahrscheinlichkeiten noch erheblich sein. Um dieses Problem zu lösen, wurde von Zeng Yi auf einen Ansatz von Bongaarts (1987) zurückgegriffen. Hierbei wird davon ausgegangen, dass bestimmte Zustandswechsel zu verschiedenen Zeitpunkten im Verlaufe eines Jahres stattfinden und bestimmte Wechsel von Merkmalen unabhängig voneinander sind (vgl. Bongaarts 1987: 209). Genauer wird davon ausgegangen, dass alle Zustandswechsel in der Mitte eines Jahres stattfinden, bis auf Zustandswechsel der Parität (vgl. Zeng Yi et al. 1997: 194). Diese sind in der ersten und zweiten Hälfte des Jahres möglich, wobei sich die verwendeten Wahrscheinlichkeiten auf ein halbes Jahr beziehen und vom Alter, Geschlecht, der Parität und dem Familienstand abhängen (vgl. ebd.: 194ff.). Der Wechsel des Familienstandes wird entsprechend einer vereinfachten Multi-State Tafel berechnet, die neben dem Familienstand, Alter und Geschlecht keine weiteren Merkmale, dafür aber das Sterberisiko berücksichtigt (vgl. ebd.: 198). Zustandswechsel in der Koresidenz mit Eltern beziehungsweise Kindern werden als hiervon unabhängig aufgefasst und jeweils separat, ebenfalls mit einfachen, eigenen Multi-State-Tafeln berechnet (vgl. ebd.: 198). Veränderungen der Koresidenz mit Eltern berücksichtigen die Wahrscheinlichkeit, dass Personen mit ihren Eltern beziehungsweise mit einem Elternteil zusammenziehen, die Sterbe- und Scheidungswahrscheinlichkeit der Eltern sowie die Wahrscheinlichkeit für eine Wiederheirat von geschiedenen beziehungsweise verwitweten Elternteilen (vgl. Zeng Yi et al. 1997: 198). Bei der Koresidenz mit Kindern wird die Sterbewahrscheinlichkeit von Kindern sowie deren Auszugswahrscheinlichkeit berücksichtigt (vgl. ebd.: 198).<sup>90</sup>

---

90 Für eine genauere Beschreibung des Vorgehens vgl. Zeng Yi et al. (1997: 211ff.) sowie Bongaarts (1987: 209ff.).

Ein weiteres Problem ist die Konsistenz von Übergängen und der Zahl der Personen in bestimmten Zuständen (vgl. Zeng Yi et al. 1997: 194). Beispielsweise sollte die Zahl verheirateter Frauen der Zahl verheirateter Männer entsprechen. Ebenso sollte die Zahl der Übergänge von Frauen vom Familienstand verheiratet zu geschieden mit der Zahl der Übergänge der Männer zwischen diesen Merkmalen übereinstimmen. Um Konsistenz bezüglich des Familienstandes zu erreichen, wird der *harmonic means*-Ansatz verwendet.<sup>91</sup> Ein abgewandelte Variante wird angewendet, um Konsistenz bezüglich der Koresidenz von Eltern und Kindern zu erreichen (vgl. ebd.: 195). Geht man beispielsweise von einem Paarhaushalt mit einem Kind aus, muss bei einem Auszug des Kindes der Status von drei Personen geändert werden: der Zustand des Kindes bezüglich der Koresidenz mit den Eltern und der Zustand von Vater und Mutter bezüglich der Koresidenz mit Kindern.

Weiterer Konsistenzbedarf ergibt sich bei der Zahl der Geburten, die Männern zugeordnet werden, und der Zahl der Geburten, die Frauen zugeordnet werden. Diese sollten identisch sein. Bei Zeng Yi et al. (1997) wird davon ausgegangen, dass für Männer *parity progression ratios* fehlen, so dass Veränderungen in der Parität der Männer nicht ohne weiteres modelliert werden können (vgl. ebd.: 196).<sup>92</sup> Um diese beiden Probleme zu lösen, wird zunächst die Zahl der Geburten in der Bevölkerung insgesamt über die Übergangswahrscheinlichkeiten von Frauen berechnet (vgl. ebd.: 196). Fertilitätsraten für Männer werden dann über die sich ergebenden Geburtenraten für Frauen unter Verwendung des durchschnittlichen Altersunterschiedes zwischen Männern und Frauen bei der Geburt eines Kindes berechnet, so dass die Gesamtzahl der Geburten von Frauen der Gesamtzahl der Geburten von Männern entspricht (vgl. ebd.: 196). Dieses Vorgehen zur Erreichung der Konsistenz der Geburtenzahlen führt allerdings nur zu einer Übereinstimmung der Gesamtzahl der Geburten von Männern und Frauen. Wie die Geburten auf einzelne Männer verteilt werden – wie sie also die Verteilung der Paritäten von Männern beeinflussen – wird in keiner der hier genannten Publikationen genauer beschrieben.

Schließlich muss noch Konsistenz bezüglich der Zahl von (koresidierenden) Kindern bei Scheidungen und Wiederheiraten erreicht werden. Bei Scheidungen werden die Kinder bei einer geraden Zahl an Kindern zu gleichen Teilen auf die Elternteile aufgeteilt, bei einer ungeraden Kinderzahl wird der Frau ein Kind mehr zugesprochen (vgl. Zeng Yi et al. 1997: 196). Bei einer Wiederheirat – also einem Zustandswechsel von „geschieden“ oder „verwitwet“ zu „verheiratet“ – wird die Zahl der durch den neuen Partner hinzukommenden, koresidierenden Kinder über die Verteilung der Zahl an koresidierenden Kindern aller im entsprechenden Jahr neu verheirateten errechnet (vgl. ebd.: 197).<sup>93</sup> Wenn beispielsweise 10% aller neu verheirateten Frauen bereits vor der Heirat mit einem Kind zusammenleben, beträgt bei wiederverheirateten Männern die Wahrscheinlichkeit, aus der neu eingegangenen Ehe mit einem zusätzlichen Kind zusammen zu leben, 10%.

---

91 Dieser und weitere Ansätze zur Lösung des Zwei-Geschlechter-Problems bei Multi-State-Ansätzen werden von Schoen (1988a: 120ff.) beschrieben. Allgemeine Konsistenzalgorithmen zur Einhaltung von Randbedingungen werden von Imhoff (1992) und Keilman (1985) diskutiert.

92 Unter *parity progression ratios* bzw. paritätsspezifischen Geburtenraten versteht man altersspezifische Geburtenraten, bei denen nach der Kinderzahl (Parität) differenziert wird (vgl. Preston et al. 2001: 104). Beispielsweise kann man Geburtenraten für Frauen, die bereits ein Kind haben, angeben. Diese beziehen sich somit nur auf die Geburt eines zweiten Kindes.

93 Der Begriff der „neu verheirateten“ umfasst sowohl erstmalig verheiratete Personen wie auch wiederverheiratete Personen (vgl. Zeng Yi et al. 1997: 187).

Zur Modellierung von Migration werden alters- und geschlechtsspezifische Ein- und Auswanderungsraten verwendet. Die Verteilung der weiteren Merkmale von Ein- und Auswanderern entspricht der Verteilung dieser Merkmale in den einzelnen Altersgruppen der beiden Geschlechter in der betrachteten Bevölkerung (vgl. Zeng Yi et al. 1997: 201). Soll interne Migration berücksichtigt werden, werden für die beiden berücksichtigten Gebiete (Stadt/Land) zusätzlich zustandsspezifische Migrationsraten benötigt (vgl. ebd.: 201). Im Fall der Modellierung interner Migration werden die in den vorausgegangenen Absätzen angegebenen Übergangswahrscheinlichkeiten zusätzlich auf das Gebiet konditioniert (vgl. ebd.: 201).<sup>94</sup>

Durch die geschilderten Lösungsansätze und Verfahrensweisen werden etliche Modellannahmen eingeführt: Zustandswechsel finden in der Jahresmitte statt, während Geburten in der ersten und zweiten Jahreshälfte auftreten können. Diese sind unabhängig voneinander, was zwei Geburten innerhalb eines Jahres erlaubt. Zustandswechsel sind ferner teilweise voneinander unabhängig, beispielsweise ist die Koresidenz mit Kindern unabhängig vom Familienstand. Die Zahl der Geburten hängt von der Zahl der Frauen und ihren zustandsspezifischen *parity progression ratios* ab, während Geburten von Männern aus diesen abgeleitet werden. Zudem wurden Annahmen zur Aufteilung von Kindern bei Scheidungen und Wiederheiraten getroffen. Zur Berechnung der Zahl der Haushalte ist es nötig, dass (Ehe-)Paare grundsätzlich zusammen leben. Schließlich geht dieses Modell von Homogenität von Personen in demselben Zustand aus. Die Markov-Eigenschaft wird hingegen – trotz des Verweises von Zeng Yi et al. (vgl. 1997: 200) auf diese – teilweise durch die Umsetzung der Konsistenzbedingungen gelockert. Beispielsweise hängen Zustandswechsel bezüglich des Merkmals Familienstand durch die Verwendung des *harmonic means*-Ansatzes vom Verhältnis von Männern und Frauen in der Bevölkerung ab – also nicht nur vom momentan besetzten Zustand, sondern auch von der Verteilung der Geschlechter in der Bevölkerung.

### 4.5.3 Anwendungsbeispiel

Anwendungen des im vorherigen Abschnitt vorgestellten Modells für Deutschland wurden von Hullen (2002, 2004, 2006) und Hullen et al. (2004) vorgestellt. Beispielhaft seien hier die Ergebnisse aus Hullen (2006) vorgestellt, da diese dem Thema der vorliegenden Arbeit am nächsten sind.

Gegenstand ist die zukünftige Entwicklung der „Lebensformen“ von Kindern (vgl. Hullen 2006: 1f.). Unter dem Begriff „Lebensformen“ versteht Hullen Kombinationen von Merkmalsausprägungen folgender Personenmerkmale (vgl. Hullen 2002: 79):

- Vorhandensein eines eigenen Haushaltes
- Generationenzusammensetzung des Haushaltes
- Familienstand
- Kinderzahl

Als Kinder werden minderjährige Personen aufgefasst, die ledig sind und mit ihren Eltern

---

94 Unter „interner Migration“ versteht man allgemein Wanderungen innerhalb eines bestimmten Gebietes, während Wanderungen über Gebietsgrenzen (bspw. Staatsgrenzen) hinweg als „externe Migration“ bezeichnet werden (vgl. bspw. White & Lindstrom 2006).

*Tabelle 4.3:* Anzahl der Geschwister von Kindern, die mit einer (Stief-)Mutter zusammen im Haushalt leben (in Prozent)

Jahr	West			Ost		
	0	1	2 $\geq$	0	1	2 $\geq$
2004	24	46	30	33	43	25
2015	27	44	29	28	43	29
2030	22	46	32	14	39	47

Quelle: Hullen (2006: 23)

zusammenleben (vgl. Hullen 2006: 3). Mit dem Auszug aus dem elterlichen Haushalt werden Personen nicht mehr als Kinder betrachtet und bei der Analyse nicht weiter berücksichtigt (vgl. ebd.: 3).

Zur Bestimmung der Ausgangsbevölkerung wurde auf das Scientific-Use-File des Mikrozensus aus dem Jahr 2004 zurückgegriffen, der Vorausberechnungszeitraum reicht bis 2030 (vgl. Hullen 2006: 10). Bei der Vorausberechnung wurde zwischen der demographischen Entwicklung in West- und Ostdeutschland unterschieden (vgl. ebd.: 13). Annahmen über die Entwicklung der demographischen Komponenten wurden aus der zehnten koordinierten Bevölkerungsvorausberechnung des Statistischen Bundesamtes übernommen.<sup>95</sup> Interne Migration zwischen West- und Ostdeutschland wird von Hullen nicht erwähnt, weshalb davon ausgegangen wird, dass sie nicht berücksichtigt wurde. Weitere benötigte Modellparameter wurden aus dem „Generations and Gender Survey“, dem „Fertility and Family Survey“ und dem Mikrozensus-Panel geschätzt (vgl. ebd.: 11f.). Hierzu zählen beispielsweise die Heirats- und die Scheidungswahrscheinlichkeit (vgl. ebd.: 12). Für diese und weiteren Größen wurde angenommen, dass sie sich während des Zeitraums der Vorausberechnung nicht ändern (vgl. ebd.: 12f.).

Bezogen auf Kinder stellt Hullen unter anderem Ergebnisse für die Zahl der Elternteile, mit denen diese zusammenleben, sowie die Zahl der Geschwister im selben Haushalt vor (vgl. Hullen 2006: 15ff.). Die Zahl der Kinder, die nur mit einem Elternteil zusammenleben, wird zwar relativ gesehen steigen, allerdings leben nichtsdestotrotz immer noch 90% aller Kinder mit zwei Elternteilen zusammen (vgl. Hullen 2006: 15f.). Hier sei allerdings darauf hingewiesen, dass diese von Hullen angegebenen Zahlen leicht irreführend sein können. Das verwendete Modell erlaubt keine Unterscheidung zwischen leiblichen Eltern und Stiefeltern, weshalb das Vorhandensein von zwei Elternteilen nicht zwingend das Zusammenleben mit den leiblichen Eltern erfasst.

Ergebnisse zur Geschwisterzahl finden sich in Tabelle 4.3. Ausgewiesen ist die prozentuale Verteilung der Geschwisterzahl aller Kinder, die mit einer (Stief-)Mutter zusammen im Haushalt leben. Vergleichbare Ergebnisse werden von Hullen zwar auch für Kinder ausgewiesen, die mit einem (Stief-)Vater im Haushalt leben, allerdings unterscheiden sich diese nur geringfügig von den hier angegebenen Resultaten (vgl. Hullen 2006: 18). Es sei nochmals darauf hingewiesen, dass lediglich Geschwister betrachtet werden, die zusammen in einem Haushalt leben. Die Zahl der Kinder, die zumindest zeitweise ohne Geschwister aufwachsen, wird in

<sup>95</sup> Für Details hierzu s. Hullen (2006: 10f.) und Statistisches Bundesamt (2003).

Westdeutschland bis 2030 leicht sinken, während hingegen die Zahl der Kinder mit zwei oder mehr Geschwistern leicht zunimmt. Wesentlich deutlichere Veränderungen zeigen sich für Ostdeutschland. Die Zahl der Kinder, die zumindest zeitweise ohne Geschwister im selben Haushalt aufwachsen, sinkt deutlich von 33% auf 14%. Der Anteil der Kinder mit mindestens 2 Geschwisterkindern hingegen verdoppelt sich fast.

#### 4.5.4 Diskussion

Ein grundsätzlicher Vorteil von Multi-State Modellen besteht in der Möglichkeit, die betrachtete Bevölkerung nach bestimmten Merkmalen – neben Alter und Geschlecht – zu differenzieren und somit die Homogenitätsannahme dahingehend abzuschwächen, dass sie nicht mehr für die Bevölkerung als Ganzes, sondern für einzelne Bevölkerungsgruppen gilt (vgl. Willekens 2005: 83). Eine weitergehende Lockerung lässt sich beispielsweise durch die Mischung von Markov-Prozessen für einzelne Bevölkerungssteile erreichen (vgl. Kitsul & Philipov 1982). Zudem lässt sich auch eine Abhängigkeit der Übergangswahrscheinlichkeiten von der bisherigen Verweildauer in einem Zustand (*duration-dependence*) einführen, indem eine Erweiterung des Zustandsraumes vorgenommen wird (vgl. Wolf 1988b). Allerdings basieren sowohl die Lockerung der Homogenitätsannahme als auch die Einführung von *duration-dependence* auf einer weitergehenden Disaggregation der Bevölkerung und führen zu einer Erhöhung der Zahl der zu schätzenden Übergangswahrscheinlichkeiten. Dabei kann in der Regel aufgrund eingeschränkter Datenverfügbarkeit keine beliebig weitgehende Disaggregation vorgenommen werden, so dass Heterogenität unter Umständen nicht hinreichend abgebildet werden kann (vgl. Hoem & Jensen 1982: 247ff.).<sup>96</sup> Dabei führt bereits die von Zeng Yi et al. (1997) vorgeschlagene Unterteilung der Bevölkerung anhand von sechs bis sieben Merkmalen zu einem immensen Datenbedarf und macht die Einführung etlicher Annahmen notwendig, wie die „Verteilung“ des möglichen Eintretens von Zustandsänderungen über ein Jahr und einer eingeschränkten Berücksichtigung von Merkmalen bei einzelnen Übergängen. Ferner ist unklar, wie Unsicherheit bezüglich der zukünftigen Entwicklung der Modellparameter berücksichtigt werden kann. Selbst die Erstellung von einzelnen Szenarien kann sich aufgrund der Vielzahl der Parameter als schwierig erweisen.<sup>97</sup> Bei Hullen (2006) beispielsweise wird bei der Vorausberechnung bei etlichen Übergangswahrscheinlichkeiten lediglich davon ausgegangen, dass sie konstant bleiben.

---

96 Als Ausweg werden zuweilen parametrische Modelle vorgeschlagen, allerdings besteht bei diesen ebenfalls das Problem, dass nicht beliebige viele Variablen berücksichtigt werden können und somit nicht beobachtbare Heterogenität bleiben kann (vgl. Hoem & Jensen 1982: 247ff.). Hannan (1984) schlägt die Verwendung der Ereignisanalyse als mögliche parametrische Alternative vor. Auf der Ereignisanalyse aufbauende Ansätze, die diese mit den Überlegungen des Multi-State-Ansatzes kombinieren, finden sich bei Fiocco et al. (2008), Gill (1992) und Putter et al. (2007). Ein Modell, welches nicht beobachtete Heterogenität berücksichtigen kann, wurde beispielsweise von Heckman & Singer (1982) vorgeschlagen. Dieses und ähnliche Modelle basieren allerdings oft auf restriktiven Annahmen – beispielsweise Verteilungsannahmen eines Modellterms für nicht beobachtbare Heterogenität (vgl. ebd.: 569) – so dass Galler (1995: 206) konstatiert, dass dieses Problem nicht zufriedenstellend gelöst sei.

97 Unter einem „Szenario“ versteht man allgemein die konkreten Annahmen einer Vorausberechnung bezüglich der zukünftigen demographischen Entwicklung. In der Literatur ist es nicht unüblich, nicht nur ein Szenario, sondern mehrere Vorausberechnungsszenarien zu berechnen und die Ergebnisse der einzelnen Varianten miteinander zu vergleichen (s. Abschnitt 5.1.1).

Weitere Vorteile des Multi-State Ansatzes sind im Vorhandensein von Lösungen des Zwei-Geschlechter-Problems und von generellen Konsistenzalgorithmen, der einfachen Berücksichtigung von Migration sowie den umfangreichen Ergebnissen, die dieser Ansatz liefert, zu sehen. Beispielsweise liefert das in den beiden vorhergegangenen Unterabschnitten beschriebene Modell nicht nur umfassende Ergebnisse zur Haushaltsentwicklung sondern auch zur Entwicklung der Bevölkerung insgesamt. Zudem könnte man beispielsweise die erwartete Lebenszeit in oder die Häufigkeit von Wechseln zwischen bestimmten Zuständen berechnen (vgl. hierzu Keyfitz & Caswell 2005: 452ff.).

Allerdings wird bei diesem Modell das Vorhandensein von bestimmten Verwandten nur im Haushaltskontext verfolgt. Leben zum Beispiel Eltern und ihre erwachsenen Kinder nicht zusammen in einem Haushalt, gibt das Modell keine Auskunft darüber, ob die Eltern noch leben. Zudem werden lediglich Großeltern, Eltern und Kinder berücksichtigt, so dass über die Zahl anderer Verwandter nichts ausgesagt werden kann. Dies gilt sowohl für weitere Verwandte im Haushalt als auch für das Vorhandensein von Verwandten allgemein. Auch weitere, nicht verwandte Personen im Haushalt werden nicht berücksichtigt. Diese beiden Faktoren führen dazu, dass Zeng Yi et al. (1998: 65) bei einer Validierung der Haushaltstypisierung mit dem Mikrozensus 1995 deutliche Abweichungen in der Verteilung der Größe der Haushalte feststellen. Verglichen wurde die tatsächlich im Mikrozensus festgehaltene Haushaltsgröße mit einer Anwendung der Haushaltstypisierung über *marker* auf diese Daten. Es zeigte sich, dass zum Beispiel die Zahl der Ein-Personen-Haushalte deutlich überschätzt wurde (vgl. ebd.: 65), wovon ausgehend die Autoren ein Korrekturverfahren entwickelten. Überraschenderweise stellt Hullen (2002: 95) bei gleichem Vorgehen und Verwendung des Mikrozensus 1996 nur geringfügige Abweichungen fest. Allerdings ist hierbei unklar, ob das von Zeng Yi et al. (1998) vorgeschlagene Korrekturverfahren angewendet wurde.

Ein weiteres Problem des Ansatzes ist, dass sich nicht erkennen lässt, ob es sich bei der im Haushalt erfassten Verwandtschaft um leibliche Verwandte oder Stiefverwandte handelt. Ferner wird lediglich das Zusammenleben mit bis zu zwei Großeltern erfasst, wobei einerseits prinzipiell eine höhere Zahl an Großeltern im Haushalt möglich ist und zudem nicht zwischen Großeltern als Eltern der Frau oder als Eltern des Mannes im Haushalt unterschieden wird. Aufgrund der geringen Verbreitung von Drei-Generationen Haushalten in Deutschland (vgl. Hullen 2002: 96; sowie Unterabschnitt 2.2.1) dürfte dieses Problem aber keinen Einfluss auf die Güte der Ergebnisse haben.

Der Fokus auf Haushalte ist insofern problematisch, als dass der Haushaltsbegriff in der Literatur verschieden definiert wird und allgemein keine klare Abgrenzung von einzelnen Haushalten untereinander beziehungsweise eine trennscharfe Zuordnung von Personen zu bestimmten Haushalten erlaubt (vgl. bspw. Ott 1995: 178ff.). Beispielsweise ist der Auszug von Kindern aus dem elterlichen Haushalt teils schwierig zu erfassen, da ein Auszug nicht unbedingt ein fixiertes Ereignis ist, sondern schrittweise erfolgen kann (vgl. ebd.: 178). Im hier vorgestellten Modell wird diese Unschärfe nicht berücksichtigt. Eine umfassendere Diskussion des Problems der Abgrenzung von Haushalten findet sich zum Beispiel bei Keilman (1995).

Auch bei der Erreichung von Konsistenz innerhalb des Modells bei der hier vorgestellten Variante stellen sich mehrere Fragen. Zum einen wird in keiner der verwendeten Publikationen expliziert, wie beispielsweise Ein- und Auswanderung von verheirateten Personen und Personen, die in nicht-ehelichen Lebensgemeinschaften leben, berücksichtigt wird. Wenn dies

nicht beachtet wird, kann die Zahl der Frauen in Partnerschaften gegebenenfalls nicht mehr der Zahl der Männer in Partnerschaften entsprechen. Zum anderen ist nicht klar, wie genau Paare „weiterverfolgt“ werden, da das beschriebene Vorgehen nur zu Konsistenz der Zahl der Partnerschaften und der Zahl der Geburten insgesamt führt. Ein vereinfachendes Beispiel: Angenommen, eine verheiratete Frau bekommt ihr zweites Kind und wechselt somit von Parität 1 zu 2. Bei welchem Mann erhöht sich dann ebenfalls die Parität? Es scheint naheliegend, dass dieser ebenfalls verheiratet ist (was allerdings keine zwingende Voraussetzung ist), allerdings ist dessen Parität unklar. Denn dieser Mann könnte zum Beispiel in zweiter Ehe verheiratet sein und bereits aus erster Ehe 2 Kinder haben, so dass seine Parität nicht unbedingt 1 betragen muss.

Schließlich seien noch zwei grundsätzliche Probleme, die bei der Verwendung von Markov-Prozessen zur Modellbildung auftreten können, erwähnt. Der Multi-State-Ansatz basiert auf der Annahme eines zeitstetigen Markov-Prozesses, der als mathematisches Fundament dient. Für die Ermittlung der Übergangswahrscheinlichkeiten verwendbare Daten sind aber grundsätzlich zeitdiskret. Dabei muss nicht notwendigerweise ein stetiger Markov-Prozess existieren, der zu den beobachteten Daten hätte führen können (vgl. Hoem & Jensen 1982: 200). Die Frage nach dem Vorhandensein eines oder mehrerer „passender“ Prozesse wird als Einbettbarkeit bezeichnet (vgl. Hoem & Jensen 1982: 200).<sup>98</sup> Das also die für die Formulierung des Modells getroffenen Annahmen im Anwendungsfall zutreffen ist nicht immer gewährleistet. Erschwerend kommt hinzu, dass die Modellparameter oftmals über Stichprobendaten ermittelt werden müssen und deshalb mit einem Schätzfehler behaftet sind (vgl. Singer & Spilerman 1976a: 4). Sollte eine Übergangsmatrix nicht einbettbar sein, ist unklar, ob dies auf falsche Annahmen oder aber auf Schätzfehler zurückzuführen ist (vgl. ebd.: 4). Sollte die empirisch geschätzte Übergangsmatrix einbettbar sein, ist weiterhin nicht garantiert, dass sich ein bestimmter, zugrunde liegender Prozess eindeutig identifizieren lässt (vgl. Singer & Spilerman 1976b: 463). Gegebenenfalls existieren mehrere stochastische Prozesse, die die beobachteten Daten hätten erzeugen können (vgl. ebd.: 463).

## 4.6 Mikrosimulationen

### 4.6.1 Grundlagen

Im Gegensatz zu den anderen hier vorgestellten Modellierungsstrategien setzen Mikrosimulationen direkt an der Mikroebene an. Die Grundidee wurde von Orcutt (1957) eingeführt.<sup>99</sup> Modelliert werden einzelne Mikroeinheiten, wie z.B. Personen oder auch Unternehmen, wobei in demographischen Anwendungen in der Regel Personen betrachtet werden (vgl. Ruggles 1990: 24). Ausgehend von einer realen oder hypothetischen Ausgangsmenge von Personen,

---

98 engl.: *embeddability*; eine ausführliche Darstellung dieses Problems und des Identifikationsproblems finden sich bei Singer & Spilerman (1976b,a). Ein Test auf Einbettbarkeit für 2-Zustands-Prozesse, der mögliche Schätzfehler berücksichtigt, wurde von Singer & Cohen (1980) vorgeschlagen. In der empirischen Literatur wird in der Regel nicht auf diese Probleme eingegangen (eine Ausnahme findet sich bei Keilman 1995: 284ff.).

99 Vgl. auch Orcutt et al. (1965) und Orcutt et al. (1976).

für die bestimmte interessierende Merkmale bekannt sind, lässt sich durch die Verwendung von probabilistischen Regeln eine Fortschreibung vornehmen. Dabei wird für jedes Mitglied der betrachteten Menge anhand der Ziehung einer Zufallszahl aus einer Wahrscheinlichkeitsverteilung ermittelt, ob ein Ereignis beziehungsweise Zustandswechsel eintritt oder nicht (vgl. Orcutt 1957: 117). Nachdem eine Mikroeinheit eine Zeiteinheit fortgeschrieben wurde, lässt sich dieses Vorgehen abermals durchführen, so dass durch häufiges Wiederholen ganze Lebensverläufe simuliert werden können. Diese simulierten Lebensverläufe einzelner Personen sind dabei zwar zufälligen Schwankungen unterworfen, allerdings gleichen sich diese bei einer großen Zahl von Lebensverläufen aus (vgl. Galler 1990: 72). Nach der Durchführung der Fortschreibung hat man dann Daten über Personen am Ende des Vorausberechnungszeitraums, wie auch über individuelle Verläufe von Beginn bis zum Ende der Vorausberechnung, welche sich sowohl in einer Querschnitts- als auch Längsschnittperspektive auswerten lassen, wobei auch eine Aggregation der Daten stattfinden kann (vgl. Ruggles 1990: 25).<sup>100</sup>

Mikrosimulationen können je nach Fragestellung sehr unterschiedlich gestaltet sein und in einzelnen Punkten auch von der obigen Beschreibung abweichen (vgl. Galler 1994a: 371). Eine wichtige Unterscheidung kann zwischen statischen und dynamischen Mikrosimulationen getroffen werden (vgl. bspw. Gilbert & Troitzsch 1999: 56). Bei statischen Mikrosimulationen werden zwei oder mehr Situationen miteinander verglichen, allerdings nicht die Übergänge zwischen diesen (vgl. Galler 1994a: 372). Unterschiedliche Situationen werden erzeugt, indem die Merkmalsverteilung in den Ausgangsdaten durch Gewichtung verändert wird oder durch Variation in exogenen Variablen (vgl. Gilbert & Troitzsch 1999: 56). Beispielsweise könnte der eigentliche Datensatz Informationen zu Einkommen enthalten, welche unverändert bleiben, während Steuersätze, aus denen die individuelle Steuerlast berechnet wird, variiert werden, wodurch das Steueraufkommen bei verschiedenen Sätzen simuliert werden kann (vgl. Galler 1994a: 372). Für ein Beispiel für eine statische Simulation sei auf Merz (1994) verwiesen.

Dynamische Mikrosimulationen lassen sich ferner in Querschnitts- und Längsschnittsimulationen untergliedern (vgl. Gilbert & Troitzsch 1999: 57). Erstere benutzen als Ausgangsdaten Querschnittsstichproben, die fortgeschrieben werden, während bei Längsschnittsimulationen Lebensverläufe innerhalb einer einzelnen Kohorte simuliert werden (vgl. Galler 1994a: 373). Die Längsschnittvariante ist zwar bezogen auf Aussagemöglichkeiten auf die simulierte Kohorte beschränkt, welche in der Regel synthetisch ist, allerdings lässt sich hierdurch der nötige Rechenaufwand reduzieren (vgl. Helberger & Wagner 1994: 444).

Spielauer & Vencatasawamy (2003) nennen als weitere Unterscheidungsmöglichkeit Datenbasierung und Theoriebasierung. Datenbasierte Mikrosimulationen stützen sich bei der Modellierung auf empirische Daten – also empirisch gewonnene Stichproben und Parameter (vgl. ebd.: 144). Theoriebasierte Modelle hingegen basieren auf abstrakten Überlegungen und zielen nicht auf die Simulation einer realen Bevölkerung, sondern auf die Analyse der Implikationen bestimmter Regeln (vgl. ebd.: 144). Als Beispiel sind hier insbesondere Agentenbasierte Modelle zu nennen (vgl. hierzu Gilbert & Troitzsch 1999: Kap. 8).

Für Mikrosimulationen von Verwandtschaft gilt grundsätzlich, dass sie datenbasierte, dynamische Querschnittssimulationen sind (s. die Quellen im nächsten Unterabschnitt). Ansonsten unterscheiden sich verschiedene Ansätze aber oftmals relativ deutlich.

---

100 Dies macht allerdings wie bei der Auswertung von Umfragedaten gegebenenfalls Hochrechnungsfaktoren notwendig; für eine Berechnung solcher vgl. bspw. Nakamura & Nakamura (1978).

#### 4.6.2 Varianten der Mikrosimulation von Verwandtschaft

Die Mikrosimulation von Verwandtschaft wurde zuerst im Rahmen von Fragestellungen der Anthropologie und der historischen Demographie eingesetzt, wobei in beiden Fällen nicht die Vorausberechnung von Verwandtschaftsnetzwerken im Mittelpunkt stand.<sup>101</sup> In den anthropologischen Arbeiten lag das Interesse auf den Effekten unterschiedlicher Heiratssysteme und insbesondere Cousinenheiraten in kleinen Bevölkerungen (vgl. bspw. Kunstader et al. 1963; Levy & Westoff 1965; Gilbert & Hammel 1966; Howell & Lehotay 1978; Hammel 2005), während bei historisch orientierten Analysen vor allem auf den Zusammenhang von Verwandtschafts- und Haushaltsstrukturen abgehoben wurde, wobei als prominentestes Beispiel das präindustrielle England zu sehen ist (vgl. bspw. Wachter et al. 1978; Smith 1987; Smith & Oeppen 1993; Zhongwai Zhao 1996; Murphy 2004; Zhongwai Zhao 2006). Prospektive Simulationen, die zur Vorausberechnung von Verwandtschaft dienen, wurden vor allem für die Vereinigten Staaten (vgl. u.a. Hammel et al. 1981; Reeves 1987; Wachter 1997; Wachter et al. 1997) und den asiatischen Raum (vgl. Hammel et al. 1991; Jiang 1995; Tu et al. 1993; Wachter et al. 2002), aber auch für Deutschland (vgl. Galler 1990), die Niederlande (vgl. Wolf 1988a; Imhoff & Post 1998) und Italien (vgl. Bertino et al. 1988; Tomassini & Wolf 2000) durchgeführt.

Anhand der im letzten Absatz vorgestellten Literatur lässt sich eine von Tomassini & Wolf (2000: 357) vorgenommene Unterteilung von Simulationsmodellen für Verwandtschaft in dynamisch-stabile, dynamisch-historische und dynamisch-prospektive Modelle leicht nachvollziehen.<sup>102</sup> Dynamisch-stabile Modelle gehen von konstanten Raten aus, wie beispielsweise die meisten oben genannten anthropologischen Arbeiten (vgl. ebd.: 357). Dynamisch-historische Modelle korrespondieren mit den Arbeiten der historischen Demographie und dynamisch-prospektive Simulationsmodelle wurden im vorausgegangenen Absatz bereits als solche bezeichnet (vgl. ebd.: 357). Im weiteren Verlauf dieses Unterabschnitts wird der Schwerpunkt der Darstellung entsprechend der Fragestellung der vorliegenden Arbeit vor allem auf dynamisch-prospektive Ansätze gelegt und deren wesentlichen Charakteristika und Unterscheidungsmerkmale diskutiert. Allerdings werden teils auch Überlegungen zu den anderen beiden Typen mit einbezogen.

##### 4.6.2.1 Zeitstetige und zeitdiskrete Simulation

Eine grundsätzliche Entscheidung, die bei dynamischen Mikrosimulationen getroffen werden muss, ist die Festlegung auf zeitdiskrete oder zeitstetige Simulation (vgl. Gilbert & Troitzsch 1999: 61). Bei zeitdiskreter Simulation wird ausgehend von einer vorab festgelegten Zeiteinheit, wie zum Beispiel einem Jahr, die Fortschreibung in Schrittgröße dieser Zeiteinheit vorgenommen und über Wahrscheinlichkeiten, die sich auf diese Zeiteinheit beziehen, fest-

---

101 Für Übersichten siehe Dyke (1981) und Zhongwai Zhao (2006).

102 Tomassini & Wolf (2000) führen zusätzlich noch statische Modelle auf und nennen als Beispiel einen Ansatz, der von Goldstein (1996) vorgeschlagen wurde. Dieser Ansatz ist allerdings ein Matching-Verfahren, um detaillierte Personendaten aus einem Datensatz mit Informationen über Verwandtschaftsbeziehungen aus einem anderen Datensatz zusammenspielen, und stellt keine statische Mikrosimulation im hier vorgestellten Sinne dar, sondern ist allgemein als Imputationsansatz zu sehen (vgl. auch Goldstein 1996: 1). Eine allgemeine Diskussion der Kombination von Matching-Verfahren und Mikrosimulationen bietet Cohen (1991b).

gelegt, ob ein bestimmtes Ereignis eintritt. Bei zeitstetiger Simulation wird hingegen durch eine Ziehung aus einer passenden Verteilung festgelegt, wann ein Ereignis stattfindet (vgl. ebd.: 61). Zum Beispiel könnte bei einer zeitdiskreten Simulation für ein Individuum jede Vorausberechnungseinheit überprüft werden, ob diese Person stirbt. Bei zeitstetiger Simulation könnte beispielsweise zur Geburt einer Person aus einer Verteilung von Lebensdauern ein Wert gezogen werden. Wird in der Simulation dieser Zeitpunkt erreicht, wird der Status der Person geändert.

Der Vorteil zeitdiskreter Simulation ist vor allem darin zu sehen, dass für die Simulation verwendbare Daten oftmals für eine Periodenlänge von einem Jahr vorliegen (vgl. Galler & Ott 1994: 401). Darüber hinaus gilt die zeitdiskrete Simulation allgemein als leichter umsetzbar als zeitstetige (vgl. ebd.: 402), was unter anderem daran liegt, dass wesentlich leichter Konsistenz erreicht werden kann, wenn Ereignisse mehrere Personen betreffen, wie zum Beispiel Scheidungen (vgl. Imhoff & Post 1998: 117). Problematisch ist allerdings die Simulation von mehreren möglichen Ereignissen in einer Periode (vgl. ebd.: 115). Hierbei muss eine Reihenfolge für die Ereignisse festgelegt oder diese aber in Ereigniskombinationen, die an Stelle der einzelnen Ereignisse simuliert werden, zusammengefasst werden (vgl. ebd.: 115).

Zeitstetige Simulation macht dies hingegen überflüssig (vgl. Gilbert & Troitzsch 1999: 61). Für alle möglichen Ereignisse wird aus entsprechenden Verteilungen die Dauer bis zum Eintritt gezogen und in einem „Ereigniswettbewerb“ (Wachter 1997: 92) gewinnt das Ereignis, welches als erstes eintritt. Zum festgelegten Eintrittszeitpunkt wird das Ereignis ausgeführt und ein neuer Ereigniswettbewerb beginnt (vgl. ebd.: 92). Hierdurch lassen sich konkurrierende Ereignisse vergleichsweise einfach umsetzen und es muss nicht während jedes Simulations-schrittes für alle Individuen überprüft werden, ob ein Ereignis eintritt. Allerdings schwindet der letzte Vorteil etwas, wenn man bedenkt, dass einzelne Ereignisse mehrere Individuen treffen können. Hierdurch ist es nicht möglich, komplette Lebensläufe „vorauszusimulieren“, und anschließend nur noch an den passenden Zeitstellen Ereignisse umzusetzen (vgl. Hammel et al. 1991: 246).

#### 4.6.2.2 Simulierte Prozesse

Die für Verwandtschaftsbeziehungen grundlegenden demographischen Prozesse – Mortalität, Fertilität und Nuptialität – werden in allen hier bisher erwähnten Mikrosimulationen berücksichtigt. Allerdings gibt es hierbei teils auch beträchtliche Unterschiede. Beispielsweise werden zuweilen nur Ehen, zuweilen zusätzlich auch nicht-eheliche Lebensgemeinschaften berücksichtigt (vgl. bspw. Galler 1990; Jiang 1995). Bei der Mikrosimulation von Smith (1987: 258) wiederum werden zwar Eheschließungen abgebildet, allerdings keine Scheidungen.

Grundsätzlich ergibt sich sowohl bei der Simulation des Eingehens als auch der Lösung von Paarbeziehungen das Problem, auf welche Raten beziehungsweise Wahrscheinlichkeiten man sich bezieht (vgl. Imhoff & Post 1998: 117). Angenommen, es soll für ein simuliertes Ehepaar überprüft werden, ob eine Scheidung eintritt. Dann kann eine Scheidungswahrscheinlichkeit verwendet werden, die sich auf die Ehe bezieht – die also die Frage beantwortet, wie wahrscheinlich es ist, dass eine Ehe (gegebenenfalls mit einer bestimmten Ehedauer oder anderen Merkmalen) geschieden wird (vgl. ebd.: 118). Man kann aber auch eine Wahrscheinlichkeit

verwenden, die sich auf die Frau bezieht und angibt, wie wahrscheinlich es ist, dass die Ehe einer Frau (gegebenenfalls mit bestimmten Merkmalen) geschieden wird (vgl. ebd.: 117). Zudem könnte eine ähnliche Wahrscheinlichkeit für den Mann oder aber eine Kombination aus beiden verwendet werden (vgl. ebd.: 118). Letztere Variante wurde von Imhoff & Post (1998) umgesetzt. Ausgegangen wurde dabei zunächst von individuellen Scheidungswahrscheinlichkeiten für Männer und Frauen. Diese wurden halbiert und dann sowohl für den Mann als auch für die Frau eine Zufallszahl gezogen (vgl. ebd.: 127).<sup>103</sup> Hingegen werden von Galler (1990) Wahrscheinlichkeiten verwendet, die sich auf Ehen beziehen und von Hammel et al. (1976) werden Trennungsraten für Frauen verwendet.

Bei Eheschließungen treten prinzipiell ähnliche Probleme auf (vgl. Imhoff & Post 1998: 1118). Hinzu kommt, dass bei Schließung von Ehen und Partnerschaften simulierte Personen zu einem Paar zusammengeführt werden müssen (vgl. bspw. Wachter et al. 1997: 93). Eine Möglichkeit, diese Problemstellung zu umgehen, besteht darin, dass einfach ein nach bestimmten Kriterien „passender“ Partner erzeugt wird (diese Möglichkeit wird in Abschnitt 4.6.2.4 genauer besprochen). Bei vielen Ansätzen wird hingegen ein Heirats- oder Beziehungsmarkt simuliert (vgl. bspw. Hammel et al. 1976; Galler 1990; Wachter et al. 1997; Imhoff & Post 1998). Zunächst wird für einzelne Individuen festgelegt, ob sie am Heiratsmarkt teilnehmen, wobei hierfür Heiratswahrscheinlichkeiten verwendet werden (vgl. bspw. Galler 1990: 376). Anschließend wird von einem Geschlecht ausgehend nach „passenden“ Partnern gesucht (vgl. ebd.: 376). Beispielsweise kann über vorgegebene „ideale“ Merkmale und ein Distanzmaß ein am besten „passender“ Partner ausgewählt werden (vgl. bspw. Galler & Ott 1994: 411). Eine andere Möglichkeit wird von Imhoff & Post (1998) verwendet. Bei dieser wird ausgehend von einer Person und dem Alter  $x$  derselben ein Partner zufällig aus den potenziellen Partnern gezogen, wobei die Ziehungswahrscheinlichkeit über die bedingte Altersverteilung für Partner von Personen im Alter  $x$  festgelegt wird (vgl. ebd.: 128). Insbesondere bei der letztgenannten Variante ist es möglich, dass einer Person am Heiratsmarkt kein Partner zugespielt werden kann. Entweder wird diese Person nicht heiraten, oder aber man spielt ihr zufällig einen Partner zu, der ebenfalls noch am Heiratsmarkt vorhanden ist (vgl. Hammel et al. 1976: 16). Allerdings können auch bei der letzteren Variante noch immer Personen am Heiratsmarkt „übrig bleiben“, was insbesondere gilt, wenn es ein deutliches Ungleichgewicht zwischen der Zahl der heiratswilligen Männer und Frauen gibt.

Eine Frage, die sich nicht nur für die Modellierung von Partnerschaften stellt, sondern auch für Mortalität und Fertilität, ist die Frage nach der Berücksichtigung von Heterogenität.<sup>104</sup> Bei der Modellierung von Mortalität wird zumeist auf alters- und geschlechtsspezifische Sterbewahrscheinlichkeiten zurückgegriffen (vgl. z.B. Galler & Ott 1994; Imhoff & Post 1998; Tomassini & Wolf 2000). Bei Fertilität werden oftmals paritäts- und altersspezifische Geburts-

---

103 Dieses Verfahren führt dazu, dass die Wahrscheinlichkeit, dass eine Ehe geschieden wird, grundsätzlich nicht den individuellen Wahrscheinlichkeiten entspricht, selbst wenn beide Wahrscheinlichkeiten gleich hoch sind. Sei  $Pr(S_w = 1) = 0.1$  die Scheidungswahrscheinlichkeit für die Frau und  $Pr(S_m = 1) = 0.1$  die Scheidungswahrscheinlichkeit für den Mann. Dann ist die Scheidungswahrscheinlichkeit der Ehe  $Pr(S = 1) = 1 - Pr(S_w = 0, S_m = 0) = 0.0975$ , wenn man das Vorgehen von Imhoff & Post (1998) verwendet. Die „gemeinsame“ Trennungswahrscheinlichkeit unterschätzt in diesem Beispiel die individuellen Trennungswahrscheinlichkeiten also leicht.

104 Hier wird lediglich auf häufig in der Literatur zu findende Differenzierungen und Differenzierungsmöglichkeiten eingegangen und entsprechende Beispiele aufgeführt. Weitere Differenzierungen sind sicherlich leicht vorstellbar und zuweilen auch in der Literatur zu finden.

wahrscheinlichkeiten verwendet, wobei teilweise noch nach Familienstand (vgl. Imhoff & Post 1998: 126) oder Ehedauer (vgl. Galler & Ott 1994: 410) differenziert wird. Wachter et al. (1997) beschreiben ein Modell, welches zusätzlich noch einen individuellen Multiplikator aufweist, der Geburtswahrscheinlichkeiten erhöhen oder senken und zumindest teilweise an Nachfahren vererbt werden kann (vgl. ebd.: 93). Die in den obigen Abschnitten erwähnten Wahrscheinlichkeiten für die Schließung und Trennung von Partnerschaften werden in der Regel nach Alter, Geschlecht und Familienstand und gegebenenfalls noch Beziehungsdauer differenziert, wenn individuelle Wahrscheinlichkeiten verwendet werden, und nach Beziehungsdauer, wenn sich Trennungswahrscheinlichkeiten auf Partnerschaften beziehen (vgl. bspw. Galler & Ott 1994; Wachter et al. 1997; Imhoff & Post 1998).

Migration wird zumeist nicht explizit modelliert (für eine Ausnahme vgl. Wachter et al. 1997: 93). Neben diesen „demographischen“ Prozessen besteht noch die Möglichkeit, weitere Prozesse, wie zum Beispiel Erwerbstätigkeit oder Bildungspartizipation, zu berücksichtigen. Dies erlaubt eine weitergehende, modellendogene Abbildung von Heterogenität. Allerdings wird hiervon in allen Modellen, die explizit nur auf die Modellierung von Verwandtschaft ausgerichtet sind, abgesehen. Beispiele finden sich in breiter angelegten Simulationen, wie zum Beispiel beschrieben von Orcutt et al. (1976) und Galler & Ott (1994).

#### 4.6.2.3 Ausgangsstichprobe und weitere Datenquellen

Eine grundlegende Frage, die sich bei jeder Mikrosimulation stellt, ist, was für ein Ausgangsdatensatz verwendet wird. Sowohl bei historischen als auch prospektiven Simulationen ergibt sich das Problem, dass oftmals keine Daten vorhanden sind, die genaue Informationen zu Verwandtschaftsverhältnissen liefern, die sich Fortschreiben lassen (vgl. bspw. Hammel et al. 1981: 14) beziehungsweise zumindest nicht in dem Datensatz vorhanden sind, der verwendet werden soll (vgl. bspw. Galler 1990: 74). Eine Möglichkeit besteht darin, dieses Problem zu übergehen und auf Informationen zu Verwandtschaft im Ausgangsdatensatz zu verzichten (vgl. ebd.: 74). Durch das Fortschreiben der Daten werden dann nach und nach Verwandtschaftsverhältnisse generiert. Allerdings können diese auch nach einem längeren Vorausberechnungszeitraum noch immer unvollständig sein. Beispielsweise benutzte Galler (1990) Daten aus dem SOEP aus dem Jahr 1984 als Grundlage für eine Vorausberechnung bis zum Jahr 2050. In diesen Daten sind bezogen auf einzelne Personen allerdings nur Verwandte im selben Haushalt erfasst (vgl. ebd.: 74). Somit sind für Personen, die 1984 nicht mit ihren Eltern zusammenlebten und 2050 noch leben, keine Informationen über Geschwister vorhanden (vgl. ebd.: 76). Zwar dürften 2050 nichtsdestotrotz Verwandtschaftsnetzwerke größtenteils vollständig sein, allerdings lässt sich die simulierte Entwicklung nicht beschreiben. Zudem ist für den Vergleich mit der Ausgangssituation eine weitere Datenquelle nötig. Galler (1990) verwendet Daten aus der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) für das Jahr 1986. Allerdings ist hier die Frage, ob die beiden Datensätze vergleichbar sind und die Ergebnisse zu Verwandten aus dem ALLBUS direkt auf das SOEP übertragbar sind.<sup>105</sup>

---

105 Beispielsweise ist bekannt, dass der erfasste Anteil an kinderlosen Frauen zwischen verschiedenen Stichproben aufgrund unterschiedlicher Ursachen variiert (vgl. bspw. Kreyenfeld & Konietzka 2007; Scharein & Unger 2005).

Ein anderer Ansatz wird unter anderem von Reeves (1987) und Imhoff & Post (1998) verwendet. Die Grundidee besteht darin, zunächst eine Art „Vorsimulation“ durchzuführen, die für das Startjahr der Vorausberechnung einen synthetischen Datensatz erzeugt. Imhoff & Post (1998) verwenden als Grundlage eine Stichprobe des niederländischen Zensus aus dem Jahr 1947. Diese wird über historische Raten bis zum Startjahr 1994 der eigentlichen Vorausberechnung fortgeschrieben (vgl. ebd.: 125). Für das Startjahr steht dann ein Datensatz mit größtenteils vollständigen Verwandtschaftsnetzwerken zur Verfügung. Reeves (1987: 230) bildet auf Grundlage von Daten der amtlichen US-amerikanischen Statistik einen synthetischen Datensatz, dessen Zusammensetzung bezogen auf die Verteilung von Geschlecht und Alter der US-Bevölkerung im Jahr 1900 entspricht, welche dann ebenfalls über historische Raten bis zum eigentlichen Ausgangsjahr 1980 fortgeschrieben wird. Da es dabei allerdings zu Abweichungen zwischen der Altersstruktur der synthetischen Stichprobe und der realen Bevölkerung kommen kann, schlägt Reeves vor, mehrere Vorsimulationen durchzuführen und anschließend die synthetische Stichprobe zu verwenden, die der realen Bevölkerung des Ausgangsjahres am „ähnlichsten“ ist (vgl. ebd.: 231). Als Möglichkeit, die Ähnlichkeit von realer Bevölkerung und synthetischer Stichprobe bezüglich der Verteilung bestimmter Merkmale festzustellen, schlägt er einen  $\chi^2$ -Test vor (vgl. ebd.: 232).

Neben einer Ausgangsbevölkerung sind noch weitere Informationen notwendig. Hiermit sind insbesondere Übergangsraten beziehungsweise -wahrscheinlichkeiten gemeint. Oftmals werden diese aus Aggregatdaten der amtlichen Statistik gewonnen (vgl. bspw. Tomassini & Wolf 2000: 359) und teilweise durch die Schätzung von Modellparametern aus Stichproben ergänzt. Zuweilen werden auch vollständig hypothetische Parameter verwendet, die bestimmte demographische Regime möglichst gut widerspiegeln sollen (für ein Beispiel s. Smith & Oeppen 1993; Hammel 2005). Darüber hinaus wird in einem Großteil der Simulationen von konstanten Raten ausgegangen, die denen des Ausgangsjahres entsprechen (vgl. bspw. Galler 1990; Tu et al. 1993). Vorausschätzungen der künftigen Entwicklung von demographischen Raten aus der amtlichen Statistik werden allerdings ebenso verwendet (vgl. Imhoff & Post 1998; Tomassini & Wolf 2000), wie eigene Annahmen und Vorausschätzungen (vgl. bspw. Reeves 1987; Jiang 1995).

#### 4.6.2.4 Offene versus geschlossene Simulation

Die Unterscheidung zwischen offenen und geschlossenen Mikrosimulationen bezieht sich auf die Frage, ob unter bestimmten Umständen zusätzliche Individuen erzeugt werden (vgl. bspw. Wachter et al. 1997: 92). Bei geschlossenen Simulationen ist dies nicht der Fall – „neue“ Individuen können nur durch Geburt in die simulierte Population gelangen (vgl. Smith 1987: 253). Hingegen wird bei offenen Mikrosimulationen bei bestimmten Ereignissen ein zusätzliches Individuum erzeugt (vgl. ebd.: 253). Beispielsweise könnte für ein simuliertes Individuum das Ereignis Heirat eintreten. In einer geschlossenen Simulation wird der Heiratspartner aus den in der Population vorhandenen Personen ausgewählt (s. bspw. Reeves 1987), während hingegen bei offenen Simulationen ein Heiratspartner mit bestimmten Eigenschaften erzeugt wird (s. bspw. Smith 1987).

Dabei kann diese Unterscheidung nicht nur auf Nuptialität, sondern auch auf andere Modellaspekte bezogen werden. Ein Extrem besteht darin, ausgehend von einer Person Verwandte

für diese zu generieren. Beispiele für diese Vorgehensweise wurden unter anderem von Smith (1987) und Zhongwai Zhao (1996) vorgelegt. Um beispielsweise die Mutter eines Egos zu erzeugen, wird ein weibliches Individuum generiert und dessen Lebenslauf partiell simuliert, unter der Bedingung, dass mindestens ein Kind geboren werden muss (vgl. Smith 1987: 253). Werden mehrere Kinder geboren, wird zufällig festgelegt, welches davon Ego ist (vgl. ebd.: 253). Egos Mutter und gegebenenfalls bereits vorhandene Geschwister werden dann ab der Geburt Egos parallel simuliert (vgl. ebd.: 253). Für andere Verwandte, die nicht direkt von Ego abstammen wird analog vorgegangen (vgl. ebd.: 254). Hierdurch entstehen egozentrierte Netzwerke, wobei sich die Netze verschiedener Egos nicht überlappen, was zu einer Überschätzung der Varianz der Zahl der Verwandten führt (vgl. ebd.: 257).<sup>106</sup> Zudem erfordert dieses Vorgehen konstante demographische Raten (vgl. ebd.: 254). Ein Vorteil ist allerdings darin zu sehen, dass für die Ausgangsstichprobe keine Informationen über Verwandtschaftsbeziehungen notwendig sind und nichtsdestotrotz ausgehend von einzelnen Egos direkt Verwandtschaftsnetzwerke erzeugt werden können.

#### 4.6.3 Anwendungsbeispiel

Eine Anwendung des Mikrosimulationsansatzes zur Vorausberechnung von Verwandtschaft liefert Galler (1990). Als Ausgangsstichprobe wurde wie erwähnt die erste Welle des SOEP aus dem Jahre 1984 verwendet und um Personen, die in Anstaltshaushalten leben, ergänzt (vgl. ebd.: 74). Hier ist anzumerken, dass der Datensatz aus historischen Gründen lediglich Personen aus westdeutschen Haushalten enthält, allerdings keine Personen aus Ostdeutschland. Der Stichprobenumfang wurde durch mehrfache „Verwendung“ von Fällen von circa 15000 auf etwa 60000 Personen erhöht, um Variabilität in den Ergebnissen durch die Zufälligkeit der Fortschreibung zu reduzieren (vgl. ebd.: 82, Endnote 22). Diese Personen sind der Ausgangspunkt der Simulation, allerdings werden auch Haushalte und Haushaltsbildung beziehungsweise -auflösung simuliert (vgl. ebd.: 72).

Modelliert wurden Mortalität, Fertilität, die Schließung und Lösung von Ehen und nicht-ehelichen Partnerschaften sowie der Auszug von Kindern aus dem elterlichen Haushalt und Übergänge in und aus Anstalten (vgl. Galler 1990: 72ff.). Etliche der benötigten Parameter wurden direkt oder indirekt aus Tafelrechnungen der amtlichen Statistik gewonnen, andere wurden aus Bevölkerungsstichproben, wie zum Beispiel der Lebensverlaufsstudie des Sfb 3, geschätzt (für Details hierzu s. Ott 1986b). Alle Parameter beziehen sich auf den Zeitraum der Jahre 1982 bis 1984 und wurden über den Verlauf der Vorausberechnung hinweg konstant gehalten (vgl. Galler 1990: 75).<sup>107</sup>

Neben Ergebnissen zur durchschnittlichen Zahl von Verwandten und dem Anteil an Personen, die bestimmte Verwandte haben, gibt der Autor „Belastungsindizes“ in diversen Varianten an. Hier sei lediglich auf eine der Varianten dieser Belastungsindizes eingegangen. Die Grundidee dieser Indizes ähnelt der von Altenquotienten und verwandten Maßen, bei denen das

---

106 Die Bezeichnung „egozentriert“ stammt aus dem Bereich der Netzwerkanalyse und bedeutet, dass Netzwerke ausgehend von einer Person betrachtet werden (vgl. Wasserman & Faust 1998: 41f.). Ein Beispiel wäre die Betrachtung von Verwandtschaft nach dem Verwandtschaftsschema nach Atkins in Tabelle 2.1 auf Seite 23.

107 Für eine Übersicht über alle verwendeten Parameter und deren Datenquellen, Informationen zur Schätzung mancher dieser Parameter sowie weiteren Ergebnissen des Sfb 3-Modells s. Ott (1986a,b).

numerische Verhältnis zweier Bevölkerungsgruppen berechnet wird. Für 30 bis 59-jährige Personen wird die durchschnittliche Belastung zum einen durch Eltern, zum anderen durch Kinder sowie durch diese beiden Gruppen zusammen ermittelt (vgl. Galler 1990: 78). Zur Ermittlung der individuellen Belastung durch Eltern wurde die Zahl der lebenden Elternteile geteilt durch die Zahl der Geschwister plus 1 berechnet, unter der Annahme, dass Geschwister sich zu gleichem Teil um pflegebedürftige Eltern kümmern (vgl. ebd.: 78).<sup>108</sup> Zur Berechnung der Belastung durch Kinder wurde die Zahl der im selben Haushalt lebenden Kinder herangezogen (vgl. ebd.: 79). Diese Werte wurden für alle Personen in der betrachteten Altersgruppe errechnet und das arithmetische Mittel bestimmt, wobei sich die Gesamtbelastung aus der Summe der beiden Mittelwerte ergibt (vgl. ebd.: 79). Da im SOEP für das Jahr 1984 wie erwähnt nur eingeschränkte Informationen zur Verwandtschaft vorliegen, wurden die Kennwerte für das Ausgangsjahr der Vorausberechnung mit Daten des ALLBUS ermittelt (vgl. ebd.: 78).

Der Belastungsindex für Eltern liegt sowohl für das Jahr 1986 als auch für das Jahr 2050 bei einem Wert von 0.46 (vgl. Galler 1990: 79). Ursache hierfür sind zwei gegenläufige Entwicklungen. Zum einen nimmt zwar die Zahl der Geschwister ab, so dass die individuellen Werte im Nenner sinken, andererseits nimmt durch spätere Geburten der Altersabstand zwischen der Eltern und der Kindergeneration zu, so dass es weniger Personen mittleren Alters gibt, bei denen beide Eltern oder ein Elternteil noch leben (vgl. ebd.: 78). Werden nur Personen mit mindestens einem Elternteil berücksichtigt, zeigt sich ein leichter Anstieg des Index von 0.67 auf 0.71 (vgl. ebd.: 79). Der Index für Kinder sinkt von einem Wert von 1.30 auf 0.94, wobei diese Abnahme deutlich moderater ausfällt, wenn nur Personen mit Kindern bei der Berechnung berücksichtigt werden (vgl. ebd.: 79). Dies wird von Galler vor allem auf den Anstieg der Zahl kinderloser Personen zurückgeführt (vgl. ebd.: 79). Im Ausgangsjahr beträgt der Kennwert für die Gesamtbelastung 1.76, im Jahr 2050 1.40 (vgl. ebd.: 79). Werden die Indizes für Belastung durch Eltern und Kinder, die von mindestens einem Elternteil beziehungsweise mindestens einem Kind ausgehen, aufaddiert, betragen die Werte 1.98 und 1.71 (vgl. ebd.: 79).

Zwar merkt Galler an, dass durch die verwendeten Indizes nur eine grobe Abbildung der Belastung durch Kinder und pflegebedürftige Eltern möglich ist, da Pflegebedürftigkeit nicht modelliert wird (vgl. Galler 1990: 77). Und ebenso wird auf weitere Probleme hingewiesen, wie beispielsweise die Annahme der gleichmäßigen Verteilung der Belastung durch Eltern auf Geschwister, die nicht vorhandene Berücksichtigung von räumlichen Distanzen sowie möglicherweise abnehmende Unterstützungspotentiale durch eine Zunahme der Erwerbstätigkeit von Frauen (vgl. ebd.: 80). Nichtsdestotrotz kommt er insgesamt zu dem Schluss, dass der Umfang von Verwandtschaftsnetzwerken sinken wird, allerdings durch die Abnahme von zu betreuenden Kindern insgesamt keine Abnahme des Unterstützungspotentials von Verwandten für pflegebedürftige Ältere zu erwarten ist (vgl. ebd.: 80).

---

108 Der Altenquotient, an den diese Größen angelehnt sind, wird ermittelt als das Verhältnis der Bevölkerung im Rentenalter zur Bevölkerung im erwerbsfähigen Alter, wobei diese beiden Altersbereiche unterschiedlich abgegrenzt werden können (vgl. Statistisches Bundesamt 2010c: 47). Für eine kritische Diskussion solcher Maße s. Bacon (1997), Clark & Sprenger (1980), Kleimann (1967) und Sanderson & Scherbov (2007).

#### 4.6.4 Diskussion

Die im Verlauf der Erörterung von Mikrosimulationen aufgeführten Themen und Beispiele machen deutlich, dass Mikrosimulationsansätze insgesamt sehr heterogen sind. Ein wichtiger Aspekt, der allen Varianten gemein ist und bisher nur am Rande Erwähnung fand, ist die Zufälligkeit der Ergebnisse. Willekens (2005: 117f.) sieht in dieser Zufälligkeit eine Möglichkeit, Unsicherheit bei Vorausberechnungen zu berücksichtigen. Dabei entsteht die Variabilität der Ergebnisse aber primär durch die Verwendung von Zufallszahlen zur Fortschreibung der Bevölkerung (vgl. Orcutt et al. 1976: 247). Diese Variabilität kann zum einen durch die Erhöhung des Stichprobenumfangs (vgl. Imhoff & Post 1998: 110), zum anderen durch die Verwendung von Zufallsgeneratoren mit festem Saat-Wert für jeden Modelldurchlauf verringert werden (vgl. Orcutt et al. 1976: 248).<sup>109</sup> Insofern kann nicht davon gesprochen werden, dass Mikrosimulationen Unsicherheit bezüglich der demographischen Raten oder der Datengrundlage per se berücksichtigt.

Vielmehr lassen sich drei weitere Faktoren aufzählen, die ebenfalls zur Zufälligkeit der Ergebnisse beitragen und ihre Unsicherheit noch erhöhen können: die Verwendung von Stichprobendaten als Ausgangsbevölkerung, die Schätzung von Parametern aus Stichproben sowie unter Umständen der verwendete Modellierungsansatz (vgl. Imhoff & Post 1998: 110; Orcutt et al. 1976: 248). Die Verwendung einer Bevölkerungsstichprobe als Ausgangsbevölkerung kann die Unsicherheit erhöhen, da es sich in aller Regel um Zufallsstichproben handelt, bei denen die Verteilung der interessierenden Merkmale zufällig von der Verteilung in der Grundgesamtheit abweichen kann (vgl. Imhoff & Post 1998: 110).<sup>110</sup> Ähnliche Probleme ergeben sich bei der als zweites genannten Schätzung von Modellparametern aus Stichprobendaten, die zumindest bei stark disaggregierten Modellen nötig werden kann und mit einem Stichprobenfehler behaftet ist (vgl. ebd.: 110). Schließlich kann bei diskreten Mikrosimulationen ein mehrfaches Ziehen von Zufallszahlen erforderlich sein, um eine Simulationseinheit um eine Zeiteinheit fortzuschreiben, wodurch die Variabilität der Ergebnisse zusätzlich erhöht werden kann (vgl. ebd.: 110).

Weiterhin wurde bei etlichen der im Verlauf dieses Abschnitts aufgeführten Vorausberechnungen von konstanten Raten ausgegangen (s. bspw. Galler 1990), was zumindest für die langfristige Entwicklung als unangemessen erscheint. Eine mögliche Ursache hierfür dürfte sein, dass aufgrund der häufig vorliegenden Vielzahl an Modellparametern die Bildung von Szenarien sehr aufwendig ist.

Als ein Vorteil lässt sich hervorheben, dass Mikrosimulationen sehr detaillierte Auswertungen erlauben, da Mikrodatensätze erzeugt werden, die für die Analyse beliebig aggregiert und sowohl in Längsschnitt- als auch Querschnittsperspektive ausgewertet werden können (vgl. Imhoff & Post 1998: 107; Ruggles 1990: 24). Weiterhin lassen sich auch vergleichsweise große Merkmalsräume effizient modellieren und sowohl stetige Variablen als auch Interaktionseffekte zwischen Variablen lassen sich vergleichsweise einfach berücksichtigen (vgl. Spielauer &

109 Dass Variabilität auf der Mikroebene bei großen Populationen keinen oder kaum Einfluss auf die Bevölkerungsentwicklung hat, wird in Unterabschnitt 5.2.3 nochmals kurz diskutiert.

110 Es kann zwar versucht werden, die Unsicherheit durch Verwendung von möglichst großen Stichproben oder von Hochrechnungsfaktoren zu reduzieren (vgl. Galler & Ott vgl. 1994: 422; Imhoff & Post 1998: 110; Orcutt et al. 1976: 247). Nichtsdestotrotz kann ein Einfluss auf die Ergebnisse in aller Regel nicht ausgeschlossen werden und eine Abschätzung dieses Einflusses gilt als schwierig (für eine Diskussion s. Cohen 1991c).

Vencatasawamy 2003: 145; Imhoff & Post 1998: 104ff.). Allerdings gehen diese Vorteile in der Regel mit einem vergleichsweise hohen Rechenaufwand und Datenbedarf einher (vgl. Gilbert & Troitzsch 1999: 55). Zusätzlich sind insbesondere bei sehr detaillierter Modellierung viele Annahmen nötig, wie beispielsweise die Wahl von Regressionsmodellen und unabhängigen Variablen für diese (vgl. Ruggles 1990: 25). Zudem ist es bei der detaillierten Modellierung von vielen Prozessen komplex herauszufinden, welche Einflussgrößen ein Ergebnis genau beeinflusst haben, da die Eigenschaften solcher komplexen Modelle schwer zu beschreiben sind (vgl. Galler 1994b: 428), was eine Interpretation unter Umständen schwierig macht (vgl. Ruggles 1990: 25; Hammel et al. 1991: 244). Aufgrund dieser Komplexität gilt auch die Programmierung von Mikrosimulationen am PC als aufwendig und fehleranfällig (vgl. Ruggles 1990: 25; Hammel et al. 1991: 244).<sup>111</sup> Dabei ist allerdings zu beachten, dass in etlichen der aufgeführten Arbeiten nur gering differenzierte Daten aus der amtlichen Statistik als Grundlage dienten (vgl. bspw. Galler 1990; Imhoff & Post 1998), so dass die zu Beginn dieses Absatzes genannten Vorteile nicht oder nur eingeschränkt genutzt wurden.

Bei den vorgestellten Arbeiten fällt ferner auf, dass Migration zumeist nicht berücksichtigt wurde (für eine Ausnahme s. Wachter et al. 1997). Zudem wird von einer Ausnahme abgesehen in allen Arbeiten von Unabhängigkeit von Personen ausgegangen (vgl. ebd.). Ruggles (1993: 162ff.) weist darauf hin, dass dies auch bei Mikrosimulationen zu einer Verzerrung der Ergebnisse führen kann, selbst bei geringer „Korrelation“ zwischen den Generationen.

Bezogen auf die Modellierung von Verwandtschaft ist positiv hervorzuheben, dass theoretisch beliebig umfassende Verwandtschaftsnetzwerke generiert und Ereignisse, die mindestens zwei Personen betreffen, vergleichsweise einfach simuliert werden können (vgl. Imhoff & Post 1998: 117). Ein bereits erwähntes Beispiel ist die Simulation von Eheschließungen, die aber auch einige Probleme birgt. Werden Ehepartner bei einer offenen Mikrosimulation je nach Bedarf erzeugt, entsteht zum einen das Problem, dass die Bevölkerung „künstlich“ vergrößert wird, zum anderen sind für den generierten Partner keine Informationen über Verwandte vorhanden, so dass affinale Verwandtschaft nur schlecht simuliert werden kann (vgl. Wachter et al. 1997: 92). Bei geschlossenen Simulationen, bei denen grundsätzlich keine zusätzlichen Personen erzeugt werden, besteht das Problem, dass für eine Person gegebenenfalls kein passender Heiratspartner gefunden werden kann (vgl. Hammel et al. 1976: 16). Dies kann dazu führen, dass die Simulation nicht mehr vorgegebenen Heiratsraten beziehungsweise -wahrscheinlichkeiten folgt, so dass Auswertung und Interpretation der Ergebnisse erschwert werden.

Ein weiteres Problem ist das Vorhandensein von Informationen zu Verwandtschaftsverhältnissen in der Ausgangsstichprobe. Wenn solche Informationen nicht oder nur begrenzt vorhanden sind und darauf gesetzt wird, dass während der Simulation nach und nach Verwandtschaftsverhältnisse generiert werden, dürfte in aller Regel wie bei Galler (1990) nur noch ein Querschnittsvergleich zwischen dem Ausgangs- und dem Zieljahr der Vorausberechnung möglich sein. Für diesen Vergleich ist eine weitere Stichprobe mit detaillierteren Informationen zu Verwandtschaft notwendig, so dass sich zum einen die Frage nach der Vergleichbarkeit dieser

---

111 Die Entwicklungszeit der DYNASIM Mikrosimulation, die eine Vielzahl von Prozessen berücksichtigt, allerdings nicht für die Simulation von Verwandtschaft eingesetzt wurde, betrug 7 Jahre (s. Orcutt et al. 1976: iv). Das von Galler (1990) verwendete Mikrosimulationsmodell wurde im Rahmen des Sonderforschungsbereich 3 der Deutschen Forschungsgemeinschaft entwickelt, geht aber auf Arbeiten aus den 1970ern zurück (wie bspw. Galler 1977).

weiteren Stichprobe mit der Ausgangsstichprobe stellt. Zum anderen müsste in diesem Fall begründet werden, warum die Vergleichsstichprobe nicht als Ausgangsstichprobe verwendet wird, zumal sie mehr Informationen über Verwandtschaft enthält. Die Verwendung einer Vorsimulation zur Erzeugung einer hypothetischen Ausgangsstichprobe ist ebenfalls nicht unproblematisch. Zum einen können benötigte historische Raten nicht vorhanden sein (vgl. Reeves 1987: 229). Zum anderen ist unklar, wann eine vorsimierte, hypothetische Bevölkerung als „realistisch“ angesehen werden kann. Denn zumindest die in dieser vorliegenden Verwandtschaftsverhältnisse können nicht mit denen der Bevölkerung, die man eigentlich simulieren möchte, verglichen werden, da ansonsten die Vorsimulation überflüssig wäre.

Schließlich liegen bis auf eine Ausnahme keinerlei Validierungsversuche der Simulationen vor.<sup>112</sup> Wachter (1997) und Wachter et al. (1997) kommen bei ihrer Validierung der Vorausberechnung von Reeves (1987) zu dem Schluss, dass zwar grundsätzlich eine gute Abschätzung von Verwandtschaftsnetzwerken mittels Mikrosimulation möglich ist, allerdings wie bei anderen Modellen auch eine starke Abhängigkeit von den getroffenen Annahmen besteht und zudem teilweise keine Rückschlüsse aus der Validierung für die Verbesserung von Modellen getroffen werden können, da manche Abweichungen der simulierten von der tatsächlichen Entwicklung nicht auf einzelne, isolierte Fehler zurückzuführen sind (vgl. Wachter et al. 1997: 102).

## 4.7 Zwischenfazit

Für eine zusammenfassende Bewertung der vier verschiedenen Ansätze wird zunächst wieder auf die zu Beginn des Kapitels genannten drei zentralen Probleme von Vorausberechnungen zurückgegriffen.

Die Unsicherheit der zukünftigen Entwicklung der demographischen Komponenten wird bei keiner der erwähnten Arbeiten berücksichtigt. In aller Regel werden noch nicht einmal unterschiedliche Vorausberechnungsszenarien erstellt, wie sie bei „normalen“ Bevölkerungsvorausberechnungen üblich sind. Stattdessen wird oftmals von konstanten demographischen Raten ausgegangen.

Bezüglich der Unsicherheit der verwendeten Daten hat sich gezeigt, dass diese eng mit anderen Problemen wie der Homogenitätsannahme verknüpft ist, da je nach Differenzierungsgrad gegebenenfalls mit Unsicherheit behaftete Stichprobendaten zur Parameterschätzung verwendet werden müssen. Bei Mikrosimulationen ergibt sich zusätzlich das Problem der Unsicherheit der Ausgangsstichprobe, welches in der Literatur allerdings kaum Beachtung findet.

Die theoretischen Überlegungen bei Verzweigungsprozessen und analytischen Ansätzen haben gezeigt, dass die Homogenitäts- und die Unabhängigkeitsannahme wahrscheinlich nicht unproblematisch sind. Für Lösungsversuche dieser beiden Probleme sowie des Zweigeschlechter-Problems kann festgehalten werden, dass sie in aller Regel den Datenbedarf deutlich erhöhen und weitere Annahmen nötig machen. Diese zusätzlichen Annahmen werden in der Literatur in aller Regel nicht überprüft und zumeist auch nicht begründet. Zudem

---

<sup>112</sup> Für einen allgemeinen Überblick über Validierung von Mikrosimulationen s. Cohen (1991a).

können die Lösungsansätze häufig nur eingeschränkt genutzt werden. Beispielsweise ist die Zahl weiterer Merkmale, die zur Lockerung der Homogenitätsannahme betrachtet werden, aufgrund der zur Verfügung stehenden Daten oftmals deutlich begrenzt.

Insofern verwundert es, dass die Zahl der in der Literatur zu findenden Validierungen der verschiedenen Ansätze sehr niedrig ist. Die durchgeführten Validierungen zeigen grundsätzlich mehr oder weniger deutliche Abweichungen zwischen Modellrechnungen und tatsächlichen Verwandtschaftsverhältnissen (vgl. Wachter 1997; Wachter et al. 1997; Zeng Yi et al. 1998; Dudel 2008). Dies weist aber darauf hin, dass die Annahmen, auf denen die Modellrechnungen basieren, nicht unproblematisch sind. Zudem wird deutlich, dass Validierungsversuche notwendig sind, um die Qualität von Ergebnissen zu überprüfen.

Insgesamt lässt sich festhalten, dass Verzweigungsprozesse und Multi-State-Ansätze, wie sie momentan in der Literatur verwendet werden, aufgrund ihrer Betrachtung von Verwandtschaft eher ungeeignet für die Fragestellung dieser Arbeit sind: Multi-State-Ansätze gehen von einer Einschränkung auf den Haushalt aus, während Verzweigungsprozesse lediglich jemals vorhandene Verwandtschaft betrachten, unabhängig davon, wann oder wie lange Verwandte zusammenleben.

Bei bevölkerungsmathematischen Ansätzen ist die Betrachtung von lateraler Verwandtschaft relativ problematisch, zudem fehlt affinale Verwandtschaft, während der Datenbedarf hingegen relativ gering ist. Mikrosimulationen erlauben zwar eine detaillierte Modellierung, allerdings bestehen die grundsätzlichen Probleme der Ausgangsstichprobe und des hohen Datenbedarfs, wenn die Vorzüge dieser Methode genutzt werden sollen. Insofern sind sowohl bevölkerungsmathematische Ansätze als auch Mikrosimulationen jeweils aus einer bestimmten Perspektive attraktiv.

Neben den in diesem Abschnitt bisher genannten Schwierigkeiten wurden im Verlauf des Kapitels weitere genannt, auf welche hier aber nicht mehr explizit eingegangen sei. In allen Fällen ergibt sich die Kritik aus dem angestrebten Ziel der vorliegenden Arbeit, wobei die beschriebenen Ansätze zumindest teilweise nicht auf dieses ausgerichtet sind und auf andere Problemstellungen zielen. Ausgehend von anderen Zielsetzungen kann die Bewertung somit auch anders ausfallen.

## 4.8 Eingrenzung der Fragestellung und Ausblick

Ausgehend von den bisher gemachten Überlegungen soll die in dieser Arbeit angestrebte Vorausberechnung von Verwandtschaft in mehrerer Hinsicht ein- und abgegrenzt werden:

1. Es soll Verwandtschaft aus einer „Ego-zentrierten“ Perspektive betrachtet werden (s. Unterabschnitt 2.1.1)
2. Es soll lediglich lineare Verwandtschaft betrachtet werden
3. Es soll die aktuell lebende Verwandtschaft betrachtet werden

Die erste Einschränkung bedeutet insbesondere, dass der Fokus nicht auf Haushalten oder auf Familie im Sinne der Definition in Unterabschnitt 2.1.2 liegt. Bei einer Beschränkung auf zusammenlebende Verwandte könnten außerhalb des Haushalts lebende Verwandte nicht

berücksichtigt werden, was eine unnötige Beschränkung ist und zudem mit methodischen Problemen verbunden wäre (s. Unterabschnitt 4.5.4; sowie Bien & Marbach 1991).

Für die Einschränkung auf lineare Verwandtschaft sprechen mehrere Gründe. Zunächst ist für laterale Verwandtschaft und einen Großteil der affinalen Verwandtschaft nicht bekannt, inwiefern sie von Bedeutung ist (s. Unterabschnitt 2.2.1), so dass Ergebnisse einer Vorausberechnung kaum interpretiert werden könnten. Ferner ist bei affinaler Verwandtschaft tendenziell eine historische Bedeutungsabnahme zu vermuten (abnehmende Heiratsneigung, steigende Scheidungsraten, Zunahme nichtehelicher Geburten; s. Abschnitt 3.2) Hinzu kommen methodische Schwierigkeiten. Beispielsweise zeigen die Überlegungen in Paragraph 4.4.2.4, dass Ergebnisse zu lateraler Verwandtschaft bei nicht ausreichender Berücksichtigung von Heterogenität verzerrt sein können (s. auch den Hinweis zu Ruggles 1993 in Unterabschnitt 4.6.4).

Die Beschränkung auf lineare Verwandtschaft umfasst auch den Ausschluss von Ehepartnern. Sollten diese mit berücksichtigt werden, müssten mehrere zusätzliche Prozesse eingeführt werden: zum einen das Eingehen von Ehen und zum anderen Ehescheidungen. Ferner müssten dann eigentlich auch nicht-eheliche Beziehungen mit betrachtet werden. Dies würde zum einen den Datenbedarf deutlich erhöhen und zum anderen die Einführung spezieller Annahmen notwendig machen, was sich beispielsweise bei den Ausführungen zu Verzweigungsprozessen und Mikrosimulationen zeigte. Somit wird zwar ein Teilbereich von Verwandtschaft ausgeschlossen, der mit Ehepartnern bedeutsame Verwandte enthält. Nichtsdestotrotz zeigen die Beispiele aus Abschnitt 2.2 aber auch, dass lineare Verwandtschaft ebenfalls bedeutsam ist, womit zumindest eine wichtige Teilmenge von Verwandtschaft betrachtet werden kann.

Die Betrachtung der aktuell lebenden Verwandtschaft ist sicherlich nicht die einzige interessante Möglichkeit. Beispielsweise wäre auch eine Betrachtung von jemals geborener Verwandtschaft vorstellbar. Denkt man aber an den in Unterabschnitt 2.2.2 beschriebenen Themenkomplex der Pflege, wird deutlich, dass sich über die aktuelle Verwandtschaft am besten angeben lässt, wie groß verwandtschaftliche Unterstützungspotenziale sind. Auch die Beispiele aus Unterabschnitt 2.2.1 sprechen eher für eine Betrachtung der aktuellen Verwandtschaft. Andere Betrachtungsweisen werden im weiteren aber nicht gänzlich ausgeschlossen.

Ausgehend von diesen Einschränkungen scheint ein Verfahren sinnvoll, dass dem bevölkerungsmathematischen Ansatz ähnelt. Dies würde die Betrachtung von lebender, linearer Verwandtschaft aus egozentrierter Perspektive erlauben und würde zudem mit einem relativ geringen Datenbedarf einhergehen. Allerdings wäre zum einen eine Erweiterung um die Berücksichtigung von Unsicherheit und zum anderen eine weitergehende Erweiterung auf die beiden Geschlechter notwendig, ohne dadurch den Datenbedarf zu stark zu erhöhen. Entsprechend der bisherigen Ausführungen scheint zudem unabhängig von der genauen Ausgestaltung des Vorgehens eine Validierung der Ergebnisse angebracht zu sein.

Im folgenden Kapitel wird ein Ansatz vorgestellt, der die im vorausgegangenen Abschnitt vorgestellten Punkte umfasst und von einer Erweiterung des Galton-Watson-Prozesses aus Abschnitt 4.3 als theoretischer Grundlage ausgeht. Bevor dieser Ansatz diskutiert wird, werden zu Beginn des nächsten Kapitels verschiedene in der Literatur zu findende Vorschläge diskutiert, wie Unsicherheit bei demographischen Vorausberechnungen berücksichtigt werden kann. Eine Modellvalidierung folgt in einem späteren Kapitel.

## 5 Ein Ansatz zur Vorausberechnung linearer Verwandtschaft

### 5.1 Probabilistische Vorausberechnungen

#### 5.1.1 Probleme „klassischer“ Vorausberechnungen als Motivation für probabilistische Vorausberechnungen

Wie zu Beginn von Kapitel 4 dargestellt, zeigen sich bei rückblickenden Vergleichen von vorausberechneten und tatsächlichen Bevölkerungsentwicklungen oftmals deutliche Unterschiede, die auf Unsicherheit in der Entwicklung der demographischen Komponenten zurückzuführen sind. Um dieser Unsicherheit zu begegnen, wird oft auf die Berechnung mehrerer Szenarien zurückgegriffen. Diese sollen unterschiedliche zukünftige Verläufe abbilden und hierüber Unsicherheit erfassen (vgl. Alho & Spencer 1985: 306). Häufig werden drei Szenarien gebildet, die jeweils einen zukünftigen Verlauf mit hohem, mittlerem und niedrigem Bevölkerungswachstum wiedergeben sollen (vgl. ebd.: 306).

Allerdings ist auch dieses Vorgehen mit etlichen Problemen behaftet. Zunächst erlaubt die durch die verschiedenen Szenarien aufgespannte Schwankungsbreite interessierender Größen keine probabilistische Interpretation – wie wahrscheinlich es beispielsweise ist, dass der Bevölkerungsumfang zwischen dem höchsten und dem niedrigsten errechneten Wert liegt, wird nicht angegeben und somit eine Abschätzung der Unsicherheit nicht ermöglicht (vgl. bspw. Lutz & Scherbov 1998: 84).<sup>113</sup>

Zudem kann die Verwendung von Szenarien je nach Zusammenstellung der Annahmen in einzelnen Szenarien zu inkonsistenten Ergebnissen führen. Lee (1998) führt als Beispiel eine Bevölkerungsvorausberechnung für die USA für das Jahr 2070 an. Bei dieser liegt die Schwankungsbreite um das mittlere Szenario bei der Größe der Bevölkerung im Alter von 0 bis 19 Jahren bei  $\pm 34\%$ , die Bandbreite für die Bevölkerung im Alter von 20 bis 64 Jahren bei  $\pm 20\%$  und für die Zahl der Personen im Alter von 65 und mehr Lebensjahren bei  $\pm 9\%$  (vgl. ebd.: 166). Die Bandbreite des aus diesen Bevölkerungsbeständen berechneten Gesamtquotienten beträgt allerdings nur  $\pm 5\%$  (vgl. ebd.: 166). Ursache hierfür ist, dass Szenarien mit bestimmten Kombinationen von Annahmen für Mortalität, Fertilität und Migration entweder die Unsicherheit des Bevölkerungsumfangs beziehungsweise des Umfangs einzelner Altersgruppen oder aber daraus abgeleiteter Kennwerte wie dem Gesamtquotienten unterschätzen (vgl. ebd.: 166). Wird beispielsweise ein Szenario mit hoher Fertilität, hoher Lebenserwartung und hoher Migration, sowie eines, bei dem diese drei niedrig sind, und eines, bei dem diese im mittleren Bereich liegen, berechnet, zeigen sich beim Vergleich der aus diesen drei Szenarien resultierenden Bevölkerungsgrößen relativ unterschiedliche Werte. Kennwerte wie

---

113 Sieht man von diesem Problem ab und nimmt an, dass Größen wie der Bevölkerungsumfang zumindest mit „großer Wahrscheinlichkeit“ zwischen dem höchsten und dem niedrigsten Wert liegen, kann ein Vergleich mit Verfahren durchgeführt werden, die eine probabilistische Interpretation der Ergebnisse erlauben. Es zeigt sich, dass aus Szenarien resultierende Intervalle oftmals vergleichsweise eng ausfallen und Unsicherheit keinesfalls über- sondern eher unterschätzt wird (Alho & Spencer 1990a,b; Alho 1997). Ein solcher Vergleich für die zehnte koordinierte Bevölkerungsvorausberechnung des Statistischen Bundesamtes findet sich bei Lipps & Betz (2005).

der Gesamtquotient werden hingegen nicht stark variieren, da die Verhältnisse der einzelnen Altersgruppen zueinander durch die Wahl der Szenarien nicht im gleichen Maße variieren wie deren absolute Werte.<sup>114</sup>

Werden Szenarien wie gerade beschrieben gewählt („Hoch-Mittel-Tief“), wird weiterhin zumindest implizit von einer hohen Korrelation der verschiedenen demographischen Komponenten ausgegangen (Lipps & Betz 2005: 6). Zwar lässt sich dies durch eine entsprechende Wahl der Szenarien umgehen, allerdings wird auch wenn dies der Fall ist oftmals von einer „perfekten“ Autokorrelation bezogen auf die einzelnen Prozesse ausgegangen (vgl. bspw. Lee 1998: 167). Dies meint, dass für einzelne Szenarien entweder Komponenten konstant gehalten werden oder aber ein konstanter Anstieg (bzw. Abnahme) angenommen wird (vgl. ebd.: 167). So können Szenarien bezüglich der einzelnen Komponenten eindeutig in eine Ordnung gebracht werden – also zum Beispiel nach der Höhe der Fertilität geordnet werden (vgl. ebd.: 167). Szenarien hingegen, bei denen beispielsweise die Fertilität erst niedrig ist, dann deutlich ansteigt und schließlich wieder abnimmt, werden häufig nicht verwendet (vgl. ebd.: 167). Von vielen prinzipiell denkbaren Verläufen wird also nur eine kleine, selektive Auswahl betrachtet.

Insgesamt lassen sich bei szenariobasierten Bevölkerungsvorausberechnungen somit vier Probleme ausmachen:

1. Keine probabilistische Interpretation
2. Mögliche Inkonsistenzen
3. Hohe Korrelation der Komponenten untereinander
4. Jeweils perfekte Autokorrelation bei den einzelnen Komponenten

Hierdurch ist Unsicherheit nicht genau quantifizierbar und kann unterschätzt werden.<sup>115</sup>

### 5.1.2 Grundidee probabilistischer Vorausberechnungen

Um die genannten Probleme zu lösen, wurden probabilistische Bevölkerungsvorausberechnungen eingeführt. Zwar existieren diverse Varianten (s. Unterabschnitt 5.1.3), die zugrunde liegende Idee ist allerdings bei allen diesen Varianten identisch.

---

114 Betrachtet man die 12. koordinierte Bevölkerungsvorausberechnung des Statistischen Bundesamtes für das Jahr 2060 (s. Statistisches Bundesamt 2010c), zeigen sich ähnliche Ergebnisse. Lässt man die drei explizit als „Modellrechnungen“ bezeichneten Szenarien außen vor und bezieht man die Bandbreite der in den übrigen 13 Szenarien resultierenden Größen mangels eines „mittleren“ Szenarios auf deren Mittelwert, ergibt sich für die Zahl der Personen unter 20 Jahren eine Spannweite von etwa  $\pm 20\%$ , für Personen von 20 bis einschließlich 64 Jahren von etwa  $\pm 10\%$  und für Personen im Alter von 65 und mehr Jahren von circa  $\pm 6\%$  (alle Angaben gerundet; bei auftretenden Asymmetrien der Größenordnung der Abweichungen vom Mittelwert nach oben und unten wurde immer der kleinere Wert angegeben). Die Spannweite des Gesamtquotienten reicht hingegen von  $-7.5\%$  bis  $+6.4\%$ . Werden zusätzlich die als Modellrechnungen bezeichneten Szenarien bei diesen Berechnungen berücksichtigt, werden die Spannbreiten für die drei Altersgruppen deutlich breiter, während sich diese für den Gesamtquotienten leicht reduziert. Für Finnland finden sich ähnliche Berechnungen bei Keilman et al. (2002), ein weiteres auf die USA bezogenes Beispiel liefern Lee & Tuljapurkar (1994).

115 Die 12. koordinierte Bevölkerungsvorausberechnung des Statistischen Bundesamtes umfasst insgesamt 15 verschiedene Szenarien (vgl. Statistisches Bundesamt 2010c: 38). Für die Gültigkeit der genannten Argumente ist die Zahl der Szenarien aber unerheblich, solange diese nicht deutlich höher liegt.

Grundsätzlich wird davon ausgegangen, dass zukünftige Werte von benötigten Modellparametern – wie demographischen Raten oder bestimmte Kennziffern – als Zufallsvariablen aufgefasst werden können. Dies meint, dass für solch einen Parameter für einen bestimmten Zeitpunkt nicht ein bestimmter Wert oder einige wenige Werte angegeben werden, sondern eine Wahrscheinlichkeitsverteilung spezifiziert wird. Zum Beispiel könnte man für die zusammengefasste Geburtenziffer im Jahr 2020 eine Normalverteilung mit Erwartungswert 1.4 und Standardabweichung 0.1 angeben, oder eine stetige Gleichverteilung über das Intervall [1.3, 1.5]. Die Implikationen dieser Annahme für den Bevölkerungsumfang oder andere interessierende Größen lassen sich anschließend analytisch oder über Simulationen bestimmen. Hiervon ausgehend können dann beispielsweise Aussagen darüber getroffen werden, zwischen welchen Werten der Bevölkerungsumfang mit einer Wahrscheinlichkeit von 90% liegt.

Der Vorteil dieses Vorgehens ist vor allem darin zu sehen, dass auf diese Weise Vorausberechnungsintervalle für Ergebnisse angegeben werden können.<sup>116</sup> Auf diese Weise kann das Ausmaß der Unsicherheit der Ergebnisse zugänglich gemacht werden. Modelle, deren Ergebnisse eine solche Interpretation und somit eine Quantifizierung von Unsicherheit zulassen, werden im Anschluss an Lee (1998) als probabilistische Modelle bezeichnet; als stochastische Modelle werden probabilistische Modelle bezeichnet, bei denen die Modellparameter als stochastische Prozesse modelliert werden (ebenfalls in Anlehnung an Lee 1998).

Die anderen oben aufgelisteten Probleme szenariobasierter Bevölkerungsvorausberechnungen, die zu einer Unterschätzung von Unsicherheit führen, lassen sich bei entsprechender Modellierung ebenfalls beheben – allerdings ist dies bei probabilistischen Modellen nicht allgemein der Fall.

### 5.1.3 Varianten probabilistischer Vorausberechnungen

Die erste probabilistische Bevölkerungsvorausberechnung wird in das Jahr 1949 datiert und auf Törnqvist zurückgeführt, allerdings wurde dessen Ansatz aufgrund eines sehr hohen Bedarfs an Daten und eines hohen Rechenaufwands zunächst nicht weiter verfolgt (vgl. Alho 1984: 100). Ende der sechziger Jahre gab es weiterhin einige erste theoretische Überlegungen zu probabilistischen Vorausberechnungen (vgl. Pollard 1968; Sykes 1969). Eine vollständige theoretische Ausarbeitung und umfangreiche praktische Anwendung fand allerdings erst ab den achtziger Jahren statt. Seitdem haben sich vier Varianten der probabilistischen Vorausberechnung herausgebildet, die sich darin unterscheiden, wie die Verteilungen der Modellparameter bestimmt werden (vgl. Lee 1998: 170ff.):

1. Ex-Post-Evaluation
2. Statistische Modelle für Aggregate

---

116 In der zumeist englischsprachigen Literatur zu probabilistischen Vorausberechnungen sind die Begriffe *confidence interval*, *probability interval* und *prediction interval* gebräuchlich (vgl. Sanderson 1995: 259). Allerdings werden die beiden zuerst genannten Begriffe auch in anderen Kontexten verwendet (vgl. Chatfield 1993), weshalb sie hier nicht benutzt werden sollen. Die übliche Übersetzung von *prediction interval* mit „Prognoseintervall“ soll hier ebenfalls nicht verwendet und stattdessen der Begriff „Vorausberechnungsintervall“ benutzt werden, da der Begriff „Prognose“ entsprechend der Ausführungen in Kapitel 3 mit einem Anspruch verbunden ist, der nicht unbedingt von probabilistischen Vorausberechnungen erfüllt wird.

3. Experten-basierte Ansätze
4. Statistische Modelle für demographische Raten

Der erste, der zweite und der vierte Ansatz lassen sich nach der hier verwendeten Terminologie als stochastische Ansätze bezeichnen.<sup>117</sup>

Der erste Ansatz wurde von Keyfitz (1981) und Stoto (1983) verfolgt. Die grundlegende Idee besteht darin, für eine interessierende Größe, wie beispielsweise den Bevölkerungsumfang, zunächst die Fehler älterer Vorausberechnungen zu evaluieren (vgl. Keyfitz 1981: 579). Anschließend wird angenommen, dass auch bei aktuellen Vorausberechnungen Fehler im gleichen Umfang gemacht werden (vgl. ebd.: 580). Liegt eine solche aktuelle Vorausberechnung vor und werden weiterhin Verteilungsannahmen für die Fehler getroffen, lassen sich für die gewählte Größe Vorausberechnungsintervalle berechnen (vgl. Stoto 1983: 17). Insofern wird die interessierende Größe direkt als Zufallsvariable aufgefasst und nicht aus demographischen Raten abgeleitet. Gleiches gilt bei der zweiten Gruppe von Ansätzen, bei denen die interessierende Größe allerdings über statistische Modelle abgebildet wird. Beispielsweise verwendeten Cohen (1986), Pflaumer (1992) und Saboia (1974) statistische Zeitreihenmodelle zur Modellierung des Bevölkerungsumfangs.

Der dritten Gruppe von Ansätzen dient als Grundlage eine Einschätzung der Verteilung bestimmter Modellparameter durch Experten. Beispielsweise erfragten Lutz et al. (1996a) von Experten eine subjektive Einschätzung, zwischen welchen Werten die zusammengefasste Geburtenziffer in einem zukünftigen Jahr  $t$  mit einer Wahrscheinlichkeit von 95% liegt, sowie einen mittleren Wert (vgl. ebd.: 37). Solche Einschätzungen wurden für mehrere Zeitpunkte sowie ebenfalls für Mortalität und Migration erfragt und dies für insgesamt 13 Weltregionen (vgl. Lutz et al. 1997: 803). Die Experteneinschätzungen wurden in Normalverteilungen überführt und aus diesen zufällige Werte gezogen (vgl. Lutz et al. 1996b: 403 u. 427). Unter Verwendung von Interpolationsverfahren wurden anschließend Bevölkerungsvorausberechnungen für die einzelnen Weltregionen durchgeführt, und aus diesen Berechnungen Verteilungen für bestimmte Kennwerte abgeleitet (vgl. ebd.: 427). Die Möglichkeit, Expertenmeinungen aus vorhandenen Vorausberechnungen beziehungsweise der Fachliteratur zu gewinnen und diese dann in Wahrscheinlichkeitsverteilungen zu überführen, wurde von Alho (1984, 1985, 1997), Lutz & Scherbov (1998) und Pflaumer (1984, 1986, 1988) genutzt.

Statistische Modelle können auch für die Vorausberechnung von demographischen Raten benutzt werden, die anschließend mit der Kohorten-Komponenten-Methode zur Vorausberechnung und Erstellung von Vorausberechnungsintervallen kombiniert werden können. Der letztgenannte Ansatz wurde vor allem von Lee & Tuljapurkar (1994) populär gemacht und dürfte mittlerweile die am meisten genutzte Variante sein, was vermutlich auch auf die Vielzahl

---

117 Neben den bereits genannten Arbeiten von Pollard (1968) und Sykes (1969) sind vor allem die folgenden Arbeiten von grundsätzlicher theoretischer Bedeutung: Alho & Spencer (1991); Cohen (1977); Tuljapurkar & Orzack (1980); Tuljapurkar (1986); eine Übersicht bietet Tuljapurkar (1989, 1992). Anwendungen probabilistischer Bevölkerungsvorausberechnungen für Deutschland finden sich bei Babel (2007); Betz & Lipps (2004); Bomsdorf et al. (2008); Härdle & Mysickova (2009); Lipps & Betz (2005); Lutz & Scherbov (1998); Pflaumer (1984, 1986). Arbeiten, die auf probabilistische Bevölkerungsvorausberechnungen aufsetzen, um ökonomische Fragestellungen zu untersuchen, sind von Lee & Tuljapurkar (1998) und Fehr & Habermann (2006) erstellt worden. Eine probabilistische Haushaltsvorausberechnung für Norwegen findet sich bei Alho & Keilman (2010) und die Berücksichtigung probabilistischer Bevölkerungsvorausberechnungen im Rahmen von Fragen des Klimawandels bei O'Neill (2004). Probabilistische Vorausberechnungen für den Hochschulbereich wurden von Boes (2004) vorgelegt.

an speziellen Zeitreihenmodellen für Fertilität und insbesondere Mortalität zurückzuführen ist (für eine weitergehende Diskussion dieser s. Kapitel 8). Die grundlegende Idee besteht darin, dass demographische Raten oder bestimmte Kennwerte, die diese Raten beschreiben, als stochastische Prozesse aufgefasst werden, die über statistische Modelle abgebildet werden können (vgl. ebd.: 1177). Beispielsweise gehen Härdle & Mysickova (2009) unter Verwendung eines Ansatzes von Lee (1993) davon aus, dass sich die zusammengefasste Geburtenziffer in der BRD ab dem Jahr 1976 als *random walk* modellieren lässt (vgl. Härdle & Mysickova 2009: 12 f.). Hierbei ergibt sich die zusammengefasste Geburtenziffer für eine Zeitstelle  $t + 1$  als Summe des vorhergegangenen Wertes zu  $t$  und eines zufälligen, normalverteilten Wertes, wobei die Parameter der Normalverteilung aus der historischen Zeitreihe der zusammengefassten Geburtenziffer geschätzt werden (vgl. ebd.: 13). Ausgehend von der letzten Zeitstelle der historischen Zeitreihe kann dann über wiederholtes hinzu addieren des normalverteilten Wertes ein möglicher, zukünftiger Verlauf simuliert werden. Durch die Erzeugung vieler solcher Verläufe, die in altersspezifische Geburtenraten überführt und in Kombination mit der Kohorten-Komponenten-Methode zur Berechnung einzelner Entwicklungsverläufe verwendet werden, lassen sich die Verteilungen interessierender Größen herleiten.

#### 5.1.4 Diskussion und Übertragung auf die Vorausberechnung von Verwandtschaft

Umfassende, allgemeine Diskussionen der Vor- und Nachteile der Ansätze findet man bei Lee (1998), Lutz et al. (1996b) und Tuljapurkar et al. (2004). Bezogen auf die Vorausberechnung von Verwandtschaft ist festzuhalten, dass die Ex-Post-Evaluation und die statistische Modellierung interessierender Größen nicht in Frage kommen, da lediglich eine ältere Vorausberechnung von Verwandtschaft vorliegt (Galler 1990) und historische Zeitreihen zu Verwandtschaftsverhältnissen fehlen. Die Verwendung von Expertenmeinungen scheint dem Ziel der Berücksichtigung von Unsicherheit entgegenzulaufen, zumal sich die bei älteren Vorausberechnungen eingeflossenen Einschätzungen bezüglich der Entwicklung demographischer Komponenten wie zu Beginn dieses und des letzten Kapitels erwähnt oftmals als falsch erwiesen und gerade als Motivation zur Entwicklung von probabilistischen Vorausberechnungen gedient haben.<sup>118</sup> Insofern scheinen lediglich statistische Modelle für demographische Komponenten in Frage zu kommen.

Allerdings müssen auch bei Verwendung statistischer Modelle mehrere Aspekte beachtet werden. Zunächst werden bei diesem Vorgehen Expertenmeinungen zumindest implizit einen Einfluss haben und können nicht gänzlich ausgeschlossen werden, wie beispielsweise bei der Modellwahl (vgl. Booth 2006: 552). Ferner sollte das verwendete Modell die Daten, auf deren Grundlage es geschätzt wurde, gut repräsentieren (vgl. bspw. Lee 1998: 181). Dies lässt sich in

---

118 In der Literatur wird bezweifelt, dass Experten Unsicherheit verlässlich einschätzen können (s. insb. Tversky & Kahneman 1974: 1129ff.). Auch die Überführung von Expertenmeinungen in Wahrscheinlichkeitsverteilungen erscheint zweifelhaft, zumal diese ad hoc gesetzt und nicht begründet wird. Insbesondere besteht im Gegensatz zur Ex-Post-Evaluation oder zu Zeitreihenmodellen keine Möglichkeit, diese Verteilungsannahmen in irgendeiner Art und Weise empirisch zu überprüfen. Des Weiteren ist die Frage, wie von Experten gegebene Intervalle zu interpretieren sind. Werden Experten beispielsweise explizit nach 90%-Intervallen gefragt, ist unklar, inwiefern diese sinnvoll von 95%-Intervallen unterschieden werden können (vgl. ebd.: 1129ff.). Werden Intervalle aus der Literatur gewonnen ist gänzlich unklar, wie zu einer Wahrscheinlichkeitsinterpretation gelangt wird, da in der Regel keine Wahrscheinlichkeiten für unterschiedliche Szenarioannahmen gegeben werden.

mehrere Aspekte unterteilen, wie beispielsweise die Modellspezifikation oder die Schätzung von Parametern (vgl. ebd.: 181). Die Überprüfung dieser Unterpunkte kann sich dabei als relativ aufwendig erweisen (vgl. Lutz et al. 1996b: 399). Schließlich muss beachtet werden, dass demographische Raten allein noch keine Aussagen über Verwandtschaft erlauben und mit einem Ansatz kombiniert werden müssen, der entsprechende Aussagen möglich macht.

Es wird also ein Ansatz zur Betrachtung linearer Verwandter benötigt, der von demographischen Raten ausgeht und entsprechend der Ausführungen aus Abschnitt 4.8 möglichst wenige zusätzliche Daten benötigt und beide Geschlechter berücksichtigt. Ein solcher Ansatz wird im weiteren Verlauf dieses Kapitels vorgestellt. Zunächst wird der theoretische Rahmen und dessen Bezug zu Bevölkerungsvorausberechnungen erklärt und gezeigt, wie ausgehend von statistischen Modellen für demographische Komponenten Unsicherheit berücksichtigt werden kann. Anschließend wird erläutert, wie in diesem Rahmen die Zahl lebender Verwandter aus demographischen Raten abgeleitet werden kann. Hierauf folgt eine Diskussion der Annahmen des Modells. Die Modellierung der demographischen Komponenten wird in Kapitel 8 vorgestellt und diskutiert.

## 5.2 Diskrete Multityp-Galton-Watson-Prozesse

### 5.2.1 Definition und Eigenschaften von diskreten Multityp-Galton-Watson-Prozessen

#### 5.2.1.1 Definition

Ein Multityp-Galton-Watson-Prozess ist ein vektorwertiger, homogener und diskreter Markov-Prozess  $\mathbf{n}_0, \mathbf{n}_1, \mathbf{n}_2, \dots$  für eine Folge von Zeitpunkten  $t = 0, 1, \dots$ , wobei allgemein  $\mathbf{n}_t \in \mathbb{N}^k$  gilt (vgl. Harris 1963: 36). Im weiteren wird angenommen, dass  $\mathbf{n}_0$  gegeben und somit nicht zufällig ist (vgl. ebd.: 36). Ferner seien die betrachteten Zeitpunkte äquidistant, so dass zwischen zwei aufeinander folgende Zeitpunkte genau eine Zeiteinheit liegt, wie beispielsweise ein Jahr. Die einzelnen Vektoren  $\mathbf{n}_t = (n_{1,t}, \dots, n_{k,t})$  können als Populationsvektoren aufgefasst werden, bei denen der Eintrag  $n_{i,t}$  die Zahl von Individuen vom Typus  $i$  wiedergibt, wobei es insgesamt  $k$  Typen gibt.<sup>119</sup>

Die Verteilung von  $\mathbf{n}_{t+1}$  ist, gegeben  $\mathbf{n}_t$ , durch  $k$  wahrscheinlichkeitserzeugende Funktionen

$$f_i(s_1, \dots, s_k) = \sum_{z_1=0}^{\infty} \cdots \sum_{z_k=0}^{\infty} p_i(z_1, \dots, z_k) s_1^{z_1} \cdots s_k^{z_k} \quad (5.1)$$

mit  $|s_1|, \dots, |s_k| < 1$  und  $0 \leq p_i(z_1, \dots, z_k) \leq 1$  sowie

$$\sum_{z_1=0}^{\infty} \cdots \sum_{z_k=0}^{\infty} p_i(z_1, \dots, z_k) = 1 \quad (5.2)$$

<sup>119</sup> Im Gegensatz zur Definition des Bevölkerungsvektors bei der Kohorten-Komponenten-Methode in (3.15) wird hier mit  $\mathbf{n}_t = (n_{1,t}, \dots, n_{k,t})$  vorerst davon ausgegangen, dass es sich bei  $\mathbf{n}_t$  um einen Zeilenvektor handelt.

und  $z_j \in \mathbb{N}$  eindeutig charakterisiert (vgl. Harris 1963: 36). Hierbei ist  $p_i(z_1, \dots, z_k)$  die Wahrscheinlichkeit, dass ein Individuum vom Typus  $i$   $z_1$  Nachfahren vom Typus 1 hat,  $z_2$  Nachfahren vom Typus 2 und so fort bis  $z_k$  Nachfahren vom Typus  $k$  (vgl. ebd.: 36).

Wird von einem inhomogenen Markov-Prozess ausgegangen, müssen für jede Zeitstelle insgesamt  $k$  wahrscheinlichkeitserzeugende Funktionen

$$f_{i,t}(s_1, \dots, s_k) = \sum_{z_1=0}^{\infty} \cdots \sum_{z_k=0}^{\infty} p_{i,t}(z_1, \dots, z_k) s_1^{z_1} \cdots s_k^{z_k} \quad (5.3)$$

angegeben werden, wobei diese entweder deterministisch gegeben sind oder als zufällig aufgefasst werden können (vgl. Haccou et al. 2005: 46f.). Bis auf weiteres wird angenommen, dass diese Funktionen deterministisch gegeben sind.

Die Übergangswahrscheinlichkeiten des Prozesses leiten sich sowohl im homogenen als auch inhomogenen Fall aus Wahrscheinlichkeitsverteilungen her, die sich auf Individuen beziehen. Dabei hängen diese Verteilungen nur vom aktuellen Zustand eines Individuums und nicht von vergangenen Zuständen ab. Zudem sind die Individuen unabhängig voneinander. Hieraus folgt, dass der Prozess die Markov-Eigenschaft aufweist – die Wahrscheinlichkeit, dass auf einen bestimmten Vektor  $\mathbf{n}_t$  ein bestimmter Vektor  $\mathbf{n}_{t+1}$  folgt, hängt nur von  $\mathbf{n}_t$  und den Übergangswahrscheinlichkeiten ab und nicht von vorhergegangenen Vektoren  $\mathbf{n}_{t-1}, \mathbf{n}_{t-2}, \dots, \mathbf{n}_0$ .

### 5.2.1.2 Interpretation

Im weiteren Verlauf wird wie bisher in dieser Arbeit von einzelnen Jahren als Zeiteinheit ausgegangen, wobei sich einzelne Zeitpunkte  $t$  wie bereits in Abschnitt 3.1 erläutert immer auf den ersten Januar eines Jahres beziehen sollen. Betrachtet wird eine Bevölkerung unterteilt nach den Merkmalen Alter und Geschlecht. Diese Eigenschaften werden durch die Mengen  $A = \{0, 1, 2, \dots, \rho\}$  und  $G = \{m, w\}$  repräsentiert, wobei  $m$  für männlich und  $w$  für weiblich steht und  $\rho \in \mathbb{N}$  das höchste erreichbare Lebensalter ist. Die einzelnen  $a \in A$  repräsentieren Intervalle exakten Alters  $[a, a + 1)$  und werden im weiteren auch als Altersklassen bezeichnet. Die Elemente aus  $A \times G$  sind die Typen des betrachteten Prozesses, die entsprechend der bisher eingeführten Terminologie auch als Zustände bezeichnet werden. Somit ist  $k = \#\{A \times G\}$  die Zahl der möglichen unterschiedlichen Typen in der Bevölkerung. Ferner wird davon ausgegangen, dass die Zustände in irgendeiner Art und Weise nummeriert sind, so dass zur Identifikation eines Zustandes die Angabe eines  $i \in \{1, \dots, k\}$  ausreicht. Neben dieser vereinfachenden Schreibweise wird im folgenden aber auch direkt auf Paare  $(a, g) \in A \times G$  Bezug genommen.<sup>120</sup>

Bei der Modellierung menschlicher Populationen werden nur solche Wahrscheinlichkeiten  $p_i(\dots)$  größer Null sein, bei denen Individuen in Altersklasse  $a$  und mit Geschlecht  $g$  Nachfahren in Altersklasse  $a + 1$  mit demselben Geschlecht  $g$  oder aber Nachfahren beliebigen Geschlechts in Altersklasse 0 haben. Der Einfachheit halber wird zunächst angenommen, dass alle  $p_{(a,m)}(\dots)$  mit  $z_{(0,m)} > 0$  oder  $z_{(0,w)} > 0$  gleich Null sind – also

<sup>120</sup> Im weiteren werden Merkmalskombinationen auch bei den bereits eingeführten Symbolen in Klammern gesetzt, also beispielsweise  $J_{(x,w),t}$  anstelle von  $J_{x,w,t}$ . Wird lediglich ein Merkmal betrachtet, wird auch weiterhin auf Klammern verzichtet und beispielsweise  $J_{x,t}$  benutzt.

Männer keine Nachfahren bekommen. Gleiches wird für Frauen in Altersklassen, die nicht in der reproduktiven Phase liegen, gesetzt. Für diese beiden Gruppen sind lediglich die Wahrscheinlichkeiten  $p_{(a,g)}(z_{(0,g)} = 0, \dots, z_{(a,g)} = 0, z_{(a+1,g)} = 1, z_{(a+2,g)} = 0, \dots)$  und  $p_{(a,g)}(z_{(0,g)} = 0, \dots, z_{(a,g)} = 0, z_{(a+1,g)} = 0, z_{(a+2,g)} = 0, \dots)$  größer Null. Für die Altersklasse  $\rho$  sind alle  $p_{(\rho,g)}(\dots)$  gleich Null.

Je nach Werten für  $z_j$  lassen sich die Wahrscheinlichkeiten  $p_i(\dots)$  dann als Überlebenswahrscheinlichkeiten, als Geburtswahrscheinlichkeiten oder als Kombinationen aus diesen beiden interpretieren. Zwar ist das Überleben eines Individuums formal nicht von der Erzeugung eines neuen Individuums zu unterscheiden. Um ein Jahr zu überleben, müsste beispielsweise ein Individuum in Altersklasse 40 mit Geschlecht  $g$  ein Individuum in Altersklasse 41 mit demselben Geschlecht „erzeugen“ – ist dies nicht der Fall, ist dies gleichbedeutend mit dem Tod des Individuums. Wie dieses Beispiel aber verdeutlicht, lässt sich je nachdem, zu welchem Typus erzeugte Individuen gehören durchaus von Geburten und Überleben im üblichen Sinne reden. Ferner ist zu beachten, dass Individuen nicht identifiziert sind, wenn lediglich Bevölkerungsvektoren bekannt sind: wenn beispielsweise zu  $t$  drei Individuen in Altersklasse  $x$  sind und zu  $t + 1$  zwei Individuen in Altersklasse  $x + 1$  und eines in Altersklasse 0, lässt sich nicht feststellen, welches der drei Individuen zu  $t$  gestorben ist und welches Individuum der Vorfahr des Individuums in Altersklasse 0 ist.

### 5.2.1.3 Beispiel

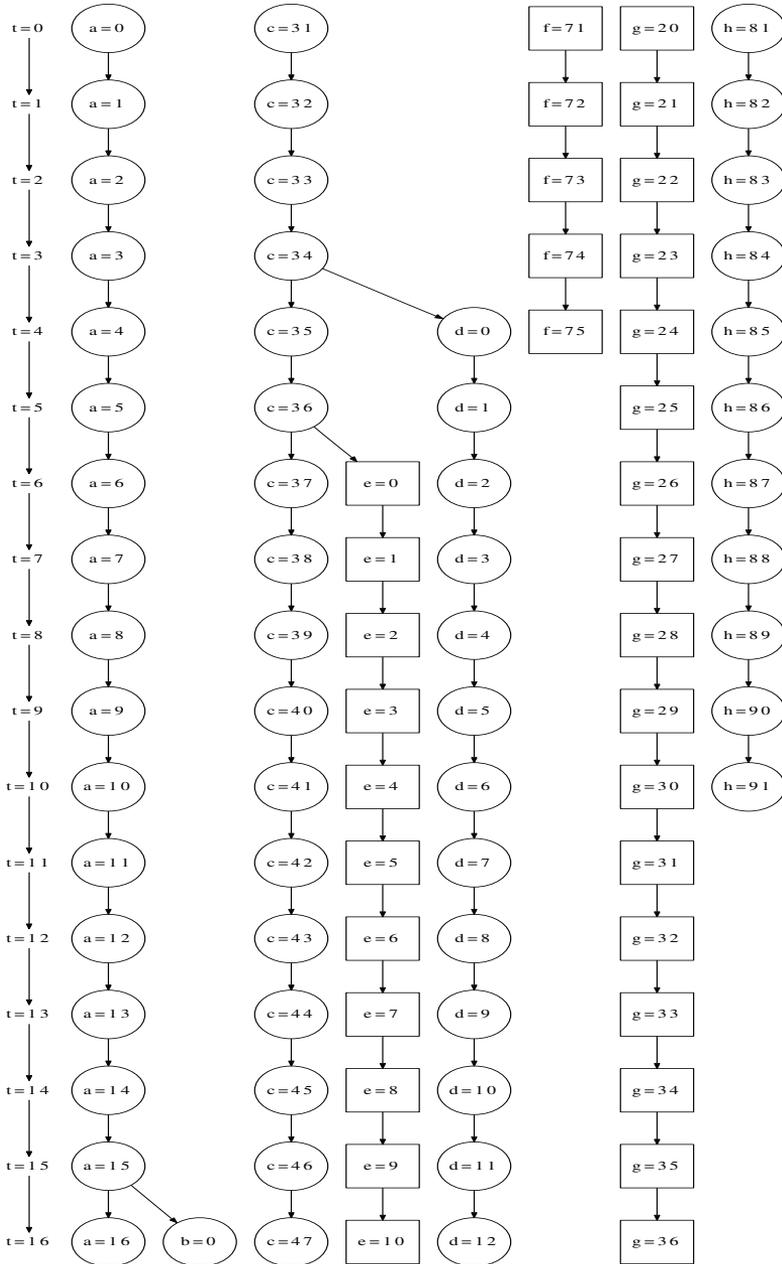
Zum besseren Verständnis findet sich ein Beispiel für eine mögliche Realisierung eines Multityp-Verzweigungsprozesses ausgehend von einer bestimmten Ausgangsbevölkerung in Abbildung 5.1. Abgetragen ist die Zusammensetzung der Bevölkerung zu einer Folge von Zeitpunkten  $t = 0, 1, \dots, 16$ . Runde Symbole stehen für weibliche Individuen und rechteckige für männliche. Die Buchstaben dienen als Personenbezeichner, die dahinter angegebenen Werte geben das Alter wieder. Der Anschaulichkeit halber wird in dieser Darstellung die Nicht-Identifikation der Individuen somit aufgehoben.<sup>121</sup>

Zu  $t = 0$  besteht die Bevölkerung aus den drei Frauen  $a$ ,  $c$  und  $h$  in den Altersklassen 0, 31 und 81 sowie zwei Männern  $f$  und  $g$  in den Altersklassen 20 und 71. Zwischen  $t = 3$  und  $t = 4$  bekommt Individuum  $c$  eine Tochter (Individuum  $d$ ). Individuum  $f$  überlebt bis zum Zeitpunkt  $t = 4$  und erreicht die Altersklasse 75. Zwischen  $t = 5$  und  $t = 6$  bekommt  $c$  einen Sohn (Individuum  $e$ ). Individuum  $h$  überlebt bis zum Zeitpunkt  $t = 10$ , erreicht die Altersklasse 91 und stirbt zwischen  $t = 10$  und  $t = 11$ . Zwischen  $t = 15$  und  $t = 16$  bekommt Individuum  $a$  eine Tochter. Zum letzten Zeitpunkt besteht die Bevölkerung aus den Individuen  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$  und  $g$ .

Die Grafik verdeutlicht, dass zwar Informationen darüber vorhanden sind, ob ein Ereignis zwischen  $t$  und  $t + 1$  eintritt oder nicht, allerdings ist nicht bekannt, wann genau das Ereignis zwischen  $t$  und  $t + 1$  eintritt. Hierdurch sind auch keine genauen Aussagen über exakte Alter von Personen möglich. Beispielsweise kann man für Individuum  $d$  für  $t = 4$  nur mit

<sup>121</sup> Durch die Wahl der Startbevölkerung und des Verlaufs bei diesem Beispiel entstehen durch die Identifikation kaum Probleme bei der Darstellung des Modells. Nimmt man wie erwähnt an, dass nur Frauen Nachwuchs zeugen, die reproduktive Phase mit der Altersklasse 15 beginnt und mit der Altersklasse 49 endet und ferner keine Migration vorliegt, ist lediglich die Zuweisung von  $b$  zu  $a$  als Tochter problematisch, da  $b$  unter den genannten Annahmen auch Tochter von  $c$  sein könnte.

Abbildung 5.1: Vereinfachte Darstellung eines beispielhaften Verlaufs eines Multityp-Galton-Watson-Prozesses



Quelle: eigene Darstellung

Sicherheit angeben, dass das erste Lebensjahr noch nicht vollendet wurde und das exakte Alter im Intervall  $[0, 1)$  liegen muss.

Wie oben erwähnt, ist der dargestellte Verlauf ausgehend von der Bevölkerung zu  $t = 0$  nur einer von vielen möglichen Verläufen. Beispielsweise hätte Individuum  $g$  nur bis zum Zeitpunkt  $t = 4$  und Altersklasse 24 überleben können. Für die möglichen Folgebestände ausgehend von einem Ausgangsbestand lassen sich Wahrscheinlichkeiten angeben, die sich aufgrund der Unabhängigkeit der Individuen voneinander als Produkte der einzelnen Wahrscheinlichkeiten  $p_i(z_1, \dots, z_k)$  ergeben. Hiervon ausgehend lassen sich auch Erwartungswerte für Bevölkerungsbestände betrachten.

### 5.2.2 Bezug zum Kohorten-Komponenten-Modell

Sei  $\mathbf{n}_t = (n_{1,t}, n_{2,t}, \dots, n_{k,t})$  ein Bevölkerungsvektor, der zunächst als fix angenommen wird. Gegeben  $\mathbf{n}_t$  lässt sich der Zufallsvektor  $\mathbf{n}_{t+1}$  als Summe von  $|\mathbf{n}_t|$  unabhängigen Zufallsvektoren schreiben (vgl. Harris 1963: 37). Von diesen haben  $n_{1,t}$  Zufallsvektoren die wahrscheinlichkeitserzeugende Funktion  $f_1(\cdot)$  beziehungsweise  $f_{1,t}(\cdot)$ ,  $n_{2,t}$  haben die wahrscheinlichkeitserzeugende Funktion  $f_2(\cdot)$  beziehungsweise  $f_{2,t}(\cdot)$  und so fort. Betrachtet wird zunächst der homogene Fall.

Sei  $\mathbf{e}_i$  ein Vektor, der an der Stelle  $i$  den Eintrag 1 aufweist, 0 sonst. Ferner sei  $\mathbf{A}$  eine Matrix mit Einträgen  $a_{ij}$ , wobei diese gegeben sind durch (vgl. Harris 1963: 36)

$$a_{ij} = E(n_{j,t+1} | \mathbf{n}_t = \mathbf{e}_i) = \frac{\delta f_i(1, \dots, 1)}{\delta s_j} \quad (5.4)$$

$a_{ij}$  ist somit der Erwartungswert der Zahl an Nachfahren des Typus  $j$  eines Individuums des Typus  $i$ . Der Erwartungswert von  $\mathbf{n}_{t+1}$  gegeben  $\mathbf{n}_t$  lässt sich unter Verwendung von  $\mathbf{A}$  schreiben als (vgl. ebd.: 37):

$$E(\mathbf{n}_{t+1} | \mathbf{n}_t) = \mathbf{n}_t \mathbf{A} \quad (5.5)$$

Transponiert man beide Seiten, und lässt die sich aus der stochastischen Formulierung ergebende Variabilität außer Betracht, erhält man das Kohorten-Komponenten-Modell in der Matrixnotation von Leslie (1945: 184) beziehungsweise je nach Zustandsraum die allgemeinere Formulierung aus Abschnitt 4.5.<sup>122 123</sup>

$$\mathbf{n}'_{t+1} = \mathbf{A}' \mathbf{n}'_t \quad (5.6)$$

Im weiteren Verlauf wird der Einfachheit halber auf die Angabe der Transponierung verzichtet. Für den inhomogenen Fall gilt (vgl. Caswell 2001: 502)

$$\mathbf{n}_{t+1} = \mathbf{A}_t \mathbf{n}_t \quad (5.7)$$

122 Den umgekehrten Fall einer Konstruktion eines Multityp-Galton-Watson-Prozesses ausgehend von einer Leslie-Matrix beschreibt Caswell (2001: 479ff.).

123 Im folgenden Unterabschnitt 5.2.3 wird begründet, warum die Variabilität nicht berücksichtigt werden muss. Die Variante in Formel (5.5) mit dem Erwartungswert auf der linken Seite hat gegenüber der ursprünglichen Formulierung von Leslie in (5.6) bzw. (3.14) und (5.7) den Vorteil, dass Dezimalzahlen als Einträge von  $E(\mathbf{n}_t)$  ohne weiteres zu interpretieren sind, aber für  $\mathbf{n}_t$  selbst wenig sinnvoll sind. Beispielsweise wäre  $E(n_{(50,w),t}) = 100.5$  unproblematisch, während  $n_{(50,w),t} = 100.5$  darauf hinauslaufen würde, dass es 100 und eine halbe Frau in Altersklasse 50 gibt.

Geht man allgemeiner davon aus, dass nur  $\mathbf{n}_0$  gegeben ist, erhält man

$$\mathbf{n}_{t+1} = \prod_{i=0}^t \mathbf{A}_i \mathbf{n}_0 \quad (5.8)$$

Insofern gibt die deterministische Kohorten-Komponenten-Methode den Erwartungswert des Bevölkerungsvektors des hier beschriebenen stochastischen Modells wieder, wofür die Einträge von  $\mathbf{A}$  beziehungsweise  $\mathbf{A}_t$ ,  $a_{ij}$ , als Einträge der Leslie-Matrix – also  $F_x$  und  $P_x$  – interpretiert werden (für Details s. 5.2.4).

Soll zusätzlich noch Migration berücksichtigt werden, lässt sich ganz analog zum Kohorten-Komponenten-Modell ein Vektor  $\mathbf{m}_t$  verwenden (vgl. Mode 1971: 83):

$$E(\mathbf{n}_{t+1} | \mathbf{n}_t) = \mathbf{A}_t \mathbf{n}_t + E(\mathbf{m}_t) \quad (5.9)$$

Hierbei korrespondieren die einzelnen Einträge von  $\mathbf{m}_t$  zu einzelnen Typen und erfassen die Nettomigration für diese, wobei  $\mathbf{m}_t \in \mathbb{Z}^k$  gilt. In dieser Formulierung ist  $\mathbf{m}_t$  zufällig, wobei für die weiteren Betrachtungen unerheblich ist, durch was für eine Art von stochastischen Prozess  $\mathbf{m}_t$  gegeben ist. Voraussetzung für (5.9) ist, dass Migration nicht von der aktuellen Bevölkerungsgröße oder -struktur abhängt (vgl. ebd.: 83). Wird Migration deterministisch eingebracht, so dass  $E(\mathbf{m}_t) = \mathbf{m}_t$  gilt, oder aber die mögliche Variabilität von  $\mathbf{m}_t$  nicht berücksichtigt, erhält man die Kohorten-Komponenten-Methode mit Migration (s. Unterabschnitt 3.1.5).

Lässt man Variabilität auf der Individualebene nicht unberücksichtigt, lässt sich zeigen, dass Multityp-Galton-Watson-Prozesse ohne Migration als Kohorten-Komponenten-Modell mit einem additiven, stochastischen Fehlervektor  $\boldsymbol{\epsilon}_{t+1}$  geschrieben werden können (vgl. Deistler & Feichtinger 1974):

$$\mathbf{n}_{t+1} = \mathbf{A}_t \mathbf{n}_t + \boldsymbol{\epsilon}_{t+1} \quad (5.10)$$

### 5.2.3 Mikrovariabilität

Das vorgestellte Modell geht von gegebenen Wahrscheinlichkeiten  $p_i(z_1, \dots, z_k)$  beziehungsweise  $p_{i,t}(z_1, \dots, z_k)$  aus. Variabilität ergibt sich allein auf der Ebene der Individuen, da für diese Ereignisse nicht deterministisch festgelegt sind. Allerdings kann diese Variabilität für große  $|\mathbf{n}_t|$  vernachlässigt werden, weswegen  $\boldsymbol{\epsilon}_t$  in Gleichung (5.10) wegfallen kann. Hierfür lässt sich in Anlehnung an Pollard (1966, 1968) folgende Argumentation nennen. Sei  $n_{(30,w),t} = 100000$ . Berechnet werden soll die Varianz von  $n_{(31,w),t+1}$ . Die hierfür nötige Überlebenswahrscheinlichkeit sei 0.99. Geht man davon aus, dass diese für alle Personen gilt und diese unabhängig voneinander sind, ist  $n_{(31,w),t+1}$  binomialverteilt, womit sich  $E(n_{(31,w),t+1}) = 99000$  und  $\text{Var}(n_{(31,w),t+1}) = 990$  ergibt. Errechnet man hieraus den Variationskoeffizienten erhält man den Wert 0.0003.<sup>124</sup> Somit ist die zufällige

<sup>124</sup> Der Variationskoeffizient kann berechnet werden über

$$\frac{\sqrt{\text{Var}(n_{(31,w),t+1})}}{E(n_{(31,w),t+1})}.$$

Streuung auf der Individualebene bezogen auf den Erwartungswert vernachlässigbar klein (zum gleichen Ergebnis kommt Sykes 1969).<sup>125</sup> Da im weiteren Verlauf von großen Bevölkerungen ausgegangen wird, kann Mikrovariabilität insofern vernachlässigt werden (s. auch Kurtz 2005).

#### 5.2.4 Modellparameter

Der Maximum-Likelihood-Schätzer für die Erwartungswerte  $a_{ij}$  eines Multityp-Galton-Watson-Prozesses ist bei Beobachtung von zwei Zeitpunkten  $t$  und  $t + 1$  gegeben durch (vgl. Maaouia & Touati 2005: 2659):

$$\hat{a}_{ij} = \frac{n_{i,t}^{j,t+1}}{n_{i,t}} \quad (5.11)$$

Hierbei ist  $n_{i,t}^{j,t+1}$  die beobachtete Zahl der Individuen vom Typus  $j$  zu  $t + 1$ , die von Individuen des Typus  $i$  „erzeugt“ werden (vgl. ebd.: 2658). Dies entspricht dem Maximum-Likelihood-Schätzer für die Einträge der Leslie-Matrix, der auf Daten für identifizierte Individuen basiert (vgl. Caswell 2001: 134ff.).<sup>126</sup> Dieser Schätzer setzt voraus, dass für beide Zeitpunkte  $t$  und  $t + 1$  Mikrodaten vorliegen, aus denen für die Individuen zu  $t + 1$  hervorgeht, zu welchen Individuen zu  $t$  sie zugeordnet werden können. Je nach  $i$  und  $j$  werden die Schätzer für  $\hat{a}_{ij}$  dann entweder als  $P_x$  oder  $F_x$  verwendet.

Eine andere Möglichkeit besteht in der Verwendung von Sterbe- und Geburtenraten, aus denen dann die Parameter  $a_{ij}$  abgeleitet werden müssen. Das Problem hierbei ist, dass Sterbe- und Geburtenraten für exakte Alter  $x \in \mathbb{N}$  beziehungsweise Altersintervalle  $[x, x + 1)$  erstellt werden (vgl. Keyfitz & Caswell 2005: 56). Individuen in einem Zustand  $(a, g)$  weisen aber nicht unbedingt das exakte Alter  $a$  auf, sondern sind zwischen  $a$  und  $a + 1$  Jahren alt, weshalb beispielsweise Überlebenswahrscheinlichkeiten aus Sterbetafeln nicht direkt verwendet werden können (vgl. Caswell 2001: 23).

Um  $P_x$  jeweils für Männer und Frauen aus geschlechtsspezifischen Sterbetafelfunktionen zu gewinnen, schlagen Keyfitz & Caswell (2005: 56) folgende Schätzung vor:

$$P_x = \frac{\int_{x+1}^{x+2} l_i di}{\int_x^{x+1} l_i di} \quad (5.12)$$

Die Idee hierbei ist, dass Individuen, die zum Zeitpunkt  $t$  in Altersklasse  $x$  sind, ein exaktes Alter zwischen  $x$  und  $x + 1$  aufweisen (vgl. ebd.: 56). Überleben diese Personen eine Zeiteinheit, befinden sie sich im Altersbereich von  $x + 1$  bis  $x + 2$  (vgl. ebd.: 56). Ist  $l_i$  über

<sup>125</sup> Zwar kann der Variationskoeffizient für kleine  $n_{(x,g),t}$  auch vergleichsweise hoch ausfallen. Allerdings beeinflussen diese dann Größe und Struktur der Gesamtbevölkerung nur geringfügig.

<sup>126</sup> Werden keine Geburten berücksichtigt und nur Übergänge von Individuen zwischen Zuständen betrachtet, ist dieser Schätzer ebenfalls zum Maximum-Likelihood-Schätzer für Übergangswahrscheinlichkeiten von Markov-Ketten äquivalent (vgl. Anderson & Goodman 1957).

das gesamte Intervall  $[x, x + 2)$  bekannt, kann für alle  $i \in [x, x + 1)$  die Überlebenswahrscheinlichkeit  $l_{i+1}/l_i$  berechnet und anschließend über alle  $i$  das arithmetische Mittel der Überlebenswahrscheinlichkeiten gebildet werden. Da aber  $l_i$  in aller Regel nur für  $x, x + 1$  und  $x + 2$  bekannt ist, wird (5.12) geschätzt über (vgl. ebd.: 56)

$$P_x \approx \frac{l_{x+1} + l_{x+2}}{l_x + l_{x+1}} \quad (5.13)$$

Die rechte Seite von (5.13) lässt sich umformulieren zu

$$\frac{l_{x+1}}{l_x} \cdot \frac{1}{1 + \frac{l_{x+1}}{l_x}} + \frac{l_{x+2}}{l_{x+1}} \cdot \frac{1}{1 + \frac{l_x}{l_{x+1}}} \quad (5.14)$$

Da  $l_{x+1}/l_x = p_x$  (s. bspw. Chiang 1984: 205) erhält man durch einsetzen<sup>127</sup>

$$p_x \frac{1}{1 + p_x} + p_{x+1} \frac{1}{1 + p_x^{-1}} \quad (5.15)$$

Nimmt man an, dass  $p_x \approx 1$ , oder dass  $p_x \approx p_{x+1}$ , lässt sich  $P_x$  schätzen über

$$P_x \approx \frac{1}{2} (p_x + p_{x+1}) \quad (5.16)$$

Diese Variante wird ebenfalls von Keyfitz & Caswell (2005: 56) angegeben. Die numerischen Unterschiede, die sich bei der Schätzung von  $P_x$  durch die beiden Varianten ergeben, sind in der Regel marginal (vgl. ebd.: 56).

Für  $F_x^g$  geben Keyfitz & Caswell (2005: 57) folgenden Ausdruck an:

$$\frac{1}{2} (f_x^g + P_x f_{x+1}^g) \quad (5.17)$$

Hierbei ist  $f_x^g$  die altersspezifische Geburtenrate für das Alter  $x$ , bei der lediglich Kinder mit dem Geschlecht  $g$  betrachtet werden. Diese Formel berücksichtigt, dass Personen, die zu  $t$  im Altersintervall  $[x, x + 1)$  sind, zwischen  $t$  und  $t + 1$  in das Altersintervall  $[x + 1, x + 2)$  wechseln. Ferner wird durch  $P_x$  berücksichtigt, dass ein Wechsel der Altersklassen nicht unbedingt stattfinden muss, sondern Personen auch sterben können. Wenn die Geburtenraten  $f_x^g$  ab dem Alter 15 vorliegen, kann bei dieser Formulierung  $F_{14}^g > 0$  sein.

Zusätzlich schlagen Keyfitz & Caswell (2005: 57) vor, (5.17) noch mit  $l_{(0.5,g)}$  zu gewichten, womit man

$$F_x^g \approx l_{(0.5,g)} \frac{1}{2} (f_x^g + P_x f_{x+1}^g) \quad (5.18)$$

erhält. Die zugrunde liegende Idee ist, dass zwischen  $t$  und  $t + 1$  geborene Individuen zunächst noch bis  $t + 1$  überleben müssen. Unter der Annahme, dass Geburten über das Intervall  $t$  bis  $t + 1$  gleichmäßig verteilt sind, muss das „durchschnittliche“ neugeborene Individuum bis zum Alter 0.5 überleben (vgl. ebd.: 57).  $l_{(0.5,g)}$  wiederum kann entweder über lineare Interpolation

<sup>127</sup>  $p_x$  ist hier Sterbetafel-funktion und keine Wahrscheinlichkeit des Multityp-Prozesses wie bspw. in Formel (5.1).

oder aber bei vergleichsweise hoher neonataler Mortalität über logarithmische Interpolation geschätzt werden, wobei letztere zu folgendem Ausdruck führt (vgl. ebd.: 58):

$$l_{(0.5,g)} \approx l_{(0,g)} \sqrt{l_{(1,g)}} = \sqrt{l_{(1,g)}} \quad (5.19)$$

Ist  $l_{(0,g)} = 1$ , gilt  $l_{(1,g)} = p_{(0,g)}$ .<sup>128</sup> Die Schätzung von  $F_x^g$  wird im nächsten Abschnitt nochmal aufgegriffen.

Für die Einträge  $m_{i,t}$  des Vektors der Nettomigration  $\mathbf{m}_t$  finden sich in der Literatur kaum Vorschläge, wie diese in einem formal strengen Sinn hergeleitet werden können. Das in der vorliegenden Arbeit gewählte Vorgehen wird in Abschnitt 8.6.1 beschrieben.

### 5.2.5 Stochastische Modellparameter

Um demographische Unsicherheit zu berücksichtigen, können die Wahrscheinlichkeiten  $p_i(z_1, \dots, z_k)$  als stochastisch aufgefasst werden. Die Idee, diese Wahrscheinlichkeiten und hierüber die daraus resultierenden Matrixeinträge  $a_{ij}$  als Zufallsvariablen aufzufassen, geht auf Pollard (1968) zurück. Sykes (1969) ging von einer Leslie-Matrix mit stochastischen Einträgen aus. In beiden Fällen wird angenommen, dass die Einträge  $a_{ij}$  von  $\mathbf{A}$  nicht fix gegeben sind, sondern variieren können, so dass ein inhomogener Prozess vorliegt. Dabei hängen die Werte  $a_{ij}$  in irgendeiner Art und Weise von einem oder mehreren stochastischen Prozessen ab. Sie können sich beispielsweise aus der statistischen Modellierung der demographischen Komponenten in Kombination mit der im vorausgegangenen Unterabschnitt diskutierten Herleitung der Einträge von  $\mathbf{A}$  über demographische Raten ergeben. Solche stochastischen Parameter werden im weiteren mit  $\tilde{a}_{ij}$  und entsprechende Matrizen mit  $\tilde{\mathbf{A}}$  bezeichnet.

Für eine konkrete Folge von Realisierungen von  $\tilde{\mathbf{A}}_t$  gilt noch immer Gleichung (5.6):

$$\mathbf{n}_{t+1} = \tilde{\mathbf{A}}_t \mathbf{n}_t \quad (5.20)$$

Ist keine konkrete Realisierung gegeben, kann der Erwartungswert von  $\mathbf{n}_{t+1}$  mit Bezug auf die stochastischen Parameter betrachtet werden. Geht man weiterhin wie oben angegeben davon aus, dass nur  $\mathbf{n}_0$  fix gegeben ist, lässt sich der Erwartungswert von  $\mathbf{n}_{t+1}$  schreiben als (vgl. Lee & Tuljapurkar 1994: 1178)

$$\mathbb{E}(\mathbf{n}_{t+1}) = \mathbb{E} \left[ \left( \prod_{x=0}^t \tilde{\mathbf{A}}_x \right) \mathbf{n}_0 \right] \quad (5.21)$$

Daneben kann man sich für weitere Größen und Kennwerte interessieren, wie beispielsweise  $\mathbb{E}(|\mathbf{n}_{t+1}|)$ ,  $\text{Var}(|\mathbf{n}_{t+1}|)$  oder Vorausberechnungsintervalle für  $|\mathbf{n}_{t+1}|$ .

Für diesen Ausdrücke lassen sich nicht ohne weiteres einfache, analytische Lösungen finden (vgl. Tuljapurkar 1992: 388). Zwar existieren Approximationen (vgl. bspw. Alho & Spencer 1991; Lee & Tuljapurkar 1994), allerdings basieren diese auf stark vereinfachenden Annahmen und sind nur mit großem Aufwand zu berechnen, weshalb wie in Unterabschnitt

<sup>128</sup> Für  $p_{(0,g)} \geq 0.95$ , also Werte, wie sie etwa seit den fünfziger Jahren für die BRD vorliegen, sind die Abweichungen zwischen linearer und logarithmischer Interpolation marginal.

5.1.3 bereits angedeutet in aller Regel auf Monte-Carlo-Simulationen zurückgegriffen wird (vgl. Alho & Spencer 2005: 286f.). Bei diesen werden entsprechend eines stochastischen Modells viele zufällige Folgen von Matrizen  $\tilde{\mathbf{A}}_t$  erzeugt und für jede einzelne Folge (5.6) angewendet. Aus den Ergebnissen lassen sich dann die oben genannten Kennwerte herleiten. Hat man beispielsweise 100 Folgen von Modellparametern erzeugt, um eine Vorausberechnung ausgehend von  $t$  zu  $t + 10$  vorzunehmen, kann für jede dieser Folgen  $\mathbf{n}_{t+10}$  berechnet werden. Aus diesen 100 Ergebnissen kann dann zum Beispiel die durchschnittliche Bevölkerungsgröße zu  $t + 10$  hergeleitet werden.

### 5.3 Lineare Verwandtschaft im diskreten Multityp-Galton-Watson-Prozess

Ausgehend von dem im letzten Abschnitt vorgestellten Modell des diskreten Multityp-Galton-Watson-Prozesses soll hier auf die Berechnung von linearer Verwandtschaft im Rahmen dieses Modells eingegangen werden.<sup>129</sup> Dabei wird davon ausgegangen, dass von einer im weiteren noch zu erläuternden Ausnahme lediglich Sterbe- und Geburtenraten sowie Bevölkerungsbestände bekannt sind.

#### 5.3.1 Vorüberlegungen

Angenommen zu Zeitpunkt  $t$  wird eine Population betrachtet, die aus nur einer Frau in Altersklasse  $x$  besteht. Gesucht wird die Zahl der weiblichen Personen in Altersklasse 0 zu  $t + 1$ . Der Erwartungswert dieser Größe ist entsprechend der Überlegungen im vorausgegangenen Abschnitt gegeben über

$$E(n_{(0,w),t+1} | \mathbf{n}_t = \mathbf{e}_{(x,w)}) = a_{(x,w),(0,w)} = \frac{\delta f_{(x,w)}(1, \dots, 1)}{\delta s_{(0,w)}} = F_x^w \quad (5.22)$$

Wenn  $x$  die erste Altersklasse ist, für die  $F_x^w > 0$  ist, dann kann das Ergebnis der obigen Formel als der Erwartungswert der Zahl der lebenden Töchter der betrachteten Frau interpretiert werden. Hierbei wird bei Verwendung von (5.18) berücksichtigt, dass die Töchter zwischen  $t$  und  $t + 1$  geboren werden und dann bis  $t + 1$  überleben müssen.

Allerdings wird nicht darauf geachtet, dass die betrachtete Frau selbst zu  $t + 1$  noch lebt. Formel (5.18), die zur Schätzung von  $F_x^g$  verwendet werden kann, lässt sich so formulieren, dass dies direkt ersichtlich wird:

$$F_x^w \approx l_{(0.5,w)} \frac{1}{2} (f_x^w + f_{x+1}^w) P_x + l_{(0.5,w)} \frac{1}{2} f_x^w (1 - P_x) \quad (5.23)$$

Der erste Term auf der rechten Seite gilt für Frauen, die von  $t$  bis  $t + 1$  überleben, und der zweite Term entsprechend für Frauen, die zwischen  $t$  und  $t + 1$  sterben, wobei durch

<sup>129</sup> Die Herleitung einiger Ausdrücke aus der Arbeit von Goodman et al. (1974) im Rahmen eines stetigen Multityp-Prozesses mit fixen Wahrscheinlichkeiten findet sich bei Jagers (1982) sowie Joffe & Waugh (1985, 1986).

$P_x$  beziehungsweise  $(1 - P_x)$  die jeweilige relative Häufigkeit der entsprechenden Gruppe berücksichtigt wird.

Eine analoge Formulierung lässt sich für den ML-Schätzer finden:

$$\frac{n_{(x,w),t}^{(0,w),t+1}}{n_{(x,w),t}} = \frac{n_{(x,w,+),t}^{(0,w),t+1}}{n_{(x,w,+),t}} \frac{n_{(x,w,+),t}}{n_{(x,w),t}} + \frac{n_{(x,w,-),t}^{(0,w),t+1}}{n_{(x,w,-),t}} \frac{n_{(x,w,-),t}}{n_{(x,w),t}} \quad (5.24)$$

Hier wird genutzt, dass  $n_{(x,w),t} = n_{(x,w,+),t} + n_{(x,w,-),t}$  ist, wobei  $n_{(x,w,+),t}$  die Frauen erfasst, die zu  $t$  in Altersklasse  $x$  sind und bis zu  $t + 1$  und Altersklasse  $x + 1$  überleben.  $n_{(x,w,-),t}$  erfasst entsprechend Frauen, die zu  $t$  in Altersklasse  $x$  sind und zwischen  $t$  und  $t + 1$  sterben.<sup>130</sup>  $n_{(x,w,+),t}^{(0,w),t+1}$  und  $n_{(x,w,-),t}^{(0,w),t+1}$  erfassen die weiblichen Nachfahren, die jeweils von diesen beiden Frauengruppen zwischen  $t$  und  $t + 1$  geboren werden und bis  $t + 1$  überleben. Da bei Nicht-Berücksichtigung von Migration  $n_{(x,w,+),t} = n_{(x,w),t}^{(x+1,w),t+1}$  ist, lässt sich (5.24) schreiben als

$$\frac{n_{(x,w),t}^{(0,w),t+1}}{n_{(x,w),t}} = \frac{n_{(x,w,+),t}^{(0,w),t+1}}{n_{(x,w,+),t}} P_x + \frac{n_{(x,w,-),t}^{(0,w),t+1}}{n_{(x,w,-),t}} (1 - P_x) \quad (5.25)$$

Die Berücksichtigung von gestorbenen Frauen bei (5.23) beziehungsweise (5.25) ist ungünstig, da dies zu einer hypothetischen Betrachtung führt: wie viele lebende Töchter haben lebende Frauen in Altersklasse  $x + 1$  und bis zu Altersklasse  $x + 1$  gestorbene Frauen, wenn sie noch leben würden? Insofern scheint es in Hinblick auf die Fragestellung der vorliegenden Arbeit angebracht, gestorbene Frauen von der Betrachtung auszuschließen und folgenden Ausdruck zu verwenden:

$$a_{(x,w,+), (0,w)} = \frac{n_{(x,w,+),t}^{(0,w),t+1}}{n_{(x,w,+),t}} \quad (5.26)$$

Hierüber kann

$$E(n_{(0,w),t+1} | \mathbf{n}_t = \mathbf{e}_{(x,w,+)}) = a_{(x,w,+), (0,w)} \quad (5.27)$$

berechnet werden. Somit wird in der Bedingung des Erwartungswertes explizit berücksichtigt, dass die betrachtete Frau bis zu  $t + 1$  überlebt. Es stellt sich dann allerdings die Frage, wie und unter welchen Bedingungen (5.26) beziehungsweise (5.27) über demographische Raten geschätzt werden kann. Und auch für den allgemeinen ML-Schätzer (5.24) stellt sich die Frage nach der Äquivalenz zu (5.23).

Unter den Annahmen, dass

1. die Wahrscheinlichkeit für Mehrlingsgeburten unabhängig vom späteren Überleben der Mutter und zwischen  $t$  und  $t + 1$  über das gesamte Intervall  $[x, x + 1)$  identisch ist,
2. die Geburtenrate  $f_{x,t}^w$  von  $t$  bis  $t + 1$  für das gesamte Intervall  $[x, x + 1)$  gilt und unabhängig von Sterblichkeit ist

<sup>130</sup> Die Unterteilung in überlebende und sterbende Frauen läuft auf eine Erweiterung des Zustandsraums auf  $A \times G \times \{+, -\}$  hinaus.

3. Ereignisse durchschnittlich in der Mitte des Zeitintervalls  $t$  bis  $t + 1$  stattfinden lässt sich zeigen, dass

$$\begin{aligned} a_{(x,w,+),(0,w)} &= \frac{1}{2} l_{(0.5,w),t} (f_{x,t}^w + f_{x+1,t}^w) \\ &= F_{x,t}^{w*} \end{aligned} \quad (5.28)$$

gilt. Zudem entspricht unter diesen Bedingungen  $F_{x,t}^w$  nun

$$l_{(0.5,w),t} \frac{1}{2} (f_{x,t}^w + f_{x+1,t}^w) P_x + l_{(0.5,w),t} \left( \frac{3}{8} f_{x,t}^w + \frac{1}{8} f_{x+1,t}^w \right) (1 - P_{x,t}) \quad (5.29)$$

Allerdings unterscheiden sich mit (5.28) und (5.29) berechnete Ergebnisse praktisch nicht von über (5.18) ermittelte Resultate, weshalb es letztlich unerheblich ist, welche der Formeln genutzt wird, und die formale Herleitung und ein Vergleich der Varianten im Anhang B zu finden ist. Im weiteren werden aber nichtsdestotrotz (5.28) und (5.29) verwendet, um zu verdeutlichen, ob das Überleben der Person, die betrachtet wird, vorausgesetzt wird oder nicht.

### 5.3.2 Direkte Nachfahren von Frauen

#### 5.3.2.1 Kinder

Der Einfachheit halber werden bei der Betrachtung von Verwandtschaft zunächst fixe Modellparameter und somit ein homogener Prozess angenommen. Zeitveränderliche und stochastische Modellparameter werden in Unterabschnitt 5.3.6 eingeführt. Ebenfalls vorerst unberücksichtigt bleibt Migration, deren möglicher Einfluss auf die im weiteren zu entwickelnden Ausdrücke in Unterabschnitt C.5 besprochen wird (s. auch die Diskussion am Ende dieses Kapitels).

Betrachtet wird eine Frau in einer Altersklasse  $a$ . Angegeben werden soll die durchschnittliche Zahl an jemals geborenen sowie lebenden Söhnen und Töchtern. In allen genannten Fällen muss zunächst betrachtet werden, wie viele Nachfahren eines Geschlechtes diese Frau geboren hat. Angenommen die reproduktive Phase beginnt ab dem exakten Alter 15, die betrachtete Frau war zu  $t$  in der Altersklasse  $\alpha = 14$  und  $a$  sei nun gleich  $\alpha + 1$ , womit  $t + 1$  betrachtet wird. Dann ist der Erwartungswert der Anzahl der lebenden Töchter gleich<sup>131</sup>

$$\begin{aligned} E(n_{(0,w),t+1} | \mathbf{n}_t = \mathbf{e}_{(\alpha,w,+)}) &= a_{(\alpha,w,+),(0,w)} \\ &= \frac{\delta f_{(\alpha,w,+)}(1, \dots, 1)}{\delta S_{(0,w)}} \\ &= F_{\alpha}^{w*} \end{aligned} \quad (5.30)$$

Hierbei wird durch die Schätzung von  $F_x^{w*}$  über (5.28) berücksichtigt, dass diese Töchter zwischen  $t$  und  $t + 1$  geboren wurden und bis  $t + 1$  überleben müssen. Soll die Zahl der jemals geborenen Töchter ermittelt werden, kann  $\frac{1}{2}(f_{\alpha}^w + f_{\alpha+1}^w)$  anstelle von  $F_{\alpha}^{w*}$  verwendet werden.

131 Im weiteren wird anstelle der Formulierung „Erwartungswert der Anzahl“ vereinfacht auch von der „erwarteten Anzahl“ gesprochen.

Ist  $a = \alpha + 2$  wird  $t + 2$  betrachtet und es gilt

$$\begin{aligned} \mathbb{E}(n_{(0,w),t+1} | \mathbf{n}_t = \mathbf{e}_{(\alpha,w,+)} ) &+ \mathbb{E}(n_{(0,w),t+2} | \mathbf{n}_{t+1} = \mathbf{e}_{(\alpha+1,w,+)} ) \\ &= a_{(\alpha,w,+),(0,w)} + a_{(\alpha+1,w,+),(0,w)} \\ &= F_{\alpha}^{w*} + F_{\alpha+1}^{w*} \end{aligned} \quad (5.31)$$

aufgrund der Markov-Eigenschaft.

Für eine beliebige Altersklasse  $a$  lässt sich auf analoge Weise folgender Ausdruck herleiten:

$$\sum_{x=0}^{a-\alpha-1} \mathbb{E}(n_{(0,w),t+x+1} | \mathbf{n}_{t+x} = \mathbf{e}_{(\alpha+x,w,+)} ) = \sum_{x=\alpha}^{a-1} F_x^{w*} \quad (5.32)$$

Hierbei wird nur die Mortalität für das erste halbe Lebensjahr der Töchter berücksichtigt. Um die Zahl der jemals geborenen oder der lebenden Töchter zu berechnen, sind noch Modifikationen notwendig. Soll die Zahl der jemals geborenen Töchter betrachtet werden, kann

$$\frac{1}{l_{(0.5,w)}} \sum_{x=\alpha}^{a-1} F_x^{w*} = \sum_{x=\alpha}^{a-1} \frac{1}{2} (f_x^w + f_{x+1}^w) \quad (5.33)$$

benutzt werden. Dies ist der linearen Approximation des stetigen Ansatzes von Goodman et al. (1974) von Dudel (2008) nicht unähnlich. Für die Zahl an jemals geborenen Söhnen ergibt sich analog zu jemals geborenen Töchtern

$$\sum_{x=\alpha}^{a-1} \frac{1}{2} (f_x^m + f_{x+1}^m) \quad (5.34)$$

Geht man davon aus, dass es für ein Alter  $x$  allgemeine Wahrscheinlichkeiten für  $0, 1, \dots$  Geburten gibt und bei diesen das Geschlecht der Kinder mit bestimmten Wahrscheinlichkeiten zufällig ausfällt, lässt sich die erwartete Zahl der jemals geborenen Kinder schreiben als<sup>132</sup>

$$\sum_{x=\alpha}^{a-1} \frac{1}{2} (f_x^w + f_x^m + f_{x+1}^w + f_{x+1}^m) \quad (5.35)$$

---

132 Sei  $\Pr(k)$  die Wahrscheinlichkeit für genau  $k$  Kinder und  $K$  die Zahl der Kinder insgesamt. Sei  $\Pr(w)$  die Wahrscheinlichkeit, dass ein Neugeborenes weiblichen Geschlechts ist und  $\Pr(m) = 1 - \Pr(w)$  die Wahrscheinlichkeit für das männliche Geschlecht. Sei  $K_w$  die Zahl der Töchter und  $K_m$  die Zahl der Söhne. Dann gilt:

$$\mathbb{E}(K_w) = \sum_k k \Pr(k) \Pr(w)$$

$$\mathbb{E}(K_m) = \sum_k k \Pr(k) \Pr(m)$$

Womit:

$$\mathbb{E}(K_w) + \mathbb{E}(K_m) = \mathbb{E}(K).$$

Für die Zahl der lebenden Töchter einer Frau in Altersklasse  $a$  muss neben der Fertilität dieser Frau noch das Überleben der Töchter berücksichtigt werden. Geht man davon aus, dass  $a = \alpha + 1$  ist, dann ist die erwartete Zahl an lebenden Töchtern wie bereits gezeigt gleich

$$\begin{aligned} \mathbb{E} \left( n_{(0,w),t+1} | \mathbf{n}_t = \mathbf{e}_{(\alpha,w,+)} \right) &= a_{(\alpha,w,+),(0,w)} \\ &= \frac{\delta f_{(\alpha,w,+)}(1, \dots, 1)}{\delta s_{(0,w)}} \\ &= F_{\alpha}^{w*} \end{aligned} \quad (5.36)$$

Ist  $a = \alpha + 2$  und werden zunächst nur Töchter betrachtet, bei deren Geburt die Mutter in Altersklasse  $\alpha$  war, gilt aufgrund der Markov-Eigenschaft, der Unabhängigkeit und der Homogenität von Personen

$$\begin{aligned} \mathbb{E} \left( n_{(1,w),t+2} | \mathbf{n}_t = \mathbf{e}_{(\alpha,w,+)} \right) &= \mathbb{E} \left( n_{(0,w),t+1} | \mathbf{n}_t = \mathbf{e}_{(\alpha,w,+)} \right) \\ &\cdot \mathbb{E} \left( n_{(1,w),t+2} | \mathbf{n}_t = \mathbf{e}_{(0,w)} \right) \end{aligned} \quad (5.37)$$

was

$$a_{(\alpha,w,+),(0,w)} a_{(0,w),(1,w)} = F_{\alpha}^{w*} P_{(0,w)} \quad (5.38)$$

entspricht.

Sei

$$P_x^* = \begin{cases} 1 & \text{wenn } x = 0 \\ \prod_{y=0}^{x-1} P_y & \text{wenn } 0 < x \leq \rho \\ 0 & \text{wenn } x > \rho \end{cases} \quad (5.39)$$

die Wahrscheinlichkeit, von Altersklasse 0 bis zur Altersklasse  $x$  zu überleben, wobei  $P_0^* = 1$  gesetzt wird.  $P_{(x,w)}^*$  bezeichnet die entsprechende Wahrscheinlichkeit für Frauen und  $P_{(x,m)}^*$  für Männer. Dann lässt sich für eine beliebige Altersklasse  $a > \alpha$  der Mutter und Töchter, die von der Mutter in Altersklasse  $\alpha$  geboren wurden, schreiben

$$\mathbb{E} \left( n_{(a-\alpha-1,w),t+a-\alpha} | \mathbf{n}_t = \mathbf{e}_{(\alpha,w,+)} \right) = F_{\alpha}^{w*} P_{(a-\alpha-1,w)}^* \quad (5.40)$$

Für Töchter, die von der Mutter in einem Alter  $x$ ,  $a > x > \alpha$ , geboren wurden, lassen sich auf analoge Weise Ausdrücke herleiten, wobei  $\alpha$  in (5.40) durch  $x$  ersetzt wird.

Unter den oben genannten gilt allgemein für die Zahl lebender Töchter einer Frau in Altersklasse  $a > \alpha$

$$\begin{aligned} &\sum_{x=0}^{a-\alpha-1} \mathbb{E} \left( n_{(a-\alpha-x-1,w),t+a-\alpha} | \mathbf{n}_{t+x} = \mathbf{e}_{(\alpha+x,w,+)} \right) \\ &= \sum_{x=\alpha}^{a-1} F_x^{w*} P_{(a-x-1,w)}^* \end{aligned} \quad (5.41)$$

Für Söhne ergibt sich mit einer analogen Argumentation

$$\begin{aligned}
& \sum_{x=0}^{a-\alpha-1} \mathbb{E} \left( n_{(a-\alpha-x-1,m),t+a-\alpha} | \mathbf{n}_{t+x} = \mathbf{e}_{(\alpha+x,w,+)} \right) \\
&= \sum_{x=\alpha}^{a-1} F_x^{m*} P_{(a-x-1,m)}^* \tag{5.42}
\end{aligned}$$

Für Kinder insgesamt ergibt sich unter der oben gemachten Annahmen der fixen Wahrscheinlichkeiten für Geschlechter von Nachfahren und Unabhängigkeit von Personen formal

$$\begin{aligned}
& \sum_{x=0}^{a-\alpha-1} \mathbb{E} \left( n_{(a-\alpha-x-1,w),t+a-\alpha} | \mathbf{n}_{t+x} = \mathbf{e}_{(\alpha+x,w,+)} \right) \\
&+ \sum_{x=0}^{a-\alpha-1} \mathbb{E} \left( n_{(a-\alpha-x-1,m),t+a-\alpha} | \mathbf{n}_{t+x} = \mathbf{e}_{(\alpha+x,w,+)} \right) \tag{5.43}
\end{aligned}$$

was geschrieben werden kann als

$$\sum_{x=\alpha}^{a-1} \left( F_x^{w*} P_{(a-x-1,w)}^* + F_x^{m*} P_{(a-x-1,m)}^* \right) \tag{5.44}$$

Wie in Paragraph 5.2.1.2 diskutiert, erlaubt der Multityp-Galton-Watson-Prozess eigentlich keine Identifikation von Individuen und somit eigentlich auch keine Zuordnung von Kindern zu Müttern. Aufgrund der Modellannahmen und durch die Verwendung von den in diesem Unterabschnitt vorgestellten Bedingungen der Erwartungswerte kann dieses Problem aber umgangen werden (s. auch Fußnote 121 auf Seite 102). Die Bedingungen der Erwartungswerte stellen formal sicher, dass die betrachtete Frau im Alter  $a$  noch lebt und niemand anderes als diese Frau zur Zahl der Kinder beiträgt. Dies wird der Einfachheit halber im weiteren teilweise vorausgesetzt. Ist aus dem Kontext ersichtlich, worauf Bezug genommen wird, wird auf die Angabe des Erwartungswertes und entsprechender Bedingungen über  $\mathbb{E}(\cdot|\cdot)$  verzichtet.

### 5.3.2.2 Enkelkinder

Für die erwartete Zahl an jemals geborenen und lebenden Enkelkindern ist die Zahl der Söhne und Töchter, deren Überleben und deren Fertilität entscheidend. Dies läuft darauf hinaus, dass im vorherigen Unterabschnitt entwickelte Schema zur Berechnung der Zahl der Kinder gewissermaßen doppelt anzuwenden – einmal auf Ego und anschließend mit kleineren Veränderungen auf Egos Kinder.

Ausgegangen wird zunächst von einem weiblichen Ego, dass zwischen  $t$  und  $t + 1$  in der Altersklasse  $x$  eine Tochter geboren hat und sich nun in der Altersklasse  $a$  befindet, womit der Zeitpunkt  $t + a - x$  betrachtet wird. Wenn die Tochter bis zu diesem Zeitpunkt überlebt, befindet sie sich in der Altersklasse  $a - x - 1$ , wobei ihr Überleben aber nicht vorausgesetzt

wird. Die erwartete Zahl an jemals geborenen Töchtern dieser einen Tochter ist gleich

$$\begin{aligned}
& \sum_{y=\alpha}^{a-x-2} \mathbb{E} \left( n_{(0,w),t+y+2} \mid \mathbf{n}_{t+y+1} = \mathbf{e}_{(y,w)} \right) \\
& \cdot \mathbb{E} \left( n_{(y,w),t+y+1} \mid \mathbf{n}_{t+1} = \mathbf{e}_{(0,w)} \right) \\
& = \sum_{y=\alpha}^{a-x-2} P_{(y,w)}^* F_y^w
\end{aligned} \tag{5.45}$$

beziehungsweise wenn Mortalität in der Altersklasse 0 nicht beachtet wird

$$\sum_{y=\alpha}^{a-x-2} P_{(y,w)}^* \frac{1}{l_{(0.5,w)}} F_y^w \tag{5.46}$$

Hierbei wird zum einen ganz analog zur erwarteten Zahl der Töchter die Fertilität berücksichtigt. Zudem wird beachtet, dass das Überleben der Tochter bis zu den einzelnen Altersklassen  $y$ , die sie bis zum Alter  $a - x - 1$  durchläuft, nicht gesichert ist.

Allerdings können in einem Alter  $x$  der betrachteten Frau auch keine oder mehrere Töchter geboren werden und Fertilität ist nicht nur auf ein Alter  $x$  beschränkt. Nimmt man an, dass einzelne Töchter voneinander und auch von Ego unabhängig sind, kann (5.45) mit dem Erwartungswert für  $n_{(0,w),t+1}$  gewichtet werden. Dann lässt sich

$$\begin{aligned}
& \sum_{x=0}^{a-\alpha-1} \mathbb{E} \left( n_{(0,w),t+x+1} \mid \mathbf{n}_{t+x} = \mathbf{e}_{(\alpha+x,w,+)} \right) \\
& \cdot \sum_{y=\alpha}^{a-x-2} \mathbb{E} \left( n_{(0,w),t+x+y+2} \mid \mathbf{n}_{t+x+1+y} = \mathbf{e}_{(y,w)} \right) \\
& \cdot \mathbb{E} \left( n_{(y,w),t+x+y+1} \mid \mathbf{n}_{t+x+1} = \mathbf{e}_{(0,w)} \right) \\
& = \sum_{x=\alpha}^{a-1} F_x^{w*} \sum_{y=\alpha}^{a-x-2} P_{(y,w)}^* F_y^w
\end{aligned} \tag{5.47}$$

schreiben. Soll Mortalität in der Altersklasse 0 abermals nicht beachtet und somit die Zahl der jemals geborenen Enkeltöchter über Töchter betrachtet werden, ergibt sich analog zu (5.46)

$$\sum_{x=\alpha}^{a-1} F_x^{w*} \sum_{y=\alpha}^{a-x-2} P_{(y,w)}^* \frac{1}{l_{(0.5,w)}} F_y^w \tag{5.48}$$

Berücksichtigt man hingegen das volle Sterberisiko der Enkeltöchter und somit, dass sie überleben müssen bis Ego in Altersklasse  $a$  ist, erhält man die Zahl der lebenden Enkeltöchter über Töchter über

$$\sum_{x=\alpha}^{a-1} F_x^{w*} \sum_{y=\alpha}^{a-x-2} P_{(y,w)}^* F_y^w P_{(a-x-y-2,w)}^* \tag{5.49}$$

Die erwartete Zahl der lebenden Enkelsöhne über Töchter wird analog hergeleitet und ist gleich

$$\sum_{x=\alpha}^{a-1} F_x^{w*} \sum_{y=\alpha}^{a-x-2} P_{(y,w)}^* F_y^m P_{(a-x-y-2,m)}^* \quad (5.50)$$

Möchte man die erwartete Zahl der Enkeltöchter und Enkelsöhne über Söhne berechnen, werden auch für Männer um Säuglingssterblichkeit korrigierte Geburtenraten benötigt. Ansonsten kann analog zu Enkelkindern über Töchter vorgegangen werden.

### 5.3.3 Direkte Nachfahren von Männern

Entsprechend der Ausführungen in Unterabschnitt 5.2.1.2 wurde bisher davon ausgegangen, dass Männern keine Geburten zugeordnet werden können. Diese Annahme soll hier aufgehoben werden. Genauer sollen Geburtenraten für Männer berechnet werden, so dass die bisher vorgestellten Formeln für diese verwendet werden können. Die Zahl der Geburten ergibt sich aber weiterhin lediglich aus der Zahl der Frauen und deren Fertilität, die für die Bevölkerungsfortschreibung weiterhin als maßgeblich angenommen werden. Fertilität von Männern wird also nur zur Berechnung von Verwandtschaft verwendet.

Es wird davon ausgegangen, dass für Geburten, die Frauen im Altersintervall  $[x, x + 1)$  zugeordnet werden können, die Verteilung der Alter  $[y, y + 1)$  der zu diesen Geburten gehörenden Väter bekannt ist. Somit sind also bedingte Anteilswerte  $\phi(y|x)$  bekannt, so dass

$$\sum_y \phi(y|x) n_{(x,w),t}^{(0,g)} = n_{(x,w),t}^{(0,g)} \quad (5.51)$$

wobei  $n_{(x,w),t}^{(0,g)}$  die Zahl von Geburten zwischen  $t$  und  $t + 1$  von Frauen im Altersintervall  $[x, x + 1)$  erfasst, unabhängig davon in welcher Altersklasse diese Frauen zu  $t$  waren. Hieraus lässt sich die Zahl der Geburten von Männern im Altersintervall  $[y, y + 1)$  ermitteln als

$$\sum_x \phi(y|x) n_{(x,w),t}^{(0,g)} = n_{(y,m),t}^{(0,g)} \quad (5.52)$$

Dann können die zur Berechnung von  $F_{(y,m)}^g$  und  $F_{(y,m)}^{g*}$  benötigten Geburtenraten  $f_{(y,m)}^g$  ermittelt werden über<sup>133</sup>

$$f_{(y,m)}^g = \frac{n_{(y,m),t}^{(0,g)}}{J_{(y,m),t}} \quad (5.53)$$

wobei  $J_{(y,m),t}$  für die zwischen  $t$  und  $t + 1$  von Männern im Altersintervall  $[y, y + 1)$  gelebten Personenjahre steht. Entsprechend der Überlegungen in Unterabschnitt 5.3.1 und den daran

133 Ein ähnliches Vorgehen zur Berechnung von altersspezifischen Geburtenraten von Männern, dass auf einer „Verteilung“ von Geburten auf einzelne Alter von Männern basiert, hat Brouard (1977) angewandt, um Geburtenraten von französischen Männern für den Zeitraum von 1899 bis 1974 zu rekonstruieren.

anschließenden Ausführungen in Anhang B kann dieser Wert über

$$J_{(y,m),t} = \frac{1}{2}P_{(y,m)}n_{(y,m),t} + \frac{3}{8}(1 - P_{(y,m)})n_{(y,m),t} \\ + \frac{1}{2}P_{(y-1,m)}n_{(y-1,m),t} + \frac{1}{8}(1 - P_{(y-1,m)})n_{(y-1,m),t} \quad (5.54)$$

bestimmt werden.<sup>134 135</sup>

Liegen  $F_{(y,m)}^g$  und  $F_{(y,m)}^{g*}$  für alle interessierenden Alter vor, kann analog zur oben beschriebenen Vorgehensweise bei Frauen die erwartete Zahl an Nachfahren von Männern berechnet werden. Hierfür muss einfach an den entsprechenden Stellen  $F_{(x,w)}^g$  durch  $F_{(x,m)}^g$  beziehungsweise  $F_{(x,w)}^{g*}$  durch  $F_{(x,m)}^{g*}$  und gegebenenfalls  $P_{(x,w)}^*$  durch  $P_{(x,m)}^*$  ersetzt werden. Beispielsweise ergibt sich die Zahl der erwarteten Enkelkinder über Söhne eines Mannes in Altersklasse  $a$  als

$$\sum_{x=\alpha}^{a-1} F_{(x,m)}^{m*} \sum_{y=\alpha}^{a-x-2} P_{(y,m)}^* F_{(y,m)}^w P_{(a-x-y-2,w)}^* \quad (5.55)$$

Nun kann man auch die erwartete Zahl der lebenden Enkelkinder einer Frau im Alter  $a$  angeben. Diese setzt sich zusammen aus Enkelkinderern und -söhnen über Töchter und Enkelkinderern und -söhnen über Söhne. Dies erhält man über

$$\sum_{x=\alpha}^{a-1} \left[ F_x^{w*} \sum_{y=\alpha}^{a-x-2} P_{(y,w)}^* \left( F_y^w P_{(a-x-y-2,w)}^* + F_y^m P_{(a-x-y-2,m)}^* \right) \right] \\ + \sum_{x=\alpha}^{a-1} \left[ F_x^{m*} \sum_{y=\alpha}^{a-x-2} P_{(y,m)}^* \left( F_{(y,m)}^w P_{(a-x-y-2,w)}^* + F_{(y,m)}^m P_{(a-x-y-2,m)}^* \right) \right] \quad (5.56)$$

Die erwartete Zahl der lebenden Enkelkinder eines männlichen Egos lässt sich dann berechnen, indem die um Mortalität korrigierten Geburtenraten, die auf die öffnenden eckigen Klammern folgen,  $F_x^{w*}$  und  $F_x^{m*}$ , durch  $F_{(x,m)}^{w*}$  und  $F_{(x,m)}^{m*}$  ersetzt werden.

Im weiteren wird davon ausgegangen, dass die Verteilungen  $\phi(y|x)$  nicht stochastisch und gegeben  $x$  konstant für alle  $t$  sind, sich also nicht über die Zeit hinweg ändern. Ferner wird im weiteren Verlauf auch  $\phi^*(y|x)$  bei den Berechnungen zu Vorfahren benutzt. Dies erfasst den Anteil von Geburten zwischen  $t$  und  $t + 1$  von Frauen, die zu  $t$  in Altersklasse  $x$  sind, denen ein Vater zugeordnet werden kann, der zu  $t$  in Altersklasse  $y$  ist. Sei zunächst

$$\phi'(y|x) = \phi(y|x)\kappa(y|x) + \phi(y|x+1)\kappa(y|x+1) \quad (5.57)$$

eine entsprechende Zuordnung von Frauen, die zu  $t$  in Altersklasse  $x$  sind, zu Männern, die zwischen  $t$  und  $t + 1$  im Altersintervall  $[y, y + 1)$  Vater werden. Dabei entspricht  $\kappa(y|x)$

$$\frac{P_{(x,w)} \frac{1}{2} f(x,w) + (1 - P_{(x,w)}) \frac{3}{8} f(x,w)}{P_{(x,w)} \frac{1}{2} (f(x,w) + f(x+1,w)) + (1 - P_{(x,w)}) (\frac{3}{8} f(x,w) + \frac{1}{8} f(x+1,w))} \quad (5.58)$$

134 Ob dieses Vorgehen zu konsistenten Geburtenzahlen für Männer und Frauen führt, wird in Anhang C diskutiert.

135 Zur konkreten Berechnung von Geburtenraten für Männer wird im weiteren Verlauf der Arbeit eine andere Schätzung von  $J_{(y,m),t}$  verwendet, die aber nur zu marginal abweichenden Ergebnissen führt. S. hierzu Anhang A.3.

und  $\kappa(y|x + 1)$  ist gleich

$$\frac{P_{(x,w)} \frac{1}{2} f_{(x+1,w)} + (1 - P_{(x,w)}) \frac{1}{8} f_{(x+1,w)}}{P_{(x,w)} \frac{1}{2} (f_{(x,w)} + f_{(x+1,w)}) + (1 - P_{(x,w)}) (\frac{3}{8} f_{(x,w)} + \frac{1}{8} f_{(x+1,w)})} \quad (5.59)$$

Die Männer sind zu  $t$  entweder in Altersklasse  $y - 1$  oder  $y$ , so dass noch entsprechend gewichtet werden muss:

$$\begin{aligned} \phi^*(y|x) = & \phi'(y|x) \frac{\frac{1}{2} P_{(y,m)} n_{(y,m),t} + \frac{3}{8} (1 - P_{(y,m)}) n_{(y,m),t}}{J_{(y,m),t}} \\ & + \phi'(y + 1|x) \frac{\frac{1}{2} P_{(y,m)} n_{(y,m),t} + \frac{1}{8} (1 - P_{(y,m)}) n_{(y,m),t}}{J_{(y+1,m),t}} \end{aligned} \quad (5.60)$$

Solange die demographischen Raten und die Bevölkerungszusammensetzung nicht konstant sind, ist auch  $\phi^*(y|x)$  nicht notwendigerweise konstant.

### 5.3.4 Direkte Vorfahren

#### 5.3.4.1 Lebende Eltern

Unter der Annahme, dass Geburten unabhängig vom Alter der Eltern mit konstanten Wahrscheinlichkeiten zufällig männlich oder weiblich sind, ist für die Betrachtung der Frage, ob die Eltern einer Person noch leben, das Geschlecht dieser Person unerheblich und es reicht, ihr Alter zu betrachten.

Ausgegangen wird von einem Ego im Alter  $a$  zum Zeitpunkt  $t$ . Wenn Egos Mutter bei der Geburt in der Altersklasse  $y$  war, ist die Wahrscheinlichkeit, dass diese noch lebt gegeben über

$$\begin{aligned} & E(n_{(y+a+1,w),t} = 1 | \mathbf{n}_{t-a-1} = \mathbf{e}_{(y,w)}) \\ & = \Pr(n_{(y+a+1,w),t} = 1 | \mathbf{n}_{t-a-1} = \mathbf{e}_{(y,w)}) \\ & = \frac{P_{(y+a+1,w)}^*}{P_{(y,w)}^*} \end{aligned} \quad (5.61)$$

Dies muss noch mit der Wahrscheinlichkeit, dass die Mutter bei der Geburt in der Altersklasse  $y$  war, gewichtet werden. Dann ist die Wahrscheinlichkeit, dass die Mutter der betrachteten Person noch lebt, gleich

$$\sum_{y=\alpha}^{\beta} W_w(y|t - a - 1) \frac{P_{(y+a+1,w)}^*}{P_{(y,w)}^*} \quad (5.62)$$

Hierbei ist  $W_w(y|t - a - 1)$  der Anteil von im Intervall  $[t - a - 1, t - a]$  gebärenden Müttern in Altersklasse  $y$  an allen gebärenden Müttern und lässt sich berechnen als:

$$W_w(y|t - a - 1) = \frac{n_{(y,w),t-a-1} F_y}{\sum_{x=\alpha}^{\beta} n_{(x,w),t-a-1} F_x} \quad (5.63)$$

Die Idee bei dieser Berechnung ist, dass Ego zum Zeitpunkt  $t - a$  in der Altersklasse 0 war und somit zwischen  $t - a - 1$  und  $t - a$  geboren wurde. Entsprechend kommen Frauen, die zwischen  $t - a - 1$  und  $t - a$  ein Kind bekommen haben und sich somit in der reproduktiven Phase befinden müssen, als Mutter von Ego in Frage. Befindet sich eine Frau zu  $t - a - 1$  in Altersklasse  $y$ , befindet sie sich zu  $t$  in der Altersklasse  $y + a + 1$  und muss somit von der Altersklasse  $y$  bis zur Altersklasse  $y + a + 1$  überleben. Allerdings ist dabei zu beachten, dass dieses Vorgehen voraussetzt, dass Fertilität und Mortalität nicht miteinander korreliert sind. Dies bedeutet, dass für Personen, die mindestens ein Kind geboren haben, die Überlebenswahrscheinlichkeiten  $P_x$  gelten. Diese Bedingung wiederum impliziert, dass die Homogenitätsannahme erfüllt sein muss, also für alle Personen die gleichen Raten und Wahrscheinlichkeiten gelten.<sup>136</sup>

Soll die Wahrscheinlichkeit berechnet werden, dass der Vater einer Person noch lebt, wird analog zu Müttern vorgegangen und es werden lediglich anstelle der benötigten Wahrscheinlichkeiten für Frauen entsprechende Wahrscheinlichkeiten für Männer verwendet:<sup>137</sup>

$$\sum_{y=\alpha'}^{\beta'} W_m(y|t-a-1) \frac{P_{(y+a+1,m)}^*}{P_{(y,m)}^*} \quad (5.64)$$

$W_m(y|t-a-1)$  kann dabei analog zu (5.63) ermittelt werden.

Möchte man die Wahrscheinlichkeit berechnen, dass sowohl die Mutter als auch der Vater leben, reicht es nicht aus, die Einzelwahrscheinlichkeiten zu multiplizieren, sondern es muss die gemeinsame Verteilung der Alter von Vätern und Müttern bei der Geburt betrachtet werden. Die im vorausgegangenen Satz zuerst genannte Wahrscheinlichkeit lässt sich bei

136 Dieses Vorgehen überschätzt die Sterbewahrscheinlichkeit der Mutter und unterschätzt somit die Wahrscheinlichkeit für ihr Vorhandensein. Nimmt man an, dass die betrachtete Person zwischen  $t$  und  $t + 1$  geboren wurde, lässt sich das Überleben der Mutter bis  $t + 1$  daraus ableiten, wie viele der den Geburten zugeordneten Mütter sterben. Der Anteil von sterbenden Frauen an Müttern ist nicht gleich  $(1 - P_x)$ , sondern ergibt sich bei Verwendung von (5.29) als

$$\frac{(1 - P_x)(\frac{3}{8}f_x + \frac{1}{8}f_{x+1})}{(1 - P_x)(\frac{3}{8}f_x + \frac{1}{8}f_{x+1}) + P_x \frac{1}{2}(f_x + f_{x+1})}$$

Wenn beispielsweise  $f_x = f_{x+1}$  gilt, erhält man

$$\frac{(1 - P_x)\frac{1}{2}}{(1 - P_x)\frac{1}{2} + P_x}$$

was immer niedriger als  $(1 - P_x)$  ist. Praktisch wird dies kaum Bedeutung haben, da die Sterbewahrscheinlichkeiten für Frauen in Altern von  $\alpha$  bis  $\beta$  sehr gering sind.

Das von Lotka (1931) andererseits angeführte erhöhte Sterberisiko von Müttern aufgrund von gestationsbedingten Sterbefällen (beispielsweise aufgrund von Komplikationen bei der Schwangerschaft oder der Geburt) hat in Deutschland seit der Nachkriegszeit kontinuierlich abgenommen und gilt als sehr gering (vgl. Welsch & Wischnik 2006: 1053f.), weshalb es nicht berücksichtigt wird.

137 Hier ergibt sich im Gegensatz zu Müttern eine Überschätzung der Wahrscheinlichkeit, da nicht mit Sicherheit davon ausgegangen werden kann, dass der Vater bei der Geburt noch lebt. Nimmt man an, dass die Geburt zwischen  $t - a - 1$  und  $t - a$  stattfindet, liegt die Zeugung und somit der letzte Zeitpunkt, an dem der Vater sicher gelebt hat, ungefähr zwischen  $t - a - 1.75$  und  $t - a - 0.75$ . Allerdings stellte bereits Lotka (1931) fest, dass dieser Aspekt kaum Einfluss auf das Ergebnis hat, weshalb er nicht berücksichtigt wird (s. auch Fußnote 75 auf S. 62).

Vernachlässigung von Fertilität formal bezeichnen als

$$\begin{aligned} \sum_{x=\alpha}^{\beta} \sum_{y=\alpha'}^{\beta'} & \Pr(n_{(x+a+1,w),t} = 1 | \mathbf{n}_{t-a-1} = \mathbf{e}_{(x,w)}) \\ & \cdot \Pr(n_{(y+a+1,m),t} = 1 | \mathbf{n}_{t-a-1} = \mathbf{e}_{(y,m)}) \\ & \cdot \Pr(\mathbf{n}_{t-a-1} = \mathbf{e}_{(x,w)} + \mathbf{e}_{(y,m)}) \end{aligned} \quad (5.65)$$

und kann berechnet werden über<sup>138</sup>

$$\sum_{x=\alpha}^{\beta} W_w(x|t-a-1) \frac{P_{(x+a+1,w)}^*}{P_{(x,w)}^*} \sum_{y=\alpha'}^{\beta'} \phi^*(y|x) \frac{P_{(y+a+1,m)}^*}{P_{(y,m)}^*} \quad (5.66)$$

Dies ist nur gleich dem Produkt der Einzelwahrscheinlichkeiten, wenn  $\phi^*(y|x)$  für alle  $x$  identisch ist. Allgemein besteht die Idee darin, zunächst von Müttern in einer Altersklasse  $x$  auszugehen und deren Anteil an Geburten und deren Überlebenswahrscheinlichkeiten zu betrachten. Dies wird kombiniert mit der Verteilung der Altersklassen von Vätern in Abhängigkeit von  $x$  und den entsprechenden Überlebenswahrscheinlichkeiten.

Die Wahrscheinlichkeit, dass nur die Mutter noch lebt und der Vater nicht, lässt sich analog herleiten, wobei nun die Wahrscheinlichkeit verwendet wird, dass der Vater nicht bis zur Altersklasse  $y + a + 1$  überlebt:

$$\sum_{x=\alpha}^{\beta} W_w(x|t-a-1) \frac{P_{(x+a+1,w)}^*}{P_{(x,w)}^*} \sum_{y=\alpha'}^{\beta'} \phi^*(y|x) \left(1 - \frac{P_{(y+a+1,m)}^*}{P_{(y,m)}^*}\right) \quad (5.67)$$

Auf gleiche Weise erhält man auch die Wahrscheinlichkeit, dass nur der Vater noch lebt

$$\sum_{x=\alpha}^{\beta} W_w(x|t-a-1) \left(1 - \frac{P_{(x+a+1,w)}^*}{P_{(x,w)}^*}\right) \sum_{y=\alpha'}^{\beta'} \phi^*(y|x) \frac{P_{(y+a+1,m)}^*}{P_{(y,m)}^*} \quad (5.68)$$

und die Wahrscheinlichkeit, dass beide Elternteile nicht mehr leben ist gleich

$$\sum_{x=\alpha}^{\beta} W_w(x|t-a-1) \left(1 - \frac{P_{(x+a+1,w)}^*}{P_{(x,w)}^*}\right) \sum_{y=\alpha'}^{\beta'} \phi^*(y|x) \left(1 - \frac{P_{(y+a+1,m)}^*}{P_{(y,m)}^*}\right) \quad (5.69)$$

#### 5.3.4.2 Lebende Großeltern

Da sich Großeltern als „Eltern der Eltern“ auffassen lassen, kann die Wahrscheinlichkeit für lebende Großeltern durch eine wiederholte Anwendung des im letzten Paragraph dargestellten Schemas berechnet werden. Hierbei muss zwischen Großeltern mütterlicher- und väterlicherseits unterschieden werden. Zwar weisen diese aufgrund der Modellannahmen nicht per se unterschiedliche Sterberisiken auf, allerdings führen unterschiedliche Altersverteilungen

<sup>138</sup> Die Verwendung von  $\phi^*(y|x)$  erfolgt relativ analog zum Vorgehen von Le Bras (1973) und Le Bras & Wachter (1978).

bei der Geburt eines Kindes bei Männern und Frauen dazu, dass die Großeltern über den durchschnittlich älteren Elternteil ebenfalls tendenziell älter sind. Bei Ego ist wie bei der Betrachtung von lebenden Eltern nur das Alter und nicht das Geschlecht von Bedeutung.

Interessiert man sich nur für das Überleben eines bestimmten Großelternanteils, muss sowohl das Alter des vermittelnden Elternteils wie auch des betrachteten Großelternanteils berücksichtigt werden. Beispielsweise lässt sich die Wahrscheinlichkeit, dass die Großmutter mütterlicherseits einer Person, die zum Zeitpunkt  $t$  in der Altersklasse  $a$  ist, noch lebt, berechnen als

$$\sum_{x=\alpha}^{\beta} W_w(x|t-a-1) \sum_{y=\alpha}^{\beta} W_w(y|t-a-x-2) \frac{P_{(y,x+a+2,w)}^*}{P_{(y,w)}^*} \quad (5.70)$$

Hierbei wird sowohl die Unsicherheit des Alters der Mutter als auch des Alters der Großmutter von Ego berücksichtigt. Zunächst werden die möglichen Alter  $x$  der Mutter betrachtet und anschließend die Alter der Großmutter  $y$ . Hierüber ergibt sich, in welcher Altersklasse die Großmutter sein muss, wenn sie bis  $t$  überlebt hat.

Auf gleiche Weise lassen sich Überlebenswahrscheinlichkeiten für die Großmutter väterlicherseits sowie Großväter herleiten. Die Wahrscheinlichkeit, dass der Großvater väterlicherseits noch lebt, lautet zum Beispiel

$$\sum_{x=\alpha'}^{\beta'} W_m(x|t-a-1) \sum_{y=\alpha'}^{\beta'} W_m(y|t-a-x-2) \frac{P_{(y,x+a+2,m)}^*}{P_{(y,m)}^*} \quad (5.71)$$

Möchte man das gleichzeitige Überleben (beziehungsweise Nicht-Überleben) aller Großeltern betrachten, müssen sowohl die Abhängigkeiten der Alter der Eltern als auch der Großeltern berücksichtigt werden. Hierzu wird von der Mutter einer Person in der Altersklasse  $a$  ausgegangen:

$$\begin{aligned} & \sum_{x=\alpha}^{\beta} W_w(x|t-x-1) \\ & \cdot \left[ \sum_{z=\alpha}^{\beta} W_w(z|t-a-x-2) \frac{P_{(z,x+a+2,w)}^*}{P_{(z,w)}^*} \sum_{z'=\alpha'}^{\beta'} \phi^*(z'|z) \frac{P_{(z'+x+a+2,m)}^*}{P_{(z',m)}^*} \right] \\ & \cdot \left[ \sum_{y=\alpha'}^{\beta'} \phi^*(y|x) \left\langle \sum_{z=\alpha}^{\beta} W_w(z|t-a-y-2) \frac{P_{(z,y+a+2,w)}^*}{P_{(z,w)}^*} \right. \right. \\ & \quad \left. \left. \cdot \sum_{z'=\alpha'}^{\beta'} \phi^*(z'|z) \frac{P_{(z'+y+a+2,m)}^*}{P_{(z',m)}^*} \right\rangle \right] \end{aligned} \quad (5.72)$$

Ausgehend von den Anteilswerten möglicher Alter der Mutter bei der Geburt von Ego werden Anteilswerte möglicher Alter der Großmutter mütterlicherseits ermittelt und in Abhängigkeit von diesen Anteilswerte möglicher Alter des Großvaters mütterlicherseits. Zudem werden

in Abhängigkeit vom Alter der Mutter Anteilswerte möglicher Alter des Vaters betrachtet und hierüber analog wie bei Großeltern mütterlicherseits Anteilswerte möglicher Alter der Großeltern väterlicherseits festgestellt. Ersetzt man in (5.72) einzelne Überlebenswahrscheinlichkeiten mit ihren Gegenwahrscheinlichkeiten, lassen sich die Wahrscheinlichkeiten berechnen, dass bestimmte Großelternanteile noch leben und andere tot sind beziehungsweise alle Großeltern nicht mehr leben, ganz analog zum Vorgehen bei Eltern.

### 5.3.5 Alter von Vor- und Nachfahren

Ausgehend von den bisher vorgestellten Formeln lassen sich relativ analog zum Vorgehen von Keyfitz & Caswell (2005) (s. Abschnitt 4.4) die Verteilungen der Altersklassen von lebenden Vor- und Nachfahren bestimmen. Beispielhaft wird eine Frau in Altersklasse  $a$  betrachtet. Der erwartete Anteil lebender Töchter in Altersklasse  $a - y - 1$ ,  $\alpha \leq y \leq a - 1$ , bezogen auf alle Töchter ist gleich

$$\frac{F_y^{w*} P_{(a-y-1,w)}^*}{\sum_{x=\alpha}^{a-1} F_x^{w*} P_{(a-x-1,w)}^*} \quad (5.73)$$

Hieraus lässt sich der Erwartungswert der Altersklasse lebender Töchter einfach bestimmen über

$$\sum_{y=\alpha}^{a-1} \frac{(a - y - 1) F_y^{w*} P_{(a-y-1,w)}^*}{\sum_{x=\alpha}^{a-1} F_x^{w*} P_{(a-x-1,w)}^*} \quad (5.74)$$

Entsprechende Ausdrücke für andere Nachfahren lassen sich analog herleiten – die erwartete Zahl der entsprechenden Nachfahren steht im Nenner, während sich im Zähler die erwartete Zahl der Nachfahren in einer bestimmten Altersklasse multipliziert mit dieser Altersklasse findet.

Bei Vorfahren kann identisch vorgegangen werden. Beispielsweise ergibt sich für eine Person in Altersklasse  $a$  der Erwartungswert der Altersklasse des Vaters unter der Voraussetzung, dass dieser noch lebt, als

$$\sum_{z=\alpha'}^{\beta'} \frac{(z + a + 1) W_m(z|t - a - 1) \frac{P_{(z+a+1,m)}^*}{P_{(y,m)}^*}}{\sum_{y=\alpha'}^{\beta'} W_m(y|t - a - 1) \frac{P_{(y+a+1,m)}^*}{P_{(y,m)}^*}} \quad (5.75)$$

Das durchschnittliche Alter des Vaters wird also berechnet über die Altersverteilung von Vätern im Geburtsjahr von Ego gewichtet mit den Überlebenswahrscheinlichkeiten. Da sich die ursprünglichen Anteilswerte gewichtet mit den Überlebenswahrscheinlichkeiten nicht zu 1 summieren, wird anschließend durch ihre Summe dividiert. Für andere Vorfahren lässt sich analog vorgehen.

### 5.3.6 Zeitveränderliche und stochastische Parameter

#### 5.3.6.1 Zeitveränderliche Parameter

Um zeitveränderliche Parameter einzuführen, werden die bisher verwendeten Modellparameter mit einem zusätzlichen Index  $t$  versehen, der die Zeitstelle erfasst, auf die sich ein Parameter bezieht, beispielsweise  $F_{x,t}^w$ .  $F_{x,t}^w$  wird nun über

$$P_{(x,w),t} \frac{1}{2} l_{(0.5,w),t} (f_{x,t}^w + f_{x+1,t}^w) + (1 - P_{(x,w),t}) l_{(0.5,w),t} \left( \frac{3}{8} f_{x,t}^w + \frac{1}{8} f_{x+1,t}^w \right) \quad (5.76)$$

und  $P_{(x,w),t}$  über  $1/2(p_{(x,w),t} + p_{(x,w)+1,t})$  sowie  $l_{(0.5,w),t}$  über  $\sqrt{p_{(0,w),t}}$  geschätzt. Entsprechendes gilt für  $F_{x,t}^{g*}$ .

Beispielsweise lässt sich nun die erwartete Zahl der lebenden Töchter einer Frau im Alter  $a$ , die zu  $t$  im Alter  $\alpha$  war, schreiben als

$$\begin{aligned} & \sum_{x=0}^{a-\alpha-1} \mathbb{E} (n_{(a-\alpha-x-1,w),t+a-\alpha} | \mathbf{n}_{t+x} = \mathbf{e}_{(\alpha+x,w,+)} ) \\ = & \sum_{x=\alpha}^{a-1} F_{x,t+x-\alpha}^{w*} P_{(a-x-1,w),t+a-\alpha}^* \end{aligned} \quad (5.77)$$

wobei

$$P_{(x,g),t}^* = \begin{cases} 1 & \text{wenn } x = 0 \\ \prod_{y=0}^{x-1} P_{(y,g),t-x+y} & \text{wenn } 0 < x \leq \rho \\ 0 & \text{wenn } x > \rho \end{cases} \quad (5.78)$$

ist.

Ein anderes Beispiel ist die Wahrscheinlichkeit, dass der Vater und die Mutter einer Person im Alter  $a$  zu  $t$  noch leben:

$$\sum_{x=\alpha}^{\beta} W_w(x|t-a-1) \frac{P_{(x+a+1,w),t}^*}{P_{(x,w),t-a-1}^*} \sum_{y=\alpha'}^{\beta'} \phi^*(x|y, t-a-1) \frac{P_{(y+a+1,m),t}^*}{P_{(y,m),t-a-1}^*} \quad (5.79)$$

Die weiteren Ausdrücke für Verwandtschaft lassen sich bei zeitveränderlichen Parametern auf gleiche Weise herleiten.

#### 5.3.6.2 Stochastische Parameter

Im weiteren wird angenommen, dass die Parameter  $F_{x,t}$  beziehungsweise  $F_{x,t}^*$  sowie  $P_{x,t}$  über einen oder mehrere zunächst nicht näher bestimmte stochastische Prozesse beziehungsweise Transformationen dieser Prozesse gegeben sind. Beim erstgenannten Fall betrachtet man  $\tilde{F}_{x,t}$

beziehungsweise  $\tilde{F}_{x,t}^*$  und  $\tilde{P}_{x,t}$  direkt, beim zweiten Fall  $\tilde{f}_{x,t}$  und  $\tilde{p}_{x,t}$ .  $\tilde{F}_{x,t}$ ,  $\tilde{F}_{x,t}^*$  und  $\tilde{P}_{x,t}$  können dann über die in diesem und den vorausgegangenen Abschnitt genannten Formeln hergeleitet werden. Im folgenden wird von der zweiten Variante ausgegangen.

Für eine konkrete Realisierung der stochastischen Raten können die bisher entwickelten Formeln ohne weiteres angewendet werden. Allgemein kann aber jeweils auch der Erwartungswert betrachtet werden. Beispielsweise ist der Erwartungswert der erwarteten Zahl lebender Töchter gegeben über

$$\mathbb{E} \left( \sum_{x=\alpha}^{a-1} \tilde{F}_{x,t+x-\alpha}^{w*} \tilde{P}_{(a-x-1,w),t+a-\alpha}^* \right) \quad (5.80)$$

Lässt man der Übersichtlichkeit halber Indizes für Zeitstellen unberücksichtigt und nimmt an, dass Fertilität und Mortalität nicht korreliert sind, lässt sich (5.80) durch einsetzen der entsprechenden Schätzer schreiben als

$$\begin{aligned} & \mathbb{E} \left( \tilde{F}_{a-1}^{w*} + \sum_{x=\alpha}^{a-2} \tilde{F}_x^{w*} \prod_{y=0}^{a-x-1} \tilde{P}_{(y,w)} \right) \\ &= \mathbb{E} \left( \frac{1}{2} \left[ \tilde{f}_{a-1}^w + \tilde{f}_a^w \right] \sqrt{\tilde{p}_{(0,w)}} \right. \\ & \quad \left. + \sum_{x=\alpha}^{a-2} \frac{1}{2} \left[ \tilde{f}_x^w + \tilde{f}_{x+1}^w \right] \sqrt{\tilde{p}_{(0,w)}} \prod_{y=0}^{a-x-1} \frac{1}{2} \left[ \tilde{p}_{(y,w)} + \tilde{p}_{(y+1,w)} \right] \right) \\ &= \frac{1}{2} \mathbb{E} \left( \tilde{f}_{a-1}^w \right) \mathbb{E} \left( \sqrt{\tilde{p}_{(0,w)}} \right) + \frac{1}{2} \mathbb{E} \left( \tilde{f}_a^w \right) \mathbb{E} \left( \sqrt{\tilde{p}_{(0,w)}} \right) \\ & \quad + \frac{1}{4} \sum_{x=\alpha}^{a-2} \left\langle \mathbb{E} \left( \tilde{f}_x^w \right) + \mathbb{E} \left( \tilde{f}_{x+1}^w \right) \right\rangle \\ & \quad + \mathbb{E} \left( \sqrt{\tilde{p}_{(0,w)}} \left[ \tilde{p}_{(0,w)} + \tilde{p}_{(1,w)} \right] \cdots \left[ \tilde{p}_{(a-x-1,w)} + \tilde{p}_{(a-x,w)} \right] \right) \end{aligned} \quad (5.81)$$

Möchte man hiermit den Erwartungswert analytisch bestimmen, erweisen sich die letzten Erwartungswerte in der letzten Zeile als problematisch, da diese Erwartungswerte für Produkte mehrerer  $\tilde{p}_{(x,w)}$  nicht ohne weiteres angegeben werden können, weil die einzelnen  $\tilde{p}_{(x,w)}$  möglicherweise korreliert sind. Insbesondere wird für den Erwartungswert eines Produktes aus  $k$  Zufallsvariablen das  $k$ -te gemeinsame Moment dieser Zufallsvariablen benötigt, welches in diesem Fall nicht ohne weiteres zu bestimmen ist. Dies gilt insbesondere, wenn kein stochastischer Prozess direkt für  $p_{(x,w)}$  gegeben ist, sondern wie oftmals üblich für  $r_{(x,w)}$ , so dass eigentlich noch der Schritt von der Sterberate zur Überlebenswahrscheinlichkeit berücksichtigt werden müsste.

Ausgehend von stochastischen Raten lassen sich Verwandtenzahlen somit zwar ohne weiteres berechnen, allerdings zeigt das Beispiel des Erwartungswertes der erwarteten Zahl der lebenden Töchter, dass eine analytische Charakterisierung der sich aus den stochastischen Raten ergebenden Verteilung schwierig sein dürfte. Um dennoch Verteilungen und Charakterisierungen dieser Verteilungen für die interessierenden Größen berechnen zu können, wird bei

der Vorausberechnung deshalb auf Monte-Carlo-Simulationen zurückgegriffen (s. auch die Anmerkungen in Unterabschnitt 5.2.5).

## 5.4 Zusammenfassung und Diskussion

In diesem Kapitel wurde ein Ansatz vorgestellt, der die Berechnung der Zahl linearer Vor- und Nachfahren und deren (durchschnittlichen) Alters mittels demographischer Raten aus einer egozentrierten Perspektive erlaubt. Dabei wurde ein Augenmerk auf die Berücksichtigung von Unsicherheit in den demographischen Raten gelegt. Ferner ist der Ansatz „kompatibel“ mit der Kohorten-Komponenten-Methode. Es kann somit nicht nur angegeben werden, wie viele lineare Verwandte für Personen eines bestimmten Geschlechts und Alters zu erwarten sind, sondern auch, wie viele dieser Personen es gibt – wie viele Personen also von den berechneten Verwandtenanzahlen „betroffen“ sind. Der Ansatz enthält zudem eine relativ simple Lösung des Zwei-Geschlechter-Problems. Von den Daten abgesehen, die für die Lösung des Zwei-Geschlechter-Problems benötigt werden, weist der Ansatz einen geringen Datenbedarf auf. Insbesondere können auch historische demographische Raten verwendet werden, womit sich prinzipiell Verwandtenzahlen zu beliebigen Zeitpunkten berechnen lassen und auch die Entwicklung zwischen zwei Zeitpunkten.

Bezogen auf die in Unterabschnitt 4.1 diskutierten Probleme von demographischen Vorausberechnungen lässt sich festhalten, dass Unsicherheit bezüglich der demographischen Raten wie erwähnt berücksichtigt werden kann. Die Unsicherheit der benötigten Daten hält sich zumindest bezogen auf Sterbe- und Geburtenraten und Bevölkerungsbestände in engen Grenzen, wobei sich eine genauere Diskussion der für die BRD zu Verfügung stehenden Daten in Anhang A findet.

Bezogen auf das Zwei-Geschlechter-Problem stellt sich die Frage, inwiefern das Vorgehen zu konsistenten Ergebnissen führt. Das dies der Fall ist, wird in Anhang C.1 gezeigt. Daneben ist zu klären, wie  $\phi(y|x)$  ermittelt werden kann und ob der Ansatz zu realistischen Geburtenraten für Männer führt. Die Datengrundlage zur Ermittlung der Verteilungen  $\phi(y|x)$ , hierüber erzielte Ergebnisse sowie eine Einschätzung dieser Ergebnisse werden im nächsten Kapitel beschrieben. Theoretische Ergänzungen finden sich ebenfalls in Anhang C.1.

Weiterhin wurden in diesem Kapitel etliche Annahmen getroffen, die als Voraussetzung dafür angesehen werden können, dass der Ansatz realistische Ergebnisse liefert. Dies umfasst die folgenden Annahmen:

- die Unterteilung der Bevölkerung nach Alter und Geschlecht ist ausreichend (Homogenität bzw. keine Heterogenität)
- kein systematischer Zusammenhang zwischen Mortalität und Fertilität (Unkorreliertheit der Komponenten)
- keine Abhängigkeit von Ereignissen untereinander<sup>139</sup>
- Unabhängigkeit von Personen untereinander

---

139 Eine Ausnahme ist natürlich der Tod einer Person, der das Eintreten weiterer Ereignisse für die tote Person ausschließt.

- kein Einfluss von Migration auf die Anzahl der Verwandten
- Ereignisse finden durchschnittlich in der Jahresmitte statt
- $\phi(y|x)$  ist jeweils für jedes  $x$  konstant

Diese Annahmen sind zwar unterschiedlich stark, aber sie sind alle nicht unproblematisch. In der Regel dürfte anzunehmen sein, dass sie verletzt sind. Beispielsweise ist nicht davon auszugehen, dass Personen voneinander unabhängig sind, sich also nicht gegenseitig beeinflussen. In der Literatur lassen sich zahlreiche Gegenbeispiele finden, die auf das Gegenteil und somit eine mehr oder weniger starke wechselseitige Beeinflussung von Personen untereinander hinweisen (vgl. bspw. Keim et al. 2009; Murphy 1999; Murphy & Wang 2001).<sup>140</sup> Theoretische Überlegungen, ob die Verletzung der genannten Annahmen zu fehlerhaften Ergebnissen führen kann, werden in Anhang C vorgestellt. Bei einigen Annahmen zeigt sich, dass ihre Verletzung zu Verzerrungen der Ergebnisse führen kann. Im übernächsten Kapitel findet sich deshalb wie bereits in Abschnitt 4.8 gefordert eine empirische Überprüfung der Validität des Ansatzes, bei der genutzt wird, dass auch historische demographische Raten verwendet und hierüber erzielte Ergebnisse mit anderen Datenquellen verglichen werden können.

Eine in der obigen Auflistung nicht genannte und bisher nicht diskutierte Annahme ist, dass eine zeitdiskrete Modellierung von Verwandtschaft angemessen ist. Diese Frage stellt sich insbesondere, da etliche der Ansätze, die in Kapitel 4 diskutiert wurden, zeitstetig sind. Einige theoretische Überlegungen hierzu und ein Vergleich von diskreter und stetiger Modellierung findet sich in Anhang D, wobei der Ansatz nach Goodman et al. (1974) zum Vergleich herangezogen wird.

Schließlich sei erwähnt, dass bei den im Modell möglichen Ereignissen eine sehr hohe Aggregation vorliegt. Es sind lediglich Geburten und Sterbefälle möglich, wobei nicht nach Parität oder Todesursache unterschieden wird. Dies führt zu einem relativ geringen Detailgrad der Ergebnisse – beispielsweise kann nur die erwartete Kinderzahl angegeben werden, nicht aber die Verteilung der Kinderzahl. Zumindest bei Fertilität wäre eine weitergehende Differenzierung zwar wünschenswert, allerdings wird von dieser aufgrund fehlender Daten abgesehen (vgl. bspw. Kreyenfeld et al. 2012).

Ausgehend von den gerade genannten Punkten gestaltet sich der weitere Verlauf der Arbeit zusammengefasst also wie folgt: Im nächsten Kapitel werden die Lösung des Zwei-Geschlechter-Problems und Geburtenraten für Männer besprochen. In Kapitel 7 findet sich eine empirische Überprüfung des hier vorgestellten Ansatzes. Wie die Unsicherheit der demographischen Raten modelliert wird, wird in Kapitel 8 besprochen. Im Anhang A findet sich eine Diskussion der Qualität der Daten. In den Anhängen C und D werden die Modellannahmen theoretisch untersucht.

---

140 Murphy & Wang (2001) untersuchen Daten für Italien, Norwegen, Polen und die USA auf einen möglichen Zusammenhang zwischen der Fertilität von Kindern und der Fertilität derer Eltern. Es wird zu dem Schluss gekommen, dass ein nicht unerheblicher Zusammenhang zwischen diesen beiden besteht (s. auch Pullum & Wolf 1991). Als mögliche Ursachen werden bei Murphy (1999) vor allem genetische Faktoren und Einflüsse der Eltern auf ihre Kinder über Sozialisation genannt. Keim et al. (2009) diskutieren den Einfluss von sozialen Netzwerken auf Fertilitätsintentionen. Über diese ist ein Zusammenhang zwischen Personen über die bloße Korrelation derer Fertilität hinaus möglich. Beispielsweise kann das Vorhandensein von lebenden (alten) Eltern die Entscheidung für oder gegen ein Kind beeinflussen, wenn Eltern als Betreuungsmöglichkeit für dieses Kind mangels institutioneller Alternativen benötigt werden (vgl. ebd.: 13). Darüber hinaus können auch andere Personen wie Geschwister oder Freunde den Kinderwunsch beeinflussen (vgl. ebd.).

## 6 Fertilität von Männern

### 6.1 $\phi(y|x)$ : Datengrundlage und Berechnung

#### 6.1.1 Die Statistik der Geburten

Angaben zu bedingten Verteilungen des Alters von Vätern nach dem Alter von Müttern wurden für Deutschland bisher nicht veröffentlicht. Zwar liegen beispielsweise für das Jahr 2007 Zahlen vor, aus denen sich diese bedingten Verteilungen berechnen ließen, allerdings beschränken sich diese Daten auf eheliche Geburten (vgl. Statistisches Bundesamt 2010b: 58-59). Zur Ermittlung von  $\phi(y|x)$  wurde deshalb auf die Daten der Statistik der Geburten zurückgegriffen, die Informationen zum Alter von Vätern und Müttern enthalten.<sup>141</sup>

Bei der Statistik der Geburten handelt es sich um eine jährliche Vollerhebung aller standesamtlich registrierten Geburten von Müttern, die zur Wohnbevölkerung der BRD zählen.<sup>142</sup> Dabei kann davon ausgegangen werden, dass diese Daten eine hohe Zuverlässigkeit aufweisen (vgl. Statistisches Bundesamt 2010d: 5f.).<sup>143</sup> Die Erfassung beginnt mit dem Jahr 1991.<sup>144</sup> Zur Auswertung standen die Erhebungsjahre 1995, 2000, 2006, 2007 und 2008 zur Verfügung. Für jedes Erhebungsjahr ist ein Datensatz aller Lebend- und Totgeburten vorhanden, wobei jeder (tot-)geborene Säugling eine Datenzeile umfasst, so dass für Mehrlingsgeburten entsprechend mehrere Datenzeilen vorliegen. Die Fallzahlen der einzelnen ausgewerteten Jahre finden sich in Tabelle 6.1.

Aus Datenschutzgründen standen nicht alle in diesen Daten enthaltene Merkmale zur Verfügung, sondern nur eine eingeschränkte Auswahl.<sup>145</sup> Diese umfasste folgende Merkmale:

- Berichtsjahr
- Monat der Geburt
- Bundesland und Kreis der Wohngemeinde der Mutter
- Lebend- oder Totgeburt

---

141 Dies war im Rahmen eines Aufenthalts als Gastwissenschaftler am Forschungsdatenzentrum der Statistischen Landesämter, Standort Düsseldorf, möglich.

142 Das Konzept der Wohnbevölkerung wurde wie bereits in Fußnote 29 auf Seite 31 erwähnt in der amtlichen Statistik lediglich bis 1983 verwendet und durch das in Abschnitt 3.1 beschriebene Konzept der „Bevölkerung am Ort der Hauptwohnung“ abgelöst (vgl. Statistisches Bundesamt 2011c: 30). Zwar gibt es Unterschiede in den Definitionen (vgl. ebd.: 33), nichtsdestotrotz werden diese Begriffe im weiteren der Einfachheit halber synonym verwendet. Für Daten vor 1983 steht der Begriff dann entsprechend für das alte Konzept der Wohnbevölkerung, für Daten ab 1983 dient der Begriff als Kurzform für „Bevölkerung am Ort der Hauptwohnung“.

143 Da nur Frauen, die zur Wohnbevölkerung der BRD gehören, erfasst sind, fehlen Geburten, bei denen zwar der Vater zur Wohnbevölkerung zählt, die Mutter allerdings nicht. Umgekehrt sind Geburten erfasst, bei denen der Vater nicht zur Wohnbevölkerung gehört. Sebald (2008: 136) vermutet, dass sich diese Unterschiede in der Erfassung teils ausgleichen.

144 Für die Jahre 1991 bis 1994 liegen Daten nicht für alle Bundesländer vor. Eine überblicksartige Beschreibung der Daten ist auf der Internetpräsenz des Forschungsdatenzentrums unter <http://www.forschungsdatenzentrum.de/bestand/geburten/index.asp> abrufbar, zudem werden die Daten der Jahre 1991 bis 2005 auch von Sebald (2008: Kap. 7.1) beschrieben.

145 Für eine vollständige Übersicht aller Merkmale siehe Forschungsdatenzentrum der Statistischen Ämter der Länder (2006, 2007, 2009).

Tabelle 6.1: Fallzahlen der Statistik der Geburten für die Jahre 1995, 2000, 2006, 2007 und 2008

Erhebungsjahr	Lebendgeburten	Totgeburten	Geburten insgesamt
1995	765221	3405	768626
2000	766999	3084	770083
2006	672724	2420	675144
2007	684862	2371	687233
2008	682514	2412	684926

Quelle: Statistik der Geburten; eigene Auswertung

- Legitimität (Eltern verheiratet/nicht verheiratet)
- Monat und Jahr der Eheschließung
- Staatsangehörigkeit der Mutter und des Vaters
- Erwerbstätigkeit der Mutter (erwerbstätig/nicht erwerbstätig)
- Alter der Mutter in vollendeten Lebensjahren und Alter des Vaters in vollendeten Jahren
- Wieviertes Kind der Mutter
- Wieviertes Kind der Ehe

Merkmale, die sich auf eine Ehe beziehen, liegen nur für verheiratete Elternpaare vor. Angaben zum Vater eines Kindes sind nur bei ehelichen Geburten verpflichtend. Bei nichtehelichen Geburten wurden vor dem Jahr 2000 keine Angaben zum Vater erhoben. Seit dem Jahr 2000 sind freiwillige Angaben möglich, aber nicht verpflichtend, so dass Angaben zum Vater bei nichtehelichen Geburten teils fehlen. Alle Altersangaben zu Eltern, die unter 16 Jahren liegen, wurden in einer Kategorie von 15 und weniger Lebensjahren zusammengefasst. Bei Frauen wurden ferner alle Alter, die über 49 liegen, in eine Kategorie von 50 und mehr Lebensjahren zusammengefasst und bei Männern wurden alle Alter über 79 in einer in einer Alterskategorie von 80 und mehr Lebensjahren erfasst.

Ausgehend von diesen Daten sollen die Verteilungen  $\phi(y|x)$  berechnet werden, wobei entsprechend der Ausführungen in Unterabschnitt 5.3.3 davon ausgegangen wird, dass diese konstant sind. Dabei sollen die Verteilungen möglichst die aktuelle Situation widerspiegeln. Bei den Auswertungen der Daten wurden deshalb im Wesentlichen vier Fragen betrachtet. Zunächst wurde untersucht, ob sich die Daten der Erhebungsjahre 2006, 2007 und 2008 zur Berechnung von  $\phi(y|x)$  zusammenfassen lassen, um etwaige temporäre Schwankungen in den Daten einzelner Jahre auszugleichen. Hierbei ließ sich auch überprüfen, ob eine starke Variabilität der Verteilungen von Jahr zu Jahr vorliegt, was gegen eine Zusammenfassung sprechen könnte. Anschließend wurde untersucht, ob die resultierenden Verteilungen  $\phi(y|x)$  längerfristigen zeitlichen Schwankungen unterliegen, indem die Daten der Jahre 1995 und 2000 zum Vergleich hinzugezogen wurden. Ebenso wurde auf räumliche Differenzen zwischen West- und Ostdeutschland geprüft. Schließlich wurden die resultierenden Verteilungen, die weiter verwendet werden sollten, geglättet. Totgeburten wurden bei allen Untersuchungen nicht berücksichtigt.

### 6.1.2 Vergleich der Jahre 2006, 2007 und 2008

Zunächst wurde der Frage nachgegangen, ob kurzfristige Schwankungen in den zu berechnenden Verteilungen auftreten. Bereits die Auswertungen von Sebald (2008) lassen vermuten, dass solche Schwankungen nicht vorliegen. Zum einen entwickelte sich im Zeitraum von 1991 bis 2005 das durchschnittliche Alter von Vätern bei der Geburt eines Kindes relativ parallel zum durchschnittlichen Alter von Müttern bei einer Geburt (vgl. Sebald 2008: 152f.), zum anderen veränderte sich die durchschnittliche Altersdifferenz zwischen Vater und Mutter in diesem Zeitraum nur leicht (vgl. ebd.: 141ff.).

Um dieser Vermutung weiter nachzugehen, wurden konditionale Verteilungen  $\phi(y|x)$  jeweils für die Jahre 2006, 2007 und 2008 berechnet und für alle Alter  $x$  paarweise Vergleiche der drei Verteilungen mittels des Dissimilaritätsindex durchgeführt.<sup>146</sup> Beobachtungen, für die das Alter des Vaters nicht vorlag, wurden bei der Ermittlung der Verteilungen nicht berücksichtigt.<sup>147</sup> Die Ergebnisse sind in Abbildung 6.1 graphisch dargestellt. Für Alter  $x$  von 20 bis 40 Lebensjahren liegt der Wert des Dissimilaritätsindex zwischen 0.015 und 0.04, was auf relativ geringe Differenzen zwischen den Verteilungen hinweist. Für Alter unter 20 und über 45 Lebensjahren zeigen sich wesentlich höhere Werte, die auf starke Unterschiede in den Verteilungen verweisen. Diese lassen sich auf die geringen Fallzahlen in diesen Altersklassen zurückführen, wodurch die Ergebnisse stark durch zufällige Schwankungen beeinflusst werden. Beispielsweise liegen für das Jahr 2007 für die Altersklasse von 48 Jahren lediglich 57 Beobachtungen vor, für das Jahr 2008 sind es 66 Beobachtungen.

Ausgehend von diesen Ergebnissen scheint eine Zusammenfassung der Daten der drei betrachteten Jahre möglich und sinnvoll, da dies zum einen bei geburtsintensiven Altern nur zu geringen Unterschieden führt und zum anderen so die geringen Fallzahlen für manche Altersklassen erhöht werden können.

### 6.1.3 Vergleich der Jahre 1995 und 2000 mit den Jahren 2006 bis 2008

Um zu überprüfen, ob es über längere Zeiträume hinweg zu Veränderungen in den Verteilungen  $\phi(y|x)$  kam, wurden die aus den zusammengefassten Daten der Jahre 2006 bis 2008 berechneten Verteilungen zum einen mit den Verteilungen des Jahres 1995 und zum anderen mit

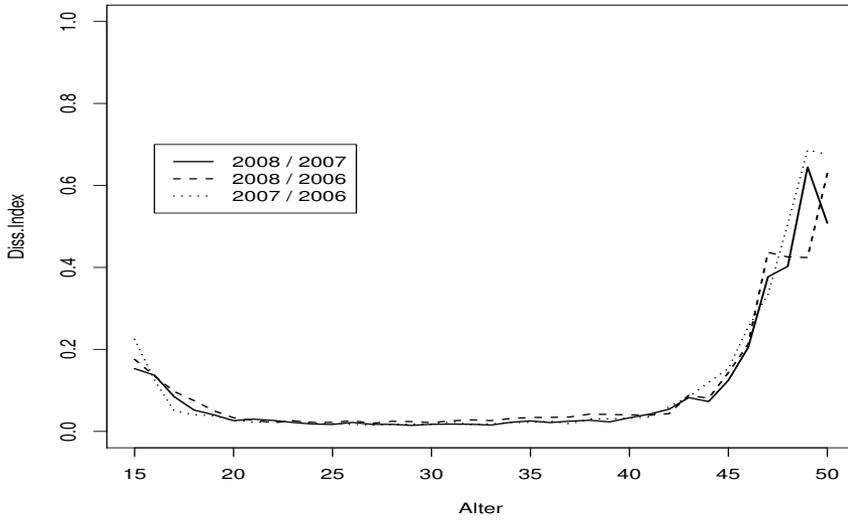
146 Dieser lässt sich für zwei Verteilungen  $\phi(y|x)$  und  $\phi'(y|x)$  berechnen als (vgl. Rowland 2003: 95f.)

$$DI(x) = \frac{1}{2} \sum_y |\phi(y|x) - \phi'(y|x)| \quad (6.1)$$

wobei  $|\cdot|$  hier für den Betrag der Differenz steht.  $DI(x)$  kann dabei Werte zwischen 0 und 1 annehmen, wobei allgemein ein Ergebnis  $z$  bedeutet, dass  $100z$  Prozentpunkte der einen Verteilung umverteilt werden müssten, um die andere zu erhalten. Für  $x$  und  $y$  wurden für die Berechnung alle in den Daten erfassten Alter berücksichtigt.

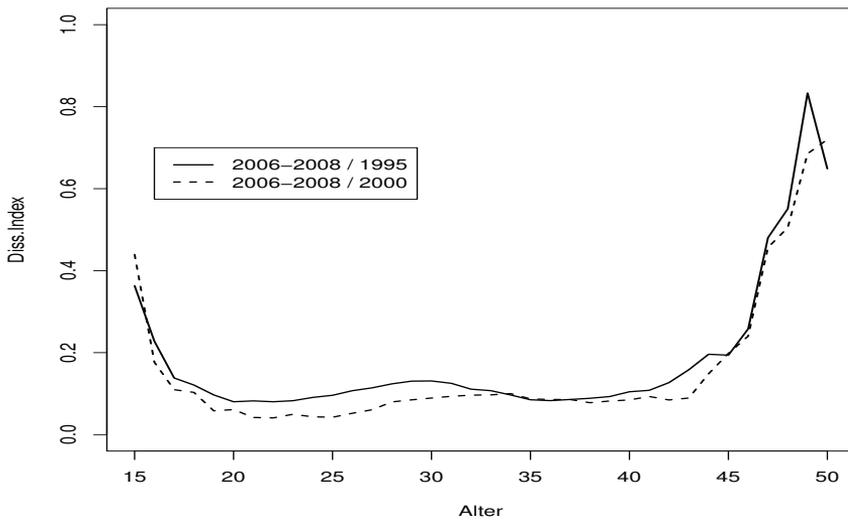
147 Dies gilt auch bei den weiteren Auswertungen in diesem Abschnitt. Es wurde zwar der Versuch unternommen, fehlende Werte durch Schätzwerte zu ersetzen, wobei sich an Sebald (2008: 139ff.) orientiert wurde. Allerdings liefern die von ihr besprochenen Ansätze bezogen auf konditionale Verteilungen  $\phi(y|x)$  relativ unplausible Ergebnisse. Bezogen auf die Altersverteilung von Vätern unabhängig vom Alter von Müttern liegen nur geringfügige Unterschiede zwischen der Verteilung mit und der Verteilung ohne Schätzung fehlender Werte vor und ähnliches gilt für die Altersdifferenz zwischen Mann und Frau (s. auch Sebald 2008: 147ff.). Aus diesen Gründen wird von einer Schätzung fehlender Werte im weiteren abgesehen.

Abbildung 6.1: Dissimilaritätsindex für  $\phi(y|x)$  für einzelne  $x$ , berechnet zwischen 2008 und 2007, 2008 und 2006 sowie 2007 und 2006



Quelle: Statistik der Geburten; eigene Berechnung

Abbildung 6.2: Dissimilaritätsindex für  $\phi(y|x)$  für einzelne  $x$ , berechnet zwischen 2006/2008 und 1995 sowie 2000



Quelle: Statistik der Geburten; eigene Berechnung

denen des Jahres 2000 verglichen. Hierfür wurde wieder der Dissimilaritätsindex verwendet. Die Auswertungen wurden auf eheliche Geburten beschränkt, da für das Jahr 1995 keine Informationen zum Alter von Vätern bei nicht-ehelichen Geburten vorliegen.

Die Ergebnisse sind in Abbildung 6.2 zu sehen. Relativ hohe Werte erreicht der Dissimilaritätsindex wieder für Alter unter 20 und über 45 Jahren, wobei für diese Alter für die Jahre 1995 und 2000 nur wenige Beobachtungen vorhanden sind, so dass die Unterschiede in den Verteilungen wieder auf zufällige Schwankungen aufgrund der niedrigen Fallzahlen zurückgeführt werden können. Für Alter von 20 bis etwa 44 Jahren hingegen fallen die Werte des Dissimilaritätsindex wieder niedriger aus und liegen für alle Vergleiche zwischen etwa 0.04 und 0.13. Während also einzelne Verteilungen nur sehr geringe Unterschiede aufweisen, sind diese bei anderen deutlicher ausgeprägt. Als Beispiel seien hier  $\phi(x|25)$  und  $\phi(x|35)$  herausgegriffen. Diese sind für alle einzelnen Erhebungsjahre in den Abbildungen 6.3 und 6.4 dargestellt.  $\phi(x|35)$  fällt für alle Erhebungsjahre relativ ähnlich aus.  $\phi(x|25)$  hingegen weist zwischen den Jahren 2000, 2006, 2007 und 2008 keine großen Unterschiede auf, während 1995 relativ klar heraussticht. Für dieses Jahr ist das durchschnittliche Alter  $M(y|x = 25) = \sum_y y\phi(y|x = 25) = 28.65$ , für 2008 beispielsweise liegt dieser Wert bei 29.46.

Somit lassen sich für einzelne Alter zwar Unterschiede konstatieren, aber von dieser begrenzten Datenlage ausgehend scheint die im letzten Kapitel getroffene Annahme, dass die Verteilungen  $\phi(y|x)$  konstant sind, zumindest für die meisten Alter  $x$  gerechtfertigt zu sein.<sup>148</sup>

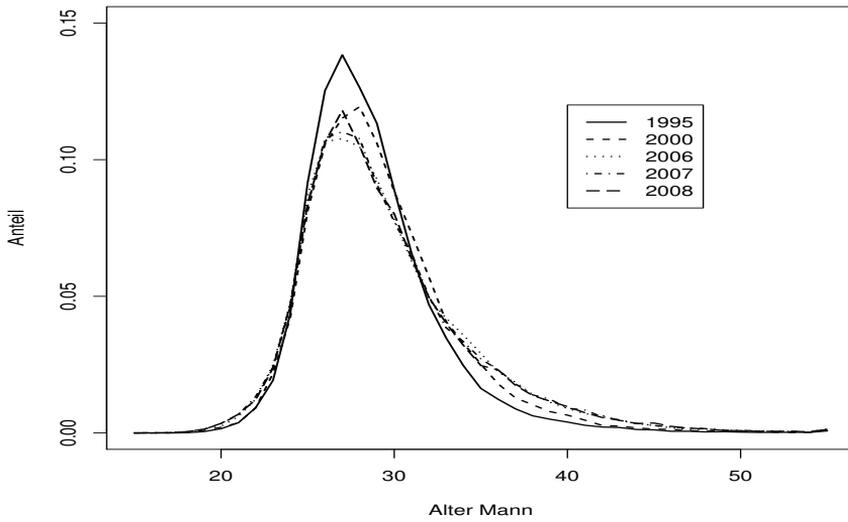
#### 6.1.4 Vergleich der Ergebnisse für West- und Ostdeutschland

Der Vergleich der konditionalen Verteilungen der Alter der Väter zwischen Ost- und Westdeutschland wurde mit den zusammengefassten Daten der Jahre 2006 bis 2008 durchgeführt. Hierbei wurden wieder alle Beobachtungen berücksichtigt, bei denen Angaben zum Alter des Vaters vorlagen. Abermals wurde für einzelne  $\phi(y|x)$  der Dissimilaritätsindex berechnet. Hier zeigten sich sehr ähnliche Ergebnisse wie bei den vorherigen Vergleichen: für den Altersbereich von 20 bis 40 Lebensjahren liegen die Werte des Dissimilaritätsindex zwischen 0.02 und 0.04, während für Alter außerhalb dieses Bereiches teils wesentlich höhere Werte erreicht werden. Zwar sollte der zusammengefasste Datensatz nicht mehr allzu stark von zufälligen Schwankungen beeinflusst sein, allerdings ist die Zahl an ostdeutschen Geburten insgesamt deutlich niedriger als die Zahl der Beobachtungen für Westdeutschland, so dass bei wenig geburtenintensiven Altern für Ostdeutschland wieder sehr niedrige Fallzahlen vorliegen, die zu zufälligen Schwankungen führen. Beispielsweise sind für die Altersklasse von 48 Jahren für Ostdeutschland lediglich 22 Beobachtungen in den Daten enthalten. Insofern scheint eine Differenzierung zwischen West- und Ostdeutschland nicht notwendig.

---

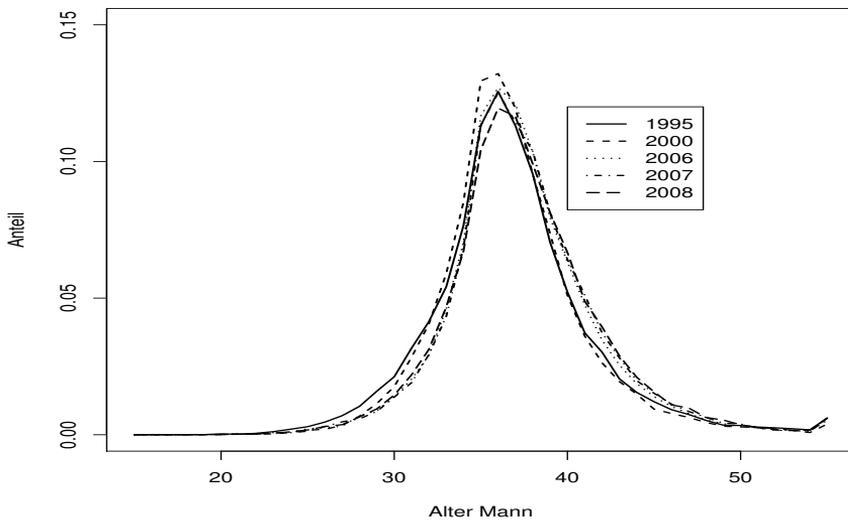
148 Stärkere Veränderungen der konditionalen Verteilungen, die außerhalb des durch die Daten abgegrenzten Zeitraums stattfanden, können allerdings nicht ausgeschlossen werden und erscheinen in Hinblick auf die in Unterabschnitt 6.2.2 noch vorzustellenden Ergebnisse nicht unplausibel.

Abbildung 6.3:  $\phi(y|25)$  für die Jahre 1995, 2000, 2006, 2007 und 2008



Quelle: Statistik der Geburten; eigene Berechnung

Abbildung 6.4:  $\phi(y|35)$  für die Jahre 1995, 2000, 2006, 2007 und 2008



Quelle: Statistik der Geburten; eigene Berechnung

### 6.1.5 Ergebnisse und Glättung der Ergebnisse

Zur Bestimmung von  $\phi(y|x)$  wurden schließlich also die Daten der Jahre 2006, 2007 und 2008 verwendet, wobei fehlende Werte nicht berücksichtigt wurden. Aus Gründen des Datenschutzes und aufgrund geringer Fallzahlen wurden alle Alter  $y$  über 54 Jahren in einer Altersklasse von 55 und mehr Jahren zusammengefasst. Bei Frauen wurden die Altersklasse von 49 Jahren und die Altersklasse von 50 und mehr Jahren zusammengefasst, da altersspezifische Geburtenraten für Frauen nur für Alter von 15 Jahren bis hin zu einer Altersklasse von 49 und mehr Jahren vorliegen (s. Anhang A.3).

Die resultierenden Verteilungen  $\phi(y|x)$  sind für ausgewählte Alter in Abbildung 6.5 dargestellt. Für die Alter von 20 bis 40 Jahren lässt sich erkennen, dass die Verteilungen relativ glatt sind. Für Alter die Alter von 15 und 45 Jahren zeigen sich Unregelmäßigkeiten sowie beim Alter von 45 Jahren ein „Deckeneffekt“, der durch die nach oben offene Altersklassen von 55 Jahren verursacht wird.

Um diese Unregelmäßigkeiten zu beseitigen, wurden die Verteilungen  $\phi(y|x)$  geglättet.<sup>149</sup> Die resultierenden Verteilungen  $\hat{\phi}(y|x)$  finden sich für ausgewählte Alter in Abbildung 6.6.

Im Wesentlichen lässt sich erkennen, dass der Effekt der Glättung bei Altern von 20 bis 40 Jahren sehr gering ist und die Verteilungen kaum verändert werden. Deutlichere Effekte zeigen sich bei den Altern 15 und 45. Beispielsweise verläuft  $\hat{\phi}(y|15)$  wesentlich regelmäßiger als  $\phi(y|15)$ . Zudem ist der Deckeneffekt bei  $\hat{\phi}(y|45)$  nicht mehr so deutlich ausgeprägt. Berechnet man für alle möglichen Alterskombinationen  $x$  und  $y$  die absolute Differenz  $|\hat{\phi}(y|x) - \phi(y|x)|$ , beträgt die höchste Differenz 0.06, also 6 Prozentpunkte. Der Median der absoluten Differenzen liegt bei 0.06 Prozentpunkten und das dritte Quartil bei 0.2 Prozentpunkten. Dies unterstützt die Einschätzung, dass die meisten Verteilungen  $\phi(y|x)$  durch die Glättung nur geringfügig verändert werden.

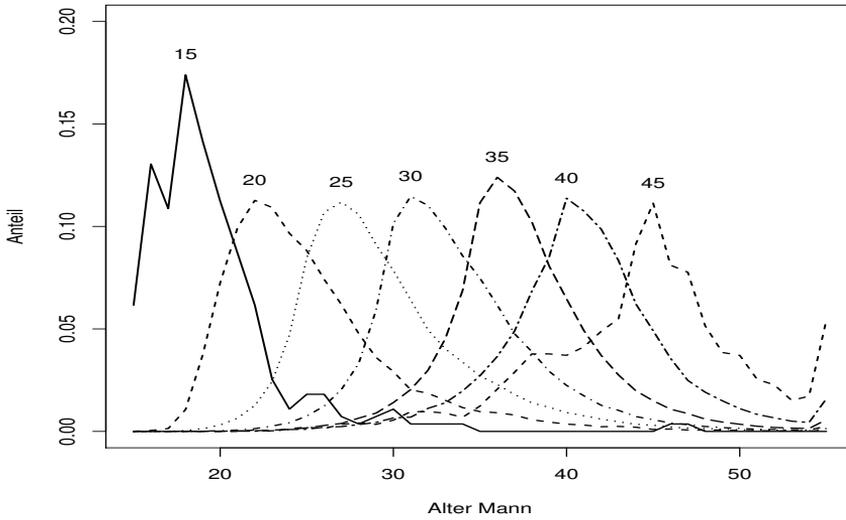
## 6.2 Geburtenraten für westdeutsche Männer ab 1951

Altersspezifische Geburtenraten für westdeutsche Männer wurden aus den im vorausgegangenen Abschnitt beschriebenen Verteilungen beginnend ab dem Jahr 1951 bis einschließlich 2008 berechnet. Die Berechnung erfolgte entsprechend der Ausführungen in 5.3.3. Einige ergänzende Anmerkungen finden sich in Anhang A.3. Die Beschränkung auf westdeutsche Männer erfolgte zum einen, um die Auswertungen in diesem Abschnitt übersichtlich zu halten.<sup>150</sup> Zum anderen wird auch die Validierung des Modells im nächsten Kapitel lediglich für Westdeutschland durchgeführt (s. Abschnitt 7.1 für eine Begründung).

149 Hierzu wurde wiederholt ein gleitender Mittelwert angewendet. Die Glättung erfolgte nicht für einzelne  $\phi(y|x)$  getrennt, sondern für  $h(x, y) = \phi(y|x)$ , so dass auch über „nebeneinander“ liegende Alter von Frauen geglättet wurde. Um  $\sum_y \hat{\phi}(y|x) = 1$  zu gewährleisten, wurden die geglätteten Werte  $\hat{\phi}(y|x)$  jeweils durch ihre Summe geteilt.

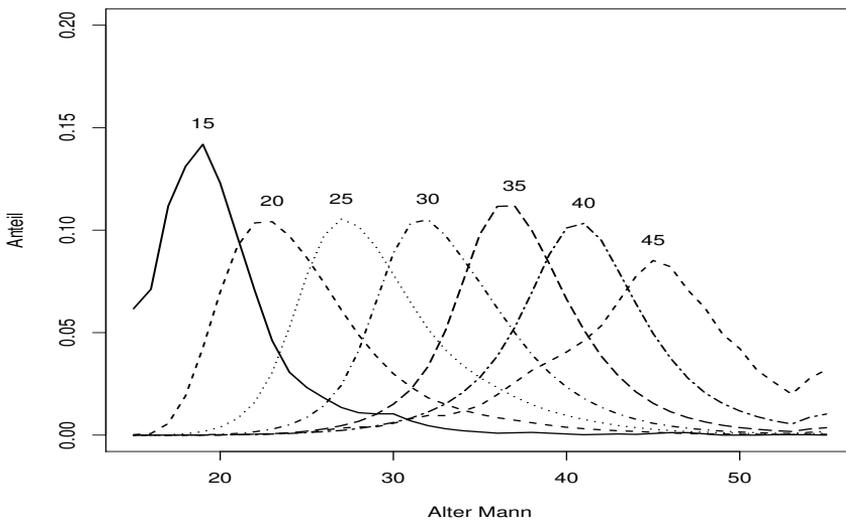
150  $\phi(y|x)$  wurde zwar ausgehend von Daten für die gesamte BRD ermittelt, allerdings zeigte sich beim Vergleich der Ergebnisse der getrennten Schätzung für Ost- und Westdeutschland in Unterabschnitt 6.1.4 sehr ähnliche Resultate, weshalb hier davon ausgegangen wird, dass  $\phi(y|x)$  auch nur auf Westdeutschland beschränkt verwendet werden kann.

Abbildung 6.5: Verteilungen  $\phi(y|x)$  für  $x = 15, 20, 25, 30, 35, 40, 45$  berechnet aus den zusammengefassten Daten 2006 bis 2008



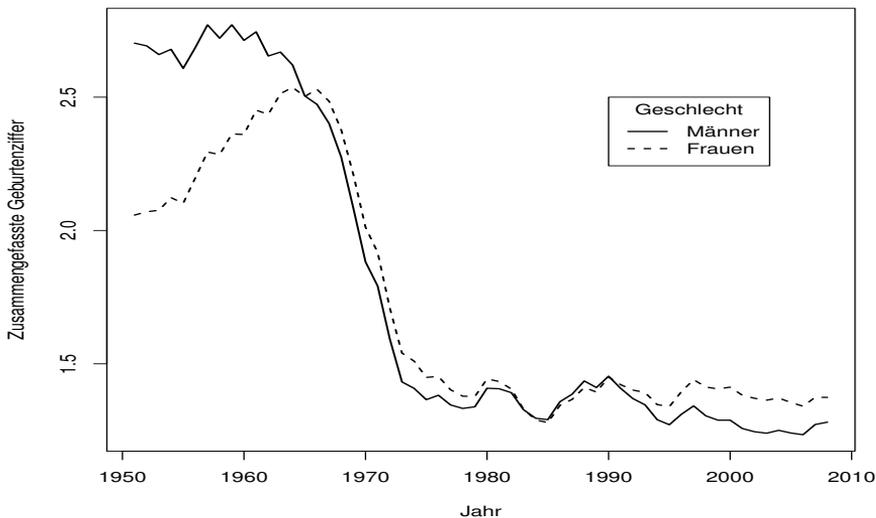
Quelle: Statistik der Geburten; eigene Berechnung

Abbildung 6.6: Geglättete Verteilungen  $\hat{\phi}(y|x)$  für  $x = 15, 20, 25, 30, 35, 40, 45$  berechnet aus den zusammengefassten Daten 2006 bis 2008



Quelle: Statistik der Geburten; eigene Berechnung

Abbildung 6.7: Zusammengefasste Geburtenziffer für Männer und Frauen für die Jahre von 1951 bis 2008



Quelle: Statistisches Bundesamt; eigene Berechnung

Die Ergebnisse werden im folgenden aus mehreren Gründen ausführlich diskutiert. Zunächst erlaubt eine genaue Auswertung der Resultate eine Einschätzung der Verlässlichkeit der hier verwendeten Lösung des Zwei-Geschlechter-Problems und somit eines wichtigen Teilaspekts des in Kapitel 5 vorgestellten Modells. Dies gilt insbesondere in Hinblick darauf, dass davon ausgegangen wird, dass  $\phi(y|x)$  konstant ist. Darüber hinaus ist für die Validierung im nächsten Kapitel von Bedeutung, ob Fertilität von Männern hinreichend gut rekonstruiert werden kann. Ist dies nicht der Fall, sind bei einem Vergleich von Modellrechnungen und empirischen Daten Abweichungen zu erwarten, die allerdings nicht die hier durchgeführte Berechnung von Verwandtschaft in Frage stellen müssen, sondern durch Fehler bei der Rekonstruktion erklärt werden können. Schließlich sind Ergebnisse zu männlicher Fertilität als solche von Interesse, da es zu dieser kaum Daten gibt. Einige der wenigen Quellen zu männlicher Fertilität werden im nächsten Abschnitt zum Vergleich herangezogen, um zu einer Einschätzung der Plausibilität der Ergebnisse zu kommen.

### 6.2.1 Ergebnisse

In Abbildung 6.7 ist die zusammengefasste Geburtenziffer jeweils für westdeutsche Männer und Frauen für die Jahre 1951 bis 2008 dargestellt, wie sie sich aus amtlichen Daten und der hier vorgenommenen Rekonstruktion ergibt.<sup>151</sup> Ab dem Jahr 1965 ist der Verlauf grob parallel, wobei die zusammengefasste Geburtenziffer von Männern zumeist etwas unter der von Frauen

<sup>151</sup> Im weiteren Verlauf dieses Abschnittes wird größtenteils auf den Hinweis verzichtet, dass sich die zugrunde gelegten Daten und die aus diesen resultierenden Ergebnisse auf Westdeutschland beziehen.

liegt. Vor 1965 gibt es hingegen erhebliche Differenzen, wobei der Wert für Männer deutlich über dem für Frauen liegt. Beispielsweise beträgt die zusammengefasste Geburtenziffer für Männer 1951 etwa 2.70, während der Wert für Frauen bei ungefähr 2.06 liegt. Bis zum Anfang der sechziger Jahre bleibt der Wert für Männer auf einem hohen Niveau, während der Wert für Frauen hingegen deutlich ansteigt.

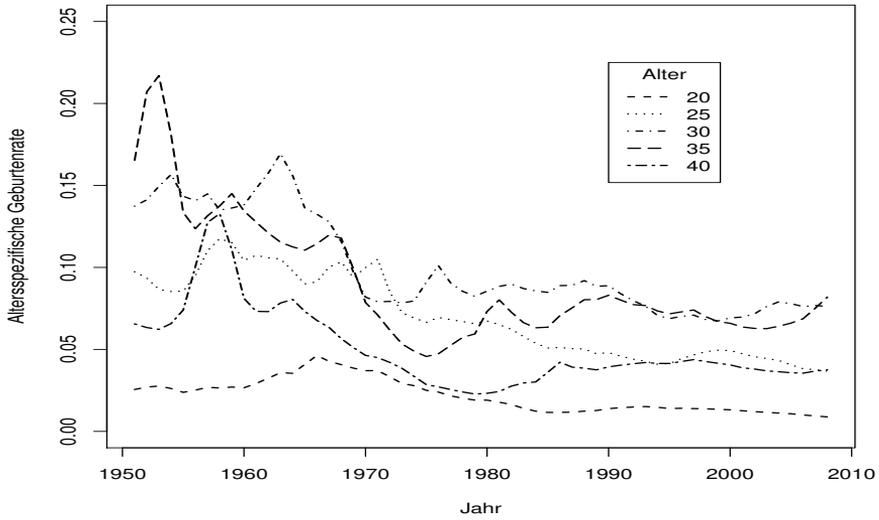
Die Abbildungen 6.8 und 6.9 zeigen die altersspezifischen Geburtenraten von Männern und Frauen in den Altern von 20, 25, 30, 35 und 40 Jahren für den Zeitraum von 1951 bis 2008. Bis etwa zum Beginn der siebziger Jahre verlaufen die altersspezifischen Geburtenraten der Männer deutlich unregelmäßiger als die der Frauen. Beispielsweise erreicht die Geburtenrate für 35-jährige Männer 1953 einen vergleichsweise hohen Wert von circa 0.22 und fällt danach relativ stark ab. 1958 ergibt sich wiederum ein sehr hoher Wert bei der Geburtenrate für die 40-jährigen Männer, der die vorausgegangenen und folgenden Werte deutlich übersteigt. Dabei beziehen sich diese beiden altersspezifischen Raten in den genannten Jahren bei Vernachlässigung von Migration und Mortalität ungefähr auf die gleichen Personen, die zu den Geburtsjahrgängen 1918 und 1919 gehören. Diese Jahrgänge „bewegen“ sich relativ deutlich durch die altersspezifischen Geburtenraten der Männer. Gleiches zeigt sich auch für andere Jahrgänge, allerdings nur in abgeschwächter Form.

In Abbildung 6.10 sind die altersspezifische Geburtenraten für Männer für die Jahre 1951, 1960, 1970, 1980 und 1990 eingezeichnet. Das Altersprofil des Jahres 1951 zeigt nochmals deutlich die bereits erwähnten Unregelmäßigkeiten. Dies betrifft insbesondere die Alter 32, 33 und 34, die klar herausstechen, während die Geburtenraten der Alter 30 und 31 deutlich niedriger sind. Für das Jahr 1960 tritt der Bruch in den altersspezifischen Geburtenraten zeitversetzt auf, wobei die Geburtenraten der Alter 39 und 40 niedriger als die der Alter 41, 42 und 43 ausfallen. Die Altersprofile der Jahre 1970 und 1980 weisen eher kleinere Schwankungen auf, während das Altersprofil für das Jahr 1990 relativ gleichmäßig verläuft.

Dies gilt auch für die Altersprofile folgender Jahre, von denen die Altersprofile der Jahre 2003 bis 2008 in Abbildung 6.11 zu sehen sind. Zusätzlich zu den altersspezifischen Geburtenraten für Männer sind die altersspezifischen Geburtenraten der selben Jahre für Frauen in grau eingezeichnet. Vergleicht man die Geburtenraten zwischen den Geschlechtern, zeigt sich, dass die Altersprofile für Männer in ähnlichem Maße regelmäßig verlaufen wie die der Frauen. Ansonsten ist die Fertilität von Männern insgesamt niedriger, erstreckt sich über mehr und spätere Alter und das Maximum der Geburtenraten wird in späteren Altern erreicht als bei Frauen. Ferner ist zu erkennen, dass alle Geburten von Männern in einem Alter von 55 und mehr Jahren den 55-jährigen zugespielt werden (s. Unterabschnitt 6.1.5), wodurch die Geburtenrate für dieses Alter etwas über der der 54-jährigen liegt. Von diesem Effekt abgesehen fallen die Geburtenraten für Männer nach dem Maximum ähnlich schnell ab wie bei Frauen und sind ab einem Alter von etwa 45 Jahren relativ niedrig.

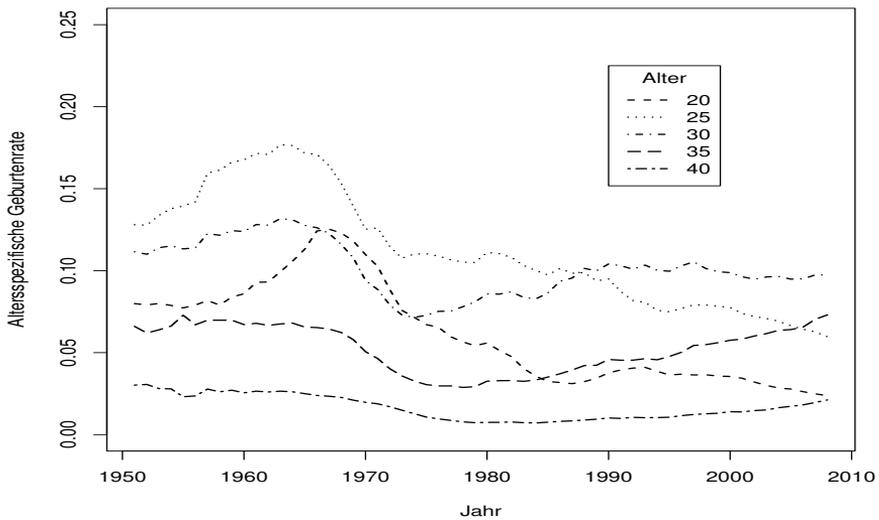
Für diese Abnahme lassen sich mehrere Erklärungen finden lassen. Eine Ursache dürfte zum einen sein, dass auch trotz zuweilen gegenteiliger Annahmen in der Literatur die Fekundität von Männern mit fortschreitendem Alter abnimmt (vgl. Hagemann & Haidl 2001). Zum anderen sind Männer durchschnittlich etwa drei Jahre älter als ihre Partnerinnen, so dass deren mit dem Alter abnehmende Fekundität ebenfalls limitierend wirkt (vgl. Huinink 1988: 371; Schmitt 2005: 21). Dieser Altersunterschied zeigt sich auch im insgesamt verschobenen Verlauf der Altersprofile. Schließlich lässt sich vermuten, dass zusätzlich soziale Normen über

Abbildung 6.8: Altersspezifische Geburtenraten für Männer in den Altern 20, 25, 30, 35 und 40 von 1951 bis 2008



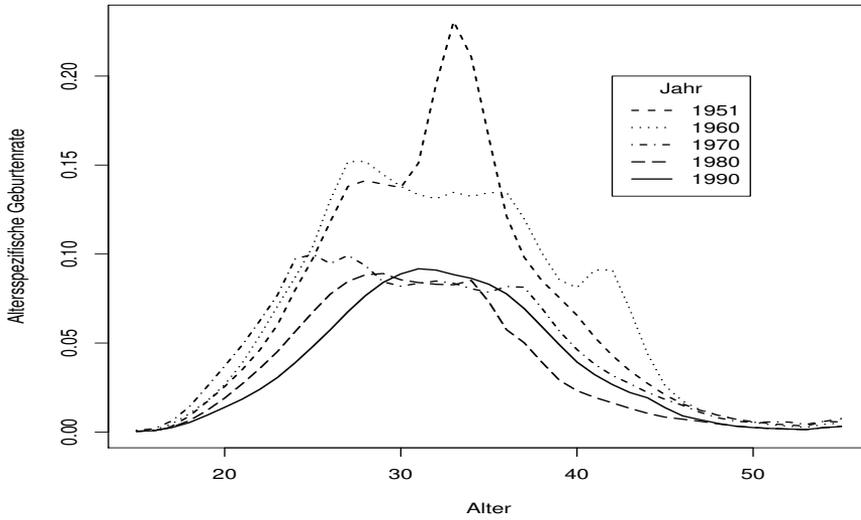
Quelle: eigene Berechnung

Abbildung 6.9: Altersspezifische Geburtenraten für Frauen in den Altern 20, 25, 30, 35 und 40 von 1951 bis 2008



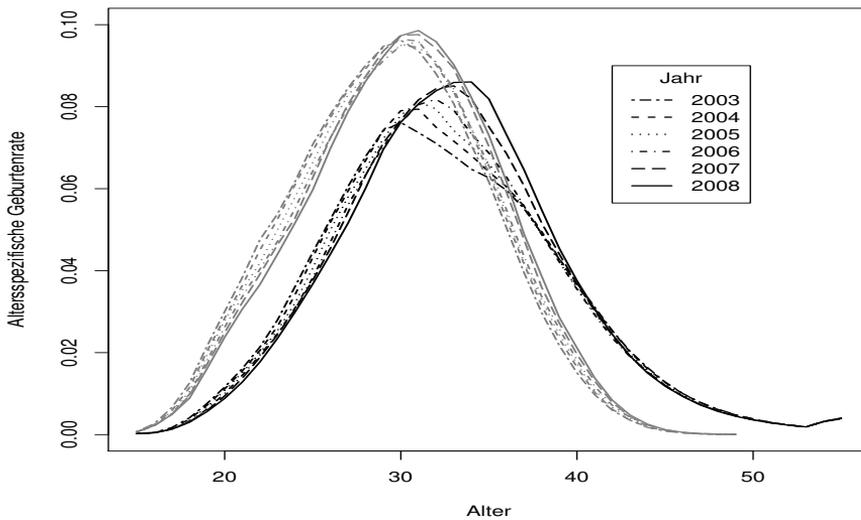
Quelle: Statistisches Bundesamt; eigene Darstellung

Abbildung 6.10: Altersspezifische Geburtenraten für Männer für die Jahre 1951, 1960, 1970, 1980 und 1990



Quelle: eigene Berechnung

Abbildung 6.11: Altersspezifische Geburtenraten für Männer (schwarze Linien) und Frauen (graue Linien) für die Jahre 2003 bis 2008



Quelle: Statistisches Bundesamt; eigene Berechnung

„angemessene“ Alter bei Vater- und Mutterschaft einen nicht unerheblichen Einfluss ausüben (vgl. Billari et al. 2011).

### 6.2.2 Mögliche Ursachen für die Ergebnisse

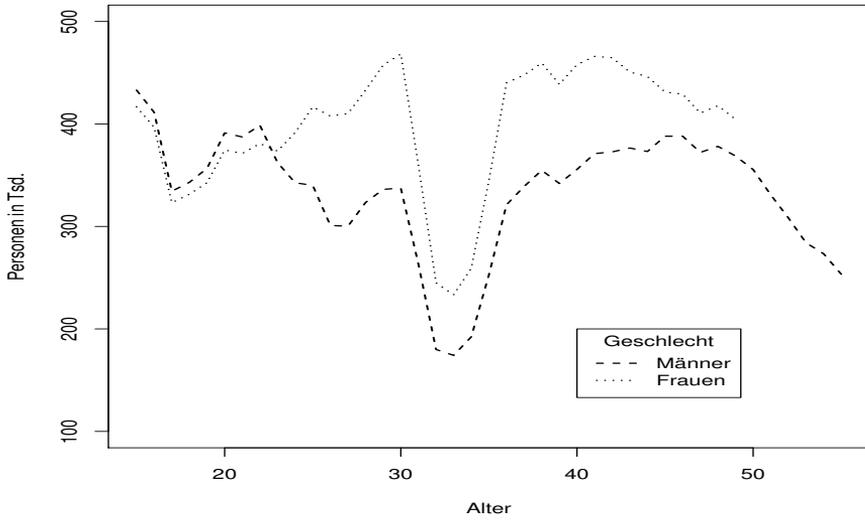
Insbesondere für die fünfziger Jahre zeigen sich bei den bisher betrachteten Werten deutliche Unregelmäßigkeiten bei der Fertilität der Männer und ausgeprägte Unterschiede zwischen den Geschlechtern. Zwar ist weder ein regelmäßiger Verlauf noch Gleichheit beziehungsweise Ähnlichkeit zwischen den Geschlechtern zu erwarten (vgl. bspw. Karmel 1947; Poston et al. 2006: 872), nichtsdestotrotz stellt sich die Frage nach den Ursachen für diese Befunde. Den Ausführungen in Anhang C folgend könnte dies an der Altersstruktur der männlichen und weiblichen Bevölkerung sowie den Verhältnissen bestimmter Altersklassen von Männern und Frauen liegen. Für die Nachkriegszeit wird auch in der Literatur davon ausgegangen, dass die Bevölkerungsstruktur die Fertilität deutlich beeinflusst hat (vgl. bspw. Dinkel & Milenovic 1992, 1993). Mögliche Hinweise hierauf finden sich in den Abbildungen 6.12 und 6.13.

In Abbildung 6.12 ist der Bestand an Personen in einzelnen Altern der reproduktiven Phase nach Geschlecht für das Jahr 1951 abgetragen. Für die Alter von 15 bis 23 Jahren gibt es mehr männliche als weibliche Personen. Für alle darauf folgenden Alter gibt es als Folge des zweiten Weltkriegs deutlich mehr Frauen als Männer (vgl. Dinkel & Milenovic 1993: 156). Ferner fällt der deutliche „Knick“ bei Altern von etwa 31 bis 34 Jahren auf, die ungefähr den Geburtsjahrgängen 1916 bis 1919 entsprechen. Ursache hierfür sind relativ starke Geburtenausfälle in diesen Jahren aufgrund des ersten Weltkriegs (vgl. Marschalck 1984: 67f. u. 84f.).

Die sich durch diese Sondereinflüsse ergebende „verzerrte“ Altersstruktur kann in Kombination mit Verteilungen  $\phi(y|x)$ , die sich aus einer „unverzerrten“ Struktur ergeben, eventuell zu Problemen bei der Rekonstruktion der Geburtenraten von Männern führen. Betrachtet man beispielsweise Geburten von 30-jährigen Frauen, werden diese durch die aus den Auswertungen aus dem vorherigen Abschnitt gewonnene Verteilung  $\phi(y|30)$  Männern zugespielt, die durchschnittlich etwa 33.5 Jahre alt sind. Bei der in Abbildung 6.12 zu sehenden Altersstruktur wird ersichtlich, dass dies zu hohen Geburtenraten von Männern führen kann und diese Geburten eventuell anders verteilt werden müssten.

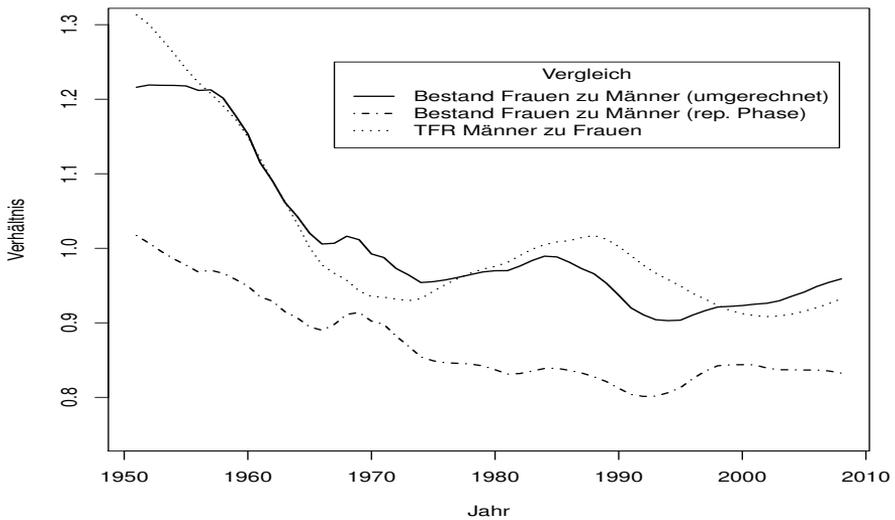
Einen Überblick über das numerische Verhältnis von Männern und Frauen von 1951 bis 2008 liefert Abbildung 6.13. In dieser ist das Verhältnis der zusammengefassten Geburtenziffer für Männer zur zusammengefassten Geburtenziffer für Frauen und das Verhältnis der Zahl von Frauen in der reproduktiven Phase (Alter 15 bis 49) zur Zahl von Männern in der reproduktiven Phase (Alter 15 bis 55) zu sehen. Ferner wurde die Zahl der Männer in der reproduktiven Phase zusätzlich als  $\sum_x \sum_y n_{(y,m),t} \phi(y|x)$  berechnet, also entsprechend der Zuordnung von Männern zu Frauen über  $\phi(y|x)$ . Die Zahl der Frauen in der reproduktiven Phase wurde dann ebenfalls zu diesem Wert ins Verhältnis gesetzt. In der Abbildung ist zu erkennen, dass sich der letztgenannte Kennwert und das Verhältnis der zusammengefassten Geburtenziffern ungefähr gleich entwickeln, insbesondere im Zeitraum von etwa 1955 bis 1965, in dem beide Kennzahlen von einem Wert von etwas über 1.2 auf einen Wert von ungefähr 1 sinken. Ab etwa 1970 liegen beide Kennwerte größtenteils zwischen 0.9 und 1. Für die letzten drei Jahre 2006, 2007 und 2008, für die die Verteilungen  $\phi(y|x)$  relativ genau zutreffen sollten, werden

Abbildung 6.12: Altersstruktur der Bevölkerung der BRD in der reproduktiven Phase im Jahr 1951 nach Geschlecht



Quelle: Human Mortality Database; eigene Darstellung

Abbildung 6.13: Ausgewählte Verhältniswerte für Fertilität und Bestand von Männern und Frauen 1951 bis 2008



Quelle: Statistisches Bundesamt; Human Mortality Database; eigene Berechnung

für die Verhältnisse der zusammengefassten Geburtenziffern Werte von 0.92 bis 0.93 erreicht, für das umgerechnete Verhältnis von Frauen zu Männern in der reproduktiven Phase Werte von 0.95 bis 0.96. Das eigentliche Verhältnis von Frauen zu Männern in der reproduktiven Phase weist aufgrund von Unterschieden im Nenner deutlich niedrigere Werte auf und zeigt einen etwas regelmäßigeren Verlauf, der den Verlauf des Verhältnisses der zusammengefassten Geburtenziffern aber zumindest grob wiedergibt.

Insgesamt hat sich über den betrachteten Zeitraum hinweg das numerische Verhältnis von Männern und Frauen stark verändert. Neben den bereits oben genannten Sondereinflüssen ist zum einen noch von einem Einfluss von Migration auszugehen (vgl. auch Zhang 2011: 18). Zum anderen dürfte das Verhältnis auch durch Unterschiede in den Größen aufeinander folgender Jahrgänge beeinflusst werden, da Männer durchschnittlich mit etwa 3 Jahre jüngeren Frauen Kinder bekommen, die somit also tendenziell zu einem „jüngeren“ Geburtsjahrgang gehören.

### 6.2.3 *Bewertung der Ergebnisse*

Die sich aus den Berechnungen ergebenden Werte der zusammengefassten Geburtenziffer scheinen insgesamt nicht unplausibel, wenn man zusätzlich die Entwicklung der Bevölkerung in der reproduktiven Phase betrachtet. Allerdings ist die Bestimmung altersspezifischer Raten für einzelne Jahre nicht unproblematisch, insbesondere für den Anfang des betrachteten Zeitraums von 1951 bis etwa 1965. Zudem ist zu bedenken, dass die hier bisher vorgenommenen Auswertungen zwar Rückschlüsse auf die „technischen“ Ursachen der Ergebnisse zulassen, wie beispielsweise die Bevölkerungsstruktur. Mögliche Veränderungen in der Partnerwahl, die als Reaktion auf „ungünstige“ Geschlechterverhältnisse bestimmter Jahrgänge denkbar sind, können aber nicht berücksichtigt werden. Zudem wird die Fertilität sowohl von Männern als auch von Frauen nicht allein durch die numerische Verfügbarkeit des anderen Geschlechts bestimmt (vgl. Zhang 2011: 18).

Für das eigentliche Ziel dieser Arbeit, der Vorausberechnung von linearer Verwandtschaft, erscheinen diese Resultate eher unproblematisch. Allerdings ist für den im nächsten Kapitel angestrebten Vergleich von Modellrechnungen und empirischen Daten die Verlässlichkeit der Resultate zur männlichen Fertilität ein deutlich größeres Problem, weshalb im folgenden Abschnitt ein Vergleich der hier vorgestellten Ergebnisse zur männlichen Fertilität mit Ergebnissen aus anderen Quellen vorgenommen wird.

## 6.3 Vergleiche mit Ergebnissen aus anderen Quellen

### 6.3.1 Mögliche Vergleichsquellen

Die Zahl der Quellen, aus denen Vergleichsmaterial für die Ergebnisse zu männlicher Fertilität gewonnen werden kann, ist sehr eingeschränkt.<sup>152</sup> Vergleichswerte für Deutschland können lediglich im *Demographic Yearbook 2008* des UN Department of Economic and Social Affairs (2010) gefunden werden.<sup>153</sup>

Daneben können auch internationale Ergebnisse zur Fertilität von Männern herangezogen werden. Zwar scheint ein direkter internationaler Vergleich von altersspezifischen Geburtenraten oder zusammengefassten Geburtenziffern zur Einschätzung der Qualität der Ergebnisse wenig sinnvoll. Allerdings lassen sich die Verhältnisse von zusammengefassten Geburtenziffern von Männern und Frauen auf die Frage hin untersuchen, ob bestimmte Verhältniswerte besonders häufig – also in gewissem Sinne „üblich“ – sind. Hier kann ebenfalls auf Ergebnisse des UN Department of Economic and Social Affairs (1966, 1983, 2001) zurückgegriffen werden. Ferner findet sich bei Brouard (1977) eine Rekonstruktion altersspezifischer Geburtenraten französischer Männer für den Zeitraum von 1899 bis 1974.

### 6.3.2 Ergebnisse der Vergleiche

Im *Demographic Yearbook 2008* finden sich altersspezifische Geburtenraten für Männer für das Jahr 2007 (vgl. UN Department of Economic and Social Affairs 2010: 379f.). Allerdings sind diese nicht für Einzelalter, sondern gruppierte Altersklassen von jeweils fünf Jahren ausgewiesen und beziehen sich zudem nicht nur auf Westdeutschland, sondern auf die gesamte BRD. Als zusammengefasste Geburtenziffer erhält man aus diesen Angaben den Wert 1.2315.<sup>154</sup> In der Modellrechnung ergibt sich ein Wert von ungefähr 1.2713, woraus eine Differenz von circa  $-0.0399$  resultiert. Nimmt man an, dass die Fertilität ostdeutscher Männer in Periodenbetrachtung niedriger als die von westdeutschen Männern war, so wie es sich 2007 für ost- und westdeutsche Frauen verhielt, erscheinen Höhe und Vorzeichen der Differenz plausibel.

In Abbildung 6.14 sind die Verhältnisse von zusammengefassten Geburtenziffern von Männern und Frauen für Westdeutschland, England und Wales, Dänemark, Frankreich und Norwegen abgetragen. Die Daten zu England und Wales, Dänemark, Norwegen und ein Teil der französischen Daten stammen aus diversen Ausgaben des *Demographic Yearbook*

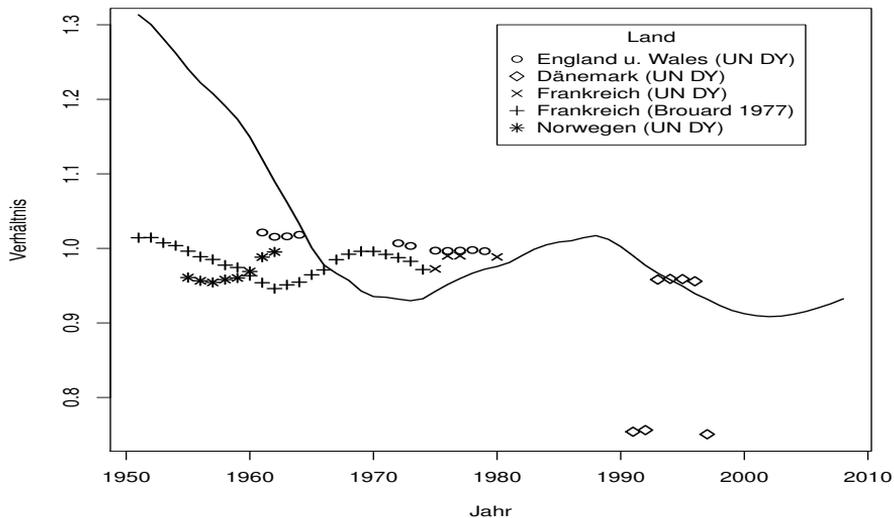
---

152 Für eine entsprechende Einschätzung für Deutschland s. Hank & Tölke (2005: 8); zum internationalen Mangel an Daten zu männlicher Fertilität vgl. Poston et al. (2006: 871f.) sowie Coleman (2000).

153 Zwar liegen Auswertungen der Statistik der Geburten zur Fertilität von Männern für den Zeitraum von 1991 bis 2005 vor (Sebald 2008). Allerdings werden die berechneten altersspezifischen Geburtenraten nur graphisch dargestellt und es fehlen Tabellen, aus denen Vergleichswerte gewonnen werden könnten. Weitere Arbeiten, die sich mit der Fertilität von Männern befassen, setzen einen Schwerpunkt auf Kinderlosigkeit und weisen keine Geburtenraten, durchschnittliche Kinderzahlen oder ähnliches aus (s. bspw. Schmitt 2005; Onnen-Isemann 2003).

154 Zur Berechnung der zusammengefassten Geburtenziffer aus gruppierten Daten s. Preston et al. (2001: 95).

Abbildung 6.14: Verhältnisse von zusammengefassten Geburtenziffern von Männern und Frauen für Westdeutschland, England und Wales, Dänemark, Frankreich und Norwegen



Quelle: Brouard (1977); UN Department of Economic and Social Affairs (1966, 1983, 2001); eigene Berechnung

(vgl. UN Department of Economic and Social Affairs 1966, 1983, 2001).<sup>155</sup> Hierbei handelt es sich in der Regel um Sonderauswertungen amtlicher Daten, so dass Ergebnisse nur für ausgewählte Jahre vorliegen (vgl. hierzu Coleman 2000: 42f.). Für Frankreich liegt ferner eine Rekonstruktion männlicher Fertilität für den Zeitraum von 1899 bis 1974 vor, die von Brouard (1977) erstellt wurde. Diese Rekonstruktion folgt einem Vorgehen, das dem hier verwendeten relativ ähnlich ist (s. auch Coleman 2000: 47f.).

Aufgrund der unterschiedlichen demographischen Entwicklung in den einzelnen Ländern sind einzelne Länderwerte für ein bestimmtes Jahr nicht direkt miteinander vergleichbar. Nichtsdestotrotz lässt sich für alle Länder bis auf Westdeutschland relativ deutlich ein Wertebereich des Verhältnisses von männlicher zu weiblicher zusammengefasster Geburtenziffer erkennen, der zwischen 0.95 und 1.02 liegt. Für Dänemark finden sich in den neunziger Jahren drei relativ niedrige Werte von jeweils ungefähr 0.75 (1991, 1992 und 1997), die auf niedrige Werte der zusammengefassten Geburtenziffer für Männer zurückzuführen sind. Die Ursache hierfür ist unklar. Möglicherweise sind in diesem Fall die Angaben aus dem *Demographic Yearbook* ungenau oder fehlerhaft.<sup>156</sup> Es lässt sich aber festhalten, dass das Verhältnis in West-

<sup>155</sup> Eine Auswertung aller im *Demographic Yearbook* 2001 enthaltenen Daten zu männlicher Fertilität, die insgesamt 48 Länder und die Jahre 1990 bis 1998 umfassen, findet sich bei Zhang (2011: 46f.). Auf eine so detaillierte Auswertung wird in diesem Unterabschnitt verzichtet, da viele dieser Länder eine deutlich höhere Fertilität aufweisen als die BRD (TFR>2.2), wie beispielsweise Mexiko, und zudem oftmals nur Informationen zu ehelichen Geburten vorhanden sind.

<sup>156</sup> Die zusammengefassten Geburtenziffern für Frauen für diese Jahre stimmen zwischen *Demographic Yearbook* und Angaben aus der Online-Datenbank von *Danmarks Statistik* (<http://www.statbank.dk>, abgerufen am 20.04.2011) relativ genau überein, so dass die geäußerte Vermutung nahe liegt, dass die Angaben zur Fertilität von Männern im *Demographic Yearbook* eventuell nicht korrekt sind.

deutschland ab der Mitte der sechziger Jahre gemessen an diesen Ergebnissen als plausibel erscheint.

Verhältniswerte, die wie in Westdeutschland während der fünfziger Jahre deutlich über 1 lagen, wurden in Frankreich kurz nach dem ersten Weltkrieg erreicht (vgl. Brouard 1977: 1157). Beispielsweise beträgt der Wert für das Jahr 1919 etwa 1.26. Für die Jahre von 1918 bis 1925 liegt der Durchschnitt bei etwa 1.21. Dies erklärt sich durch die hohe Zahl an insbesondere männlichen Kriegstoten während des ersten Weltkriegs, so dass zum Kriegsende eine deutlich verzerrte Altersstruktur der französischen Bevölkerung vorlag (vgl. Coleman 2000: 49).

Insofern scheinen die Ergebnisse für Westdeutschland in den fünfziger Jahren nicht gänzlich unplausibel zu sein. Hierbei ist allerdings zu bedenken, dass die Rekonstruktion von Geburtenraten für Frankreich von Brouard (1977) einem sehr ähnlichen Vorgehen wie dem hier gewählten folgt, so dass auch für die französischen Daten unklar ist, inwieweit sie die „tatsächliche“ historische Entwicklung von männlicher Fertilität wiedergeben.

### 6.3.3 *Bewertung der Vergleiche*

Abschließend lässt sich aufgrund der Vergleiche in Kombination mit den im vorausgegangenen Abschnitt angestellten Auswertungen folgendes festhalten:

- Im internationalen Vergleich scheinen die Werte der zusammengefassten Geburtenziffer für Männer ab Mitte der sechziger Jahre nicht auffällig und relativ plausibel zu sein. Inwieweit die zusammengefasste Geburtenziffer für die fünfziger Jahre und die erste Hälfte der sechziger Jahre als realistisch eingeschätzt werden können, ist hingegen unklar.
- Die periodenspezifischen Altersprofile der männlichen Fertilität sind zumindest für die fünfziger und sechziger Jahre nicht als verlässlich einzustufen. Altersprofile, die sich auf Jahre ab 1995 beziehen, können als verlässlicher gelten als solche Profile, die vor diesem Zeitraum liegen, zumal die Verteilungen  $\phi(y|x)$  ab diesem Jahr untersucht und als Grundlage für die Rekonstruktion verwendet wurden.
- Die Vergleiche verdeutlichen, dass die Rekonstruktion männlicher Fertilität bei einer stark durch Sondereinflüsse beeinflussten Altersstruktur problematisch sein kann (s. auch Anhang C.1).

Hieraus folgt, dass bei der im nächsten Kapitel folgenden Validierung des Ansatzes Vorsicht geboten ist. Abweichungen zwischen Modell und Vergleichsdaten sind aufgrund der gerade genannten Einschränkungen nicht unbedingt auf eine Fehlerhaftigkeit der Modellierung von Verwandtschaft, sondern möglicherweise auf Ungenauigkeiten bei der Rekonstruktion männlicher Fertilität zurückzuführen.

## 7 Validierung des Ansatzes

### 7.1 Vorbemerkungen

Zur Validierung des in Kapitel 5 vorgestellten Ansatzes werden zum einen historische demographische Raten benötigt, zum anderen Vergleichsdaten. Die Vergleichsdaten werden dem Sozio-oekonomischen Panel entnommen, das Informationen zum Vorhandensein von Verwandten enthält. Eine genauere Beschreibung der Daten findet sich im nächsten Abschnitt.

Historische Raten für die BRD sind nur eingeschränkt vorhanden. Durch die beiden Weltkriege gibt es immer wieder Brüche im Vorliegen von statistischen Daten, was insbesondere für die Zeit während und kurz nach dem zweiten Weltkrieg gilt. Hiermit einhergehend gab es etliche Wechsel im Gebietsstand, die die Analyse erschweren. Hierzu zählt insbesondere die Trennung in BRD und DDR, die mit unterschiedlichen demographischen Entwicklungen in den beiden Teilen Deutschlands einherging (vgl. Dinkel 1992), wobei auch nach der Wiedervereinigung Differenzen auftraten (vgl. Kreyenfeld & Konietzka 2004; sowie Abschnitt 3.2).

Um die Validierung übersichtlich zu halten wird daher eine Einschränkung auf Westdeutschland, also das frühere Bundesgebiet, vorgenommen. Anhand der in Kapitel 6 und Anhang A beschriebenen Daten wurden Berechnungen für das Jahr 2009 durchgeführt, wobei diese im weiteren als „Modellrechnungen“ beziehungsweise einfach als „Modellrechnung“ bezeichnet werden. Entsprechend sind aus dem SOEP Ergebnisse für das selbe Jahr zu ermitteln.<sup>157 158</sup>

Bei der Modellrechnung konnten nicht alle möglichen Altersklassen Egos berücksichtigt werden, was auf das eingeschränkte Vorliegen demographischer Raten zurückgeführt werden kann. Werden die in Anhang A beschriebenen demographischen Raten verwendet, beziehen sich die Ergebnisse auf den 1. Januar 2009. Um beispielsweise die Wahrscheinlichkeit einer lebenden Großmutter mütterlicherseits einer Person in einer Altersklasse  $a$  zu berechnen, müsste zunächst das Geburtsjahr dieser Person und die für dieses geltende Verteilung der Alter von Müttern betrachtet werden. Nun müsste in einem zweiten Schritt vom Alter der Mutter  $b$  zum Geburtsjahr dieser zurückgegangen werden, um zu ermitteln, in welchem Alter die Großmutter bei der Geburt der Mutter war. Hierbei müssten teils Jahre betrachtet werden, für die keine Daten vorliegen. Setzt man beispielsweise  $a = 40$  und  $b = 30$ , müsste das Jahr 1937 betrachtet werden.

Dieses Problem betrifft alle bei der Modellrechnung betrachteten Verwandtschaftsgrade. Die erwartete Zahl an Kindern kann für Altersklassen von 0 bis 72 berechnet werden. Da aus den Daten des SOEP nur Rückschlüsse auf die Zahl jemals geborener Kinder möglich sind (s. Unterabschnitt 7.2.2.1), werden diese betrachtet. Die erwartete Zahl lebender Enkel kann ebenfalls für Altersklassen von 0 bis 72 Jahren berechnet werden. Die Wahrscheinlichkeit

---

157 Hierbei ist eigentlich zu beachten, dass sich die Ergebnisse des Modells, wie in Kapitel 5 erläutert, immer auf den ersten Januar eines Jahres beziehen. Die Erhebung des SOEP fand allerdings während des Jahres 2009 statt (vgl. Siegel et al. 2010: 38), so dass sich Modellrechnung und SOEP nicht auf den selben Zeitpunkt beziehen. Da aber ein Großteil der Befragungen bis Mitte 2009 stattfand (vgl. ebd.: 38), wird davon ausgegangen, dass dieser Unterschied vernachlässigt werden kann.

158 Die Aufbereitung der demographischen Raten ist ebenfalls in Anhang A beschrieben.

lebender Eltern kann für das Jahr 2009 nur für Altersklassen von 0 bis 57 berechnet werden. Das Vorhandensein von Großeltern lässt sich eigentlich nur für die Altersklasse 0 betrachten. Um die Berechnungen für einige weitere Altersklassen durchführen zu können, wurde die reproduktive Phase für diese Berechnungen bei Frauen auf Alter von 15 bis 45 Jahren begrenzt, bei Männern auf Alter von 15 bis 50 Jahren. Bei Frauen wurden Geburtenraten in höheren Altern ignoriert, bei Männern wurden bei den Verteilungen  $\phi(y|x)$  höhere Alter entfernt und die verbleibenden Anteilswerte durch ihre Summe geteilt.

Ferner wird lediglich die Zahl der Verwandten zwischen Modellrechnung und SOEP verglichen, nicht deren (durchschnittliches) Alter. Diese Entscheidung wurde getroffen, da der Schwerpunkt der Arbeit auf der Zahl der Verwandten liegt und die Validierung übersichtlich gehalten werden sollte. Zudem ändern sich die mittels des SOEP berechneten durchschnittlichen Alter von Verwandten aufgrund geringer Fallzahlen teils sehr stark von Altersklasse zu Altersklasse, so dass sie eher wenig verlässlich scheinen und der Nutzen eines Vergleichs wahrscheinlich eingeschränkt wäre.<sup>159</sup>

## 7.2 Das Sozio-oekonomische Panel

### 7.2.1 Datensatzbeschreibung

Das Sozio-oekonomische Panel (SOEP) ist eine seit dem Jahr 1984 jährlich durchgeführte Längsschnitterhebung von privaten Haushalten in der BRD (vgl. Frick et al. 2005: 16). Alle Mitglieder eines erfassten Haushaltes im Alter von 16 oder mehr Lebensjahren werden befragt und zudem allgemeine Informationen zum Haushalt erhoben (vgl. ebd.: 21). Dabei gibt es einen Kern an Fragen, der etliche Themengebiete wie Demographie, Bildung, Erwerbstätigkeit und Einkommen abdeckt und jährlich erhoben wird, sowie Fragen aus wechselnden Themenbereichen (vgl. ebd.: 16).

Seit 1984 wurden insgesamt acht Stichproben gezogen, deren Mitglieder jeweils weiterverfolgt wurden (vgl. Siegel et al. 2010: 4). Das sogenannte Sample A umfasst die Wohnbevölkerung der BRD ohne Personen aus „Gastarbeiterländern“ (bspw. Italien, Griechenland, Türkei) und wird seit der ersten Erhebungswelle 1984 weiterverfolgt (vgl. Frick et al. 2005: 19). Ebenfalls in diesem Jahr wurde das Sample B, welches aus Personen aus „Gastarbeiterländern“ besteht, zum ersten mal erhoben (vgl. ebd.: 19). Diese Stichprobe ist verglichen mit dem Sample A überproportional groß (*oversampling*), um eine für statistische Auswertungen ausreichend große Fallzahl zu gewährleisten und Ausfälle durch Remigration auszugleichen (vgl. Frick & Schupp 2001: 5). Das Sample C wurde 1990 zum ersten Mal erhoben und umfasst die Wohnbevölkerung der ehemaligen DDR als Grundgesamtheit, wobei auch für diese Stichprobe *oversampling* vorliegt (vgl. Frick et al. 2005: 19). Das in zwei Schritten 1994 und 1995 erhobene Sample D besteht aus Haushalten, bei denen mindestens ein Haushaltsmitglied

---

<sup>159</sup> Die an sich bereits relativ geringen Fallzahlen pro Altersklasse (s. bspw. Unterabschnitt 7.2.2.1) wurden durch fehlende oder unplausible Angaben zu den Altern von Verwandten nochmals verringert. Eine Glättung über Altersklassen hinweg, wie sie beispielsweise im Unterabschnitt 7.2.2.1 beschrieben wird, hätte die Resultate teilweise sehr stark verändert.

nach 1984 in die BRD eingewandert ist (vgl. ebd.: 20). Sample E und Sample F (1998 und 2000) dienten zum einen der Auffrischung der Erhebung, zum anderen ging Sample F mit veränderten Erhebungsmethoden einher (vgl. ebd.: 20). Beide Stichproben weisen die privaten Haushalte in der BRD als Grundgesamtheit auf (vgl. ebd.: 20). 2002 wurde das Sample G zum ersten mal erhoben, wobei die Stichprobe Privathaushalte mit einem Einkommen von über 7500 DM erfasst (vgl. ebd.: 20). Mit dem Sample H wurde 2006 eine weitere Aufstockungsstichprobe gezogen (vgl. Siegel et al. 2010: 4). Schließlich wurde 2009 das Sample I zum ersten mal erfasst (vgl. Kroh 2010: 1).

## 7.2.2 Datenaufbereitung

Informationen zum Vorhandensein von Kindern, Enkelkindern, Eltern und Großeltern bei einzelnen Personen sind im SOEP zwar prinzipiell vorhanden, allerdings wurden diese zu unterschiedlichen Zeitpunkten und unterschiedlich umfassend ermittelt. Bei der Aufbereitung der Daten wurde versucht, Vergleichbarkeit mit den Modellergebnissen zu erreichen. Probleme und Einschränkungen des Vergleichs von Daten des SOEP mit Ergebnissen der Modellrechnungen, die teils zu den im weiteren geschilderten Entscheidungen bei der Datenaufbereitung geführt haben, werden erst im nächsten Abschnitt zusammengefasst diskutiert.

### 7.2.2.1 Kinder

Für die letzte betrachtete, im Jahr 2009 erhobene Welle des SOEP liegen für etwa 71% der insgesamt 20869 befragten Personen valide Informationen zur Geburtsbiographie vor. Diese erfasst die Anzahl der jemals geborenen Kinder sowie deren Geburtsjahr und Geschlecht (für Details s. Frick & Schmitt 2010; Schmitt 2010).<sup>160</sup> Solange die in der Geburtsbiographie erfassten Kinder nicht als Beobachtungen im SOEP enthalten sind, kann nicht ohne weiteres festgestellt werden, ob sie noch leben oder nicht. Dies trifft auf etwa 21% der Kinder zu.<sup>161</sup> Da also für einen nicht unerheblichen Teil der Kinder keine Angaben zum Überleben vorliegen, wurde die Auswertung auf jemals geborene Kinder beschränkt. Zudem wurde nur die Gesamtzahl an Kindern betrachtet und nicht nach Söhnen und Töchtern differenziert.

Von den 14856 Beobachtungen mit valider Geburtsbiographie wurden vor der Auswertung etliche Fälle aus dem Datensatz entfernt, da lediglich eine Validierung für Westdeutschland durchgeführt werden sollte. Zunächst wurden Personen ausgeschlossen, die nicht in Deutschland geboren wurden und in einem Alter von mehr als 25 Jahren in die BRD eingewandert sind.<sup>162</sup> Anschließend wurden Personen, die 1989 in Ostdeutschland lebten, aus dem Datensatz

160 Für Frauen werden Daten zur Geburtsbiographie seit Beginn des SOEP erhoben, für Männer hingegen erst seit dem Jahr 2001 (vgl. Schmitt 2010: 78). Dabei werden die Daten zur Geburtsbiographie von Männern nur für Personen erhoben, die seit dem Jahr 2000 in die SOEP Stichprobe aufgenommen wurden, so dass für Männer, die bereits vor diesem Jahr Teil des SOEP waren, keine Angaben vorliegen (vgl. ebd.: 79). Dies führt dazu, dass für etwa 47% der im Jahr 2009 befragten Männer keine Informationen zur Geburtsbiographie vorliegen, während dies lediglich bei 12% der Frauen der Fall ist.

161 Hierfür gibt es diverse mögliche Ursachen. Beispielsweise könnte ein Kind zu dem Zeitpunkt, zu dem der elterliche Haushalt zum ersten mal im SOEP erfasst wurde, bereits aus diesem ausgezogen sein (vgl. Frick & Schmitt 2010: 73). Ob eines der Kinder einer befragten Person seit der letzten Befragung verstorben ist, wird erst seit dem Erhebungsjahr 2007 erfasst.

162 Diese Altersgrenze wurde von Pötzsch (2010) übernommen, die Geburtenraten und Mikrozensus vergleicht.

entfernt. Diese beiden Schritte reduzierten die Fallzahl auf 9810 Beobachtungen. Ferner wurden Personen ausgeschlossen, die seit 1990 in mehreren Wellen erfasst sind und bei mindestens einer Erhebung auf dem Gebiet der ehemaligen DDR gelebt haben. Hiernach verblieben 9512 Fälle. Für die Kinder der betrachteten Personen ist oftmals kein Aufenthaltsort bekannt, so dass diesbezüglich keine Einschränkungen vorgenommen werden konnten.

Anschließend wurde jeweils für Frauen und Männer die durchschnittliche Anzahl an jemals geborenen Kindern nach dem Alter berechnet. Die Alter der Befragten wurden dabei als die Differenz von Erhebungs- und Geburtsjahr generiert. Im weiteren werden alle Alter auf diese Weise berechnet und als generierte Alter bezeichnet. Bei diesem Vorgehen ist das jüngste erfasste Alter, für das die Zahl der jemals geborenen Kinder berechnet werden könnte, gleich 17. Allerdings werden Informationen zur Biographie und somit auch zur Geburtsbiographie oftmals nicht bei der ersten Erfassung einer Person erhoben (vgl. Frick & Schmitt 2010: 70), wodurch für das Erhebungsjahr 2009 und die betrachtete Teilpopulation erst ab einem generierten Alter von 18 Jahren Informationen zu Geburtsbiographien vorliegen. Das höchste berücksichtigte Alter beträgt 89 Jahre, da für höhere Alter nur wenige Fälle vorhanden sind.

Weil sich für etliche generierte Alter relativ niedrige Fallzahlen ergeben, wurden gleitende Mittelwerte berechnet, die neben den Beobachtungen für ein Alter  $x$  auch die Beobachtungen der angrenzenden Alter  $x - 2$ ,  $x - 1$ ,  $x + 1$  und  $x + 2$  berücksichtigen.<sup>163</sup> Die Ergebnisse basieren trotz der gleitenden Mittelwerte teils dennoch auf relativ wenigen Beobachtungen. Für das Alter von 85 Jahren wurden beispielsweise unter Berücksichtigung der angrenzenden Alter nur 156 Fälle betrachtet.

Die Berechnungen wurden sowohl mit den im SOEP enthaltenen Hochrechnungsfaktoren für das Jahr 2009 als auch ohne diese Hochrechnungsfaktoren durchgeführt (zur Hochrechnung im SOEP s. bspw. Pannenberg et al. 2005; Pischner 2007), so dass zwei Varianten der Resultate vorliegen. Die Motivation für die Erstellung von zwei Varianten ist, dass unklar ist, inwieweit die Verwendung von Hochrechnungsfaktoren zu valideren Ergebnissen führt. Eine entsprechende Diskussion der Verwendung von Gewichten findet sich im folgenden Abschnitt.<sup>164</sup>

### 7.2.2.2 Enkelkinder

Im SOEP gibt es für Enkelkinder keine genaue „Buchführung“ wie für Kinder. Zwar könnte man versuchen, die Zahl der Enkelkinder einer Person indirekt als die Summe der Kinder von im SOEP erfassten Kindern dieser Person zu ermitteln. Es ergibt sich aber das oben bereits erwähnte Problem, dass nicht alle Kinder einer befragten Person selber im SOEP erfasst sind.

---

163 Für Alter  $x$ , die an den „Grenzen“ des betrachteten Altersbereich liegen, sind die oben definierten angrenzenden Alter teils nicht vorhanden. In solchen Fällen wurden die benötigten Beobachtungen der „fehlenden Alter“ durch die Beobachtungen des letzten vorhandenen Alters vor beziehungsweise nach  $x$  ersetzt.

164 Die in den Daten des SOEP enthaltenen Hochrechnungsfaktoren werden im Wesentlichen in zwei Schritten erstellt. Bei der ersten Erhebung einer Teilstichprobe werden für Mitglieder dieser Stichprobe Startgewichte festgelegt, die unterschiedliche Ziehungswahrscheinlichkeiten der einzelnen Stichprobenelemente berücksichtigen (vgl. Pischner 2007: 6ff.). Bei weiteren Wellen erfolgt eine Anpassung der Randverteilungen bestimmter Merkmale an die Verteilungen des jeweils aktuellen Mikrozensus (vgl. ebd.: 5). Seit 2005 sind dies die Merkmale „Haushaltsgröße“, „Alter der Personen in Privathaushalten“, „Geschlecht der Personen in Privathaushalten“ und „Nationalität der Personen in Privathaushalten“, wobei eine getrennte Gewichtung für Ost- und Westdeutschland vorgenommen wird (vgl. ebd.: 5).

Allerdings wurde in den Erhebungsjahren 1991, 1996, 2001 und 2006 erfasst, ob eine Person Enkel hat, und wenn dies der Fall ist, wie viele Enkel vorhanden sind.

Zum Vergleich mit den Modellergebnissen wurden die Daten des Jahres 2006 gewählt, wobei davon ausgegangen wurde, dass die erfasste Zahl der Enkel der Zahl der lebenden Enkel entspricht.<sup>165</sup> Die Stichprobe wurde dann wie bei Kindern bereits beschrieben über etliche Merkmale eingeschränkt. In einigen Fällen haben Befragte keine Angaben zur Zahl der Enkel gemacht und konnten somit nicht berücksichtigt werden. Ferner wurden Beobachtungen, bei denen unplausible Angaben zur Enkelzahl vorlagen, von den Auswertungen ausgeschlossen. Als unplausibel wurden dabei Angaben von 20 oder mehr Enkeln angesehen. Zudem wurden nur Personen in einem generierten Alter von 30 oder mehr Jahren betrachtet, da dies als das jüngste Alter angesehen werden kann, in dem Enkelkinder vorhanden sein können.

Insgesamt wurden 13376 Personen bei den Auswertungen berücksichtigt, was etwa 59% der Fälle der Stichprobe des Jahres 2006 entspricht. Auch hier konnten keine Einschränkungen bezüglich des Aufenthaltsortes von Enkeln beziehungsweise der Eltern der Enkel getroffen werden, da die in den Daten enthaltenen Informationen hierzu abermals nicht ausreichten.<sup>166</sup>

Jeweils für Frauen und Männer wurde die durchschnittliche Anzahl an lebenden Enkeln nach generierten Altern berechnet, wobei wie bei der Zahl der Kinder gleitende Mittelwerte verwendet wurden und nicht nach dem Geschlecht der Enkel differenziert wurde. Hierbei wurden nur generierte Alter bis einschließlich 85 Jahren betrachtet, da für höhere Alter die Fallzahlen sehr niedrig sind. Die Berechnungen wurden wieder in zwei Varianten durchgeführt: einer Variante mit Hochrechnungsfaktoren für das Jahr 2006 und einer Variante ohne Hochrechnungsfaktoren.

### 7.2.2.3 Eltern

Ob Mutter und Vater einer befragten Person noch leben, wurde ebenfalls 2006 zuletzt abgefragt. Zusätzlich wurde in den folgenden Erhebungswellen erfasst, ob die Mutter oder der Vater seit der letzten Befragung gestorben sind, so dass sich für das Jahr 2009 teils rekonstruieren lässt, ob die Eltern noch leben. Dies setzt voraus, dass für eine betrachtete Person Informationen zu den Erhebungsjahren 2006, 2007, 2008 und 2009 vorliegen. Aus diesem Grund können lediglich Personen in einem generierten Alter von 20 oder mehr Jahren betrachtet werden.<sup>167</sup> Für diese wurden zwei Indikatorvariablen gebildet, die erfassen, ob die Mutter beziehungsweise der Vater 2009 noch leben.

---

165 Bei der Erhebung wurde folgende Frage gestellt: „Welche der folgenden Verwandten haben Sie? Für diese Personen geben Sie bitte die Anzahl an und zusätzlich, ob sie in Ihrem Haushalt leben bzw. wie weit entfernt sie wohnen.“ Hierauf folgt im Fragebogen eine Auflistung diverser Verwandtschaftsgrade, darunter auch Enkelkinder. Es ist unklar, wie genau diese Abfrage ist und ob nicht beispielsweise von manchen Befragten auch gegebenenfalls vorhandene Urenkel mitgezählt wurden. Für eine Diskussion s. auch Unterabschnitt 4.4.3.

166 Bei Vorhandensein eines oder mehrerer Enkelkinder wurde bei der Befragung lediglich für das am nächsten bei der befragten Person lebende Enkelkind erfasst, in welcher Entfernung es zu dieser Person lebt.

167 Bei generierten Altern unter 17 Jahren besteht das Problem, dass Personen in diesen Altern nicht direkt befragt werden, sondern nur indirekt über ihre Eltern erfasst werden. Um für diese Kinder zu betrachten, ob ein Elternteil noch lebt, müsste ein gegebenenfalls eingetretener Todesfall im SOEP erfasst werden. Allerdings ist davon auszugehen, dass Sterbefälle im SOEP untererfasst sind (vgl. Schnell & Trappmann 2006). Von daher scheint es fraglich, ob für diese Alter eine verlässliche Schätzung möglich ist. Da für diese relativ jungen Alter anzunehmen ist, dass beide Elternteile mit großer Wahrscheinlichkeit noch leben, scheint ein Verzicht auf die Berechnung mit dem SOEP und auf den Vergleich mit der Modellrechnung eher unproblematisch zu sein.

Die bei den Berechnungen zu berücksichtigenden Beobachtungen wurden restriktiver als bei Kindern und Enkelkindern ausgewählt. Zunächst wurden Personen ausgeschlossen, die nicht in Deutschland geboren wurden und erst nach 1949 eingewandert sind. Anschließend wurden Beobachtungen, bei denen keine deutsche Staatsangehörigkeit vorliegt, aus dem Datensatz entfernt. Ferner wurden Personen, die 1989 oder zu einem späteren Zeitpunkt in Ostdeutschland lebten, nicht berücksichtigt. Zudem wurden Angaben zum Wohnort der Eltern genutzt, die ebenfalls 2006 erhoben wurden: Personen, bei denen mindestens ein Elternteil im Ausland lebte, wurden von den Berechnungen ausgeschlossen. Durch diese Einschränkungen und fehlende Angaben verblieben insgesamt 7112 Beobachtungen für das Jahr 2009, wobei für diese trotz der getroffenen Einschränkungen nicht ausgeschlossen werden kann, dass die Eltern im Ausland oder Ostdeutschland leben. Da für das Alter von 20 Jahren lediglich 12 Beobachtungen vorlagen, wurde es nicht bei den Auswertungen berücksichtigt.

Für die verbliebenen Fälle wurde jeweils die Wahrscheinlichkeit, dass die Mutter noch lebt, und die Wahrscheinlichkeit, dass der Vater noch lebt, nach generierten Altern differenziert berechnet, wobei abermals gleitende Mittelwerte verwendet wurden. Dabei ist das höchste berücksichtigte Alter gleich 80 Jahren. Von einer gemeinsamen Auswertung der Wahrscheinlichkeiten für die beiden Elternteile wurde abgesehen, da dies auf die Betrachtung einer Verteilung mit vier Merkmalswerten hinausläuft, bei der die Fallzahlen für einzelne Merkmalswerte oftmals klein sind.

Bei der Berechnung wurde eine Gewichtung über die Hochrechnungsfaktoren für das Jahr 2006 multipliziert mit inversen Bleibewahrscheinlichkeiten von 2006 zu 2007, 2007 zu 2008 sowie 2008 zu 2009 verwendet (vgl. Pannenberg et al. 2005: 179f.).<sup>168</sup> In einer zweiten Variante wurden lediglich die inversen Bleibewahrscheinlichkeiten berücksichtigt.

#### 7.2.2.4 Großeltern

Ob die Großeltern einer Person noch leben, wurde in den Jahren 1991, 1996, 2001 und 2006 erfasst. Allerdings wurde lediglich die Zahl lebender Großmütter insgesamt und die Zahl lebender Großväter insgesamt abgefragt, so dass nicht zwischen Großeltern mütterlicher- und väterlicherseits unterschieden werden kann. Ferner liegen diese Angaben nur für Personen ab einem generierten Alter von 17 Jahren vor. Anstatt diese Daten zu verwenden, wurden folgende Überlegungen angestellt, die auch durch Einschränkungen der Modellrechnungen motiviert sind, die im folgenden Abschnitt diskutiert werden.

Für das Jahr 2009 wurde wie oben beschrieben für Personen in einem generierten Alter von 20 oder mehr Jahren betrachtet, ob die Eltern dieser Personen noch leben. Wenn diese Personen Kinder haben, ist für diese Kinder bekannt, ob die Großeltern mütterlicher- beziehungsweise väterlicherseits noch leben, je nach Geschlecht der betrachteten Person.

Personen in einem generierten Alter von unter 17 Jahren werden im SOEP zwar nicht befragt, allerdings werden ihr Vorhandensein in einem Haushalt sowie einige Merkmale wie

---

<sup>168</sup> Die in den Daten des SOEP enthaltenen Bleibewahrscheinlichkeiten werden in zwei Schritten hergeleitet. Zunächst wird betrachtet, wie wahrscheinlich es ist, dass Kontakt zu einem Haushalt aufgenommen werden kann, und anschließend wird die Wahrscheinlichkeit betrachtet, dass die Teilnahme verweigert wird (vgl. Kroh 2010: 37). Zur Berechnung der Wahrscheinlichkeiten werden Logit-Modelle verwendet, wobei im Falle der Kontaktwahrscheinlichkeit lediglich eine geringe Anzahl an „erklärenden“ Variablen mit in die Modellschätzung eingeht (vgl. ebd.: 27f.).

Geburtsjahr und Geschlecht erfasst und in einem eigenen Datensatz festgehalten. Ferner ist in diesem Datensatz die Personnummer der Mutter bis auf wenige Ausnahmen angegeben.<sup>169</sup> Lässt sich für eine Frau mit Kindern für das Jahr 2009 rekonstruieren, ob ihre Eltern noch leben, kann diese Information den Kindern über die Personnummer zugespielt werden und Auskunft über das Vorhandensein Großeltern mütterlicherseits geben.

Informationen von Vätern lassen sich nicht ohne weiteres hinzufügen, da für Kinder die Personnummer des Vaters nicht direkt festgehalten wird. Im Kinderdatensatz ist allerdings die Personnummer des Haushaltsvorstandes des Haushalts, in dem ein Kind lebt, sowie die Beziehung zu diesem Haushaltsvorstand erfasst. Eine mögliche Beziehung ist „Sohn, Tochter des Haushaltsvorstandes“. Ist der Haushaltsvorstand in einem solchen Fall männlich, kann dieser als Vater identifiziert werden und wie bei Müttern vorgegangen werden. Ist der Vater aber nicht Haushaltsvorstand, sondern beispielsweise die Mutter, muss auf andere Informationen zurückgegriffen werden. Hier können Daten zu Geburtsbiographien für Männer verwendet werden, bei denen Personennummern einzelner Kinder dieser Männer erfasst sind. Hierüber lässt sich eine zusätzliche Verknüpfung von Vätern und deren Kindern und hierüber von Großeltern vornehmen.

Ausgangspunkt war also ein Datensatz, der für Personen in einem generierten Alter von 20 und mehr Lebensjahren für das Jahr 2009 das Vorhandensein von Eltern erfasst. Dieser wurde wie bei Eltern beschrieben erstellt. Anschließend wurde der Kinderdatensatz hinzugenommen und über das beschriebene Vorgehen versucht, Informationen von Eltern und Kindern miteinander zu verknüpfen. Insgesamt konnte so 2226 von insgesamt 4429 Kindern Informationen über die Großeltern zugespielt werden. Vollständige Informationen zu den väterlichen und mütterlichen Großeltern weisen 1415 Beobachtungen auf. Informationen zu Großeltern mütterlicherseits gibt es bei 1975 Fällen und Informationen zu Großeltern väterlicherseits bei 1666 Beobachtungen. Diese Differenz zwischen Großeltern mütterlicher- und väterlicherseits zeigt, dass über das oben beschriebene Vorgehen nicht allen Kindern ein Vater zugespielt werden konnte.

Die Auswertung erfolgte separat für Großeltern mütterlicher- und väterlicherseits. Bei den Berechnungen zu Großeltern mütterlicherseits wurden Kinder mit den inversen Bleibewahrscheinlichkeiten der Mutter von Welle 2006 nach Welle 2009 gewichtet. Bei Großeltern väterlicherseits wurde analog vorgegangen, nur dass in diesem Fall die inversen Bleibewahrscheinlichkeiten des Vaters verwendet werden.

Da Kinder, die erst nach der Erhebung 2006 geboren wurden, für das Jahr 2006 einen Hochrechnungsfaktor von 0 aufweisen, könnten sie bei Verwendung von Hochrechnungsfaktoren nicht bei den Auswertungen berücksichtigt werden. Dies betrifft alle Kinder, die 2009 ein generiertes Alter von 0, 1 oder 2 Jahren aufweisen. Aus diesem Grund wurde lediglich eine Variante mit inversen Bleibewahrscheinlichkeiten erstellt und auf die Verwendung von Hochrechnungsfaktoren verzichtet.

---

169 Die Personnummer ist eine eindeutige Identifikationsnummer, die jede im SOEP erfasste Person erhält (vgl. Göbel et al. 2008: 26). Für den Kinderdatensatz des Jahres 2009 fehlt die Identifikationsnummer der Mutter bei 54 von 4429 Fällen.

## 7.3 Vergleich der Ergebnisse der Modellrechnung mit Daten des SOEP

### 7.3.1 Vorüberlegungen

Beim Vergleich der Ergebnisse des Modells, die sich bei der Verwendung historischer demographischer Raten ergeben, und den Daten des SOEP sind möglicherweise auftretende Differenzen zwischen diesen im Wesentlichen auf vier potenzielle Fehlerquellen zurückzuführen:

1. Mangelnde Vergleichbarkeit der Modellergebnisse und der empirischen Daten
2. Fehler in den empirischen Mikrodaten
3. Fehler in den verwendeten demographischen Raten
4. Verletzung der Modellannahmen

Im weiteren wird diskutiert, inwieweit diese Fehlerquellen vorliegen und ob sie behoben werden können.

#### 7.3.1.1 Vergleichbarkeit von Modellrechnung und empirischen Daten

Zunächst ist bezüglich der Vergleichbarkeit zu bedenken, dass sich die Ergebnisse der Modellrechnung auf Altersklassen beziehen, die Ergebnisse des SOEP auf generierte Alter. Als Beispiel wird ein Jahr  $t$  betrachtet, eine Altersklasse  $a$  und ein generiertes Alter  $a'$ , wobei  $a = a'$  sei. Eine Person aus Altersklasse  $a$  befand sich zu  $t - a$  in Altersklasse 0 und wurde im Jahr  $t - a - 1$  beziehungsweise im Intervall  $[t - a - 1, t - a)$  geboren. Eine Person im generierten Alter  $a'$  wurde hingegen im Jahr  $t - a'$  beziehungsweise im Intervall  $[t - a', t - a' + 1)$  geboren. Deshalb werden die Ergebnisse des SOEP für ein generiertes Alter  $x$  den Ergebnissen der Modellrechnung für die Altersklasse  $x - 1$  zugeordnet.

Des Weiteren liegen die Ergebnisse des SOEP entsprechend der Ausführungen des vorausgegangenen Abschnitts nur ab beziehungsweise bis zu bestimmten Altern vor. Gleiches gilt wie zu Beginn dieses Kapitels beschrieben für die Ergebnisse der Modellrechnungen, so dass immer nur Vergleiche für bestimmte Altersbereiche vorgenommen werden können. In Kombination mit der oben genannten „Verschiebung“ der generierten Alter sind Vergleiche für folgende Altersklassen möglich:

- Bei Kindern für die Altersklassen 17 bis 72
- Bei Enkelkindern für die Altersklassen 29 bis 72
- Bei Eltern für die Altersklassen 21 bis 57
- Bei Enkelkindern für die Altersklassen 0 bis 5

Dabei ist aber zu beachten, dass die Ergebnisse des SOEP wie erwähnt größtenteils durch gleitende Mittelwerte geglättet wurden. Insofern beziehen sich die Ergebnisse des SOEP, die entsprechend des gerade vorgestellten Schemas einer Altersklasse  $x$  zugespült werden, auf die generierten Alter  $x - 1, x, x + 1, x + 2$  und  $x + 3$ . Inwieweit dies die Vergleichbarkeit einschränkt ist unklar.

Vergleiche von Modellrechnung und SOEP sind ferner durch Unterschiede in den Populationen nur eingeschränkt aussagekräftig. Die den Modellrechnungen zugrunde liegenden

demographischen Raten beziehen sich einerseits in allen Fällen auf die Wohnbevölkerung Westdeutschlands. Andererseits gelten die Ergebnisse der Modellrechnung nur für eine geschlossene Bevölkerung mit genau diesen Raten, womit also keine Migration stattfinden darf (s. auch Anhang C.5). Hier stellt sich für den Vergleich die Frage, ob dieser Umstand bei der für das SOEP vorgenommenen Fallauswahl hinreichend beachtet wurde.<sup>170</sup> Wie im vorausgegangenen Abschnitt beschrieben wurde versucht, bei der Fallauswahl einer geschlossenen Bevölkerung möglichst nahe zu kommen, indem beispielsweise bei der Betrachtung von Vorfahren Personen ausgeschlossen wurden, die nicht in Deutschland geboren wurden. Allerdings wird so nur ein Teil der eigentlichen Wohnbevölkerung betrachtet.

Weitere Abweichungen gibt es beispielsweise bei der Unterscheidung von West- und Ostdeutschland. Während bei den demographischen Raten Westberlin nach 1989 nur teilweise bei Westdeutschland berücksichtigt wird (s. Anhang A), wird im SOEP anhand der bis zum 3. Oktober 1990 geltenden Grenzen zwischen Ost- und Westdeutschland differenziert, wobei im Falle Berlins bekannt ist, ob eine Person im ehemaligen West- oder Ostteil der Stadt lebt. Hier wurde sich dafür entschieden, Beobachtungen aus Westberlin bei den Auswertungen zu berücksichtigen.

Somit sind die Ergebnisse der Auswertung des SOEP und der Modellrechnung zum einen bezogen auf die Alter und zum anderen bezogen auf Abgrenzungen der Bevölkerung nicht vollständig deckungsgleich.

### 7.3.1.2 Fehler in den empirischen Mikrodaten

Auch bei Gleichheit der Populationen ist nicht davon auszugehen wäre, dass die Modellrechnungen exakt mit den Daten des SOEP übereinstimmen. Ursache hierfür sind Stichprobenfehler und möglichen Ungenauigkeiten der Daten des SOEP.<sup>171</sup>

Ferner kommen Probleme bei der Gewichtung und durch Panelfälle hinzu. Zwar wird in der Literatur darauf hingewiesen, dass bei deskriptiven Auswertungen eine Gewichtung der Daten des SOEP über Hochrechnungsfaktoren unerlässlich sei (vgl. Pischner 2007: 1). Allerdings ist fraglich, inwiefern die Hochrechnungsfaktoren und inversen Bleibewahrscheinlichkeiten zu einer „Verbesserung“ der Ergebnisse führen, da unklar ist, ob die bei der Gewichtung berücksichtigten Merkmale relevant für den hier gewählten Untersuchungsgegenstand sind und ob nicht weitere relevante Merkmale unberücksichtigt bleiben. Schnell & Trappmann (2006) vergleichen beispielsweise mit dem SOEP gewonnene Ergebnisse zur Mortalität von Männern und Frauen mit Daten der amtlichen Statistik. Bei Frauen führt eine Gewichtung zu deutlicheren Abweichungen zwischen SOEP und amtlichen Daten als ohne diese Gewichtung, während es sich bei Männern umgekehrt verhält (vgl. Schnell & Trappmann 2006: 7ff.).

---

170 Das selbe Problem wird von Pötzsch (2010: 187ff. u. 194) im Rahmen eines Vergleichs von Ergebnissen des Mikrozensus und amtlichen Daten zur Fertilität von Frauen diskutiert.

171 Insbesondere bei retrospektiven Befragungen ist davon auszugehen, dass Männer bei Fragen nach der Kinderzahl nichteheliche Kinder und Kinder aus vergangenen Beziehungen teils nicht angeben (vgl. Rendall et al. 1999; Vere 2008). In wenigen Fällen werden Väter nicht über die Geburt eines Kindes informiert sein (vgl. Schmitt 2005: 22). Schließlich kann es umgekehrt der Fall sein, dass für ein Kind fälschlicherweise von einer Vaterschaft ausgegangen wird (vgl. Anderson 2006; Coleman 2000: 46f.). Eine Übersicht über Literatur zur Qualität von Umfragedaten zur Fertilität von Männern findet sich bei Zhang (2011: 29-32).

Bei den Vergleichen zwischen Modellrechnung und SOEP werden deshalb sowohl die Ergebnisse des SOEP mit Hochrechnung, als auch die Ergebnisse ohne Hochrechnung berechnet, wobei nicht ohne weiteres gesagt werden kann, welche der beiden Varianten zu „besseren“ Ergebnissen führt. Die Gewichtung von Beobachtungen aufgrund von Panelausfällen scheint einerseits ebenfalls problematisch, andererseits sollten „nicht-neutrale“ Ausfälle auf jeden Fall berücksichtigt werden, weshalb die Längsschnittgewichtung bei allen in Frage kommenden Auswertungen verwendet wird.

#### 7.3.1.3 Fehler in den demographischen Raten

Neben Problemen der Vergleichbarkeit von Modellrechnungen und empirischen Daten sowie möglichen Problemen mit letzteren können Fehler in den verwendeten demographischen Raten ebenfalls zu Abweichungen zwischen Modell und Daten führen. Für die meisten verwendeten demographischen Raten ist aber den Ausführungen in Anhang A entsprechend davon auszugehen, dass sie als relativ fehlerfrei gelten können. Eine Ausnahme sind die Verteilungen  $\phi(y|x)$ , die aus aktuellen Daten gewonnen wurden, allerdings für den gesamten Zeitraum der Modellrechnung angewandt werden. Dabei zeigte sich in Kapitel 6, dass die Verwendung dieser Verteilungen zur Rekonstruktion der Fertilität von Männern zumindest für die Nachkriegsjahre problematisch ist. Insofern wären bei Ergebnissen der Modellrechnung, die auf  $\phi(y|x)$  und relativ weit in der Vergangenheit liegenden Daten basieren, Abweichungen zum SOEP wenig überraschend.

#### 7.3.1.4 Verletzung der Modellannahmen

Differenzen, die auf die bisher diskutierten Ursachen zurückzuführen sind, stellen das Modell und die Modellannahmen selbst nicht in Frage und lassen sich im Wesentlichen als „Datenprobleme“ auffassen. Nur Abweichungen, die sich aus Modellannahmen wie der Unkorreliertheit von Personen ergeben, weisen auf eine eingeschränkte Aussagekraft des Modells hin. Bei einem Vergleich von Modellrechnungen und Ergebnissen des SOEP müssen Abweichungen aus diesem Grunde möglichst genau untersucht werden, um einzugrenzen, welche der Fehlerquellen die Abweichungen verursacht.

Aufgrund der in diesem Unterabschnitt genannten Ursachen und mit Hinblick auf in der Fachliteratur zu findende Ergebnisse zum Vergleich von Daten der amtlichen Statistik und Umfragedaten sind Abweichungen zwischen SOEP und Modellrechnung allerdings zu erwarten.<sup>172</sup> Nichtsdestotrotz sollten die Ergebnisse zumindest grob übereinstimmen und keine systematischen Abweichungen auftreten. Beispielsweise sollte die Zahl bestimmter Verwandter nach der Modellrechnung nicht durchgängig deutlich niedriger oder höher liegen als im SOEP.

---

172 Dinkel & Milenovic (1992: 60) stellen fest, dass Schätzungen aus Surveydaten 3% bis 5% unter Schätzungen aus Geburtenraten der amtlichen Statistik liegen. Pötsch (2010: 193) errechnet absolute Differenzen zwischen aus amtlichen Geburtenraten und aus Umfragedaten gewonnener Kohortenfertilität von bis zu 0.11. Ein von Huinink (1988: 363f.) vorgenommener inferenzstatistischer Vergleich verweist auf signifikante Differenzen zwischen den beiden Datenquellen.

## 7.3.2 Ergebnisse

### 7.3.2.1 Nachfahren

In Abbildung 7.1 ist die durchschnittliche Zahl jemals geborener Kinder von Frauen für das Jahr 2009 abgetragen, wie sie sich aus der Modellrechnung sowie den Daten des SOEP ergibt. Abbildung 7.2 enthält entsprechende Resultate für Männer. Hier sei daran erinnert, dass die aus dem SOEP gewonnenen Ergebnisse unter Verwendung eines gleitenden Mittelwerts berechnet wurden.

Bei Frauen liegen die Ergebnisse der Modellrechnung bis zur Altersklasse von etwa 35 Jahren relativ nah an den Ergebnissen des SOEP. Ab diesem Alter liegen die Ergebnisse zwischen den beiden mit den SOEP berechneten Varianten und unterschreiten sie lediglich in wenigen Altern leicht. Insgesamt scheint der Verlauf und auch der Umfang der Kinderzahl von Frauen relativ gut erfasst zu werden, insbesondere für Altersklassen, in denen die reproduktive Phase bereits abgeschlossen ist. Die durchschnittliche absolute Abweichung zwischen Modellrechnung und hochgerechneten Ergebnissen des SOEP beträgt 0.08. Werden die Hochrechnungsfaktoren nicht verwendet liegt die durchschnittliche absolute Differenz ebenfalls bei 0.08.

Bei Männern ist die Übereinstimmung der Verläufe geringer. Zwar liegen die Ergebnisse der Modellrechnung bis zur Altersklasse von etwa 35 Jahren ebenfalls relativ nah an den Ergebnissen des SOEP. Allerdings sind die Ergebnisse der Modellrechnung für die folgenden Altersklassen bis zu 45 Jahren niedriger als die Werte der empirischen Daten. Bis zu einem Alter von ungefähr 60 Jahren liegen die Ergebnisse zwischen den beiden mit dem SOEP berechneten Varianten. Die sich bei der Modellrechnung ergebenden Schwankungen ab der Altersklasse von 60 Jahren lassen sich aus den Ergebnissen des SOEP nicht erkennen. Die durchschnittliche Differenz zwischen Modellrechnung und empirischen Daten beträgt 0.14, beziehungsweise 0.12 bei Verwendung der Hochrechnungsfaktoren. Zwar sind die Abweichungen höher als bei Frauen, aber zumindest für höhere Altersklassen in Anbetracht der Auswertungen aus Kapitel 6 wie bereits mehrfach erwähnt zu erwarten gewesen. Hierbei ist zusätzlich zu bedenken, dass gerade für höhere Alter die Fallzahlen im SOEP sehr niedrig ausfallen, so dass die Ergebnisse nicht unbedingt verlässlich sind (s. auch Fußnote 160 auf Seite 147).<sup>173</sup>

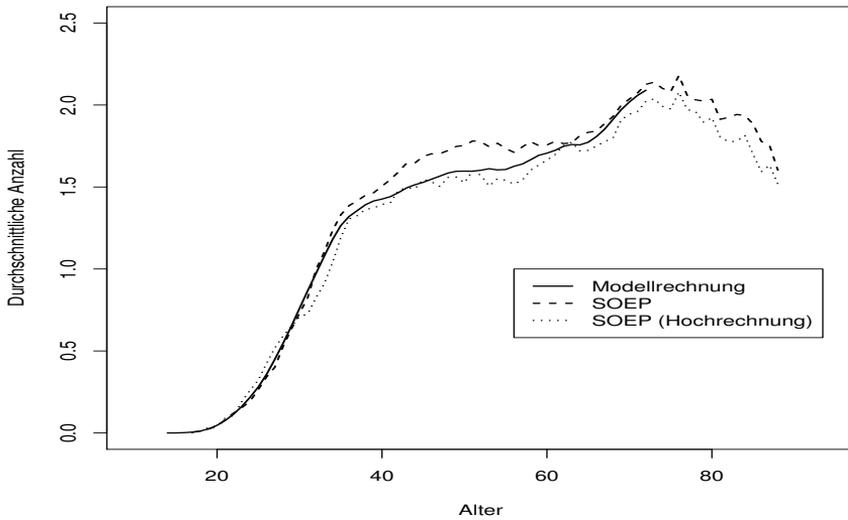
Die durchschnittliche Zahl lebender Enkel von Frauen ist in Abbildung 7.3 dargestellt, die durchschnittliche Zahl lebender Enkel von Männern ist in Abbildung 7.4 zu sehen. Die Ergebnisse des SOEP beziehen sich wie im vorausgegangenen Abschnitt beschrieben auf das Jahr 2006, die Ergebnisse der Modellrechnung auf das Jahr 2009.<sup>174</sup>

Bei beiden Geschlechtern beginnt die Zahl der Enkel im Wesentlichen erst ab der Altersklasse von 40 Jahren zu steigen. Während bei Männern die durchschnittliche Zahl lebender Enkel über alle Alter mit durchschnittlichen absoluten Abweichungen von 0.07 beziehungsweise 0.06 (Hochrechnung) relativ gut geschätzt erscheint, zeigen sich bei Frauen ab der Altersklasse

<sup>173</sup> Beispielsweise liegen nur 51 Beobachtungen für das Alter von 70 Jahren vor. Unter Verwendung des gleitenden Mittelwerts werden für diese Altersklasse 285 Beobachtungen genutzt.

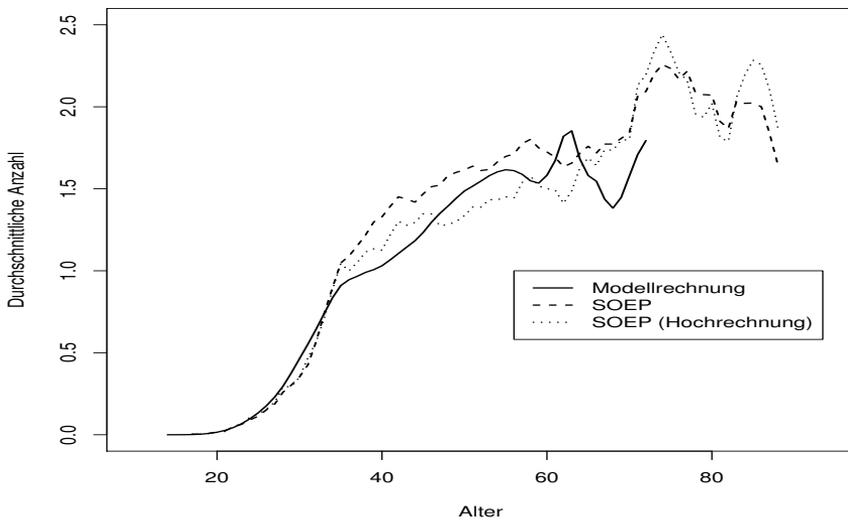
<sup>174</sup> Bei der Verwendung der Ergebnisse der Modellrechnung für das Jahr 2006 ergeben sich die selben Interpretationen, wobei allerdings Ergebnisse für weniger Alter zur Verfügung stehen, weshalb der Vergleich mit den Ergebnissen für 2009 durchgeführt wird.

Abbildung 7.1: Durchschnittliche bzw. erwartete Zahl jemals geborener Kinder von Frauen 2009



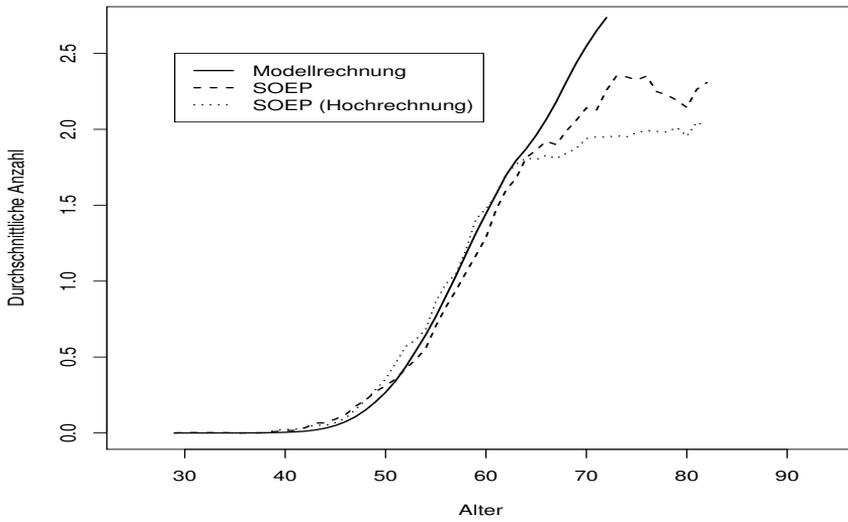
Quelle: SOEP; eigene Berechnung

Abbildung 7.2: Durchschnittliche bzw. erwartete Zahl jemals geborener Kinder von Männern 2009



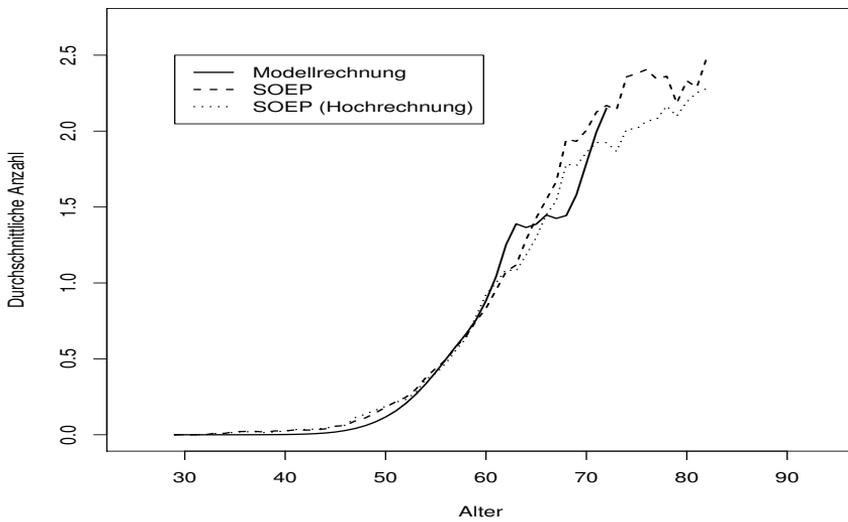
Quelle: SOEP; eigene Berechnung

Abbildung 7.3: Durchschnittliche bzw. erwartete Zahl lebender Enkel von Frauen 2006 und 2009



Quelle: SOEP; eigene Berechnung

Abbildung 7.4: Durchschnittliche bzw. erwartete Zahl lebender Enkel von Männern 2006 und 2009



Quelle: SOEP; eigene Berechnung

von 65 Jahren hingegen deutliche Abweichungen, wobei die Ergebnisse der Modellrechnung über den Ergebnissen des SOEP liegen. Beispielsweise ergibt sich in der Modellrechnung für die Altersklasse von 70 Jahren ein Wert von etwa 2.55. Die Ergebnisse des SOEP liegen für diese Altersklasse bei 1.94 (Hochrechnung) beziehungsweise 2.14. Betrachtet man die ursprünglichen Ergebnisse für das Einzelalter von 70 Jahren ohne Verwendung des gleitenden Mittelwerts zeigen sich Werte von 2.03 (Hochrechnung) und 2.21 bei insgesamt 89 Fällen. Dabei beträgt die durchschnittliche Zahl jemals geborener Kinder von Frauen in diesem Alter nach der Modellrechnung 2.02 und nach Ergebnissen des SOEP 1.95 (Hochrechnung) beziehungsweise 2.10 bei insgesamt 105 Beobachtungen.

Die betrachteten Frauen haben somit durchschnittlich etwa 2 Kinder. Bedenkt man, dass diese Kinder die reproduktive Phase zumindest größtenteils durchlaufen haben und somit durchschnittlich wahrscheinlich mehr als 1 Kind haben sollten, erscheinen die aus dem SOEP gewonnenen Resultate relativ niedrig zu sein. Es lässt sich in diesem Fall aber nicht definitiv entscheiden, ob die Ergebnisse der Modellrechnung die Zahl lebender Enkel für höhere Alter überschätzen, oder aber die aus dem SOEP gewonnenen Ergebnisse aufgrund der genannten Einschränkungen unzuverlässig beziehungsweise nicht mit der Modellrechnung zu vergleichen sind. Insgesamt beträgt die durchschnittliche absolute Abweichung bei Enkelkindern von Frauen 0.10 beziehungsweise 0.12 (Hochrechnung), was darauf zurückzuführen ist, dass für Altersklassen bis zu etwa 65 Jahren die Ergebnisse der Modellrechnung und die Ergebnisse des SOEP relativ gut übereinstimmen.

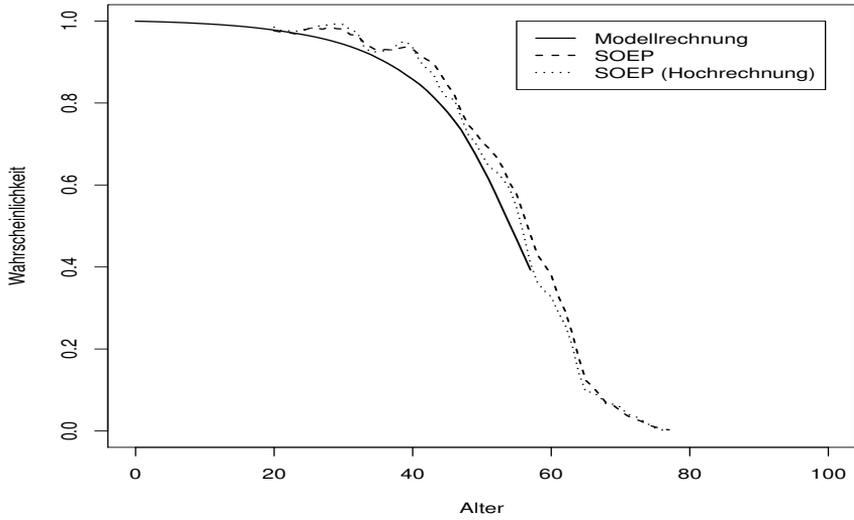
Bei den Modellergebnissen für Nachfahren bleibt festzuhalten, dass sich zwar Abweichungen zeigen, die Modellrechnungen insgesamt aber eine grobe Abschätzung der aus dem SOEP berechneten Ergebnisse erlauben, insbesondere in Anbetracht der Einschränkungen der Vergleiche.

### 7.3.2.2 Vorfahren

In Abbildung 7.5 ist die Wahrscheinlichkeit, dass die Mutter einer Person lebt, zu sehen. Die Ergebnisse der Modellrechnung liegen fast durchgängig unter den aus dem SOEP berechneten Resultaten, allerdings stimmen die Ergebnisse in der Verlaufsform überein. Die durchschnittliche absolute Differenz zwischen SOEP und Modellrechnung beträgt ohne Berücksichtigung der Hochrechnungsfaktoren 0.05, also 5 Prozentpunkte. Die größte absolute Abweichung liegt bei 0.10. Werden die Hochrechnungsfaktoren berücksichtigt, betragen diese Werte 0.04 und 0.09.

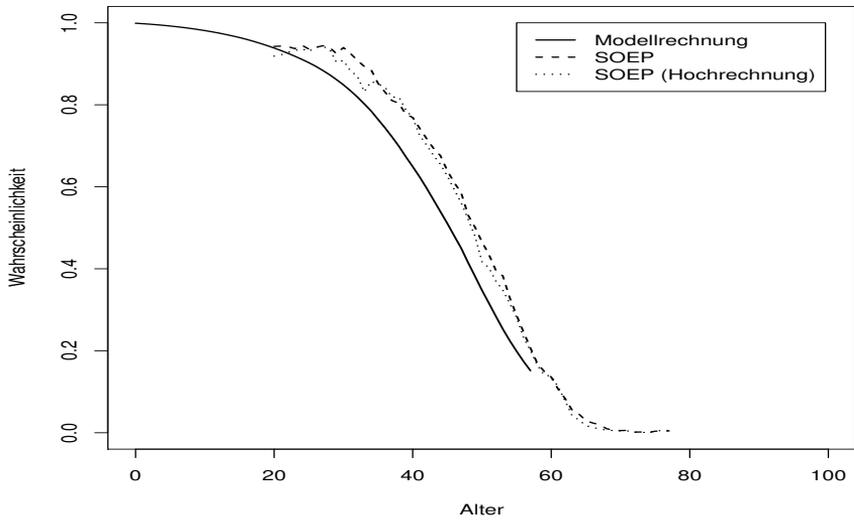
Deutlichere Abweichungen zeigen sich bei der Wahrscheinlichkeit, dass der Vater noch lebt, die in Abbildung 7.6 dargestellt ist. Auch hier liegen die Ergebnisse der Modellrechnung für fast alle Alter unter den Ergebnissen des SOEP. Ohne Hochrechnungsfaktoren ergibt sich eine durchschnittliche absolute Differenz von 0.09 und mit Hochrechnungsfaktoren von 0.08. Die größte absolute Differenz beträgt bei diesen beiden Varianten 0.17 beziehungsweise 0.16. Geht man davon aus, dass die Ergebnisse des SOEP die „tatsächlichen“ Verhältnisse widerspiegeln, und bedenkt, dass die bei der Modellrechnungen verwendeten Verteilungen der Alter von Frauen bei der Geburt eines Kindes und Überlebenswahrscheinlichkeiten als relativ genau gelten können, lässt sich vermuten, dass die Abweichungen bei Müttern auf eine Verletzung

Abbildung 7.5: Wahrscheinlichkeit für eine lebende Mutter 2009



Quelle: SOEP; eigene Berechnung

Abbildung 7.6: Wahrscheinlichkeit für einen lebenden Vater 2009



Quelle: SOEP; eigene Berechnung

von Modellannahmen zurückzuführen sind.<sup>175</sup> Zwar sind die Abweichungen zumindest teilweise eher moderat, allerdings zeigt sich eine systematische Unterschätzung der Ergebnisse des SOEP durch die Modellrechnung. Für die Unterschätzung der Wahrscheinlichkeit, dass der Vater noch lebt, könnten zum einen ebenfalls Modellannahmen verantwortlich sein, aber auch ein falsches Festlegen von Altern von Vätern über  $\phi(y|x)$ .

Eine Möglichkeit, diesen Überlegungen weiter nachzugehen, bieten die Tabellen 7.1 und 7.2, in denen die Wahrscheinlichkeiten für lebende Großeltern abgetragen sind. Die Ergebnisse des SOEP beziehen sich auf gruppierte Altersklassen von 0 bis 1 Jahr, 2 bis 3 Jahren und 4 bis 5 Jahren. Die Ergebnisse für diese Altersgruppen wurden jeweils den beiden entsprechenden Altersklassen der Modellrechnung zugeordnet. Auf die Verwendung eines gleitenden Mittelwertes wurde hier aufgrund des relativ „schmalen“ betrachteten Altersbereichs verzichtet.

Tabelle 7.1 enthält zum einen die Wahrscheinlichkeit, dass die Großmutter mütterlicherseits noch lebt, und zum anderen die entsprechende Wahrscheinlichkeit für den Großvater mütterlicherseits. Die Ergebnisse der Modellrechnung liegen unter denen des SOEP. Bei der Wahrscheinlichkeit, dass die Großmutter noch lebt, reichen die absoluten Abweichungen von 0.02 bis 0.04. Beim Großvater liegen die Abweichungen zwischen 0.08 und 0.15.

In Tabelle 7.2 sind die Wahrscheinlichkeiten zu sehen, dass die Großeltern väterlicherseits noch leben. Auch hier fallen die Ergebnisse der Modellrechnung durchgängig niedriger als die aus dem SOEP berechneten Resultate aus. Bei der Großmutter liegen die absoluten Differenzen zwischen 0.02 und 0.06, beim Großvater zwischen 0.12 und 0.18.

Sowohl bei Großeltern mütterlicher- als auch väterlicherseits liegen die Abweichungen zwischen Modellrechnung und SOEP in der Größenordnung der Abweichungen bei den Wahrscheinlichkeiten für lebende Eltern. Dies spricht dafür, dass Alter von Eltern nicht per se falsch zugeordnet werden. Denn sollte dies der Fall sein, müsste bei Großeltern eine „doppelt falsche“ Zuordnung vorliegen, zumal die Wahrscheinlichkeit für deren Vorhandensein über die Eltern einer Person vermittelt berechnet wird. Allerdings werden bei Großeltern nur sehr wenige Alter betrachtet, so dass die gerade genannten Schlussfolgerungen nur für diese Gültigkeit haben. Zudem ist beispielsweise denkbar, dass bei den betrachteten Altern die Alter der Eltern richtig zugespielt werden, allerdings nicht die Alter der Großeltern, was eine mögliche Erklärung für den Umfang der Abweichungen wäre.<sup>176</sup>

Deshalb lässt sich für die Ergebnisse zu lebenden Vorfahren insgesamt festhalten, dass bezogen auf die Resultate des SOEP eine systematische Unterschätzung der Wahrscheinlichkeiten

---

175 Sterbefälle von Befragten scheinen im SOEP untererfasst (vgl. Schnell & Trappmann 2006). Inwieweit sich dies auf „indirekte“ Erfassung der Sterblichkeit von Eltern über die Befragung ihrer Kinder überträgt ist unklar. Man könnte spekulieren, dass der Tod eines Elternteils als Stressor wirkt, bei dessen auftreten die Bereitschaft zur (wiederholten) Teilnahme an der Befragung sinkt. In diesem Fall würden die Ergebnisse des SOEP zu einer Überschätzung des Vorhandenseins von Eltern führen, zumindest wenn die Ausfälle dauerhaft sind.

Damit temporär ausfallende Beobachtungen berücksichtigt werden können, müssen für diese zumindest Informationen für die Jahre 2006 und 2009 vorhanden sein. Zwar gibt es für nur bei einzelnen Befragungen ausfallende Beobachtungen Nachbefragungen, allerdings wurde bei der Nachbefragung für das Jahr 2006 nicht die Frage nach dem Todesjahr der Eltern gestellt. Für Fälle, für die 2009 keine Informationen vorhanden sind, kann nicht bestimmt werden, ob der Ausfall nur temporär ist oder nicht. Fehlende Informationen zu Eltern für die Jahre 2007 und 2008 können aber über Nachbefragungen zugespielt werden. Dies betrifft allerdings nur 5 Beobachtungen.

176 Allerdings ist bei Großeltern väterlicherseits zu bedenken, dass entsprechend der Ausführungen in Paragraph 7.2.2.4 nicht allen Kindern ein Vater und hierüber Großeltern väterlicherseits zugespielt werden konnten. Inwieweit dies einen Effekt auf die Ergebnisse hat, ist unklar.

*Tabelle 7.1: Wahrscheinlichkeiten für lebende Großeltern mütterlicherseits 2009*

Alter	GM (M)	GM (SOEP)	GV (M)	GV (SOEP)
0	0.93	0.97	0.81	0.89
1	0.92	0.97	0.80	0.89
2	0.92	0.94	0.79	0.93
3	0.91	0.94	0.77	0.93
4	0.90	0.92	0.75	0.84
5	0.89	0.92	0.73	0.84

Abkürzungen: GM=Großmutter, GV=Großvater, M=Modellrechnung  
 Quelle: SOEP; eigene Berechnung

*Tabelle 7.2: Wahrscheinlichkeiten für lebende Großeltern väterlicherseits 2009*

Alter	GM (M)	GM (SOEP)	GV (M)	GV (SOEP)
0	0.90	0.94	0.76	0.89
1	0.89	0.94	0.74	0.89
2	0.88	0.90	0.72	0.88
3	0.87	0.90	0.70	0.88
4	0.86	0.91	0.68	0.80
5	0.85	0.91	0.66	0.80

Abkürzungen: GM=Großmutter, GV=Großvater, M=Modellrechnung  
 Quelle: SOEP; eigene Berechnung

für lebende Vorfahren vorliegt, die bei männlichen Vorfahren stärker ausfällt als bei Frauen. Nimmt man an, dass die Ergebnisse des SOEP „richtig“ sind, kann vermutet werden, dass die dem Modell zugrunde liegenden Annahmen eine mögliche Ursache für diese Unterschätzung sind.

## 7.4 Zusammenfassung und Diskussion

Beim Vergleich der Ergebnisse der Modellrechnungen mit empirischen Daten des SOEP wurde herausgestellt, dass dieser mit etlichen Einschränkungen behaftet ist. Wird er nichtsdestotrotz durchgeführt, zeigen sich bei der Berechnung von Nachfahren insgesamt keine übermäßigen Abweichungen zwischen Modell und Umfragedaten. Hiervon ausgenommen ist die Zahl der Kinder von Männern in hohen Altern und die Zahl der Enkel von Frauen in hohen Altern, bei denen sich relativ deutliche Unterschiede zeigen. Jedoch lassen diese das Modell nicht grundsätzlich als ungeeignet erscheinen.

Werden die Ergebnisse der Modellrechnung zu Vorfahren mit aus dem SOEP gewonnenen Resultaten verglichen, zeigen sich systematische Abweichungen, dergestalt dass die Ergebnisse des Modells fast durchgängig unter den Resultaten des SOEP liegen. Nimmt man das SOEP trotz Einschränkungen der Vergleichbarkeit als Maßstab, bedeutet dies, dass das Vorhandensein von Vorfahren durch das Modell unterschätzt wird. Von dieser Beobachtung ausgehend wird in Anhang F ein einfaches Korrekturverfahren entwickelt, welches die Modellrechnungen an die empirischen Daten angleichen kann.

Allerdings ist aufgrund etlicher, in diesem Kapitel genannter Argumente nicht unbedingt davon auszugehen, dass die Ergebnisse des SOEP als „korrekt“ gelten können und mit denen der Modellrechnung vergleichbar sind. Daneben geht das Korrekturverfahren von diversen zusätzlichen Annahmen aus, weshalb im weiteren Verlauf auf die eigentlichen, nicht korrigierten Ergebnisse Bezug genommen wird und die Ergebnisse des Korrekturverfahrens lediglich als Ergänzung aufgefasst und ebenfalls in Anhang F besprochen werden.

Bevor die Ergebnisse der Vorausberechnung linearer Verwandter vorgestellt werden, folgt im nächsten Kapitel zunächst noch eine Beschreibung der Verfahren, die verwendet wurden, um demographische Raten vorauszuberechnen.

## 8 Vorausberechnung der demographischen Komponenten und der Bevölkerung

### 8.1 Vorbemerkungen

Wie in Kapitel 5 beschrieben soll als Grundlage für die Vorausberechnung von Verwandtschaft auf stochastische Vorausberechnungen der demographischen Komponenten zurückgegriffen werden, um Unsicherheit der zukünftigen demographischen Entwicklung abbilden zu können. Vorausberechnet werden müssen Geburten- und Sterberaten, die in die Modellparameter  $P_x$  und  $F_x$  umgewandelt werden.<sup>177</sup> Daneben ist auch eine Vorausberechnung der Nettomigration notwendig. Diese geht zwar nicht explizit in die Modellgleichungen aus Kapitel 5 ein, beeinflusst aber die Bevölkerungsstruktur. Letztere muss für die Vorausberechnung von Verwandtschaft bekannt sein, weshalb die Vorausberechnungen der Mortalität, Fertilität und Migration zu einer Bevölkerungsvorausberechnung kombiniert werden müssen.

Aufgrund des hier gewählten Kontextes der BRD scheint zudem eine regionale Unterteilung in Ost- und Westdeutschland angebracht, da sich die demographischen Komponenten in diesen beiden Landesteilen unterschiedlich entwickelt haben (s. Abschnitt 3.2).

Um das Kapitel übersichtlich zu halten, erfolgt keine detaillierte Validierung der Ergebnisse, wie in Unterabschnitt 4.1.1 eigentlich gefordert. Stattdessen werden diese mit anderen stochastischen Vorausberechnungen aus der Literatur verglichen und auf ihre Plausibilität überprüft. Zudem werden die Ergebnisse den Annahmen der zwölften koordinierten Bevölkerungsvorausberechnung des Statistischen Bundesamtes gegenübergestellt, um eine weitergehende Einordnung zu ermöglichen.

Im nächsten Abschnitt folgt eine Darstellung der verwendeten Vorausberechnungsmethode. In den anschließenden Abschnitten werden die Vorausberechnungen der demographischen Komponenten jeweils kurz diskutiert und ihre Ergebnisse präsentiert. Darauf werden das Vorgehen bei der Bevölkerungsvorausberechnung und deren Ergebnisse dargestellt. Das Kapitel schließt mit einer kurzen Zusammenfassung.

### 8.2 Nichtparametrisch-stochastische Vorausberechnung demographischer Komponenten

#### 8.2.1 Parametrische und nichtparametrische Vorausberechnungsverfahren

Eine umfassende, allgemeine Übersicht zu Verfahren zur Vorausberechnung von demographischen Komponenten findet sich bei Booth (2006). Bei den meisten in der Literatur vorgeschla-

---

<sup>177</sup> Prinzipiell ist wie zu Beginn von Kapitel 4 und am Ende von Kapitel 5 angedeutet noch eine weitere Disaggregation denkbar. Beispielsweise könnten Sterberaten nach Todesursache oder paritätsspezifische Geburtenraten betrachtet werden. Diese sind aber zum einen für das verwendete Modell für Verwandtschaft nicht notwendig, zum anderen wird durch diese Disaggregation in aller Regel keine Verbesserung der Vorausberechnung erreicht (vgl. Alho & Spencer 1990a; Bell 1997: 283; Booth & Tickle 2008: 19f.; de Beer 2000: 7f.).

genen Verfahren handelt es sich um parametrische Modelle, bei denen demographische Raten in Abhängigkeit von einer Kombination der Faktoren Alter, Periode und Kohorte modelliert werden (vgl. ebd. 549). Vereinfacht dargestellt werden hierbei oftmals Dummy-Variablen für einzelne Jahre, Alter oder Kohorten in einer Regressionsgleichung verwendet. Die Vorausberechnung erfolgt dann, indem ein Modell an historische Daten angepasst wird und die resultierenden Parameter vorausberechnet werden (vgl. Congdon 1993: 240), wobei in der Regel Verfahren für Zeitreihen Verwendung finden (vgl. Lee 1998: 176).<sup>178</sup>

Als Alternative zur Kombination parametrischer Modelle mit Zeitreihenverfahren nennt Lee (1998: 176) zwei Möglichkeiten. Zum einen kann für einzelne Raten oder Kennwerte von demographischen Komponenten eine theoretische Verteilung spezifiziert werden, aus der dann zufällig gezogen wird. Zum anderen können zufällige Werte aus historischen Zeitreihen gezogen werden (*resampling*), wobei Lee (1998: 176) zu diesem letzten Ansatz anmerkt, dass er bisher nicht verwendet worden sei.<sup>179</sup> Allerdings sind hier Arbeiten von Keyfitz (1985, 1989) als Gegenbeispiel zu nennen.

Bei Keyfitz' Ansatz wird davon ausgegangen, dass die zukünftige Variabilität der einzelnen Komponenten der Variabilität in der Vergangenheit entspricht (vgl. Keyfitz 1989: 1). Um Mortalität vorauszuberechnen, wird für die historische Entwicklung der Lebenserwartung berechnet, um welchen Wert sich diese zwischen einzelnen Jahren verändert hat (vgl. Keyfitz 1985: 180). Aus der Zeitreihe dieser Veränderungen wird nun zufällig eine Abfolge von Veränderungen mit Zurücklegen gezogen und angenommen, dass diese Abfolge eine mögliche zukünftige Entwicklung der Veränderung der Lebenserwartung wiedergibt (vgl. ebd.: 180).<sup>180</sup> Für Geburtenraten werden Vorausberechnungen erstellt, indem zufällig die Geburtenraten aus einem Jahr gezogen werden und für fünf Jahre der zukünftigen Entwicklung gesetzt werden (vgl. Keyfitz 1989: 2). Bei Migration wird zufällig aus der Zeitreihe des Umfangs der Nettomigration gezogen und die resultierende Abfolge mit der Altersstruktur der Nettomigration aus dem letzten beobachteten Jahr kombiniert (vgl. ebd.: 5). Durch wiederholtes simulieren möglicher Entwicklungen der einzelnen Komponenten und einer Einspeisung dieser in Vorausberechnungen mit der Kohorten-Komponenten-Methode lässt sich hierdurch eine stochastische Bevölkerungsvorausberechnung vornehmen (vgl. Keyfitz 1985: 188).

Ein ähnliches Vorgehen für die Vorausberechnung von Mortalität wurde von Denton et al. (2005) entwickelt. Ausgangspunkt ist allerdings nicht die historische Veränderung der Lebenserwartung, sondern die historische Veränderung von logarithmierten Sterberaten (vgl.

---

178 Innerhalb dieses Rahmens existiert sowohl zur Modellierung von Mortalität als auch von Fertilität eine umfangreiche Literatur, die auch etliche Erweiterungen dieses allgemeinen Ansatzes enthält. Übersichten geben Booth (2006) und Tabeau (2001). Eine bei weitem nicht vollständige Auswahl spezifischer Modelle umfasst Bell (1997); Cheng & Lin (2010); Carter & Lee (1992); Carter (2010); Congdon (1993); de Beer (1985, 1989); Denton et al. (2005); Girosi & King (2007); Goldstein et al. (1994); Hyndman & Booth (2008); Keilman & Pham (2000); Koissia et al. (2006); Lee & Carter (1992); Lee (1993, 2000); Lee & Miller (2001); Li & Wu (2003); Li & Lee (2005); Li et al. (2011); Liu & Braun (2010); Lundström & Qvist (2004); Pullum (1980); Renshaw & Haberman (2003); Thompson et al. (1989); Willekens & Baydar (1984); Wilmoth (1993). Evaluationen und Vergleiche einiger dieser Ansätze finden sich bei Bell (1997); Booth et al. (2005); Booth & Tickle (2008); Cairns et al. (2009); Dowd et al. (2008); Haberman & Renshaw (2011).

179 Der Begriff *resampling* wird teils auch im Kontext von Bootstrap-Verfahren für parametrische Vorausberechnungsansätze verwendet (s. bspw. Goldstein et al. 1994). Im weiteren wird diese Bezeichnung ausschließlich für das im weiteren noch genauer vorzustellende Verfahren benutzt.

180 Es findet sich in keiner der beiden zugrunde gelegten Publikationen (Keyfitz 1985, 1989) ein Hinweis darauf, wie aus der Lebenserwartung Sterberaten abgeleitet wurden.

ebd.: 218). Für alle aufeinander folgenden Zeitstellen werden hierbei die Differenzen der logarithmierten altersspezifischen Sterberaten gebildet, so dass Vektoren mit Differenzen vorliegen (vgl. ebd.: 218). Ein zukünftiger Verlauf der Entwicklung der Sterberaten wird nun simuliert, indem zufällig aus diesen Vektoren gezogen wird (vgl. ebd.: 218). Hierbei werden allerdings nicht einzelne Vektoren gezogen, sondern Blöcke von jeweils zeitlich aufeinander folgenden Vektoren (vgl. ebd.: 218). Dieses Vorgehen hat den Vorteil, dass zumindest innerhalb der Blöcke eine mögliche Autokorrelation der Differenzen über mehrere Zeitstellen hinweg und eine mögliche Korrelation der Differenzen zwischen Altern erhalten bleibt (vgl. ebd.: 218).<sup>181</sup> Da keine Parameter geschätzt werden müssen, ist der Ansatz als nichtparametrisch aufzufassen, und es sind keine Verteilungsannahmen notwendig, wie sie bei parametrischen Modellen benötigt werden (vgl. ebd.: 218).<sup>182</sup>

Gerade diese zuletzt genannte Eigenschaft des nichtparametrischen *resampling* Verfahrens macht dieses attraktiv. Das Vorgehen von Denton et al. (2005) soll im weiteren auf Mortalität angewendet und auf Fertilität und Migration übertragen werden. Im nächsten Unterabschnitt erfolgt zunächst eine formale Beschreibung des Ansatzes sowie eine kurze, allgemeine Diskussion möglicher Varianten.

### 8.2.2 Formale Definition des Ansatzes

Gegeben sei eine Matrix  $\mathbf{C}$  mit Einträgen  $c_{x,t}$ , wobei ein Eintrag eine demographische Rate für das Alter  $x$  im Jahr  $t$  repräsentiert. Gegebenenfalls werden Transformationen  $\mathbf{C}^* = (c_{x,t}^*) = (f[c_{x,t}])$  betrachtet, wobei  $g(\cdot) = f^{-1}(\cdot)$  sei. Die Zahl der Jahre sei gleich  $T$  und die Zahl der Alter gleich  $A$ , so dass  $\mathbf{C}$  beziehungsweise  $\mathbf{C}^*$   $A$  Zeilen und  $T$  Spalten aufweist.  $\mathbf{D}$  sei eine  $(A \times T - 1)$  Matrix mit Einträgen  $d_{x,t} = c_{x,t+1}^* - c_{x,t}^*$ .

Sei  $\mathbf{d}_t$  der Spaltenvektor  $t$  der Matrix  $\mathbf{D}$  und entsprechend  $\mathbf{c}_t^*$  ein Spaltenvektor von  $\mathbf{C}^*$ . Soll eine Vorausberechnung bis einschließlich zum Jahr  $T + \tau$  durchgeführt werden, wird im einfachsten Fall eine Folge von  $\tau$  Zufallszahlen  $(s_1, \dots, s_j, \dots, s_\tau)$  aus der Menge  $\{1, \dots, T\}$  mit Zurücklegen und gleichen Ziehungswahrscheinlichkeiten generiert. Die Spalte  $i$  der Matrix der vorausberechneten Raten  $\mathbf{C}^\tau$  ergibt sich dann als

$$\mathbf{c}_i^\tau = g \left( \mathbf{c}_T^* + \sum_{j=1}^i \mathbf{d}_{s_j} \right) \quad (8.1)$$

Bei diesem Vorgehen ist zu beachten, dass die Transformation  $f(\cdot)$  einen Einfluss darauf hat, wie die Berechnung zu interpretieren ist. Wenn beispielsweise  $f(\cdot) = \log(\cdot)$  und somit  $g(\cdot) = \exp(\cdot)$  ist, dann erhält man

$$\mathbf{c}_i^\tau = \mathbf{c}_T \prod_{j=1}^i \frac{\mathbf{c}_{s_j+1}}{\mathbf{c}_{s_j}} \quad (8.2)$$

181 Eine ähnliche Idee wird beim sogenannten *block bootstrap* Verfahren verwendet, welches im Rahmen von parametrischen Zeitreihenmodellen Verwendung findet (vgl. Berkowitz & Kilian 2000; Li & Maddala 1996).

182 Entsprechend der in den Unterabschnitten 5.1.2 und 5.1.3 eingeführten Terminologie kann der Ansatz als stochastisch bezeichnet werden, da von der Grundidee her ein stochastischer Prozess angenommen und nichtparametrisch modelliert wird. Bei der Verwendung von parametrischen Ansätzen wird der stochastische Prozess hingegen über konkrete Verteilungsannahmen modelliert.

wenn man  $\mathbf{c}_{s_j+1}/\mathbf{c}_{s_j}$  für elementweise Division schreibt. Die letzten beobachteten Raten  $\mathbf{C}_T$  werden also einfach mit Veränderungsraten multipliziert. Bei anderen Transformationen  $f(\cdot)$  erhält man gegebenenfalls andere Interpretationen.

Ausgehend von diesem einfachen Ansatz sind zwei Variationen denkbar. Zunächst kann man nicht nur einzelne Spalten  $\mathbf{d}_t$  ziehen, sondern mehrere aufeinander folgende Spalten – also Spaltenblöcke – zugleich (vgl. Denton et al. 2005: 218). Ausgegangen wird von Blöcken der Länge  $k$ , wobei  $k \leq \tau$  gilt. Wenn  $(\tau \bmod k) = 0$  ist die Zahl der zu ziehenden Blöcke gleich  $\tau/k$ . Für den Fall, dass  $(\tau \bmod k) \neq 0$  ist, werden  $\lceil \tau/k \rceil$  Blöcke gezogen, wobei der letzte zu ziehende Block die Länge  $(\tau \bmod k)$  aufweist. Sei  $z$  die Zahl der zu ziehenden Blöcke, mit Blocklänge  $k_1, \dots, k_z$ . Dann werden  $z$  Zufallszahlen  $(s_1, \dots, s_j, \dots, s_z)$  gezogen und zwar in Abhängigkeit von  $k_j$  aus der Menge  $\{1, \dots, T - k_j\}$  mit Zurücklegen und gleichen Ziehungswahrscheinlichkeiten. Für die Vorausberechnung werden die Spalten  $(s_1, s_1 + 1, \dots, s_1 + k_1 - 1, \dots, s_z, s_z + 1, \dots, s_z + k_z - 1)$  verwendet, wobei ansonsten wie bisher vorgegangen werden kann. Die letzte Spalte der Matrix  $\mathbf{C}^\tau$  ergibt sich nun beispielsweise als

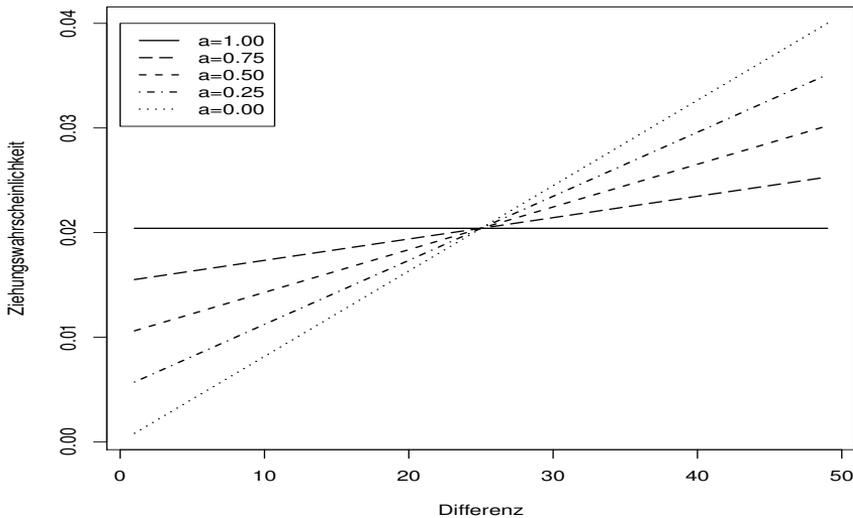
$$g \left( \mathbf{c}_T^* + \sum_{j=1}^z \sum_{l=0}^{k_j-1} \mathbf{d}_{s_j+l} \right) \quad (8.3)$$

Eine zweite Variation besteht darin, unterschiedliche Ziehungswahrscheinlichkeiten für die einzelnen Spalten zu verwenden. Ein Vorschlag, wie solche Ziehungswahrscheinlichkeiten berechnet werden können, wurde von Keyfitz (1989: 5) gemacht. Die Ziehungswahrscheinlichkeit für Spalte  $i$ ,  $w_i$ , ergibt sich bei diesem als

$$w_i = \frac{a}{S} + \frac{2i(1-a)}{S(S+1)} \quad \text{für } 0 \leq a \leq 1 \quad (8.4)$$

wobei  $S$  der Zahl der Spalten der Matrix entspricht, aus der gezogen werden soll. Mit  $a = 1$  erhält man gleiche Ziehungswahrscheinlichkeiten für alle Spalten. Für  $a = 0$  erhält man Ziehungswahrscheinlichkeiten, die umso höher sind, je näher eine Spalte an  $S$  liegt, wobei die Wahrscheinlichkeit für Spalte 1,  $w_1$ , am niedrigsten und die Wahrscheinlichkeit für Spalte  $S$ ,  $w_S$ , am höchsten ausfällt und  $w_S/w_1 = S$  gilt. Die Gewichte für die dazwischen liegenden Spalten steigen linear an. Für  $0 < a < 1$  erhält man Gewichte, die zwischen diesen beiden Extremen liegen. In Abbildung 8.1 sind als Beispiel die Ziehungswahrscheinlichkeiten eingezeichnet, die sich bei unterschiedlichen Werten für  $a$  ergeben, wenn  $T = 50$  Jahre und somit  $T - 1 = 49 = S$  Differenzen betrachtet werden. Dabei gibt ein Wert  $t$  an der Abszisse an, dass es sich um die Differenz zwischen den Jahren  $t$  und  $t + 1$  handelt.

Abbildung 8.1: Ziehungswahrscheinlichkeiten bei 49 Differenzen und unterschiedlichen Werten für  $a$



Quelle: eigene Berechnung

### 8.3 Nichtparametrische Vorausberechnung von Sterberaten

#### 8.3.1 Umsetzung des Ansatzes bei Sterberaten

Bei der Vorausberechnung der Sterberaten wurde zunächst von der Entwicklung in Westdeutschland ausgegangen. Als Datengrundlage wurden hierbei die in Anhang A.1 beschriebenen Sterberaten aus der Human Mortality Database (HMD) für den Zeitraum von 1956 bis 2008 verwendet. Diese Sterberaten wurden Logit-transformiert, so dass bei der Vorausberechnung lediglich Raten zwischen 0 und 1 auftreten können.<sup>183</sup> Für die Vorausberechnung zeigte sich, dass ein Wert von  $k = 15$  und  $a = 1$  – also eine Blocklänge von 15 Differenzen bei gleicher Ziehungswahrscheinlichkeit der Blöcke – plausible Ergebnisse liefert (s. Unterabschnitt 8.3.3).<sup>184</sup>

<sup>183</sup> Entsprechend der Definition demographischer Raten in Abschnitt 3.1 sind zwar auch Sterberaten über 1 denkbar, scheinen aber unrealistisch. Die höchste im Zeitraum von 1956 bis 2008 bei westdeutschen Frauen aufgetretene Sterberate beträgt beispielsweise 0.83.

<sup>184</sup> Bei anderen Werten für  $k$  und  $a$  ändert sich verglichen mit  $k = 15$  und  $a = 1$  primär die Breite der Vorausberechnungsintervalle. Für niedrigere Werte für  $k$  und  $a$  ist diese höher, allerdings ergibt sich dann ein relativ unregelmäßiger, unrealistischer Verlauf von Sterberaten über Alter hinweg. Ähnliche Effekte der Werte von  $k$  und  $a$  wurden auch von Denton et al. (2005: 224) festgestellt. Eine weitergehende Validierung erfolgte, indem Vorausberechnungen durchgeführt wurden, bei denen lediglich ein Teil der Daten bis zu einem Jahr  $t < 2008$  verwendet wurde. Beispielsweise wurde unter anderem eine Vorausberechnung durchgeführt, bei der lediglich die Daten der Jahre 1956 bis 1986 berücksichtigt wurden. Die Resultate solcher Vorausberechnungen wurden dann anschließend mit der tatsächlichen Entwicklung verglichen. Für  $k = 15$  zeigte sich, dass bei einer Datengrundlage ab etwa 30 Jahren insgesamt relativ geringe Abweichungen zwischen den Vorausberechnungen und der tatsächlichen Entwicklung vorliegen.

Insgesamt wurde das im vorausgegangenen Abschnitt beschriebene Verfahren 1000 mal durchgeführt, so dass 1000 Simulationsinstanzen vorlagen, die jeweils bis einschließlich zum Jahr 2059 reichten.<sup>185 186 187</sup> Die Vorausberechnung von Sterberaten für westdeutsche Männer und Frauen erfolgte dabei jeweils „parallel“, indem für eine Simulationsinstanz die gleichen Zufallszahlen  $(s_1, \dots, s_z)$  für Männer und Frauen verwendet wurden, so dass eine mögliche Korrelation der Entwicklungen der Lebenserwartungen der beiden Geschlechter erhalten bleibt (s. auch Denton et al. 2005: 218).

Da sich aus den Sterberaten für ostdeutsche Frauen aus dem Jahr 2008 praktisch die selbe Lebenserwartung ergibt wie für westdeutsche Frauen, wurde davon ausgegangen, dass die Sterblichkeit west- und ostdeutscher Frauen in Zukunft identisch ist. Bei ostdeutschen Männern hingegen ergibt sich für das Jahr 2008 eine Lebenserwartung von 76.49 Jahren, während man bei westdeutschen Männern einen Wert von 77.55 Jahren erhält, was einer absoluten Differenz von 1.06 Jahren entspricht. Aus diesem Grund wurden west- und ostdeutsche Männer zunächst getrennt betrachtet. Bei der Vorausberechnung der Sterberaten ostdeutscher Männer wurden nur HMD-Daten ab dem Jahr 1990 berücksichtigt, so dass im Wesentlichen nur die zum Westen aufholende Entwicklung berücksichtigt wurde, wobei  $k = 1$  und  $a = 1$  gesetzt wurde. Die Vorausberechnung wurde dann zunächst unabhängig von westdeutschen Männern durchgeführt. Ab dem ersten Jahr der Vorausberechnung, bei dem sich für ostdeutsche Männer eine Lebenserwartung gleich oder höher der der westdeutschen Männer ergab, wurde von identischen Sterberaten ausgegangen und die westdeutschen Raten für Ostdeutschland übernommen.<sup>188</sup>

Für alle 1000 Simulationen wurde die Entwicklung der Lebenserwartung nach Geschlecht und Landesteil berechnet und für jedes Jahr der Vorausberechnung die Mediane und 90%-Vorausberechnungsintervalle der resultierenden Verteilungen erstellt, wobei das Vorausberechnungsintervall auf dem 5%- und dem 95%-Quantil der Verteilungen basiert. Soweit nicht anders angegeben, werden Vorausberechnungsintervalle im weiteren immer auf diesen Quantilen basieren. Bei der Betrachtung der Mediane über die einzelnen Jahre der Vorausberechnung ist dabei zu beachten, dass sich diese aus Ergebnissen mehrerer Simulationsinstanzen zusammensetzen und nicht einer konkreten Simulation zugeordnet werden können. Gleiches gilt für die Vorausberechnungsintervalle.

### 8.3.2 Ergebnisse

Die vorausberechnete Entwicklung der ferneren Lebenserwartung bei der Geburt für west- und ostdeutsche Frauen bis 2059 ist aus Tabelle 8.1 zu ersehen. Die Entwicklung der Lebens-

---

185 Entsprechend der Ausführungen in Unterabschnitt 3.1.5 und Kapitel 5 reicht es für eine Vorausberechnung der Bevölkerung und der Verwandtschaft bis zu einem Jahr  $t$ , die demographischen Raten bis einschließlich für das Jahr  $t - 1$  vorliegen zu haben.

186 Der Begriff „Simulationsinstanz“ wird im weiteren für eine Realisierung der Simulation, also für einen „Simulationsdurchlauf“, verwendet.

187 Dies entspricht einem Vorausberechnungszeitraum von grob 50 Jahren, wie bei anderen demographischen Vorausberechnungen nicht unüblich (vgl. Bretz 2000: 644f.).

188 Bei stochastischen Vorausberechnungen für Deutschland wurden bisher in der Regel die westdeutschen Raten vollständig für Ostdeutschland übernommen (vgl. Betz & Lipps 2004: 10; Härdle & Mysickova 2009: 5; Lipps & Betz 2005: 22f.).

*Tabelle 8.1: Ausgewählte Ergebnisse für die stochastische Vorausberechnung der Mortalität*

Größe	Kennwert	2008	2020	2030	2040	2050	2059
Lebenserwartung (Frauen)	5%-Quantil	—	83.25	84.72	85.97	87.11	88.16
	Median/Ausgangswert	82.44	84.42	85.90	87.27	88.50	89.56
	95%-Quantil	—	85.25	86.78	88.38	89.68	90.82
Lebenserwartung (Männer/West)	5%-Quantil	—	77.47	78.50	79.39	80.56	81.68
	Median/Ausgangswert	77.55	79.72	81.05	82.35	83.50	84.77
	95%-Quantil	—	80.50	82.38	84.19	85.63	86.96
Lebenserwartung (Männer/Ost)	5%-Quantil	—	77.34	78.44	79.32	80.47	81.61
	Median/Ausgangswert	76.49	79.65	81.04	82.34	83.50	84.75
	95%-Quantil	—	80.36	82.34	84.19	85.63	86.96

Quelle: eigene Berechnung

erwartung west- und ostdeutscher Männer findet sich ebenfalls in der selben Tabelle. Wird die Entwicklung der Mediane der Verteilungen betrachtet, ergibt sich sowohl bei Frauen als auch bei westdeutschen Männern ausgehend vom letzten beobachteten Wert im Jahr 2008 ein Anstieg der Lebenserwartung bis 2059 von etwa 7 Jahren, bei ostdeutschen Männern ein Anstieg von ungefähr 8 Jahren. Bei Frauen korrespondiert die untere Grenze des 90%-Vorausberechnungsintervalls zu einem Anstieg von 6 Jahren und das obere zu einem Anstieg von etwas über 8 Jahren. Bei westdeutschen Männern liegen diese Werte bei ungefähr 4 und 8 Jahren, bei ostdeutschen Männern bei etwa 5 und 9 Jahren.

### 8.3.3 Vergleich mit Ergebnissen aus der Literatur

Stochastische, parametrische Vorausberechnungen der Mortalität für Deutschland wurden unter anderem von Babel et al. (2008) vorgelegt. Diese wurden ausgehend vom Jahr 2002 bis zum Jahr 2050 durchgeführt, wobei keine Unterscheidung zwischen West- und Ostdeutschland vorgenommen wurde. Die Autoren verwendeten ihr Verfahren in mehreren Varianten, wobei die Medianwerte der Lebenserwartung im Jahr 2050 sowohl für Frauen als auch für Männer relativ stabil sind und 88.3 und 82.4 Jahre betragen (vgl. ebd.: 548). Der Wert für Frauen entspricht ungefähr dem beim hier verwendeten Verfahren erreichten Wert, während der Wert bei Männern um etwas mehr als ein Jahr niedriger liegt. Je nach Variante fallen die Vorausberechnungsintervalle unterschiedlich breit aus, wobei die Autoren ebenfalls 90%-Intervalle ausweisen (vgl. ebd.: 548). Bei Frauen reicht die Breite der Intervalle von 0.4 bis 7.8 Jahre, bei Männern von 0.6 bis 9.4 Jahre. Die Breite der Intervalle bei der nichtparametrischen Vorausberechnung von etwa 2.6 Jahren bei Frauen entspricht grob der Variante *B1.II*, bei der eine Breite von 2.9 Jahren erreicht wird (vgl. ebd.: 548). Bei Männern liegt die Breite von circa 5.1 Jahren zwischen den Varianten *B1.II* (4.0 Jahre) und *B2.I* (5.8 Jahre) (vgl. ebd.: 548). Betrachtet man die Bandbreite und die in den einzelnen Varianten konkret realisierten Werte der von Babel et al. (2008) durchgeführten Vorausberechnungen, scheint die durch die nichtparametrischen Methode abgebildete Unsicherheit im „mittleren Bereich“ zu liegen.

Weitere Vergleichswerte liefern Betz & Lipps (2004), die ein von Lee & Carter (1992) vorgeschlagenes, parametrisches Verfahren verwenden. Die Vorausberechnungen beginnen ab dem Jahr 2003 und wurden bis zum Jahr 2050 durchgeführt. Auch hier liegen wieder keine nach West- und Ostdeutschland differenzierten Ergebnisse vor, zudem sind lediglich die Mittelwerte der stochastischen Vorausberechnungen ausgewiesen (vgl. Betz & Lipps 2004: 11). Für Frauen ergibt sich für das Jahr 2050 durchschnittlich eine Lebenserwartung von 86.7 Jahren, bei Männern wird ein Wert von 80.3 Jahren erreicht (vgl. Betz & Lipps 2004: 11). Diese Resultate liegen etwas unter den unteren Vorausberechnungsintervallen der nichtparametrischen Vorausberechnung. Bei diesem Vergleich erscheinen die Ergebnisse der nichtparametrischen Methode eine vergleichsweise hohe Steigerung der Lebenserwartung zu implizieren.

Das Statistische Bundesamt verwendete bei der zwölften koordinierten Bevölkerungsvorausberechnung zwei unterschiedliche Annahmen für die zukünftige Entwicklung der Mortalität. Beim ersten Szenario wird bis 2060 von einem Anstieg der Lebenserwartung von Frauen auf 89.2 Jahre und bei Männern auf 85 Jahre ausgegangen (vgl. Statistisches Bundesamt 2010c: 30). Beim zweiten Szenario werden Werte von 91.2 Jahren (Frauen) und 87.7 Jahren (Männer)

angenommen (vgl. Statistisches Bundesamt 2010c: 30). Vergleicht man diese Annahmen des Statistischen Bundesamtes für das Jahr 2060 mit den Werten der nichtparametrischen Vorausberechnung für das Jahr 2059, zeigt sich, dass die niedrigere Variante des Statistischen Bundesamtes grob mit dem Medianverlauf und die obere Variante grob mit der oberen Grenze des 90%-Vorausberechnungsintervalls korrespondiert.<sup>189</sup>

Insgesamt lässt sich festhalten, dass eine Einschätzung, ob die Ergebnisse des nichtparametrischen Ansatzes eher hoch oder niedrig ausfallen, von den zum Vergleich herangezogenen Verfahren beziehungsweise Annahmen abhängt, allerdings erscheinen sie nicht unplausibel.

## 8.4 Nichtparametrische Vorausberechnung von Geburtenraten

### 8.4.1 Umsetzung des Ansatzes bei Geburtenraten

Bei der Vorausberechnung von Geburtenraten wurde zunächst von westdeutschen Frauen ausgegangen. Für diese wurden die in Anhang A.3 beschriebenen Geburtenraten ab dem Jahr 1950 verwendet. Da die westdeutschen Daten für die Jahre ab 1950 bis einschließlich 1989 nur für die Alter von 15 bis 44 Jahren vorliegen, wurden die Alter von 45 bis 49 Jahren bei den Vorausberechnungen zunächst nicht berücksichtigt und erst in einem zweiten Schritt durch eine Verkettung von Regressionsschätzungen ergänzt. Diese Ergänzung erfolgte, indem die Geburtenrate für das Alter von 45 Jahren mittels einer mit den Daten der Jahre 1990 bis 2008 geschätzten Regressionsgleichung in Abhängigkeit von der Rate im Alter von 44 Jahren bestimmt wurde. Ausgehend von der Rate für das Alter 45 wurde auf gleiche Weise die Rate für das Alter von 46 Jahren geschätzt. Für die Alter von 47, 48 und 49 Jahren wurde analog vorgegangen. Beispielsweise wurde die Geburtenrate für das Alter von 47 Jahren über die Geburtenrate für das Alter von 46 Jahren geschätzt.<sup>190</sup>

Da sich bei den Vorausberechnungen in einzelnen Fällen relativ hohe beziehungsweise niedrige Werte für die zusammengefasste Geburtenziffer ergaben, wurden Simulationsverläufe, bei denen diese Werte unter 1 und über 3 erreichte, verworfen.<sup>191</sup> Wenn dann im weiteren

---

189 Bei diesem und den im weiteren folgenden Vergleichen wird davon ausgegangen, dass die Unterschiede in den Bezugsjahren zwischen den Annahmen und Berechnungen des Statistischen Bundesamtes (2060) und der nichtparametrischen Vorausberechnung (Komponenten bis 2059, Bevölkerung bis 2060) nur einen unwesentlichen Einfluss haben und vernachlässigt werden können.

190 Im Falle der Regression der Geburtenrate für das Alter von 44 Jahren auf die Geburtenrate für das Alter von 45 Jahren ergibt sich ein Determinationskoeffizient von  $R^2 = 0.896$ . Bei den weiteren Regressionsrechnungen nimmt der Determinationskoeffizient immer weiter ab, bis bei der Regression der Geburtenrate für das Alter von 48 auf die Geburtenrate für das Alter von 49 Jahren ein Wert von  $R^2 = 0.324$  vorliegt. Zwar ist somit die Anpassungsgüte bei höheren Altern teilweise nicht sonderlich hoch, allerdings sind die resultierenden Schätzwerte plausibel und fallen sehr gering aus, womit sie keinen wesentlichen Einfluss auf die Ergebnisse haben. Aus diesem Grunde wurde auch die Unsicherheit, die aus der nicht perfekten Anpassung resultiert, nicht weiter berücksichtigt.

191 Das Problem unrealistischer Werte der zusammengefassten Geburtenziffer tritt auch bei parametrischen Vorausberechnungen auf. Keilman & Pham (2000: 50) verwerfen wie hier Simulationsverläufe, die ihnen zu extrem erscheinen, während Lee (1993: 192ff.) durch eine mathematische Transformation erreicht, dass die Ergebnisse in einem bestimmten Intervall liegen (s. auch Keilman & Pham 2004: 6). Beim nichtparametrischen Vorgehen bietet sich diese Transformation allerdings nicht an, da für jede altersspezifische Geburtenrate eine Ober- und Untergrenze der zulässigen Werte festgelegt werden müsste und nicht klar ist, wie diese gewählt werden sollten.

Verlauf die Rede davon ist, dass eine bestimmte Zahl  $x$  an Simulationen durchgeführt wurde, meint dies, dass  $x$  Simulationen erzeugt wurden, deren Resultate im zulässigen Bereich liegen.

Während die Obergrenze der zusammengefassten Geburtenziffer von 3 nicht allzu restriktiv scheint, ist die Untergrenze von 1 problematischer. Wie in Abschnitt 3.2 beschrieben, lag beispielsweise die zusammengefasste Geburtenziffer in Ostdeutschland im Jahr 1994 bei etwa 0.77 und somit deutlich unter 1. Allerdings ist zu bedenken, dass die Grenzen nur für die Fertilität der Frauen gelten. Die zusammengefasste Geburtenziffer für Männer wird bei der gewählten Modellierung neben der Fertilität der Frauen auch noch von Migration beeinflusst. Wie die Ergebnisse aus Kapitel 6 vermuten lassen, kann hierdurch bei einer zusammengefassten Geburtenziffer von Frauen etwas über 1 und vergleichsweise starker Migration die zusammengefasste Geburtenziffer von Männern deutlich unter 1 liegen. Somit ist zwar für Frauen eine zusammengefasste Geburtenziffer unter 1 durchaus denkbar, allerdings verhindert die Untergrenze, dass die Fertilität von Männern allzu leicht extrem niedrige Werte annimmt. Zudem wurden mit den gewählten Grenzen relativ plausible Ergebnisse erzielt (s. Fußnote 192 und Unterabschnitt 8.4.3).

Für die eigentliche Vorausberechnung wurden die Geburtenraten logarithmiert und  $k = 1$  und  $a = 0.5$  gesetzt, womit relativ plausible Ergebnisse bezogen auf die Entwicklung der zusammengefassten Geburtenziffer und des durchschnittlichen Alters bei der Geburt erzielt werden konnten. Bei anderen Werten für  $k$  und  $a$  war dies oftmals aber nicht der Fall.<sup>192</sup> Bei den gewählten Werten zeigte sich wiederum das Problem, dass die resultierenden Altersprofile der Geburtenraten teils sehr unregelmäßig waren. Um dieses Problem zu beheben, wurde wie folgt vorgegangen. Zunächst wurden für jedes vorausberechnete Jahr standardisierte Geburtenraten  $f_x^* = f_x / \sum_x f_x$  berechnet. Diese wurden anschließend Logit-transformiert zu  $z_x = \log[f_x^* / (1 - f_x^*)]$ . Die resultierenden Werte wurden über kubische Splines geglättet.<sup>193</sup> Die geglätteten Werte  $z_x^*$  wurden anschließend über  $\exp[z_x^*] / (1 + \exp[z_x^*]) = f_x^z$  zurück transformiert. Um zu gewährleisten, dass  $\sum_x f_x^z = 1$  gilt, wurden die Werte eines Jahres durch ihre Summe dividiert. Die resultierenden Werte wurden dann schließlich mit der vorausberechneten zusammengefassten Geburtenziffer des entsprechenden Jahres multipliziert. Hierdurch werden die vorausberechneten zusammengefassten Geburtenziffern nicht verändert.

Mit diesem Vorgehen wurden 1000 Simulationsinstanzen erzeugt, die bis zum Jahr 2059 reichen.

Bei der Übertragung des Verfahrens auf ostdeutsche Geburtenraten stellte sich das Problem, dass die historische Entwicklung der Fertilität in der DDR beziehungsweise Ostdeutschland unregelmäßiger verlief als in Westdeutschland. Beispielsweise entwickelte sich die zusammengefasste Geburtenziffer in beiden Teilen Deutschlands recht unterschiedlich. Nach dem starken Absinken der Fertilität nach der Wiedervereinigung hat sich die zusammengefasste Geburtenziffer für Ostdeutschland mittlerweile in etwa dem westdeutschen Wert angeglichen. Allerdings gibt es noch Unterschiede im durchschnittlichen Alter bei der Geburt, welches

---

192 Das in Fußnote 184 auf Seite 167 beschriebenen Vorgehen zur Validierung der gewählten Werte für  $a$  und  $k$  wurde auch bei Geburtenraten verwendet, wobei die vorausberechnete zusammengefasste Geburtenziffer mit der tatsächlichen Entwicklung verglichen wurde. Hier zeigten sich deutlichere Differenzen als bei der Mortalität. Nichtsdestotrotz erwiesen sich die 90%-Vorausberechnungsintervalle als eher konservativ – sie deckten immer alle historischen Werte der zusammengefassten Geburtenziffer ab.

193 Kubische Splines gelten allgemein als geeignetes Mittel, um Altersprofile demographischer Raten zu glätten (vgl. McNeil et al. 1977; Hoem et al. 1981). Hier wurden Splines mit fünf freien Knoten verwendet.

im Jahr 2008 im Westen bei 30.18 Jahren und im Osten bei 29.09 Jahren lag, was einer Differenz von 1.09 Jahren entspricht. Dabei ist das durchschnittliche Alter bei der Geburt seit der Wiedervereinigung aber in Ostdeutschland stärker gestiegen als in Westdeutschland.

Aufgrund dieser Entwicklung wurde ein Vorgehen verwendet, bei dem die zusammengefasste Geburtenziffer für Ostdeutschland der für Westdeutschland entspricht, das durchschnittliche Alter bei der Geburt sich aber in gewissen Grenzen unterschiedlich entwickeln kann. Hierfür wurden zunächst wie auch für Westdeutschland nichtparametrische Vorausberechnungen erstellt, die auf den Daten des Zeitraums 1990 bis 2008 basierten. Die Daten entsprechen bezüglich der zugrunde liegenden Definitionen den westdeutschen Daten für den selben Zeitraum (s. Anhang A.3). Es wurden Vorausberechnungen bis zum Jahr 2059 durchgeführt, wobei für alle Differenzen gleiche Ziehungswahrscheinlichkeiten, eine Blocklänge von  $k = 1$  und keine Ober- und Untergrenze für die resultierende zusammengefasste Geburtenziffer verwendet wurden. Da die Daten für die Altersklassen von 15 bis 49 Jahren vorliegen, ist die bei westdeutschen Daten nötige Schätzung von Geburtenraten für die Alter von 44 bis 49 Jahren nicht notwendig.

Insgesamt wurden wieder 1000 Simulationsinstanzen erzeugt. Diese Instanzen wurden der Reihe nach den Simulationsdurchläufen für Westdeutschland zugeordnet, also beispielsweise die erste Simulationsinstanz für Ostdeutschland der ersten Simulationsinstanz für Westdeutschland. Die sich bei einer Simulationsinstanz ergebenden zusammengefassten Geburtenziffern für Westdeutschland wurden auf Ostdeutschland übertragen, indem die altersspezifischen Geburtenraten für Ostdeutschland mit dem Verhältnis der zusammengefassten Geburtenziffer zwischen Ost- und Westdeutschland multipliziert wurden. Anschließend wurden die ostdeutschen Altersprofile wie im Falle Westdeutschlands geglättet. Schließlich wurde die vorausberechnete Entwicklung des durchschnittlichen Alters bei der Geburt zwischen den einander zugeordneten west- und ostdeutschen Simulationsinstanzen verglichen. Erreichte dieser Kennwert für eines oder mehrere der vorausberechneten Jahre für Ostdeutschland einen Wert, der höher oder gleich dem Westdeutschen ist, so wurden ab dem zeitlich ersten Jahr, für welches dies der Fall war, die westdeutschen Altersprofile für Ostdeutschland übernommen. Wenn also beispielsweise bei einer Vorausberechnung für das Jahr 2031 ein höherer Wert für Ostdeutschland erreicht wurde, so wurden für alle Jahre von einschließlich 2031 an die westdeutschen altersspezifischen Geburtenraten verwendet.

Dieses Vorgehen führt dazu, dass die Entwicklung des durchschnittlichen Alters bei der Geburt in Westdeutschland eine Obergrenze für die Entwicklung in Ostdeutschland bildet. Aufgrund der historischen Entwicklung dieses Kennwerts in Ostdeutschland in den Jahren von 1990 bis 2008, die durch einen starken Anstieg charakterisiert ist, ist bei den Vorausberechnungen also eine Angleichung an Westdeutschland sehr wahrscheinlich. Bei allen der 1000 Simulationen trat eine solche Angleichung ein, durchschnittlich etwa ab dem Jahr 2016.

#### 8.4.2 *Ergebnisse*

Der Median und das 90%-Vorausberechnungsintervall für die zusammengefasste Geburtenziffer bis 2059 finden sich in Tabelle 8.2. Betrachtet man den Median, ergibt sich ausgehend vom letzten beobachteten Wert von 1.37 bis zum Jahr 2059 ein leichter Anstieg auf 1.40. Die untere Grenze des 90%-Vorausberechnungsintervalls von 1.08 korrespondiert zu einer

Tabelle 8.2: Ausgewählte Ergebnisse für die stochastische Vorausberechnung der Fertilität

Größe	Kennwert	2008	2020	2030	2040	2050	2059
Zusammengefasste Geburtenziffer	5%-Quantil	—	1.15	1.11	1.12	1.10	1.08
	Median/Ausgangswert	1.37	1.38	1.37	1.38	1.40	1.40
	95%-Quantil	—	1.64	1.73	1.86	1.95	2.05
Durchschnittliches Alter (West)	5%-Quantil	—	30.23	30.57	30.82	31.10	31.36
	Median/Ausgangswert	30.18	30.78	31.22	31.62	31.96	32.24
	95%-Quantil	—	31.32	31.93	32.38	32.76	33.08
Durchschnittliches Alter (Ost)	5%-Quantil	—	30.23	30.57	30.82	31.10	31.36
	Median/Ausgangswert	29.09	30.77	31.22	31.62	31.96	32.24
	95%-Quantil	—	31.29	31.93	32.38	32.76	33.08

Quelle: eigene Berechnung

absoluten Abnahme von 0.29, die obere Grenze von 2.0 entspricht einer deutlichen Zunahme von 0.68.

In Tabelle 8.2 sind zudem der Median sowie die Ober- und Untergrenze des 90%-Vorausrechnungsintervalls für das durchschnittliche Alter bei der Geburt in West- und Ostdeutschland zu sehen. In Westdeutschland liegt der Anstieg des durchschnittlichen Alters bei der Geburt bis 2059 entsprechend der Ober- und Untergrenze des Vorausrechnungsintervalls in 90% der Simulationen zwischen 1.18 und 2.9 Jahren, der Median entspricht einem Anstieg von 2.06 Jahren. Wie zu sehen ist, entsprechen der Median und das Vorausrechnungsintervall für ostdeutsche Frauen bereits ab 2020 relativ genau den Werten für Westdeutschland. Dies liegt an der relativ schnellen Angleichung des Gebäralters an westdeutsche Verhältnisse, welche wie bereits erwähnt durchschnittlich im Jahr 2016 abgeschlossen ist.

#### 8.4.3 Vergleich mit Ergebnissen aus der Literatur

Bei der parametrisch-stochastischen Vorausberechnung von Betz & Lipps (2004: 28) wird für das Jahr 2050 eine durchschnittliche zusammengefasste Geburtenziffer von 1.41 erreicht und das 90%-Vorausrechnungsintervall reicht von 0.995 bis 1.814. Bomsdorf et al. (2008: 128), die ebenfalls ein parametrisches Verfahren einsetzen, weisen für dieses Jahr einen Median von 1.35 aus, das 5%-Quantil liegt bei 0.86 und das 95%-Quantil bei 1.83.

Der Median der nichtparametrischen Vorausberechnungen entspricht in etwa den genannten mittleren Werten. Allerdings fallen die Grenzen des Vorausrechnungsintervalls recht unterschiedlich aus und sind bei den parametrischen Vorausberechnungen niedriger, was wahrscheinlich mit auf die gesetzten Grenzen der zusammengefassten Geburtenziffer bei den nichtparametrischen Simulationen zurückzuführen ist. Die Intervallbreite von 0.85 entspricht dabei aber in etwa der von Betz & Lipps (2004), während sie bei Bomsdorf et al. (2008) mit einem Wert von 0.97 höher ausfällt.<sup>194</sup>

Das Statistische Bundesamt geht bei der zwölften koordinierten Bevölkerungsvorausberechnung von drei möglichen Entwicklungsszenarien für Fertilität aus. In der ersten Variante wird von einer über den gesamten Vorausrechnungszeitraum gleichbleibenden zusammengefassten Geburtenziffer von 1.4 ausgegangen (vgl. Statistisches Bundesamt 2010c: 27). Bei der zweiten Variante wird ein Anstieg der zusammengefassten Geburtenziffer bis 2025 auf einen Wert von 1.6 angenommen, der anschließend konstant bleibt (vgl. ebd.: 27). Die dritte Variante geht von einer Abnahme der zusammengefassten Geburtenziffer auf einen Wert von 1.2 aus, die sich bis zum Jahr 2060 erstreckt (vgl. ebd.: 27).

Die mittlere dieser drei Varianten entspricht im Wesentlichen dem Median, der sich bei der nichtparametrischen Vorausberechnung ergibt. Das untere und das obere Szenario des Statistischen Bundesamtes liegen aber deutlich innerhalb des Vorausrechnungsintervalls. In etwa 52% der Simulationen werden im Jahr 2059 Werte der zusammengefassten Geburtenziffer zwischen 1.2 und 1.6 erreicht.

Beim Szenario, bei dem die zusammengefasste Geburtenziffer konstant auf dem Wert 1.4 bleibt, wird von einem Anstieg des durchschnittlichen Alters bei der Geburt von 1.6 Jahren

---

194 Mit dem Begriff „Intervallbreite“ wird im weiteren die Differenz zwischen den Werten des 95%-Quantils und des 5%-Quantils bezeichnet.

bis zum Jahr 2020 ausgegangen (vgl. Statistisches Bundesamt 2010c: 27). Das Szenario einer langfristigen Abnahme der Fertilität auf eine zusammengefasste Geburtenziffer von 1.2 bis zum Jahr 2060 geht mit einem Anstieg des durchschnittlichen Alters um 2.0 Jahre einher, wobei keine Angabe gemacht wird, wie sich dieser Anstieg auf den Zeitraum von 2010 bis 2060 verteilt (vgl. ebd.: 27). Im dritten Szenario, bei dem die zusammengefasste Geburtenziffer auf den Wert von 1.6 steigt, wird eine Zunahme des durchschnittlichen Alters um 1.1 Jahre bis 2025 angenommen (vgl. ebd.: 27). Geht man davon aus, dass sich in der Variante mit dem hohen Anstieg des durchschnittlichen Alters bei der Geburt der Anstieg bis zum Jahr 2060 erstreckt, entspricht dieses Szenario in etwa dem Median, der sich für das Jahr 2059 bei der nichtparametrischen Vorausberechnung ergibt. Ein Anstieg von 1.6 Jahren bis 2020 liegt außerhalb des 90%-Intervalls und ist verglichen mit dem nichtparametrischen Modell relativ hoch, während der Anstieg von 1.1 Jahren bis 2025 zwar über dem Median, aber innerhalb des Intervalls liegt.

Fasst man die durchgeführten Vergleiche zusammen, deckt sich die Entwicklung des Medians der zusammengefassten Geburtenziffer mit anderen Vorausberechnungen und die Vorausberechnungsintervalle scheinen weder besonders breit noch schmal auszufallen.

## **8.5 Nichtparametrische Vorausberechnung von Ein- und Auswanderung**

### *8.5.1 Vorbemerkungen*

Die Modellierung von Migration gilt in der demographischen Forschung verglichen mit Mortalität und Fertilität als eher vernachlässigt (vgl. Booth 2006: 562). Ursache hierfür ist zum einen, dass die Messung und Operationalisierung von Migration als relativ schwierig gilt (vgl. Brown & Bean 2006: 347). Beispielsweise lassen sich für die Beschreibung von Einwanderung in ein Land keine „Einwanderungsraten“ ermitteln, da für diese keine Nenner, also keine konkreten Risikomengen, bestimmt werden können (vgl. Bijak & Wisniowski 2010: 777). Zum anderen lassen sich der Umfang und die Struktur zukünftiger Migration kaum vorausberechnen, da diese von einer Vielzahl von Faktoren abhängen, wie der politischen, ökonomischen und sozialen Entwicklung im In- und Ausland (vgl. Bretz 2000: 675). de Beer (1988) untersuchte beispielsweise die Genauigkeit der Vorausberechnung des Umfangs von Immigration über Zeitreihenverfahren für die Niederlande für Vorausberechnungszeiträume von 1 bis 3 Jahren. Selbst bei diesen kurzen Vorausberechnungszeiträumen erwies sich Einwanderung als kaum vorausschätzbar (vgl. de Beer 1988: 287f.). Bijak & Wisniowski (2010) kamen bei der Analyse von Vorausberechnungen, die Zeitreihenverfahren mit Expertenwissen kombinieren, um Einwanderung vorauszuberechnen, zu dem ernüchternden Schluss, dass „[...] the process of inflows under study appeared to be barely predictable [...]“ (Bijak & Wisniowski 2010: 793).

Deshalb sind bei der Modellierung und Vorausberechnung von Migration allgemein Verfahren vorzuziehen, die zu relativ breiten Vorausberechnungsintervallen führen, da dies die große Unsicherheit bezüglich zukünftiger Migration widerspiegelt (vgl. Bijak & Wisniowski 2010: 786). Darüber hinaus sollten Ein- und Auswanderung getrennt und nicht als Nettomigration

zusammengefasst modelliert werden, da es sich um zwei substantiell unterschiedliche Prozesse handelt (vgl. ebd.: 776). Zudem kann sich ein bestimmter Wert der Nettomigration aus unterschiedlichen Werten für Ein- und Auswanderung ergeben, die unterschiedliche Effekte auf die Altersstruktur einer Bevölkerung haben können (vgl. Babel 2007: 64). Bezüglich der Struktur der Ein- und Auswanderung in die BRD kann wiederum von keinem klaren Trend ausgegangen werden (vgl. Härdle & Mysickova 2009: 18).

### 8.5.2 Anwendung des Ansatzes auf Migration

Als Grundlage der Vorausberechnung dienen die in Anhang A.4 beschriebenen Daten zur Ein- und Auswanderung aus dem Zeitraum von 1950 bis 2008, die sich vor dem Jahr 1989 lediglich auf Westdeutschland beziehen. Bezogen auf Einwanderung erscheint dies eher unproblematisch, da auch vor der Wiedervereinigung vergleichsweise hohe Werte erreicht wurden, die teils deutlich über manchen Werten aus dem Zeitraum von 1990 bis 2008 lagen. Bei der Auswanderung hingegen ist die Verwendung der Daten ab 1950 wegen der Veränderung des Gebietsstandes eher ungünstig, zumal dies zu einer Veränderung der Definition der Risikomenge führt. Trotz dieses Problems wurden die vollständigen Zeitreihen ab 1950 als Datengrundlage verwendet.<sup>195</sup>

Da keine Raten berechnet werden können, wird im weiteren wie bei anderen Arbeiten auch lediglich der Umfang der Ein- und Auswanderung berücksichtigt (vgl. bspw. Härdle & Mysickova 2009; Lipps & Betz 2005). Die Altersstruktur und Aufgliederung der Wanderungen nach Geschlecht wird aus dem Jahr 2008 übernommen.

Zunächst wurde versucht, die Vorausberechnung relativ analog zu Mortalität und Fertilität ausgehend von Differenzen der logarithmierten Zeitreihen vorzunehmen. Die Zeitreihen der Ein- und Auswanderungen wurden zunächst entsprechend eines von Thompson (1989) vorgeschlagenen Verfahrens transformiert, das es erlaubt, Ober- und Untergrenzen für die zulässigen Werte zu setzen. Als Untergrenze wurde in beiden Fällen 0 gewählt, als Obergrenze wurde beim Umfang der Einwanderung der Wert 1600000 gesetzt und bei der Auswanderung 1000000. Diese Werte liegen um etwa 100000 beziehungsweise 200000 Personen höher als die höchsten historischen Werte.

Trotz dieser Einschränkung konnten mit diversen Werten für  $k$  und  $a$  keine plausiblen Ergebnisse erzielt werden. Als Alternative wurde deshalb direkt aus den historischen Werten des Umfangs der Ein- und der Auswanderung gezogen, ähnlich wie bei Keyfitz (1989: 5). Durchgeführt wurden 1000 Simulationen, bei denen  $k = 1$  und  $a = 0$  gesetzt und die Werte für Ein- und Auswanderung unabhängig voneinander gezogen wurden. Dabei lässt sich festhalten, dass mit diesem Vorgehen Ergebnisse erzeugt werden können, deren Vorausberechnungsintervalle zumindest grob die historische Entwicklung abdecken. Ein Problem ist allerdings, dass durch das gewählte Vorgehen keine Autokorrelation berücksichtigt wird, was sich beispielsweise in

195 Entsprechend der Ausführungen in Unterabschnitt 4.1.4 könnte man zudem fragen, inwieweit die Migration nach Zusammenbruch des Ostblocks zu Beginn der neunziger Jahre nicht auf „besondere Einflüsse“ zurückgeführt werden kann und somit bei der Modellierung besonders beachtet bzw. ausgeschlossen werden sollte. Allerdings lassen sich für alle in Unterabschnitt 3.2.4 angeführten Wanderungswellen „besondere Einflüsse“ geltend machen. Beispielsweise kann die erste Wanderungswelle auf das „Wirtschaftswunder“ als besonderen Einfluss zurückgeführt werden. Dementsprechend wird wie in Unterabschnitt 4.1.4 bereits erwähnt auf eine gesonderte Modellierung verzichtet.

mehr oder weniger konstanten Vorausberechnungsintervallen zeigt (s. den folgenden Unterabschnitt). Von diesem Aspekt abgesehen sind die Ergebnisse aber relativ plausibel und spiegeln einerseits wieder, dass Deutschland in der Entwicklung seit 1950 ein Einwanderungsland war, und sie zeigen andererseits, dass die Unsicherheit bezüglich der künftigen Entwicklung der Migration relativ hoch ist. Verglichen mit anderen Vorausberechnungen von Migration (s. 8.5.4) zeigen sich zudem keine systematischen Unterschiede, die auf eine grundsätzliche Über- oder Unterschätzung hinweisen.

Problematisch dürfte ferner noch sein, dass von einer konstanten Alters- und Geschlechterstruktur der Ein- und Auswanderung ausgegangen wird. In Anbetracht der wie gerade erwähnt ansonsten aber zufriedenstellenden Ergebnisse und der Tatsache, dass die Modellierung von Migration allgemein als relativ schwierig gilt (s. die Vorbemerkungen zu diesem Abschnitt) wurde auf die Berechnung weiterer Varianten verzichtet.

Schließlich stellt sich die Frage, wie im Kontext von Migration mit der Differenzierung zwischen West- und Ostdeutschland bei der Vorausberechnung umgegangen werden kann. Zum einen muss Außenwanderung auf die beiden Gebiete aufgeteilt werden, zum anderen ist auch Binnenwanderung zwischen diesen Gebieten möglich. Bezüglich der Aufteilung der Außenwanderung wurde hier das Vorgehen von Lipps & Betz (2005) variiert. Die Autoren wiesen 20% der Nettomigration dem deutschen Osten zu und 80% Westdeutschland, was grob den Anteilen der Bevölkerung dieser Gebiete an der Gesamtbevölkerung der BRD seit der Wiedervereinigung entspricht (vgl. Lipps & Betz 2005: 10). Hier wurde im weiteren eine ähnliche Zuordnung vorgenommen, wobei für das jeweils zuletzt vorausberechnete Jahr der Anteil der west- und ostdeutschen Bevölkerung an der Gesamtbevölkerung berechnet wurde und die Ein- und Auswanderung entsprechend dieser Anteile auf die beiden Gebiete aufgeteilt wurde.

Binnenwanderung wurde nicht berücksichtigt, was ebenfalls dem Vorgehen von Lipps & Betz (vgl. 2005: 39) entspricht. Hierfür spricht zum einen, dass Daten für Binnenmigration zwischen Ost- und Westdeutschland entsprechend der deutschen Geschichte erst ab 1990 vorliegen, so dass die Zeitreihe relativ kurz ist. Zudem liegen die Daten nur in einer relativ groben Altersklassierung mit sechs Altersbereichen vor. Schließlich berücksichtigt auch das Statistische Bundesamt in der aktuellen Vorausberechnung für die einzelnen Bundesländer Binnenwanderung lediglich bis zum Jahr 2030, so dass bei den offiziellen Vorausberechnungen ab 2031 keine Binnenwanderung mehr auftritt (vgl. Statistisches Bundesamt 2010a).

### 8.5.3 *Ergebnisse*

Ergebnisse zu den Vorausberechnungen von Ein- und Auswanderungen sowie der sich aus diesen ergebenden Nettomigration und kumulierten Nettomigration sind in Tabelle 8.3 zu sehen. In allen Fällen sind der Median und das Vorausberechnungsintervall für ausgewählte Jahre abgetragen. Bei der Einwanderung, der Auswanderung und der Nettowanderung sind diese zumindest bezogen auf die Größenordnung der Werte relativ konstant, wie durch das gewählte Vorgehen auch zu erwarten ist. Bei der kumulierten Wanderung ergibt sich für das Jahr 2059 ein Median von 9.98 Millionen. Das Vorausberechnungsintervall reicht von 6.11 bis 13.47 Millionen, was einer Breite des Intervalls von über 7 Millionen entspricht.

Table 8.3: Ausgewählte Ergebnisse für die stochastische Vorausberechnung der Migration

Größe	Kennwert	2008	2020	2030	2040	2050	2059
Einwanderung (in Tsd.)	5%-Quantil	—	395	398	395	395	355
	Median/Ausgangswert	682	769	769	753	769	769
	95%-Quantil	—	1256	1256	1256	1277	1277
Auswanderung (in Tsd.)	5%-Quantil	—	395	395	399	395	399
	Median/Ausgangswert	738	623	606	606	606	606
	95%-Quantil	—	768	768	768	755	768
Nettowanderung (in Tsd.)	5%-Quantil	—	-288	-269	-274	-281	-294
	Median/Ausgangswert	-55	175	169	168	173	166
	95%-Quantil	—	702	741	717	712	764
Kumulierte Wanderung (in Mio.)	5%-Quantil	—	0.51	1.97	3.33	4.69	6.11
	Median	—	2.36	4.35	6.16	8.13	9.98
	95%-Quantil	—	4.17	6.70	9.03	11.40	13.47

Quelle: eigene Berechnung

#### 8.5.4 Vergleich mit Ergebnissen aus der Literatur

Bei der parametrischen Vorausberechnung der Migration von Betz & Lipps (2004: 12) wird bis zum Jahr 2050 durchschnittlich eine kumulierte Migration von 14.96 Millionen Personen erreicht, wobei dieser Wert relativ deutlich über der in Tabelle 8.3 ausgewiesenen Obergrenze des Vorausberechnungsintervalls liegt. Allerdings lagen den Autoren nur Daten bis zum Jahr 2002 vor, so dass die niedrige beziehungsweise negative Nettomigration der Jahre nach 2002 nicht berücksichtigt werden konnte, was eine mögliche Erklärung für diesen deutlichen Unterschied ist.

Bomsdorf et al. (2008: 28) weisen für ihre parametrische Vorausberechnung für das Jahr 2050 den Median, das 5%- und das 95%-Quantil der Nettowanderung aus. Die Werte betragen der Reihe nach 152000 Personen, –108000 Personen und 422000 Personen, womit eine Intervallbreite von 530000 Personen erreicht wird. Die Höhe des Median entspricht in etwa dem bei der nichtparametrischen Vorausberechnung erreichten Wert, während die Intervallbreite deutlich schmaler ausfällt, zumal bei der nichtparametrischen Vorausberechnung eine Breite von etwa 993000 Personen erreicht wird.

Bei der zwölften koordinierten Bevölkerungsvorausberechnung gibt es zwei Szenarien zur Migration. Beim ersten wird von einem Anstieg der jährlichen Nettomigration bis 2020 auf 100000 Personen und anschließend einem gleichbleibenden Wert ausgegangen (vgl. Statistisches Bundesamt 2010c: 34). Das zweite Szenario nimmt einen Anstieg der jährlichen Nettomigration bis 2020 auf 200000 Personen an (vgl. Statistisches Bundesamt 2010c: 34). Im ersten Fall ergibt sich bis zum Jahr 2060 eine kumulierte Wanderung von 4.86 Millionen Personen, beim zweiten Fall wird ein Wert von 9.36 Millionen Personen erreicht. Hiermit decken die Annahmen des Statistischen Bundesamtes eher den unteren Bereich der nichtparametrischen Vorausberechnung ab und liegen zudem nochmals wesentlich deutlicher unter der Vorausberechnung von Betz & Lipps (2004).

Diese Vergleiche zeigen, dass Ergebnisse von Vorausberechnungen der Migration sehr unterschiedlich ausfallen können. Insgesamt scheinen die Resultate des *resampling*-Ansatzes aber nicht unrealistisch zu sein.

## 8.6 Fortschreibung der Bevölkerung

### 8.6.1 Vorgehen

Die stochastischen Vorausberechnungen der demographischen Komponenten können zu einer stochastischen Vorausberechnung der Bevölkerung bis zum Jahr 2060 verwendet werden. Die jeweils 1000 vorliegenden Simulationsinstanzen pro Komponente werden dabei der Reihe nach kombiniert – also beispielsweise die erste Simulationsinstanz der Fertilität mit den ersten Simulationsinstanzen der Mortalität und Migration. Für die Fortschreibung der Bevölkerung kann man sich an den Ausführungen in den Unterabschnitten 5.2.2, 5.2.4 und 5.3.1 orientieren.

Die Bevölkerung in Altersklasse  $x + 1$  zu  $t + 1$  ergibt sich als

$$n_{x+1,t+1} = n_{x,t}P_{x,t} + m_{x,t} \quad (8.5)$$

wobei  $m_{x,t}$  wie in Abschnitt 5.2.2 beschrieben Nettomigration erfasst.<sup>196</sup>

Bisher wurde für Nettomigration nicht festgelegt, wie diese genau berechnet werden kann. Für die Vorausberechnung der Bevölkerung wurde wie folgt vorgegangen. Sei  $M_{x,t}^i$  die Zahl der Personen, die zwischen  $t$  und  $t + 1$  einwandern und sich zum Zeitpunkt der Einwanderung in Altersklasse  $x$  befinden.  $M_{x,t}^o$  sei die Zahl der zwischen  $t$  und  $t + 1$  auswandernden Personen, die sich zum Zeitpunkt der Auswanderung in Altersklasse  $x$  befinden. Dann ergibt sich  $n_{x+1,t+1}$  als

$$\begin{aligned} n_{x+1,t+1} = & n_{x,t}P_{x,t} + \frac{1}{2}(1 + P_x)\frac{1}{2}(M_{x,t}^i + M_{x+1,t}^i) \\ & - \frac{1}{2}(M_{x,t}^o + M_{x+1,t}^o) \end{aligned} \quad (8.6)$$

Sowohl bei der Ein- als auch bei der Auswanderung werden die Altersklassen  $x$  und  $x + 1$  berücksichtigt. Bei der Auswanderung liegt dies darin begründet, dass beispielsweise eine Person, die zu  $t$  in Altersklasse  $x$  ist, zwischen  $t$  und  $t + 1$  in Altersklasse  $x + 1$  wechseln kann. Wandert diese Person aus, so wird sie von  $M_{x+1,t}^o$  erfasst. Ebenso könnte eine Person, die ebenfalls zu  $t$  in Altersklasse  $x$  ist, zwar zwischen  $t$  und  $t + 1$  in die Altersklasse  $x + 1$  wechseln, allerdings vor diesem Wechsel auswandern, so dass sie von  $M_{x,t}^o$  erfasst wird. Dabei erfasst  $M_{x,t}^o$  aber auch Personen, die zu  $t$  in Altersklasse  $x - 1$  sind und zwischen  $t$  und  $t + 1$  in die Altersklasse  $x$  wechseln, so dass sie zu  $t + 1$  in Altersklasse  $x$  wären. Ebenso werden Personen erfasst, die zum Zeitpunkt der Auswanderung in Altersklasse  $x$  sind und die Altersklasse bis  $t + 1$  nicht mehr wechseln.

Analog verhält es sich mit  $M_{x+1,t}^o$  und Personen, die beispielsweise von Altersklasse  $x + 1$  in die Altersklasse  $x + 2$  wechseln. In beiden Fällen wird davon ausgegangen, dass die beiden jeweils erfassten Altersgruppen gleich stark in  $M_{x,t}^o$  beziehungsweise  $M_{x+1,t}^o$  vertreten sind, unter der Annahme, dass Wanderungen und Wechsel der Altersklassen gleichmäßig über den Zeitraum von  $t$  bis  $t + 1$  verteilt sind.

Bei Einwanderung lässt sich für die Berücksichtigung von  $M_{x,t}^i$  und  $M_{x+1,t}^i$  relativ analog zur Auswanderung argumentieren. Beispielsweise erfasst  $M_{x,t}^i$  sowohl Personen, die in Altersklasse  $x$  einwandern und bis zu  $t + 1$  in Altersklasse  $x + 1$  wechseln, als auch solche, die sich zu  $t + 1$  noch in Altersklasse  $x$  befinden.

Zusätzlich wird bei der Einwanderung noch die halbe Sterbewahrscheinlichkeit für die Altersklasse  $x$  berücksichtigt, da Personen, die zwischen  $t$  und  $t + 1$  einwandern, noch bis zu  $t + 1$  sterben können. Nimmt man entweder wie oben an, dass Wanderungen gleichmäßig über das Intervall von  $t$  bis  $t + 1$  verteilt sind, oder aber, dass Wanderungen nur exakt zu  $t$  und  $t + 1$  stattfinden, sind Einwanderer durchschnittlich dem halben Sterberisiko ausgesetzt, was die Verwendung der halben Sterbewahrscheinlichkeit begründet (für die Verwendung der ersten Argumentation s. Bretz 2000: 649; für die zweite vgl. Preston et al. 2001: 124f.). Verglichen

<sup>196</sup> Eigentlich müsste eine weitergehende Differenzierung der Bevölkerung nach Geschlecht und Region erfolgen, also  $n_{(x,g,r),t}$  anstelle von  $n_{x,t}$  betrachtet werden. Da das Vorgehen aber für alle Gruppen identisch ist, wird die durch eine weitergehende Differenzierung notwendig werdende erweiterte Indizierung nicht verwendet.

mit nicht migrierenden Personen führt dies zu einer höheren Überlebenswahrscheinlichkeit für diese Personengruppe, denn wenn  $Q_{x,t} = 1 - P_{x,t}$ , dann ist  $1 - 1/2Q_{x,t} = 1/2(1 + P_{x,t}) \geq P_{x,t}$ .

Der Bestand an Personen mit Geschlecht  $g$  in Altersklasse 0 zu  $t + 1$ , wird berechnet über

$$n_{(0,g),t+1} = \left( \sum_{\alpha}^{\beta} F_{x,t}^g n_{(x,w),t} + \frac{1}{2} F_{x,t}^g M_{(x,w),t}^i \right) + \frac{1}{2} (1 + P_{(0,g),t}) \frac{1}{2} M_{(0,g),t}^i - \frac{1}{2} M_{(0,g),t}^o \quad (8.7)$$

Hierbei wird zunächst die Fertilität von Frauen über alle Altersklassen von  $\alpha$  bis  $\beta$  berücksichtigt, wobei  $\alpha$  die erste und  $\beta$  die letzte Altersklasse ist, für die  $F_{x,t}^g$  größer 0 ist. Für einwandernde Frauen,  $M_{(x,w),t}^i$ , wird die halbe Fertilität veranschlagt. Von Personen, die in Altersklasse 0 sind und ein- beziehungsweise auswandern, wird analog zur oben genannten Argumentation die Hälfte berücksichtigt, wobei bei einwandernden Personen zusätzlich noch die halbe Sterbewahrscheinlichkeit berücksichtigt wird.

Bei dieser Fortschreibung der Bevölkerungsbestände wird wie bisher von einer Unabhängigkeit der Komponenten ausgegangen. Dies zeigt sich beispielsweise darin, dass für einwandernde Personen direkt die Geburtenraten der einheimischen Bevölkerung verwendet werden, was impliziert, dass keine Heterogenität vorliegt.

## 8.6.2 Ergebnisse

Eine Übersicht über zentrale Ergebnisse der Bevölkerungsvorausberechnung findet sich in Tabelle 8.4. Unter anderem ist der Median und das 90%-Vorausrechnungsintervall für den Bevölkerungsbestand eingetragen, dessen Entwicklung ebenfalls in Abbildung 8.2 dargestellt ist. Betrachtet man den Verlauf des Medians, zeigt sich zunächst bis zum Jahr 2020 ein leichter Anstieg der Bevölkerung. Bis 2030 findet eine leichte Abnahme unter den Ausgangswert der Vorausberechnung statt. Anschließend beschleunigt sich diese Abnahme, wobei im Jahr 2060 schließlich ein Bevölkerungsbestand von 72.73 Millionen erreicht wird, was einer Abnahme um 8.91 Millionen Personen beziehungsweise etwa 11% entspricht.

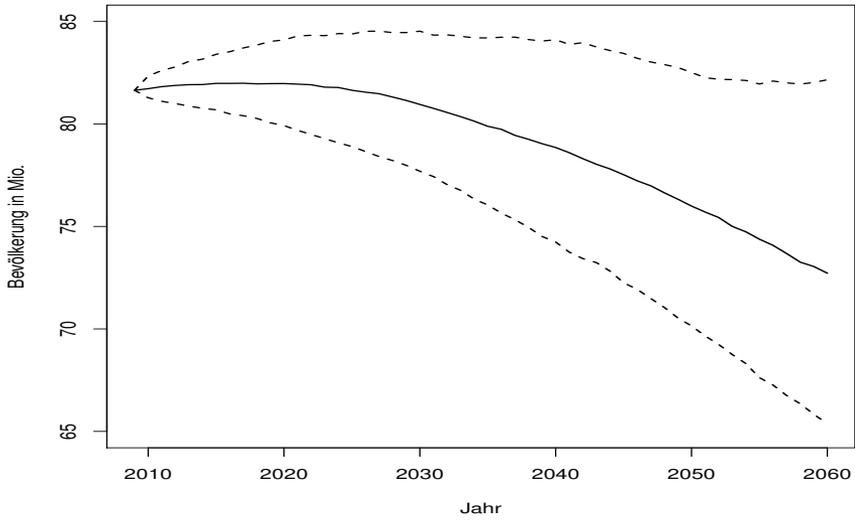
Betrachtet man den Verlauf des Vorausrechnungsintervalls, lässt sich festhalten, dass bis 2020 keine übermäßig starke Abnahme der Bevölkerung zu erwarten ist. Bis 2060 allerdings erreicht die Untergrenze des Intervalls einen Wert von 65.41 Millionen Personen, was zu einer Abnahme um etwa 16.23 Millionen Personen beziehungsweise circa 20% korrespondiert. Die obere Grenze des Vorausrechnungsintervalls liegt bei 82.15 Millionen Personen, was um etwa eine halbe Millionen Personen über dem Ausgangswert liegt. Somit wird bei 90% der Simulationen für das Jahr 2060 ein Bevölkerungsbestand erreicht, der in etwa grob dem Ausgangswert entspricht oder unter diesem liegt. Ein langfristiges Bevölkerungswachstum scheint ausgehend von diesen Ergebnissen sehr unwahrscheinlich zu sein. Eine Ursache dürfte sein, dass für ein nennenswertes Wachstum der Bevölkerung eine langfristig hohe Fertilität, hohe Migration und gegebenenfalls niedrige Sterblichkeit notwendig wäre, was unter anderem aufgrund der Unabhängigkeit der Komponenten nur eingeschränkt wahrscheinlich ist. Aufgrund des Vorgehens bei der Vorausberechnung könnte sich zudem ein relativ deutlicher Anstieg

Table 8.4: Ausgewählte Ergebnisse für die stochastische Bevölkerungsvorausberechnung

Größe	Kennwert	2009	2020	2030	2040	2050	2060
Bevölkerungsbestand (in Mio.)	5%-Quantil	—	79.93	77.70	74.24	70.16	65.41
	Median/Ausgangswert	81.64	81.97	80.95	78.86	76.00	72.73
	95%-Quantil	—	84.11	84.53	84.12	82.55	82.15
Bevölkerungsbestand (West)	5%-Quantil	—	63.81	62.17	59.62	56.44	52.73
	Median/Ausgangswert	65.47	65.46	64.79	63.31	61.14	58.67
	95%-Quantil	—	67.15	67.64	67.55	66.40	66.23
Bevölkerungsbestand (Ost)	5%-Quantil	—	16.10	15.52	14.63	13.72	12.65
	Median/Ausgangswert	16.17	16.52	16.17	15.53	14.87	14.09
	95%-Quantil	—	16.95	16.87	16.55	16.13	15.91
Altenquotient	5%-Quantil	—	35.36	45.93	53.99	54.67	56.15
	Median/Ausgangswert	32.06	36.66	48.75	58.37	60.40	63.67
	95%-Quantil	—	37.87	51.02	62.01	65.97	71.31
Durchschnittsalter (in Jahren)	5%-Quantil	—	44.22	45.40	46.40	46.54	45.70
	Median/Ausgangswert	42.27	44.72	46.48	48.08	49.07	49.24
	95%-Quantil	—	45.19	47.44	49.48	51.01	51.65

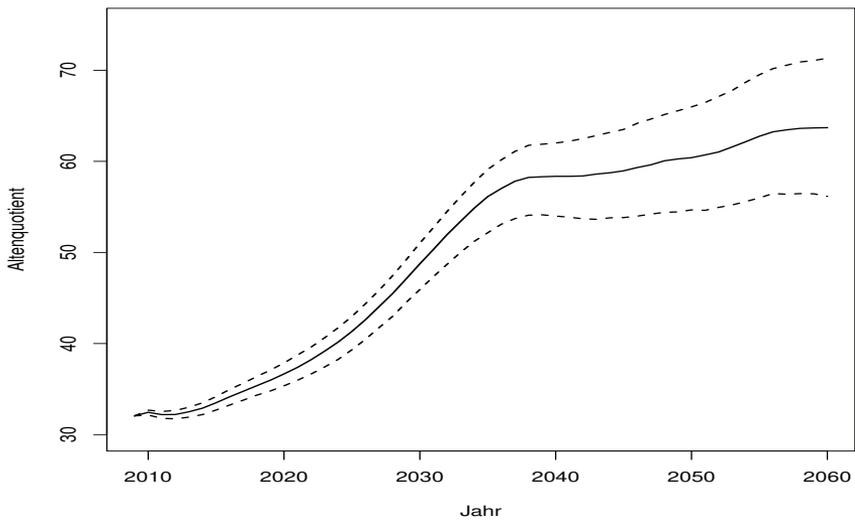
Quelle: eigene Berechnung

Abbildung 8.2: Median (durchgezogene Linie) und 90%-Vorausberechnungsintervall (gestrichelte Linien) des Bevölkerungsbestandes bis 2060



Quelle: eigene Berechnung

Abbildung 8.3: Median (durchgezogene Linie) und 90%-Vorausberechnungsintervall (gestrichelte Linien) des Altenquotienten bis 2060



Quelle: eigene Berechnung

der Fertilität in der Regel erst mittelfristig ergeben. Da aber die Geburtsjahrgänge, die sich mittelfristig im gebärfähigen Alter befinden, vergleichsweise klein sind, wie beispielsweise der Geburtsjahrgang von 2008 mit 682514 Personen, würde auch ein deutlicher Anstieg der Fertilität nicht zu einem direkten, deutlichen Effekt auf das Bevölkerungswachstum führen, sondern seine volle Wirkung erst zeitverzögert zeigen.

Die Breite des Vorausberechnungsintervalls beträgt für das Jahr 2060 16.74 Millionen Personen. In diesem Wert spiegelt sich die relativ große Unsicherheit bezüglich der langfristigen Bevölkerungsentwicklung wieder, die die Unsicherheit der Entwicklung aller einzelnen Komponenten enthält.

Die Entwicklung des Altenquotienten ist ebenfalls in Tabelle 8.4 zu sehen und zusätzlich in Abbildung 8.3 dargestellt. Der Altenquotient wurde berechnet als das Verhältnis der Bevölkerung in den Altersklassen von 20 bis einschließlich 64 Jahren zur Bevölkerung in den Altersklassen ab 65 Jahren, multipliziert mit 100.<sup>197</sup>

Betrachtet man zunächst den Median, zeigt sich, dass dieser bis zum Ende der 2030er Jahre deutlich ansteigt und anschließend stagniert beziehungsweise nur noch stark abgeschwächt zunimmt. Zumindest bis zum Jahr 2030 liegen die Grenzen des Vorausberechnungsintervalls dabei relativ nah am Median. Dieser Anstieg des Altenquotienten wird durch die geburtenstarken Jahrgänge der 1950er und 1960er verursacht, die nach und nach die Altersgrenze von 65 Jahren überschreiten, und gewissermaßen vom Nenner des Altenquotienten in den Zähler wechseln. Da diese Personen im Ausgangsjahr der Vorausberechnung bereits in der Bevölkerung sind, sind die einzigen Quellen der Unsicherheit Wanderungen und Mortalität. Dabei spielen Wanderungen aber mit fortschreitendem Alter eine immer kleinere Rolle für die Entwicklung des Umfangs der Altersgruppe und Sterblichkeit ist zumindest bis zur Altersgrenze von 65 Jahren sowohl für Männer als auch für Frauen relativ gering, so dass es bei den Übergängen aus der Nennerpopulation in die Zählerpopulation nur ein relativ geringes Maß an Unsicherheit gibt. Zugleich sind außerdem die in den Nenner nachrückenden Jahrgänge vergleichsweise klein, was zu einer Beschleunigung des Anstiegs führt. Dabei sind diese nachrückenden Jahrgänge zumindest mittelfristig bereits in der Ausgangsbevölkerung enthalten, so dass sich bei diesen ebenfalls nur eine eingeschränkte Unsicherheit ergibt, die vor allem von der Unsicherheit in der Migration geprägt wird.

Nach den 2030er Jahren steigt dann der Median des Altenquotienten wesentlich langsamer an und das Vorausberechnungsintervall verbreitert sich deutlich. Für das Jahr 2060 beträgt der Median 63.67 und das Vorausberechnungsintervall erreicht mit einer Untergrenze von 56.15 und einer oberen Grenze von 71.31 eine Breite von ungefähr 15.16. Dieses fällt somit relativ breit aus, weist allerdings darauf hin, dass bei einem Großteil der Simulationen ein relativ deutlicher Anstieg des Altenquotienten stattfindet.

Ähnlich verhält es sich bei der Entwicklung des durchschnittlichen Alters der Bevölkerung, welches als ergänzender Kennwert für die Entwicklung der Altersstruktur herangezogen wird. Ergebnisse sind abermals in Tabelle 8.4 erfasst. Der Median steigt bis zum Jahr 2050 kontinuierlich an und liegt danach relativ konstant bei dem auch im letzten Vorausberechnungsjahr 2060 erreichten Wert von 49.24 Jahren, was einem Anstieg des durchschnittlichen Alters bis

---

197 Der Altenquotient soll hier nur als ein grober Indikator für die Entwicklung der Altersstruktur der Bevölkerung verwendet werden. Kritische Anmerkungen zum Altenquotienten finden sich in den in Fußnote 108 auf Seite 89 angegebenen Quellen.

Tabelle 8.5: Ergebnisse zur Bevölkerungsgröße und zum Altenquotienten im Jahr 2050 aus der Literatur

Quelle	Kennwert	5%	50%	95%	Breite
Betz & Lipps (2004)	Bevölkerung (in Mio.)	68.3	75.8	83.3	15.0
Babel (2007)	Bevölkerung (in Mio.)	62.5	71.5	80.7	18.2
Babel (2007)	Altenquotient	51.1	61.4	73.0	21.9
Härdle & Mysickova (2009)	Bevölkerung (in Mio.)	73.5	77.4	81.1	7.1

Quelle: Betz & Lipps (2004: 13); Babel (2007: 183); Härdle & Mysickova (2009: 23); eigene Berechnung

2060 von etwas weniger als 7 Jahren entspricht. Das Vorausberechnungsintervall für das Jahr 2060 reicht von 45.70 bis zu 51.65 Jahren, womit in der großen Mehrheit der Simulationen ein langfristiger Anstieg des durchschnittlichen Alters zu erwarten ist. Ein interessantes, nicht in Tabelle 8.4 enthaltenes Detail ist, dass die Untergrenze des Intervalls bis etwa zur Mitte der 2040er Jahre ansteigt und anschließend wieder leicht abnimmt. Dies deutet auf Simulationen hin, bei denen die Fertilität zwar langfristig steigt, allerdings aus den oben bereits genannten Gründen erst zeitverzögert einen verjüngenden Effekt auf die Bevölkerungsentwicklung hat.

Als Ergänzung zu den bisher vorgestellten Ergebnissen findet sich in Tabelle 8.4 noch die Bevölkerungsentwicklung getrennt für West- und Ostdeutschland. Vergleicht man die Ergebnisse, ist insbesondere auffällig, dass der Median der westdeutschen Bevölkerungsentwicklung bis 2020 nur geringfügig abnimmt, während der Median für Ostdeutschland leicht steigt. Eine mögliche Ursache könnte die Aufteilung der Nettomigration auf die beiden Teile Deutschlands sein, die dem Osten eventuell Anteilswerte für Ein- und Auswanderung zuweist, die bezogen auf tatsächliche historische Werte unrealistisch sind. Ansonsten weichen die Ergebnisse allerdings nicht vom bei der Entwicklung der Gesamtbevölkerung gewonnenen Eindruck ab, so dass sie hier nicht weitergehend kommentiert werden.

### 8.6.3 Vergleich mit Ergebnissen aus der Literatur

In Tabelle 8.5 sind Ergebnisse zum Bevölkerungsbestand und zum Altenquotienten im Jahr 2050 von parametrisch-stochastischen Vorausberechnungen aus der Literatur zu finden. Die Vorausberechnung von Betz & Lipps (2004) beginnt mit dem Jahr 2002, die Vorausberechnung von Babel (2007) weist 2005 als Startjahr auf und die Vorausberechnung von Härdle & Mysickova (2009) beginnt ab dem Jahr 2007.

Der sich bei der nichtparametrischen Vorausberechnung ergebende Median des Bevölkerungsbestandes von 76.00 Millionen Personen entspricht von der Größenordnung her relativ gut den Ergebnissen von Betz & Lipps (2004) und Härdle & Mysickova (2009). Gleiches gilt für die obere Grenze des Vorausberechnungsintervalls von 82.55 Millionen. Bei der Untergrenze fallen die Unterschiede etwas deutlicher aus, wobei der Wert von 70.16 Millionen, der sich bei der nichtparametrischen Vorausberechnung ergibt, zwischen den Resultaten von Betz & Lipps (2004) und Härdle & Mysickova (2009) liegt. Entsprechendes lässt sich für die Breite des Vorausberechnungsintervalls von 12.39 Millionen feststellen. Die Ergebnisse von Babel (2007) weisen bezüglich der Bevölkerungsgröße das breiteste Vorausberechnungsintervall und die niedrigsten Werte für das 5%-Quantil und das 95%-Quantil auf. Verglichen mit den sich

bei der nichtparametrischen Vorausberechnung ergebenden Resultaten für den Altenquotienten fallen die Ergebnisse von Babel (2007) ebenfalls wieder relativ breit aus. Die Breite der Vorausberechnungsintervalle ist mit einem Wert von 11.3 beim nichtparametrischen Verfahren fast nur halb so hoch.

Bei den Szenarien der zwölften koordinierten Bevölkerungsvorausberechnung des Statistischen Bundesamtes ergeben sich für das Jahr 2060 Bevölkerungsbestände zwischen 61.77 Millionen und 74.52 Millionen Personen. Etwa 63% der Simulationen der stochastischen Vorausberechnungen liegen zwischen diesen beiden Werten. Der obere vom Statistischen Bundesamt ausgewiesene Wert liegt um etwas weniger als 2 Millionen Personen höher als der Median der Simulationen, der untere Wert liegt etwas weniger als 4 Millionen Personen unter der Untergrenze des Vorausberechnungsintervalls. Insgesamt fallen die Ergebnisse des Statistischen Bundesamtes also vergleichsweise niedrig aus. Gleiches gilt verglichen mit den Werten aus Tabelle 8.5 auch in eingeschränktem Maße für die Resultate für das Jahr 2050, wobei die Resultate des Statistischen Bundesamtes zwischen 67.41 und 78.67 Millionen Personen liegen.

Der Altenquotient liegt in den amtlichen Szenarien zwischen 59.8 und 77.1. 80% der Simulationen liegen in diesem Bereich. Dabei beträgt beispielsweise die Grenze des Vorausberechnungsintervalls 56.15 und die obere Grenze 71.31. Bei diesem Vergleich fallen die Ergebnisse des Statistischen Bundesamts vergleichsweise hoch aus.

Insgesamt weichen die Ergebnisse des nichtparametrischen Verfahrens nicht allzu stark von den Resultaten von parametrischen Verfahren ab. Die Ergebnisse des Statistischen Bundesamtes fallen bezogen auf den Bevölkerungsumfang eher niedrig und bezogen auf den Altenquotienten eher hoch aus und scheinen vergleichsweise „pessimistisch“ zu sein.<sup>198</sup>

## 8.7 Zusammenfassung

In diesem Kapitel wurde ein nichtparametrisches Verfahren für die Vorausberechnung demographischer Komponenten vorgestellt und angewendet. Bei der Anwendung auf Mortalität ergibt sich sowohl für Männer als auch für Frauen ein relativ „sicherer“ Anstieg der Lebenserwartung. Bei der Fertilität wird die durch Vorausberechnungsintervalle ausgedrückte Unsicherheit mit fortschreitender Länge der Vorausberechnung zunehmend größer, so dass langfristig viele unterschiedliche Entwicklungen möglich sind. Bei der Anwendung des Verfahrens auf Migration mussten etliche Einschränkungen gemacht werden, die dazu führen, dass die Unsicherheit bezüglich zukünftiger Migration extrem groß ist, was insbesondere bereits ab dem ersten Vorausberechnungsjahr gilt.

Bei allen drei Komponenten ist zu beachten, dass die Vorausberechnungen primär auf westdeutschen Daten basieren. Die Vorausberechnungen für Ostdeutschland setzten auf unter-

---

<sup>198</sup> Mit „pessimistisch“ ist gemeint, dass die Ergebnisse des Statistischen Bundesamtes am ehesten auf eine Entwicklung hindeuten, die als ungünstig für das System der sozialen Sicherheit aufgefasst werden kann. Eine solche Interpretation findet sich teils auch in den Publikationen des Statistischen Bundesamtes: „Somit werden im Jahr 2030 nicht einmal zwei Menschen im Erwerbsalter einem potenziellen Rentenbezieher gegenüberstehen, während im Jahr 2008 noch knapp drei Personen im Erwerbsalter für einen Rentner aufkommen konnten.“ (Statistisches Bundesamt 2011a: 26).

schiedliche Art und Weise an die Vorausberechnungen für Westdeutschland an. Hierbei wurden jeweils zusätzliche Annahmen getroffen, so dass die Vorausberechnungen für Ostdeutschland an mehr Bedingungen geknüpft sind als die Vorausberechnungen für Westdeutschland.

Die nichtparametrischen Vorausberechnungen der demographischen Komponenten wurden genutzt, um eine stochastische Vorausberechnung der Bevölkerung durchzuführen. Die Ergebnisse legen nahe, dass eine Abnahme der Bevölkerungsgröße sowie starke Zunahmen des durchschnittlichen Alters in der Bevölkerung und des Altenquotienten langfristig sehr wahrscheinlich sind, obwohl langfristig ein Anstieg der Fertilität nicht unwahrscheinlich ist und auch Einwanderung hoch ausfallen kann.

Vergleicht man die Ergebnisse des nichtparametrischen Ansatzes mit Resultaten von parametrischen Vorausberechnungen, zeigen sich insgesamt keine übermäßigen Unterschiede. Die Ergebnisse und Annahmen zum Bevölkerungsumfang, zur zusammengefassten Geburtenziffer und zur Höhe der Migration der zwölften koordinierten Bevölkerungsvorausberechnung fallen verglichen mit den Ergebnissen der stochastischen Verfahren eher niedrig aus, wobei beim Bevölkerungsumfang aber die Tendenz der Entwicklung die selbe ist. Die Annahmen zur Lebenserwartung und die Ergebnisse zum Altenquotienten des Statistischen Bundesamtes scheinen hingegen vergleichsweise hoch zu sein.

Im nächsten Kapitel werden die stochastischen Vorausberechnungen der demographischen Komponenten und der Bevölkerung in Kombination mit dem Modell aus 5 und den Ergebnissen aus Kapitel 6 genutzt, um eine stochastische Vorausberechnung linearer Verwandtschaft durchzuführen.

## 9 Ausgewählte Ergebnisse der Vorausberechnung linearer Verwandtschaft

### 9.1 Vorbemerkungen

In diesem Kapitel werden die Ergebnisse der Vorausberechnung linearer Verwandtschaft bis zum Jahr 2060 präsentiert. Dabei werden lediglich die Resultate der Simulationen beschrieben. Wenn also beispielsweise davon gesprochen wird, dass ein bestimmter Kennwert bis zu einem in der Zukunft liegenden Jahr  $x$  deutlich abnimmt, bedeutet dies lediglich, dass dies in den Simulationen der Fall ist. Dies gilt auch für die in Kapitel 10 noch vorzustellenden Resultate. Unter welchen Bedingungen weitergehende Schlussfolgerungen aus den Ergebnissen gezogen werden können, wird am Ende von Kapitel 10 diskutiert.

Bei allen Resultaten ist wie zuvor bei den Ergebnissen aus Kapitel 8 angemerkt grundsätzlich zu beachten, dass sich die dargestellten Mediane und Vorausberechnungsintervalle immer aus vielen Simulationsinstanzen zusammensetzen und die sich aus diesen Werten ergebenden zeitlichen Verläufe keinen bestimmten Verlauf einer Simulation oder einen wie auch immer definierten „typischen“ Verlauf darstellen. Werden diese Kennwerte nicht gegen den zeitlichen Verlauf, sondern Altersklassen abgetragen handelt es sich immer um Querschnittsdarstellungen für ein bestimmtes Jahr – so dass Personen, die in Altersklasse 40 sind, andere Personen sind, als die in Altersklasse 50, und sich unterschiedliche Werte für diese Altersklassen nicht als Entwicklungsverläufe für eine Personengruppe interpretieren lassen.

Ferner sei darauf hingewiesen, dass nur ein Teil der Ergebnisse vorgestellt wird, da viele Perspektiven auf die Resultate möglich sind. Beispielsweise können die Ergebnisse im Querschnitt über alle Alter betrachtet werden. Dies ist für die Jahre 2009 bis 2060, Männer und Frauen sowie West- und Ostdeutschland möglich. Daneben sind aber auch Längsschnittbetrachtungen denkbar, wie beispielsweise die Verwandtenzahl für eine bestimmte Altersklasse über die Jahre 2009 bis 2060. Hierbei können wieder beliebige Altersklassen, Frauen und Männer und West- und Ostdeutschland betrachtet werden.

Im weiteren Verlauf werden sowohl die Ergebnisse für Frauen und Männer als auch West- und Ostdeutschland vorgestellt, wobei die Resultate für West- und Ostdeutschland jeweils in getrennten Unterabschnitten diskutiert werden, da sie auf unterschiedlich starken Annahmen basieren (s. Kapitel 8). Ansonsten wird bei den betrachteten Verwandten jeweils nur eine Perspektive gewählt (Quer- oder Längsschnitt), die kurz begründet wird. Ergebnisse zum (durchschnittlichen) Alter von Vor- und Nachfahren werden nur am Rande behandelt. Umfangreiche Tabellen zu allen Ergebnissen finden sich in Anhang E.

Zunächst wird auf Ergebnisse zu Kindern und Enkelkindern eingegangen, anschließend werden Ergebnisse zu Eltern und Großeltern präsentiert. Ferner werden die Resultate einer Sensitivitätsanalyse vorgestellt, die durchgeführt wurde, um den Einfluss der einzelnen demographischen Komponenten auf die Ergebnisse zu ermitteln. Das Kapitel schließt mit einer Zusammenfassung.

## 9.2 Ergebnisse zur Zahl der Kinder

### 9.2.1 Ergebnisse für Westdeutschland

Bei der Darstellung der Ergebnisse zu Nachfahren wird im weiteren grundsätzlich eine Querschnittsperspektive gewählt. Es werden jeweils die Ergebnisse für das Jahr 2060 vorgestellt und mit den Ergebnissen der Modellrechnung für 2009 verglichen (s. Kapitel 7). Die Entwicklung von 2009 bis 2060 ist in den Ergebnissen für 2060 zumindest implizit enthalten. Beispielsweise waren Frauen, die 2060 in Altersklasse 80 sind, 2030 in Altersklasse 50, also am Ende der reproduktiven Phase. Da Mortalität für die Nachfahren dieser Frauen eher zu vernachlässigen sein dürfte, sollten die Ergebnisse für Altersklasse 80 im Jahr 2060 ungefähr den Ergebnissen für Altersklasse 50 im Jahr 2030 entsprechen. Hierdurch würde eine Betrachtung im Längsschnitt zusätzliche Redundanzen schaffen: Ergebnisse für eine Altersklasse  $x$  am Ende oder nach der reproduktiven Phase zu  $t$  entsprechen zumeist ungefähr den Ergebnissen der Altersklasse  $x + y$  zu  $t + y$ .

In Abbildung 9.1 sind der Median, das 5%- und das 95%-Quantil für die Verteilung der erwarteten Anzahl lebender Kinder westdeutscher Frauen nach Altersklassen im Jahr 2060 zu sehen. Zusätzlich sind die sich aus der Modellrechnung für das Jahr 2009 ergebenden Werte abgetragen.

Die Abbildung kann wie folgt gelesen werden. Für die Altersklasse von 50 Jahren ergibt sich für das Jahr 2060 ein Median über alle 1000 Simulationsinstanzen von 1.44 Kindern (durchgezogene Linie). Das Vorausberechnungsintervall, das 90% aller Simulationen umfasst, reicht von 1.18 bis 1.92 (gestrichelte Linien). Im Jahr 2009 lag die erwartete Kinderzahl westdeutscher Frauen in Altersklasse 50 bei ungefähr 1.58 Kindern (gepunktete Linie). Für die anderen Altersklassen und die im weiteren noch folgenden Abbildungen zu Nachfahren können Ergebnisse auf die gleiche Art und Weise abgelesen werden.

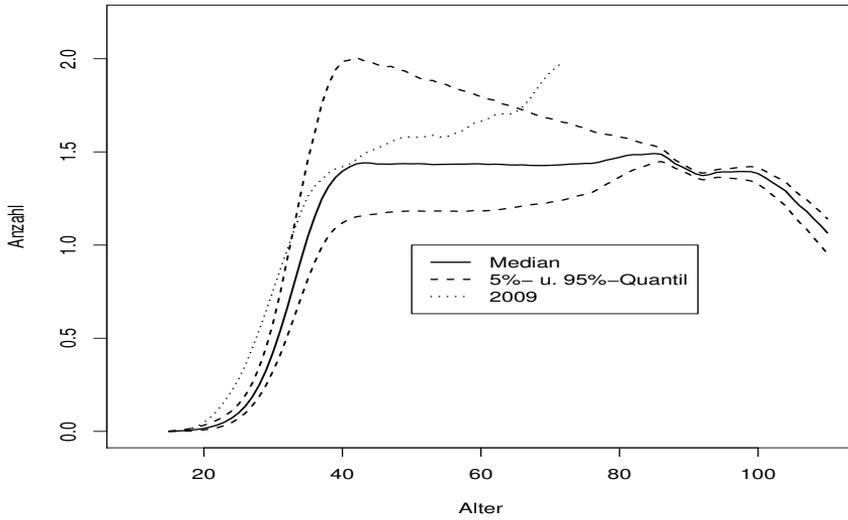
Betrachtet man in Abbildung 9.1 die Entwicklung des Medians beginnend ab den Altersklassen, die am Anfang der reproduktiven Phase liegen, zeigt sich, dass dieser zunächst eher langsam, dann ab der Altersklasse von etwa 25 Jahren beschleunigt ansteigt. Ab der Altersklasse von 40 Jahren verläuft der Median bis ungefähr zur Altersklasse von 80 Jahren relativ konstant mit kleineren Schwankungen, wobei Werte zwischen 1.396 und 1.471 erreicht werden. Nach einem leichten Anstieg fällt der Median hierauf folgend bis zur Altersklasse von 92 Jahren ab und verläuft anschließend bis zur Altersklasse von 100 Jahren relativ konstant.<sup>199</sup>

Das 5%- und das 95%-Quantil steigen zunächst relativ parallel zum Median und werden nur langsam breiter. Bis zur Altersklasse von 40 Jahren allerdings nimmt die Intervallbreite deutlich zu und liegt für diese Altersklasse bei 0.865. Ab der Altersklasse von 40 Jahren nimmt die Intervallbreite wieder langsam ab und erreicht für die Altersklasse von 91 Jahren einen Wert von 0.034. Anschließend nimmt die Intervallbreite wieder leicht zu.

---

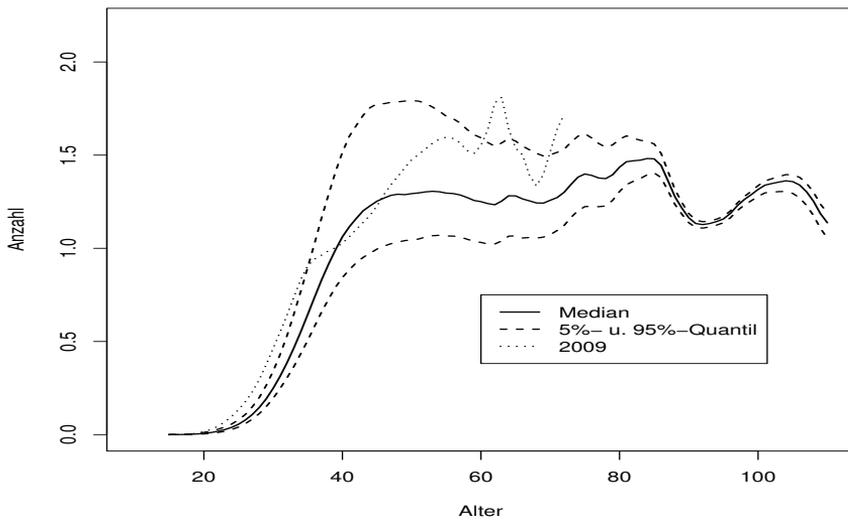
199 Für Altersklassen über 100 nimmt die Zahl lebender Kinder ab, da sich diese ebenfalls in vergleichsweise hohen Altersklassen mit hoher Sterblichkeit befinden. Allerdings ist der Bevölkerungsbestand in Altersklassen über 100 sehr gering. Aus diesem Grunde wird bei den weiteren Ausführungen nicht explizit auf diese eingegangen. Nichtsdestotrotz werden auch bei den weiteren Abbildungen Altersklassen über 100 der Vollständigkeit halber dargestellt.

Abbildung 9.1: Erwartete Anzahl lebender Kinder westdeutscher Frauen 2060 nach Altersklassen, sowie Ergebnisse für 2009



Quelle: eigene Berechnung

Abbildung 9.2: Erwartete Anzahl lebender Kinder westdeutscher Männer 2060 nach Altersklassen, sowie Ergebnisse für 2009



Quelle: eigene Berechnung

Ursache für die Zu- und Abnahme der Breite des Vorausberechnungsintervalls ist, dass sich zunächst Unsicherheit in den der Berechnung zugrunde liegenden Raten „akkumuliert“. Die hierauf folgende Abnahme der Intervallbreite erklärt sich dadurch, dass die Unsicherheit in den für die Berechnung verwendeten Raten abnimmt. Beispielsweise gehören Frauen, die 2060 in Altersklasse 80 sind, zum Geburtsjahrgang 1979. Für diesen ist aber zumindest ein Teil der altersspezifischen Geburtenraten historisch beobachtet. Die vorausberechneten Geburtenraten liegen größtenteils vergleichsweise nah am Beginn der Vorausberechnung und weisen somit eine relativ geringe Unsicherheit auf (s. Kapitel 8). Für Frauen, die 2060 in Altersklasse 40 sind, gehören alle Geburtenraten zur Vorausberechnung und liegen vergleichsweise weit in der Zukunft, so dass sie mit relativ großer Unsicherheit behaftet sind. Für Frauen in sehr hohen Altern ergibt sich die Unsicherheit in der durchschnittlichen Anzahl lebender Kinder nur oder fast nur aus der Unsicherheit der zukünftigen Entwicklung der Mortalität.

Beim Vergleich mit den Werten der Modellrechnung für das Jahr 2009 sei zunächst daran erinnert, dass sich Median und Quantile aus den Werten vieler einzelner Vorausberechnungsinstanzen ergeben, während die Werte für 2009 einer konkreten Realisierung entsprechen. Nimmt man einen Vergleich der Ergebnisse vor, fällt zunächst auf, dass bei allen Altersklassen die erwartete Anzahl lebender Kinder bei den vorausberechneten Ergebnissen des Jahres 2060 niedriger ausfällt, wobei der 2009 zu sehende Anstieg der Kinderzahl über die Altersklassen bis etwas über 30 Jahren außerhalb der Vorausberechnungsintervalle liegt. Dies ist auf eine Zunahme des durchschnittlichen Alters bei der Geburt zurückzuführen. Für die Altersklassen um 40 liegen die Werte des Jahres 2009 und der Median für das Jahr 2060 relativ eng zusammen. Für folgende Altersklassen steigen die Werte für 2009 allerdings und liegen etwa ab der Altersklasse von 65 Jahren wieder außerhalb des Vorausberechnungsintervalls, womit die vergleichsweise hohe Anzahl lebender Kinder von älteren Frauen im Jahr 2009 für das Jahr 2060 gegeben dieser Resultate eher unwahrscheinlich scheint.

Der entsprechend den Ergebnissen der Vorausberechnung der Fertilität zu erwartende Anstieg des durchschnittlichen Alters bei der Geburt wirkt sich auch beim durchschnittlichen Alter der Kinder aus, welches durch diesen Anstieg abnimmt. Beispielsweise ergibt sich in der Modellrechnung für 2009 bei Frauen in Altersklasse 50 ein durchschnittliches Alter der lebenden Kinder von 21.8. 2060 liegt der Median bei 18.0, das 5%-Quantil beträgt 17.2 und das 95%-Quantil 18.8.

Analog zu Abbildung 9.1 ist in Abbildung 9.2 die erwartete Anzahl lebender Kinder westdeutscher Männer in den Jahren 2009 und 2060 nach Altersklassen eingezeichnet. Die erwartete Kinderzahl fällt bei allen Altersklassen tendenziell niedriger aus als bei Frauen der selben oder der vorausgehenden Altersklassen. Zudem findet der Anstieg der durchschnittlichen Kinderzahl mit steigendem Alter verglichen mit Frauen zeitversetzt statt. Schließlich sind die Ergebnisse deutlich unregelmäßiger. Dies dürfte der Berechnungsmethode der Fertilität von Männern geschuldet sein, bei der das quantitative Verhältnis von Männern und Frauen einen maßgeblichen Einfluss hat, wodurch beispielsweise die relative Größe aufeinander folgender Geburtsjahrgänge einen deutlichen Effekt auf die Fertilität von Männern hat (s. Kapitel 6).

Für die Altersklasse von 30 Jahren ergibt sich ein Median von 0.248, für die Altersklasse von 40 Jahren ein Median von 1.062 und für die Altersklasse von 50 Jahren liegt dieser Kennwert bei 1.294. Die höchste auftretende Breite des Vorausberechnungsintervalls beträgt 0.781 und gilt für die Altersklasse von 44 Jahren. Vergleicht man diesen Wert mit der Intervallbreite der

selben und der vorausgehenden Altersklassen für Frauen, fällt die Unsicherheit bei Männern niedriger aus. Allerdings gilt dies nur bezogen auf die absolute Breite des Vorausberechnungsintervalls. Berechnet man die Varianz und den Mittelwert der sich aus den Vorausberechnungen ergebenden Verteilungen und berechnet hierüber wiederum den Variationskoeffizienten, ergibt sich beispielsweise für Männer in der Altersklasse von 40 Jahren für das Jahr 2060 ein Wert von 0.186. Für Frauen in der Altersklasse von 37 Jahren wird mit 0.183 ein sehr ähnlicher erzielt. Für Frauen in der Altersklasse von 40 Jahren liegt der Wert mit 0.180 nochmals etwas niedriger, weicht allerdings immer noch nicht allzu stark vom Wert der Männer in der selben Altersklasse ab. Insofern kann festgehalten werden, dass die sich bei Männern ergebende Unsicherheit in absoluten Zahlen (Intervallbreite) niedriger ausfällt als bei Frauen, relativ betrachtet allerdings in etwa der der Frauen entspricht.

Betrachtet man die durchschnittlichen Alter der lebenden Kinder, zeigen sich Veränderungen, die in etwa denen für Frauen entsprechen. Beispielsweise beträgt das durchschnittliche Alter von Kindern von Männern in Altersklasse 50 in der Modellrechnung für 2009 17.6 Jahre. 2060 liegt der Median bei 15.1. Die Unterschiede in der Höhe der Werte verglichen mit denen für Frauen der selben Altersklasse ergeben sich durch den Altersunterschied von Mann und Frau bei der Geburt eines Kindes, durch den Männer und Frauen der selben Altersklasse nicht direkt miteinander vergleichbar sind.

### 9.2.2 *Ergebnisse für Ostdeutschland*

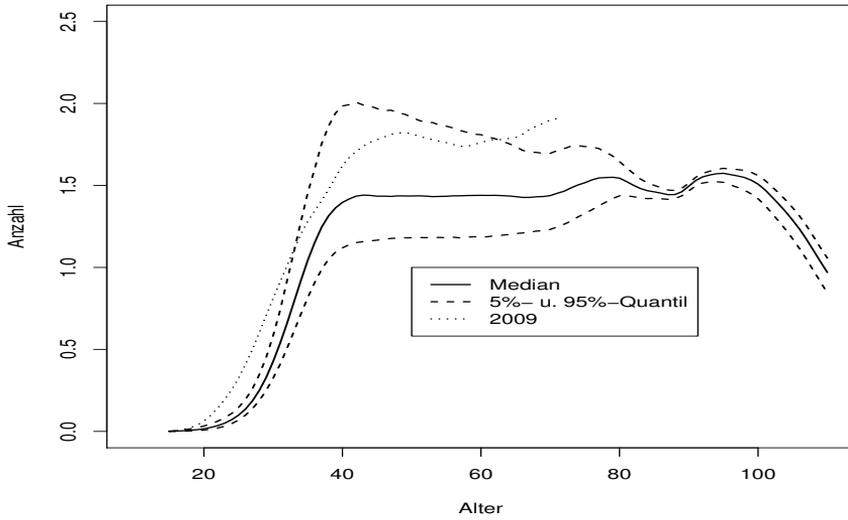
In der Abbildung 9.3 sind die Ergebnisse für ostdeutsche Frauen für das Jahr 2060 dargestellt. Zusätzlich sind noch Ergebnisse einer Modellrechnung für das Jahr 2009 eingezeichnet, die analog zur in Kapitel 7 vorgestellten Modellrechnung für Westdeutschland erstellt wurde. Dies gilt auch für alle im weiteren folgenden Ergebnisse für Ostdeutschland für das Jahr 2009.

Vergleicht man die in 9.3 abgebildeten Ergebnisse für ostdeutsche Frauen mit denen für westdeutsche Frauen, zeigen sich einige Unterschiede. Zunächst unterscheidet sich die Ausgangssituation, also die erwartete Anzahl lebender Kinder nach Alter im Jahr 2009 nach den Ergebnissen der Modellrechnung. Beispielsweise liegt der Wert für ostdeutsche Frauen in der Altersklasse von 30 Jahren bei 0.811, während bei westdeutschen Frauen ein Wert von 0.758 vorliegt. Bei der Altersklasse von 50 Jahren beträgt der Wert für Ostdeutschland 1.815 und der für Westdeutschland 1.580. Es werden also für die selben Altersklassen überwiegend höhere Werte erreicht als im Westen (vgl. auch Goldstein & Kreyenfeld 2011).

Für das Jahr 2060 entsprechen die Werte des Medians und des Vorausberechnungsintervalls für etliche der Altersklassen den Werten für Westdeutschland. Ursache für diese Ähnlichkeit ist die bei der Vorausberechnung der Fertilität angenommene Angleichung von Ost und West. In Kombination mit den bereits genannten Werten für das Jahr 2009 zeigt sich insgesamt, dass in Ostdeutschland stärkere Veränderungen zu erwarten sind als für Westdeutschland, wenn man den Median als Vergleichswert benutzt.

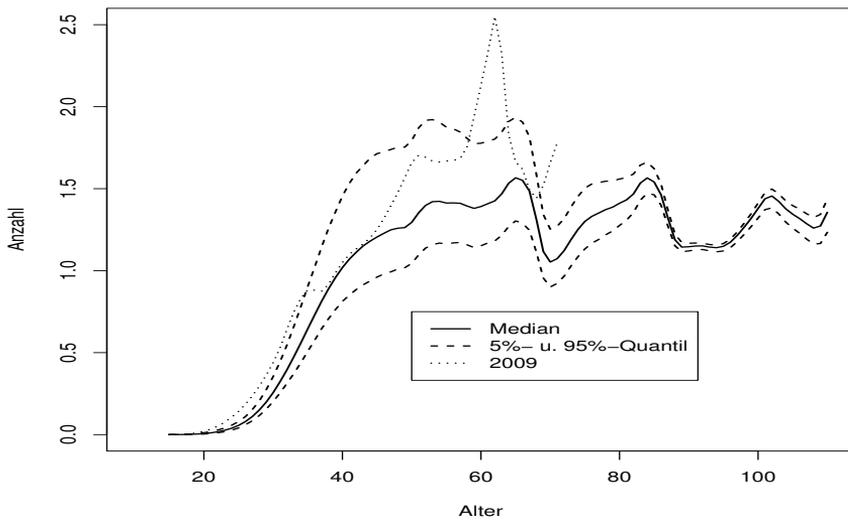
Dies gilt auch für das durchschnittliche Alter der lebenden Kinder. Für das Jahr 2009 ergibt sich für lebende Kinder von ostdeutschen Frauen in Altersklasse 50 ein Wert von 25.8, 2060 liegt der Median wie bei westdeutschen Frauen bei 18.0. Der Ausgangswert für 2009 liegt also um 4 Jahre höher als in Westdeutschland.

Abbildung 9.3: Erwartete Anzahl lebender Kinder ostdeutscher Frauen 2060 nach Altersklassen, sowie Ergebnisse für 2009



Quelle: eigene Berechnung

Abbildung 9.4: Erwartete Anzahl lebender Kinder ostdeutscher Männer 2060 nach Altersklassen, sowie Ergebnisse für 2009



Quelle: eigene Berechnung

Betrachtet man die in Abbildung 9.4 dargestellten Ergebnisse für ostdeutsche Männer, zeigen sich wie bei westdeutschen Männern deutliche Unregelmäßigkeiten. Insbesondere sticht der Wert der Altersklasse von 62 Jahren für 2009 mit einem Wert von 2.550 heraus. Der Wert für westdeutsche Männer der selben Altersklasse und für das selbe Jahr liegt bei 1.780. Als Ursache für diesen sehr hohen Wert kann angegeben werden, dass in Ostdeutschland die relativen Größenverhältnisse aufeinander folgender Geburtsjahrgänge von denen in Westdeutschland verschieden waren. Personen, die 2009 in Altersklasse 62 waren, wurden im Jahr 1947 geboren. Verwendet man die ersten für Ost und West zur Verfügung stehenden Bevölkerungsdaten und bildet über diese das Verhältnis der Zahl der Frauen dieses Jahrgangs zur Zahl der Männer des drei Jahre jüngeren Jahrgangs, ergibt sich für Westdeutschland ein Wert von 1.14, für Ostdeutschland aber ein Wert von 1.56. Insofern ist der Hohe Wert für Männer in Altersklasse 62 bei der für männliche Fertilität verwendeten Berechnungsmethode wenig überraschend.

Vergleicht man die Werte für das Jahr 2060 zwischen Ost und West zeigt sich wie bei Frauen, dass für etliche der Altersklassen sehr ähnliche Werte für den Median und das Vorausberechnungsintervall resultieren, wobei die Unterschiede allerdings aufgrund der gewählten Modellierung männlicher Fertilität etwas höher ausfallen als bei Frauen. Lediglich bei höheren Altersklassen zeigen sich wieder deutlichere Unterschiede, wobei teils für ostdeutsche und teils für westdeutsche Männer die jeweils höheren Werte vorliegen. Beispielsweise liegen die Werte für die Altersklasse von 30 Jahren bei 0.248 (West) und 0.256 (Ost), während bei der Altersklasse von 60 Jahren Werte von 1.250 und 1.391 erreicht werden. Insgesamt gilt ausgehend von den Werten für 2009 auch für Männer, dass für Ostdeutschland mit einer stärkeren Veränderung bis 2060 als im Westen zu rechnen ist.

Wie bei ostdeutschen Frauen zeigt sich dies im durchschnittlichen Alter der Kinder. In der Modellrechnung für 2009 ergibt sich für Altersklasse 50 ein durchschnittliches Alter der Kinder von 21.8 Jahren, während 2060 ein Median von 15.1 vorliegt, so das auch hier der Ausgangswert höher liegt als im Westen.

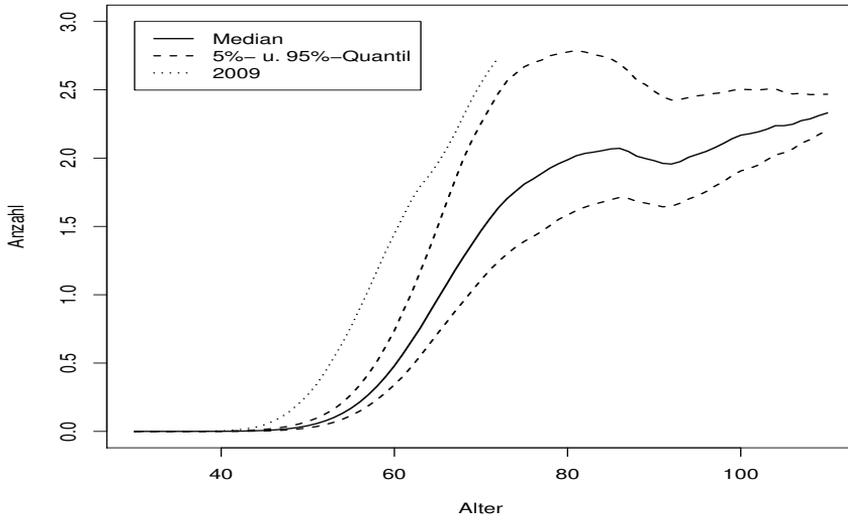
## **9.3 Ergebnisse zur Zahl der Enkelkinder**

### *9.3.1 Ergebnisse für Westdeutschland*

Der Median und das Vorausberechnungsintervall der erwarteten Zahl lebender Enkel westdeutscher Frauen nach Altersklasse im Jahr 2060 ist in Abbildung 9.5 zu sehen. Zusätzlich ist der über die Modellrechnung für das Jahr 2009 berechnete Verlauf dargestellt.

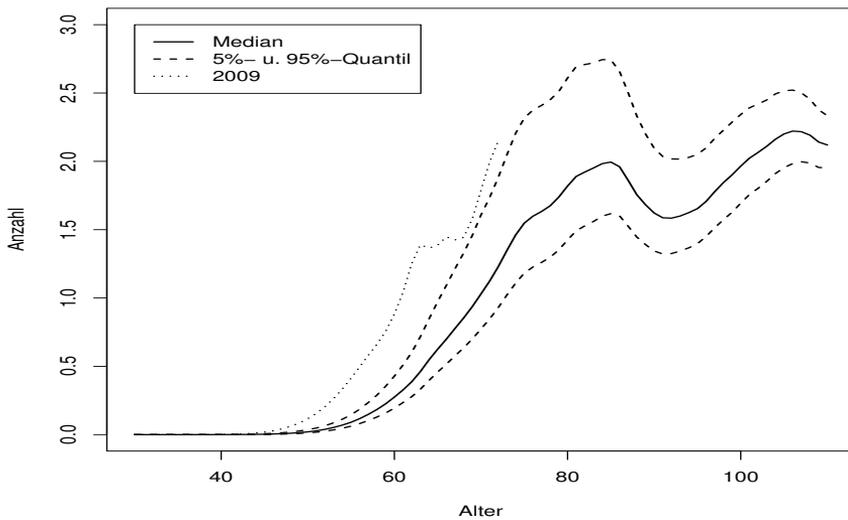
Entsprechend der Modellannahmen ist die Altersklasse von 30 Jahren die erste, bei der Werte größer Null erreicht werden. Bis zur Altersklasse 50 steigen die Werte im Jahr 2060 langsam an, wobei für diese Altersklasse ein Median von 0.042 erreicht wird. Anschließend steigt der Median bis zur Altersklasse von 86 Jahren bis auf einen Wert von 2.071, wobei sich dieser Anstieg zunächst bis zur Altersklasse von 60 Jahren beschleunigt und etwa ab der Altersklasse von 70 Jahren wieder langsamer ausfällt. Bis zur Altersklasse von 92 Jahren fällt der Median wieder auf einen Wert von 1.957 ab.

Abbildung 9.5: Erwartete Anzahl lebender Enkel westdeutscher Frauen 2060 nach Altersklassen, sowie Ergebnisse für 2009



Quelle: eigene Berechnung

Abbildung 9.6: Erwartete Anzahl lebender Enkel westdeutscher Männer 2060 nach Altersklassen, sowie Ergebnisse für 2009



Quelle: eigene Berechnung

Die Breite des Vorausberechnungsintervalls nimmt zunächst mit steigendem Alter zu. Die höchste Intervallbreite wird mit 1.278 für die Altersklasse von 70 Jahren erreicht, wobei das 5%-Quantil bei 1.390 und das 95%-Quantil bei 2.668 liegt. Anschließend nimmt die Intervallbreite wieder ab und erreicht beispielsweise bei der Altersklasse von 90 Jahren einen Wert von 0.833. Insgesamt kann die Unsicherheit somit als relativ hoch angesehen werden.

Ursache hierfür ist, dass für Personen in eher niedrigen Altern zwar die Unsicherheit über die Zahl der (lebenden) Kinder relativ groß ist, zugleich aber kaum davon auszugehen ist, dass diese Kinder wiederum selbst Kinder haben. Hierdurch fallen die Werte des Medians und des Vorausberechnungsintervalls sehr gering aus. Mit steigendem Alter steigt nun die Zahl der Kinder der Kinder und die Unsicherheit dieser Größe nimmt zu. Ab der Altersklasse von 65 Jahren allerdings nimmt nun langsam die Unsicherheit bezüglich der Zahl der Kinder der betrachteten Personen ab, da deren Fertilität nun mit steigendem Alter zu immer größeren Teilen historisch beobachtet wurde. Die Unsicherheit der Zahl der Kinder und die Unsicherheit der Zahl der Kinder der Kinder weisen nun unterschiedliche Entwicklungsrichtungen auf. Dabei nimmt die Unsicherheit der Zahl der Kinder zunächst nur langsam ab, da beispielsweise für Frauen in der Altersklasse von 70 Jahren lediglich die ersten Jahre der reproduktiven Phase beobachtet wurden, die allgemein mit relativ niedriger Fertilität einhergehen. Zugleich kumuliert sich bei dieser Altersklasse allerdings die Unsicherheit der Zahl der Kinder der Kinder, so dass die Unsicherheit insgesamt größer ausfällt als in den vorausgegangenen Altersklassen. Anschließend nimmt die Unsicherheit in der Zahl der Kinder allerdings in zunehmenden Maße ab, so dass Unsicherheit insgesamt zunehmend kleiner ausfällt. Bei hohen Altern kommt zudem hinzu, dass für diese auch die Fertilität der Kinder zumindest teilweise historisch beobachtet wurde und somit auch für die Zahl der Kinder der Kinder die Unsicherheit abnimmt.

Vergleicht man die Resultate der Vorausberechnung mit den Werten der Modellrechnung für 2009, zeigt sich zunächst, dass der Anstieg der erwartete Anzahl lebender Enkel mit steigendem Alter um etwa 10 Altersklassen nach hinten versetzt stattfindet, wenn man den Median für das Jahr 2060 als Vergleichswert heranzieht. Beispielsweise war im Jahr 2009 die Altersklasse von 57 Jahren die erste Altersklasse, für die ein Wert über 1 realisiert wurde, und die Altersklasse 62 die erste Altersklasse, bei der ein Wert über 1.5 erreicht wurde. 2060 sind dies die Altersklassen 66 und 71.

Zudem fallen beim direkten Vergleich der Ergebnisse einzelner Altersklassen alle für 2009 berechneten Werte höher aus als für 2060. Beispielsweise betrug der höchste erreichte Wert im Jahr 2009 2.736 und galt für die Altersklasse 72. Der höchste erreichte Median 2060 liegt bei 2.330 (Altersklasse 110). Betrachtet man nur die Altersklassen bis 95 Jahre, liegt der höchste Wert bei 2.071 (Altersklasse 86). Das höchste 2060 erreichte 95%-Quantil bezieht sich auf die Altersklasse 81 und liegt bei 2.788, womit es mit einer absoluten Differenz von 0.052 nur geringfügig höher liegt als der höchste Wert des Jahres 2009. Betrachtet man das Vorausberechnungsintervall insgesamt, zeigt sich, dass der Verlauf des Jahres 2009 für alle relevanten Altersklassen, bei denen die Anzahl lebender Enkel nicht gleich 0 oder nur sehr gering ist, vollständig außerhalb des Intervalls liegt, so dass eine Veränderung gegeben dieser Ergebnisse sehr wahrscheinlich scheint.

Dies gilt auch für die durchschnittlichen Alter der Enkelkinder. Beispielsweise ergibt sich für 2009 für das durchschnittliche Alter von lebenden Enkelkindern von westdeutschen Frauen

in Altersklasse 70 ein Wert von 12.9 Jahren. 2060 liegt der Median bei 7.4, das 5%-Quantil bei 6.6 und das 95%-Quantil bei 8.2. Dieser Unterschied ist darauf zurückzuführen, dass sich das steigende durchschnittliche Alter bei der Geburt gewissermaßen „doppelt“ auswirkt, nämlich sowohl bei der Geburt der Kinder als auch bei Geburt der Kinder der Kinder.

Der Median und das Vorausberechnungsintervall für die erwartete Anzahl lebender Enkel westdeutscher Männer nach Altersklassen im Jahr 2060 sind in Abbildung 9.6 zu sehen. Wie bei Frauen sind zusätzlich die Ergebnisse der Modellrechnung für das Jahr 2009 eingezeichnet.

Zunächst zeigt sich 2060 mit steigendem Alter ein Verlauf, der verglichen mit dem Verlauf westdeutscher Frauen größtenteils um etwa 3 bis 4 Altersklassen „nach hinten versetzt“ ist und etwas niedriger ausfällt. Beispielsweise ist die erste Altersklasse, bei der ein Median größer 1 erreicht wird, bei Frauen die Altersklasse 66 und bei Männern ist dies die Altersklasse von 70 Jahren.

Dabei zeigt sich nach einem anfänglichen Anstieg des Medians, der bis zur Altersklasse von 85 Jahren und einem Wert von 1.995 reicht, eine Abnahme der Anzahl lebender Enkel, die deutlich stärker ausfällt als bei Frauen. Dabei sinkt der Wert des Medians bis zur Altersklasse 92 auf einen Wert von 1.584.

Bezogen auf die Vorausberechnungsintervalle für einzelne Alter zeigt sich das selbe Bild wie bei Frauen: zunächst nimmt die Intervallbreite zu – hier bis zur Altersklasse von 81 Jahren und einer Breite von 1.202 – und anschließend wieder ab, so dass auch die „absolute“ Unsicherheit bei der Zahl der Enkel bei Männern zumindest grob dem Ergebnis für Frauen entspricht.

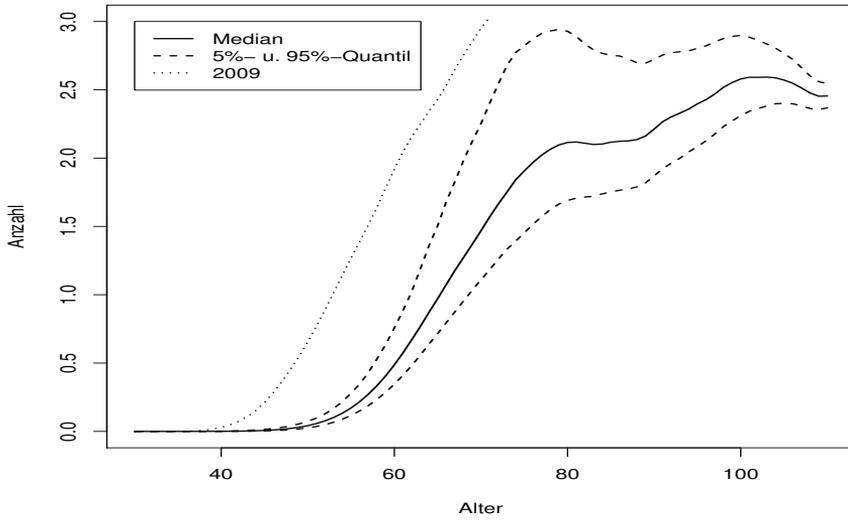
Vergleicht man die Resultate für das Jahr 2060 mit den Ergebnissen der Modellrechnung für 2009 zeigen sich wieder Ähnlichkeiten zu den Ergebnissen für westdeutsche Frauen. Insbesondere ist der Verlauf ebenfalls um einige Altersklassen „verschoben“. Beispielsweise ist die erste Altersklasse, bei der ein Wert von 1 überschritten wird, im Jahr 2009 die Altersklasse 61 und im Jahr 2060 die Altersklasse 70 (Median). Zudem liegt der Verlauf des Jahres 2009 für alle relevanten Alter außerhalb des Vorausberechnungsintervalls. Allerdings beträgt der höchste im Jahr 2009 erreichte Wert 2.150. Die höchste erreichte Obergrenze des Vorausberechnungsintervalls beträgt 2.744 und gilt für die Altersklasse von 84 Jahren, so dass zumindest in höhere Alter verlagert ähnlich hohe Werte wie 2009 nicht unplausibel erscheinen, wobei allerdings zu beachten ist, dass die Werte der Modellrechnung für 2009 bei Männern deutlich niedriger ausfallen als bei Frauen.

Betrachtet man die durchschnittliche Altersklasse von Enkelkindern von westdeutschen Männern, zeigen sich allerdings deutlichere Veränderungen, die in etwa den Veränderungen bei Frauen entsprechen. Beispielsweise beträgt die durchschnittliche Altersklasse von Enkeln von Männern in Altersklasse 70 für 2009 10.5 Jahre. 2060 wird ein Median von 5.8 Jahren erreicht, die Intervallbreite beträgt ungefähr 1.4 Jahre.

### 9.3.2 Ergebnisse für Ostdeutschland

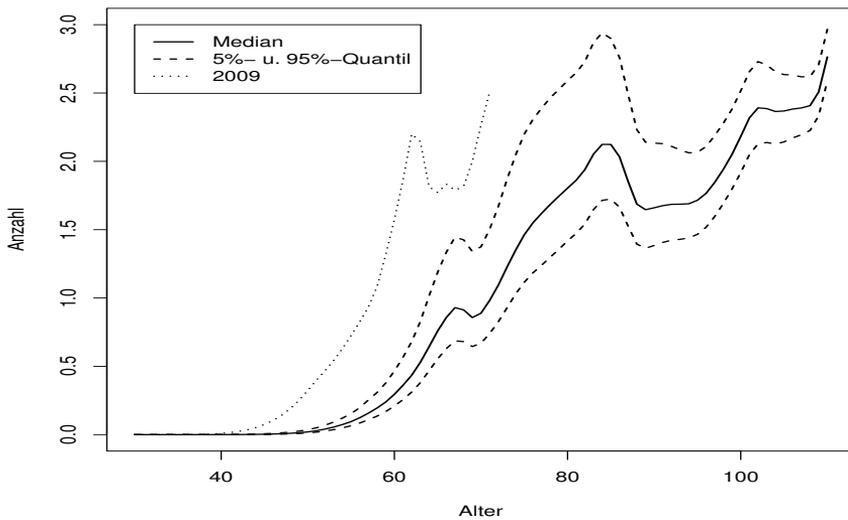
Median und Vorausberechnungsintervall der Verteilung der erwarteten Anzahl lebender Enkelkinder ostdeutscher Frauen im Jahr 2060 nach Alter finden sich in Abbildung 9.7. Entsprechende Ergebnisse für Männer sind in Abbildung 9.8 eingezeichnet. In beiden Fällen sind zusätzlich noch die Ergebnisse der Modellrechnung für 2009 zu sehen.

Abbildung 9.7: Erwartete Anzahl lebender Enkel ostdeutscher Frauen 2060 nach Altersklassen, sowie Ergebnisse für 2009



Quelle: eigene Berechnung

Abbildung 9.8: Erwartete Anzahl lebender Enkel ostdeutscher Männer in den Jahren 2009 und 2060 nach Altersklassen



Quelle: eigene Berechnung

Zunächst steigt der Median der erwarteten Anzahl lebender Enkel ostdeutscher Frauen wie bei westdeutschen Frauen an, wobei sehr ähnliche Werte erreicht werden. Beispielsweise liegt der Median für die Altersklasse von 60 Jahren bei einem Wert von 0.487 (westdeutsche Frauen: 0.480). Die erste Altersklasse, bei der ein Median über 1 erreicht wird, ist die Altersklasse 66 und die erste Altersklasse, bei der ein Median über 1.5 erreicht wird, ist die Altersklasse 71, wie bei westdeutschen Frauen. Dies ist wie bei den Ergebnissen zu Kindern auf die Annahmen bei der Vorausberechnung der Fertilität für Ostdeutschland zurückzuführen.

Für höhere Alter werden allerdings zunehmend deutlicher abweichende Werte realisiert. Beispielsweise liegt der Median für Altersklasse 80 bei 2.115 und bei Altersklasse 90 bei 2.213 (westdeutsche Frauen: 1.987 und 1.981). Zudem findet in den auf Altersklasse 81 folgenden Altersklassen zwar eine Abnahme des Medians statt, allerdings ist diese nur marginal und betrifft nur vier Altersklassen, während bei westdeutschen Frauen eine zwar immer noch nicht extreme, aber doch deutlichere Abnahme stattfindet. Nach dieser Abnahme steigt der Median bis zur Altersklasse von 101 auf einen Wert von 2.592 an und nimmt anschließend nochmals ab. Somit sind für ostdeutsche Frauen in höheren Altern insgesamt höhere Werte der Anzahl lebender Enkelkinder zu erwarten als für westdeutsche Frauen.

Der Verlauf des Vorausberechnungsintervalls entspricht im Wesentlichen dem bei westdeutschen Frauen: das Intervall wird zuerst langsam, dann anschließend relativ schnell breit und wird anschließend wieder etwas schmaler. Die höchste Breite des Intervalls liegt bei 1.380 und gilt für die Altersklasse von 74 Jahren. Dies liegt relativ nah an der Altersklasse von 70 Jahren, bei der für westdeutsche Frauen der höchste Wert erreicht wird, wobei dieser mit 1.278 niedriger ausfällt.

Vergleicht man die Ergebnisse für das Jahr 2060 mit den Ergebnissen der Modellrechnung für das Jahr 2009, zeigt sich, dass eine stärkere Verschiebung des Anstiegs der Kurve in höhere Alter als für Westdeutschland festzustellen ist. Beispielsweise wurde im Jahr 2009 der Wert 1 zum ersten Mal in Altersstufe 53 überschritten, für Westdeutschland war dies die Altersklasse von 57 Jahren. Für das Jahr 2060 trifft dies sowohl für Ost- als auch für Westdeutschland auf die Altersklasse von 66 Jahren zu, wenn der Median betrachtet wird. Dies bedeutet für Ostdeutschland eine Verschiebung um 13 Altersklassen, für Westdeutschland hingegen nur eine Verschiebung um 9 Altersklassen.

Wie bei westdeutschen Frauen liegen die für 2009 berechneten Werte teils deutlich über dem Median der Vorausberechnung. Beispielsweise lag der höchste für 2009 berechnete Wert bei 3.030, für 2060 liegt der höchste berechnete Median bei 2.593 und das höchste 95%-Quantil bei 2.943. Insofern scheinen auch für Ostdeutschland ähnlich hohe Werte wie 2009 dauerhaft eher unwahrscheinlich. Zudem liegt der Verlauf des Jahres 2009 für alle relevanten Alter außerhalb des Vorausberechnungsintervalls für 2060, so dass auch für ostdeutsche Frauen eine Veränderung des Verlaufs gegeben dieser Ergebnisse als sehr wahrscheinlich angesehen werden kann.

Dies zeigt sich auch bei der durchschnittlichen Altersklasse der Enkelkinder. Für 2009 ergibt sich für die Enkelkinder von ostdeutschen Frauen in Altersklasse 70 ein Wert von 18.5 Jahre und für 2060 wird ein Median von 7.5 Jahre erreicht. Der zuletzt genannte Wert entspricht in etwa dem Resultat für westdeutsche Frauen in der selben Altersklasse, während der Wert für 2009 deutlich höher ausfällt (West: 12.9 Jahre).

Der in Abbildung 9.8 zu sehende Verlauf des Medians und der Vorausberechnungsintervalle über alle Altersklassen im Jahr 2060 für ostdeutsche Männer fällt unregelmäßiger aus als der Verlauf bei westdeutschen Männern. Dabei ist der Verlauf aber ähnlich: zunächst findet ein einmal kurz unterbrochener Anstieg bis zur Altersklasse von 86 Jahren statt, anschließend kommt es zu einer deutlichen Abnahme und anschließend steigen die Werte wieder. Dabei werden sowohl höhere und niedrigere als auch ähnliche Werte wie bei westdeutschen Männern erreicht. Beispielsweise betragen die Mediane der Altersklassen 70, 80 und 90 der Reihe nach 0.889, 1.804 und 1.660. Für Westdeutschland liegen sie wie teils schon erwähnt bei 1.030, 1.819 und 1.622.

Bezogen auf die Vorausberechnungsintervalle zeigt sich ein ähnlicher Verlauf wie für die anderen bisher betrachteten Personengruppen, bei denen die Vorausberechnungsintervalle zunächst breiter und anschließend schmaler werden. Die höchste Breite eines Vorausberechnungsintervalls liegt bei 1.223, was grob dem Wert für westdeutsche Männer von 1.202 entspricht.

Vergleicht man die Werte der Modellrechnung für das Jahr 2009 mit den vorausberechneten Werten, zeigt sich wie bei Frauen und westdeutschen Männern eine Verschiebung des Anstiegs der erwarteten Enkelzahl in höhere Alter. Beispielsweise wird bei der Modellrechnung für 2009 zum ersten mal bei Altersklasse 58 der Wert 1 überschritten. 2060 ist dies beim Median für Altersklasse 72 der Fall, was einer Verschiebung um 14 Altersklassen entspricht. Bei westdeutschen Männern waren dies die Altersklassen 61 und 70, was einer Verschiebung um 9 Altersklassen entspricht. Insofern zeigen sich hier für Ostdeutschland stärkere Effekte als für Westdeutschland, ähnlich wie bei Frauen, was sich auch beim durchschnittlichen Alter der Enkelkinder zeigt: 2009 beträgt das durchschnittliche Alter von Enkelkindern von ostdeutschen Männern in Altersklasse 70 14.8 Jahre, 2060 beträgt der Median 6.1 Jahre.

Die höchste 2009 für eine Altersklasse erreichte Zahl an Enkeln betrug 2.506, was relativ deutlich unter dem höchsten 2060 auftretenden Wert des 95%-Quantils von 2.764 liegt. So scheinen Werte, wie sie 2009 aufgetreten sind, für 2060 nicht gänzlich unplausibel zu sein. Aber auch hier ist wieder zu beachten, dass der Verlauf des Jahres 2009 für alle relevanten Alter vollständig außerhalb des Vorausberechnungsintervalls liegt, so dass abermals eine Veränderung des Verlaufs als sehr wahrscheinlich angesehen werden kann.

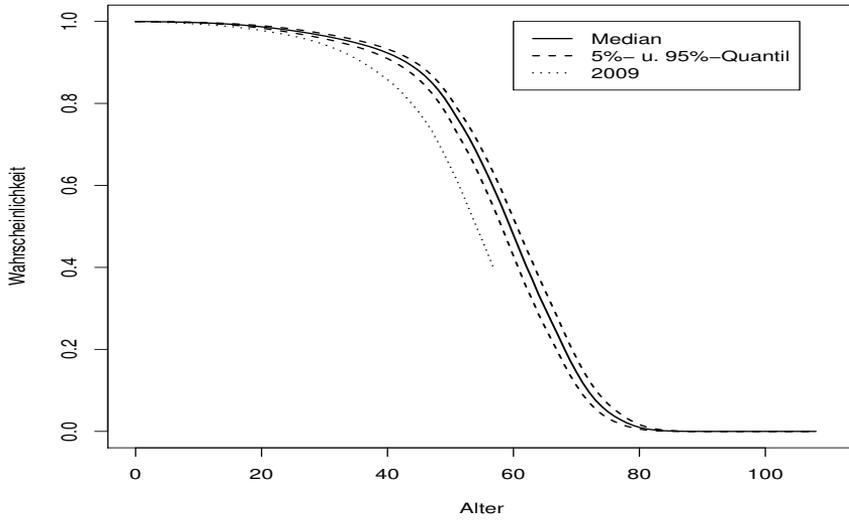
## **9.4 Ergebnisse zur Zahl lebender Eltern**

### *9.4.1 Ergebnisse für Westdeutschland*

Im Gegensatz zu Nachfahren wird bei Eltern eine Längsschnittperspektive gewählt, da die Entwicklung über alle Altersklassen relativ einheitlich ist. Als Beispiel sind in Abbildung 9.9 für Westdeutschland im Jahr 2060 Median und Vorausberechnungsintervall für die Verteilung der Wahrscheinlichkeit, dass die Mutter einer Person noch lebt, nach Altersklassen abgetragen. Zusätzlich sind die Ergebnisse der Modellrechnung für 2009 eingezeichnet.

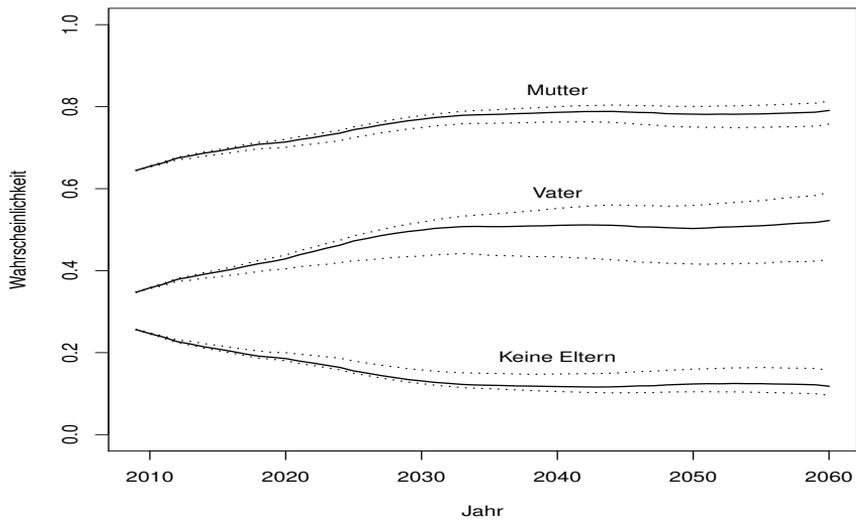
Wie zu sehen ist, findet im Wesentlichen eine „Verschiebung“ der Kurve statt. Betrachtet man beispielsweise die Altersklasse, für die die Wahrscheinlichkeit für eine lebende Mutter

Abbildung 9.9: Wahrscheinlichkeit einer lebenden Mutter nach Alter in den Jahren 2009 und 2060, Westdeutschland



Quelle: eigene Berechnung

Abbildung 9.10: Wahrscheinlichkeit einer lebenden Mutter, eines lebenden Vaters und für keine lebenden Eltern in Altersklasse 50 in den Jahren 2009-2060, Westdeutschland



Quelle: eigene Berechnung

zum ersten mal unter 0.75 sinkt, ist dies für das Jahr 2009 die Altersklasse 47 und für 2060 die Altersklasse 52. Verwendet man statt dem Wert 0.75 den Wert 0.50 findet man die Altersklassen 55 (2009) und 60 (2060). In beiden Fällen zeigt sich also eine Verschiebung um 5 Jahre.

Bei der Längsschnittbetrachtung werden lediglich Ergebnisse für die Altersklasse von 50 Jahren dargestellt, da zum einen bei deutlich jüngeren Altersklassen nur geringe Veränderungen auftreten (s. Abbildung 9.9). Zum anderen können für deutlich ältere Altersklassen keine Ausgangswerte für das Jahr 2009 berechnet werden (s. Kapitel 7). Zudem gibt die Altersklasse von 50 Jahren einen guten Eindruck der Entwicklung der Unsicherheit und der Entwicklung insgesamt. Zu beachten ist allerdings, dass sich bei dieser Altersklasse die vorausberechnete Entwicklung des durchschnittlichen Alters bei der Geburt praktisch nicht auswirkt, sondern nur die historische. Beispielsweise wurden Personen, die 2030 in Altersklasse 50 sind, im Jahr 1979 geboren. Das Alter der Eltern dieser Personen hängt also von der historischen Fertilität ab. Dies gilt auch für alle anderen Jahre der Vorausberechnung mit Ausnahme von 2060.<sup>200</sup>

In Abbildung 9.10 ist der Median und das Vorausberechnungsintervall der Verteilung der Wahrscheinlichkeit, dass die Mutter einer Person in Altersklasse 50 noch lebt, über die Jahre 2009 bis 2060 für Westdeutschland eingezeichnet (unabhängig vom Überleben des Vaters). Dabei ergibt sich der Wert des Jahres 2009 aus der Modellrechnung aus Kapitel 7. Ebenso sind die selben Kennwerte für den Vater eingezeichnet und die Wahrscheinlichkeit, dass weder Mutter noch Vater leben. In allen Fällen sind die Mediane jeweils als durchgezogene Linien eingezeichnet und die Vorausberechnungsintervalle als gestrichelte Linien.

Als Ausgangswert der Modellrechnung ergibt sich für 2009 bei der Wahrscheinlichkeit, dass die Mutter noch lebt, ein Wert von 0.645. Die Wahrscheinlichkeit, dass der Vater noch lebt, liegt bei 0.348. Hiervon ausgehend zeigt sich sowohl bei der Mutter als auch beim Vater eine relativ parallele Entwicklung: der Median der Verteilung der Wahrscheinlichkeit steigt zunächst in beiden Fällen bis etwa 2030 an. Für dieses Jahr liegt der Wert für die Mutter bei 0.770 und den Vater bei 0.499. Anschließend steigen beide Mediane nur noch leicht mit geringen Schwankungen. 2060 beträgt der Median für die Mutter 0.791 und für den Vater 0.522.

Als Ursache für dieses „Abflachen“ können die gegenläufigen Effekte der historischen Entwicklung des Alters bei der Geburt und der Lebenserwartung genannt werden. Das durchschnittliche Alter von lebenden Müttern von Personen in Altersklasse 50 beträgt in der Modellrechnung für 2009 beispielsweise 77.7 Jahre. 2030 liegt der Median bei 77.4 Jahren, wobei das 5%- und das 95%-Quantil nur geringfügig andere Werte aufweisen, und 2060 wird ein Median von 81.2 Jahren und eine Intervallbreite von ungefähr 0.4 Jahren erreicht. Wie aus diesen Werten zu erkennen ist, nimmt das durchschnittliche Alter der Mutter über den Vorausberechnungszeitraum zunächst leicht ab, während zugleich aber die Lebenserwartung steigt. Dies ist wiederum auf die historische Entwicklung des durchschnittlichen Alters bei der Geburt zurückzuführen, welches bis in die Mitte der 1970er Jahre zunächst gesunken ist. Frauen, die 2009 in Altersklasse 50 sind, wurden im Jahr 1958 geboren und Frauen, die 2030

---

200 Lediglich bei der Altersklasse 0 könnte sich der Effekt direkt zu Beginn der Vorausberechnung zeigen – wie aus Abbildung 9.9 aber zu erkennen ist, ist dies aufgrund der bereits im Ausgangsjahr sehr niedrigen Mortalität während der reproduktiven Phase nicht der Fall. Auch bei anderen, jüngeren Altersklassen zeigt sich dieses Problem.

in derselben Altersklasse sind, wurden 1979 geboren. Dabei war das durchschnittliche Alter bei der Geburt 1979 niedriger als 1958. Bis zum Ausgangsjahr 2009 und über den Vorausberechnungszeitraum steigt das durchschnittliche Alter bei der Geburt dann aber tendenziell an, so dass Mütter von Personen in Altersklasse 50 tendenziell älter werden, was zu einem der Abnahme der Sterblichkeit entgegengesetzten Effekt führt. Beim durchschnittlichen Alter des Vaters zeigt sich eine sehr ähnliche Entwicklung: 2009 beträgt das durchschnittliche Alter von Vätern von Personen in Altersklasse 50 81.0 Jahre, 2030 wird ein Median von 80.7 erreicht und 2060 schließlich liegt der Median bei 83.7 Jahren.

Bei der Wahrscheinlichkeit, dass weder Mutter noch Vater leben, zeigt sich, wie aufgrund der gerade genannten Ergebnisse zu erwarten ist, ein entgegengesetzter Verlauf. Der Ausgangswert der Modellrechnung für 2009 ist gleich 0.256. Bis 2030 sinkt der Median auf einen Wert von 0.131 ab. Bis 2060 nimmt dieser Wert mit leichten Schwankungen weiter ab und liegt schließlich bei 0.118.

Bezüglich der Vorausberechnungsintervalle zeigt sich in allen drei Fällen eine relativ stetige Zunahme. Das 5%-Quantil liegt für die Mutter im Jahr 2060 bei 0.758 und das 95%-Quantil bei 0.814. Für den Vater betragen diese Werte 0.427 und 0.591. Für die Wahrscheinlichkeit, dass beide Eltern nicht mehr leben, liegen sie bei 0.096 und 0.157. Insofern liegen in allen drei Fällen die oberen Grenzen des Vorausberechnungsintervalls näher am Median als die unteren Grenzen. Zudem ist das Vorausberechnungsintervall für den Vater mit einer Breite von etwa 0.164 deutlich ausgeprägter als das für die Mutter, das eine Breite von 0.056 aufweist. Vergleicht man die unteren Intervallgrenzen mit den Ausgangswerten, zeigt sich, dass für die Mutter ein Anstieg von etwas über 10 Prozentpunkten und für den Vater ein Anstieg von etwas unter 10 Prozentpunkten als relativ wahrscheinlich gelten kann.

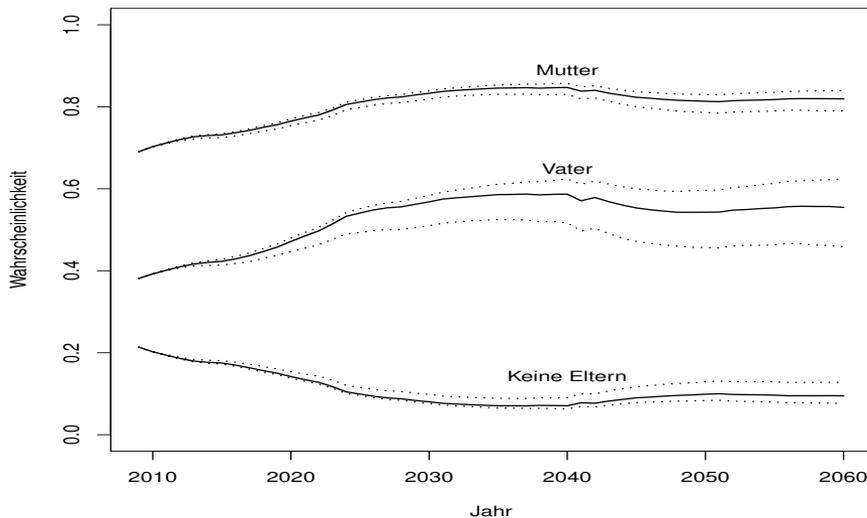
#### 9.4.2 *Ergebnisse für Ostdeutschland*

In Abbildung 9.11 sind der Median und das Vorausberechnungsintervall der Verteilung der Wahrscheinlichkeit, dass die Mutter einer Person in Altersklasse 50 in Ostdeutschland noch lebt, über die Jahre 2010 bis 2060 eingezeichnet, ebenso wie der Wert der Modellrechnung für 2009. Zudem finden sich in dieser Abbildung entsprechende Ergebnisse für die Wahrscheinlichkeit, dass der Vater noch lebt, und die Wahrscheinlichkeit, dass beide Elternteile nicht mehr leben.

Die Ausgangswerte der Modellrechnung für das Jahr 2009 betragen 0.691 (Mutter), 0.381 (Vater) und 0.214 (keine Eltern). Die Wahrscheinlichkeiten für eine lebende Mutter und einen lebenden Vater sind für Ostdeutschland also etwas höher als für Westdeutschland (0.645 und 0.348). Ursache hierfür ist wahrscheinlich das historisch niedrigere durchschnittliche Alter bei der Geburt. Personen, die 2009 in Altersklasse 50 sind, wurden im Jahr 1958 geboren. In der DDR lag für dieses Jahr das durchschnittliche Alter bei der Geburt bei 25.63 und in der BRD lag es bei 27.95. Ein gegenläufiger Effekt ist von der höheren Sterblichkeit in der DDR zu erwarten, wodurch sich der genannte Altersunterschied wahrscheinlich nur eingeschränkt bemerkbar macht.

Von diesen Werten ausgehend steigen sowohl die Wahrscheinlichkeit für die Mutter als auch die für den Vater mit geringen Schwankungen bis 2040 an. Für dieses Jahr liegt der Median bei der Mutter bei einem Wert von 0.847 und für den Vater bei 0.587. Anschließend

Abbildung 9.11: Wahrscheinlichkeit einer lebenden Mutter, eines lebenden Vaters und für keine lebenden Eltern in Altersklasse 50 in den Jahren 2009-2060, Ostdeutschland



Quelle: eigene Berechnung

nehmen die Werte zunächst wieder ab, wobei sich für 2041 sowohl bei Männern als auch bei Frauen ein „Knick“ zeigt, der bei Männern allerdings etwas ausgeprägter ist. Ursache hierfür ist der vergleichsweise drastische Anstieg des durchschnittlichen Alters bei der Geburt im Jahr 1990, der dem deutlichen Absinken der zusammengefassten Geburtenziffer nach der Wiedervereinigung vorausging (s. auch Fußnote 52 auf S. 42). 1989 lag das durchschnittliche Alter bei der Geburt bei etwa 24.23, während es 1990 bei 25.09 lag. Dies ist aber genau das Jahr, in dem Personen, die 2041 in Altersklasse 50 sind, geboren wurden.

Bis zum Jahr 2042 steigt sowohl bei der Mutter als auch beim Vater die Wahrscheinlichkeit wieder etwas an, nimmt in den folgenden Jahren zunächst leicht ab und steigt wieder geringfügig an. Auch hier zeigen sich wieder gegenläufige Effekte der Entwicklung des durchschnittlichen Alters bei der Geburt und der Lebenserwartung. Beispielsweise beträgt das durchschnittliche Alter einer lebenden Mutter einer Person in Altersklasse 50 in der Modellrechnung für 2009 75.2 Jahre. 2030 wird ein Median von 74.3 Jahren erreicht, 2060 ein Wert von 80.1 Jahren. Die Veränderungen fallen somit zwar stärker aus als für Westdeutschland, allerdings gelten die selben Erklärungen.

Im Jahr 2060 liegt der Median der Wahrscheinlichkeit einer lebenden Mutter bei 0.819 und der Median der Wahrscheinlichkeit eines lebenden Vaters bei 0.555. Diese Werte liegen zwar etwas über den Resultaten für Westdeutschland, sind diesen aber relativ ähnlich.

Bei den Vorausberechnungsintervallen zeigen sich insgesamt ähnliche Entwicklungen wie für Westdeutschland: zum einen nimmt die Intervallbreite für die Mutter und den Vater kontinuierlich zu, zum anderen liegt das 95%-Quantil näher am Median als das 5%-Quantil. Im Jahr 2060 liegt das 5%-Quantil der Wahrscheinlichkeit für die Mutter bei 0.790 und das 95%-Quantil bei 0.839. Für den Vater betragen die Werte 0.459 und 0.623. Insofern fallen

auch die Werte für die Quantile etwas höher als die Werte, die für Westdeutschland errechnet wurden. Die höchste Intervallbreite wird sowohl bei der Mutter als auch beim Vater im Jahr 2060 erreicht, ergibt sich somit aus den gerade genannten Werten und beträgt 0.049 (Mutter) beziehungsweise 0.164 (Vater). Dies entspricht in etwa den Intervallbreiten für Westdeutschland.

Die Entwicklung des Medians der Verteilung der Wahrscheinlichkeit, dass beide Eltern nicht mehr leben, verläuft der bei Mutter und Vater entgegengesetzt. 2060 liegt der Median bei 0.095, das 5%-Quantil bei 0.077 und das 95%-Quantil bei 0.128. In diesem Fall liegen die Werte also unter denen für Westdeutschland. Die Intervallbreite beträgt 0.051, was wie bei der Mutter und dem Vater in etwa dem westdeutschen Wert entspricht.

## 9.5 Ergebnisse zur Zahl lebender Großeltern

### 9.5.1 Ergebnisse für Westdeutschland

Bei der Darstellung der Ergebnisse zu Großeltern wird wie bei Eltern eine Längsschnittperspektive gewählt, allerdings werden Ergebnisse für die Altersklasse von 0 Jahren dargestellt. Die Altersklasse 0 wurde hier gewählt, weil für das Ausgangsjahr 2009 lediglich Wahrscheinlichkeiten für die Altersklassen 0 bis 5 berechnet werden konnten (s. Abschnitt 7.1), so dass für höhere Alter keine vollständige Betrachtung von 2009 bis 2060 möglich ist. Bei den Altersklassen 0 bis 5 liegen zwar für das Ausgangsjahr recht unterschiedliche Werte vor, allerdings ist die Entwicklung in allen sechs Fällen bezogen auf die Richtung der Entwicklung und das Ausmaß an Unsicherheit insgesamt relativ ähnlich, so dass die Altersklasse von 0 Jahren beispielhaft herausgegriffen wird.

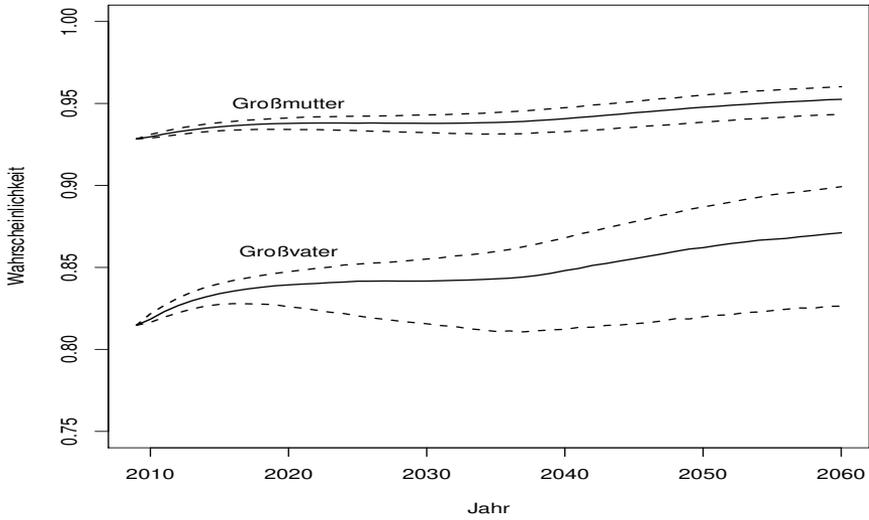
In Abbildung 9.12 sind für die Jahre 2010 bis 2060 der Median und das Vorausberechnungsintervall der Verteilung für die Wahrscheinlichkeit zu sehen, dass die Großmutter mütterlicherseits einer Person in Altersklasse 0 noch lebt (unabhängig vom Überleben anderer Großeltern), genauso wie der Wert der Modellrechnung für 2009. Ebenso sind entsprechende Resultate für den Großvater mütterlicherseits eingezeichnet. In Abbildung 9.13 wiederum finden sich Ergebnisse für Großeltern väterlicherseits.

Für die Wahrscheinlichkeit, dass die Großmutter mütterlicherseits noch lebt, ergibt sich in der Modellrechnung für das Jahr 2009 ein Wert von 0.928. Beim Großvater mütterlicherseits liegt dieser Wert bei 0.815. Für Großeltern väterlicherseits betragen die Werte 0.901 (Großmutter) und 0.756 (Großvater).

Der Median der Verteilung der Wahrscheinlichkeit für eine lebende Großmutter mütterlicherseits steigt bis 2020 zunächst leicht auf einen Wert von 0.938 an. Anschließend schwankt der Median länger um diesen Wert. Ab 2036 steigt der Median wieder an, wobei 2060 schließlich ein Wert von 0.953 erreicht wird. Das Vorausberechnungsintervall für dieses Jahr reicht von 0.943 bis 0.960, fällt also mit einer Breite von 0.017 relativ schmal aus.

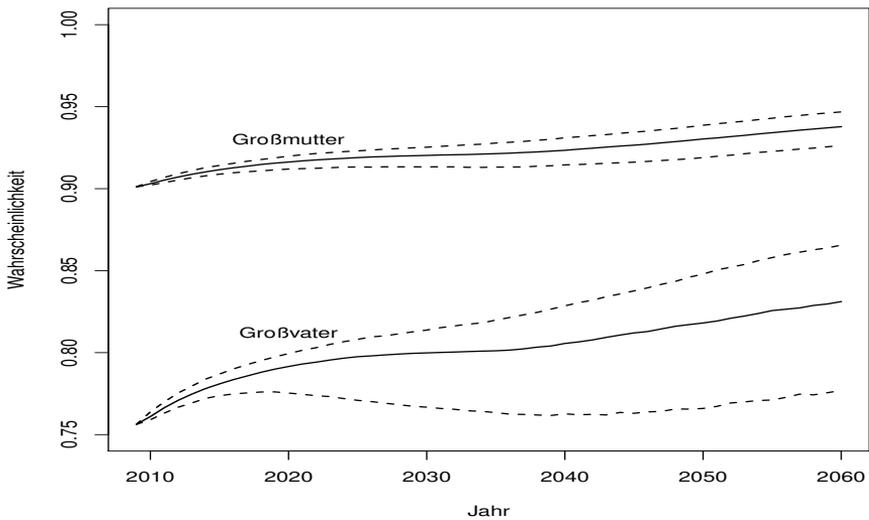
Beim Großvater mütterlicherseits zeigen sich wesentlich deutlichere Veränderungen. Zunächst findet wie bei der Großmutter mütterlicherseits ein Anstieg des Medians bis 2025 auf einen Wert von 0.842 statt. Anschließend schwankt der Median leicht um diesen Wert,

Abbildung 9.12: Wahrscheinlichkeiten für lebende Großeltern mütterlicherseits in Altersklasse 0 von 2009 bis 2060, Westdeutschland



Quelle: eigene Berechnung

Abbildung 9.13: Wahrscheinlichkeiten für lebende Großeltern väterlicherseits in Altersklasse 0 von 2009 bis 2060, Westdeutschland



Quelle: eigene Berechnung

steigt ab 2034 wieder an und liegt 2060 schließlich bei 0.871. Dabei fallen für den gesamten Vorausberechnungszeitraum die Vorausberechnungsintervalle deutlich breiter aus als bei der Großmutter. 2060 liegt das 5%-Quantil bei 0.826 und das 95%-Quantil bei 0.899, was einer Intervallbreite von 0.073 entspricht. Das 5%-Quantil liegt somit relativ nah am Wert des Jahres 2009.

Bei Großeltern väterlicherseits zeigt sich ausgehend von niedrigeren Werten für das Jahr 2009 eine sehr ähnliche Entwicklung, bei der sich insbesondere auch ähnliche Unterschiede zwischen Großvater und Großmutter zeigen. Bis 2060 steigt der Median der Wahrscheinlichkeit für die Großmutter auf 0.938 und die Wahrscheinlichkeit für den Großvater auf 0.831, so dass sich die Differenzen zu Großeltern mütterlicherseits verglichen mit 2009 etwas verringern. Die Intervallbreiten betragen 0.021 (Großmutter) und 0.089 (Großvater), fallen somit also etwas größer aus als bei Großeltern mütterlicherseits.

Insgesamt lässt sich festhalten, dass sich bei Großmüttern über den Vorausberechnungszeitraum hinweg nur geringfügige Veränderungen zeigen, wobei nur geringe Unsicherheit besteht. Bei Großvätern fallen die Veränderungen etwas deutlicher aus, sind aber ebenfalls nicht übermäßig hoch und zudem mit einer größeren Unsicherheit behaftet. Als Ursache für das geringe Ausmaß an Veränderung können die bereits bei den Ergebnissen zu Eltern genannten gegenläufigen Effekte des tendenziellen Anstiegs des durchschnittlichen Alters bei der Geburt und der Abnahme der Sterblichkeit geltend gemacht werden. Beispielsweise beträgt das durchschnittliche Alter der Großmutter mütterlicherseits von Personen in Altersklasse 0 für die Modellrechnung für 2009 59.3 Jahre, während der Median für 2060 bei 65.8 liegt. Zwar steigt die Lebenserwartung, zugleich müssen Großeltern aber aufgrund Veränderungen im durchschnittlichen Alter bei der Geburt älter werden, um zeitgleich mit ihren Enkeln zu leben.

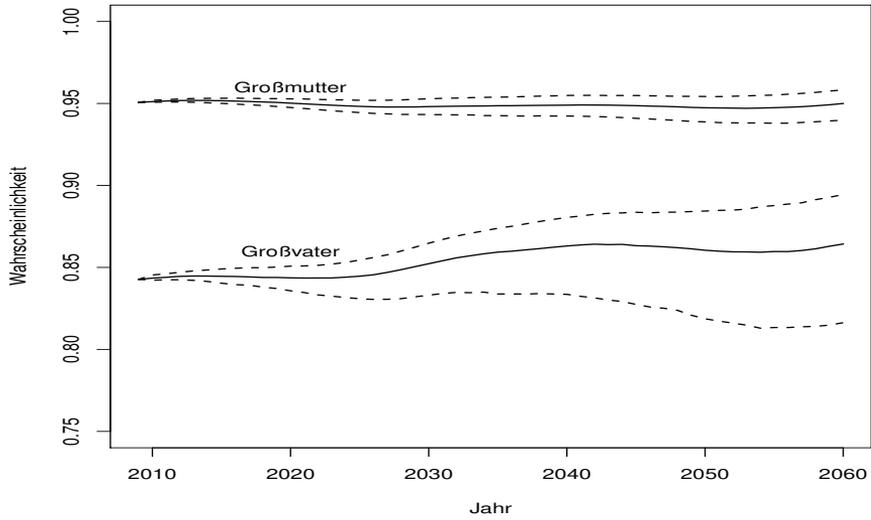
### 9.5.2 *Ergebnisse für Ostdeutschland*

In den Abbildungen 9.14 und 9.15 sind für die Jahre 2010 bis 2060 die Mediane und Vorausberechnungsintervalle der Verteilungen der Wahrscheinlichkeiten, dass die Großeltern mütterlicher- und väterlicherseits einer Person in Altersklasse 0 noch leben, für Ostdeutschland eingezeichnet. Zusätzlich sind die Werte aus der Modellrechnung für 2009 zu sehen.

Für alle Großeltern liegen die Werte der Modellrechnung des Jahres 2009 über den Werten für Westdeutschland. Beispielsweise beträgt der Wert für die Großmutter väterlicherseits 0.951 und für den Großvater mütterlicherseits 0.843, wobei letzterer sich deutlich vom Wert für Westdeutschland (0.756) unterscheidet.

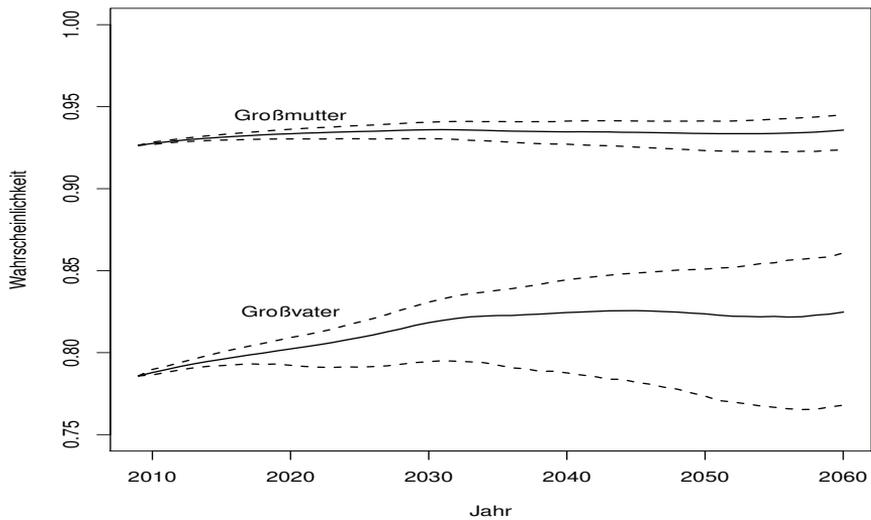
Bei der Großmutter mütterlicherseits zeigen sich bis 2060 lediglich geringfügige Veränderungen. Im Jahr 2060 wird ein Median von 0.950 erreicht, was ungefähr dem Wert des selben Jahres für Westdeutschland und dem Ausgangswert für 2009 entspricht. Mit einer Intervallbreite von 0.019 liegt nur eine relative geringe Unsicherheit vor, deren Umfang ebenfalls ungefähr dem Resultat für Westdeutschland entspricht. Somit ist hier also im Wesentlichen mit keiner Veränderung zu rechnen. Ursache hierfür sind wieder die zwei bereits genannten gegenläufige Effekte des Anstieg des durchschnittlichen Alters bei der Geburt, welcher für Ostdeutschland ausgehend von 2009 stärker ausfällt als im Westen und durch den die Wahrscheinlichkeit einer lebenden Großmutter mütterlicherseits abnimmt, und der Abnahme

Abbildung 9.14: Wahrscheinlichkeiten für lebende Großeltern mütterlicherseits in Altersklasse 0 von 2009 bis 2060, Ostdeutschland



Quelle: eigene Berechnung

Abbildung 9.15: Wahrscheinlichkeiten für lebende Großeltern väterlicherseits in Altersklasse 0 von 2009 bis 2060, Ostdeutschland



Quelle: eigene Berechnung

der Mortalität, durch welche die Wahrscheinlichkeit wiederum zunimmt. Betrachtet man beispielsweise das durchschnittliche Alter der Großmutter mütterlicherseits ergibt sich bei der Modellrechnung für 2009 ein Wert von 54.5 Jahren. 2060 liegt der Median aber bei 66.3 Jahren, so dass für Ostdeutschland ein deutlicherer Anstieg als für Westdeutschland festgestellt werden kann.

Bei der Wahrscheinlichkeit, dass der Großvater mütterlicherseits noch lebt, zeigen sich etwas deutlichere Veränderungen als bei der Großmutter mütterlicherseits. Ausgehend vom oben genannten Wert des Jahres 2009 bleibt der Median für die ersten Vorausberechnungsjahre ungefähr konstant, steigt anschließend etwas und schwankt schließlich leicht um einen Wert von etwa 0.860. 2060 liegt der Median bei 0.864, was in etwa dem westdeutschen Wert entspricht. Dies verweist somit auf einen leichten Anstieg, der sich wahrscheinlich dadurch erklärt, dass für ostdeutsche Männer insgesamt die deutlichsten Veränderungen in der Mortalität zu erwarten sind. Allerdings liegt das 5%-Quantil bei 0.816 und das 95%-Quantil bei 0.894, womit eine Intervallbreite von etwa 0.078 erreicht wird und auch der Wert des Jahres 2009 in das Vorausberechnungsintervall fällt, wodurch insgesamt keine eindeutige Entwicklungsrichtung ausgemacht werden kann.

Für das Jahr 2009 wird in der Modellrechnung für die Großmutter väterlicherseits ein Wert von 0.926 erreicht. 2060 liegt der Median bei 0.936, das 5%-Quantil bei 0.924 und das 95%-Quantil bei 0.945. Auch hier weisen die Ergebnisse auf eine nur niedrige Veränderung hin, die mit geringer Unsicherheit verbunden ist. Beim Großvater väterlicherseits zeigt sich wiederum eine etwas deutlichere Veränderung: für 2009 wird ein Wert von 0.786 erreicht, für 2060 beträgt der Median 0.825. Dabei ist allerdings mit einer Intervallbreite von 0.093 die Unsicherheit verglichen mit den anderen Großeltern relativ hoch und der Ausgangswert liegt innerhalb der Intervallgrenzen, die von 0.768 bis 0.861 reichen. Insgesamt liegen somit auch bei Großeltern väterlicherseits die Werte des Jahres 2009 über denen für Westdeutschland, wobei für 2060 aber die Werte des Medians und die Breite der Vorausberechnungsintervalle in etwa den Werten für Westdeutschland entsprechen.

## 9.6 Sensitivitätsanalysen

Um zu überprüfen, wie stark die vorgestellten Resultate von der Entwicklung einzelner Komponenten abhängen, wurden Sensitivitätsanalysen durchgeführt. Bei diesen wurden immer eine oder mehrere Komponenten konstant auf den Ausgangswerten gehalten und die Ergebnisse neu berechnet. Diese neuen Resultate können anschließend mit den ursprünglichen Simulationsergebnissen und auch untereinander verglichen werden. Dies erlaubt es zum einen, den Einfluss einzelner Komponente auf die Ergebnisse festzustellen. Zum anderen können die Resultate der Sensitivitätsanalyse dazu dienen, die bei den bisher vorgestellten Ergebnissen aufgezeigten Effekte einzelner Komponenten weitergehend zu plausibilisieren.

Insgesamt wurden sechs Varianten erstellt. Bei der ersten Variante wurden die Geburtenraten für alle 1000 Simulationsinstanzen konstant auf den Werten von 2008 gehalten, während bei Mortalität und Migration die Ergebnisse der stochastischen Vorausberechnungen benutzt

wurden.<sup>201</sup> Bei der zweiten Variante wurden die standardisierten Geburtenraten des Jahres 2008 für alle Vorausberechnungsjahre übernommen, wobei aber die zusammengefasste Geburtenziffer aus der stochastischen Vorausberechnung der Fertilität verwendet wurde.<sup>202</sup> So kann sich die Höhe der altersspezifischen Geburtenraten zwar ändern, allerdings bleibt das durchschnittliche Alter bei der Geburt konstant. Für Mortalität und Migration wurden wieder die stochastischen Vorausberechnungen verwendet. Letzteres gilt auch für die dritte Variante, bei der die zusammengefasste Geburtenziffer konstant auf dem Wert von 2008 gehalten wurde, während die standardisierten Geburtenraten aus der stochastischen Vorausberechnung übernommen wurden, so dass sich das durchschnittliche Alter bei der Geburt, nicht aber der Umfang der Fertilität ändern kann.

Bei der vierten Variante wurden alle Sterberaten auf den Werten des Jahres 2008 gehalten und Fertilität und Migration aus den stochastischen Vorausberechnungen übernommen. Bei der fünften Variante wurde sowohl Ein- als auch Auswanderung auf den Wert 0 gesetzt, so dass keine Wanderungen stattfinden. Dieses Vorgehen wurde gewählt, weil hierdurch keine Veränderung der Altersstruktur der Bevölkerung aufgrund von Migration stattfinden kann. Dies ist bei dem in dieser Arbeit verwendeten Modell die einzige Möglichkeit, über die Migration die Ergebnisse für Verwandtschaft beeinflussen kann. Fertilität und Mortalität wurden aus der stochastischen Vorausberechnung übernommen. Bei der sechsten Variante wurden schließlich alle Geburten- und Sterberaten auf den Werten von 2008 konstant gehalten und Ein- und Auswanderung auf den Wert 0 gesetzt.<sup>203</sup>

Insofern gibt es drei Varianten der Sensitivitätsanalyse für Fertilität, eine für Mortalität, eine für Migration sowie eine „Status-Quo“-Variante. Für die ersten fünf Varianten wurden jeweils die ursprünglichen 1000 Simulationsinstanzen mit den teils konstant gehaltenen Werten berechnet, bei der sechsten Variante ist nur eine Instanz notwendig. Dabei wurden wieder sowohl Ergebnisse für die Entwicklung von Verwandtschaft als auch für die Entwicklung der Bevölkerung erstellt. Die Ergebnisse der Sensitivitätsanalyse für die Bevölkerungsentwicklung finden sich in Anhang E.5, ebenso wie detailliertere Tabellen zu den Ergebnissen für Verwandtschaft, die hier nur anhand einiger Beispiele knapp zusammengefasst besprochen werden sollen. Im weiteren werden die sechs Varianten der Einfachheit halber als S1, S2, S3, S4, S5 und S6 bezeichnet. Einen Überblick über die Varianten gibt Tabelle 9.1.

### 9.6.1 Ergebnisse für Nachfahren

In Tabelle 9.2 sind ausgewählte Ergebnisse der Sensitivitätsanalyse für Nachfahren zu sehen. Zum einen sind jeweils der Median und die Grenzen des Vorausberechnungsintervalls für die

201 Wenn davon die Rede ist, dass eine demographische Komponente auf den Werten von 2008 konstant gehalten wurde, wurden für West- und Ostdeutschland jeweils die entsprechenden historischen Werte verwendet.

202 Die standardisierten Geburtenraten für ein Jahr  $t$  ergeben sich als

$$f_{x,t}^s = \frac{f_{x,t}}{\sum_{i=\alpha}^{\beta} f_{i,t}}$$

203 Entsprechend dieses Schemas des „Konstant-haltens“ von Komponenten, sind keine Sensitivitätsanalysen für die Verteilungen  $\phi(y|x)$  möglich, da diese bereits bei den ursprünglichen Vorausberechnungen als konstant angenommen werden.

*Tabelle 9.1: Übersicht der Varianten der Sensitivitätsanalyse*

Variante	Veränderungen ggü. der eigentlichen Vorausberechnung
S1	Geburtenraten konstant
S2	Standardisierte Geburtenraten konstant
S3	Zusammengefasste Geburtenziffer konstant
S4	Sterberaten konstant
S5	keine Migration
S6	Alle Raten konstant, keine Migration

Quelle: eigene Darstellung

Verteilung der erwarteten Anzahl lebender Kinder westdeutscher Frauen in Altersklasse 50 im Jahr 2060 bei den sechs Varianten der Sensitivitätsanalyse eingetragen.<sup>204</sup> Zum anderen finden sich entsprechende Ergebnisse für die erwartete Anzahl lebender Enkel von Frauen in Altersklasse 70.<sup>205</sup>

#### 9.6.1.1 Kinder

Vergleicht man bei den Resultaten für Kinder zunächst die Varianten S1, S4 und S5 mit den eigentlichen Ergebnissen der Vorausberechnung, zeigt sich, dass die deutlichsten Unterschiede auftreten, wenn Fertilität konstant gehalten wird, während dies bei Mortalität nur einen sehr geringen und bei Migration keinen feststellbaren Effekt hat. Dies ist wenig überraschend, da zum einen die Kinder von Frauen in Altersklasse 50 in Altersklassen sein werden, für die Mortalität vernachlässigbar ist. Zum anderen wird bei den formalen Ausdrücken für die Kinderzahl von Frauen Migration weder direkt noch indirekt über die Bevölkerungsstruktur berücksichtigt. Die Mediane der Status-Quo-Variante S6 liegen sowohl bei Kindern als auch bei Enkelkindern jeweils unter den Resultaten der eigentlichen Vorausberechnung und unter dem Ausgangswert des Jahres 2009.

Betrachtet man den Einfluss der Fertilität auf die Resultate differenzierter und vergleicht die Varianten S2 und S3, lässt sich deutlich erkennen, dass bei S3 ein Median erreicht wird, der nur geringfügig unter dem der eigentlichen Vorausberechnung liegt, während die Intervallbreite deutlich geringer ausfällt (0.106 verglichen mit 0.734). Bei S2 hingegen ist die Abweichung vom Median absolut betrachtet größer, wenn auch noch immer relativ niedrig ( $-0.051$ ), während die Breite des Vorausberechnungsintervalls deutlich näher an der ursprünglichen Vorausberechnung liegt (0.673). Insofern hat bei der hier betrachteten Altersklasse vor allem der Umfang der Fertilität einen starken Einfluss auf die Ergebnisse, wohingegen der Einfluss des Timing nur sehr gering ausfällt. Da die betrachtete Altersklasse am Ende der reproduktiven

<sup>204</sup> Für die Variante S6 lassen sich keine Vorausberechnungsintervalle berechnen. Ferner ist es bei dieser Variante nicht korrekt, vom „Median“ zu sprechen. Der Einfachheit und Einheitlichkeit halber wird dieser Begriff aber dennoch verwendet.

<sup>205</sup> Für die Betrachtung von Altersklasse 50 spricht, dass bei der zukünftigen Entwicklung die Zahl der lebenden Kinder insgesamt bedeutsamer ist als die genauen Alter, in denen diese geboren werden. Altersklasse 70 wurde verwendet, da einerseits für deutlich jüngere Altersklassen nur eine sehr geringe Anzahl an Enkeln zu erwarten ist, während für höhere Altersklassen teils noch die historische Entwicklung einen stärkeren Einfluss auf die Ergebnisse ausübt.

*Table 9.2: Kennwerte für die Verteilung der erwarteten Anzahl lebender Kinder und Enkelkinder westdeutscher Frauen 2060 bei den sechs Varianten der Sensitivitätsanalyse und der eigentlichen Vorausberechnung*

Kennwert	V	S1	S2	S3	S4	S5	S6
Kinder (AK 50)							
5%-Quantil	1.183	1.370	1.143	1.369	1.177	1.183	—
Median	1.436	1.372	1.385	1.425	1.428	1.436	1.365
95%-Quantil	1.916	1.373	1.816	1.475	1.906	1.917	—
Enkel (AK 70)							
5%-Quantil	1.107	1.456	1.119	1.379	1.098	1.114	—
Median	1.468	1.466	1.475	1.455	1.455	1.480	1.463
95%-Quantil	2.254	1.475	2.258	1.530	2.235	2.264	—

Abkürzungen: AK=Altersklasse, V=Vorausberechnung, Sx=Variante x der Sensitivitätsanalyse  
 Quelle: eigene Berechnung

Phase liegt, ist dieses Ergebnis aber wenig überraschend, da unwesentlich ist, in welchem Alter genau die Kinder geboren wurden (s. auch Fußnote 205 auf S. 212).

#### 9.6.1.2 Enkelkinder

Betrachtet man als nächstes die in Tabelle 9.2 zu sehenden Ergebnisse zu Enkelkindern von westdeutschen Frauen in Altersklasse 70, zeigt sich zunächst bei allen der sechs Varianten, dass die Werte des Medians nur geringfügig vom Median der eigentlichen Vorausberechnung abweichen. Für S2, S4 und S5 fallen auch die Vorausberechnungsintervalle ähnlich aus. Bei S1 und S3 hingegen sind die Vorausberechnungsintervalle deutlich schmaler. Hier zeigt sich, dass die Unsicherheit über die erwartete Anzahl lebender Enkel vor allem auf die Unsicherheit über die Höhe der Fertilität zurückzuführen ist. Ein nennenswerter Unterschied zu den Ergebnissen für Kinder ist, dass bei S5 nun nicht mehr die selben Ergebnisse wie bei der eigentlichen Vorausberechnung erzielt werden und der Median verglichen mit dieser geringfügig höher ist (+0.012). Die Ursache hierfür dürfte sein, dass für die Anzahl an Enkeln auch die Zahl der Kinder von Söhnen mit entscheidend ist. Die Fertilität von Männern kann aber wie in Kapitel 6 angedeutet aufgrund der hier gewählten Berechnung durch Wanderungen beeinflusst werden.

Vergleicht man die Ergebnisse von Variante S6 mit den Ausgangswerten der Modellrechnung für das Jahr 2009, zeigt sich, dass die Resultate dieser Variante deutlich unter dem Ausgangswert liegen. Vergleicht man sie hingegen mit dem Median der eigentlichen Vorausberechnung, zeigt sich kaum ein Unterschied.

### 9.6.2 Ergebnisse für Vorfahren

In Tabelle 9.2 sind ausgewählte Ergebnisse der Sensitivitätsanalyse für Vorfahren zu sehen. Eingetragen sind die Mediane und Vorausberechnungsintervalle der Wahrscheinlichkeit, dass die Mutter einer westdeutschen Person in Altersklasse 50 noch lebt. Ebenso finden sich entsprechende Ergebnisse für die Wahrscheinlichkeit, dass die Großmutter mütterlicherseits einer Person in Altersklasse 0 noch lebt.<sup>206</sup> Detailliertere Ergebnisse finden sich im Anhang in den Tabellen E.34, E.35 und E.36.

#### 9.6.2.1 Eltern

Betrachtet man die Ergebnisse für die Wahrscheinlichkeit einer lebenden Mutter, fällt zunächst bei den Varianten S1, S2, S3 und S5 auf, dass sich die Mediane und Vorausberechnungsintervalle wenn überhaupt nur marginal von denen der ursprünglichen Vorausberechnung unterscheiden, wobei beispielsweise der Median der eigentlichen Vorausberechnung bei 0.791 liegt. Insofern sind die Ergebnisse zu lebenden Eltern bei der betrachteten Altersklasse im Jahr 2060 wie zu erwarten praktisch nicht durch zukünftige Veränderungen der Fertilität und Migration beeinflusst. Denn die Höhe der Fertilität hat keinen Einfluss auf die Wahrscheinlichkeit lebender Eltern. Da Personen, die 2060 in Altersklasse 50 sind, im Jahr 2009

---

<sup>206</sup> Die Altersklassen 50 und 0 wurden aus den selben Gründen gewählt wie bei der Darstellung der Ergebnisse der eigentlichen Vorausberechnung.

**Table 9.3: Kennwerte für die Verteilung der Wahrscheinlichkeit einer lebenden Mutter und einer lebenden Großmutter mütterlicherseits in Westdeutschland 2060 bei den sechs Varianten der Sensitivitätsanalyse und der eigentlichen Vorausberechnung**

Kennwert	V	S1	S2	S3	S4	S5	S6
Mutter (AK 50)							
5%-Quantil	0.758	0.759	0.759	0.758	0.608	0.758	—
Median	0.791	0.792	0.792	0.791	0.612	0.791	0.616
95%-Quantil	0.814	0.814	0.814	0.814	0.620	0.814	—
GM (AK 0)							
5%-Quantil	0.943	0.954	0.953	0.943	0.894	0.943	—
Median	0.953	0.961	0.961	0.953	0.904	0.952	0.920
95%-Quantil	0.960	0.966	0.966	0.960	0.914	0.960	—

Abkürzungen: AK=Altersklasse, V=Vorausberechnung, Sx=Variante x der Sensitivitätsanalyse, GM=Großmutter  
 Quelle: eigene Berechnung

geboren wurden, ist wie in Abschnitt 9.4.1 bereits erwähnt die Unsicherheit bezüglich des durchschnittlichen Alters bei der Geburt sehr gering. Das durchschnittliche Alter weicht bei der Vorausberechnung also im Wesentlichen kaum von dem Wert des Jahres 2008 ab.

Bei S4 sieht man deutlichere Abweichungen von der eigentlichen Vorausberechnung. Vergleicht man den in Tabelle 9.3 angegebenen Median mit dem Median der Vorausberechnung ergibt sich eine Differenz von  $-0.179$ , womit der Median von S4 deutlich niedriger ausfällt. Gleiches gilt für das Vorausberechnungsintervall, dessen Breite verglichen mit dem eigentlichen Wert von  $0.056$  ebenfalls deutlich niedriger ist. Verglichen mit dem Wert der Modellrechnung für das Jahr 2009 ergibt sich eine Differenz von  $-0.033$ .

Dies deutet darauf hin, dass der in der eigentlichen Vorausberechnung festgestellte Anstieg der Wahrscheinlichkeit für lebende Eltern wenig überraschend primär auf Veränderungen in der Mortalität zurückzuführen ist und die Unsicherheit dieses Anstiegs sich aus der Unsicherheit der Entwicklung der Mortalität ergibt. Zugleich zeigt die Differenz mit der Modellrechnung für 2009 klar den gegenläufigen Effekt des Anstiegs des durchschnittlichen Alters bei der Geburt: hält man die Sterblichkeit konstant, ist wie oben erwähnt lediglich ein Effekt des historischen Anstiegs des durchschnittlichen Alters bei der Geburt möglich, der negativ ausfällt. Ähnliche Effekte zeigen sich auch bei der Variante S6.

#### 9.6.2.2 Großeltern

Betrachtet man die in Tabelle 9.3 dargestellten Ergebnisse für die Großmutter mütterlicherseits, zeigen sich bei den Varianten S3 und S5 keine Unterschiede zu den ursprünglichen Ergebnissen der Vorausberechnung.

Bei S1 und S2 zeigt sich ein relativ geringer Effekt. Die Differenz vom Median von S1 zum Median der ursprünglichen Vorausberechnung beträgt beispielsweise  $0.008$ . Der insgesamt stärkste Effekt zeigt sich wie zu erwarten bei S4. Beispielsweise weicht der Median um den Wert  $-0.049$  von der ursprünglichen Vorausberechnung ab. Vergleicht man die Intervallbreiten ergibt sich eine Differenz von  $0.003$  – die Intervallbreite fällt bei S4 also marginal höher aus als bei der eigentlichen Vorausberechnung. Die Differenz des Medians zur Modellrechnung für 2009 beträgt  $-0.025$ . Diese Ergebnisse lassen sich wie die Ergebnisse zur Mutter interpretieren: den stärksten Effekt auf die zukünftige Entwicklung hat die Entwicklung der Sterblichkeit. Zugleich geht aber von der Entwicklung des durchschnittlichen Alters bei der Geburt ein gegenläufiger Effekt aus. Ähnliches zeigt sich bei der Variante S6.

## 9.7 Zusammenfassung

**Kinder** Vorausberechnet wurde die erwartete Kinderzahl nach Alter und die Ergebnisse für die Jahre 2009 und 2060 wurden gegenübergestellt. Es lässt sich festhalten, dass der Vergleich der Ergebnisse der Modellrechnung für 2009 und der Medianwerte der Vorausberechnung auf eine langfristige Abnahme hindeutet, insbesondere für höhere Altersklassen. Dabei fällt diese Abnahme allerdings nicht für alle Altersklassen gleich stark aus und ist zudem mit einer relativ hohen Unsicherheit behaftet.

Vergleicht man die Resultate für Männer und Frauen, zeigt sich, dass sich die erwartete Anzahl lebender Kinder bei Männern wesentlich erratischer verhält als bei Frauen. Beim Vergleich der Ergebnisse für West- und Ostdeutschland kann eine Angleichung bis 2060 festgestellt werden. Dabei sind allerdings die unterschiedlichen Ausgangswerte für das Jahr 2009 zu bedenken, so dass für Ostdeutschland von einer stärkeren Abnahme als für Westdeutschland gesprochen werden kann.

**Enkelkinder** Vorausberechnet wurde die erwartete Anzahl an Enkelkindern nach Alter und wie bei Kindern wurden die Ergebnisse für die Jahre 2009 und 2060 miteinander verglichen. Für Männern und Frauen und für West- und Ostdeutschland liegt der „historische“ Verlauf der Modellrechnung 2009 für alle relevanten Alter außerhalb des Vorausberechnungsintervalls. Insofern scheint eine Veränderung des Verlaufs gegeben dieser Ergebnisse relativ wahrscheinlich zu sein.

Bei Frauen zeigt sich für 2060 zudem, dass zum einen nicht so hohe Werte wie für 2009 zu erwarten sind und zum anderen der Anstieg der Zahl der Enkel auf höhere Lebensalter verlagert wird, wobei sich letzteres bei ostdeutschen Frauen deutlich stärker bemerkbar macht als bei westdeutschen Frauen. Dabei ist allerdings die Unsicherheit über die Zahl der Enkel, die sich in den Vorausberechnungsintervallen ausdrückt, sehr groß. Bei Männern zeigt sich ebenfalls eine Verschiebung des Anstiegs der Enkelzahl in höhere Altersklassen, mit dem selben Unterschied zwischen West- und Ostdeutschland wie bei Frauen. Allerdings sind dabei mit dem Jahr 2009 vergleichbare Enkelanzahlen wahrscheinlicher als bei Frauen, wobei allerdings sowohl west- als auch ostdeutsche Männer für 2009 deutlich weniger Enkel aufwiesen als Frauen.

**Eltern** Vorausberechnet wurden die Wahrscheinlichkeiten, dass Mutter und Vater einer Person noch leben. Bei der Darstellung der Ergebnisse wurde von einer Person im Alter von 50 Jahren ausgegangen und betrachtet, wie sich die Wahrscheinlichkeiten im Zeitraum von 2009 bis 2060 ändern. Vergleicht man die Ergebnisse der Vorausberechnung für das Jahr 2060 mit denen der Modellrechnung für 2009 lässt sich festhalten, dass die Wahrscheinlichkeit, dass zumindest noch ein Elternteil einer Person in Altersklasse 50 noch lebt, deutlich zunimmt und die Wahrscheinlichkeit, keine lebenden Eltern mehr zu haben, deutlich abnimmt. Allerdings verdeckt diese Entwicklung zwei gegenläufige Effekte, die auf die genannten Wahrscheinlichkeiten einwirken: zum einen der Anstieg des durchschnittlichen Alters bei der Geburt, der einen negativen Effekt aufweist; und zum anderen die Abnahme der Mortalität, von der ein positiver Effekt ausgeht.

Unterschiede zwischen den Geschlechtern, die sich für das Jahr 2009 zeigen, bleiben größtenteils bestehen. Die durch die Vorausberechnungsintervalle ausgedrückte Unsicherheit fällt bei der Mutter insgesamt relativ niedrig aus. Beim Vater hingegen ist die Unsicherheit größer. Beim Vergleich der Resultate für West- und Ostdeutschland zeigen sich einige Unterschiede, die zum einen auf Unterschiede in den demographischen Komponenten vor Beginn des Vorausberechnungszeitraums zurückzuführen sind und zum anderen aus den in Kapitel 8 gemachten Annahmen resultieren.

**Großeltern** Vorausberechnet wurden die Wahrscheinlichkeiten, dass die einzelnen Großeltern einer Person noch leben. Bei der Darstellung der Ergebnisse wurde wie bei der Betrachtung

der Ergebnisse zu Eltern nur eine einzelne Altersklasse betrachtet, nämlich die Altersklasse von 0 Jahren, und dargestellt, wie sich die Wahrscheinlichkeiten im Zeitraum von 2009 bis 2060 ändern. Es zeigte sich, dass für die Altersklasse von 0 Jahren nur geringfügige beziehungsweise keine Veränderungen bezüglich der Wahrscheinlichkeiten für lebende Großmütter zu erwarten sind. Bei Großvätern zeigen sich zwar deutlichere Veränderungen, allerdings halten sich auch diese überwiegend in gewissen Grenzen und sind zudem mit wesentlich höherer Unsicherheit als bei Großmüttern behaftet. Ursache für die geringe Veränderung ist wie bei Eltern der Effekt des Anstiegs des durchschnittlichen Alters bei der Geburt, der sich bei Großeltern gewissermaßen doppelt auswirkt und der dem Effekt der Abnahme der Sterblichkeit entgegenwirkt.

Beim Vergleich der Ergebnisse für West- und Ostdeutschland zeigen sich zwar Unterschiede in den Ergebnissen der Modellrechnung für 2009 und den Ergebnissen der Vorausberechnung, die noch relativ stark durch historische Unterschiede beeinflusst sind, allerdings kann bis 2060 eine Angleichung festgestellt werden.

**Sensitivitätsanalyse** Bei den Ergebnissen der Sensitivitätsanalyse lässt sich wenig überraschend festhalten, dass bei den Ergebnissen für Nachfahren die Höhe der Fertilität die mit Abstand wichtigste Einflussgröße ist. Bei Vorfahren ist vor allem Mortalität bedeutsam. Gleichzeitig zeigt sich aber ein Einfluss der Entwicklung des durchschnittlichen Alters bei der Geburt. Zu beachten ist dabei, dass hier nur einige ausgewählte Altersklassen betrachtet wurden. Bei anderen Altersklassen ergeben sich gegebenenfalls andere Veränderungen, wenn einzelne Komponenten konstant gehalten werden.

**Überblick** Zusammenfassend und stark vereinfachend lassen sich folgende Resultate festhalten:

- weniger Nachfahren, verbunden mit großer Unsicherheit
- deutlich höhere Wahrscheinlichkeiten für lebende Eltern, verbunden mit mäßiger Unsicherheit (je nach Geschlecht)
- leicht höhere beziehungsweise gleichbleibende Wahrscheinlichkeit für lebende Großeltern, verbunden mit mäßiger Unsicherheit (je nach Geschlecht)
- teils starker Einfluss der historischen Entwicklung auf die zukünftige Entwicklung
- langfristige Angleichung der Verhältnisse in West- und Ostdeutschland

Dabei zeigte sich im Verlauf dieses Kapitels aber auch, dass manche Ergebnisse relativ deutlich auf die Annahmen aus den Kapiteln 6 und 8 zurückzuführen sind, wie beispielsweise die Angleichung der Verwandtschaftsstrukturen in West- und Ostdeutschland.

Inwieweit Schlussfolgerungen möglich sind, wird im nächsten Kapitel diskutiert. Zudem werden ergänzende Ergebnisse vorgestellt.

## 10 Zwei Anwendungsbeispiele

### 10.1 Mehr gemeinsame Lebenszeit von Generationen?

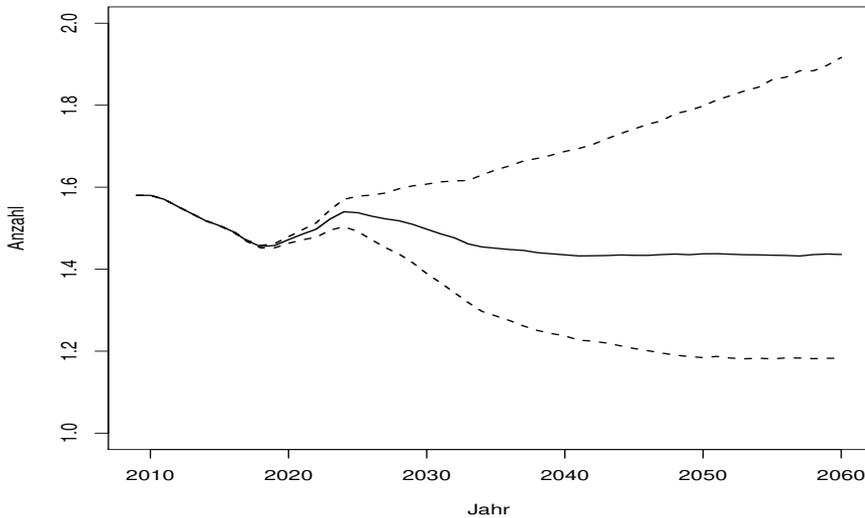
In Abschnitt 3.3 wurde die These der Vertikalisierung von Verwandtschaftsbeziehungen diskutiert. Kurz zusammengefasst besagt diese, dass aufgrund niedriger Fertilität die Zahl der Nachfahren und der Verwandtschaft in Seitenarmen abnimmt, zugleich aber aufgrund niedriger Mortalität die geteilte Lebenszeit von Generationen zunimmt. Im Folgenden werden die im vorausgegangenen Kapitel vorgestellten Resultate sowie einige zusätzliche Ergebnisse diskutiert, die eine genauere Einschätzung dieser These erlauben. Dabei werden die Entwicklungen bei Vor- und Nachfahren getrennt voneinander betrachtet und anschließend ein kurzes, zusammenfassendes Fazit gezogen.

#### 10.1.1 Nachfahren

Betrachtet man die im vorausgegangenen Kapitel beschriebenen Ergebnisse zur Zahl der Nachfahren, zeigt sich eine hohe Unsicherheit. Zwar scheint langfristig eine Verschiebung von Geburten in höhere Alter sehr wahrscheinlich zu sein, bezüglich des Umfangs der Fertilität lassen sich aber keine definitiven Aussagen treffen, so dass der erste Teil der Vertikalisierungsthese (weniger Nachfahren) zumindest langfristig sehr spekulativ ist und nur auf einen Teil der möglichen zukünftigen Entwicklungen abhebt. Dies gilt lediglich nicht für relativ hohe Lebensalter, bei denen die Zahl der Nachfahren stark durch die historische Entwicklung gekennzeichnet ist. Für diese Alter sind aber keine Spekulationen über eine mögliche Vertikalisierung nötig – wenn diese vorliegen sollte, wäre sie historisch beobachtet worden. Hier zeigen sich, wie in Abschnitt 3.3 bereits angesprochen, wechselhafte Ergebnisse.

Betrachtet man die Entwicklung über den Zeitraum von 2009 bis 2060, zeigt sich zudem, dass auch mittelfristig nicht unbedingt von einer weitergehenden Vertikalisierung auszugehen ist. In Abbildung 10.1 sieht man Median und Vorausberechnungsintervall der erwarteten Zahl lebender Kinder westdeutscher Frauen in Altersklasse 50 in den Jahren 2009 bis 2060 – also die Entwicklung der Kinderzahl am Ende der reproduktiven Phase. Für relativ weit in der Zukunft liegende Jahre der Vorausberechnung zeigt sich das bereits erwähnte hohe Ausmaß der Unsicherheit. Interessant ist die Entwicklung bis etwa zur Mitte der 2020er Jahre. Für 2009, dem letzten Jahr, für welches die erwartete Anzahl lebender Kinder über historische Raten berechnet werden kann, liegt diese für westdeutsche Frauen bei 1.58. In den folgenden Jahren sinkt die erwartete Kinderzahl bis 2018 auf einen Wert von 1.46 ab, wobei der Median der Vorausberechnungen und das 5%- beziehungsweise 95%-Quantil zunächst praktisch identisch sind und nur geringfügig voneinander abweichen. Bis 2024 steigt der Median der erwarteten Kinderzahl wieder an und erreicht einen Wert von 1.54. Das 5%-Quantil liegt bei 1.50 und das 95%-Quantil bei 1.57. Somit nimmt die Breite des Vorausberechnungsintervalls zwar zu und zeigt eine zunehmende Unsicherheit an, nichtsdestotrotz weisen auch die Intervallgrenzen auf einen Anstieg hin. Da die Zahl lebender Kinder in der Altersklasse von 50 Jahren kaum durch Mortalität beeinflusst wird, ist diese Zunahme auf einen Anstieg der Fertilität zurückzuführen.

Abbildung 10.1: Median (durchgezogene Linie) und Vorausberechnungsintervall (gestrichelte Linie) der Verteilung der erwarteten Anzahl lebender Kinder westdeutscher Frauen in Altersklasse 50, 2009 bis 2060



Quelle: eigene Berechnung

Dieser Anstieg lässt sich wie folgt erklären. Frauen, die 2024 in Altersklasse 50 sein werden, waren im Jahr 1974 in Altersklasse 0 und wurden im Jahr 1973 geboren. Zu Beginn des Jahres 2009 waren sie in Altersklasse 35 bei einer erwarteten Anzahl lebender Kinder von 1.26. Frauen, die im Jahr 2018 in der Altersklasse 50 sind, wurden im Jahr 1967 geboren. In der Altersklasse von 35 Jahren waren sie im Jahr 2003, mit einer erwarteten Anzahl lebender Kinder von 1.23, was bereits etwas unter dem Wert der Frauen des Jahrgangs 1973 liegt. Bei der vorausberechneten Fertilität wirkt sich zusätzlich der Anstieg des durchschnittlichen Alters bei der Geburt aus, der Frauen des Jahrgangs 1973 stärker trifft als Frauen des Jahrgangs 1967.

Gegeben dieser Resultate scheint die Annahme einer linearen Vertikalisierung ungerechtfertigt zu sein.

### 10.1.2 Vorfahren

Bereits in Abschnitt 3.3 wurde eine in der Literatur zu findende Kritik bezogen auf die Annahme eines Anstiegs der geteilten Lebenszeit von Generationen genannt: Einem Anstieg der Lebenserwartung kann ein Anstieg des Alters bei der Geburt entgegenwirken. Dies zeigte sich bei den im vorausgegangenen Kapitel vorgestellten Ergebnissen zu lebenden Vorfahren. Zumindest für die betrachteten Altersklassen (Altersklasse 50 bei Eltern; Altersklasse 0 bei Großeltern) lässt sich dieser gegenläufige Effekt klar erkennen. Insbesondere bezogen auf die geteilte Lebenszeit von Enkeln und Großeltern ist aufgrund der bisher vorgestellten Ergebnisse nur von einem sehr geringen Anstieg auszugehen.

Für eine ergänzende Betrachtung finden sich in Tabelle 10.1 zusätzliche Ergebnisse. In dieser ist eine grobe Abschätzung der erwarteten bisher geteilten Lebenszeit (in Jahren) mit einzelnen Großeltern zu sehen, die eine Person in Altersklasse 20 aufweist.

Dabei wird zum einen das Jahr 2029 und zum anderen das Jahr 2060 betrachtet.<sup>207</sup> Verglichen werden also die Geburtsjahrgänge 2008 und 2039. Für Personen, die 2029 in Altersklasse 20 sind, bezieht sich die geteilte Lebenszeit auf den Zeitraum von 2009 bis 2029 und für Personen, die 2060 in Altersklasse 20 sind, auf den Zeitraum vom Jahr 2040 bis zum Jahr 2060.<sup>208</sup> Wie bei den bisherigen Ergebnissen auch ist der Median der 1000 Simulationen ausgewiesen. Zudem ist noch die Intervallbreite zu sehen. Beispielsweise beträgt der Median der erwarteten, bisher geteilten Lebenszeit für Altersklasse 20, das Jahr 2029 sowie die Großmutter mütterlicherseits in Westdeutschland 16.83 Jahre und für Ostdeutschland 17.87 Jahre. Eine 20-jährige Person hat im Jahr 2029 in Ostdeutschland also tendenziell ein Jahr länger zeitgleich mit der Großmutter mütterlicherseits zusammengelebt als in Westdeutschland.

Zusätzlich ist in Tabelle 10.1 noch die Differenz der Mediane der Jahre 2060 und 2029 eingezeichnet. Orientiert man sich für einen Vergleich der Jahre 2029 und 2060 (bzw. Jahrgänge 2008 und 2039) an dieser, zeigen sich für Westdeutschland eher geringe Differenzen, wobei die Werte des Jahres 2060 leicht höher liegen als für 2029. Bei Ostdeutschland sind die Unterschiede nochmals geringer und für Großeltern mütterlicherseits sind die Werte für 2060 sogar leicht niedriger als für 2029. Dies stützt die eingangs getroffene Einschätzung, dass eine Zunahme der geteilten Lebenszeit von Enkeln und Großeltern nur in geringem Ausmaß beziehungsweise nicht zu erwarten ist.<sup>209</sup>

### 10.1.3 Fazit

Insgesamt lässt sich festhalten, dass ausgehend von den hier vorgestellten Ergebnissen in Zukunft nicht von einer Vertikalisierung beziehungsweise weitergehenden Vertikalisierung ausgegangen werden kann. Bei Nachfahren lässt sich aufgrund der hohen Unsicherheit der zukünftigen Entwicklung keine definitive Aussage treffen. Die Ergebnisse für Vorfahren weisen darauf hin, dass der positive Effekt der steigenden Lebenserwartung auf die geteilte

207 Für Personen, die 2029 in Altersklasse 20 sind, wurden ab dem Jahr 2009 und ab Altersklasse 0 alle altersspezifischen Wahrscheinlichkeiten betrachtet, dass ein Großelternanteil noch lebt. Sei  $p_x$  eine solche Wahrscheinlichkeit für eine Altersklasse  $x$ . Dann wurde die erwartete bisher geteilte Lebenszeit berechnet als  $\sum_1^{20} p_x + (p_x - p_{x-1})0.5$ . Für Großeltern, die zwischen zwei Jahren  $t$  und  $t + 1$  beziehungsweise Altersklassen  $x$  und  $x + 1$  sterben, wurde also angenommen, dass sie durchschnittlich ein halbes Jahr zwischen  $t$  und  $t + 1$  leben. Die geteilte Lebenszeit im Geburtsjahr der betrachteten Personen wird bei diesem Vorgehen nicht berücksichtigt. Für Personen in Altersklasse 20 im Jahr 2060 wurde analog vorgegangen.

208 Die Altersklasse von 20 Jahren wurde gewählt, um zwei relativ weit auseinander liegende Geburtsjahrgänge betrachten zu können. Denn für das Jahr 2009 liegen bei Großeltern lediglich Ergebnisse für die Altersklassen von 0 bis 5 Jahren vor (s. Abschnitt 7.1), bei denen eine Betrachtung der geteilten Lebenszeit eher uninteressant sein dürfte. Für höhere Altersklassen müssen somit aber weiter in der Zukunft liegende Jahre betrachtet werden. Dabei ergibt sich dann umgekehrt das Problem, dass keine hohen Alter betrachtet werden können, bei denen davon ausgegangen werden kann, dass alle Großeltern tot sind und die bisherige geteilte Lebenszeit der gesamten Lebenszeit mit den Großeltern entspricht.

209 Verwendet man andere Altersklassen, zeigen sich sehr ähnliche Ergebnisse. Vergleicht man beispielsweise westdeutsche Personen in Altersklasse 40 in den Jahren 2049 und 2060 miteinander, betragen die Differenzen der Mediane 0.44 (Großmutter mütterlicherseits) und 0.43 (Großvater väterlicherseits).

Tabelle 10.1: Geteilte Lebenszeit mit einzelnen Großeltern von Altersklasse 0 bis Altersklasse 20

	2009 bis 2029		2040 bis 2060		Differenz (2)-(1)
	Median (1)	Intervallbreite	Median (2)	Intervallbreite	
West					
Großmutter mütterlicherseits	16.83	0.06	17.19	0.32	0.36
Großvater mütterlicherseits	13.08	0.11	13.53	0.92	0.45
Großmutter väterlicherseits	15.73	0.08	16.29	0.36	0.57
Großvater väterlicherseits	11.50	0.12	12.16	0.92	0.67
Ost					
Großmutter mütterlicherseits	17.87	0.04	17.65	0.26	-0.22
Großvater mütterlicherseits	14.36	0.09	14.16	0.85	-0.20
Großmutter väterlicherseits	16.90	0.06	16.92	0.31	0.02
Großvater väterlicherseits	12.75	0.10	12.87	0.88	0.12

Quelle: eigene Berechnung

Lebenszeit durch den negativen Effekt des Anstiegs des durchschnittlichen Alters bei der Geburt zumindest bezogen auf Enkel und Großeltern größtenteils aufgehoben wird.

## **10.2 Pflege und Verwandtschaft: Entwicklung verwandtschaftlicher Unterstützungspotenziale am Beispiel von Kindern**

Für das in Unterabschnitt 2.2.2 besprochene Beispiel der Pflege sind Veränderungen der Zahl der Verwandten unter anderem in Hinblick auf das verwandtschaftliche Unterstützungspotenzial von Pflegebedürftigen interessant. Hierzu gibt es in der Literatur etliche Modellrechnungen, deren Fokus vor allem auf dem Vorhandensein beziehungsweise der Zahl von Ehepartnern und Kindern als „informellen Unterstützungspotenzial“ liegt (vgl. Blinkert & Klie 2004; Blinkert & Gräf 2009; Doblhammer & Ziegler 2010; Hackmann & Moog 2008; Hofmann 2006; Schulz 2010). Diese werden im folgenden Unterabschnitt kurz diskutiert. Anschließend wird eine alternative Berechnung vorgestellt. Ergänzende Bemerkungen und ein Fazit schließen diesen Abschnitt ab.

### *10.2.1 Die Vorausberechnung verwandtschaftlicher Unterstützungspotenziale in der Literatur*

In der Literatur wird bei Modellrechnungen zu verwandtschaftlichen Unterstützungspotenzialen in aller Regel zunächst eine Vorausberechnung der Zahl der Pflegebedürftigen vorgenommen (vgl. bspw. Blinkert & Gräf 2009: 9ff.). Anschließend wird die Zahl potentieller Pflegepersonen vorausberechnet und das Verhältnis dieser beiden Größen als Indikator für die Entwicklung des Unterstützungspotenzials verwendet (vgl. bspw. Schulz 2010: 77). Dabei erfolgt die Berechnung der Zahl potenzieller Pflegepersonen in aller Regel sehr undifferenziert. Beispielsweise berechnen Blinkert & Gräf (2009: 15f.) das Pflegepotenzial von Kindern von Pflegebedürftigen als die Summe aller nicht erwerbstätigen Personen im Alter von 30 bis 69 Jahren, womit auch eine Vorausberechnung der Erwerbstätigkeit nötig ist. Andere Vorausberechnungen berücksichtigen teils noch zusätzliche Einflussfaktoren (vgl. bspw. Hackmann & Moog 2008), allerdings folgt die Berechnung der Unterstützungspotenziale von Kindern bei den genannten Publikationen von einer Ausnahme abgesehen (s.u.) dieser Logik, während die Vorausberechnung des Vorhandenseins und des Unterstützungspotenzials von Ehepartnern teils differenzierter erfolgt (vgl. Doblhammer & Ziegler 2010; Schulz 2010).

Diese Berechnung des Unterstützungspotenzials von Kindern ist aus mehreren Gründen problematisch. Ein wesentliches Problem ist, dass lediglich die Größe zweier Bevölkerungsgruppen miteinander verglichen wird – inwiefern pflegebedürftige Personen zu 30- bis 69-jährigen nicht erwerbstätigen Personen in irgendeinem Verhältnis stehen, wird nicht betrachtet. Ein zweites Problem ist, dass neben der zukünftigen Entwicklung der demographischen Komponenten noch die zukünftige Entwicklung weiterer Einflussfaktoren wie Pflege- und Erwerbsquoten berücksichtigt werden muss. Dies führt zu einer erhöhten Unsicherheit der Vorausberechnung, wenn diese zusätzlichen Faktoren einen maßgeblichen Einfluss auf die

Resultate haben. Insbesondere wenn primär der Einfluss der demographischen Entwicklung interessiert, scheint dies unnötig. Schließlich ist bei den bisher genannten Arbeiten problematisch, dass alle auf Vorausberechnungsszenarien basieren und Unsicherheit nicht berücksichtigen. Die zukünftige Entwicklung der Bevölkerung wird beispielsweise teils aus den offiziellen Vorausberechnungen des Statistischen Bundesamtes übernommen, wobei wiederum in aller Regel nur ein oder zwei der vom Statistischen Bundesamt berechneten Entwicklungsvarianten berücksichtigt werden (dies gilt für Blinkert & Gräf 2009; Hackmann & Moog 2008; Hofmann 2006). Dabei sind aber zumindest die Ergebnisse der letzten amtlichen Vorausberechnung als eher pessimistisch einzustufen (s. Unterabschnitt 8.6.3).

### 10.2.2 Alternative Berechnungen der Unterstützungspotenziale durch Kinder

Hier soll im weiteren eine alternative Berechnung des Unterstützungspotenzials von Kindern vorgestellt werden. Um einen Vergleich zu ermöglichen, finden sich in Abbildung 10.2 zunächst die Ergebnisse einer Berechnung des Unterstützungspotenzials, die wie im vorausgegangenen Unterabschnitt beschrieben durchgeführt wurde. Zunächst wurde für jede der 1000 Simulationen der Bevölkerungsentwicklung die Zahl der Pflegebedürftigen berechnet. Hierfür wurde davon ausgegangen, dass die amtlichen Pflegequoten des Jahres 2009 konstant bleiben.<sup>210</sup> Diese wurden dann jeweils auf die gesamtdeutsche Bevölkerung angewendet, wobei Pflegequoten für Alter unter 60 Jahren nicht berücksichtigt wurden, so dass der Fokus auf älteren Pflegebedürftigen liegt. Das Unterstützungspotenzial der Kinder wurde dann abgeschätzt als die Zahl der Personen in Altersklasse 20 bis Altersklasse 59.<sup>211</sup> Werden andere Altersgrenzen verwendet, verändern sich zwar die absoluten Ergebnisse, die relative Entwicklung sieht aber zumeist sehr ähnlich aus: bei Betrachtung des Medians der Entwicklung zeigt sich bis zur Mitte der 2050er Jahre ein mehr oder weniger linearer Anstieg des Verhältnisses. Der Median für das Jahr 2060 entspricht etwa dem 3.4-fachen des Wertes für das Jahr 2009. Die Vorausberechnungsintervalle sind für einen Großteil des Vorausberechnungszeitraums recht schmal. Diese Resultate verweisen insgesamt auf eine deutliche Zunahme der „Belastung“ von Unterstützungspotenzialen.

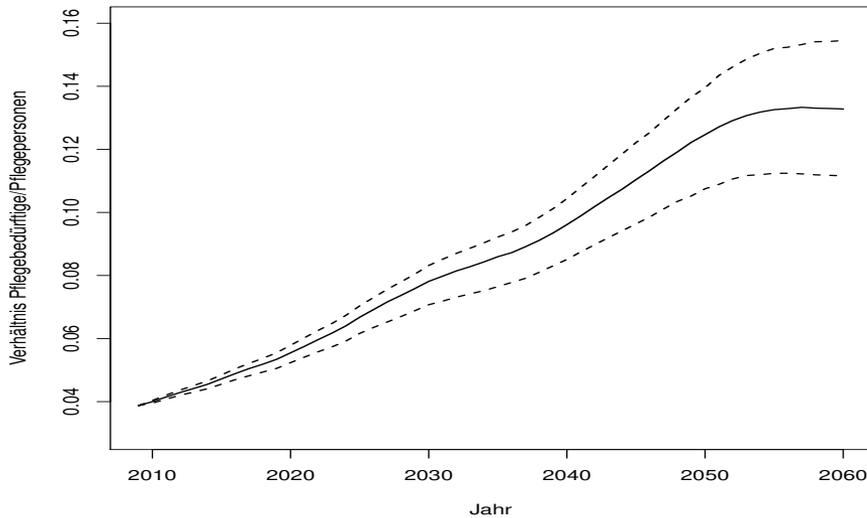
Als Alternative zu diesen Berechnungen soll hier der tatsächliche Verhältniswert von Pflegebedürftigen zu ihren Kindern betrachtet werden. Sei  $n_{x,P}$  die Zahl der Pflegebedürftigen in einer Altersklasse  $x$ . Die erwartete Zahl an Kindern dieser Personen ergibt sich unter den Annahmen des in dieser Arbeit verwendeten Modells einfach als Erwartungswert der Kinderzahl einer Person multipliziert mit der Zahl der betrachteten Personen – also aus dem Produkt von  $n_{x,P}$  mit einer erwarteten Anzahl an Kindern, wie sie in Abschnitt 9.2 beschrieben wurde. Berechnet man das Verhältnis der Zahl der Pflegebedürftigen und der Zahl ihrer Kinder und nimmt an, dass Pflegebedürftigkeit und Kinderzahl unabhängig voneinander sind, ergibt sich dieses Verhältnis einfach als Kehrwert der erwarteten Kinderzahl:

$$\frac{q_x n_x}{q_x n_x E(K|X = x)} = \frac{1}{E(K|X = x)} \quad (10.1)$$

210 Diese Pflegequoten wurden übernommen aus Statistisches Bundesamt (2011b).

211 Variationen dieser Altersabgrenzung – bspw. die Verwendung der Altersklassen von 20 bis 64 – führen im Wesentlichen zu denselben Ergebnissen.

Abbildung 10.2: Median (durchgezogene Linie) und Vorausberechnungsintervall (gestrichelte Linie) des Verhältnisses der Zahl der Pflegebedürftigen zur Zahl der potenziellen Pflegepersonen von 2009 bis 2060



Quelle: eigene Berechnung

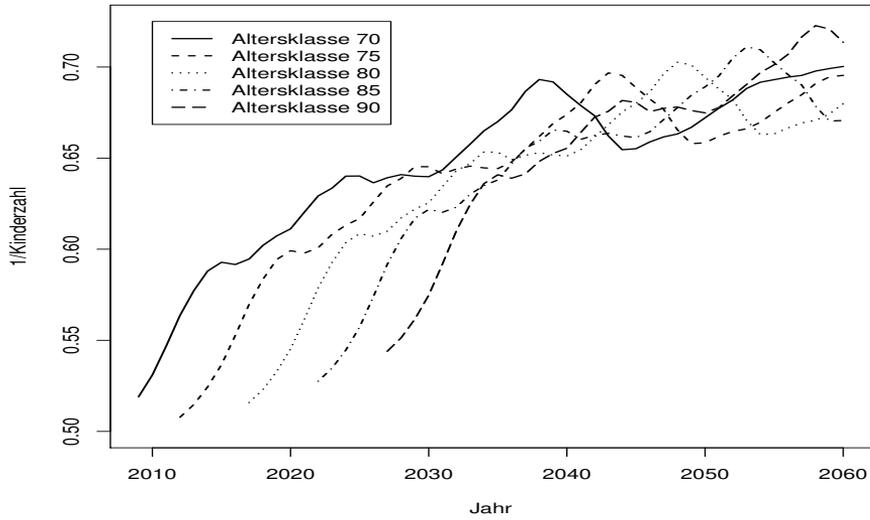
wobei  $q_x$  die Pflegequote für Altersklasse  $x$  ist,  $n_x$  der Bevölkerungsbestand in Altersklasse  $x$  und  $E(K|X = x)$  die erwartete Kinderzahl in Altersklasse  $x$ . Wenn die Annahme der Unabhängigkeit von Pflegebedürftigkeit im Alter und Kinderzahl zutreffend ist, hat dieser Kennwert den Vorteil, nicht von Annahmen über die zukünftige Entwicklung der Pflegequoten abzuhängen.

Der gerade vorgestellte Kennwert kann prinzipiell für beliebige Altersklassen, Männer und Frauen sowie West- und Ostdeutschland betrachtet werden. Hier werden im weiteren lediglich kurz Ergebnisse für west- und ostdeutsche Frauen für die Altersklassen von 70, 75, 80, 85 und 90 vorgestellt. Dabei ist zu bedenken, dass entsprechend der Ausführungen in Abschnitt 7.1 je nach Altersklasse Ergebnisse nicht für alle Vorausberechnungsjahre vorliegen.

Ergebnisse für Westdeutschland sind in Abbildung 10.3 zu sehen, Ergebnisse für Ostdeutschland in Abbildung 10.4. Dabei ist für die einzelnen Altersklassen lediglich der Median der Entwicklung eingezeichnet, um die Darstellung übersichtlich zu halten.

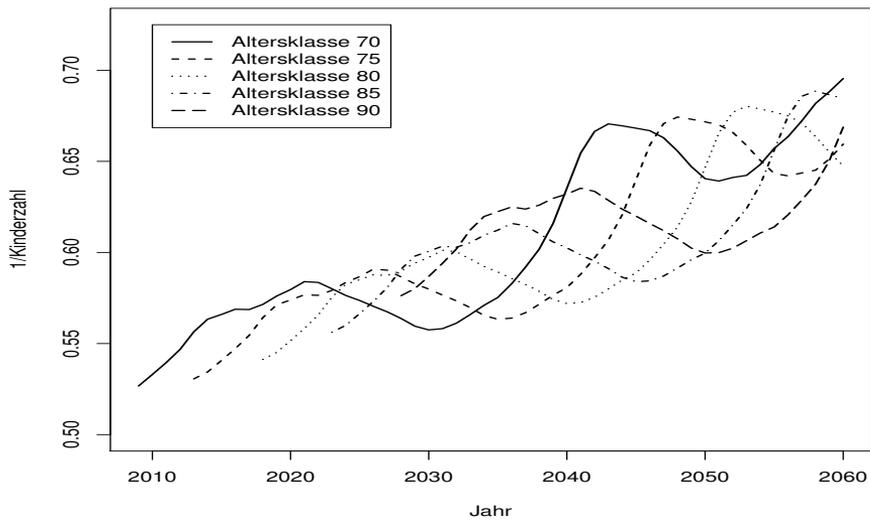
In beiden Abbildungen ist zu sehen, dass der Verlauf für Altersklasse 75 in etwa dem Verlauf für Altersklasse 70 um 5 Jahre verschoben entspricht. Entsprechendes gilt auch für die anderen Kurven (s. auch Abschnitt 9.2). Für Westdeutschland zeigt sich entsprechend bei allen Altersklassen zunächst ein relativ deutlicher Anstieg, der allerdings schnell abflacht. Die Ergebnisse für Ostdeutschland weisen zunächst auf einen leichten Anstieg, gefolgt von einer leichten Abnahme hin. Anschließend folgt eine deutliche Zunahme. Diese dürfte auf die Entwicklung der Fertilität in Ostdeutschland seit etwa 1980 zurückzuführen sein – ab 1980 nahmen die Geburtenraten in Ostdeutschland tendenziell ab und zu Beginn der 1990er Jahre folgte ein drastischer Einbruch der Geburtenraten (s. Abschnitt 3.2).

Abbildung 10.3: Median der erwarteten Kinderzahl westdeutscher Frauen in den Altersklassen 70, 75, 80, 85 und 90 von 2009 bis 2060



Quelle: eigene Berechnung

Abbildung 10.4: Median der erwarteten Kinderzahl ostdeutscher Frauen in den Altersklassen 70, 75, 80, 85 und 90 von 2009 bis 2060



Quelle: eigene Berechnung

Die Breite der Vorausberechnungsintervalle hängt deutlich von der betrachteten Altersklasse ab, wobei aber bei allen Altersklassen die höchste Unsicherheit für das Jahr 2060 zu verzeichnen ist. Für die Altersklasse von 90 Jahren beträgt die Intervallbreite sowohl für West- als auch für Ostdeutschland 0.02, während für Altersklasse 70 jeweils ein Wert von 0.22 erreicht wird.

Interessant ist hier aber vor allem die relative Entwicklung. Bei der Entwicklung des Verhältnisses der Pflegebedürftigen zu potentiellen Pflegepersonen über Bevölkerungsbestände, die in Abbildung 10.2 abgebildet ist, ergab sich wie erwähnt ein Anstieg von 2009 bis 2060 auf das 3.4-fache beziehungsweise um 240 Prozentpunkte. Betrachtet man hingegen die Ergebnisse in den Abbildungen 10.3 und 10.4, zeigt sich ein anderes Bild. Da nicht für alle Altersklassen ein Wert für das Jahr 2009 berechnet werden kann, wird der Wert der Altersklasse 70 für dieses Jahr als Maßstab genommen. Wird dieser Wert mit allen anderen Medianwerten verglichen, ergibt sich für Westdeutschland der höchste relative Anstieg bei Altersklasse 90 und dem Jahr 2058. Der entsprechende Wert liegt um etwa 39 Prozentpunkte höher als der des Ausgangsjahres. Für Ostdeutschland beträgt der höchste Anstieg 32 Prozentpunkte und tritt bei Altersklasse 70 und dem Jahr 2060 auf. Vergleicht man nicht die Medianwerte, sondern die oberen Grenzen des Vorausberechnungsintervalls mit dem Ausgangswert für 2009 findet man Anstiege von etwa 57 Prozentpunkten (West) und 54 Prozentpunkten. Vergleicht man diese Resultate mit dem oben genannten Anstieg des einfachen Indikators um 240 Prozentpunkte, kommt man zu einer wesentlich anderen Einschätzung der Entwicklung der „Belastung“ des Unterstützungspotenzials.<sup>212</sup> Zwar zeigt sich ebenfalls ein recht deutlicher Anstieg, allerdings fällt dieser bei weitem nicht so extrem aus. Zudem ist diese Einschätzung zumindest für einige der betrachteten Altersklassen wie erwähnt mit einer eher geringen Unsicherheit behaftet, so dass ein extremer Anstieg eher unwahrscheinlich wirkt.

### 10.2.3 Ergänzung und Fazit

Bei den gerade vorgestellten Ergebnissen ist zu bedenken, dass bezogen auf das Unterstützungspotenzial der Anteil kinderloser Personen als bedeutsamer gilt als die durchschnittliche Kinderzahl (vgl. Uhlenberg 1995: 21).<sup>213</sup> Eine Betrachtung der mittelfristigen Entwicklung des Anteils älterer, kinderloser Pflegebedürftiger liefern Doblhammer & Ziegler (2010). Sie nutzen aus, dass für viele der zukünftig alten Menschen die Kinderzahl bereits feststeht. Beispielsweise lässt sich für Personen, die im Jahr 2030 im Alter von 80 Jahren sind, die Kinderzahl bereits heute ermitteln. Da diese für die meisten Alter im Wesentlichen kaum durch Mortalität beeinflusst wird (s. Abschnitt 9.2), kann davon ausgegangen werden, dass sie sich bis 2030 kaum verändern wird (vgl. ebd.: 42). Dies nutzen sie in Kombination mit einer Vorausberechnung der Pflegebedürftigkeit und des Familienstands, um Unterstützungspotenziale im Jahr 2030 einzuschätzen. Dabei kommen sie zu folgendem Schluss: „In future, a larger proportion of the elderly will be married with at least one child than is the case now and has ever been experienced in recent history.“ (ebd.: 56). Betrachtet man dies in Kombination mit den zuvor

212 Betrachtet man die mittelfristige Entwicklung bis 2030 zeigt sich dasselbe Bild. Die einfache Abschätzung des Unterstützungspotenzials weist auf einen Anstieg der Belastung um etwa 96% hin. Für westdeutsche Frauen ergibt sich beim Indikator (10.1) der höchste Anstieg bei Altersklasse 75, welcher bei 24% liegt.

213 Wie am Ende von Kapitel 5 erwähnt, kann paritätsspezifische Fertilität und darüber der Anteil kinderloser Personen aufgrund fehlender Daten nicht im hier verwendeten Modell berücksichtigt werden.

vorgestellten Ergebnissen, scheint der „demographische Druck“ auf Unterstützungspotenziale zumindest mittelfristig eher mäßig auszufallen.

Zu beachten ist bei dieser Einschätzung, dass sie sich allein auf die demographische Entwicklung bezieht. Weitere mögliche Einflussfaktoren, wie die weiter oben bereits erwähnte Entwicklung der Erwerbstätigkeit, werden bei dieser Betrachtung nicht berücksichtigt und können nochmals deutliche Effekte auf das Unterstützungspotenzial haben. Für eine weitergehende Interpretation der vorgestellten Resultate müsste man annehmen, dass alle anderen Einflussfaktoren konstant bleiben, was mehr als unrealistisch erscheint. Insofern sollten die vorliegenden Ergebnisse nicht überbewertet werden.

### 10.3 Bedingungen für die Gültigkeit der Ergebnisse

Zu Beginn von Kapitel 4 wurden drei potentielle Fehlerquellen bei demographischen Vorausberechnungen benannt:

1. Unsicherheit bezüglich der zukünftigen Entwicklung der demographischen Komponenten und gegebenenfalls weiterer Parameter
2. Unsicherheit von Ausgangsgrößen
3. Unsicherheit bezüglich der den verwendeten Verfahren zugrunde liegenden Annahmen

Wenn für alle dieser drei Fehlerquellen davon ausgegangen werden kann, dass sie die im vorausgegangenen Kapitel und den vorausgegangenen Abschnitten vorgestellten Ergebnisse nicht beeinflussen beziehungsweise dass sie hinreichend berücksichtigt wurden, können die bisher vorgestellten Resultate als eine realistische Beschreibung der zukünftigen Entwicklungsmöglichkeiten aufgefasst und sinnvoll interpretiert werden. Allerdings kann dies trotz der Auswertungen und Überlegungen der letzten Kapitel nicht abschließend beurteilt werden – inwieweit beispielsweise die zukünftige demographische Entwicklung bis zu einem Jahr  $t$  „richtig“ modelliert wurde, wird erst im Jahr  $t$  selbst festgestellt werden können. Nichtsdestotrotz können allgemeine Bedingungen genannt werden, die Voraussetzung für die „Richtigkeit“ der Ergebnisse sind. Etliche von diesen wurden bereits im Laufe der Arbeit genannt und sollen hier nur kurz zusammengefasst werden, während einige weitere Bedingungen ergänzt werden. Dabei wird nicht auf die den Vorausberechnungen zugrunde liegenden Daten eingegangen, die in den entsprechenden Kapiteln beziehungsweise in Anhang A diskutiert werden.

Bezogen auf die in Kapitel 8 vorgenommenen Vorausberechnungen zur Berücksichtigung der Unsicherheit der demographischen Komponenten müssen etliche Voraussetzungen erfüllt sein. Eine zentrale Bedingung ist, dass die Entwicklung einer demographischen Komponente in der Vergangenheit auch Aufschluss über die zukünftige Entwicklung geben kann – die demographischen Komponenten müssen sich also „wie bisher“ entwickeln (vgl. Lee 1998: 158f.). Hiervon ausgehend muss der nichtparametrische Ansatz, der in dieser Arbeit zur Modellierung der Komponenten verwendet wird, die historische Entwicklung adäquat abbilden können (vgl. Cairns et al. 2009: 2). Dies umfasst auch die Wahl geeigneter Modellparameter (vgl. de Beer 2000: 23). Die in Kapitel 8 verwendeten Parameter liefern zwar

eine recht plausible Beschreibung der historischen Entwicklung der Komponenten, nichtsdestotrotz ist die Verwendung von anderen Parameterwerten, die zumindest bei langfristigen Vorausberechnungen zu anderen Ergebnissen führen könnten, nicht gänzlich ausgeschlossen. Insofern müssen die gewählten Werte als „richtig“ oder zumindest als näherungsweise „richtig“ gelten können. Ähnliches muss auch für weitere vorgenommene Anpassungen gelten, wie beispielsweise die Glättung von Geburtenraten über Spline-Funktionen, die die Ergebnisse nicht verfälschen dürfen. Schließlich wurden etliche ergänzende Annahmen getroffen, wie beispielsweise eine Unter- und Obergrenze der zusammengefassten Geburtenziffer. Dabei basiert insbesondere die Vorausberechnung der Komponenten für Ostdeutschland stark auf solchen zusätzlichen Annahmen. Auch diese müssen zutreffen, um zu einer Interpretation der Ergebnisse zu kommen.

Die Annahmen des verwendeten Modells wurden bereits am Ende von Kapitel 5 beschrieben. Für eine sinnvolle Interpretation der Ergebnisse darf eine Verletzung der Annahmen – von der auszugehen ist – keinen übermäßig verzerrenden Einfluss auf die Ergebnisse haben. Dies wurde in Kapitel 7 überprüft. Bezüglich der gefundenen Abweichungen zwischen Modellrechnung und empirischen Daten muss angenommen werden, dass diese vernachlässigt werden können und keine Korrektur der Ergebnisse, wie in Anhang F beschrieben, notwendig ist. Wenn man den Vergleich von Modellrechnung und empirischen Daten entsprechend deutet, muss für die Interpretierbarkeit der Ergebnisse der Vorausberechnung ferner gelten, dass sich die Verletzung der Modellannahmen in Zukunft nicht stärker auswirkt.

Aus Kapitel 5 ergibt sich des Weiteren auch das Vorgehen zur Vorausberechnung der Bevölkerung sowie die Handhabung von  $\phi(y|x)$ . Ersteres entspricht zwar einer Bevölkerungsfortschreibung nach der Buchführungsgleichung (s. Abschnitt 3.1), allerdings sind nichtsdestotrotz einige Annahmen notwendig, wie beispielsweise bei der Handhabung von Migration (s. Unterabschnitt 8.6.1). Ausgehend von den „richtigen“ demographischen Raten sollte eine Verletzung dieser Annahmen aber nur einen vernachlässigbaren Effekt haben.

Bezüglich der konditionalen Verteilungen  $\phi(y|x)$  wird in Kapitel 5 angenommen, dass sie konstant sind. Zumindest über kürzere Zeiträume scheint dies historisch gegeben zu sein (s. Kapitel 6). Ferner wird der sich aus den verwendeten Verteilungen ergebende durchschnittliche Altersabstand zwischen Männern und Frauen bei der Geburt eines Kindes von etwa drei Jahren durch die Literatur gestützt (vgl. Coleman 2000: 51; Dinkel & Milenovic 1992: 65; Dinkel & Milenovic 1993: 148; Huinink & Tuma 1987: 514; Huinink 1988: 361). Insofern ist diese Annahme für die zukünftige Entwicklung zwar nicht unplausibel, nichtsdestotrotz könnte eine Veränderung der Verteilungen  $\phi(y|x)$  durchaus zu deutlich anderen Ergebnissen führen, weshalb die Richtigkeit der eingangs genannten Annahme vorausgesetzt werden muss. Für die vergangene Entwicklung seit den fünfziger Jahren scheint diese Annahme aber wie in Kapitel 6 gezeigt nur in Teilen angemessen zu sein. Dies dürfte zumindest für weiter in der Zukunft liegende Jahre der Vorausberechnung aber keinen oder nur einen geringen Einfluss auf die Ergebnisse haben.

Insgesamt zeigt sich also, dass etliche Bedingungen nötig sind, damit die Ergebnisse sinnvoll interpretiert werden können. Die im vorausgegangenen Kapitel und in den vorausgegangenen Abschnitten dargestellten Schlussfolgerungen basieren auf der Annahme, dass die Bedingungen alle erfüllt sind.



## **11 Zusammenfassung, Diskussion und Ausblick**

### **11.1 Ziel und Motivation**

Das Ziel der vorliegenden Arbeit ist die Vorausberechnung des Vorhandenseins linearer Verwandtschaft. Unter dem Begriff „Verwandtschaft“ wird eine Verbindung von Personen entweder über gemeinsame Abstammung oder über Heirat verstanden. Die „lineare Verwandtschaft“ einer Person umfasst direkte Vor- und Nachfahren, wobei Großeltern, Eltern, Kinder und Enkelkinder betrachtet werden. Das „Vorhandensein“ wird aus einer statistischen Perspektive beleuchtet und meint beispielsweise die durchschnittliche Kinderzahl oder die Wahrscheinlichkeit, dass die Großeltern einer Person noch leben. Die Vorausberechnung erfolgt bis einschließlich zum Jahr 2060.

Inhaltliche Motivation für die Vorausberechnung linearer Verwandtschaft ist zum einen der mögliche Einfluss des demographischen Wandels auf deren Vorhandensein und zum anderen die besondere gesellschaftliche Bedeutung von Verwandtschaft. Bezogen auf den Effekt des demographischen Wandels auf das Vorhandensein von Verwandten findet sich in der Literatur beispielsweise die Überlegung, dass ein zukünftiger Anstieg der Lebenserwartung dazu führen wird, dass die gemeinsame Lebenszeit von Enkeln und Großeltern steigt. Allerdings ist dieser Zusammenhang nicht so eindeutig, wie es auf den ersten Blick scheint. Denn ein Anstieg der Lebenserwartung würde nur bei ansonsten gleichen Bedingungen zu einem Anstieg der geteilten Lebenszeit führen. Würde aber zugleich auch das Alter bei der Geburt steigen, würde hiervon ein gegenläufiger Effekt ausgehen. Welcher der Effekte überwiegt, lässt sich nicht ohne weiteres bestimmen. Ausgehend von einer Vorausberechnung linearer Verwandtschaft können hingegen konkretere Angaben über zukünftige Entwicklungstendenzen gemacht werden.

Ein Beispiel für die Bedeutung von Verwandtschaft ist der Bereich der Pflege pflegebedürftiger Menschen. So wurden Ende 2011 laut amtlicher Pflegestatistik von den etwa 2.5 Millionen Pflegebedürftigen rund 1.2 Millionen ausschließlich durch Angehörige versorgt. Zwar bilden lineare Verwandte wie insbesondere Kinder nur eine Teilgruppe der pflegenden Angehörigen, allerdings eine vergleichsweise große. Eine Vorausberechnung linearer Verwandtschaft erlaubt dann eine Abschätzung, inwieweit sich verwandtschaftliche Unterstützungspotentiale in Zukunft ändern werden und ob beispielsweise von einer steigenden Nachfrage nach professioneller Pflege auszugehen ist.

### **11.2 Methodisches Vorgehen**

Ausgehend von dieser Problemstellung wird ein eigener Vorausberechnungsansatz entwickelt, der verglichen mit bisher in der Literatur verwendeten Ansätzen einen geringen Datenbedarf aufweist und mit Verfahren zur Berücksichtigung demographischer Unsicherheiten kombiniert werden kann (s.u.). Dabei folgt das Vorgehen einer ähnlichen Logik wie bei Bevölkerungsvorausberechnungen nach der Kohorten-Komponenten-Methode und basiert im Wesentlichen auf demographischen Raten zu Mortalität und Fertilität.

Zur Ermittlung der Nachfahren von Männern werden Geburtenraten für diese benötigt, die im Gegensatz zu den anderen für die Berechnungen notwendigen Daten nicht durch die amtliche Statistik zur Verfügung gestellt werden. Um diese Geburtenraten zu ermitteln, wird auf die Statistik der Geburten zurückgegriffen, bei der es sich um eine Vollerhebung aller standesamtlich registrierten Geburten in der BRD handelt. Für die Auswertungen stehen die Daten der Jahre 1995, 2000, 2006, 2007 und 2008 zur Verfügung, auf die ein einfacher Ansatz zur Rekonstruktion der historischen Entwicklung der Fertilität von Männern angewendet wird. Berechnet werden altersspezifische Geburtenraten für Männer für den Zeitraum von 1951 bis 2008. Die Ergebnisse deuten darauf hin, dass der Rekonstruktionsansatz zwar grundsätzlich eine Ermittlung der Fertilität von Männern erlaubt, allerdings einige der Resultate für historisch weiter zurückliegende Zeiträume mit Vorsicht zu betrachten sind.

Da der Vorausberechnungsansatz von einigen vereinfachenden Modellannahmen ausgeht, wird eine Modellvalidierung durchgeführt, bei der zunächst vorliegende historische Daten in Verbindung mit dem Berechnungsansatz genutzt werden, um das Vorhandensein von Verwandtschaft für das Jahr 2009 zu ermitteln. Diese Ergebnisse werden mit Daten des Sozio-oekonomischen Panels (SOEP) verglichen, wobei der Vergleich eine Einschätzung erlaubt, ob die Resultate des Ansatzes trotz der vereinfachenden Annahmen realistische Ergebnisse liefern. Bezogen auf Nachfahren zeigen sich keine systematischen Unterschiede zwischen SOEP und Modellrechnung, bei Vorfahren hingegen schon. Allerdings ist unklar, ob diese Abweichungen auf zu stark vereinfachende Annahmen des verwendeten Ansatzes zurückzuführen sind und es gibt gute Gründe, die gegen die Verlässlichkeit der Ergebnisse des SOEP sprechen, zumindest im Hinblick auf Vorfahren. Aufgrund dieser Unklarheit werden in der Arbeit sowohl die eigentlichen Ergebnisse des Modells dargestellt als auch zusätzlich ein Korrekturverfahren vorgeschlagen und angewendet.

Die weiter oben bereits angedeutete Berücksichtigung der Unsicherheit der zukünftigen demographischen Entwicklung erfolgt, indem die für die Vorausberechnung benötigten demographischen Größen stochastisch fortgeschrieben werden. Hierzu wird ein nichtparametrischer Resampling-Ansatz verwendet. Dies erlaubt es, für interessierende Größen einen Ergebnisbereich anzugeben, in dem die zukünftige Entwicklung mit großer Wahrscheinlichkeit liegen wird. Hierfür werden insgesamt 1000 mögliche zukünftige Entwicklungsverläufe zufällig erzeugt. Im Anschluss kann dann beispielsweise angegeben werden, in welchem Bereich 90% der Ergebnisse liegen oder auch der Median über die 1000 Simulationen als eine Art mittlerer Tendenz der Entwicklung.

## **11.3 Zentrale Ergebnisse**

### *11.3.1 Bevölkerungsentwicklung*

Der stochastische Vorausberechnungsansatz erlaubt es zunächst, eine Bevölkerungsvorausberechnung vorzunehmen. Zentrale Ergebnisse finden sich in Tabelle 11.1. Zu sehen sind der Bevölkerungsbestand und der Altenquotient, jeweils für die Jahre 2009, 2030 und 2060. Die Werte für das Jahr 2009 sind die Ausgangswerte der Vorausberechnung. Für die Jahre 2030

Table 11.1: Zentrale Ergebnisse der stochastischen Bevölkerungsvorausberechnung

Größe	Kennwert	2009	2030	2060
Bevölkerungsbestand (Mio.)	5%-Quantil	—	77.70	65.41
	Median/Ausgangswert	81.64	80.95	72.73
	95%-Quantil	—	84.53	82.15
Altenquotient	5%-Quantil	—	45.93	56.15
	Median/Ausgangswert	32.06	48.75	63.67
	95%-Quantil	—	51.02	71.31

Quelle: eigene Berechnung

und 2060 ist zum einen der Median der 1000 Simulationen eingetragen, sowie die Unter- und Obergrenze des Bereichs, der 90% aller Simulationen umfasst (5%- und 95%-Quantil).

Die Entwicklung des Medians des Bevölkerungsbestandes zeigt eine recht deutliche Abnahme der Bevölkerung von 2009 bis 2060 um etwa 9 Millionen Personen. Betrachtet man allerdings die Unter- und Obergrenze des Bereichs, in dem die zukünftige Entwicklung mit hoher Wahrscheinlichkeit liegen wird, so zeigt sich eine große Unsicherheit. Einerseits liegt die untere Grenze nochmals wesentlich deutlicher unter dem Ausgangswert und deutet auf die Möglichkeit einer noch drastischeren Abnahme hin, andererseits liegt die Obergrenze leicht über dem Ausgangswert des Jahres 2009. Insgesamt scheint somit ein langfristiger Bevölkerungsrückgang zwar sehr wahrscheinlich zu sein, allerdings kann die Stärke dieses Rückgangs nur schwer beziffert werden.

Der Altenquotient ist ein grober Indikator für die Alterung der Bevölkerung und wird berechnet als das Verhältnis des Bevölkerungsbestandes in den Altern von 65 und mehr Lebensjahren zum Bevölkerungsbestand in den Altern von 20 bis 64 Jahren multipliziert mit 100. Ein Wert des Altenquotienten von 35 würde beispielsweise bedeuten, dass 100 Personen im erwerbsfähigen Alter (20 bis 64 Jahre) 35 Personen im Rentenalter gegenüberstehen. Vergleicht man den Ausgangswert des Jahres 2009 mit dem Median für 2030 und dem Median für 2060, so zeigt sich ein deutlicher Anstieg. Bis 2060 verdoppelt sich der Wert fast. Die Unter- und Obergrenze der Entwicklung deuten ebenfalls auf einen starken Anstieg hin, womit eine langfristige, deutliche Alterung der Bevölkerung als relativ sicher gelten kann.

### 11.3.2 Entwicklung linearer Verwandtschaft

Die Vorausberechnung der linearen Verwandtschaft erfolgt differenziert nach Vor- und Nachfahren, nach Geschlecht und nach West- und Ostdeutschland. Genauer wird bei Nachfahren sowohl die erwartete („durchschnittliche“) Kinderzahl als auch die erwartete Anzahl an Enkelkindern ermittelt. Die betrachteten Vorfahren umfassen beide Elternteile sowie alle Großeltern mütterlicher- und väterlicherseits, für die jeweils die Wahrscheinlichkeit berechnet wird, dass sie noch leben. Wie für die Bevölkerungsentwicklung gilt, dass 1000 Simulationen berechnet wurden und zusätzlich zum Median wieder der Bereich angegeben werden kann, in dem 90% der Ergebnisse der Simulationen liegen.

Insgesamt ist festzuhalten, dass die Ergebnisse auf eine leichte Abnahme der Zahl der Nachfahren hinweisen, die allerdings mit einer großen Unsicherheit behaftet ist. Beispielhaft lässt sich dies an Abbildung 11.1 erkennen. In dieser ist die erwartete Kinderzahl westdeutscher Frauen nach Alter eingetragen, wobei zum einen Ergebnisse für das Jahr 2009 und zum anderen Ergebnisse für 2060 zu sehen sind. Für das Alter von 50 Jahren ergibt sich 2060 beispielsweise ein Median von 1.44 (durchgezogene Linie). Die untere und obere Grenze des Bereichs, der 90% der Ergebnisse umfasst, liegen für dieses Alter bei 1.18 und 1.92 (gestrichelte Linien). Der Ausgangswert des Jahres 2009 beträgt 1.58 (gepunktete Linie).

Vergleicht man für einzelne Alter den Median für 2060 und die Ergebnisse für 2009, so liegen die Resultate für 2060 durchgängig unter denen des Ausgangsjahres. Dies gilt insbesondere für Alter ab 50, bei denen die Unterschiede teils sehr deutlich sind. Dabei ist aber zumindest für mittlere Alter die Unsicherheit extrem hoch. Beispielsweise beträgt die Differenz der unteren und oberen Grenze des Bereichs, in dem 90% der Simulationen liegen, für das Alter von 40 Jahren fast 0.9. Ähnliche Befunde zeigen sich für die Kinderzahl ostdeutscher Frauen, die Kinderzahl west- und ostdeutscher Männer und auch bei der Anzahl der Enkel.

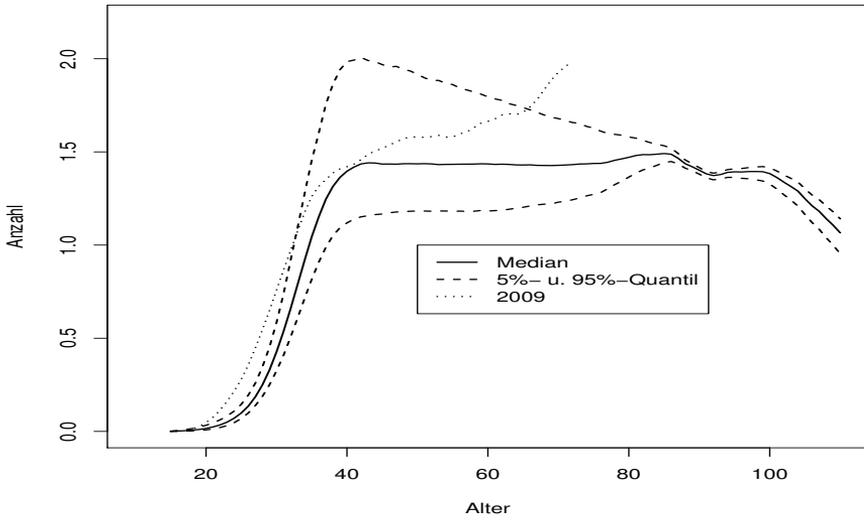
Die Wahrscheinlichkeit, dass die Eltern einer Person leben, wird den Ergebnissen folgend in Zukunft zunehmen, während die Wahrscheinlichkeit lebender Großeltern sich nur geringfügig oder gar nicht ändert. Dabei sind die Ergebnisse für Vorfahren mit geringerer Unsicherheit behaftet als die für Nachfahren. Als Beispiel werden die Großeltern mütterlicherseits betrachtet. In Abbildung 11.2 sind die Wahrscheinlichkeiten eingezeichnet, dass Großeltern mütterlicherseits einer Person im Alter von 0 Jahren noch leben, wobei sich diese Resultate auf Ostdeutschland beziehen. Zu sehen ist die Entwicklung dieser Wahrscheinlichkeiten von 2009 bis 2060. Beispielsweise beträgt die Wahrscheinlichkeit, dass die Großmutter mütterlicherseits einer Person im Alter 0 im Jahr 2009 noch lebt, etwa 95%. Die gestrichelten Linien geben wieder die Bereiche an, in denen die Ergebnisse mit einer Wahrscheinlichkeit von 90% liegen.

Über den Zeitraum von 2009 bis 2060 zeigt sich für die Großmutter mütterlicherseits praktisch keine Veränderung, wenn der Median mit dem Ausgangswert verglichen wird. Gleichzeitig ist die Unsicherheit sehr niedrig. Wird für den Großvater mütterlicherseits der Median betrachtet, deutet dieser zwar auf einen Anstieg hin, allerdings umfasst der Unsicherheitsbereich zum einen den Ausgangswert und zum anderen liegt die Untergrenze unter diesem Wert. Zudem ist auch der Anstieg, der sich bei Betrachtung des Medians andeutet, eher moderat. Bei den Großeltern väterlicherseits und für andere Alter ergeben sich recht ähnliche Ergebnisse. Für Westdeutschland ist zwar von einem leichten Anstieg der Wahrscheinlichkeiten auszugehen, allerdings fällt dieser sehr niedrig aus.

### 11.3.3 Anwendungsbeispiele: Geteilte Lebenszeit und Pflege

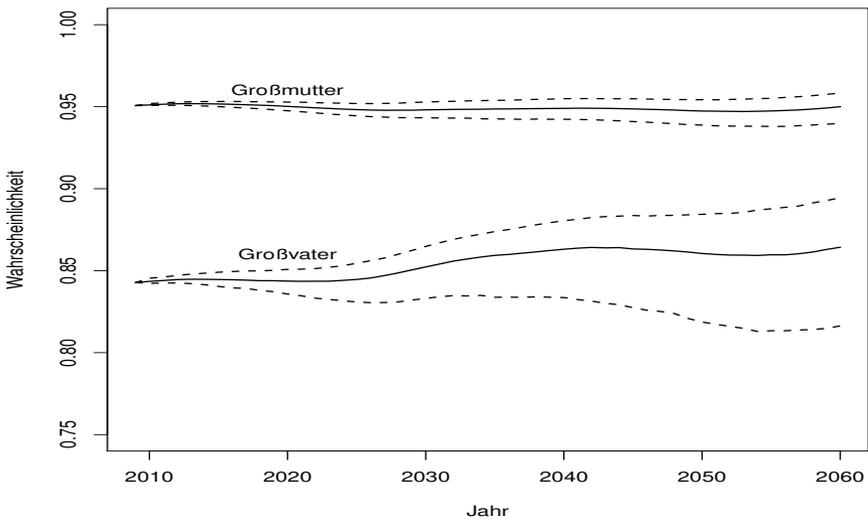
Die Ergebnisse der Vorausberechnung können genutzt werden, um die Eingangs formulierten, weitergehenden Fragen zu beantworten. Betrachtet man die geteilte Lebenszeit von Großeltern und Enkeln, deuten die Ergebnisse entgegen der Überlegungen in der Literatur nur auf marginale Veränderungen hin, da sich die Effekte des Anstiegs der Lebenserwartung einerseits und des Anstiegs des Alters bei der Geburt andererseits ungefähr aufheben, wie es sich bereits bei den Ergebnissen in Abbildung 11.2 andeutet.

Abbildung 11.1: Erwartete Anzahl lebender Kinder westdeutscher Frauen 2060 nach Altersklassen, sowie Ergebnisse für 2009



Quelle: eigene Berechnung

Abbildung 11.2: Wahrscheinlichkeiten für lebende Großeltern mütterlicherseits in Altersklasse 0 von 2009 bis 2060, Ostdeutschland



Quelle: eigene Berechnung

Um der Frage nachzugehen, inwieweit sich verwandtschaftliche Unterstützungspotentiale ändern, wird betrachtet, wie sich das Verhältnis von Pflegebedürftigen zu potentiell pflegenden Nachfahren entwickelt. Zum einen werden Berechnungen durchgeführt, die auf den bereits vorgestellten Ergebnissen aufsetzen, und zum anderen wird eine bisher in der Literatur übliche Vorgehensweise angewendet, bei der nicht von einer direkten Betrachtung von Verwandtschaft ausgegangen und eine Art Überschlagsrechnung durchgeführt wird.

Es lässt sich festhalten, dass zwar mit einer Abnahme verwandtschaftlicher Unterstützungspotentiale zu rechnen ist, sich die Ergebnisse der beiden unterschiedlichen Vorgehensweisen aber deutlich unterscheiden. Werden die Ergebnisse des in dieser Arbeit entwickelten Vorausberechnungsansatzes betrachtet, so zeigt sich eine eher moderate Abnahme, die deutlich weniger drastisch ausfällt, als beim bisher in der Literatur üblichen Vorgehen. Dies deutet darauf hin, dass die Verwendung von Überschlagsrechnungen zu einer Überschätzung der Abnahme verwandtschaftlicher Unterstützungspotentiale führt.

### *11.3.4 Überblick*

Ausgehend von dieser kurzen Vorstellung zentraler Ergebnisse können die wesentlichen Resultate bezogen auf die zukünftige Entwicklung des Vorhandenseins linearer Verwandtschaft überblicksartig wie folgt aufgelistet werden:

- Eine Abnahme linearer Nachfahren, verbunden mit einer großen Unsicherheit
- Eine steigende Wahrscheinlichkeit für lebende Eltern und eine gleich bleibende Wahrscheinlichkeit für lebende Großeltern, jeweils verbunden mit geringer Unsicherheit
- Eine ungefähr gleichbleibende geteilte Lebenszeit von Großeltern und Enkelkindern
- Eine vergleichsweise moderate Abnahme verwandtschaftlicher Unterstützungspotentiale pflegebedürftiger Personen

## **11.4 Diskussion**

Betrachtet man zunächst die Ergebnisse der zwei Anwendungsbeispiele, so zeigt sich in beiden Fällen, dass einfache Spekulationen oder Überschlagsrechnungen zu Fehleinschätzungen führen und wenig geeignet sind, Überlegungen zur zukünftigen Entwicklung des Vorhandenseins linearer Verwandtschaft anzustellen. Im Umkehrschluss bestätigen diese Resultate den Bedarf nach und die Nützlichkeit von Vorausberechnungen, die Verwandtschaft explizit modellieren.

Keht man zu den in der Einleitung dieser Arbeit angestellten Überlegungen zurück, so kann vermutet werden, dass in Zukunft nicht oder nur eingeschränkt von einem Bedeutungszuwachs intergenerationaler Beziehungen auszugehen ist und die Nachfrage nach professioneller Pflege nicht so stark steigen wird, wie in der Literatur oftmals dargestellt. Dabei ist aber auch zu bedenken, dass in dieser Arbeit lediglich demographische Einflussfaktoren auf diese beiden Themengebiete betrachtet wurden, obwohl davon auszugehen ist, dass weitere Aspekte Einfluss auf die zukünftige Entwicklung haben werden. Dies reicht von eher normativen Aspekten,

wie beispielsweise der normativen Bedeutung von und dem damit verbundenen Grad an sozialer Verpflichtung zur Pflege von Angehörigen, bis hin zu rechtlich-politischen Fragen, wie beispielsweise der Vereinbarkeit von Pflege und Beruf. Insofern liefern die Resultate zwar wichtige Hinweise über die zukünftige Entwicklung, sind aber zugleich mit inhaltlichen Einschränkungen verbunden.

Daneben ist das vorgestellte Vorgehen an etliche Annahmen geknüpft. Insbesondere ist zu bedenken, dass „Probleme“ eines Modellierungsansatzes in aller Regel nicht einfach beseitigt werden können und sich nur zum Preis neuer „Probleme“ lösen lassen, so dass nicht ohne weiteres ein „optimaler“ Ansatz entworfen werden kann. Deshalb ist ein Vergleich von Modell und Realität notwendig, wenn die Ergebnisse eines Modells interpretatorisch belastbar sein sollen. Bei der in dieser Arbeit durchgeführten Validierung zeigte sich allerdings, dass kaum eindeutige Schlussfolgerungen gezogen werden können, da unklar ist, wie die bei den Ergebnissen zu Vorfahren auftretenden Abweichungen zwischen Modell und Realität erklärt werden können – irrt das Modell oder sind die verwendeten Vergleichsdaten fehlerbehaftet? Dieses Problem tritt auch bei den wenigen anderen vorliegenden Validierungsstudien auf (vgl. Cohen 1991a; Wachter 1997). Trotz dieser Problematik scheint ein Vergleich sinnvoll und sollte grundsätzlich bei der Anwendung demographischer Modelle durchgeführt werden, da bei einem Auftreten von Differenzen zwischen Modell und Vergleichsdaten begründet werden muss, woher diese rühren und Schwächen des Modells oder der Vergleichsdaten aufgedeckt werden können.

Die explizite Berücksichtigung demographischer Unsicherheit erlaubt es, Annahmen über die zukünftige Entwicklung zumindest in Teilen zu umgehen, insbesondere verglichen mit einer rein szenariobasierten Vorgehensweise, bei der lediglich einige wenige Entwicklungsverläufe in den Fokus genommen werden. Zudem zeigt die Modellierung von Unsicherheit, wie belastbar die Ergebnisse sind und wie weit man mit Schlussfolgerungen gehen kann. Beispielsweise kann für die Bevölkerungsentwicklung festgehalten werden, dass ein Rückgang sehr sicher ist, während hingegen das Ausmaß dieses Rückgangs unklar und eine große Bandbreite an Entwicklungen denkbar ist. Ein mögliches Problem der Berücksichtigung von Unsicherheit ist darin zu sehen, dass die Ergebnisse einen höheren Abstraktionsgrad aufweisen als Ergebnisse einfacher Szenarien und damit schwieriger zu interpretieren sind.

Daneben basiert auch die Vorausberechnung der demographischen Komponenten, bei der die Berücksichtigung der Unsicherheit erfolgt, auf einigen Annahmen. Dies gilt insbesondere für die Entwicklung in Ostdeutschland – sowohl bei Mortalität als auch bei Fertilität wird eine Angleichung an Westdeutschland erwartet. Bezogen auf Mortalität erscheint dies nicht gänzlich unplausibel, bezogen auf Fertilität gibt es aber etliche Hinweise, dass sich West und Ost immer noch substantiell unterscheiden (vgl. Goldstein & Kreyenfeld 2011; Kreyenfeld & Konietzka 2004), was eine Angleichung fragwürdig erscheinen lässt. Insofern sind die Resultate über die Entwicklung in Ostdeutschland mit Vorsicht zu betrachten.

Anders verhält es sich mit den Ergebnissen für Westdeutschland: Diese basieren zwar ebenfalls auf etlichen Annahmen, allerdings wurde für einen Großteil dieser im Laufe der Arbeit gezeigt, dass sie als realistisch gelten können oder aber keine übermäßigen Fehler entstehen, wenn sie nicht zutreffen. Von dieser Feststellung ausgehend können die Ergebnisse in Anschluss an Keyfitz & Caswell (2005: 271) als Prognose aufgefasst werden, zumal sie von realistischen Annahmen ausgehen. Auch wenn man die Ergebnisse anders bewerten sollte

können sie als Hilfe dienen, über die zukünftige Entwicklung von linearer Verwandtschaft nachzudenken: „Models, of course, are never true, but fortunately it is only necessary that they be useful.“ (Box 1979: 2).

## 11.5 Ausblick

Das vorgestellte Vorgehen bietet etliche weitere Anwendungsmöglichkeiten. Beispielsweise sind weitergehende Berechnungen zu verwandtschaftlichen Unterstützungspotentialen denkbar, bei denen die bisherigen Analysen um zusätzliche wichtige Bestimmungsgrößen der Unterstützungspotentiale ergänzt werden könnten. Eine Möglichkeit wäre etwa die Berücksichtigung der zukünftigen Entwicklung der Erwerbstätigkeit insbesondere von Frauen, unter der Annahme, dass Erwerbstätigkeit die Übernahme einer Pflegeaufgabe verhindert oder zumindest stark einschränkt.

Eine weitere mögliche, auf die Pflege bezogene Fragestellung geht von einem Perspektivenwechsel aus und fragt nicht mehr nach den Unterstützungspotentialen pflegebedürftiger Personen, sondern nimmt umgekehrt in den Fokus, wie viele pflegebedürftige Verwandte junge Personen und Personen mittleren Alters aufweisen und wie sich in Zukunft die Wahrscheinlichkeit entwickelt, im Lebensverlauf pflegebedürftige Verwandte zu haben.

Daneben sind noch diverse zusätzliche Anwendungsmöglichkeiten des Ansatzes über die geteilte Lebenszeit von Generationen und Pflege hinaus denkbar. Interessant könnte zum einen eine international vergleichende Anwendung sein. Hier wäre beispielsweise ein Vergleich von Vorausberechnungen für mehrere europäische Länder denkbar. Ebenso denkbar wäre eine historische Anwendung, beispielsweise in Hinblick auf die in dieser Arbeit mehrmals angesprochene Vertikalisierungsthese. Hierfür wären allerdings vergleichsweise lange Datenreihen über die demographische Entwicklung notwendig, wie sie für Deutschland nicht vorliegen.

Aus methodischer Sicht wäre es interessant, das vorgestellte Vorgehen so zu erweitern, dass neben linearer Verwandtschaft auch laterale Verwandtschaft, wie beispielsweise Geschwister, und angeheiratete Verwandtschaft berücksichtigt werden können. Hierfür müsste der Familienstand mit in das Modell aufgenommen werden und eine genauere Betrachtung der Fertilität erfolgen.

Das Modell kann prinzipiell einfach um die genannten Merkmale erweitert werden. Allerdings würden hierdurch neue Probleme entstehen. Bezüglich des Familienstands müsste beispielsweise Konsistenz zwischen der Zahl der verheirateten Männer und der verheirateten Frauen erreicht werden, was sich als relativ komplexes Problem erweist. Zudem müssten zusätzliche Raten (bspw. Heiratsraten) vorausberechnet und die bereits berücksichtigten demographischen Größen stärker differenziert betrachtet werden (bspw. eheliche und nicht-eheliche Fertilität). Hiermit einhergehen würde ein deutlich erhöhter Datenbedarf, der zumindest bei der Fertilität momentan nicht gedeckt werden kann (s. auch Unterabschnitt 5.4). Zwar gibt es in der Literatur für einen Teil dieser Probleme Lösungsansätze (s. Kapitel 4), eine theoretische und praktische Umsetzung im Rahmen des verwendeten Modells geht allerdings über diese Arbeit hinaus.

Weiterhin wäre die Kombination von stochastischen Vorausberechnungen demographischer Komponenten mit Mikrosimulationen eine interessante Alternative zum hier gewählten Vorgehen, welches sich in der Literatur bisher nicht findet (vgl. Lee 1998: 157). Eine weitere methodische Frage, die die Übertragbarkeit des Ansatzes auf andere Länder und Kontexte betrifft, ist, wie vorgegangen werden kann, wenn die Fertilität von Männern nicht wie in der vorliegenden Arbeit aus einer Vollerhebung ermittelt werden kann. Denkbar wäre hier zum einen die Schätzung aus Umfragedaten und zum anderen eine Abwandlung des in dieser Arbeit verwendeten Vorgehens zur Ermittlung der historischen Entwicklung der Fertilität von Männern, bei der von fixen Altersabständen von Müttern und Vätern bei der Geburt ausgegangen wird.

Insgesamt bietet das Modell also etliche Weiterentwicklungspotentiale. Aber auch in der in dieser Arbeit vorgestellten Form ergeben sich bereits wichtige Anwendungsmöglichkeiten. Vor allem aus wissenschaftlicher Sicht interessant ist die Betrachtung der geteilten Lebenszeit von Generationen. Die Untersuchung verwandtschaftlicher Unterstützungspotentiale in Bezug auf die Betreuung Pflegebedürftiger besitzt daneben auch eine hohe gesellschaftspolitische Relevanz.



## Literaturverzeichnis

- Ahlburg, D. A., Lutz, W., 1998: Introduction: The Need to Rethink Approaches to Population Forecasts. *Population And Development Review* 24 (Supplement): 1-14
- Ahmad, O. B., Lopez, A. D., Inove, M., 2000: The decline in child mortality: a reappraisal. *Bulletin of the World Health Organization* 78: 1175-1191
- Akers, D. S., 1965: Cohort Fertility Versus Parity Progression as Methods of Projecting Births. *Demography* 2: 414-428
- Alho, J., 1984: Probabilistic Forecasts. The Case of Population Projections. *Scandinavian Housing and Planning Research* 1: 99-105
- Alho, J. M., 1985: Interval Estimates for Future Population. In: *Proceedings of the Social Statistics Section / American Statistical Association*, 1985, S. 44-51
- Alho, J. M., 1997: Scenarios, Uncertainty and Conditional Forecasts of the World Population. *Journal of the Royal Statistical Society* 160: 71-85
- Alho, J. M., Keilman, N., 2010: On future household structure. *Journal of the Royal Statistical Society* 173: 117-143
- Alho, J. M., Spencer, B. D., 1985: Uncertain Population Forecasting. *Journal of the American Statistical Association* 80: 306-314
- Alho, J. M., Spencer, B. D., 1990a: Effects of targets and aggregation on the propagation of error in mortality forecasts. *Mathematical Population Studies* 2: 209-227
- Alho, J. M., Spencer, B. D., 1990b: Error Models for Official Mortality Forecasts. *Journal of the American Statistical Association* 85: 609-616
- Alho, J. M., Spencer, B. D., 1991: A Population Forecast as a Database: Implementing the Stochastic Propagation of Error. *Journal of Official Statistics* 7: 295-310
- Alho, J. M., Spencer, B. D., 2005: *Statistical Demography and Forecasting*. New York: Springer
- Anderson, K. G., 2006: How Well Does Paternity Confidence Match Actual Paternity? Evidence from Worldwide Nonpaternity Rates. *Current Anthropology* 47: 513-520
- Anderson, T. W., Goodman, L. A., 1957: Statistical Inference about Markov Chains. *Annals of Mathematical Statistics* 28: 89-110
- Astone, N. M., Nathanson, C. A., Schoen, R., Kim, Y. J., 1999: Family Demography, Social Theory, and Investment in Social Capital. *Population And Development Review* 25: 1-31
- Atkins, J. R., 1974: On the Fundamental Consanguineal Numbers and Their Structural Basis. *American Ethnologist* 1: 1-31
- Attias-Donfut, C., Ogg, J., Wolff, F.-C., 2005: Family Support. In: Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegrist, J., Weber, G. (Hg.) *Health, Ageing and Retirement in Europe. First Results from the Survey of Health, Ageing and Retirement in Europe*, S. 171-178. Mannheim
- Babel, B., 2007: *Bevölkerungsvorausberechnung unter besonderer Berücksichtigung von Unsicherheiten*. Lohmar: Eul
- Babel, B., Bomsdorf, E., Schmidt, R., 2008: Forecasting German mortality using panel data procedures. *Journal of Population Economics* 21: 541-555
- Bacon, B. R., 1997: Work, retirement and dependency. *People and Place* 5: o. S.
- Baier, D., Hadjar, A., 2004: Wie wird Leistungsorientierung von den Eltern auf die Kinder übertragen? Ergebnisse einer Längsschnittstudie. *Zeitschrift für Familienforschung* 15: 156-177
- Bartlema, J., 1987: *Developments in Kinship Support Networks for the Aged in the Netherlands*. Dissertation, Universität Tilburg
- Bartlema, J., 1988: Modelling Step-Families. Exploratory Findings. *European Journal of Population* 4: 197-221
- Bartlema, J., 1989: Modelling Stepfamilies: First Results. NIDI Report 3
- Bartlema, J., de Jong, P., 1985: Shrinking Kinship-Support Networks in the Netherlands. The numerical results of a simulation. Progress Report, Catholic University of Tilburg
- Bartlema, J., Winkelbauer, L., 1986: Modelling Kinship with LISP. A two-sex model of kin-counts. IIASA Working Paper 86-69

- Bell, W. R., 1997: Comparing and Assessing Time Series Methods for Forecasting Age-Specific Fertility and Mortality Rates. *Journal of Official Statistics* 13: 279-303
- Bengtson, V. L., 2001: Beyond the Nuclear Family: The Increasing Importance of Multigenerational Bonds. *Journal of Marriage and the Family* 63: 1-16
- Bengtson, V. L., Martin, P., 2001: Families and intergenerational relationships in aging societies: comparing the United States with German-speaking countries. *Zeitschrift für Gerontologie und Geriatrie* 34: 207-217
- Bengtson, V. L., Roberts, R. E. L., 1991: Intergenerational Solidarity in Aging Families: An Example of Formal Theory Construction. *Journal of Marriage and the Family* 53: 856-870
- Berkowitz, J., Kilian, L., 2000: Recent developments in bootstrapping time series. *Econometric Reviews* 19: 1-48
- Bertino, S., Pinnelli, A., Vichi, M., 1988: Two Models for Microsimulation of Family Life Cycle and Family Structure. *Genus* 44: 1-23
- Betz, F., Lipps, O., 2004: Stochastic Population Projection for Germany Based on the Quadratic Spline Approach to Modelling Age Specific Fertility Rates. MEA Discussion Paper 59-2004
- Bien, W., Marbach, W., 1991: Haushalt – Verwandtschaft – Beziehungen. Familie als Netzwerk. In: Bertram, H. (Hg.) *Die Familie in Westdeutschland*, S. 3-44. Opladen: Leske + Budrich
- Bijak, J., Wisniowski, A., 2010: Bayesian forecasting of immigration to selected European countries by using expert knowledge. *Journal of the Royal Statistical Society* 173: 775-796
- Billari, F. C., Goisis, A., Liefbroer, A. C., Settersten, R. A., Aassve, A., Hagestad, G., Spéder, Z., 2011: Social age deadlines for the childbearing of women and men. *Human Reproduction* 26: 616-622
- Blinkert, B., Gräf, B., 2009: Deutsche Pflegeversicherung vor massiven Herausforderungen. Deutsche Bank Research 442
- Blinkert, B., Klie, T., 2004: Gesellschaftlicher Wandel und demographische Veränderungen als Herausforderung für die Sicherstellung der Versorgung von pflegebedürftigen Menschen. *Sozialer Fortschritt* 11/12-04: 319-325
- Blome, A., Keck, W., Alber, J., 2008: *Generationbeziehungen im Wohlfahrtsstaat. Lebensbedingungen und Einstellungen von Altersgruppen im internationalen Vergleich*. Wiesbaden: VS Verlag für Sozialwissenschaft
- Boes, S., 2004: *Die Anwendung der Konzepte probabilistischer Bevölkerungsmodelle auf Prognosen für den Hochschulbereich*. Norderstedt: Kube
- Bomsdorf, E., Babel, B., Schmidt, R., 2008: Zur Entwicklung der Bevölkerung, der Anzahl der Schüler, der Studienanfänger und der Pflegebedürftigen. *Sozialer Fortschritt* 57: 125-132
- Bongaarts, J., 1987: The Projection of Family Composition over the Life Course with Family Status Life Tables. In: Bongaarts, J., Burch, T. K., Wachter, K. W. (Hg.) *Family Demography. Methods and their Application*, S. 189-212. Oxford: Clarendon Press
- Bongaarts, J., Feeney, G., 1998: On the Quantum and Tempo of Fertility. *Population and Development Review* 24: 271-291
- Booth, H., 2006: Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting* 22: 547-581
- Booth, H., Tickle, L., 2008: Mortality Modelling and Forecasting: A Review of Methods. *Annals of Actuarial Science* 3: 3-43
- Booth, H., Tickle, L., Smith, L., 2005: Evaluation of the Variants of the Lee-Carter Method of Forecasting Mortality: A Multi-Country Comparison. *New Zealand Population Review* 31: 13-34
- Bowley, A. L., 1924: Births and Population in Great Britain. *Journal of the Royal Economic Society* 34: 188-192
- Box, G. E. P., 1979: Some Problems of Statistics and Every Day Life. *Journal of the American Statistical Association* 74: 1-4
- Brandt, M., 2009: *Hilfe zwischen Generationen. Ein europäischer Vergleich*. Wiesbaden: VS Verlag für Sozialwissenschaft
- Brandt, M., Deindl, C., Haberkern, K., Szydlik, M., 2008: Reziprozität zwischen erwachsenen Generationen: Familiäre Transfers im Lebenslauf. *Zeitschrift für Gerontologie und Geriatrie* 41: 374-381

- Brass, W., 1983: The formal demography of the family: an overview of the proximate determinants. OPCS Occasional Paper 31
- Bretz, M., 2000: Methoden zur Bevölkerungsvorausberechnung. In: Mueller, U., Nauck, B., Diekmann, A. (Hg.) *Handbuch der Demographie*, S. 643-681. Berlin: Springer
- Bretz, M., 2002: Treffsicherheit von Bevölkerungsvorausberechnungen. In: Bundesinstitut für Bevölkerungswissenschaft (Hg.) *Demographische Vorausschätzungen – Grenzen und Möglichkeiten, Methoden und Ziele*, Materialien zur Bevölkerungswissenschaft 104, S. 1-38
- Breyer, F., Felder, S., 2004: Life Expectancy and Health Care Expenditures in the 21st Century: A New Calculation for Germany Using the Costs of Dying. DIW Discussion Paper 452
- Brouard, N., 1977: Évolution de la fécondité masculine depuis le début du siècle. *Population* 32: 1123-1158
- Brown, S. K., Bean, F. D., 2006: International Migration. In: Poston, D. L., Micklin, M. (Hg.) *Handbook of Population*, S. 347-382. New York: Springer
- Börsch-Supan, A., 1991: Aging population: problems and policy options in the US and Germany. *Economic Policy* 6: 103-139
- Börsch-Supan, A., Wilke, C., 2007: Szenarien zur mittel- und langfristigen Entwicklung der Anzahl der Erwerbspersonen und der Erwerbstätigen in Deutschland. MEA Discussion Paper 153
- Bühler, W. J., 1971: Generations and Degree of Relationship in Supercritical Markov Branching Processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 18: 141-152
- Bühler, W. J., 1974: On the family structure of populations. *Advances in Applied Probability* 6: 192-193
- Bujard, M., 2011: *Geburtenrückgang und Familienpolitik*. Baden-Baden: Nomos
- Bundesministerium des Innern (Hg.) 2011: *Demografiebericht. Bericht der Bundesregierung zur demografischen Lage und künftigen Entwicklung des Landes*. Berlin
- Burch, T., 1995a: Theories of Household Formation: Progress and Challenges. In: Imhoff, E. v., Kuijsten, A., Hooimeijer, P., Wissen, L. v. (Hg.) *Household Demography and Household Modeling*, S. 85-108. New York: Plenum Press
- Burch, T. K., 1979: Household and Family Demography: A Bibliographic Essay. *Population Index* 45: 173-195
- Burch, T. K., 1995b: Estimating the Goodman, Keyfitz, Pullum kinship equations: An alternative procedure. *Mathematical Population Studies* 5: 161-170
- Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., Balevich, I., 2009: A Quantitative Comparison of Stochastic Mortality Models Using Data from England and Wales and the United States. *North American Actuarial Journal* 13: 1-35
- Cannan, E., 1895: The Probability of a Cessation of Growth of Population in England and Wales During the Next Century. *Journal of the Royal Economic Society* 5: 505-515
- Carter, L. R., 2010: Long-Run Relationships in Differential U.S. Mortality Forecasts by Race and Sex: Tests for Co-integration. In: Tuljapurkar, S., Ogawa, N., Gauthier, A. H. (Hg.) *Ageing in Advanced Industrialized States*, S. 47-75. New York: Springer
- Carter, L. R., Lee, R. D., 1992: Modelling and forecasting US sex differentials in mortality. *International Journal of Forecasting* 8: 393-411
- Caswell, H., 2001: *Matrix Population Models. Construction, Analysis, and Application*. Sunderland: Sinauer, 2. Aufl.
- Chatfield, C., 1993: Calculating Interval Forecasts. *Journal of Business & Economic Statistics* 11: 121-135
- Cheng, P. C. R., Lin, E. S., 2010: Completing incomplete cohort fertility schedules. *Demographic Research* 23: 223-256
- Chiang, C. L., 1984: *The Life Table and Its Applications*. Malabar: Krieger
- Clark, R. L., Sprengler, J. J., 1980: Dependency Ratios: Their Use in Economic Analysis. *Research in Population Economics* 2: 63-73
- Cohen, J. E., 1977: Ergodicity of Age Structure in Populations with Markovian Vital Rates, III: Finite-State Moments and Growth Rate. *Advances in Applied Probability* 9: 462-475
- Cohen, J. E., 1986: Population Forecasts and Confidence Intervals for Sweden: A Comparison of Model-Based and Empirical Approaches. *Demography* 23: 105-126

- Cohen, M. L., 1991a: Evaluation of Microsimulation Models: Literature Review. In: Citro, C. F., Hanushek, E. A. (Hg.) *Improving Information for Social Policy Decisions. The Use of Microsimulation Modeling*, S. 255-275. Washington: National Academy Press
- Cohen, M. L., 1991b: Statistical Matching and Microsimulation Models. In: Citro, C. F., Hanushek, E. A. (Hg.) *Improving Information for Social Policy Decisions. The Use of Microsimulation Modeling*, S. 62-85. Washington: National Academy Press
- Cohen, M. L., 1991c: Variance Estimation of Microsimulation Models Through Sample Reuse. In: Citro, C. F., Hanushek, E. A. (Hg.) *Improving Information for Social Policy Decisions. The Use of Microsimulation Modeling*, S. 237-254. Washington: National Academy Press
- Coleman, D., 2000: Male Fertility Trends in Industrial Countries: Theories in Search for Some Evidence. In: Bledsoe, C., Lerner, S., Guyer, J. (Hg.) *Fertility and the Male Life Cycle in the Era of Fertility Decline*, S. 29-60. Oxford: Oxford University Press
- Congdon, P., 1993: Statistical Graduation in Local Demographic Analysis and Projection. *Journal of the Royal Statistical Society* 156: 237-270
- Coresh, J., Goldman, N., 1988: The Effect of Variability in the Fertility Schedule on Numbers of Kin. *Mathematical Population Studies* 1: 137-156
- Daatland, S. O., Herlofson, K., 2003: Lost solidarity or changed solidarity: a comparative European view of normative family solidarity. *Ageing & Society* 23: 537-560
- Day, A. T., 1989: Kinship Networks and Informal Support in the Later Years. In: Grebenik, E., Höhn, C., Makensen, R. (Hg.) *Later Phases of the Family Life Cycle: Demographic Aspects*, S. 183-207. Oxford: Clarendon Press
- de Beer, J., 1985: A Time Series Model for Cohort Data. *Journal of the American Statistical Association* 80: 525-530
- de Beer, J., 1988: Predictability of Demographic Variables in the Short Run. *European Journal of Population* 4: 283-296
- de Beer, J., 1989: Projecting Age-Specific Fertility Rates by Using Time-Series Methods. *European Journal of Population* 5: 315-346
- de Beer, J., 2000: Dealing with uncertainty in population forecasting. Voorburg: Statistics Netherlands
- Deistler, M., Feichtinger, G., 1974: The Linear Model Formulation of a Multitype Branching Process Applied to Population Dynamics. *Journal of the American Statistical Association* 69: 662-664
- Denton, F. T., Feaver, C. H., Spencer, B. G., 2005: Time series analysis and stochastic forecasting: An econometric study of mortality and life expectancy. *Journal of Population Economics* 18: 203-227
- DeVos, S., Palloni, A., 1989: Formal Models and Methods for the Analysis of Kinship and Household Organization. *Population Index* 55: 174-198
- Diewald, M., 1989: Private Netzwerke. In: Statistisches Bundesamt (Hg.) *Datenreport 1989*, S. 454-461. Stuttgart: Bonn Aktuell
- Dinkel, R. H., 1992: Kohortensterbetafeln für die Geburtsjahrgänge 1900 bis 1962 in den beiden Teilen Deutschlands. *Zeitschrift für Bevölkerungswissenschaft* 18: 95-116
- Dinkel, R. H., Hartmann, K., Lebok, U., 1997: Langfristige Veränderungen in der Verfügbarkeit häuslicher Unterstützungspotentiale aufgrund familiärer Strukturverschiebungen – Eine Modellrechnung. *Das Gesundheitswesen* 59, Sonderheft 1: 49-54
- Dinkel, R. H., Milenovic, I., 1992: Die Kohortenfertilität von Männern und Frauen in der Bundesrepublik Deutschland. *Zeitschrift für Bevölkerungswissenschaft* 44: 55-75
- Dinkel, R. H., Milenovic, I., 1993: Male and Female Fertility: A comparison of age-specific and cohort-fertility of both sexes in Germany. *Genus* 49: 147-158
- Doblhammer, G., Ziegler, U., 2010: Care Need Projections by Marital Status and Childlessness for Germany 2000-2030 based on the FELICIE Project. In: Doblhammer, G., Scholz, R. (Hg.) *Ageing, Care Need and Quality of Life. The Perspective of Care Givers and People in Need of Care*, S. 42-60. Wiesbaden: VS
- Dorbritz, J., 2004: Nur Tempoeffekte, aber kein neuer Babyboom. *BiB-Mitteilungen* 2/2004: 10-14
- Dorbritz, J., 2008: Germany: Family diversity with low actual and desired fertility. *Demographic Research* 19: 557-598

- Dorbritz, J., Hullen, G., Schiener, R., 1997: Prognose der Haushalts- und Familienstrukturen bis zum Jahr 2030. Abschlussbericht des gleichnamigen Projektes, Wiesbaden
- Dowd, K., Cairns, A. J. G., Blake, D., Coughlan, G. D., Epstein, D., Khalaf-Allah, M., 2008: Backtesting Stochastic Mortality Models: An Ex-Post Evaluation of Multi-Period-Ahead Density Forecasts. Pensions Institute Discussion Paper PI0803
- Dudel, C., 2008: Verwandtschaft heute und morgen. Bestimmung des Umfangs verwandtschaftlicher Netzwerke 2006 und 2050. Diskussionspapiere der Fakultät für Sozialwissenschaft 08-2, Bochum
- Dyke, B., 1981: Computer Simulation in Anthropology. *Annual Review of Anthropology* 10: 193-207
- Ecarius, J., 2007: Verwandtschaft. In: Ecarius, J. (Hg.) *Handbuch Familie*, S. 220-238. Wiesbaden: VS Verlag für Sozialwissenschaft
- El-Khorazaty, M. N., 1997: Family Life Cycle and Fertility in Germany Before Unification: 1947-1989. *The History of the Family* 2: 309-330
- Espenshade, T. J., Guzman, J. C., Westoff, C. F., 2003: The surprising global variation in replacement fertility. *Population Research and Policy Review* 22: 575-583
- Farkas, J. L., Hogan, D. P., 1995: The Demography of Changing Intergenerational Relationships. In: Bengtson, V. L., Schaie, W., Burton, L. (Hg.) *Adult Intergenerational Relations. Effects of Societal Change*, S. 1-19. New York: Springer
- Fehr, H., Habermann, C., 2006: Pension reform and demographic uncertainty: the case of Germany. *Journal of Pension Economics and Finance* 5: 69-90
- Feichtinger, G., 1977: Methodische Probleme der Familienlebenszyklus-Statistik. In: Albach, H., Helmstädter, E., Henn, R. (Hg.) *Quantitative Wirtschaftsforschung*, S. 171-183. Tübingen: Mohr Siebeck
- Feichtinger, G., 1987: The Statistical Measurement of the Family Life Cycle. In: Bongaarts, J., Burch, T. K., Wachter, K. W. (Hg.) *Family Demography. Methods and their Application*, S. 81-101. Oxford: Clarendon Press
- Feller, W., 1977: Extinction Probabilities in Branching Processes. In: Smith, D., Keyfitz, N. (Hg.) *Mathematical Demography*, S. 407-409. New York: Springer
- Fiocco, M., Putter, H., Houwelingen, H. C. v., 2008: Reduced-rank proportional hazards regression and simulation-based prediction for multi-state models. *Statistics in Medicine* 27: 4340-4358
- Forschungsdatenzentrum der Statistischen Ämter der Länder, 2006: Statistik der Geburten 2000. Merkmalsdefinitionen. Abrufbar unter <http://www.fdz-metadaten.de>
- Forschungsdatenzentrum der Statistischen Ämter der Länder, 2007: Statistik der Geburten 1995. Merkmalsdefinitionen. Abrufbar unter <http://www.fdz-metadaten.de>
- Forschungsdatenzentrum der Statistischen Ämter der Länder, 2009: Statistik der Geburten 2008. Merkmalsdefinitionen. Abrufbar unter <http://www.fdz-metadaten.de>
- Frick, J., Schupp, J., 2001: Das Sozio-oekonomische Panel (SOEP). Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI)
- Frick, J. R., Haisken-DeNew, J. P., Spiess, M., Wagner, G. G., 2005: Overview of the SOEP. In: Haisken-DeNew, J. P., Frick, J. R. (Hg.) *Desktop Companion to the German Socio-Economic Panel (SOEP)*
- Frick, J. R., Schmitt, C., 2010: BIOBIRTH: A Data Set on the Birth Biography of Female Respondents. In: Frick, J. R., Lohmann, H. (Hg.) *Biography and Life History Data in the German Socio Economic Panel (SOEP, v26, 1984-2009)*, S. 68-77. DIW Data Documentation 52
- Fuchs, J., Söhnlein, D., Weber, B., 2011: Projektion des Arbeitskräfteangebots bis 2050: Rückgang und Alterung sind nicht mehr aufzuhalten. IAB-Kurzbericht 16/2011
- Galler, H. P., 1977: Das Bevölkerungsmodell des SPES-Projekts – ein Modell auf Individualdatenbasis zur Vorausschätzung von Familien und Haushaltsstrukturen. *Allgemeines Statistisches Archiv* 61: 371-390
- Galler, H. P., 1990: Verwandtschaftsnetzwerke im demographischen Wandel. *Acta demographica* 1: 63-84

- Galler, H. P., 1994a: Mikrosimulationsmodelle in der Forschungsstrategie des Sonderforschungsbereich 3. In: Hauser, R., Ott, N., Wagner, G. (Hg.) *Mikroanalytische Grundlagen der Gesellschaftspolitik. Ergebnisse aus dem gleichnamigen Sonderforschungsbereich an den Universitäten Frankfurt und Mannheim*, Bd. 2, S. 369-379. Berlin: Akademie Verlag
- Galler, H. P., 1994b: Prognoseeigenschaften dynamischer Mikrosimulationsmodelle. In: Hauser, R., Ott, N., G., W. (Hg.) *Mikroanalytische Grundlagen der Gesellschaftspolitik. Ergebnisse aus dem gleichnamigen Sonderforschungsbereich an den Universitäten Frankfurt und Mannheim*, Bd. 2, S. 428-452. Berlin: Akademie Verlag
- Galler, H. P., 1995: Competing Risks and Unobserved Heterogeneity, with Special Reference to Dynamic Microsimulation Models. In: Imhoff, E. v., Kuijsten, A., Hooimeijer, P., Wissen, L. v. (Hg.) *Household Demography and Household Modeling*, S. 203-224. New York: Plenum Press
- Galler, H. P., Ott, N., 1994: Das dynamische Mikrosimulationsmodell des Sonderforschungsbereich 3. In: Hauser, R., Ott, N., Wagner, G. (Hg.) *Mikroanalytische Grundlagen der Gesellschaftspolitik. Ergebnisse aus dem gleichnamigen Sonderforschungsbereich an den Universitäten Frankfurt und Mannheim*, Bd. 2, S. 399-427. Berlin: Akademie Verlag
- Galton, F., Watson, H. W., 1874: On the Probability of the Extinction of Families. *Journal of the Royal Anthropological Institute* 4: 138-144
- Galton, F., Watson, H. W., 1977: On the Probability of the Extinction of Families. In: Smith, D., Keyfitz, N. (Hg.) *Mathematical Demography. Selected Papers*, S. 399-406. New York: Springer
- Göbel, J., Krause, P., Pischner, R., Sieber, I., Wagner, G. G., 2008: Daten- und Datenbankstruktur der Längsschnittstudie Sozio-oekonomisches Panel (SOEP). SOEPPapers on Multidisciplinary Panel Research 89
- Gilbert, J. P., Hammel, E. A., 1966: Computer Simulation and Analysis of Problems in Kinship and Social Structure. *American Anthropologist* 68: 71-93
- Gilbert, N., Troitzsch, K. G., 1999: *Simulation for the Social Scientist*. Buckingham: Open University Press
- Gill, R. D., 1992: Multistate Life Table and Regression Models. *Mathematical Population Studies* 3: 259-276
- Girosi, F., King, G., 2007: Understanding the Lee-Carter Mortality Forecasting Method
- Goldman, N., 1986: Effects of Mortality Levels on Kinship. In: *Consequences of Mortality Trends and Differentials*, S. 79-87. New York: United Nations
- Goldstein, J. R., 1996: *The Demography of Family and Kinship in an Age of Divorce and Remarriage*. Dissertation, University of California, Berkeley
- Goldstein, J. R., Kreyenfeld, M., 2011: Has East Germany Overtaken West Germany? Recent Trends in Order-Specific Fertility. *Population and Development Review* 37: 453-472
- Goldstein, J. R., Lutz, W., Pflug, G., 1994: Estimating the Uncertainty in Population Projections by Resampling Methods. IIASA Working Paper WP-94-106
- Goldstein, J. R., Lutz, W., Testa, M. R., 2003: The Emergence of Sub-Replacement Family Size Ideals in Europe. *European Demographic Research Papers* 2
- Goodman, L. A., Keyfitz, N., Pullum, T. W., 1974: Family Formation and the Frequency of Various Kinship Relations. *Theoretical Population Biology* 5: 1-27
- Goodman, L. A., Keyfitz, N., Pullum, T. W., 1975: Addendum: Family Formation and the Frequency of Various Kinship Relations. *Theoretical Population Biology* 8: 376-381
- Grünheid, E., Scharein, M., 2011: On Developments in the Mean Joint Lifetimes of Three- and Four-Generation Families in Western and Eastern Germany – A Model Calculation. *Comparative Population Studies* 36: 41-76
- Grummer-Strawn, L., Espenshade, T. J., 1991: Evaluating the Accuracy of U.S. Population Projection Models. In: Citro, C. F., Hanushek, E. A. (Hg.) *Improving Information for Social Policy Decisions. The Use of Microsimulation Modeling*, S. 305-329. Washington: National Academy Press
- Haber Kern, K., 2009: *Pflege in Europa. Familie und Wohlfahrtsstaat*. VS Verlag für Sozialwissenschaft
- Haber Kern, K., Szydlik, M., 2008: Pflege der Eltern – Ein europäischer Vergleich. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60: 78-101

- Haberman, S., Renshaw, A., 2011: A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics* 48: 35-55
- Haccou, P., Jagers, P., Vatutin, V. A., 2005: *Branching Processes. Variation, Growth, and Extinction of Populations*. Cambridge University Press
- Hackmann, T., Moog, S., 2008: Pflege im Spannungsfeld von Angebot und Nachfrage. Forschungszentrum Generationenverträge Diskussionsbeiträge 33
- Hagemann, T., Haidl, G., 2001: Veränderungen der männlichen Fertilität im Alter. *Reproduktionsmedizin* 17: 254-260
- Hammel, E. A., 2005: Demographic dynamics and kinship in anthropological populations. *Proceedings of the National Academy of Sciences* 102: 2248-2253
- Hammel, E. A., Hutchinson, D. W., Wachter, K. W., Lundy, R. T., Deuel, R. Z., 1976: *The SOCSIM Demographic-Sociological Microsimulation Program. Operating Manual*. Research Series Nr. 27. Berkeley: Institute for International Studies
- Hammel, E. A., Mason, C., Wachter, K., Wang, F., Yang, H., 1991: Rapid population change and kinship: The effects of unstable demographic changes on Chinese kinship networks, 1750-2250. In: Tapinos, G. (Hg.) *Consequences of rapid population growth in developing countries*, S. 243-271. New York: Taylor & Francis
- Hammel, E. A., Wachter, K. W., McDaniel, C. K., 1981: The Kin of the Aged in A.D. 2000: The Chickens Come Home to Roost. In: Kiesler, S. B., Morgan, J. N., Oppenheimer, V. K. (Hg.) *Aging. Social Change*, S. 11-39. New York: Academic Press
- Hank, K., Buber, I., 2009: Grandparents Caring for their Grandchildren: Findings from the 2004 Survey of Health, Ageing, and Retirement in Europe. *Journal of Family Issues* 30: 53-73
- Hank, K., Tölke, A., 2005: Das vernachlässigte Geschlecht in der Familienforschung: Untersuchungen zu Partnerschaft und Elternschaft bei Männern. In: Tölke, A., Hank, K. (Hg.) *Männer – das „vernachlässigte“ Geschlecht in der Familienforschung*, S. 7-17. Wiesbaden: VS Verlag für Sozialwissenschaft
- Hannan, M. T., 1984: Multistate Demography and Event History Analysis. In: Diekmann, A., Mitter, P. (Hg.) *Stochastic Modeling of Social Processes*, S. 39-88. Orlando: Academic Press
- Hanslúvka, H., 1976: Mortality and the life cycle of the family. Some implications of recent research. *World health statistics report* 29: 220-227
- Härdle, W., Mysickova, A., 2009: Stochastic Population Forecast for Germany and its Consequence for the German Pension System. SFB 649 Discussion Paper 2009-009
- Harris, T. E., 1963: *The Theory of Branching Processes*. Berlin: Springer
- Heckman, J. J., Singer, B., 1982: Population Heterogeneity in Demographic Models. In: Land, K. C., Rogers, A. (Hg.) *Multidimensional Mathematical Demography*, S. 567-599. New York: Academic Press
- Heilig, G., 1985: Die Heiratsneigung lediger Frauen in der Bundesrepublik Deutschland: 1950-1984. *Zeitschrift für Bevölkerungswissenschaft* 11: 519-547
- Helberger, C., Wagner, G., 1994: Die Längsschnitt-Version des Sonderforschungsbereich 3-Mikrosimulationsmodells. In: Hauser, R., Ott, N., Wagner, G. (Hg.) *Mikroanalytische Grundlagen der Gesellschaftspolitik. Ergebnisse aus dem gleichnamigen Sonderforschungsbereich an den Universitäten Frankfurt und Mannheim*, Bd. 2, S. 443-452. Berlin: Akademie Verlag
- Herlofson, K., Hagestad, G. O., 2011: Challenges in moving from macro to micro: Population and family structures in ageing societies. *Demographic Research* 25: 337-370
- Hill, P. B., Kopp, J., 2006: *Familiensoziologie. Grundlagen und theoretische Perspektiven*. Wiesbaden: VS, 4. Aufl.
- Hoem, J. M., 1977: A Markov chain model of working life tables. *Scandinavian Actuarial Journal* 1977: 1-20
- Hoem, J. M., Jensen, F. U., 1982: Multistate Life Table Methodology: A Probabilist Critique. In: Land, K. C., Rogers, A. (Hg.) *Multidimensional Mathematical Demography*, S. 155-264. New York: Academic Press
- Hoem, J. M., Madsen, D., Nielsen, J. L., Ohlsen, E.-M., Hansen, H. O., Rennermalm, B., 1981: Experiments in Modelling Recent Danish Fertility Curves. *Demography* 18: 231-244

- Hoff, A., 2007: Patterns of intergenerational support in grandparent-grandchild and parent-child relationships in Germany. *Ageing & Society* 27: 643-665
- Hofferth, S. L., 1987: Recent trends in living arrangements of children: a cohort life table analysis. In: Bongaarts, J., Burch, T. K., Wachter, K. W. (Hg.) *Family Demography: Methods and their Application*, S. 168-188. Oxford: Clarendon Press
- Hofmann, H., 2006: Altenpflege und Soziale Dienste – wachsender Bedarf bei fehlenden Mitteln. Ifo Schnelldienst 15/2006
- Höhn, C., 1982: *Der Familienzyklus – Zur Notwendigkeit einer Konzepterweiterung*. Boppard a. R.: Boldt
- Howell, N., Lehotay, V. A., 1978: AMBUSH: A Computer Program for Stochastic Microsimulation of Small Human Populations. *American Anthropologist* 80: 905-922
- Hsieh, J. J., 1991: Construction of Expanded Continuous Life Tables – A Generalization of Abridged and Complete Life Tables. *Mathematical Biosciences* 103: 287-302
- Huinink, J., 1988: Die demographische Analyse der Geburtenentwicklung mit Lebensverlaufsdaten. *Allgemeines Statistisches Archiv* 72: 359-377
- Huinink, J., Konietzka, D., 2007: *Familiensoziologie. Eine Einführung*. Frankfurt a. M.: Campus
- Huinink, J., Tuma, N. B., 1987: Post-War Fertility Patterns in the Federal Republic of Germany. In: Mayer, K. U. (Hg.) *Applications of Event History Analysis in Life Course Research*, S. 510-531. Berlin: Max-Planck-Institut für Bildungsforschung
- Hull, D. M., 1998: A Reconsideration of Galton's Problem. (Using a Two-Sex Population). *Theoretical Population Biology* 54: 105-116
- Hullen, G., 2002: Modellrechnungen der Lebensformen, Haushalts- und Familienstrukturen. In: Flöthmann, E.-J., Scholz, R., Gärtner, K. (Hg.) *Demographische Vorausschätzungen – Grenzen und Möglichkeiten, Methoden und Ziele*, Materialien zur Bevölkerungswissenschaft 104, S. 79-106. Wiesbaden: Bundesinstitut für Bevölkerungsforschung
- Hullen, G., 2004: Makrosimulationen der Lebensformen und Haushalte – Mehrzustandstafeln als Grundlage demographischer Szenarien. In: Hartard, S., Stahmer, C. (Hg.) *Analyse von Lebenszyklen*, S. 167-182. Wiesbaden: Statistisches Bundesamt
- Hullen, G., 2006: Lebensformen der Kinder. Makrosimulationen bis zum Jahr 2030. Expertise für die Familienberichtscommission, abrufbar unter: [www.g-hullen.de/manuskripte/lebensformen\\_4.pdf](http://www.g-hullen.de/manuskripte/lebensformen_4.pdf), recherchiert am: 01.07.2009
- Hullen, G., Hübner, J., Loichinger, E., 2004: Vorausberechnung der Brandenburger Haushalte mit ProfAmy. *Daten und Analysen* 1/2004: 50-53
- Hummel, D., 2000: *Der Bevölkerungsdiskurs. Demographisches Wissen und politische Macht*. Opladen: Leske + Budrich
- Hyndman, R. J., Booth, H., 2008: Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting* 24: 323-342
- Imhoff, E. v., 1992: A General Characterization of Consistency Algorithms in Multidimensional Projection Models. *Population Studies* 46: 159-169
- Imhoff, E. v., 1995: LIPRO: A Multistate Household Projection Model. In: Imhoff, E. v., Kuijsten, A., Hooimeijer, P., Wissen, L. v. (Hg.) *Household Demography and Household Modeling*, S. 273-291. New York: Plenum Press
- Imhoff, E. v., Keilmann, N., 1991: *LIPRO 2.0: An Application of a Dynamic Demographic Projection Model to Household Structure in the Netherlands*. Amsterdam: Swets Zeitlinger
- Imhoff, E. v., Post, W., 1998: Microsimulation Methods for Population Projection. *Population* 10: 97-138
- Jagers, P., 1982: How Probable is it to be First Born? and Other Branching-Process Applications to Kinship Problems. *Mathematical Biosciences* 59: 1-15
- Jakoby, N., 2008a: Diversität des Familien- und Verwandtschaftslebens – Zur Bedeutung des erweiterten Familienkreises. *Zeitschrift für Bevölkerungswissenschaft* 33: 207-228
- Jakoby, N., 2008b: *(Wahl-)Verwandtschaft. Zur Erklärung verwandtschaftlichen Handelns*. Wiesbaden: VS
- Jiang, L., 1995: Changing Kinship Structure and its Implications for Old-Age Support in Urban and Rural China. *Population Studies* 49: 127-145

- Joffe, A., Waugh, W. A. O., 1982: Exact Distributions of Kin Numbers in a Galton-Watson Process. *Journal of Applied Probability* 19: 767-775
- Joffe, A., Waugh, W. A. O., 1985: The Kin Number Problem in a Multitype Galton-Watson Population. *Journal of Applied Probability* 22: 37-47
- Joffe, A., Waugh, W. A. O., 1986: Exact Distribution of Kin Numbers in Multitype Galton-Watson Process. In: Jansen, J. (Hg.) *Semi-Markov Models. Theory and Applications*, S. 397-405. New York: Plenum Press
- Kalmijn, M., De Vries, J., 2009: Change and Stability in Parent-Child Contact in Five Western Countries. *European Journal of Population* 25: 257-276
- Karmel, P. H., 1947: The Relations between Male and Female Reproduction Rates. *Population Studies* 1: 249-260
- Keilman, N., 1985: Internal and external consistency in multidimensional population projection models. *Environment and Planning* 17: 1473-1498
- Keilman, N., 1995: Household Concepts and Household Definitions in Western Europe: Different Levels But Similar Trends in Household Developments. In: Imhoff, E. v., Kuijsten, A., Hooimeijer, P., Wissen, L. v. (Hg.) *Household Demography and Household Modeling*, S. 111-135. New York: Plenum Press
- Keilman, N., Pham, D. Q., 2000: Predictive Intervals for Age-Specific Fertility. *European Journal of Population* 16: 41-60
- Keilman, N., Pham, D. Q., 2004: Time Series Based Errors and Empirical Errors in Fertility Forecasts in the Nordic Countries. *International Statistical Review* 72: 5-18
- Keilman, N., Pham, D. Q., Hetland, A., 2002: Why population forecasts should be probabilistic – illustrated by the case of Norway. *Demographic Research* 6: 409-454
- Keim, S., Klärner, A., Bernardi, L., 2009: Who is relevant? Exploring fertility relevant social networks. MPIDR Working Paper WP2009-001
- Kendall, D. G., 1975: The Genealogy of Genealogy. Branching Processes Before (And After) 1873. *Bulletin of the London Mathematical Society* 7: 225-253
- Kertzner, D. I., 1983: Generation as a Sociological Problem. *Annual Review of Sociology* 9: 125-149
- Keyfitz, N., 1981: The Limits of Population Forecasting. *Population and Development Review* 7: 579-593
- Keyfitz, N., 1985: A probability representation of future population. *Zeitschrift für Bevölkerungswissenschaft* 11: 179-191
- Keyfitz, N., 1986: Canadian Kinship Patterns Based on 1971 and 1981 Data. *Canadian Studies in Population* 13: 123-150
- Keyfitz, N., 1989: Measuring in Advance the Accuracy of Population Forecasts. Working Paper 89-72, IIASA
- Keyfitz, N., Caswell, H., 2005: *Applied Mathematical Demography*. New York: Springer, 3. Aufl.
- Kimmel, M., Axelrod, D. E., 2002: *Branching Processes in Biology*. New York: Springer
- Kitsul, P., Philipov, D., 1982: High- and Low-Intensity Model of Mobility. In: Land, K. C., Rogers, A. (Hg.) *Multidimensional Mathematical Demography*, S. 505-534. New York: Academic Press
- Kleimann, E., 1967: A Standardized Dependency Ratio. *Demography* 4: 876-893
- Klein, T., 1989: Bildungsexpansion und Geburtenrückgang. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 41: 465-482
- Klein, T., Lengsfeld, W., 1985: Sozialstrukturelle Ursachen des Geburtenrückgangs. *Zeitschrift für Bevölkerungswissenschaft* 11: 57-74
- Klenke, A., 2008: *Wahrscheinlichkeitstheorie*. Berlin: Springer, 2. Aufl.
- Kohli, M., Künemund, H., Motel, A., Szydlik, M., 2000: Generationenbeziehungen. In: Kohli, M., Künemund, H. (Hg.) *Die zweite Lebenshälfte. Gesellschaftliche Lage und Partizipation im Spiegel des Alters-Survey*, S. 176-211. Opladen: Leske + Budrich
- Koissia, M.-C., Shapiro, A. F., Högnäsa, G., 2006: Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics* 38: 1-20
- Kolle, C., 2006: Rentnerboom und Babynotstand – Der demographische Wandel in Deutschland. Bonn: Informationszentrum Sozialwissenschaft

- Kreyenfeld, M., Konietzka, D., 2004: Angleichung oder Verfestigung von Differenzen? Geburtenentwicklung und Familienformen in Ost- und Westdeutschland. MPIDR Working Paper 2004-025
- Kreyenfeld, M., Konietzka, D., 2007: Kinderlosigkeit in Deutschland – theoretische Probleme und empirische Ergebnisse. In: Konietzka, D., Kreyenfeld, M. (Hg.) *Ein Leben ohne Kinder. Kinderlosigkeit in Deutschland*, S. 11-41. Wiesbaden: VS Verlag für Sozialwissenschaft
- Kreyenfeld, M., Konietzka, D., 2010: Nichteheleliche Geburten. In: Goldstein, J., Kreyenfeld, M., Huinink, J., Konietzka, D., Trappe, H. (Hg.) *Familie und Partnerschaft in Ost- und Westdeutschland*, S. 8-9. Rostock: Max-Planck-Institut für Demographische Forschung
- Kreyenfeld, M., Konietzka, D., Hornung, A., 2009: Family Diversity in France, Russian Federation, East and West Germany: Overview on Living Arrangements and Living Conditions. In: United Nations Economic Commission for Europe (Hg.) *How Generations and Gender Shape Demographic Change. Towards Policies Based on Better Knowledge*, S. 33-56. New York & Genf: United Nations
- Kreyenfeld, M., Zeman, K., Burkimsher, M., Jaschinski, I., 2012: Fertility Data for German-Speaking Countries: What is the Potential? Where are the Pitfalls? *Comparative Population Studies* im Erscheinen
- Krishnamoorthy, S., 1979: Family Formation and the Life Cycle. *Demography* 16: 121-129
- Kroh, M., 2010: Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2009). DIW Data Documentation 50
- Kunstadter, P., Buhler, R., Stephan, F. F., Westoff, C. F., 1963: Demographic Variability and Preferential Marriage Patterns. *American Journal of Physical Anthropology* 21: 511-519
- Kurtz, T. G., 2005: *Branching Processes. Variation, Growth, and Extinction of Populations*, Kap. Large Populations, S. 82-88. Cambridge: Cambridge University Press
- Land, K. C., Rogers, A., 1982: Multidimensional Mathematical Demography: An Overview. In: Land, K. C., Rogers, A. (Hg.) *Multidimensional Mathematical Demography*, S. 1-41. New York: Academic Press
- Lange, A., Lauterbach, W., 1998: Aufwachsen mit oder ohne Großeltern? Die gesellschaftliche Relevanz multilokaler Mehrgenerationenfamilien. *Zeitschrift für Soziologie der Erziehung und Sozialisation* 18: 227-249
- Lauterbach, W., 1995: Die gemeinsame Lebenszeit von Familiengenerationen. *Zeitschrift für Soziologie* 24: 22-41
- Lauterbach, W., 2002: Großelternschaft und Mehrgenerationenfamilien - soziale Realität oder demographischer Mythos? *Zeitschrift für Gerontologie und Geriatrie* 35: 540-555
- Le Bras, H., 1973: Parents, grands-parents, bisaïeux. *Population* 28: 9-38
- Le Bras, H., Wachter, K. W., 1978: *Statistical Studies of Historical Social Structure*, Kap. 10: Living Forbearers in Stable Populations, S. 163-188. New York: Academic Press
- Lee, R. D., 1993: Modeling and forecasting the time series of US fertility: Age distribution, range, and ultimate level. *International Journal of Forecasting* 9: 187-202
- Lee, R. D., 1998: Probabilistic Approaches to Population Forecasting. *Population and Development Review* 24 (Supplement): 156-190
- Lee, R. D., 2000: The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications. *North American Actuarial Journal* 4: 80-93
- Lee, R. D., Carter, L. R., 1992: Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association* 87: 659-671
- Lee, R. D., Miller, T., 2001: Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality. *Demography* 38: 537-549
- Lee, R. D., Tuljapurkar, S., 1994: Stochastic Population Forecasts for the United States: Beyond High, Medium, and Low. *Journal of the American Statistical Association* 89: 1175-1189
- Lee, R. D., Tuljapurkar, S., 1998: Uncertain Demographic Futures and Social Security Finances. *American Economic Review* 88: 237-241
- Lee, Y.-J., Palloni, A., 1992: Changes in the Family Status of Elderly Women in Korea. *Demography* 29: 69-92
- Leslie, P. H., 1945: On the use of matrices in certain population mathematics. *Biometrika* 35: 183-212

- Lesthaeghe, R., 1992: Der zweite demographische Übergang in den westlichen Ländern: Eine Deutung. *Zeitschrift für Bevölkerungswissenschaft* 18: 313-354
- Lesthaeghe, R., 2010: The Unfolding Story of the Second Demographic Transition. *Population and Development Review* 36: 211-251
- Levy, M., Westoff, C., 1965: Simulation of Kinship Systems. *New Scientist* 27: 571,572
- Li, H., Maddala, G. S., 1996: Bootstrapping Time Series Models. *Econometric Reviews* 15: 115-158
- Li, J. S.-H., Chan, W.-S., Cheung, S.-H., 2011: Structural Changes in the Lee-Carter Mortality Indexes: Detection and Implications. *North American Actuarial Journal* 15: 13-31
- Li, N., Lee, R. D., 2005: Coherent Mortality Forecasts for a Group of Populations: An Extension of the Lee-Carter Method. *Demography* 42: 575-594
- Li, N., Wu, Z., 2003: Forecasting Cohort Incomplete Fertility: A Method and an Application. *Population Studies* 57: 303-320
- Lipps, O., Betz, F., 2005: Stochastische Bevölkerungsprojektion für West- und Ostdeutschland. *Zeitschrift für Bevölkerungswissenschaft* 30: 3-42
- Liu, X., Braun, J. W., 2010: Investigating Mortality Uncertainty Using the Block Bootstrap. *Journal of Probability and Statistics* Volume 2010: 1-15
- Lotka, A. J., 1931: Orphanhood in Relation to Demographic Factors. A Study in Population Analysis. *Metron* 9: 37-109
- Lowenstein, A., 2005: Global Ageing and Challenges to Families. In: Johnson, M. J., Bengtson, V. L., Coleman, P. G., Kirkwood, T. B. L. (Hg.) *The Cambridge Handbook of Age and Ageing*, S. 403-412. Cambridge: Cambridge University Press
- Lucke, D., 1998: Verwandtschaft im Recht – Rechtssoziologische Aspekte verwandtschaftlicher Beziehungen. In: Wagner, M., Schütze, Y. (Hg.) *Verwandtschaft. Sozialwissenschaftliche Beiträge zu einem vernachlässigten Thema*, S. 59-89. Stuttgart: Enke
- Lundström, H., Qvist, J., 2004: Mortality Forecasting and Trend Shifts: an Application of the Lee-Carter Model to Swedish Mortality Data. *International Statistical Review* 72: 37-50
- Lüscher, K., Pillemer, K., 1998: Intergenerational Ambivalence: A New Approach to the Study of Parent-Child Relations in Later Life. *Journal of Marriage and Family* 60: 413-425
- Lutz, W., Goldstein, J., Prinz, C., 1996a: Alternative Approaches to Population Projection. In: Lutz, W. (Hg.) *The Future Population of the World. What Can We Assume Today?*, S. 14-44. Laxenburg: IIASA, 2. Aufl.
- Lutz, W., Goujon, A., 2001: The World's Changing Human Capital Stock: Multi-State Population Projections by Educational Attainment. *Population and Development Review* 27: 323-339
- Lutz, W., Sanderson, W., Scherbov, S., 1996b: Probabilistic Population Projections Based on Expert Opinion. In: Lutz, W. (Hg.) *The Future Population of the World. What Can We Assume Today?*, S. 397-428. Laxenburg: IIASA, 2. Aufl.
- Lutz, W., Sanderson, W., Scherbov, S., 1997: Doubling of world population unlikely. *Nature* 387: 803-805
- Lutz, W., Scherbov, S., 1998: Probabilistische Bevölkerungsprognosen für Deutschland. *Zeitschrift für Bevölkerungswissenschaft* 23: 83-109
- Luy, M., 2004: Verschiedene Aspekte der Sterblichkeitsentwicklung in Deutschland von 1950 bis 2000. *Zeitschrift für Bevölkerungswissenschaft* 29: 3-62
- Maaouia, F., Touati, A., 2005: Identification of Multi-type Branching Processes. *Annals of Statistics* 33: 2655-2694
- Mai, R., 2005: Demographische Alterung in Deutschland. Die Entwicklung von 1871 bis 2050 und der Einfluss von Sterblichkeit und Zuwanderung auf die Alterung. *Zeitschrift für Bevölkerungswissenschaft* 30: 43-80
- Marbach, J. H., 2008: Die Familie als Forschungsthema – ein Auslaufmodell? Zur Aktualität des Familienbegriffs in der Familienforschung und Ansätze zu einer Neufassung. In: Bien, W., Marbach, J. H. (Hg.) *Familiale Beziehungen. Familienalltag und soziale Netzwerke. Ergebnisse der drei Wellen des Familiensurvey*, S. 13-41. Wiesbaden: VS Verlag für Sozialwissenschaft
- Marschalck, P., 1984: *Bevölkerungsgeschichte Deutschlands im 19. und 20. Jahrhundert*. Frankfurt a. M.: Suhrkamp

- Martin, L. G., Culter, S., 1983: Mortality Decline and Japanese Family Structure. *Population and Development Review* 9: 633-649
- McNeil, D., Trussel, T. J., Turner, J. C., 1977: Spline Interpolation of Demographic Data. *Demography* 14: 242-252
- Meintrup, D., Schäffler, S., 2005: *Stochastik. Theorie und Anwendungen*. Berlin: Springer
- Menken, J. A., 1985: Age and Fertility: How Late Can You Wait? *Demography* 22: 469-483
- Merz, J., 1994: Statisches Sonderforschungsbereich 3 Mikrosimulationsmodell – Mainframe und PC-Version. In: Hauser, R., Ott, N., Wagner, G. (Hg.) *Mikroanalytische Grundlagen der Gesellschaftspolitik. Ergebnisse aus dem gleichnamigen Sonderforschungsbereich an den Universitäten Frankfurt und Mannheim*, Bd. 2, S. 380-398. Berlin: Akademie Verlag
- Mode, C. J., 1971: *Multitype Branching Processes. Theory and Applications*. New York: Elsevier
- Mode, C. J., 1982: Increment-Decrement Life Tables and Semi-Markovian Processes from a Sample Path Perspective. In: Land, K. C., Rogers, A. (Hg.) *Multidimensional Mathematical Demography*, S. 535-565. New York: Academic Press
- Mueller, U., 1993: *Bevölkerungsstatistik und Bevölkerungsdynamik*. Berlin: de Gruyter
- Mueller, U., 2000: Die Maßzahlen der Bevölkerungsstatistik. In: Mueller, U. (Hg.) *Handbuch der Demographie*. Berlin: Springer
- Münz, R., Ulrich, R., 2001: Bevölkerung. In: Joas, H. (Hg.) *Lehrbuch der Soziologie*. Frankfurt: Campus
- Murphy, M., 1999: Is the relationship between fertility of parents and children really weak? *Social Biology* 46: 122-145
- Murphy, M., 2004: Tracing Very Long-Term Kinship Networks Using SOCSIM. *Demographic Research* 10: 171-196
- Murphy, M., 2008: Variations in Kinship Networks Across Geographic and Social Space. *Population And Development Review* 34: 19-49
- Murphy, M., Martikainen, P., Penne, S., 2006: Demographic change and the supply of potential family supporters in Britain, Finland and France in the period 1911-2050. *European Journal of Population* 22: 219-240
- Murphy, M., Wang, D., 2001: Family-Level Continuities in Childbearing in Low-Fertility Societies. *European Journal of Population* 17: 75-96
- Nakamura, A., Nakamura, M., 1978: On Microanalytic Simulation and its Applications in Population Projection. *Journal of the Operational Research Society* 29: 349-360
- Namoodiri, K., Suchindran, C. M., 1987: *Life Table Techniques and Their Applications*. Orlando: Academic Press
- Narrol, R., 1958: German Kinship Terms. *American Anthropologist* 60: 750-755
- Nauck, B., Suckow, J., 2006: Intergenerational Relationships in Cross-Cultural Comparison: How Social Networks Frame Intergenerational Relations Between Mothers and Grandmothers in Japan, Korea, China, Indonesia, Israel, Germany and Turkey. *Journal of Family Issues* 27: 1159-1185
- Nave-Herz, R., 1994: *Familie heute. Wandel der Familienstrukturen und Folgen für die Erziehung*. Darmstadt: Wissenschaftliche Buchgesellschaft
- Nave-Herz, R., 2001: Die Mehrgenerationen-Familie – Eine soziologische Analyse. In: Walper, S., Pekrun, R. (Hg.) *Familie und Entwicklung. Aktuelle Perspektiven der Familienpsychologie*, S. 21-35. Göttingen: Hogrefe
- Nave-Herz, R., Feldhaus, M., 2005: Geschwisterbeziehungen. In: Busch, F. W., Nave-Herz, R. (Hg.) *Familie und Gesellschaft. Beiträge zur Familienforschung*, S. 111-123. Oldenbourg: BIS
- Nerman, O., Jagers, P., 1984: The Stable Doubly Infinite Pedigree Process of Supercritical Branching Populations. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 65: 445-460
- Neyer, F. J., Lang, F. R., 2003: Blood Is Thicker Than Water: Kinship Orientation Across Adulthood. *Journal of Personality and Social Psychology* 84: 310-321
- Ni Bhrolchain, M., 1992: Period Paramount? A Critique of the Cohort Approach to Fertility. *Population and Development Review* 18: 599-629
- O'Neill, B. C., 2004: Conditional Probabilistic Population Projections: An Application to Climate Change. *International Statistical Review* 72: 167-184

- Onnen-Isemann, C., 2003: Kinderlose Partnerschaften. In: Bien, W., Marbach, J. (Hg.) *Partnerschaften und Familiengründung: Ergebnisse der dritten Welle des Familien-Survey*, S. 97-137. Opladen: Leske + Budrich
- Orcutt, G., 1957: A New Type of Socio-Economic System. *Review of Economics and Statistics* 39: 116-123
- Orcutt, G., Caldwell, S., Wertheimer, R., 1976: *Policy Exploration Through Microanalytic Simulation*. Washington, D. C.: Urban Institute
- Orcutt, G. H., Greenberger, M., Korbel, J., Rivlin, A. M., 1965: *Microanalysis of Socioeconomic Systems: A Simulation Study*. New York: Harper & Row
- Ott, N., 1986a: Bevölkerungsentwicklung bis zum Jahr 2000 – Modellrechnungen mit der Version 86.0 des Mikrosimulationsmodells des Sfb 3. Sfb 3 Arbeitspapier 212
- Ott, N., 1986b: Die Fortschreibung demographischer Prozesse mit dem Mikrosimulationsmodell des Sfb 3, Version 86.0. Sfb 3 Arbeitspapier 213
- Ott, N., 1995: The Use of Panel Data in the Analysis of Household Structure. In: Imhoff, E. v., Kuijsten, A., Hooimeijer, P., Wissen, L. v. (Hg.) *Household Demography and Household Modeling*, S. 163-183. New York: Plenum Press
- Pannenberg, M., Pischner, R., Rendtel, U., Spiess, M., Wagner, G. G., 2005: Sampling and Weighting. In: Haisken-DeNew, J. P., Frick, J. R. (Hg.) *Desktop Companion to the German Socio-Economic Panel (SOEP)*, S. 153-186
- Pflaumer, P., 1984: Die Berücksichtigung von Unsicherheit bei der zukünftigen Entwicklung der Bevölkerung und der Rentenbeitragssätze in der Bundesrepublik Deutschland. *Zeitschrift für Bevölkerungswissenschaft* 10: 501-530
- Pflaumer, P., 1986: Stochastische Bevölkerungsmodelle zur Analyse der Auswirkungen demographischer Prozesse auf die Systeme der sozialen Sicherung. *Allgemeines Statistisches Archiv* 70: 57-74
- Pflaumer, P., 1988: Confidence Intervals for Population Projections Based on Monte Carlo Methods. *International Journal of Forecasting* 4: 135-142
- Pflaumer, P., 1992: Forecasting US population totals with the Box-Jenkins approach. *International Journal of Forecasting* 8: 329-338
- Pischner, R., 2007: Die Querschnittsgewichtung und die Hochrechnungsfaktoren des Sozio-oekonomischen Panels (SOEP) ab Release 2007 (Welle W). DIW Data Documentation 22
- Pollard, J. H., 1966: On the Use of the Direct Matrix Product in Analysing Certain Stochastic Population Models. *Biometrika* 53: 397-415
- Pollard, J. H., 1968: A Note on Multi-Type Galton-Watson Processes with Random Branching Probabilities. *Biometrika* 55: 589-590
- Post, W., Poppel, F., v. Imhoff, E., Kruse, E., 1997: Reconstructing the extended kin-network in the Netherlands with genealogical data: Methods, problems and results. *Population Studies* 51: 263-278
- Poston, D. L. j., 2006: Age and Sex. In: Poston, D. L. j., Micklin, M. (Hg.) *Handbook of Population*, S. 19-58. New York: Springer
- Poston, D. L. j., Baumle, A. K., Micklin, M., 2006: Epilogue: Needed Research in Demography. In: Poston, D. L. j., Micklin, M. (Hg.) *Handbook of Population*, S. 853-881. New York: Springer
- Preston, S. H., 1976: Family Sizes of Children and Family Sizes of Women. *Demography* 13: 105-114
- Preston, S. H., 1987: Estimation of Certain Measures in Family Demography Based upon Generalized Stable Population Relations. In: Bongaarts, J., Burch, T. K., Wachter, K. W. (Hg.) *Family Demography. Methods and their Application*, S. 40-62. Oxford: Clarendon Press
- Preston, S. H., Heuveline, P., Guillot, M., 2001: *Demography. Measuring and Modeling Population Processes*. Oxford: Blackwell
- Pötzsch, O., 2010: Cohort Fertility: A Comparison of the Results of the Official Birth Statistics and of the Microcensus Survey 2008. *Comparative Population Studies* 35: 185-204
- Pullum, T. W., 1980: Separating Age, Period, and Cohort Effects in White U.S. Fertility, 1920-1970. *Social Science Research* 9: 225-244
- Pullum, T. W., 1982: The Eventual Frequencies of Kin in a Stable Population. *Demography* 19: 549-565
- Pullum, T. W., Wolf, D. A., 1991: Correlations Between Frequencies of Kin. *Demography* 28: 391-409

- Putter, H., Fiocco, M., Geskus, R. B., 2007: Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine* 26: 2389-2430
- Puur, A., Sakkeus, L., Poldma, A., A., H., 2011: Intergenerational family constellations in contemporary Europe: Evidence from the Generations and Gender Survey. *Demographic Research* 25: 135-172
- Reeves, J. H., 1987: Projections of Numbers of Kin. In: Bongaarts, J., Burch, T. K., Wachter, K. W. (Hg.) *Family Demography. Methods and their Applications*, S. 228-248. Oxford: Clarendon Press
- Rendall, M. S., Clarke, L., Peters, E. H., Ranjit, N., Verropoulou, G., 1999: Incomplete Reporting of Men's Fertility in the United States and Britain. *Demography* 36: 135-144
- Renshaw, A., Haberman, S., 2003: Lee-Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales mortality projections. *Applied Statistics* 52: 119-137
- Rinne, H., 1994: *Wirtschafts- und Bevölkerungsstatistik*. München: Oldenbourg
- Rogers, A., 1969: On Perfect Aggregation in the Matrix Cohort-Survival Model of Interregional Population Growth. *Journal of Regional Science* 9: 417-424
- Rogers, A., 1975: *Introduction to Multiregional Mathematical Demography*. New York: Wiley
- Rogers, A., 1980: Introduction to multistate mathematical demography. *Environment and Planning* 12: 489-498
- Rogers, A., Ledent, J., 1976: Increment-Decrement Life Tables: A Comment. *Demography* 13: 287-290
- Rohwer, G., Pötter, U., 2003: *Demography of Germany. Concepts, Data, and Methods*. Bochum
- Rosenbaum, H., Timm, E., 2008: *Private Netzwerke im Wohlfahrtsstaat. Familie, Verwandtschaft und soziale Sicherheit im Deutschland des 20. Jahrhunderts*. Konstanz: UVK
- Ross-Strajhar, G., 2005: Alt und Jung in Deutschland – Sozialwissenschaftliche Generationenforschung. Bonn: Informationszentrum Sozialwissenschaft
- Rowland, D. T., 2003: *Demographic Methods and Concepts*. Oxford: Oxford University Press
- Ruggles, S., 1990: Family demography and family history. *Historical Methods* 23: 22-31
- Ruggles, S., 1993: Confessions of a Microsimulator: Problems in Modelling the Demography of Kinship. *Historical Methods* 26: 161-169
- Ryder, N. B., 1990: What is Going to Happen to American Fertility? *Population and Development Review* 16: 433-454
- Saboa, J. L. M., 1974: Modelling and Forecasting Populations by Time Series: The Swedish Case. *Demography* 11: 483-492
- Samuels, M. L., 1971: Distribution of the Branching Process Population among Generations. *Journal of Applied Probability* 8: 655-667
- Sanderson, W. C., 1995: Predictability, complexity, and catastrophe in a collapsible model of population, development, and environmental interactions. *Mathematical Population Studies* 5: 259-279
- Sanderson, W. C., Scherbov, S., 2007: A new perspective on population aging. *Demographic Research* 16: 27-58
- Sauer, L., Ette, A., 2007: Auswanderung aus Deutschland: Stand der Forschung und erste Ergebnisse zur internationalen Migration deutscher Staatsbürger. Materialien zur Bevölkerungswissenschaft 123
- Scharein, M., Unger, M., 2005: Kinderlosigkeit bei Akademikerinnen? Die Aussagekraft empirischer Daten zur Kinderlosigkeit bei Akademikerinnen. *BiB-Mitteilungen* 2/2005: 6-13
- Schmähl, W., Ulrich, V. (Hg.) 2001: *Soziale Sicherungssysteme und demographische Herausforderungen*. Tübingen: Mohr-Siebeck
- Schmitt, C., 2005: Kinderlosigkeit bei Männern – Geschlechtsspezifische Determinanten ausbleibender Elternschaft. In: Tölke, A., Hank, K. (Hg.) *Männer – „das vernachlässigte“ Geschlecht in der Familienforschung*, S. 18-43. Wiesbaden: VS Verlag für Sozialwissenschaft
- Schmitt, C., 2010: BIOBRTHM: The Birth Biography of Male Respondents in the SOEP. In: Frick, J. R., Lohmann, H. (Hg.) *Biography and Life History Data in the German Socio Economic Panel (SOEP, v26, 1984-2009)*, S. 78-82. DIW Data Documentation 52
- Schneekloth, U., 2006: Entwicklungstrends und Perspektiven in der häuslichen Pflege. Zentrale Ergebnisse der Studie Möglichkeiten und Grenzen selbstständiger Lebensführung (MuG III). *Zeitschrift für Gerontologie und Geriatrie* 39: 405-412
- Schneider, N. F., Rüger, H., 2007: Value of Marriage. Der subjektive Sinn der Ehe und die Entscheidung zur Heirat. *Zeitschrift für Soziologie* 36: 131-152

- Schnell, R., Trappmann, M., 2006: Konsequenzen der Panelmortalität im SOEP für Schätzungen der Lebenserwartung. Arbeitspapier 2/2006 Zentrum für Quantitative Methoden und Surveyforschung, Universität Konstanz
- Schoen, R., 1975: Constructing Increment-Decrement Life Tables. *Demography* 12: 313-324
- Schoen, R., 1988a: *Modeling Multigroup Populations*. New York: Plenum Press
- Schoen, R., 1988b: Practical Uses of Multistate Population Models. *Annual Review of Sociology* 14: 341-361
- Schoen, R., Land, K. C., 1979: A General Algorithm for Estimating a Markov-Generated Increment-Decrement Life Table with Applications to Marital-Status Patterns. *Journal of the American Statistical Association* 74: 761-776
- Schoen, R., Landale, N. S., Daniels, K., 2007: Family Transitions in Young Adulthood. *Demography* 44: 807-820
- Schoen, R., Nelson, V. N., 1974: Marriage, Divorce, and Mortality: A Life Table Analysis. *Demography* 11: 267-290
- Schoen, R., Woodrow, K., 1980: Labor Force Status Life Tables for the United States, 1972. *Demography* 17: 297-322
- Schulz, E., 2008: Weniger Menschen, aber Arbeitskräfteangebot bleibt bis 2025 stabil. DIW Wochenbericht 40/2008
- Schulz, E., 2010: Projection of Care Need and Family Resources in Germany. In: Doblhammer, G., Scholz, R. (Hg.) *Ageing, Care Need and Quality of Life. The Perspective of Care Givers and People in Need of Care*, S. 61-81. Wiesbaden: VS
- Schütze, Y., Wagner, M., 1998: Verwandtschaft – Begriff und Tendenzen in der Forschung. In: Wagner, M., Schütze, Y. (Hg.) *Verwandtschaft: sozialwissenschaftliche Beiträge zu einem vernachlässigten Thema*, S. 7-16. Stuttgart: Enke
- Sebald, E., 2008: *Das Zwei-Geschlechter-Problem bei Bevölkerungsprojektionen*. Dissertation, Ruhr-Universität Bochum
- Shryock, H. S., Siegel, J. S., 1976: *The Methods and Materials of Demography*. New York: Academic Press
- Siebert, H., 2003: Germany – an Immigration Country. Kiel Working Paper No. 1189
- Siegel, N. A., Huber, S., Gensicke, A., Geue, D., Stimmel, S., Stutz, F., 2010: SOEP 2009. Methodenbericht zum Befragungsjahr 2009 (Welle 26) des Sozio-oekonomischen Panels. TNS Infratest, München
- Singer, B., Cohen, J., 1980: Estimating Malaria Incidence and Recovery Rates from Panel Surveys. *Mathematical Biosciences* 49: 273-305
- Singer, B., Spilerman, S., 1976a: The Representation of Social Processes by Markov Models. *American Journal of Sociology* 82: 1-54
- Singer, B., Spilerman, S., 1976b: Some Methodological Issues in the Analysis of Longitudinal Surveys. *Annals of Economic and Social Measurement* 5: 447-474
- Smith, J. E., 1987: The Computer Simulation of Kin Sets and Kin Counts. In: Bongaarts, J., Burch, T. K., Wachter, K. W. (Hg.) *Family Demography. Methods and their Applications*, S. 249-266. Oxford: Clarendon Press
- Smith, J. E., Oeppen, J. E., 1993: Estimating Numbers of Kin in Historical England Using Demographic Microsimulation. In: Reher, D. S., Schofield, R. S. (Hg.) *Old and New Methods in Historical Demography*, S. 280-317. Oxford: Clarendon Press
- Spielauer, M., Vencatasawamy, C. P., 2003: FAMSIM: Dynamic Microsimulation of Life Course Interactions Between Education, Work, Partnership Formation and Birth in Austria, Belgium, Italy, Spain and Sweden. *Vienna Yearbook of Population Research* 2003: 143-164
- Statistisches Bundesamt (Hg.) 2003: *Bevölkerung Deutschlands bis 2050. 10. koordinierte Bevölkerungsvorausberechnung*. Wiesbaden
- Statistisches Bundesamt (Hg.) 2006: *Generationensterbetafeln für Deutschland. Modellrechnungen für die Geburtsjahrgänge 1871-2004*. Wiesbaden
- Statistisches Bundesamt (Hg.) 2008: *Demografischer Wandel in Deutschland. Auswirkungen auf Krankenhausbehandlungen und Pflegebedürftige in Bund und Ländern*. Wiesbaden

- Statistisches Bundesamt (Hg.) 2010a: *Bevölkerung in den Bundesländern, dem früheren Bundesgebiet und den neuen Ländern bis 2060. Ergebnisse der 12. koordinierten Bevölkerungsvorausberechnung*. Wiesbaden
- Statistisches Bundesamt (Hg.) 2010b: *Bevölkerung und Erwerbstätigkeit. Natürliche Bevölkerungsbewegung 2007*. Fachserie 1 Reihe 1.1. Wiesbaden
- Statistisches Bundesamt (Hg.) 2010c: *Bevölkerung Deutschlands bis 2060. 12. koordinierte Bevölkerungsvorausberechnung*. Wiesbaden
- Statistisches Bundesamt (Hg.) 2010d: *Qualitätsbericht: Statistik der Geburten*. Wiesbaden
- Statistisches Bundesamt (Hg.) 2011a: *Bevölkerungs- und Haushaltsentwicklung im Bund und in den Ländern*. Wiesbaden
- Statistisches Bundesamt (Hg.) 2011b: *Pflegestatistik 2009. Pflege im Rahmen der Pflegeversicherung. Deutschlandergebnisse*. Wiesbaden
- Statistisches Bundesamt (Hg.) 2011c: *Statistisches Jahrbuch 2011*. Wiesbaden
- Steinbach, A., 2008: Stieffamilien in Deutschland. Ergebnisse des „Generations and Gender Survey“ 2005. *Zeitschrift für Bevölkerungswissenschaft* 33: 153-180
- Stoto, M. A., 1983: The Accuracy of Population Projections. *Journal of the American Statistical Association* 78: 13-20
- Suzuki, T., 1996: A Kinship Model Based on Branching Process. *Journal of Population Problems* 52: 17-29
- Sykes, Z. M., 1969: Some Stochastic Versions of the Matrix Model for Population Dynamics. *Journal of the American Statistical Association* 64: 111-130
- Szydlik, M., 2011: Erben in Europa. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 63: 543-565
- Tabeau, E., 2001: A Review of Demographic Forecasting Models for Mortality. In: Tabeau, E., v. d. Berg Jeths, A., Heathcote, C. (Hg.) *Forecasting Mortality in Developed Countries. Insights from a Statistical, Demographic and Epidemiological Perspective*, S. 1-32. New York: Kluwer Academic
- Thompson, P. A., 1989: A Transformation Useful for Bounding a Forecast. *Statistics & Probability Letters* 8: 469-475
- Thompson, P. E., Bell, W. R., Long, J. F., Miller, R. B., 1989: Multivariate Time Series Projections of Parameterized Age-Specific Fertility Rates. *Journal of the American Statistical Association* 84: 689-699
- Tomassini, C., Wolf, D. A., 2000: Shrinking Kin Networks in Italy Due to Sustained Low Fertility. *European Journal of Population* 16: 353-372
- Tu, E. J.-C., Freedman, V. A., Wolf, D. A., 1993: Kinship and Family Support in Taiwan: A Microsimulation Approach. *Research on Aging* 15: 465-486
- Tu, E. J.-C., Liang, J., 1990: Fertility, Mortality, Kinship Structure and Its Aging Implications in Taiwan. In: {Zheng Yi} (Hg.) *Changing Family Structure and Population Aging in China. A Comparative Approach*, S. 214-238. Peking: Beijing University Press
- Tuljapurkar, S., 1986: Demography in stochastic environments, II: Growth and Convergence Rates. *Journal of Mathematical Biology* 24: 569-581
- Tuljapurkar, S., 1989: An Uncertain Life: Demography in Random Environments. *Theoretical Population Biology* 35: 227-294
- Tuljapurkar, S., 1992: Stochastic population forecasts and their uses. *International Journal of Forecasting* 8: 385-391
- Tuljapurkar, S., Lee, R. D., Li, Q., 2004: Random Scenario Forecasts Versus Stochastic Forecasts. *International Statistical Review* 72: 185-199
- Tuljapurkar, S., Orzack, S. H., 1980: Population Dynamics in Variable Environments, I: Long-Run Growth Rates and Extinction. *Theoretical Population Biology* 18: 314-342
- Tversky, A., Kahneman, D., 1974: Judgement under Uncertainty: Heuristics and Biases. *Science* 185: 1124-1131
- Uhlenberg, P., 1993: Demographic Change and Kin Relationships in Later Life. *Annual Review of Gerontology and Geriatrics* 13: 219-238

- Uhlenberg, P., 1995: Commentary: Demographic Influences on Intergenerational Relationships. In: Bengtson, V. L., Burton, L. M. (Hg.) *Adult Intergenerational Relations. Effects of Societal Change*, S. 19-26. New York: Springer
- Uhlenberg, P., 2005: Historical Forces Shaping Grandparent-Grandchildren Relationships: Demography and Beyond. In: Silverstein, M. (Hg.) *Focus on Intergenerational Relations Across Time and Space*, S. 77-97. New York: Springer
- Uhlenberg, P., 2006: Demography of Aging. In: Poston, D. L. j., Micklin, M. (Hg.) *Handbook of Population*, S. 143-167. New York: Springer
- UN Department of Economic and Social Affairs, 1966: *Demographic Yearbook 1965*. New York: United Nations
- UN Department of Economic and Social Affairs, 1983: *Demographic Yearbook 1981*. New York
- UN Department of Economic and Social Affairs, 2001: *Demographic Yearbook 1999*. New York: United Nations
- UN Department of Economic and Social Affairs, 2010: *Demographic Yearbook 2008*. New York: United Nations
- Van de Kaa, D., 1987: Europe's Second Demographic Transition. *Population Bulletin* 42: 1-57
- Vaupel, J. W., Manton, K. G., Stallard, E., 1979: The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. *Demography* 16: 439-454
- Vere, J. P., 2008: The perils of father-reported fertility data for household labour supply models. *Population Studies* 62: 235-243
- Voland, E., Paul, A., 1998: Vom „egoistischen Gen“ zur Familiensolidarität – Die soziobiologische Perspektive von Verwandtschaft. In: Wagner, M., Schütze, Y. (Hg.) *Verwandtschaft: Sozialwissenschaftliche Beiträge zu einem vernachlässigten Thema*, S. 33-58. Stuttgart: Enke
- Wachter, K. W., 1997: Kinship resources for the elderly. *Philosophical Transactions of the Royal Society London, Series B* 352: 1811-1817
- Wachter, K. W., Blackwell, D., Hammel, E. A., 1997: Testing the Validity of Kinship Microsimulation. *Mathematical and Computer Modelling* 26: 89-104
- Wachter, K. W., Hammel, E. A., Laslett, P., 1978: *Statistical Studies of Historical Social Structure*. New York: Academic Press
- Wachter, K. W., Knodel, J. E., VanLandingham, M., 2002: AIDS and the Elderly of Thailand: Projecting Familial Impacts. *Demography* 39: 25-41
- Wasserman, S., Faust, K., 1998: *Social Network Analysis. Methods and Applications*. Cambridge: Cambridge University Press
- Watkins, S. C., Menken, J. A., Bongaarts, J., 1987: Demographic Foundations of Family Change. *American Sociological Review* 52: 346-358
- Waugh, W. A. O., 1981: Application of the Galton-Watson Process to the Kin Number Problem. *Advances in Applied Probability* 13: 631-649
- Welsch, H., Wischnik, A., 2006: Müttersterblichkeit. In: Schneider, H., Husslein, P., Schneider, K.-T. (Hg.) *Die Geburtshilfe*, S. 1049-1063. Berlin: Springer
- Whelpton, P. K., 1936: An Empirical Method of Calculating Future Population. *Journal of the American Statistical Association* 31: 457-473
- Whelpton, P. K., 1946: Reproduction Rates Adjusted for Age, Parity, Fecundity and Marriage. *Journal of the American Statistical Association* 41: 501-516
- White, M. J., Lindstrom, D. P., 2006: Internal Migration. In: Poston, D. L. j., Micklin, M. (Hg.) *Handbook of Population*, S. 311-346. New York: Springer
- Wiesner, G., 2001: Der Lebensverlängerungsprozess in Deutschland. Stand – Entwicklung – Folgen. Beiträge zur Gesundheitsberichterstattung des Bundes des Robert-Koch-Instituts Berlin
- Wild, E., 1998: Adoption – Familienleben mit doppelter Elternschaft. In: Wagner, M., Schütze, Y. (Hg.) *Verwandtschaft. Sozialwissenschaftliche Beiträge zu einem vernachlässigten Thema*, S. 263-281. Stuttgart: Enke
- Willekens, F., 2005: Biographic Forecasting: Bridging the Micro-Macro Gap in Population Forecasting. *New Zealand Population Review* 31: 77-124

- Willekens, F., Baydar, N., 1984: Age-period-cohort models for forecasting fertility. NIDI Working Paper 45
- Wilmoth, J. R., 1993: Computational Methods for Fitting and Extrapolating the Lee-Carter Model of Mortality Change. Technical Report, University of California
- Wilmoth, J. R., Andreev, K., Jdanov, D., Gleijeses, D. A., 2007: Methods Protocol for the Human Mortality Database. Abgerufen am 7.9.2010 unter [www.mortality.org](http://www.mortality.org)
- Wilson, T., Rees, P., 2005: Recent Developments in Population Projection Methodology: A Review. *Population, Space and Place* 11: 337-360
- Wolf, D. A., 1988a: Kinship and Family Support in Aging Societies. In: UN Department of International Economics and Social Affairs (Hg.) *Economic and Social Implications of Population Aging*, S. 305-330. New York: United Nations
- Wolf, D. A., 1988b: The Multistate Life Table With Duration-Dependence. *Mathematical Population Studies* 1: 217-245
- Zeng Yi, 1986: Changes in Family Structure in China: A Simulation Study. *Population And Development Review* 12: 675-703
- Zeng Yi, Vaupel, J. W., Wang Zhenglian, 1997: A Multi-Dimensional Model for Projecting Family Households – With an Illustrative Numerical Application. *Mathematical Population Studies* 6: 187-216
- Zeng Yi, Vaupel, J. W., Wang Zhenglian, 1998: Household Projecting Using Conventional Demographic Data. *Population And Development Review* 24 (Supplement): 59-87
- Zeng Yi, Vaupel, W., Wang Zhenglian, 2003: Household Projection Using Conventional Demographic Data. In: Hullen, G. (Hg.) *Living Arrangements and Households - Methods and Results of Demographic Projections*, Materialien zur Bevölkerungswissenschaft 109, S. 35-57. Wiesbaden
- Zhang, L., 2011: *Male Fertility Patterns and Determinants*. New York: Springer
- Zhongwai Zhao, 1996: The demographic transition in Victorian England and changes in English kinship networks. *Continuity and Change* 11: 243-272
- Zhongwai Zhao, 2006: Computer microsimulation and historical study of social structure: A comparative review of SOCSIM and CAMSIM. *Revista de Demografia Històrica* 24: 59-88