

Ein stochastisches Prognosemodell internationaler Migration in Deutschland

Patrizio Vanella

Leibniz Universität Hannover
Center for Risk and Insurance
Demographic and Insurance Research Center
pv@ivbl.uni-hannover.de

Philipp Deschermeier

Institut der deutschen Wirtschaft Köln
Kompetenzfeld Finanzmärkte und
Immobilienmärkte
deschermeier@iwkoeln.de

30. August 2017

Zusammenfassung

Internationale Migration ist eines der gesellschaftlich am kontroversesten diskutierten Themen. Kritiker einer offenen Migrationspolitik sehen hohe Immigrationszahlen als großes Risiko für die Sicherheit und warnen vor möglichen Verdrängungseffekten am Arbeitsmarkt, während die Befürworter u.a. argumentieren, dass internationale Migration aus demografischer Sicht eine große Chance sei, die Folgen des Demografischen Wandels durch eine Erhöhung und Verjüngung der Bevölkerung auszubremsen und vor allem das Arbeitskräfteangebot in vom Fachkräftemangel bereits betroffenen Wirtschaftsbereichen zu erhöhen. Aus diesen Gründen ist es umso wichtiger, eine sachliche Diskussion auf Basis empirischer Ergebnisse zu führen. Eine quantitative Diskussionsgrundlage bildet in diesem Zusammenhang eine Prognose der zukünftigen Migrationsströme für Planungen in der Politik und dem Unternehmenskontext, was bisher nur unzureichend durchgeführt wird.

Hierfür stellen wir einen Modellansatz für die Prognose der internationalen Nettomigration zwischen Deutschland und dem Ausland, differenziert nach Geschlecht, Alter und Nationalitätsgruppen, vor. Der Beitrag liefert stochastische Prognosen der zukünftigen Nettomigrationen auf Basis eines Hauptkomponenten-Zeitreihenmodells. Bei diesem Verfahren bilden Prognoseintervalle die Unsicherheit über die zukünftige Entwicklung ab.

Danksagung

Wir möchten uns an dieser Stelle bei Christian Fußmann und Lieselotte Kathe vom Statistischen Bundesamt bedanken, die uns immer schnell, freundlich und hilfreich mit ihren Datenlieferungen unterstützt haben.

JEL-Klassifikation: C15; C38; C53; F66; J11; J18; J61

1 Einführung und Motivation

Internationale Migration ist die am schwierigsten zu prognostizierende demografische Komponente der Bevölkerungsentwicklung. Die zeitliche Entwicklung ist deutlich volatil als bei der Fertilität oder der Mortalität. Die Anwendung der von Vanella (2017) empfohlenen Hauptkomponenten-Zeitreihen-Methodik ist für vergleichbare Maße, wie altersspezifische Immigrations- und Emigrationsraten problematisch. Dabei ist die Bedeutung guter Prognosen von enormem Wert, da auf Basis dieser Vorhersagen viele volks- und betriebswirtschaftliche Planungen erstellt werden. Speziell das Thema internationale Migration wird kontrovers diskutiert. Befürworter der Begrenzung der Zuwanderung sehen ein großes Sicherheitsrisiko durch eine zunehmende Terrorgefahr oder äußern die, vor dem Hintergrund des Mindestlohns, nicht rationale Sorge, dass (günstige) ausländische Arbeiter den einheimischen Arbeitskräften die Stelle durch Lohndumping streitbar machen könnten. Befürworter einer offenen Migrationspolitik argumentieren hingegen, Migration sei eine Chance, dem Demografischen Wandel durch Erhöhung und Verjüngung der Bevölkerung entgegenwirken zu können und damit Engpässe auf Seiten des Arbeitskräfteangebots in bestimmten Berufsgruppen zu bekämpfen. Im Hinblick darauf ist es vor allem wichtig, ein Bild von der Alters- und Herkunftsstruktur der Ein- und Auswanderer zu bekommen.

Im vorliegenden Beitrag wird zunächst darauf eingegangen, mit welchen Problemstellungen ein Prognostiker bei der Vorhersage zukünftiger Migrationsströme konfrontiert ist. Dabei handelt es sich um Probleme in der Datenqualität zum ersten, nicht einheitliche Erfassung und Definitionen auf internationaler Ebene zum zweiten sowie die Migrationsbewegungen inhärente Komplexität zum dritten. Diese Problemstellungen werden im folgenden Kapitel aufgegriffen und detailliert

erläutert. Im dritten Abschnitt werden einige Lösungsansätze für die in Abschnitt 2 genannten Problemstellungen diskutiert, bevor im vierten Teil der Arbeit ein stochastischer Prognoseansatz für internationale Migration zwischen Deutschland und dem Ausland vorgeschlagen wird. Dieser ermöglicht nicht nur die Einschätzung des wahrscheinlichsten Verlaufs der zukünftigen Migrationszahlen, sondern auch eine Quantifizierung des Risikos in Form von Prognoseintervallen (*PI*). Die Ergebnisse der Prognose können in Form eines probabilistischen Kohorten-Komponenten-Modells für Bevölkerungsvorausschätzungen in Deutschland genutzt werden. Unser Modell schätzt dabei die Nettomigration detailliert nach Geschlecht, Altersjahren und Nationalitätsgruppen. So erhalten wir nicht nur einen Einblick, wie sich die Bevölkerung insgesamt entwickeln wird, sondern auch, wie sich die Altersstruktur und ethnische Herkunft der Bevölkerung durch internationale Migration verändern wird. Die Ergebnisse bieten damit das Potenzial für weiterführende ökonomische und demografische Analysen.

2 Problemstellungen bei der Modellierung und Prognose internationaler Migration

Das erste große Problem bei der Modellierung der Nettomigration besteht in der mangelnden Datenqualität. In jeder Phase der Erstellung der Statistiken gibt es potenzielle Fehlerquellen. Bereits in der Primärdatenerfassung entstehen die ersten Fehler. In Deutschland besteht für Personen generell eine Registrierungspflicht am neuen Wohnort nach einem Umzug. Analog sind wegziehende Personen dazu verpflichtet, sich vom vorigen Wohnort abzumelden. Trotzdem lassen sich diese Migrationsbewegungen, sei es innerhalb Deutschlands, oder auch grenzüberschreitend, z.B. zwischen zwei EU-Ländern, nur bedingt beobachten. In den genannten Beispielen besteht zwar die Pflicht, sich regulär an- und abzumelden, letztlich geschieht dies jedoch häufig auf freiwilliger Basis. Das führt dazu, dass Personen in manchen Fällen entweder gar nicht oder auch mehrfach in den Daten erfasst werden. Die abgeleiteten Migrationszahlen zwischen Herkunfts- und Zielort weichen daher häufig voneinander ab, teils sogar sehr stark. Kupizewska und Nowok (2008: 47) haben Vergleiche der Statistiken zu Zu- und Fortzügen zwischen den EU-Ländern gezogen und dabei festgestellt, dass in der Regel bei der Migration zwischen zwei Ländern die gemeldeten Zahlen der Migranten im Zielland höher sind als die analog vom Herkunftsland gemeldeten Werte der Fortzüge ins Zielland. In Extremfällen wurden teilweise

Abweichungen um mehr als das 100-fache identifiziert. Das zeigt, wie schwierig die korrekte Erfassung von Migrationsströmen vor allem innerhalb des Schengenraums ist. Dabei sei erwähnt, dass Deutschland im internationalen Vergleich tendenziell die besten Migrationsdaten aufweist, wenn davon ausgegangen werden kann, dass die Wahrscheinlichkeit für Mehrfachzählungen geringer ist als dafür, dass eine Migration gar nicht erfasst wird (Kupiszewska und Nowok 2008: 42-52). Schließlich variiert teilweise die Detailtiefe der Datenerfassung zwischen den Zielländern, weil nicht die gleichen Variablen nachgefragt werden oder falsche Daten aufgrund unwahrer Aussagen der Migranten, speziell in der Flüchtlingsmigration, erfasst werden (vgl. Bewarder und Leubecher 2016). Die nächste Phase betrifft die Erstellung der Statistiken. Dabei kommt es häufig zu weiteren Abweichungen der Migrationszahlen zwischen dem Herkunfts- und dem Zielland, weil unterschiedliche Definitionen eines Migranten zwischen den Ländern zu verschiedenen Zählungen führen. Es gibt seit einigen Jahren eine Empfehlung der UN zur Definition eines Migranten. Demnach sollte eine Migration erst als solche gezählt werden, wenn die fragliche Person einen Mindestaufenthalt von einem Jahr am neuen Aufenthaltsort vorweisen kann (vgl. United Nations 1998: 13). Es gibt jedoch viele Länder, die sich nicht an diese Empfehlung halten. So verlangt z.B. Belgien einen Mindestaufenthalt von drei Monaten, während in Deutschland gar kein Mindestaufenthalt verlangt wird. Hier gilt eine Person bereits ab dem Moment als Migrant, in dem sie registriert wird. Die unterschiedlichen Kriterien der Mindestaufenthaltsdauer können daher dazu führen, dass eine Immigration in einem anderen Jahr registriert wird als die zugehörige Emigration (vgl. Kupiszewska und Nowok 2008: 58-62). Die dritte und finale Phase, in der Fehler entstehen, ist die Verbreitung der Migrationsstatistiken. Dabei variiert die Verfügbarkeit und Detailtiefe sowie Vergleichbarkeit der Daten aus unterschiedlichen Datenquellen teils stark (vgl. Kupiszewska und Nowok 2008: 62-68). Die Statistiken beruhen nicht nur auf Verwaltungsdaten, sondern sind zudem angepasst durch Zensusdaten und andere Befragungen. In Bezug auf die Detailtiefe in den Daten gibt es eine breite Streuung. So fehlen häufig in den verfügbaren Statistiken wichtige Informationen, wie z.B. die Migrationsgründe. In manchen Datenbanken gibt es eine Aufteilung in Herkunfts- und Zielland, aber keine Altersangaben. In anderen Quellen gibt es Informationen zu Alter der Migranten, aber wiederum nicht zum Herkunfts- bzw. Zielland. Weiterhin wird eine Migration häufig nicht direkt gemessen, sondern nur indirekt über die Aufenthaltsorte einer Person zu zwei unterschiedlichen Zeitpunkten

approximiert. Zudem weisen manche Datenbanken gar keine Migrationszahlen auf, sondern nur Migrantenzahlen (vgl. Rogers et al. 2010: 6-7), so dass zwar die absolute Anzahl der Migranten in einem Land zu einem Zeitpunkt bekannt ist, aber nicht, wann diese migriert sind. Weiterhin unterschlägt diese Zählung die Information, wie viele Menschen immigriert und emigriert sind. Stattdessen wird nur ein Saldo erfasst.

Prognosen der zukünftigen Migration sind für politische und andere Entscheidungsprozesse eine wichtige Grundlage. Üblicherweise werden zumeist deterministische Projektionen der Nettomigration getroffen, die auf zuvor getroffenen Annahmen basieren (siehe beispielsweise Statistisches Bundesamt 2015: 40-41). Diese Methodik liefert neben eines aus statistischer Sicht sehr unwahrscheinlichen Basisszenarios eine geringe Anzahl an alternativen Szenarien, die zudem nicht mit Eintrittswahrscheinlichkeiten quantifiziert werden. Dadurch wird die Unsicherheit in den Projektionen nur unzureichend erfasst. Es gibt eine relativ geringe Anzahl probabilistischer Ansätze, die auf Basis statistischer Methoden Prognosen oder Projektionen für die zukünftige Migration in Deutschland erarbeiten (vgl. Bohk 2012; Fuchs et al. 2017; Deschermeier 2015; Bijak 2011). Diese betrachten entweder noch nicht den signifikanten Strukturbruch des Jahres 2015 oder quantifizieren die Unsicherheit der Prognose nur unzureichend. Tatsächlich werden für Prognosen zukünftiger Entwicklungen gute Vergangenheitsdaten benötigt, deren Beschaffung, wie zuvor erwähnt, schwierig ist. Zudem müssen diese Daten nicht nur kurzfristig von guter Qualität, sondern auch als Zeitreihen möglichst weit in die Vergangenheit zurückreichen. Das ist global, aber speziell für Deutschland und den Schengenraum, ein großes Problem, da sich aufgrund der deutschen Wiedervereinigung und der diversen Erweiterungen der EU und des Schengenraums Grenzen und Eintrittsbarrieren in den letzten 30 Jahren mehrfach verändert haben. Migrationsprognosen für Deutschland auf Basis von Daten vor 1990 durchzuführen, ist nur bedingt sinnvoll, da die damaligen Verhältnisse nicht repräsentativ für das heutige Migrationsgeschehen sind. Auch die heutigen Daten sind problematisch für die Prognose, da diese nicht so detailliert und kombiniert verfügbar sind, wie es für eine ausführliche Prognose nötig wäre. So verfügt das Statistische Bundesamt zwar separat über Migrationsdaten nach Alter und Geschlecht. Ebenfalls existieren nach Herkunfts- und Zielländern, nach Nationalität oder nach Aufenthaltzweck aufgeteilte Daten, nicht jedoch eine konsistente Kombination aus diesen

Merkmale. Weiterhin stimmen die Zahlen aus unterschiedlichen Datenveröffentlichungen häufig nicht überein.¹ Wünschenswert wäre eine konsistente Erfassung und Publikation der alters- und geschlechtsspezifischen Migration nach Nationalität differenziert sowie diskriminiert nach Migrationsgrund als Datengrundlage für die Prognose. Es gibt Lösungsvorschläge, um die Folgen der Datenproblematik etwas einzudämmen. Auf die für dieses Papier zentralen Vorschläge wird im nächsten Abschnitt eingegangen. Neben den Datenproblemen wird die Migrationsprognose auch dadurch erschwert, dass Migrationsbewegungen von vielen externen politischen, sozialen, wirtschaftlichen und klimatischen Faktoren beeinflusst werden, sodass ein hochkomplexes Gebilde entsteht, welches kaum prognostizierbar ist (vgl. Bohk 2012: 139). Wir werden im Rahmen dieser Arbeit allerdings einen Vorschlag auf Basis einer Hauptkomponentenanalyse erarbeiten, der diese Problematik etwas reduziert. Wir werden einen Modellansatz vorstellen, der statistisch das Risiko in der Prognose quantifiziert, was den positiven Nebeneffekt hat, dass die subjektiven Modellannahmen, die in diesem Modell notwendigerweise getroffen werden müssen, wiederum etwas gelockert werden.

3 Lösungsansätze für die Datenproblematik und Prognosemodell

An dieser Stelle soll zunächst auf mögliche Lösungsansätze für die in Abschnitt 2 angesprochenen Probleme eingegangen werden. Im Anschluss folgt ein eigener Ansatz, der die Grundlage für ein stochastisches Prognosemodell bildet, das die internationale Migration zwischen Deutschland und dem Ausland vorausberechnet.

Häufig liegen Migrationsdaten nicht in Altersjahre gegliedert vor, sondern aggregiert zu Altersgruppen. Die Frage stellt sich generell, wie mit dieser Einschränkung umzugehen ist, wenn im Hinblick auf jährliche Bevölkerungsprognosen die Migrationsdaten auf Altersjahre heruntergebrochen werden müssen.

¹ So stimmt z.B. im Bericht *Wanderungen über die Grenzen Deutschlands 2005 nach der Staatsangehörigkeit und Altersgruppen* die Zahl unter 18-jähriger männlicher Deutscher (14.510) bei weitem nicht mit der kumulierten Zahl im Bericht *Wanderungen zwischen Deutschland und dem Ausland 2005 nach Einzelaltersjahren und Geschlecht* (15.473) überein. Dies ist ein willkürlich gewähltes Beispiel, was jedoch in der Sache vielfach zu beobachten ist.

Der klassische Ansatz zur Schätzung der Altersstruktur der Migranten ist das Rogers-Castro-Modell, welches von Andrei Rogers und Luis Castro entwickelt wurde. Dabei handelt es sich um ein Modell mit bis zu 13 Parametern, welches den Migrationszyklus über alle Alter schätzt (vgl. Rogers und Castro 1981: 5-13). Das Rogers-Castro-Modell liefert empirisch eine sehr gute Anpassung an internationale Migrationszahlen sowie an Binnenmigration und bietet sich vor allem für Fälle an, in denen keine Informationen oder lediglich Hilfsdaten zu Altersverteilungen von Migranten vorliegen. Bei altersgruppierten Daten empfehlen Rogers et al. eine Interpolation der Migrationszahlen über Hilfsdaten nach Altersjahren mit Hilfe des Rogers-Castro-Modells (vgl. Rogers et al. 2010: 22-24).

Neben dem Rogers-Castro-Modell erfahren Ansätze, die auf dem Paradigma der funktionalen Datenanalyse (Ramsay und Silverman, 2001; 2005) aufbauen, zunehmend Aufmerksamkeit. Dieser Ansatz beschreibt eine Denkweise für den Umgang mit Datenreihen. Die Grundidee besagt, dass Beobachtungen einer Reihe nicht unabhängig voneinander sind, sondern einem funktionalen Zusammenhang folgen. Aggregierte demografische Raten lassen sich auf dieser Grundlage durch Spline-Regression (Nieden et al. 2016, Wood 1994: 27) auf Altersjahre disaggregieren. Dies erweist sich insbesondere bei regionalen Bevölkerungsprognosen als hilfreich, da die erforderlichen Daten meist nur grob aggregiert verfügbar sind (Deschermeier 2011: 775). Auf dem Paradigma funktionaler Daten bauen verschiedene Zeitreihenmodelle auf (Hyndman und Ullah 2007; Hyndman und Booth 2008; Hyndman et al. 2013), die speziell für demographische Raten vorgesehen sind. Dieses Verfahren und die darauf aufbauenden Zeitreihenmodelle eignen sich besonders für die Fertilitäts- und Mortalitätsraten, da beide eine hohe zeitliche Konstanz aufweisen. Da die internationale Nettomigration jedoch vielfältigen Einflüssen und Strukturbrüchen im Zeitablauf unterliegt, erscheint dieser Ansatz für das vorliegende Problem suboptimal.

Theoretisch wäre eine Prognose auf Basis der alters- und geschlechtsspezifischen Immigrations- und Emigrationsraten durchführbar. Dieser Ansatz wäre jedoch für die Immigration praktisch nicht umsetzbar, weil dafür global detaillierte Bevölkerungsdaten zu allen Ländern vorliegen müssten. Selbst bei Vorliegen entsprechender Daten wäre der mit der Datenrecherche verbundene Aufwand unverhältnismäßig hoch, was das Modell für eine regelmäßige

Aktualisierung ineffizient machen würde. Denkbar wäre auch das Aufstellen eines Prognosemodells mit exogenen Variablen (siehe z.B. Bijak 2011: 64-72; Bohk 2012: 43-45). Dies hätte allerdings nicht nur den Nachteil, dass es das Modell verkomplizieren und den Bedarf an Daten hochschrauben würde, sondern vor allem, dass auch diese exogenen Variablen prognostiziert werden müssten, was in einigen Fällen nicht nur sehr schwierig, sondern praktisch unmöglich wäre. So könnte z.B. ein ökonometrisches Prognosemodell für die Immigration erklärende Variablen wie das Bruttoinlandsprodukt des Herkunfts- und des Ziellandes enthalten. Klimatische Variablen wie Erdbebenwahrscheinlichkeiten oder Temperaturvariablen wären ebenfalls denkbar. Arbeitsmarktvariablen wie die Zahl an unbesetzten Stellen wären ebenfalls mögliche Alternativen. Für eine Prognose von Migrationsbewegungen auf Basis eines solchen erklärenden ökonometrischen oder environmetrischen Modells müssten jedoch zunächst die aufgezählten Variablen prognostiziert werden, was bei ökonomischen Variablen bereits sehr schwer und bei naturwissenschaftlichen Variablen zum Teil nahezu unmöglich ist.

Dieses Vorgehen lässt sich jedoch über die Hauptkomponentenanalyse indirekt annähern, wie noch im weiteren Verlauf erklärt werden wird. Wie bereits dargelegt wurde, verfügt das Statistische Bundesamt über einige sehr detaillierte Datensätze, die allerdings untereinander nicht immer konsistent sind. Rogers et al. empfehlen die Erstellung eines synthetischen Datensatzes aus verschiedenen Quellen (z.B. die Kombination von Zensus- mit Migrationsdaten) über ein generalisiertes lineares Modell (GLM). Dieser sollte, geometrisch gesprochen, in Form eines Würfels zusammengestellt sein, wobei dieser die drei Dimensionen *demografische Informationen* (wie z.B. Geschlecht und Alter), *geografische Informationen* (wie z.B. das Herkunftsland und die Nationalität) sowie *zeitliche Informationen* (z.B. das Jahr der Migration) enthielte (vgl. Rogers et al. 2010: 4-44). Das setzt jedoch voraus, dass die verschiedenen Datenquellen konsistent sind.

Die Kombination von Mikro- und Makrodaten ist generell eine Alternative, die mit Vorsicht anzuwenden ist, da sich generell bei Mikrodaten die Frage stellt, wie repräsentativ sie für die gesamte Population sind und ob sie durch Über- und Unterrepräsentation gewisser Gruppen verzerrt sind. Obwohl sehr gute Mikro- und Paneldatensätze (wie beispielsweise für Deutschland das Sozio-ökonomische Panel oder der Mikrozensus) zur Verfügung stehen, beruhen diese doch

immer auf Stichproben und die Grundgesamtheit ist nie bekannt, weshalb derartige Fehler nicht auszuschließen sind. Weiterhin ist ein Arbeiten mit Mikrodaten generell zeitintensiver und in manchen Fällen auch kostspieliger (wie z.B. beim Arbeiten mit Mikrozensusdaten). Wir verzichten daher auf diesen Ansatz, da gute Modelle sich auch dadurch auszeichnen, die Realität mit möglichst geringem Aufwand möglichst gut abzubilden. Der GLM-Ansatz wäre auch denkbar bei Kombination unterschiedlicher Makrodaten. So verfügt das Statistische Bundesamt über Zahlen zur altersspezifischen Migration für 0-95-jährige, allerdings nur aufgeteilt in die beiden Nationalitätsgruppen „deutsch“ und „nichtdeutsch“ bzw. „ausländisch“. Eine weitere Unterscheidung wäre allerdings von immenser Wichtigkeit, da sich die Wanderungsmotive und Eintrittsbarrieren für z.B. Italiener gegenüber Äthiopiern stark unterscheiden, diese aber in der fraglichen Statistik in einer Gruppe zusammengefasst werden. Eine Prognose auf Basis dieser Daten wäre nicht sinnvoll durchführen.

Das Statistische Bundesamt verfügt jedoch ebenfalls über Migrationszahlen, die nach exakter Nationalität unterschieden werden. Diese sind allerdings im Bestfall nach relativ groben Altersgruppen², nicht nach genauem Alter diskriminiert. Eine GLM-Synthese dieser beiden Datensätze über z.B. Mikrozensusdaten als Hilfsdaten wäre theoretisch möglich, scheitert aber daran, dass die beiden angesprochenen Datensätze nicht übereinstimmen, wie in Abschnitt 2 gezeigt wurde. Eine zu große Unterscheidung der Bevölkerung hat wiederum den Nachteil, dass die Anzahl der zu schätzenden Variablen bei der späteren Prognose schnell zu groß wird, um mit gängiger Hardware-Ausstattung Simulationen durchführen zu können. Weiterhin sorgt eine Unterscheidung nach jeder Nationalität dafür, dass Länder, die relativ geringe Migrationszahlen von und nach Deutschland haben, zu kleine Stichproben in gewissen Altern aufweisen, um aus ihnen statistisch signifikante Schätzungen abzuleiten. Aus diesen Gründen unterscheiden wir zunächst unsere Daten in die sieben Nationalitätsgruppen *Deutsche (DE)*, *Ausländer aus der EU oder anderen Schengenstaaten (EU)*, *sonstige Europäer (RE)*, *Asiaten (AS)*, *Afrikaner (AF)*, *Amerikaner und Ozeanier*, im weiteren Verlauf als *Übersee* definiert (*ÜS*) und *restlichen Staatsbürgern*³(*NA*).

² Unter 18; 18-24; 25-49; 50-64; über 65.

³ Staatenlose oder unbekannt Nationalitäten.

Die Daten wurden entsprechend dieser Diskriminierung aufgeteilt und nach Geschlechtern differenziert. Da die Daten mit dieser Detailtiefe nur für die zuvor genannten fünf Altersgruppen verfügbar sind, müssen die fehlenden Daten interpoliert werden. Dazu werden, ähnlich dem Ansatz von Rogers et al., die Übersichten zur Migration nach Altersjahren als Hilfsdatensatz genommen. Um eine möglichst genaue Schätzung der Altersstrukturen in den Untergruppen zu erreichen, wurde die Altersverteilung, die sich innerhalb der fünf Altersgruppen für Nichtdeutsche nach dem Bericht *Wanderungen zwischen Deutschland und dem Ausland 1991-2015 nach Einzelaltersjahren und Geschlecht* ergeben, für alle ausländischen Nationalitätsgruppen angenommen. Für das Jahr 1990 existieren keine so detaillierten Daten für ganz Deutschland, weshalb die Altersverteilung in 1991 auch für 1990 an die Altersgruppen angepasst wurde. Dabei wurde auf die Unterscheidung zwischen den Geschlechtern und Ein- sowie Auswanderern geachtet, sodass möglichst viele in den Daten vorhandene Informationen genutzt wurden. Es wurde jedoch darauf Wert gelegt, dass die geschätzte Verteilung für jede Altersgruppe mit den Werten aus dem Migrationsdatensatz *Wanderungen über die Grenzen Deutschlands 1990-2015 nach (ausgewählten) Staatsangehörigkeit(en) und Altersgruppen* übereinstimmt.

Der Vorteil gegenüber einer klassischen Glättung nach dem Rogers-Castro-Modell liegt darin, dass keine theoretischen Altersverteilungen bei den Migranten angenommen werden müssen, sondern sich die Verteilungen nach der Gesamtverteilung der Migranten verhalten, sodass die Schätzung über beide genutzten Datensätze relativ konsistent ist. Da keine detaillierten Daten über die altersspezifische Migration für über 94-jährige vorliegen⁴, wurden die gruppierten Daten teils über OLS-, teils über lineare Extrapolation auf die Einzelalter bis 100 Jahre geschätzt. Für höhere Alter als 100 Jahre wird die Nettomigration im Modell durchgängig als null angenommen. Nachdem auf diesem Wege unser synthetischer Migrationsdatensatz erstellt wurde, werden aus den konstruierten alters- und geschlechtsspezifischen Migrationsflüssen die direktionalen alters-, geschlechts- und nationalitätsspezifischen Nettomigrationszahlen generiert, was die von Rogers et al. geforderte dreidimensionale Struktur ergibt. Das Ergebnis ist eine 26x1414-Matrix mit den

⁴ Diese Altersgrenze variiert zwischen den Berichten etwas.

Jahren (1990-2015) in den Zeilen und den Nettomigrationszahlen nach Alter (0-100 Jahre), Geschlecht (binär) und Nationalitätsgruppe (DE, EU, RE, AS, AF, ÜS, NA) in den Spalten.⁵

Mit Hilfe der Hauptkomponentenanalyse⁶ werden nun Linearkombinationen aus den Nettomigrationszahlen generiert. Wir reduzieren das Modell auf ein quasi-zwei-Hauptkomponentenmodell, da diese bereits 75,8% der Variation in der altersspezifischen Nettomigration zwischen Deutschland und dem Ausland erklären und bereits die dritte Hauptkomponente sich auf Basis der von Vanella (2017) beschriebenen Tests als Weißes Rauschen herausstellt.⁷ Abbildungen 1 und 2 illustrieren die Korrelationsstrukturen, oder auch Komponentenladungen, zwischen den ersten zwei Hauptkomponenten und den Nettomigrationszahlen.

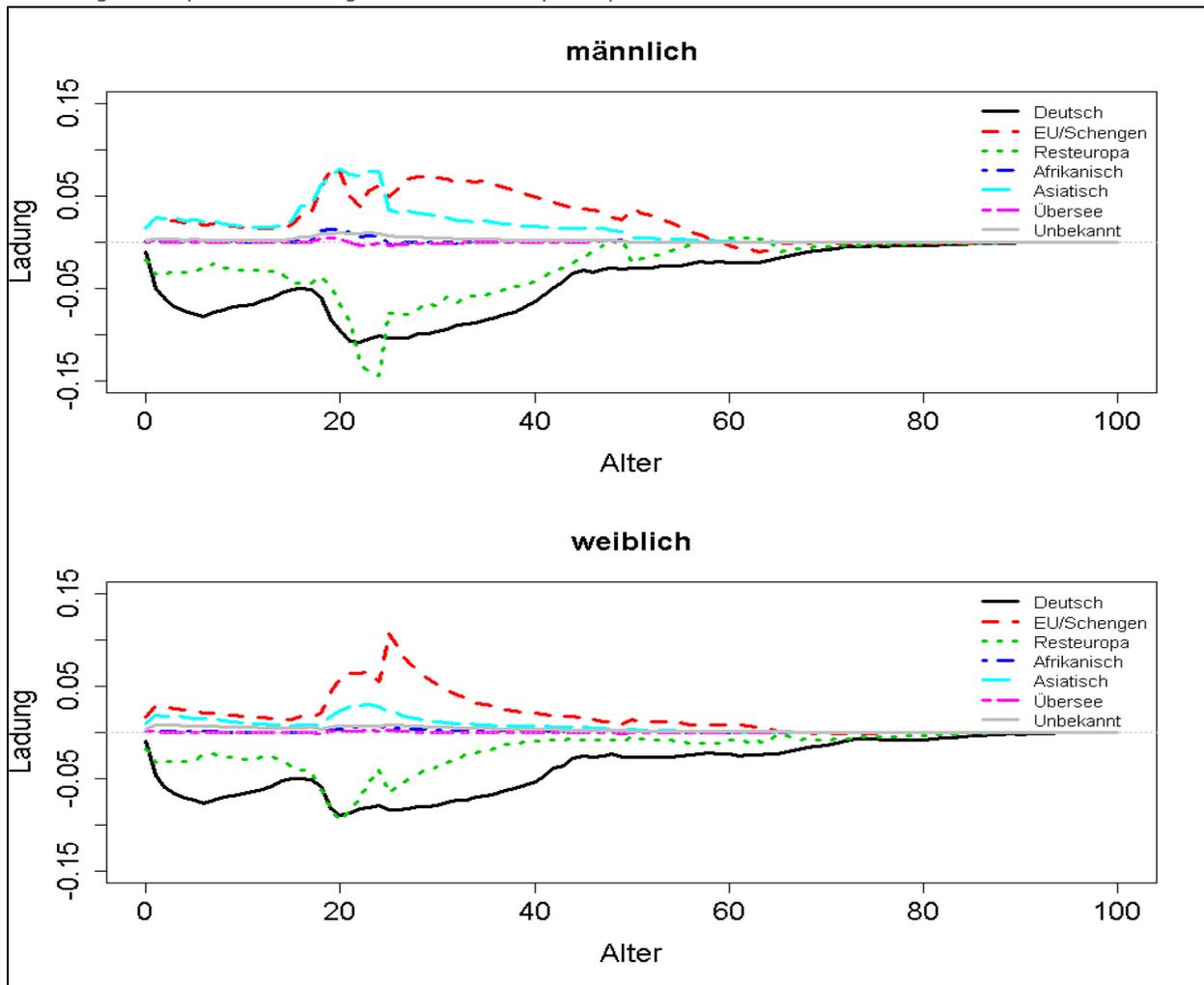
Es fällt auf, dass die Komponentenladungen besonders hoch für Personen im Alter von etwa 18 bis 50 Jahre sind, was nahelegt, dass diese Hauptkomponente vor allem die Arbeitsmigration widerspiegelt. Dabei ist zu beobachten, dass beispielsweise hohe Werte der Hauptkomponente hohe negative Nettomigrationszahlen bei Deutschen bedeuten, während ceteris paribus gleichzeitig hohe positive Nettomigrationszahlen bei EU- und Schengen-Ausländern und Asiaten generiert werden. Das legt nahe, dass Austauscheffekte auf dem Arbeitsmarkt vorliegen, wobei abwandernde deutsche Arbeitskräfte vor allem durch EU-Ausländer und Asiaten, allen voran Inder und Chinesen, ersetzt werden, wenn auch nicht eins zu eins.

⁵ Streng genommen entspricht dies sogar einer vierdimensionalen Datenstruktur mit den Dimensionen 26x101x2x7.

⁶ Siehe Vanella 2017: 3-9 für eine detaillierte Erklärung der Methodik.

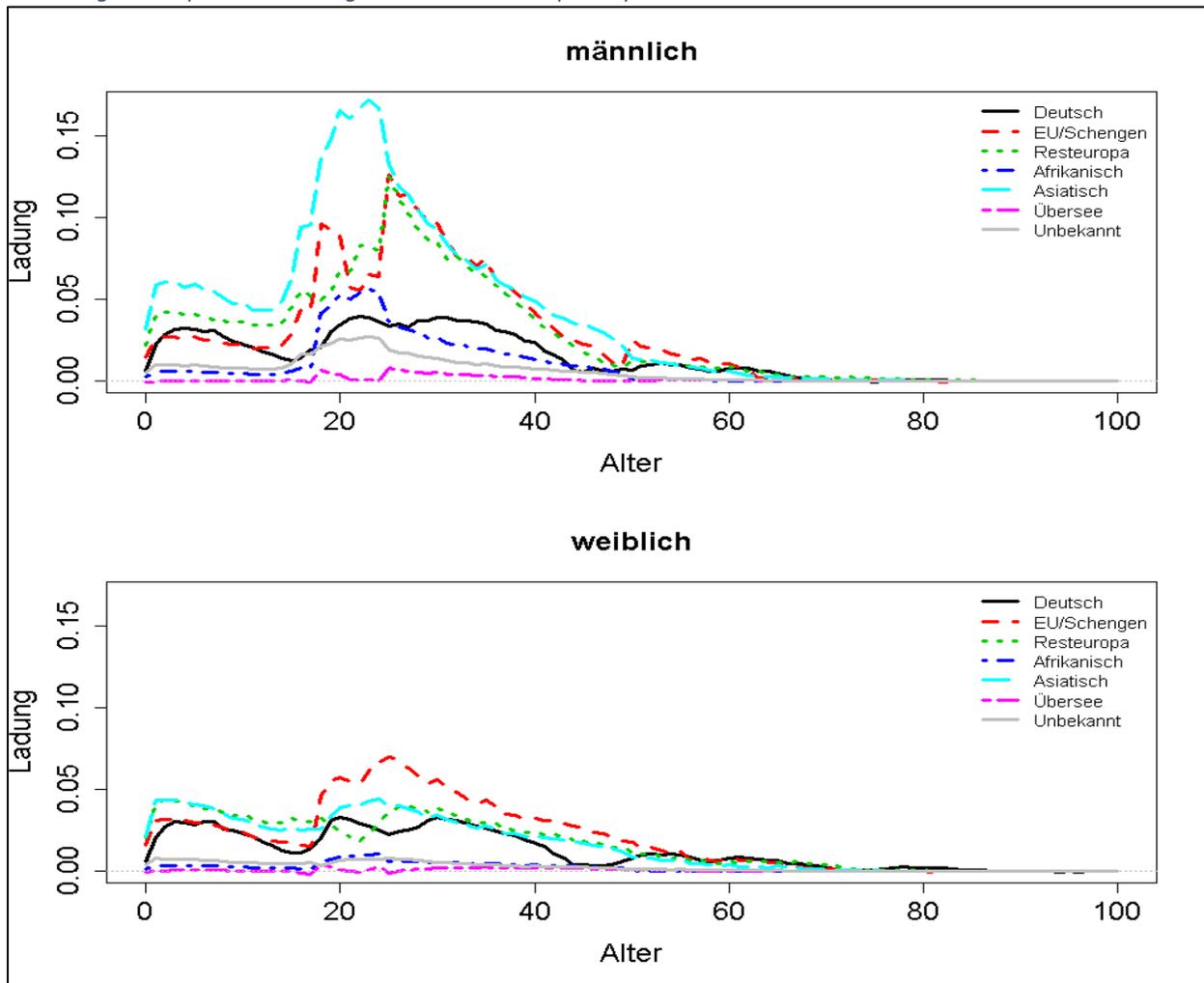
⁷ Siehe Vanella 2017: 12-17 für eine genauere Beschreibung der genutzten Tests.

Abbildung 1: Komponentenladungen der ersten Hauptkomponente



Quelle: eigene Berechnung und Darstellung

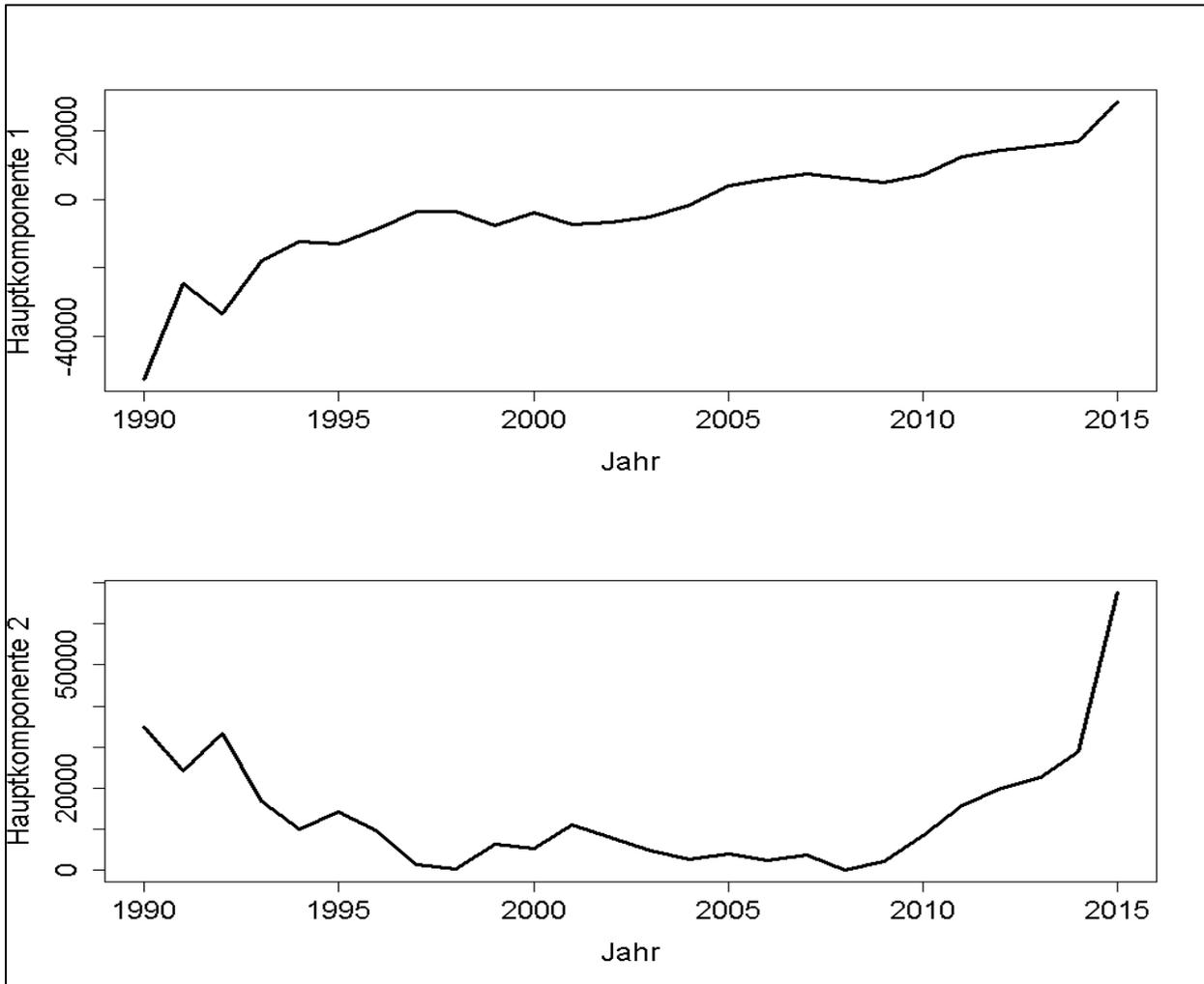
Abbildung 2: Komponentenladungen der zweiten Hauptkomponente



Quelle: eigene Berechnung und Darstellung

Die zweite Hauptkomponente ist grundsätzlich positiv mit der Nettomigration korreliert, es handelt sich scheinbar um einen generellen Indikator für das Zuwanderungsniveau. Dabei sind die Komponentenladungen besonders hoch für männliche Asiaten im Alter von ca. 20 Jahren und europäische Ausländer zwischen Mitte 20 und Anfang 30. Diese Beobachtung kann mit Hilfe der in Abbildung 3 visualisierten Ergebnisse gut in den Kontext gesetzt werden, welche den theoretischen historischen Verlauf der ersten beiden Hauptkomponenten seit 1990 wiedergibt.

Abbildung 3: Historischer Verlauf der ersten zwei Hauptkomponenten



Quelle: eigene Berechnung und Darstellung

Hauptkomponente 1 wächst im Zeitverlauf nahezu monoton, was im Kontext einer zunehmenden Globalisierung nachvollziehbar erscheint. Bei Hauptkomponente 2 ist im betrachteten Zeitraum beinahe eine U-Form zu beobachten, wobei das Niveau von Anfang bis Ende der 1990er Jahre von einem durch den Fall der Sowjetunion und die jugoslawischen Bürgerkriege sehr hohen Niveau in den frühen 90er Jahren allmählich absank und dann bis 2008 relativ stabil blieb. In der Folge lässt sich ein Anstieg durch die Eurokrise in Verbindung mit der EU-Erweiterung durch Rumänien und Bulgarien feststellen, da seitdem verstärkt junge Menschen aus den Süd- und Osteuropäischen Krisenländern nach Deutschland und generell Nordeuropa migrieren. 2014 kam ein weiterer Migrationsschub durch die letzte EU-Erweiterung durch Kroatien. 2015 markiert den Effekt der Flüchtlingskrise in Form des Maximums, die durch die von Bundeskanzlerin Merkel im Sommer 2015 für Menschen aus den durch den IS terrorisierten Ländern in Aussicht gestellte vereinfachte

Anerkennung als Flüchtling in Deutschland verursacht wurde. Dies bewegte auch Menschen aus anderen wirtschaftsschwachen Drittstaaten wie Albanien und dem Kosovo dazu, eine Anerkennung als Flüchtling in Deutschland zu suchen (vgl. Bundesamt für Migration und Flüchtlinge 2016: 118-123). Auch die Migration aus Afghanistan, wo die Sicherheitslage nach wie vor höchst prekär ist, in erster Linie durch die Taliban (vgl. Heidelberg Institute for International Conflict Research 2017: 71-73), ist in Folge der Flüchtlingskrise wieder in die Höhe geschneit. Die hohe Betonung auf männliche Migration liegt darin begründet, dass häufig zunächst die Männer aus den Krisenregionen fliehen, um hinterher über den Familiennachzug die restliche Familie gesichert nachzuholen (vgl. Leubecher 2017). Es stellt sich naturgemäß an dieser Stelle die Frage, wie dieses hochkomplexe Gebilde prognostiziert werden könnte. Im kommenden Abschnitt werden wir dazu einen Vorschlag erarbeiten.

4 Ein stochastischer Prognoseansatz für Zu- und Fortzüge zwischen Deutschland und dem Ausland

Im vorangegangenen Abschnitt wurde die Datenbasis vorgestellt, für die eine Hauptkomponentenanalyse durchgeführt wurde. Diese Zwischenergebnisse bilden nun die Grundlage für eine Prognose der internationalen Nettomigration in Deutschland. Zunächst wird auf Basis der von Vanella (2017) vorgestellten Kriterien ein parametrisches Zeitreihenmodell an die erste Hauptkomponente angepasst, wobei ein Box-Jenkins Modell⁸ zur Schätzung der Autokorrelation eingesetzt wurde (siehe Vanella 2017: 9-17). Das resultierende Modell für den fraglichen Zeitraum lautet:

$$c_1(t) = -48213,6744 + 19195,5931 \ln t + 0,3421 \delta_{t-1} + 0,6579 \delta_{t-2} + \varepsilon_t,$$

wobei $\delta_t := c_1(t) - [-48213,6744 + 19195,5931 \ln t]$ und $\varepsilon_t \sim \mathcal{N}(0, 4827,2400^2) \forall t$. Das setzt ein logarithmisches Wachstum als langfristigen Trend voraus, was angesichts der Historie realistisch ist.

Für Hauptkomponente 2 ist der zukünftige Verlauf nicht so einfach ableitbar, da dieser stark mit aktuellen Ereignissen und Krisen verbunden ist. Wir werden an dieser Stelle die Annahme treffen,

⁸ Siehe Box et al. 2016: 88-116 für eine detaillierte Beschreibung von Box-Jenkins ARIMA-Modellen.

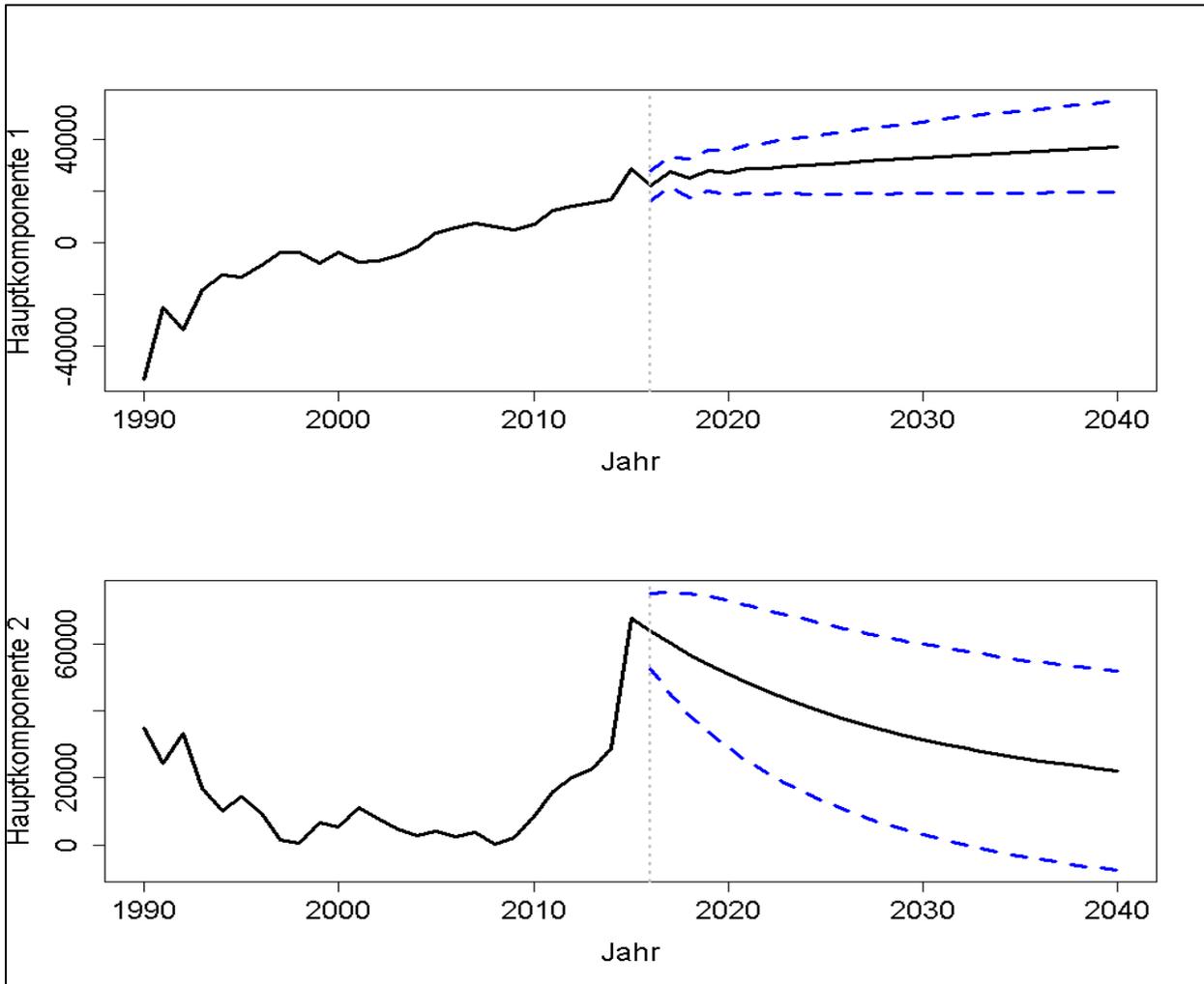
dass sich die Krisen aufgrund einer langsamen Erholung der südeuropäischen Wirtschaft, geeigneter Maßnahmen in der Flüchtlingspolitik durch das Innenministerium und militärischer Erfolge in der Terrorbekämpfung langsam stabilisieren werden und in der Folge das Niveau der Hauptkomponente langfristig zu ihrem langfristigen Mittel konvergieren wird. Das erscheint im ersten Moment als sehr starke Annahme für eine Prognose, lässt sich aber mit einem AR(1)-Modell sehr gut stochastisch modellieren, sodass das negative Wachstum der Hauptkomponente nicht willkürlich gewählt wird. Durch eine adäquate Quantifizierung von Prognoseintervallen wird die Grundannahme zum zukünftigen Verlauf schnell relativiert. Das aus der OLS-Schätzung resultierende Prognosemodell für Hauptkomponente 2 stellt sich dar als:

$$c_2(t) = 13842,6325 + 0,9281\omega_{t-1} + \xi_t,$$

wobei $\omega_t := c_2(t) - 13842,6325$ und $\xi_t \sim \mathcal{N}(0, 9767,2033^2) \forall t$.

Die Prognose für die ersten zwei Hauptkomponenten auf Basis dieser Modelle wird in Abbildung 4 illustrativ mit dem 75%-Prognoseintervall (75%-PI) dargestellt. Es ist zu beobachten, dass die Arbeitsmarkttrends einem relativ geringeren Risiko unterliegen und im Mittel weiter positiv wirken werden. Die Prognoseintervalle bei Hauptkomponente 2 sind relativ breit, sodass auch andere Verläufe als der im Mittel angenommene adäquat ins Modell einfließen und mit Wahrscheinlichkeiten belegt werden.

Abbildung 4: Prognose der ersten zwei Hauptkomponenten mit 75%-PI



Quelle: eigene Berechnung und Darstellung

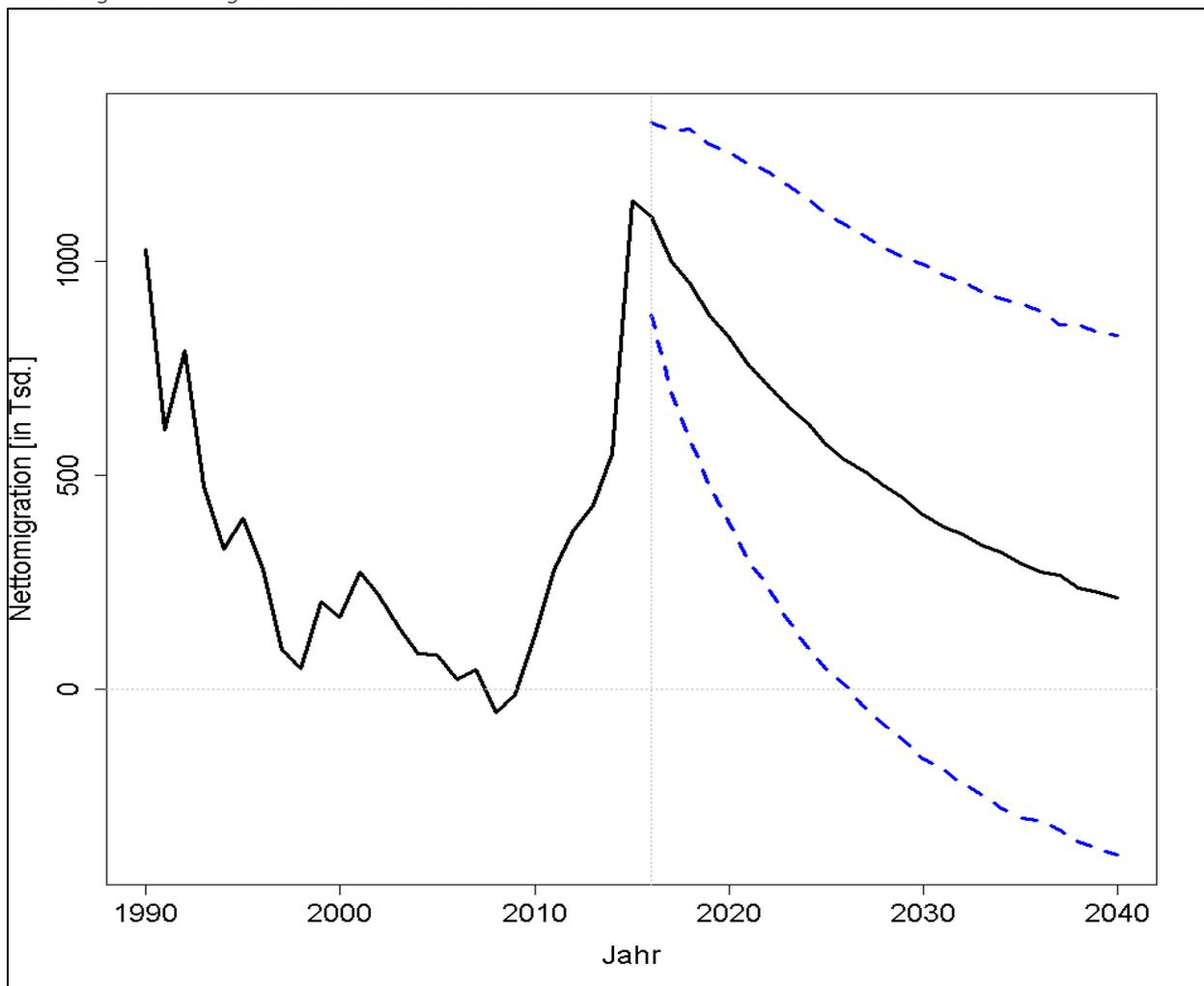
Das ist einer der großen Vorteile einer stochastischen gegenüber einer deterministischen Analyse. Selbst in Fällen, in denen die Trends nicht so eindeutig sind, werden getroffene Annahmen, die mit hoher Wahrscheinlichkeit ex-post nicht eintreten, deutlich weniger stark gewichtet. Die verbleibenden 1.412 Hauptkomponenten werden als Random Walk Prozesse⁹ angenommen und so bei der Schätzung des gesamten Migrationsrisikos mit einbezogen. Auf Basis der Modelle wird nun eine Simulationsstudie durchgeführt, wobei alle Hauptkomponenten anhand von Wiener Prozessen¹⁰ bis ins Jahr 2040 10.000-mal simuliert werden. Wie von Vanella erklärt, erfolgt nun die Rücktransformation der Simulationsergebnisse (vgl. Vanella 2017: 21), sodass 10.000 simulierte

⁹ Zur Definition eines Random Walk Prozesses siehe z.B. Vanella 2017: 12.

¹⁰ Siehe für eine detaillierte Erklärung von Wiener Prozessen z.B. Tavella 2002: 13-17.

Pfade der Nettomigration für jedes Alter von 0-100 Jahre, beide Geschlechter sowie die sieben Nationalitätsgruppen generiert werden. Auf Grund der hohen Dimensionalität der Simulationsstudie würde an dieser Stelle eine detaillierte Darstellung aller Prognoseergebnisse den Rahmen sprengen. Wir werden nachfolgend deshalb auf einige ausgewählte Ergebnisse eingehen. Abbildung 5 illustriert die Prognose der gesamten Nettomigration in Deutschland bis ins Jahr 2040 mit 75%-PI, die aus der Simulationsstudie abgeleitet werden kann.

Abbildung 5: Nettomigration in Deutschland mit 75%-PI



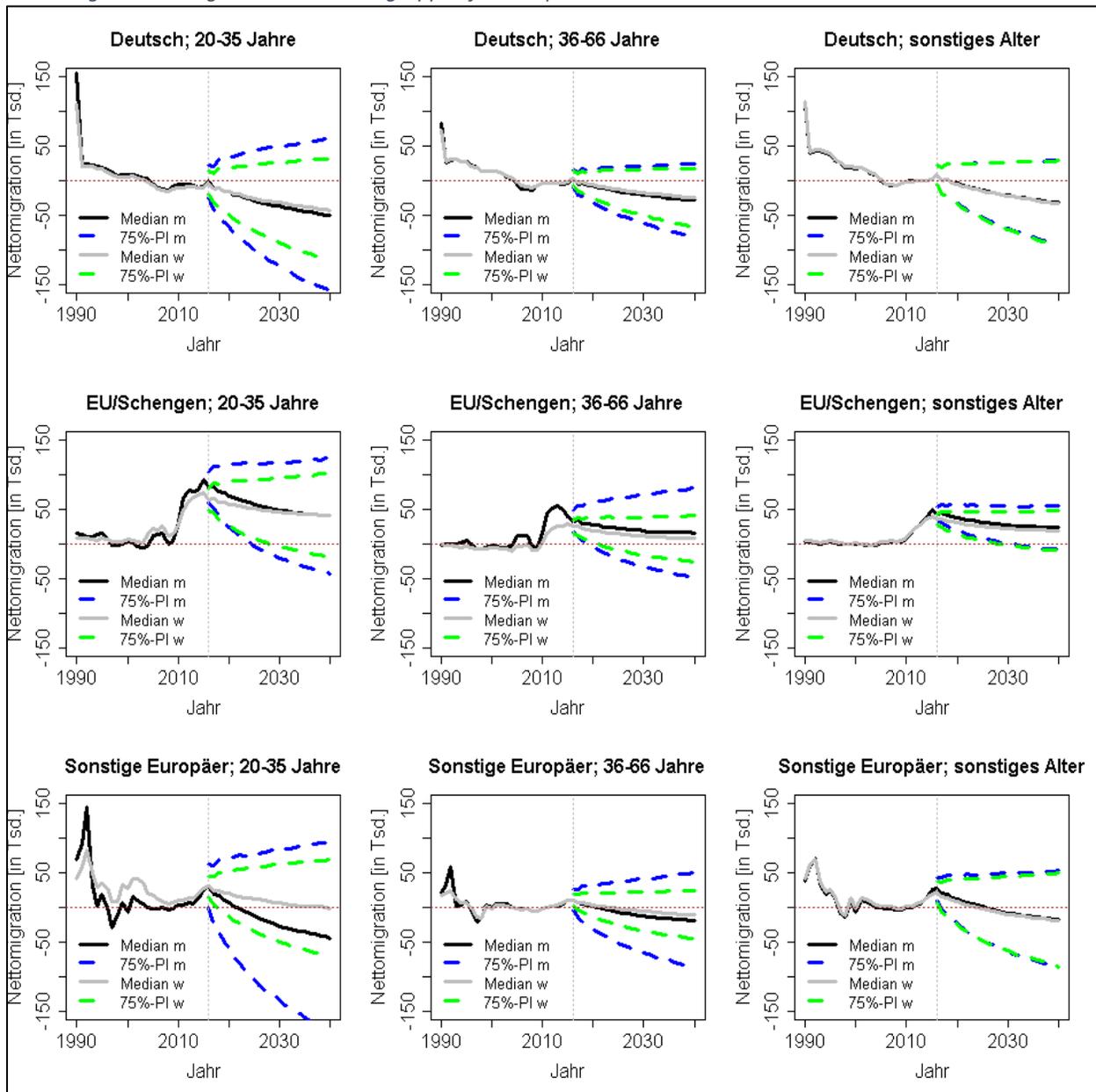
Quelle: eigene Berechnung und Darstellung

Dabei sei beachtet, dass zum gegenwärtigen Zeitpunkt noch keine detaillierten Migrationszahlen zu 2016 verfügbar sind. Daher handelt es sich an dieser Stelle um einen Prognosewert anhand des in diesem Beitrag entwickelten Modells. Die sich aus dem Ausländerzentralregister für 2016 ergebende Nettomigration liegt nach aktuellem Stand bei 642.897 (Bundesamt für Migration und

Flüchtlinge 2017: 73), was im von uns prognostizierten 97%-PI läge. Dabei können sich im Nachhinein noch durch spätere Bereinigungen Änderungen ergeben.

Es ist zu beobachten, dass im Median ein langsames Absinken des Niveaus im Zeitverlauf bis auf 212.419 im Jahr 2040 zu erwarten ist. Die Varianz in der Nettomigration ist jedoch erwartungsgemäß sehr stark, da Migration, wie erwähnt, die demografische Komponente mit der höchsten Stochastizität ist. So liegt das 75%-PI im Jahre 2040 bei [-387.809; 826.879]. Die Ergebnisse nähern sich im Median und Mittelwert asymptotisch sowohl der Prognose von Deschermeier (2016) als auch der Aktualisierung der 13. koordinierten Bevölkerungsvorausberechnung auf Basis der Daten von 2015 (Statistisches Bundesamt 2017b) an. Auf Basis unseres Modells lässt sich eine Vielzahl von spezifischen Analysen durchführen. In den Abbildungen 6 und 7 finden sich abgeleitete Prognosen für die Nettomigration nach beiden Geschlechtern, den drei Altersgruppen 20-35 Jahre, 36-66 Jahre und sonstiges Alter sowie nach den Nationalitätsgruppen. Aus Darstellungsgründen und da das Migrationsniveau zwischen der Übersee-Gruppe und Deutschland vergleichsweise niedrig ist, ist diese Gruppe an dieser Stelle mit den sonstigen Nationalitäten zusammengefasst. Jede Zeile bezieht sich in der Grafik auf eine Nationalitätsgruppe, dabei sind die Prognosen für beide Geschlechter mit 75%-PI dargestellt.

Abbildung 6: Nettomigration nach Altersgruppen für Europäer bis 2040



Quelle: eigene Berechnung und Darstellung

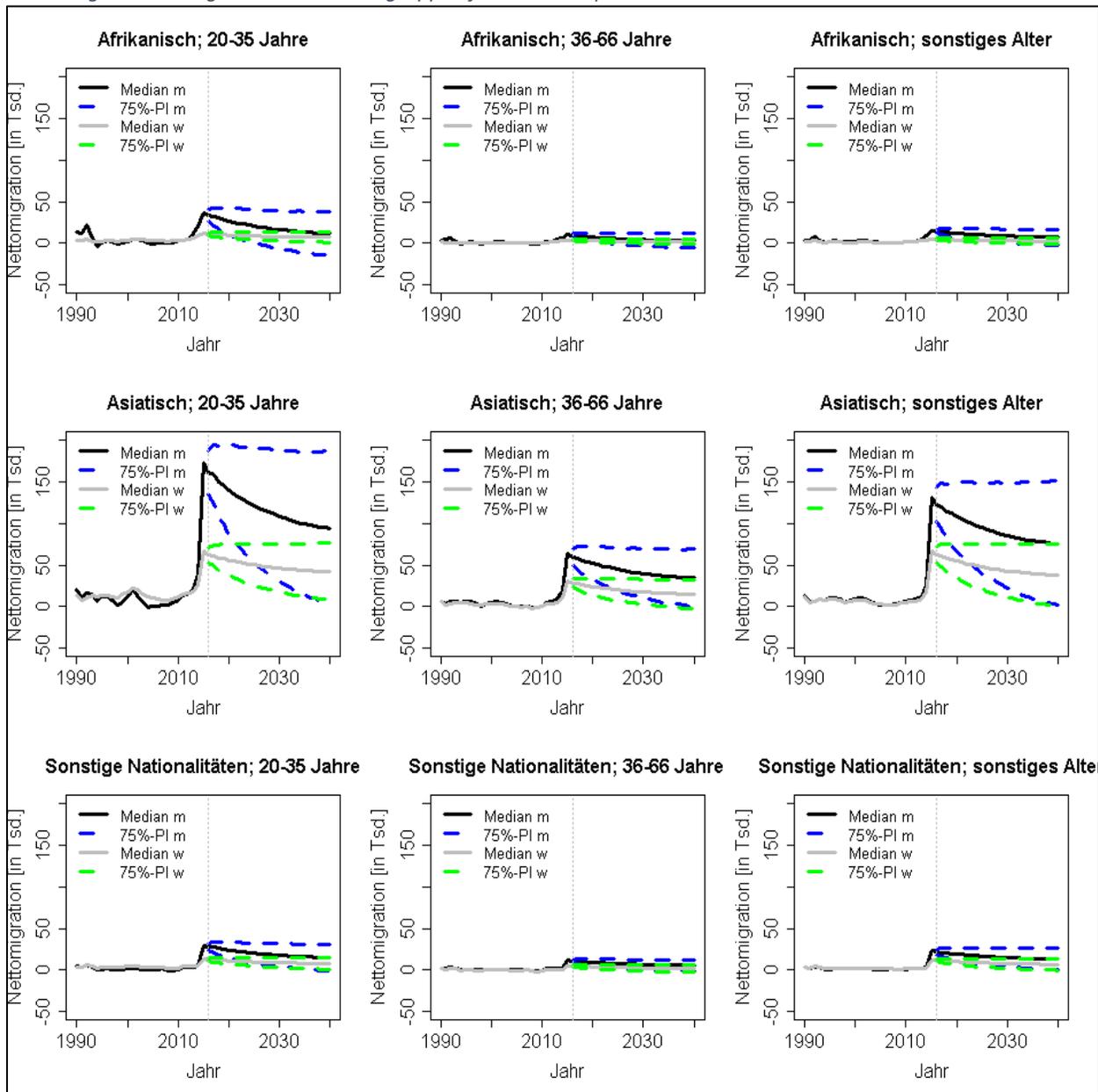
Bei den Deutschen fällt die sehr starke Korrelation in der Nettomigration der beiden Geschlechter auf, weiterhin ist eine eher sinkende Nettomigration durch stärkere Arbeitsmigration, Studienaufenthalte im Ausland und Emigration im Alter in Länder mit milderem Klima oder einer attraktiveren Landschaft zu erwarten. Dieses Phänomen wurde bereits in den 1980er Jahren von Rogers et al. ausführlich für die Migrationsbewegungen innerhalb der Vereinigten Staaten in die Südstaaten und ebenfalls für internationale Migrationen untersucht (vgl. Rogers und Watkins 1987: 490-509; Rogers und Castro 1981: 6). Dabei wurden in vielen Migrationszyklen Peaks zu

Beginn des Rentenalters identifiziert. Besonders die spanischen Inseln, Österreich und die Schweiz sind im Ruhestand für deutsche Auswanderer beliebte Ziele, da diese Regionen wärmere Temperaturen und/oder attraktivere Landschaften als Deutschland aufweisen. Zudem entstehen den deutschen Migranten vergleichsweise geringe Anpassungsprobleme beim Umzug in die angesprochenen Regionen, da entweder bereits eine hinreichend große deutsche Population dort lebt oder aber die Länder selbst deutschsprachig sind (vgl. Hildebrandt-Woeckl 2015). Zudem ist der Rentenbezug auch bei Wohnsitzverlegung in ein anderes EU-Land unproblematisch (vgl. Deutsche Rentenversicherung 2017).

Bei den EU- und Schengen-Ländern im Ausland nimmt das Modell ein moderates Absinken vom Niveau in 2015 an. Das impliziert, dass entweder vermehrt wieder in die südlichen Länder remigriert wird, oder das Auswanderungsniveau in den Krisenländern leicht abnimmt. Dabei würde davon ausgegangen werden, dass sich die wirtschaftliche Situation in den fraglichen Ländern langsam stabilisiert. Das erscheint in Spanien oder Portugal realistisch, ist aber in Italien und Griechenland fraglich, obwohl optimistische Schätzungen von einer Stabilisierung dieser beiden Volkswirtschaften ausgehen (vgl. OECD 2016: 162-238). Die hohe Unsicherheit in diesen Annahmen schlägt sich in den sehr breiten Prognoseintervallen nieder, vor allem in den Erwerbsaltern. Dabei ist zu beobachten, dass seit der Eurokrise die männliche Migration relativ zur weiblichen stark zugenommen hat.

Die Prognose impliziert, dass sich beide Geschlechter in den Erwerbsaltern wieder annähern werden. Bei den sonstigen Europäern lässt sich der Effekt von Krisen sehr deutlich sehen. So sind zu Beginn der 1990er Jahre die starken Migrationsströme in Folge des Untergangs der Sowjetunion, die Einwanderung der Sudetendeutschen und die Konsequenzen der jugoslawischen Bürgerkriege bei den jungen Altern zu beobachten. Dabei fällt auf, dass das Migrationsniveau bei den Frauen in ruhigeren Zeiten stabil etwas über dem männlichen liegt, in Krisenzeiten aber eine verstärkte männliche Migration nach Deutschland zu beobachten ist. Das ist auch in Folge der Finanzkrise in Verbindung mit der Politik der Bundesregierung in der Flüchtlingskrise 2015 zuletzt zu beobachten, was viele Männer aus den ärmeren Balkanländern veranlasste, Asylanträge in Deutschland zu stellen (vgl. Bundesamt für Migration und Flüchtlinge 2016: 118-123).

Abbildung 7: Nettomigration nach Altersgruppen für Nichteuropäer bis 2040



Quelle: eigene Berechnung und Darstellung

Bei den nachfolgenden Nationalitätsgruppen ist vor allem die Flucht vor Krieg, Terrorismus und Armut in den letzten Jahren deutlich zu beobachten, was vor allem seit dem sogenannten Arabischen Frühling und in der Folge durch das Erstarken des islamistischen Terrors in Form des IS zunehmend zu beobachten ist, wobei die Bedrohungslage durch die Taliban nach wie vor sehr hoch ist. Das erklärt, dass die, zuvor nahezu non-existenten, Unterschiede in Migrationszahlen zwischen den Geschlechtern in den letzten Jahren so hoch sind. Auch hier wird in der Prognose im Mittel von einer Stabilisierung der Situation in den Krisenregionen ausgegangen, was in einer

Normalisierung des Migrationsniveaus zwischen diesen Ländern und Deutschland resultieren würde. Das würde voraussetzen, dass sich militärische Erfolge gegen den IS, die Taliban, Boko Haram und andere Terrorgruppen einstellen würden. Das ist an einigen Stellen auch bereits erkennbar (vgl. Heidelberg Institute for International Conflict Research 2017: 189-197), allerdings ist dies ein sehr langfristiger Prozess, dessen Verbesserungen sich erst allmählich einstellen werden. Die Gefechte dürften sich noch lange hinziehen. Manche Militärexperten gehen von einem mehrere Dekaden andauernden Krieg gegen den IS und andere islamistische Gruppen aus (vgl. USA Today 2014). Selbst bei Zerschlagung des islamistischen Terrors würde sich ein lange andauernder Prozess des Wiederaufbaus der betroffenen Regionen einstellen, der immense Hilfen seitens des Westens, Russlands und des Irans erfordern dürfte (vgl. World Bank 2016). Die breiten Prognoseintervalle der Nettomigration spiegeln die Unsicherheit dieses Prozesses wider. Dies gilt insbesondere für männliche Asiaten.

5 Fazit, Limitationen und Ausblick

Der vorliegende Beitrag adressiert die Problemstellungen, mit denen Wissenschaftler im Allgemeinen bei der Modellierung und im Speziellen bei der Prognose internationaler Migration konfrontiert sind. Dabei wurde auf einige Lösungsansätze eingegangen, ein eigenes Verfahren diskutiert und dieses in einer praktischen Anwendung umgesetzt. Mit Hilfe zweier Datensätze, die vom Statistischen Bundesamt zur Verfügung gestellt wurden, wurde ein synthetischer Datensatz erstellt, der Schätzwerte für internationale Migrationsbewegungen über die drei Dimensionen Alter, Geschlecht und Nationalitätsgruppe für den Basiszeitraum 1990-2015 enthält. Dieser Datensatz wurde für die Prognose der alters-, geschlechts- und nationalitätsspezifischen Nettomigrationszahlen zwischen Deutschland und dem Ausland bis ins Jahr 2040 genutzt. Dabei wurde der von Vanella (2017) vorgeschlagene Hauptkomponenten-Zeitreihenansatz abgewandelt und zur detaillierten Simulation der Nettomigrationszahlen genutzt. Neben dem mittleren Szenario wurden durch 10.000-fache Simulation von Wiener Prozessen Prognoseintervalle für jede Nettomigrationszahl erzeugt. In der Arbeit wurden neben der gesamten Nettomigration auch 36 Untergruppen mit Prognoseintervallen illustriert. Tatsächlich schätzt das Modell jedoch den zukünftigen Verlauf von allen 1.414 Variablen simultan. Im Mittel

ist eine sinkende Nettomigration vom zuletzt historisch hohen Niveau zu erwarten, allerdings nicht so stark abfallend, wie in den meisten Studien angenommen.

Die Limitationen des Modells sind vielfältig. So ist der Basiszeitraum mit 26 Jahren relativ kurz, da Daten vor 1990 aufgrund der Teilung Deutschlands nicht mehr für die Analyse heutiger und zukünftiger Migrationsströme aussagekräftig wären. Weiterhin lagen die Daten nicht in der für die Analyse erforderlichen Konsistenz und Vollständigkeit vor, sodass auf Basis von Annahmen Interpolationen durchgeführt werden mussten. Die hohe Stochastizität und Abhängigkeit der Migration von exogenen politischen, sozialen, wirtschaftlichen und klimatischen Faktoren macht die Prognose zukünftiger Migration generell schwierig und erfordert Annahmen und Vereinfachungen, um ein Prognosemodell handhabbar zu machen. In diesem Rahmen mussten Annahmen zur zukünftigen Entwicklung der Eurokrise und der humanitären Krise in den arabischen Ländern getroffen werden. Das ist grundsätzlich ein subjektiver Prozess, welcher allerdings durch die Anpassung geeigneter statistischer Modelle etwas objektiviert wurde. Die Quantifizierung des Risikos durch Prognoseintervalle relativiert generell die ursprünglichen Annahmen etwas, sodass das Modell wieder an Aussagekraft gewinnt.

Quantitative Modelle unterliegen generell der Schwäche, sich lediglich auf die Historie zu beziehen und unprognostizierbare Schocks nicht einzubeziehen. Subjektive Annahmen könnten solche Möglichkeiten integrieren, was allerdings keine echten Wahrscheinlichkeiten und daher keine Prognoseintervalle, sondern lediglich *Credible Intervals*, wie sie im Bayesianischen Vokabular genannt werden¹¹, erzeugen würde. Da wir das Ziel verfolgen, eine möglichst gute Prognose, anstelle einer Projektion, zu erstellen, haben wir auf derartige Modelle verzichtet.

Generell lässt sich konstatieren, dass die kommenden Jahre viel Aufschluss über die weitere Entwicklung geben werden. Neben dem Hinzukommen neuer Daten und einer daraus folgenden größeren Stichprobe wird sich zeigen, ob die bisher getroffenen politischen Maßnahmen dauerhaft Früchte tragen werden oder die Flüchtlingskrise weitergehen wird. Sollte sich die Entwicklung am aktuellen Rand grundsätzlich anders gestalten als in den Annahmen oder Modellparametern festgelegt, ist es Aufgabe der Wissenschaftler die Ergebnisse zu aktualisieren

¹¹ Siehe hierzu z.B. Lee 2012: 54-55.

oder die Entwicklung neu zu modellieren. Denn bereits existierende Prognosen und Vorausberechnungen erheben nicht den Anspruch, die absolute Wahrheit vorzuberechnen. Sie dienen Entscheidern lediglich als quantitative Datengrundlage.

Literatur

Bewarder, Manuel; Leubacher, Marcel 2016: Syrische Staatsbürgerschaft wird massenhaft vorgetäuscht. Abgerufen unter <https://www.welt.de/politik/deutschland/article156496638/Syrische-Staatsbuergerschaft-wird-massenhaft-vorgetaeuscht.html> am 20.06.2017.

Bijak, Jakub 2011: Forecasting International Migration in Europe. A Bayesian View. Springer.

Bohk, Christina 2012: Ein probabilistisches Bevölkerungsprognosemodell. Entwicklung und Anwendung für Deutschland. Springer VS.

Box, George; Jenkins, Gwilym; Reinsel, Gregory; Ljung, Greta 2016: Time Series Analysis. Forecasting and Control. 5. Auflage. Wiley.

Bundesamt für Migration und Flüchtlinge 2016: Migrationsbericht 2015: Nürnberg.

Bundesamt für Migration und Flüchtlinge 2017: Das Bundesamt in Zahlen 2016. Asyl, Migration und Integration: Nürnberg.

Deschermeier, Philipp 2011: Population Development of the Rhine-Neckar Metropolitan Area: A Stochastic Population Forecast on the Basis of Functional Data Analysis. In: Comparative Population Studies 36(4): 769–806.

Deschermeier, Philipp 2016: Einfluss der Zuwanderung auf die demografische Entwicklung in Deutschland. In: IW-Trends, Vierteljahresschrift zur empirischen Wirtschaftsforschung 43(2): 21–38.

Deutsche Rentenversicherung 2017: Rente im Ausland. Abgerufen unter http://www.deutsche-rentenversicherung.de/Allgemein/de/Navigation/1_Lebenslagen/05_Kurz_vor_und_in_der_Rente/01_Kurz_vor_der_Rente/08_Rente_im_Ausland/Rente_im_Ausland_node.html am 20.06.2017.

Fuchs, Johann; Söhnlein, Doris; Weber, Brigitte; Weber, Enzo 2017: Forecasting Labour Supply and Population: An Integrated Stochastic Model. IAB Discussion Paper 1/2017: Nürnberg.

Heidelberg Institute for International Conflict Research 2017: Conflict Barometer 2016: Heidelberg.

Hildebrandt-Woeckl, Sabine 2015: Palmenstrand und deutsche Rente. Abgerufen unter <http://www.faz.net/aktuell/wirtschaft/immobilien/wie-kompliziert-ist-es-die-rente-im-ausland-zu-organisieren-13886953.html> am 20.06.2017.

Hyndman, Rob; Booth, Heather 2008: Stochastic population forecasts using functional data models for mortality, fertility and migration, in: *International Journal of Forecasting* 24(3): 323–342.

Hyndman, Rob J.; Ullah, Shahid 2007, Robust forecasting of mortality and fertility rates. A functional data approach. In: *Computational Statistics & Data Analysis*, 51(10): 4942–4956.

Hyndman, Rob J.; Booth, Heather; Yasmineen, Farah 2013, Coherent mortality forecasting: the product-ratio method with functional time series models. In: *Demography*, 50(1): 261–283.

Kupiszewska, Dorota; Nowok, Beata: Comparability of statistics on international migration flows in the European Union. In: Raymer, James; Willekens, Frans [Hrsg.] 2008: *International Migration in Europe*. Wiley.

Lee, Peter 2012: *Bayesian Statistics. An Introduction*. Wiley.

Leubecher, Marcel 2017: Familiennachzug von Migranten 2016 um 50 Prozent gestiegen. Angerufen unter <https://www.welt.de/politik/deutschland/article161307823/Familiennachzug-von-Migranten-2016-um-50-Prozent-gestiegen.html> am 20.06.2017.

Nieden, Felix zur; Rau, Roland; Luy, Marc 2016 Allgemeine Sterbetafel 2010/2012 – Neue Ansätze zur Glättung und Extrapolation der Sterbewahrscheinlichkeiten. In: *Wirtschaft und Statistik* 1: 63–74.

OECD 2016: *OECD Economic Outlook. Volume 2016 Issue 2*, OECD Publishing, Paris. http://dx.doi.org/10.1787/eco_outlook-v2016-2-en.

Ramsay, James; Silverman, Bernard 2001, *Functional Data Analysis*. In: Smelser, Neil; Baltes, Paul: *International Encyclopedia of the Social and Behavioral Sciences*: 5822–5828.

Ramsay, James; Silverman, Bernard W. 2005, *Functional Data Analysis*, Springer Series in Statistics: New York.

Raymer, James; Willekens, Frans 2008: International Migration in Europe. Wiley.

Rogers, Andrei; Castro, Luis 1981: Model Migration Schedules. Working Paper RR-81-30. International Institute for Applied Systems Analysis: Laxenburg.

Rogers, Andrei; Little, Jani; Raymer, James 2010: The Indirect Estimation of Migration. Springer.

Rogers, Andrei; Watkins, John 1987: General Versus Elderly Interstate Migration and Population Redistribution in the United States. In: Research on Aging 9(4): 483–529.

Statistisches Bundesamt 2015: Bevölkerung Deutschlands bis 2060. 13. koordinierte Bevölkerungsvorausberechnung: Wiesbaden.

Statistisches Bundesamt 2015a: Wanderungen zwischen Deutschland und dem Ausland 1991-2013 nach Einzelaltersjahren und Geschlecht. Daten auf Anfrage zur Verfügung gestellt am 02.09.2015.

Statistisches Bundesamt 2016: Wanderungen zwischen Deutschland und dem Ausland 2014 nach Einzelaltersjahren und Geschlecht. Daten auf Anfrage zur Verfügung gestellt am 14.03.2016.

Statistisches Bundesamt 2017: Wanderungen über die Grenzen Deutschlands 1990-2015 nach (ausgewählten) Staatsangehörigkeiten und Altersgruppen. Daten auf Anfrage zur Verfügung gestellt am 18.04.2017.

Statistisches Bundesamt 2017a: Wanderungen zwischen Deutschland und dem Ausland 2015 nach Einzelaltersjahren und Geschlecht. Daten auf Anfrage zur Verfügung gestellt am 24.04.2017.

Statistisches Bundesamt 2017b: Aktualisierung der 13. koordinierten Bevölkerungsvorausberechnung – Basis 2015. Abgerufen unter <https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Bevoelkerungsvorausberechnung/AktualisierungBevoelkerungsvorausberechnung.html> am 29.08.2018.

Tavella, Domingo 2002: Quantitative Methods in Derivatives Pricing. An Introduction to Computational Finance. Wiley&Sons.

United Nations 1998: Recommendations on Statistics of International Migration. Revision 1. Statistical Papers Series M, No. 58, Rev. 1, Department of Economic and Social Affairs, Statistics Division: New York.

USA Today 2014: Panetta: '30-year war' and a leadership test for Obama. Aufgerufen unter [https:// www.usatoday.com/story/news/politics/2014/10/06/leon-panetta-memoir-worthy-fights/16737615/](https://www.usatoday.com/story/news/politics/2014/10/06/leon-panetta-memoir-worthy-fights/16737615/) am 19.06.2017.

Vanella, Patrizio 2017: Stochastische Prognose demografischer Komponenten auf Basis der Hauptkomponentenanalyse. Hannover Economic Papers 597, Juni 2017.

Wood, Simon 1994: Obtaining Birth and Mortality Patterns From Structured Population Trajectories. In: Ecological Monographs 64(1): 23–44.

World Bank 2016: Transcript of World Bank Group President Jim Yong Kim's Opening Press Conference. Abgerufen unter <http://www.worldbank.org/en/news/speech/2016/04/14/transcript-world-bank-group-president-jim-yong-kim-opening-press-conference> am 20.06.2017.