

Das Forschungsdatenzentrum der gesetzlichen Rentenversicherung (FDZ-RV) im Aufbau

Bericht vom ersten Workshop des FDZ-RV
am 28. und 29. Juni 2004 in Würzburg

Oktober 2004



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

VDR ————— FNA

Verband
Deutscher
Rentenversicherungsträger

Forschungs-
Netzwerk
Alterssicherung

Das Forschungsdatenzentrum der gesetzlichen Rentenversicherung (FDZ-RV) im Aufbau

Bericht
vom ersten Workshop des FDZ-RV
am 28. und 29. Juni 2004 in Würzburg

(Bibliografische Information der Deutschen Bibliothek:

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Das Forschungsdatenzentrum der gesetzlichen Rentenversicherung (FDZ-RV) im Aufbau.
Bericht vom ersten Workshop des FDZ-RV am 28. und 29. Juni 2004 in Würzburg
Hrsg.: Verband Deutscher Rentenversicherungsträger/Forschungsnetzwerk Alterssicherung
(DRV-Schriften; Bd. 55, Sonderausgabe der DRV)
ISBN 3-926181-91-5

Herausgegeben vom Verband Deutscher Rentenversicherungsträger.
Verantwortlich für den Gesamtinhalt: Hauptschriftleiter Prof. Dr. Franz Ruland, Stellvertreter Dr. Axel Reimann,
Verband Deutscher Rentenversicherungsträger, Eysseneckstraße 55, 60322 Frankfurt am Main, Fernruf (0 69) 15 22-0,
Schriftleiter Dr. Dirk von der Heide, Verband Deutscher Rentenversicherungsträger, Berliner Büro, Albrechtstr. 10 C,
10117 Berlin, Fernruf (0 30) 28 88 65 - 11,

Verlag und Anzeigenverwaltung: wdv Gesellschaft für Medien & Kommunikation mbH & Co.OHG, Postfach 2551, 61295 Bad Homburg, Fernruf (0 61 72) 6 70 - 0, Verlagsort Bad Homburg.
Gesamtherstellung: Central-Druck Trost GmbH & Co KG, Industriestraße 2, 63131 Heusenstamm, Fernruf (0 61 04) 6 06 - 0.
Verantwortlich für den Anzeigenteil: Walter Piezonka, für Marketing und Vertrieb: Bernd Kremer.

Die Zeitschrift DEUTSCHE RENTENVERSICHERUNG mit den Fachmitteilungen der Deutschen Rentenversicherung erscheint 12mal jährlich. Preis der Einzelfolge 8,20 Euro incl. MwSt. Bestellungen nehmen entgegen: der Verlag und der Buchhandel. Abbestellungen nur mit halbjähriger Frist zum 30.06. und 31. 12. jedes Jahres. Zahlung jeweils jährlich im Voraus an: wdv, Postbank Frankfurt am Main, Konto-Nr. 773 08 603, BLZ 500 100 60, Bankkonten: Deutsche Bank AG, Hanau, Konto-Nr. 040 96 49, BLZ: 506 700 09, Frankfurter Sparkasse, Konto-Nr. 705 665, BLZ 500 502 01.

Die mit Namen gekennzeichneten Beiträge geben die Auffassung der Verfasser, aber nicht des Herausgebers wieder. Die Zeitschrift nimmt nur Originalbeiträge an. Der Nachdruck von Beiträgen ist nur mit Einwilligung der Schriftleitung unter Quellenangabe gestattet. Beiträge sind an die Schriftleitung, Eysseneckstraße 55, 60322 Frankfurt am Main, zu richten. Für unverlangt eingesandte Manuskripte und Besprechungsexemplare übernimmt die Schriftleitung keine Gewähr.

Verlag: wdv Gesellschaft für Medien und Kommunikation mbH & Co.OHG, Siemensstraße 6, 61352 Bad Homburg v.d.H. HRA 3087, Bad Homburg v.d.H., Pers.haft. Gesellschafter: Zeitschriften VVG Verlags- und Verwaltungs-Gesellschaft mbH & Co. KG, HRA 3096, Bad Homburg v.d.H. sowie VVG Gesellschaft zur Verlagsbeteiligung und Verwaltung mbH, HRB 5544, vertreten durch die Geschäftsführer Adolf Hilger, Thomas Kuhn, Rolf. M. Laufer, Klaus Tonello, jeweils Siemensstraße 6, 61352 Bad Homburg v.d.H

INHALTSVERZEICHNIS

Vorwort	5
Einleitung	7
1. FDZ-RV und Datenproduzenten	11
1.1 FDZ-RV und Rentenversicherung	11
<i>Uwe G. Rehfeld:</i> Zur Genese des Forschungsdatenzentrums der gesetzlichen Rentenversicherung	12
<i>Thomas Bütefeld:</i> Datenwege und praktischer Datenzugang	20
<i>Hilmar Luckert:</i> Statistikdaten der gesetzlichen Rentenversicherung – ein grober Überblick	24
<i>Claudia Heese:</i> Aspekte des Datenschutzes im Forschungsdatenzentrum der Rentenversicherung	41
<i>Dr. Ralf K. Himmelreicher:</i> Befunde einer schriftlichen Befragung unter potenziellen Nutzerinnen und Nutzern des Forschungsdatenzentrums der gesetzlichen Rentenversicherung	49
1.2 Forschungsdatenzentren der Statistischen Ämter und der BA	65
<i>Dr. Sylvia Zühlke, Markus Zwick, Sebastian Scharnhorst und Thomas Wende:</i> Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder	66
<i>Dr. Annette Kohlmann:</i> Datenzugang und Datenverfügbarkeit im Forschungsdatenzentrum der Bundes- agentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung	77
2. FDZ-RV und Wissenschaft	89
2.1 Erfahrungen der Wissenschaft mit Daten der RV	89
<i>Dr. Thorsten Heien:</i> Erste Erfahrungen der Wissenschaft mit Daten der RV: Die Studie „Altersvorsorge in Deutschland“ (AVID)	90
<i>Stefan Kempf und Prof. Dr. Martin Kukuk:</i> Determinanten des Renteneintritts – eine Untersuchung mit prozessproduzierten Daten	101

2.2 Projektskizzen und Datenbedarf der Wissenschaft	113
<i>Sabine Böttcher:</i> Übergänge in den Ruhestand	115
<i>Renate Büttner und PD Dr. Matthias Knuth:</i> Berichtssystem zum Altersübergang aus der Sicht von Beschäftigten und Betrieben (Altersübergangs-Monitor)	118
<i>Hans Martin von Gaudecker:</i> Intragenerationale Umverteilungswirkungen in der gesetzlichen Rentenver- sicherung	126
<i>Stefan Hupfeld:</i> Rentenversicherung und heterogene Lebenserwartung	130
<i>Prof. Dr. Barbara Riedmüller und Michaela Willert:</i> Erwerbstätigkeit und Altersmortalität	134
<i>Dr. Rembrandt Scholz:</i> Datenbedarf in der Mortalitätsforschung in Deutschland	141
<i>Dr. Rainer Unger:</i> Zielvorstellungen in der Alterssicherung: Konkretisierung, Operationalisierung und Messung als Grundlage für normative Aussagen	147
<i>Martin Grub:</i> Analysesystem Alterssicherung	152
<i>Dr. Holger Seibert:</i> Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland	156
<i>Dr. Volker Arndt, PD Dr. Dietrich Rothenbacher und Prof. Dr. Hermann Brenner:</i> Berentungen aufgrund verminderter Erwerbsfähigkeit bei Arbeitnehmern in der Bauwirtschaft	162
<i>Rainer Kaluscha:</i> Nutzung von Reha-Routinedaten für die Forschung	169
3. Zusammenfassung	173
Teilnehmerliste	183

Vorwort

Der vorliegende Band ist das Ergebnis des ersten Workshops des sich im Aufbau befindenden Forschungsdatenzentrums der Rentenversicherung (FDZ-RV). Auf Vorschlag der Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) sollen in Zukunft prozessproduzierte Daten der sozialen Sicherungssysteme der Wissenschaft zur Verfügung gestellt werden. Um den Stand der Kenntnis vor allem im Bereich der Alterssicherung durch empirisch fundierte wissenschaftliche Analysen zu verbessern, ist das FDZ-RV im Januar 2004 gegründet worden. Das Forschungsdatenzentrum wird auf Empfehlung des Gründungsausschusses des Rates für Sozial- und Wirtschaftsdaten vom Ministerium für Bildung und Forschung (BMBF) für zwei Jahre gefördert und ist während dieser Erprobungsphase beim Verband Deutscher Rentenversicherungsträger (VDR) angesiedelt. Dort ist es organisatorisch in das Forschungsnetzwerk Alterssicherung (FNA) eingebunden. Mit der Einrichtung des FDZ-RV wird angestrebt, der wissenschaftlichen Forschung einen nutzerfreundlichen und transparenten Zugriff auf Daten der Rentenversicherung zu ermöglichen.

Am 28. und 29. Juni 2004 fand der erste Workshop des FDZ-RV statt. Dieses Vorhaben war mit dem Ziel verbunden, Details der Ausgestaltung des Forschungsdatenzentrums zu diskutieren und die Kommunikation zwischen Wissenschaft und Rentenversicherung anzuregen. 40 Personen aus zahlreichen wissenschaftlichen Forschungsinstituten, den Forschungsdatenzentren der BA und denen der Statistischen Ämter des Bundes und der Länder sowie aus der Rentenversicherung kamen hierzu im Bildungszentrum des VDR in Würzburg zusammen. Die Arbeitstagung kann als Auftaktveranstaltung für einen Prozess verstanden werden, in dessen Verlauf die Möglichkeiten des Zugriffs auf Mikrodaten der Rentenversicherung sukzessive verbessert werden sollen. Die Ergebnisse dieses ersten Workshops des FDZ-RV, der insbesondere seitens der Wissenschaft ein positives Echo auslöste, werden im vorliegenden Tagungsband publiziert.

Das Forschungsdatenzentrum der Rentenversicherung fungiert als Serviceeinrichtung für die Wissenschaft: Durch diesen intermediären Charakter des FDZ-RV zwischen Datenproduzent und Wissenschaft ergeben sich zwei verschiedene Blickwinkel auf das Forschungsdatenzentrum. Der Tagungsband dokumentiert den produktiven Dialog aus beiden Sichtweisen und fasst die wichtigsten Aspekte in einem weiteren Abschnitt zusammen. Entsprechend dem Verlauf des Workshops sind die Beiträge in diesem Band demnach in drei Kapitel gegliedert.

Im ersten Kapitel wird die Verankerung des FDZ-RV in der Rentenversicherung und das Verhältnis zu Forschungsdatenzentren anderer öffentlicher Datenproduzenten beleuchtet. Zunächst stellen Mitarbeiter des VDR den Stand des Aufbaus des Forschungsdatenzentrums dar und erläutern grundlegende Rahmenbedingungen: das Datenangebot, die verschiedenen Zugangswege, wie diese Daten zu beziehen sein werden, die Regeln des Datenschutzes sowie die Befunde einer ersten Umfrage unter potenziellen Nutzern des FDZ-RV. Anschließend steuern Mitarbeiterinnen und Mitarbeiter anderer öffentlicher Forschungsdatenzentren ihre Erfahrungen im Hinblick auf den Aufbau wie auf die Weitergabe von Mikrodaten bei.

Im zweiten Kapitel steht der Datenbedarf der Wissenschaft im Vordergrund: Wissenschaftler verschiedener Forschungsinstitutionen artikulieren vor dem Hintergrund laufender Forschungsprojekte ihr Erkenntnisinteresse und ihren spezifischen Bedarf an Mikrodaten der Rentenversicherung.

Im dritten Kapitel werden die Ergebnisse des Workshops zusammengefasst. Ein Ausblick auf die weitere Vorgehensweise im „Forschungsdatenzentrum der gesetzlichen Rentenversicherung (FDZ-RV) im Aufbau“ schließt den Tagungsband ab.

Das FDZ-RV eröffnet im Bereich der Rentenversicherung ein neues Aufgabenfeld: die vom BMBF geförderte Weitergabe von anonymisierten Mikrodaten an die wissenschaftliche Forschung. Damit dieses Vorhaben gelingen kann, werden von der Rentenversicherung wie von der Wissenschaft zunächst gegenseitige Informationen benötigt, um in Zukunft erfolgreich kooperieren zu können. Hierzu soll der vorliegende Tagungsband beitragen.

An diesem Band haben neben den namentlich genannten Autoren einige Personen mitgewirkt, der ohne ihre Unterstützung nicht zustande gekommen wäre: Jonas Radl hat während seines zweimonatigen Praktikums beim VDR sämtliche redaktionellen Arbeiten am vorliegenden Tagungsband durchgeführt und durch sein Engagement maßgeblich zu einer schnellen Drucklegung beigetragen. Wichtige Anregungen hat Tatjana Mika eingebracht; Moritz Mörschel hat die Graphiken aufbereitet und ebenso wie Monika Hütter, Gabriele Moschner, Kim Poulev und Daniel Kromer die Beiträge Korrektur gelesen. Ihnen allen möchte ich danken.

*Dr. Ralf K. Himmelreicher
im Oktober 2004*

Einleitung

„Das Forschungsdatenzentrum der gesetzlichen Rentenversicherung im Aufbau“. Unter diesem Titel fand im Juni 2004 ein Workshop des VDR statt, dessen Ziel die Grundsteinlegung für den Aufbau des FDZ-RV gewesen ist.

Zum Einstieg führt *Uwe G. Rehfeld* in die Entstehungsgeschichte des Forschungsdatenzentrums der Rentenversicherung ein. Er referiert die Gründe, weshalb die Verfügbarmachung von Mikrodaten der Rentenversicherung von hoher Wichtigkeit für das System der Alterssicherung wie für die wissenschaftliche Forschung ist.

Thomas Bütetfisch illustriert, wie die Daten der Rentenversicherung über die Träger bis zum VDR gelangen, wie beim Verband das statistische Berichtswesen aufgebaut ist und auf welchen Wegen sie in Zukunft der Wissenschaft zur Verfügung gestellt werden sollen.

Hilmar Luckert gibt einen Überblick über die verschiedenen Statistiken der gesetzlichen Rentenversicherung. Er benennt spezielle Charakteristika und beschreibt die abgebildeten Informationen der Datensätze, die für eine Aufbereitung durch das FDZ-RV in Frage kommen.

Die für das Forschungsdatenzentrum besonders bedeutsamen Aspekte des Datenschutzes werden von *Claudia Heese* erörtert. Ihr Beitrag hebt darauf ab, die Grenzen der Forschungsmöglichkeiten mit Sozialdaten der Rentenversicherung abzustecken und interessierten Forschern die zu erfüllenden juristischen Kriterien darzulegen.

Im Vorfeld der Tagung zum Aufbau des FDZ-RV wurde eine Befragung zu den von der Wissenschaft gewünschten Datensätzen und Zugangswegen durchgeführt, deren Befunde von *Dr. Ralf K. Himmelreicher* dargestellt werden: Sein Beitrag, der die Datenwünsche der Wissenschaft fokussiert, vermittelt ein klares Bild von den Präferenzen der beteiligten Forschergruppen.

Dr. Sylvia Zühlke, Markus Zwick, Sebastian Scharnhorst und *Thomas Wende* berichten aus den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder, die bereits seit 2001 Mikrodaten der amtlichen Statistik für die Wissenschaft bereitstellen. Gleichzeitig führt ihre Darstellung allgemein in das Datenangebot der amtlichen Statistik in Deutschland ein.

Das *Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit* betreibt seit April 2004 ein Forschungsdatenzentrum. *Dr. Annette Kohlmann* charakterisiert in ihrem Beitrag einerseits die dort gegenwärtig und zukünftig verfügbaren Datensätze und zeigt andererseits die praktischen Zugangswege zu verschiedenen Arbeitsmarktdaten auf.

Seit 1996 werden prozessproduzierte Daten der Rentenversicherung im Rahmen der Studie „Altersvorsorge in Deutschland (AVID)“ ausgewertet. Hieraus gewonnene Erfahrungen im Umgang mit Konten von Rentenversicherten stellt *Dr. Thorsten Heien* von *TNS Infratest Sozialforschung* vor.

Im Anschluss präsentieren *Stefan Kempf* und *Prof. Dr. Martin Kukuk* von der *Universität Würzburg* Befunde ihrer Analysen mit dem Optionswertmodell, für die sie einen Pilotdatensatz der

Versichertenkontenstichprobe 2001 des VDR verwendeten. Ein zentraler Befund ist die starke Präferenz der älteren Versicherten für Freizeit, durch die vorzeitige Eintritte in den Ruhestand erklärt werden können.

Am *Zentrum für Sozialforschung Halle* wird aktuell am Projekt „ost-mobil – Analysen zur Mobilität auf dem ostdeutschen Arbeitsmarkt“ gearbeitet. *Sabine Böttcher* stellt die Konzeption von „Übergängen in den Ruhestand“ vor. Sie untersucht, inwieweit sich Arbeitslosigkeit als Übergangsstadium in den Ruhestand verfestigt hat.

Renate Büttner und *Dr. Matthias Knuth* vom *Institut Arbeit und Technik* erforschen ebenfalls solche Prozesse, die von der Erwerbstätigkeit in die Altersrente führen. Ihr Projekt eines „Altersübergangs-Monitors“ strebt den Aufbau eines Berichtssystems an, das die verschiedenen Datenquellen zum Altersübergangsgeschehen zusammenführt und kontinuierlich auswertet.

Ausgehend von einem positiven Zusammenhang zwischen Einkommenshöhe und Lebensalter untersucht *Hans-Martin von Gaudecker* vom *MEA* die Frage nach der intragenerationalen Umverteilung in der gesetzlichen Rentenversicherung mit einem ökonomischen Modell.

Stefan Hupfeld von der *Universität Konstanz* analysiert den Zusammenhang zwischen Wohlstand und Gesundheit. Ferner geht er der Frage nach, wie Reaktionen auf erwartete Umverteilungswirkungen in der Rentenversicherung sich auf die Entscheidung über den Zeitpunkt des Renteneintritts auswirken.

Der Trend zur steigenden Lebenserwartung wird besonders von der Langlebigkeit von Frauen getragen. *Prof. Dr. Barbara Riedmüller* und *Michaela Willert* von der *Freien Universität Berlin* wollen herausfinden, ob sich der geschlechtsspezifische Unterschied in der Altersmortalität durch kürzere Erwerbszeiten im Lebenslauf der Frauen erklären lässt.

Dr. Rembrandt Scholz forscht am *Max-Planck-Institut für demografische Forschung* in Rostock zur Mortalitätsentwicklung in Deutschland. Der Datenbedarf für die Erforschung von differentiellen Sterblichkeiten ergibt sich vor allem aus der eingeschränkten Validität der Daten der amtlichen Statistik im hohen Alter. Mit den Mikrodaten der Rentenversicherung könnten hierzu verlässlichere Analysen durchgeführt werden.

Am *Zentrum für Sozialpolitik der Universität Bremen* möchte *Dr. Rainer Unger* die Leistungsseite der Rentenversicherung quantifizieren. Das Absicherungsniveau soll unter Berücksichtigung des Haushaltszusammenhangs und der verschiedenen Einkommensquellen im Alter untersucht werden.

Martin Grub entwickelt am *Fraunhofer Institut für Angewandte Informationstechnik* das „Anlysesystem Alterssicherung“. Dabei handelt es sich um ein Lebenszyklusmodell, das Verteilungswirkungen und Anreize politischer Reformen abbildet und rationale Erwartungen in das Entscheidungsmodell integriert.

Mit „Ausbildungs- und Berufsverläufen der Geburtskohorten 1964 und 1971“ beschäftigt sich ein Forschungsprojekt des *Max-Planck-Instituts für Bildungsforschung*. *Dr. Holger Seibert* stellt das Projektes vor, welches theoretisch auf dem Begriff des Lebensverlaufs aufbaut und den

Wandel der institutionellen Verknüpfungen von Bildungssystem und Arbeitsmarkt in Deutschland thematisiert.

Dr. Volker Arndt, PD Dr. Dietrich Rothenbacher und Prof. Dr. Hermann Brenner erforschen am *Deutschen Zentrum für Altersforschung* sozialmedizinische Risikofaktoren für Frühverrentungen in der Bauwirtschaft. Mit diesem Projekt wird angestrebt, jene Diagnosen zu identifizieren, die besonders häufig Frühverrentungen und Rehabilitationsmaßnahmen nach sich ziehen.

Von *Rainer Kaluscha* vom *Forschungsinstitut für Rehabilitationsmedizin an der Universität Ulm* wird das im Rahmen des Projekts „Reha-Net“ entwickelte Patientenkonto erläutert. Die Forschungsdatenbank speist sich aus Daten des VDR sowie eines Qualitätssicherungsbogens und erlaubt die Untersuchung vielfältiger Fragestellungen. Sein Beitrag schließt die Reihe der Projektskizzen der Wissenschaftlerinnen und Wissenschaftler ab.

Im Zentrum des Workshops und somit auch im vorliegenden Tagungsband stand die Absicht, neben der Darstellung vor allem des Datenangebots des FDZ-RV und anderer öffentlicher Forschungsdatenzentren eine gute Basis zur Kooperation mit den Wissenschaftlerinnen und Wissenschaftlern herzustellen. Das Erfassen der wissenschaftlichen Zielsetzungen der Forschergruppen, der damit in Zusammenhang stehenden Datenwünsche und konkret die für einzelne Forschungsprojekte benötigten Variablen bilden den Schwerpunkt dieses Tagungsbandes.

Um die Verständigung zu vereinfachen sowie eine leichtere Orientierung für Nutzerinnen und Nutzer zu ermöglichen, sollen in Zukunft einheitliche Begrifflichkeiten in den Publikationen der öffentlichen Forschungsdatenzentren etabliert werden. Entsprechende Absprachen wurden im *Arbeitskreis öffentliche Forschungsdatenzentren (AKöFFFDZ)* getroffen und in diesem Tagungsband berücksichtigt. Diese Sprachregelungen beziehen sich insbesondere auf die Begriffe: Scientific Use Files, Public Use Files, Gastwissenschaftler-Arbeitsplätze, kontrollierte Datenfernverarbeitung, Campus Files und Strukturdatensätze.

1. FDZ-RV und Datenproduzenten

Im ersten Kapitel des vorliegenden Bandes werden die vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Forschungsdatenzentren der öffentlichen Datenproduzenten vorgestellt. Im ersten Abschnitt liegt der Schwerpunkt der Betrachtung auf dem sich im Aufbau befindenden Forschungsdatenzentrum der Rentenversicherung; der zweite Abschnitt beschreibt vor allem das Datenangebot der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder und das der Bundesagentur für Arbeit.

1.1 FDZ-RV und Rentenversicherung

Das Forschungsdatenzentrum der Rentenversicherung ist beim Verband Deutscher Rentenversicherungsträger angesiedelt. Im Folgenden erläutern Mitarbeiter und Mitarbeiterinnen verschiedener Abteilungen des VDR grundlegende Bedingungen, die den Rahmen für das entstehende Forschungsdatenzentrum bilden.

Zuerst wird die Entstehungsgeschichte des Forschungsdatenzentrums nachgezeichnet. Darauf folgt ein Beitrag, der die Wege der Mikrodaten der Rentenversicherung von den Trägern über den Verband bis in das FDZ-RV skizziert.

Ein Überblick über das breite statistische Datenangebot insbesondere von Mikrodaten der Rentenversicherung zeigt, mit welchen Daten in Zukunft wissenschaftliche Analysen möglich sein werden. Es folgt eine Darstellung der Grundzüge des Datenschutzrechts und insbesondere des Sozialdatenschutzrechts, letzteres bildet den unumstößlichen Rahmen der Arbeitsweise im Forschungsdatenzentrum. Abschließend werden die Resultate einer Erhebung unter potentiellen Nutzern des FDZ-RV zu ihrem Bedarf an Mikrodaten der Rentenversicherung und zu präferierten Datenzugangswegen vorgestellt.

Im Rahmen dieser Beiträge des FDZ-RV aus der Perspektive der Rentenversicherung sollen fünf Punkte deutlich werden: (1) warum ein Forschungsdatenzentrum gegründet wurde, (2) auf welche Art und Weise (3) welche Mikrodatensätze zum VDR kommen, (4) welche datenschutzrechtlichen Regelungen zu beachten sind und (5) welche(r) dieser Datensätze seitens der Wissenschaft besonders häufig nachgefragt wird bzw. werden, unter Berücksichtigung präferierter Zugangswege.

Zur Genese des Forschungsdatenzentrums der gesetzlichen Rentenversicherung

Uwe G. Rehfeld (VDR Frankfurt)

Die lange Tradition der Statistik der gesetzlichen Rentenversicherung

Mit der Gründung der gesetzlichen Rentenversicherung hat der Gesetzgeber neben Zielen und Ausgestaltung dieses Sozialversicherungszweiges von Beginn an auch ein Berichts- und Aufsichtswesen vorgesehen. Dabei standen zunächst Abrechnungsfragen, organisatorische Aspekte und sozialpolitische Zwecke im Vordergrund. Seit einigen Jahrzehnten kann jedoch ein Wandel im Verhältnis von Staat und durchführenden Behörden festgestellt werden, der auch Auswirkungen auf das statistische Berichtswesen der gesetzlichen Rentenversicherung hat.

Die zu verzeichnenden strukturellen Veränderungen im kontinuierlich betriebenen statistischen Informationswesen der Rentenversicherung lassen sich im Wesentlichen auf drei Entwicklungstrends zurückführen: Leistungsverbesserungen auf Basis des technologischen Wandels, Fortentwicklung der Gesetzgebung mit gesicherten Informationen der Verwaltung sowie die Analyse und Fortschreibung durch die Wissenschaft (vgl. Rische 1987, Rehfeld 1987/2001).

Leistungsverbesserungen auf Basis des technologischen Wandels

Während vor dem Einsatz bzw. zu Beginn des Einsatzes der elektronischen Datenverarbeitung in der Rentenversicherung die Statistik nur am Rande Nutzer der Datenverarbeitung war und noch vielfache manuelle Tätigkeiten für entsprechend bescheidene statistische Zwecke durchgeführt wurden, ist mit dem technologischen Wandel ein erweitertes, völlig neues Nutzungsfeld entstanden. Die elektronische Datenverarbeitung wird für jegliche Dokumentations- und Verwaltungsarbeiten eingesetzt. Dies macht es möglich, durch die Anwendung analytischer Konzepte und quantitativer Methoden den vorhandenen Verwaltungsdatenbestand unter Wahrung datenschutzrechtlicher Regelungen sekundär-statistisch auszuwerten. Die prozessproduzierten Daten sind in der überwiegenden Mehrheit der Fälle valide und enthalten alle für die gesetzliche Aufgabe der Rentenversicherung notwendigen Informationen. Das heutige Berichtswesen baut auf diesen Daten auf und ist für bestimmte Tatbestände die einzige qualitativ hochwertige, aktuelle und retrospektiv auswertbare Datenquelle.

Fortentwicklung der Gesetzgebung mit gesicherten Informationen der Verwaltung

Die sich zunehmend komplexer gestaltende soziale Sicherung erfordert zur Beurteilung seitens der Versicherten, der Rentner und der politischen Akteure wie auch der Selbstverwaltung und der Verwaltung, entsprechend differenzierte und tiefgegliederte Daten. Es zeigte sich, dass dabei fallgruppenbezogene Verwaltungsdaten alleine nicht ausreichten und Zusatzerhebungen erforderlich wurden. Insbesondere wurde die Perspektive vom Versicherungs- bzw. Rentenfall auf die Kumulation verschiedener Ansprüche einer Person sowie auf die Kumulation verschiedener Alterssicherungsansprüche auf die Einheit der Familie bzw. des Haushalts erweitert. Aus dieser Sichtweise folgte, dass zur Information über Reformvorhaben und zur Beurteilung der sozialen Sicherung sowohl auf prozessproduzierte Daten der Routineberichterstattungen als auch auf zusätzliche Erhebungsdaten abzustellen war. Gegenwärtig besonders erfolgreich ist

eine Kombination von Prozess- und Befragungsdaten. Initiatoren und Auftraggeber bzw. Organisatoren dieser komplizierten Erhebungen waren zunächst das frühere Bundesministerium für Arbeit und Sozialordnung im Fall der ASID-Erhebung (vgl. BMA 2001), sowie aktuell der VDR in Kooperation mit dem Bundesministerium für Gesundheit und Soziale Sicherung (BMGS) im Fall der AVID-Erhebung (vgl. Roth et al. 2002 und den Beitrag von Thorsten Heien in diesem Band).

Analyse und Fortschreibung durch die Wissenschaft

Die allseits anerkannte Notwendigkeit, die sozialen Sicherungssysteme an die erwarteten demographischen Zukunftsentwicklungen und gewünschten sozialpolitischen Reformzielen anzupassen, hat es erforderlich gemacht, dass Wissenschaft und Forschung sich ebenfalls eingehend mit diesen Fragen beschäftigen. Der Wunsch der Forschung, für diese Zwecke Verwaltungsdaten nutzen zu können, lag daher nahe und wurde durch eine entsprechende Regel im Zehnten Sozialgesetzbuch umgesetzt.¹ Auch die Rentenversicherung hat neben eigenen Grundsatzarbeiten und zahlreichen in Auftrag gegebenen Gutachten nunmehr den Weg beschritten, die Alterssicherungsforschung durch die Förderung von Wissenschaftlerinnen und Wissenschaftlern voranzutreiben (vgl. Ruland 2001). Zu diesem Zweck wurde 2001 das Forschungsnetzwerk Alterssicherung (FNA) gegründet.²

Politik und Wissenschaft als Datennutzer

Aus den drei genannten Gründen wird sich auch in Zukunft das Informationswesen der Rentenversicherung weiterentwickeln. Es kann davon ausgegangen werden, dass außerdem Politik und Wissenschaft einen sich stetig erweiternden Informationsbedarf formulieren werden. Betrachtet man die Alterssicherung als Forschungsgegenstand, so sind thematisch zwei Bereiche zu unterscheiden, auf die sich das Interesse von Politik und Wissenschaft konzentriert.

Zunächst sind Grundsatzfragen zum System der Alterssicherung ein wichtiges Forschungsfeld. Hier geht es neben strukturellen Fragen und der Definition von Zielen insbesondere auch um Fragen der Einschätzung, der Akzeptanz und einer optimierten Ausgestaltung der Finanzierung von Alterssicherung vor dem Hintergrund der verschiedenen Einkommensquellen im Alter (gesetzliche Rente, betriebliche Alterssicherung, private Vorsorge, staatliche und private Transfers sowie Arbeitseinkommen).³ Neben den Grundsatzfragen sind aktuelle Aspekte wie die gegenwärtigen demographischen Wandlungsprozesse und die Folgen von Reformen Gegenstand empirischer Analysen. Die langfristigen Perspektiven der Finanzierung der Alterssicherung sollten, auch vor dem Hintergrund der Europäischen Ebene, gründlich durchleuchtet werden (vgl. VDR 2003a).

Für Wissenschaft und Politik sind spezielle Datenerfordernisse zu formulieren, die teils auf die Verwaltungsdaten, teils auf zusätzliche Datenerhebungen abstellen. Zu beiden Themenfeldern ist außerdem anzumerken, dass sie nicht allein für die Sozialpolitikforschung relevant sind, sondern ebenfalls Arbeitsgebiete der politikberatenden Forschung sind. Schließlich sind die Daten über Alterssicherung und Rentenversicherung für eine Reihe weiterer wissenschaftlicher Fragestellungen ein einzigartiger Datenfundus.

¹ Vgl. § 75 SGB X: Die Gutachten werden regelmäßig in den DRV-Schriften publiziert.

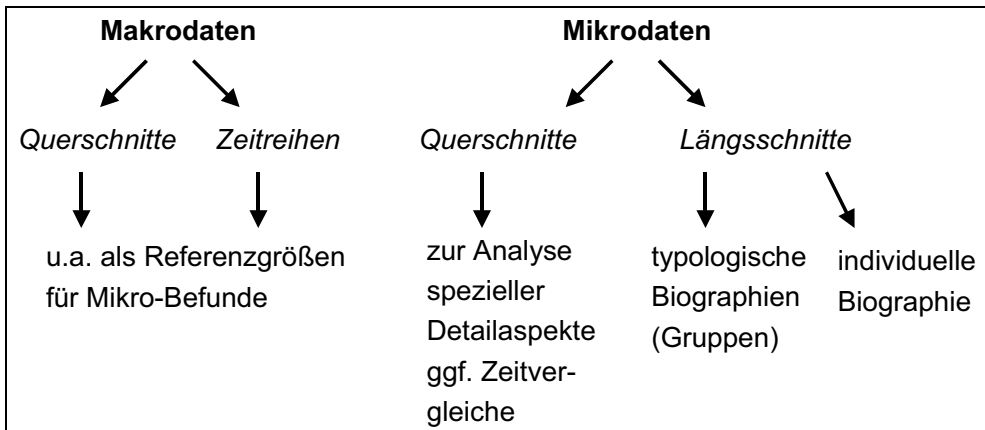
² Zum aktuellen Forschungsprogramm des Forschungsnetzwerks Alterssicherung vgl. www.vdr-fna.de.

³ Vgl. hierzu Abbildung 18 in Viebrok et al. (2004, 95).

Mikrodaten im Fokus der Empirie

Strukturiert man den Informationsbedarf von Politik und Wissenschaft unter methodischen Aspekten, so lässt sich zwischen Makrodaten und Mikrodaten⁴ unterscheiden (**vgl. Abbildung 1**). Die Makrodaten, also Strukturdaten über einzelne Perioden wie auch Zeitreihen, die Trends aufdecken können, waren bisher unerlässliche Informationsquellen und werden dies auch in Zukunft weiterhin sein. Ihre Erklärungskraft stößt jedoch an Grenzen wenn es darum geht, Wirkungszusammenhänge für einzelne Personen oder Haushalte zu erfassen, darzustellen und zu bewerten. Erforderlich sind hierzu Mikrodaten, die als Querschnittsdaten zur Analyse spezieller Detailspekte herangezogen werden. Beispielsweise sind die Mikrodaten der VDR-Statistik des Rentenzugangs geeignet, bestimmte Auswirkungen von Rechtsänderungen für verschiedene Personenkreise aufzuzeigen. Darüber hinaus sind individuelle Längsschnittsdaten von Bedeutung, wie sie z.B. im Versicherungskonto vorliegen. Hier sind die Erwerbsbiographien der Versicherten über viele Jahrzehnte präzise dokumentiert. Aus ihnen lassen sich z.B. typische biographische Muster ermitteln. Darüber hinaus lässt sich mit Biographiedaten u.a. untersuchen, inwieweit beobachtete Verhaltensweisen der Akteure durch Rechtsänderungen bedingt sind. Zukünftig werden wichtige bei der Rentenversicherung gespeicherte Verwaltungsdaten, die teilweise retrospektiv nicht in dieser Qualität erhoben werden können, für wissenschaftliche Analysen verfügbar gemacht.

Abbildung 1: Informationsbedarf unter methodischen Aspekten



Quelle: Rehfeld (2004, 66).

⁴ In diesem Band werden die Begriffe Mikrodaten, Individualdaten und Einzeldatensätze (im Unterschied zu Summendatensätzen) synonym verwendet.

Institutionelle Rahmenbedingungen des Aufbaus eines Forschungsdatenzentrums der Rentenversicherung

Über den Zugang der Wissenschaft zu Mikrodaten der Verwaltungen wurde in den vergangenen Jahren seitens der Datenhalter und Datennutzer intensiv debattiert. Das Bundesministerium für Bildung und Forschung (BMBF) hat zur Begleitung und Fokussierung dieser Diskussion die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) eingesetzt. Die KVI hat in ihrem Gutachten die Empfehlung ausgesprochen, prozessproduzierte statistische Daten öffentlicher Datenproduzenten in geeigneter Weise der Wissenschaft für Forschungszwecke zur Verfügung zu stellen (KVI 2001). Die Rentenversicherung, der finanziell bedeutendste Träger der sozialen Sicherung, ist in diese Diskussionen einbezogen worden. Sie hat die Anregungen und Überlegungen mit dem Ziel aufgegriffen, die vielfältigen Daten der GRV für Forschungszwecke zu erschließen und verfügbar zu machen. Beschlossen wurde eine zweijährige Pilotphase zum Aufbau eines Forschungsdatenzentrums der Rentenversicherung (FDZ-RV), wobei die maßgebliche finanzielle Beteiligung des Forschungsministeriums zur Voraussetzung gemacht wurde, weil die Bereitstellung der Daten für Forschungszwecke keine primäre Aufgabe der Rentenversicherung darstellt.

Der Aufbau des Forschungsdatenzentrums hat seit Januar 2004 mit finanzieller Unterstützung durch das Bundesministerium für Bildung und Forschung (BMBF) begonnen. Das FDZ-RV ist institutionell in die Infrastruktur des Verbands Deutscher Rentenversicherungsträger eingebunden. Schrittweise sollen Regelungen und Verfahren definiert und umgesetzt werden, um vorhandene Daten unter Wahrung des Datenschutzes gemäß Sozialgesetzbuch und Bundesdatenschutzgesetz und in Anlehnung an Verfahren anderer Forschungsdatenzentren für wissenschaftliche Zwecke nutzbar zu machen.

Datenquellen zur Alterssicherung

Im Kontext der bereits angesprochenen Reformdebatte hat sich sowohl in der Sozialforschung als auch in der Sozialpolitik die Sichtweise von der Betrachtung des Erwerbs und der Höhe individueller gesetzlicher Rentenansprüche auf Fragen der Gewichtung der Altersvorsorge zwischen erster, zweiter und dritter Säule sowie auf die Alterssicherung im Familienkontext verlagert. Die Rentenversicherung selbst bleibt dabei aufgrund ihres hohen Beteiligungsgrades⁵ sowie aufgrund der hohen finanzpolitischen Bedeutung⁶ regelmäßig im Fokus der Analysen.

Abbildung 2 dokumentiert die verfügbaren statistischen Daten zur Alterssicherung der ersten und zweiten Säule. Es fällt auf, dass Einzeldaten im Grundsatz lediglich im öffentlich-rechtlichen System der Alterssicherung bereitstehen. Das Tableau zeigt darüber hinaus, dass für bestimmte Bereiche - wie berufsständische Versorgung, Zusatzversorgung im öffentlichen Dienst, betriebliche Alterssicherung - die Erhebung entsprechender Daten lediglich durch zusätzliche Be-

⁵ Die gesetzliche Rentenversicherung ist das mit Abstand bedeutendste Element der Alterssicherung in Deutschland. Der Anteil der Personen mit GRV-Konto belief sich 1995 in den neuen Bundesländern bei Männern und Frauen auf 99% und in den alten Bundesländern auf 89% bei Männern und 70% bei Frauen (vgl. Kortmann & Schatz 1999, 578 sowie den Beitrag von Heien im vorliegenden Band).

⁶ Das Ausgabenvolumen entsprach im Jahr 2000 10,7% des Bruttoinlandsproduktes (VDR 2003b, 255).

fragungen von Personen und Haushalten möglich ist.⁷ Bezüglich des konkreten Datenangebots der gesetzlichen Rentenversicherung wird auf die publizierten Dokumentationen des Berichtswesens⁸ und auf den Beitrag von Hilmar Luckert verwiesen. In diesem wird auch erläutert, welche Statistiken der gesetzlichen Rentenversicherung für Stichtage oder Zeitraumerhebungen und welche Daten in Form von Zeitreihen zur Verfügung stehen. Neben der zeitlichen Abgrenzung ist aus methodischer Perspektive ein bedeutender Aspekt hervorzuheben: Einige Daten stehen im statistischen Berichtswesen in Form von Mikrodaten zur Verfügung, andere Daten sind lediglich in Form von Summendaten (das heißt in der Regel als Tabellen) verfügbar.

Abbildung 2: Verfügbare statistische Daten zur gesetzlichen und betrieblichen Alterssicherung

Sicherungssysteme	GRV	Beamtenversorgung	Landwirte	Berufsständische	Zusatzvorsorge im öffentl. Dienst	Betriebliche Alterssicherung
Datenhalter	VDR	Amtliche Statistik	Gesamtverband der Landwirtschaftlichen Alterskassen	Berufsständische Versorgungswerke	Versorgungskassen z.B. VBL, AKA	Vielzahl von Einrichtungen
Datentypen	Einzeldaten, div. Summenstatistiken	Einzeldaten gemäß FinPersG	Einzeldaten, div. Summenstatistiken	keine einheitliche Datenstruktur	keine Einzeldaten	keine einheitliche Datenstruktur
Verfügbarkeit	z.Zt. nur Tabellen neu: FDZ-RV	gemäß Bundesstatistikgesetz	z.Zt. nur Tabellen	keine einheitliche Datenstruktur		

Quelle: eigene Darstellung, in Anlehnung an Rehfeld (2004, 70).

Abbildung 2 vermittelt einen Überblick über Datenhalter und verfügbare Datentypen der verschiedenen Sicherungssysteme. Einzeldatensätze werden ausschließlich von öffentlichen Datenproduzenten bereit gestellt. Vielfältige zusätzliche Daten werden regelmäßig in Geschäftsberichten und anderen Veröffentlichungen publiziert, dann jedoch nur in Tabellenform. Neben der bereits angesprochenen Kumulationsproblematik sind auch aus diesem Grunde umfassende Erhebungen in der Bevölkerung erforderlich, um die Daten des öffentlichen Informationswesens zu vervollständigen. Soweit die Datenerhebungen mit öffentlichen Mitteln finanziert werden, sollten die Mikrodaten in Zukunft der Wissenschaft zur Verfügung stehen.

⁷ Die Erhebung von Formen der privaten Altersvorsorge in Deutschland stellt sich als besonders schwierig dar (vgl. Himmelreicher 2001, 336 ff.).

⁸ Vgl. Vorworte und methodische Anhänge in den VDR-Statistikbänden sowie Rehfeld (2001).

Ziel des Forschungsdatenzentrums der gesetzlichen Rentenversicherung

Der Aufbau des Forschungsdatenzentrums ist mit dem Zweck verbunden, Mikrodaten der gesetzlichen Rentenversicherung für Forschungszwecke zur Verfügung zu stellen. Obwohl auf vorhandenen Strukturen aufgesetzt wird, ist dies ein aufwendiges Vorhaben. Zum einen sind die bisher nur für interne Zwecke erstellten Dokumentationen derart zu verbessern, dass eine fachgerechte Nutzung der Daten durch Externe erleichtert wird. Aufgrund der komplizierten Materie erfordert die Bereitstellung der Daten zum anderen eine verlässliche Betreuung in Auswertungsfragen. Die Rentenversicherung hat die Absicht, Alterssicherung als Gegenstand empirischer Forschung zu fördern und Erkenntnisgewinne für die eigenen Aufgaben zu nutzen. Außerdem soll zur Versachlichung der sozialpolitischen Diskussion beigetragen werden. Diesen Zielen sind sowohl das Forschungsnetzwerk Alterssicherung durch die Vergabe entsprechender Forschungsprojekte und -stipendien als auch das Forschungsdatenzentrum durch die Bereitstellung entsprechender Daten verpflichtet.

Hierbei sind zur Wahrung des Datenschutzes entsprechende Vorkehrungen der Anonymisierung (vgl. IAW 2003) und vertragliche Vereinbarungen erforderlich. Zu den Fragen des Datenschutzes sei auf die Ausführungen von Claudia Heese im vorliegenden Band verwiesen.

Die Bereitstellung von prozessproduzierten statistischen Daten der Rentenversicherung soll auf verschiedenen Wegen erfolgen; sie werden in diesem Band im Beitrag von Ralf Himmelreicher dargelegt.

Start und erste Schritte

Aktuell (im Juni 2004) wurde im Aufbau des FDZ-RV mit einer Reihe wesentlicher Vorbereitungsarbeiten für die Bereitstellung von prozessproduzierten Mikrodaten begonnen. Unter anderem wird derzeit die Logistik für die Infrastruktur in der Verbandsgeschäftsstelle in Frankfurt am Main (die am 01.04.2005 in Berlin ihren Sitz nehmen wird) und in Würzburg (auch in Zukunft an diesem Ort) geplant. Am Antragsverfahren wird noch gefeilt; voraussichtlich können noch im IV. Quartal 2004 erste internetgestützte Dialoge zu Datenangeboten und Zugangswegen bereit gestellt werden. Bereits jetzt wird über das Internetportal des FDZ-RV der Kontakt zu potentiellen Nutzern hergestellt und über den Aufbau des Forschungsdatenzentrums informiert.⁹ Ein erstes Forschungsvorhaben ist als Test für die zu elaborierenden Verfahren der Datenbereitstellung und Kooperation mit der Wissenschaft erfolgreich bearbeitet worden (vgl. hierzu den Beitrag von Stefan Kempf und Martin Kukuk). In einer Umfrage unter zukünftigen Nutzern des Forschungsdatenzentrums wurden die Datenwünsche und bevorzugten Zugangswege erfragt (vgl. hierzu Himmelreicher (2004) sowie seinen Beitrag im vorliegenden Band).

Für die Weitergabe von anonymisierten Mikrodaten der Rentenversicherung (z.B. in Form von Scientific Use Files) wird ein besonderer Nutzervertrag benötigt. Insbesondere im Bereich solcher Verfahrensfragen ist eine gemeinsame Abstimmung der Vorgehensweise der öffentlichen

⁹ Vgl. www.fdz-rv.de.

Datenproduzenten im Rahmen des ‚Arbeitskreis öffentliche Forschungsdatenzentren‘ (AKöffFDZ) vorgesehen.¹⁰

Mit dem ersten Workshop zum Thema Aufbau des FDZ-RV wird die Zusammenarbeit mit den Datennutzern begonnen, weitere Workshops sind geplant. Die Publikation der auf dem Workshop gehaltenen Referate in diesem Tagungsband beabsichtigt zum einen, den Status Quo und den weiteren Aufbau des FDZ-RV zukünftigen Datennutzern mitzuteilen. Zum anderen wird angestrebt, die Nutzung statistischer Daten der Rentenversicherung für Forschungszwecke zu erleichtern.

Literatur

- Bundesministerium für Arbeit und Sozialordnung (Hg.) (2001): *Alterssicherung in Deutschland 1999 (ASID'99) – Methodenbericht*. Bonn.
- IAW (2003): *Faktische Anonymisierung wirtschaftsstatistischer Einzeldaten*. Zwischenbericht an das BMBF. In Kooperation mit den Statistischen Ämtern der Länder und des Bundes. Wiesbaden.
- Himmelreicher, R. (2001): *Soziodemographie, Erwerbsarbeit, Einkommen und Vermögen von westdeutschen Haushalten*. Berlin: Logos-Verlag.
- Himmelreicher, R. (2004): Daten der Rentenversicherung und Bedarf der Wissenschaft. Befunde einer Befragung unter potentiellen Nutzern des Forschungsdatenzentrums der gesetzlichen Rentenversicherung. *Deutsche Rentenversicherung*, 8, 472-485.
- KVI, Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (2001): *Wege zu einer besseren informationellen Infrastruktur*. Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. Nomos: Baden-Baden.
- Merz, J. & Zwick, M. (2004): Mikroanalysen und amtliche Statistik – Ein Überblick. In: Merz, J., Zwick, M. u.a.: *MIKAS - Mikroanalysen und amtliche Statistik*. Hrsg. vom Statistischen Bundesamt: Statistik und Wissenschaft, 6-12, Bd. 1. Wiesbaden.
- Kortmann, K. & Schatz, C. (1999): Altersvorsorge in Deutschland 1996 (AVID '96) – Lebensverläufe und künftige Einkommen im Alter. Verband Deutscher Rentenversicherungsträger (Hg.): *DRV-Schriften* Band 19; und Bundesministerium für Arbeit und Sozialordnung (Hg.): *BMA-Forschungsbericht*, Band 277, Frankfurt a.M./München.
- Rehfeld, U. (1987): Stand und Ausbau des statistischen Informationssystems der gesetzlichen Rentenversicherung. In: Ruland, F. (Hg.): *In Verantwortung für die Deutsche Rentenversicherung. Dr. Rudolf Kolb zum 60. Geburtstag*. VDR: Frankfurt a. M., 89-99.
- Rehfeld, U. (2001): Die Statistiken der gesetzlichen Rentenversicherung. *Deutsche Rentenversicherung*, 3-4, 169-188.
- Rehfeld, U. (2004): Datenangebot und Informationsbedarf im Bereich der Alterssicherung – Forschungsdatenzentrum der Rentenversicherung im Aufbau. *Deutsche Rentenversicherung*, 1-2, 63-75.

¹⁰ Zum Datenangebot der Forschungsdatenzentren der BA im IAB sowie des der Statistischen Ämter des Bundes und der Länder siehe die Beiträge von Annette Kohlmann und Sylvia Zühlke u.a. in diesem Band. Zu einigen Analysemöglichkeiten mit Mikrodaten insbesondere der amtlichen Statistik siehe Merz & Zwick (2004, 6 ff.).

- Rische, H. (1987): Die gesetzliche Rentenversicherung in der modernen Informationsgesellschaft - Skizze eines integrierten Informationssystems. In: Ruland, F. (Hg.): *In Verantwortung für die Deutsche Rentenversicherung. Dr. Rudolf Kolb zum 60. Geburtstag*. VDR: Frankfurt a.M. 101-107.
- Roth, M., Stegmann, M. & Bieber, U. (2002): Die Aktualisierung der Studie Altersvorsorge in Deutschland. Inhaltliche und methodische Neuerungen der AVID 2002. *Deutsche Rentenversicherung*, 11, 612-641.
- Ruland, F. (2001): Forschungsnetzwerk Alterssicherung – verstärkte Kooperation zwischen Rentenversicherung und Wissenschaft im Bereich der Alterssicherung. *Deutsche Rentenversicherung*, 3-4, 129 –136.
- Viebrok, H., Himmelreicher, R. & Schmähl, W. (2004): *Private Altersvorsorge statt gesetzlicher Rente: Wer gewinnt, wer verliert?* Herausgegeben von W. Schmähl. Beiträge zur Sozial- und Verteilungspolitik, Band 3, Lit Verlag: Münster.
- VDR (Hg.) (2003a): Rentenversicherung im internationalen Vergleich. DRV-Schriften, Band 45. Frankfurt a.M.
- VDR (Hg.) (2003b): Rentenversicherung in Zeitreihen. *DRV-Schriften*, Band 22. Frankfurt a.M.

Uwe G. Rehfeld hat an der J. W. Goethe-Universität Volkswirtschaft studiert. Schwerpunkte des Studiums waren Finanzwissenschaft, Sozialpolitik und Statistik. Nach dem Examen war er Tutor am Statistischen Seminar und zeitweilig Dozent für Volkswirtschaft und Statistik in Darmstadt, später an der FH in Würzburg für Sozialpolitik. Seit 1975 ist er beim VDR beschäftigt, zunächst als Referent, ab 1979 Referatsleiter Statistik in Würzburg. Arbeitsschwerpunkt waren dort der Auf- und Ausbau der Statistikberichterstattungen, insbesondere der Versicherungskontenstichprobe und der Reha-Renten-Verlaufsstatistik; seit 1986 Leiter der Abteilung Volkswirtschaft und Statistik. Hier waren Arbeitsschwerpunkte die Reformen, spezielle Themen (Sterblichkeit von Rentnern, Kohorten, Versicherungsfremde Leistungen, Ausländer) und die inhaltliche Fortentwicklung der Statistik durch Zusatzerhebungen (AVID). Aktuell werden im Bereich der Abteilung das Forschungsnetzwerk Alterssicherung (FNA), die Datenbank INGA (INformation zur zusätzlich Geförderten Altersvorsorge) und das Forschungsdatenzentrum der Rentenversicherung (FDZ-RV) betreut bzw. aufgebaut. Uwe G. Rehfeld vertritt den VDR und die Rentenversicherung in mehreren wissenschaftlichen und interessenpolitischen Gremien, u.a. bei der amtlichen Statistik und im Rat für Sozial- und Wirtschaftsdaten.

Datenwege und praktischer Datenzugang

Thomas Bütetisch (VDR Frankfurt)

Datenmeldungen

In § 79 SGB IV hat der Gesetzgeber eine Rahmenregelung für das Berichtswesen aller Sozialversicherungszweige getroffen. Für die Rentenversicherung werden die Inhalte der einzelnen Berichterstattungen in der „Allgemeinen Verwaltungsvorschrift für die Statistiken in der Rentenversicherung (RSVwV)“ geregelt (vgl. auch Rehfeld 2001). Beispielsweise ist in § 4 RSVwV in der noch gültigen Fassung vom 30. Januar 1992 festgelegt, dass im Bereich der durchgeführten Leistungen zur Rehabilitation die Versicherungsträger jährlich über folgende Merkmale berichten:¹

1. Alter, Geschlecht, Staatsangehörigkeit, Familienstand, Wohnort, Beruf und Versicherungsverhältnis des Rehabilitanden,
2. Diagnosen, gegliedert nach der internationalen Klassifikation der Krankheiten (ICD) in der jeweils für die Bundesrepublik Deutschland geltenden Systematik, Art, Dauer und Ort der Leistung,
3. Leistungsvermögen, Arbeit vor Antragstellung, medizinisches Entlassungsurteil und Ergebnis der beruflichen Ausbildung/Umschulung.

Die Datenmeldungen der insgesamt 26 Träger der gesetzlichen Rentenversicherung werden in der Datenstelle des VDR in Würzburg gesammelt und auf Plausibilität überprüft.

In der Regel sind die Daten den Versicherungskonten der Träger entnommen worden. Dort sind alle Informationen, die für eine spätere Leistungsgewährung (insb. Gewährung von Leistungen zur Teilhabe und Renten) an den einzelnen Versicherten oder seine Angehörigen oder Hinterbliebenen erheblich sein können, abgelegt. Zu diesen Tatbeständen gehören Informationen darüber, wann ein Versicherter in welchem Versicherungsverhältnis gestanden hat und für welche Zeiträume in welcher Höhe Beiträge entrichtet worden sind oder als entrichtet gelten. Diese Daten gelangen über das integrierte Meldeverfahren der Sozialversicherung zu den Trägern. Die Einzelheiten dieses Verfahrens sind in der Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung (DEÜV) in der Fassung zum 1. Januar 1999 geregelt.

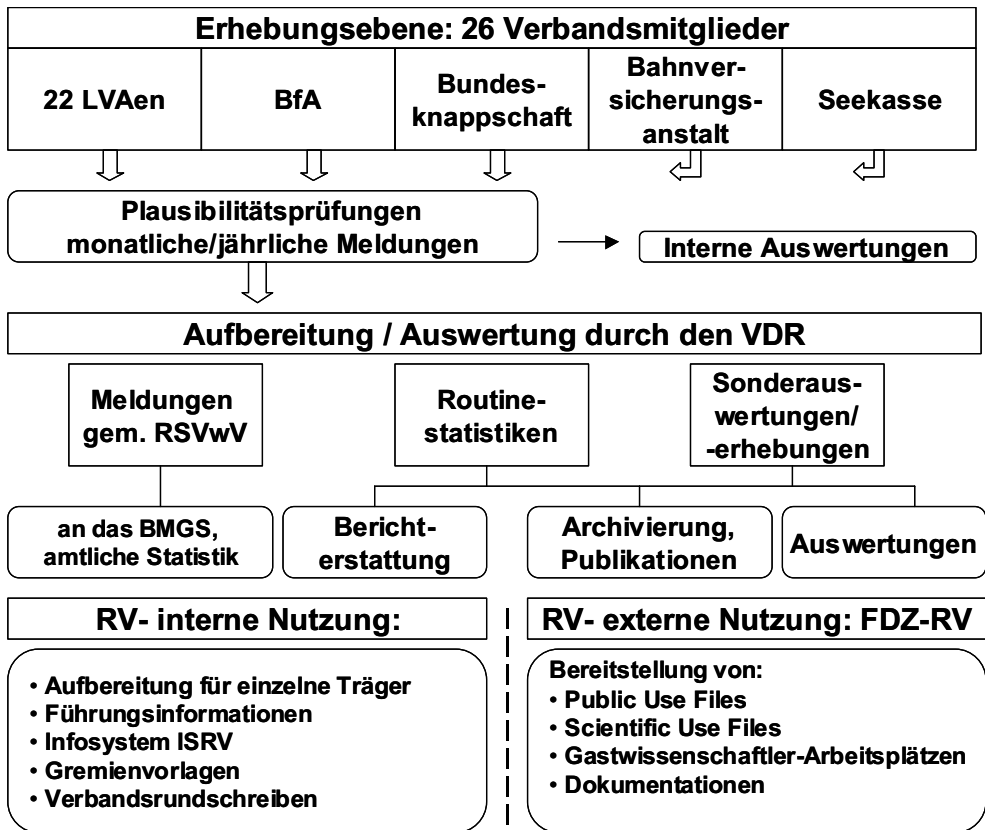
Die Arbeitgeber sollen jährlich spätestens bis Mitte April über die versicherungspflichtig Beschäftigten am 31.12. des Vorjahres eine Meldung abgeben, sofern nicht Abmeldungen (§ 8 DEÜV) oder Meldungen aus sonstigem Anlass (§ 12 DEÜV) bereits früher zu erfolgen hatten. Dabei melden die Arbeitgeber über die örtlich und sachlich zuständigen Krankenkassen. Die Arbeitsverwaltungen melden direkt an die Rentenversicherungsträger.

Diese Daten sind zum Beispiel auch in der Statistik über die aktiv Versicherten enthalten. Da im integrierten Meldeverfahren über drei Millionen Arbeitgeber und eine Vielzahl von Institutionen

¹ Vgl. z.B. VDR-Statistik Bd. 150 "Rehabilitation 2003" (2004, 330).

zusammen arbeiten, muss davon ausgegangen werden, dass Fehler unterlaufen und Meldungen verspätet eingehen. Aus diesem Grunde werden die Versicherungskonten für diese Statistik erst am 1. Oktober des Folgejahres ausgewertet.

Abbildung 1: Datenwege



Quelle: Eigene Darstellung, in Anlehnung an Rehfeld (2001, 165).

Die Versicherungsträger unterziehen vor der Übermittlung an den VDR alle Datensätze einer einheitlichen Plausibilitätsprüfung. Dabei kommen Prüfprogramme zum Einsatz, die den Trägern jeweils für die einzelnen Statistiken zur Verfügung gestellt werden. In Absprache mit den Trägern werden dann fehlerhafte Datenmeldungen gegebenenfalls erneut in bereinigter Form geliefert. Anschließend liegen die Daten dem VDR in Form von Einzeldatensätzen vor, teilweise auch in Form von Summendatensätzen.² Falls es sich um Einzeldatensätze handelt, ersetzen die Rentenversicherungsträger die Versicherungsnummer, unter der der Fall im Versicherungs-

² Vgl. hierzu den Beitrag von Hilmar Luckert in diesem Band.

konto geführt wird, vor Abgabe der Datensätze an den Verband durch eine anonyme Zählnummer.

Daten gemäß RSVwV werden vom VDR an die Aufsichtsbehörden, die amtliche Statistik und an das BMGS weiter gegeben. Die Ergebnisse von Routineauswertungen seitens des VDR fließen ferner ein in zahlreiche Berichterstattungen, die der Information innerhalb der Rentenversicherung auf verschiedenen Wegen dienen: Neben Beratungen von speziell aufbereiteten Ergebnissen in diversen Fachgremien kann jeder Rentenversicherungsträger innerhalb des elektronischen Informationssystems der Rentenversicherung (ISRV) auf verschiedene Datenbanken zugreifen und Ergebnisse von Routineauswertungen abrufen. Diese lassen sich dann bspw. mit Tabellenkalkulationsprogrammen wie MS Excel weiter verarbeiten.

Abbildung 2: Praktischer Datenzugang

Praktischer Datenzugang für ...	
... alle Interessenten	<ul style="list-style-type: none"> - durch VDR-Publikationen, Zahlen aus dem Internet (z.B. Zeitreihen im Excel-Format), Broschüren, CD-Rom - durch Sonderauswertungen (als Tabellen) nach Anfrage beim VDR in Frankfurt oder Würzburg, - zukünftig durch Bereitstellung von Public Use Files
... die wissenschaftliche Forschung nach vertraglicher Regelung	<ul style="list-style-type: none"> - durch Bereitstellung von Scientific Use Files - durch Auswertung von themenbezogenen Scientific Use Files an Gastwissenschaftler-Arbeitsplätzen - durch Fernauswertung von Daten anhand übermittelter Programmsyntax

Quelle: Eigene Abbildung.

Daten für Interessierte

Im Rahmen des Statistik-Datenservice sind die Ergebnisse der Routineberichterstattungen auch für externe Nutzer verfügbar. So sind die Statistikbände „Versicherte“, „Rehabilitation“, „Rentenzugang/-wegfall“ und „Rentenbestand“ in jährlicher Erscheinungsweise über die Presseabteilung des VDR oder über <http://www.vdr.de/statistik> zu beziehen. Auch speziell aufbereitete Ergebnisse, wie die Publikation „Rentenversicherung in Zeitreihen“, sind über die oben genannten Quellen erhältlich. Vorteil der Online-Version gegenüber der Print-Version ist hierbei, dass die Tabellen direkt in Tabellenkalkulationsprogramme wie bspw. MS Excel übertragen und weiterverarbeitet werden können.

Anfragen nach Sonderauswertungen können prinzipiell von jedem Interessierten gestellt werden. Hierbei werden in der Regel größere Datenbestände am Großrechner ausgewertet. Solche Auswertungen werden nur durchgeführt, sofern es die zeitliche Kapazität der Mitarbeiter beim VDR erlaubt. Die Ergebnisse werden bislang nicht in Form von Mikrodaten weitergegeben, sondern sie werden als Tabellen oder Summendatensätze zur Verfügung gestellt.

Daten für die wissenschaftliche Forschung

Neben den bereits oben geschilderten Datenzugangswegen für alle Interessenten bietet das FDZ-RV einen erweiterten Service für die wissenschaftliche Forschung. In ersten Schritten werden derzeit die für die einzelnen Routinestatistiken vorhandenen, teilweise komplexen Datensatzbeschreibungen überarbeitet. Ziel soll dabei sein, die ausführliche Beschreibung sämtlicher Merkmale und ihrer inhaltlichen Differenzierung sowie auch Änderungen im Zeitverlauf (bspw. bei Gesetzesänderungen) einerseits transparent zu machen, andererseits vom Fachjargon der Rentenversicherung so weit wie möglich zu befreien.

In weiteren Schritten werden zukünftig von den Mitarbeitern des FDZ-RV für spezielle Anfragen und nach vertraglicher Fixierung der Fragestellung unter Abwägung von Sozialdatenschutzgesichtspunkten Scientific Use Files erstellt. Parallel dazu wird an den Möglichkeiten der kontrollierten Datenfernverarbeitung, etwa durch Auswertung der Mikrodaten im FDZ-RV anhand übermittelter Programm-Syntax, und an der Einrichtung von Gastwissenschaftler-Arbeitsplätzen gearbeitet (vgl. Rehfeld 2004).

Fazit

In der Rentenversicherung werden sowohl Mikrodaten als auch Summendaten produziert und dem VDR übermittelt, der wiederum eine Vielzahl von Routineauswertungen durchführt. Aufgrund der hohen Fallzahlen ist die Validität der Daten und der Ergebnisse generell sehr hoch. Die Ergebnisse werden regelmäßig auf den vorgestellten Berichtswegen an die zuständigen Stellen gemeldet. Sie stehen aber auch der interessierten Öffentlichkeit zur Verfügung. Spezielle Auswertungen zu bestimmten Fragestellungen werden auf Anfrage durchgeführt. Im Rahmen des FDZ-RV besteht durch die Weitergabe von anonymisierten Mikrodaten ein erweiterter Datenservice für die wissenschaftliche Forschung.

Literatur

Rehfeld, U. (2001): Die Statistiken der gesetzlichen Rentenversicherung. *Deutsche Rentenversicherung*, 3-4, 160-188.

Rehfeld, U. (2004): Datenangebot und Informationsbedarf im Bereich der Alterssicherung – Forschungsdatenzentrum der Rentenversicherung im Aufbau. *Deutsche Rentenversicherung*, 1, 63-75.

Verband Deutscher Rentenversicherungsträger (Hg.) (2004): *VDR Statistik Rehabilitation des Jahres 2003*, Band 150. Frankfurt am Main.

Thomas Bütetfisch ist Diplom-Statistiker (Universität Dortmund). Nach Abschluss des Studiums Beschäftigung beim VDR in Frankfurt als Referent in der Abteilung „Volkswirtschaft und Statistik“. Betreuung der Versicherten- und Rehabilitationsstatistiken, insbesondere Fortentwicklung und Konzeption von Sonderauswertungen der Reha-Statistik-Datenbasis (RSD).

Statistikdaten der gesetzlichen Rentenversicherung – ein grober Überblick

Hilmar Luckert (VDR Würzburg)

1. Einleitung

Die folgenden Ausführungen konzentrieren sich weitgehend auf regelmäßig durchgeführte Routineerhebungen des statistischen Berichtswesens beim VDR. Dabei wird lediglich auf jene Statistiken näher eingegangen, die auf Individualdaten basieren und die somit im Rahmen des Forschungsdatenzentrums der Rentenversicherung von Interesse sein könnten. Solche Statistiken, zu denen auch beim VDR die Daten lediglich in aggregierter Form vorliegen, werden zur inhaltlichen Abrundung erwähnt.¹ Dieser grobe Überblick über die Daten der Rentenversicherung beschreibt die momentane Situation und lässt offen, welche Daten mit welchen Ausprägungen der Wissenschaft durch das Forschungsdatenzentrum der Rentenversicherung künftig zugänglich gemacht werden können.

Vorweg zu schicken ist, dass es sich beim statistischen Berichtswesen der Rentenversicherung nicht um eine Statistik aus einem Guss handelt, sondern vielmehr um eine Ansammlung historisch gewachsener Statistiken. Diese werden aus einem bestehenden, jedoch Veränderungsprozessen unterliegenden Verwaltungsverfahren gespeist. Infolge dessen, aber auch wegen der verschiedenartigen Anforderungen, liegen den Statistiken durchaus unterschiedliche Datenkonzepte zugrunde. Während vor allem die Leistungsstatistiken überwiegend auf der Basis von Mikrodaten durchgeführt werden, gibt es im Bereich der Verwaltungsstatistiken, bspw. über die Antragserledigung, auch beim VDR lediglich aggregierte Informationen in Form von Summendaten. Viele Statistiken haben eine feste, für alle Fälle identischen Datenstruktur, dabei können allerdings je nach Ausgestaltung des Falles durchaus unterschiedliche Merkmalsbereiche von Interesse sein. Bei Statistiken mit Verlaufscharakter, bei denen je nach Anzahl und Dauer der Episoden mehr oder weniger viele chronologisch einordenbare Daten abzubilden sind, werden auch variable Datenstrukturen benutzt. In der Regel werden Statistiken in der Rentenversicherung als Vollerhebung durchgeführt und beinhalten vorwiegend prozessproduzierte Daten. Nur an wenigen Stellen wird mit Stichproben gearbeitet, weil sich eine Vollerhebung aus arbeitsökonomischen Aspekten verbietet. Nicht übersehen werden darf, dass manche Erhebungen stichtagsbezogen sind, d.h. sie beschreiben die Verhältnisse zu einem bestimmten Zeitpunkt, während andere Statistiken Zeitraumerhebungen darstellen, also über Sachverhalte und Ereignisse berichten, die in einem fest definierten Zeitraum - meist ein Kalendermonat oder ein Kalenderjahr - liegen. Schließlich gibt es auch Statistiken, die sowohl zeitraumbezogene als auch stichtagsbezogene Informationen vereinen, wobei je nach Fragestellung die besser geeigneten Variablen zu verwenden sind.

Generell gilt jedoch, dass die Datensätze verschiedener statistischer Erhebungen in der Rentenversicherung schon so weit anonymisiert sind, dass Individualdaten auf der Personenebene nicht verbunden werden können. D.h., wenn über eine Person in einer bestimmten Statistik in verschiedenen Zeiträumen oder zu verschiedenen Stichtagen berichtet wird, lassen sich diese

¹ Für einen umfassenden Überblick des Informationswesens des VDR siehe Rehfeld (2001, 160 ff.) bzw. im Downloadbereich in www.fdz-rv.de.

Informationen nicht miteinander verknüpfen. Solche Informationen zu einer Person, die in verschiedenen Statistiken enthalten sind, selbst wenn sie den gleichen Stichtag oder Berichtszeitraum betreffen, lassen sich ebenfalls nicht verknüpfen.

2. Welche Statistiken gibt es?

Wie eingangs erwähnt, beschränkt sich dieser Überblick auf die Routinestatistiken und lässt bereichsspezifische Erhebungen, Umfragen etc. außen vor. Generell können die Statistiken nach ihrem Inhalt in folgende sachlichen Bereiche untergliedert werden:

- Versichertenstatistiken,
- Rehabilitationsstatistiken,
- Rentenstatistiken,
- Finanzstatistiken und
- sonstige Statistiken.

Hinzu kommen noch die Sondererhebungen, die unter Umständen mehrere dieser Bereiche gleichzeitig betreffen.

2.1 Versichertenstatistiken

Insgesamt werden beim VDR sieben verschiedene Versichertenstatistiken geführt: drei Summendatensätze² und vier Datensätze, die auf Individualdaten basieren. Auf letztere wird hier näher eingegangen.

Die Stammsatzbestandsstatistik (**vgl. Abbildung 1**) wird zweimal jährlich, zum 1. Januar und zum 1. Juli eines jeden Jahres, erhoben. Gegenstand dieser Erhebung sind die vergebenen Versicherungsnummern, welche den bei den Rentenversicherungsträgern geführten Versicherungskonten entsprechen. Für die Zählung in dieser Statistik ist es völlig unerheblich, ob und in welchem Umfang Versicherungszeiten vorliegen. Neben dem Überblick über die Gesamtheit der von der Rentenversicherung zu verwaltenden Versicherungskonten ist diese Statistik insbesondere im Hinblick auf die Kontoführungswechsel und auf die Neuvergabe von Versicherungsnummern von Interesse. Da sie aus dem Stammsatzbestand der Datenstelle der Rentenversicherung gewonnen wird, beschränken sich ihre Inhalte auf eine kleine Auswahl der wenigen dort verfügbaren Merkmale.

² Die Statistik der latent Versicherten, die Statistik über unmittelbare Beitragserstattungen und die neue Beitragsrückstandstatistik werden an den VDR in summarischer Form gemeldet.

Abbildung 1: Stammsatzbestandsstatistik

Termine	Jährlich zum 1.1. und 1.7.
Datenform	Einzeldatensätze
Gegenstand	Vergebene Versicherungsnummern = Anzahl der Versicherungskonten
Quelle	Stammsatzbestand der Datenstelle der Rentenversicherung
Wichtige Merkmale	Geburtsjahr und -monat, Geschlecht, Staatsangehörigkeit, Wohnort (Kreis), Vergabeanstalt, aktueller Kontoführer (Versicherungsträger), Vergabedatum, Kontoführungswechsel, Stilllegungs-, Renten- und Lebensmerkmal
Wozu?	Struktur der betreuten Versicherungskonten, Kontoführungswechsel, Neuvergaben von Versicherungsnummern
Probleme	Keine Informationen über Umfang und Zeitpunkt der zurückgelegten Versicherungszeiten bzw. ob überhaupt versicherungsrelevante Zeiten vorhanden „Karteileichen“

Quelle: Eigene Zusammenstellung.

Wichtige Merkmale sind: das Geburtsjahr, das Geschlecht und die Staatsangehörigkeit des Versicherten, welcher Versicherungsträger aktueller Kontoführer ist, das Vergabedatum und welcher Versicherungsträger die Versicherungsnummer vergeben hat, die Tatsache, ob im vergangenen Jahr Kontoführungswechsel stattgefunden haben, sowie Informationen darüber, ob der Datensatz technisch stillgelegt ist, ob ein Rentenbezug vorliegt oder ob der betreffende Versicherte verstorben ist. Bei der Stammsatzbestandsstatistik handelt es sich primär um eine Verwaltungsstatistik mit Mengengerüsten zu Versicherten nach deren Alter, Geschlecht und Staatsangehörigkeit sowie zur Durchführung von Kontoführungswechseln und zu den im Berichtsjahr neu vergebenen Versicherungsnummern. Diese Statistik beinhaltet keine Angaben darüber, ob und wann Beiträge zur Rentenversicherung entrichtet wurden und welchen zeitlichen Umfang oder welche Höhe diese ggf. aufweisen. Bei latent Versicherten ohne Leistungsbezug besteht ein so genanntes „Karteileichenproblem“.

Wesentlich informativer ist die Statistik der aktiv Versicherten (**vgl. Abbildung 2**), die jährlich zum 1. Oktober des Folgejahres durchgeführt wird und zu der erfahrungsgemäß bis Ende des Folgejahres erste Ergebnisse vorliegen.

Der betrachtete Personenkreis ist deutlich enger gefasst als bei der Stammsatzbestandsstatistik: Im Wesentlichen beinhaltet diese Statistik Versicherte, die aktuell Rentenanwartschaften aufbauen, sei es durch Beitragszahlung oder dadurch, dass sie Anrechnungszeiten erwerben. Die dieser Statistik zugrunde liegenden Daten werden weitgehend bestimmt vom Meldefluss im integrierten Verfahren zur Datenmeldung an die Sozialversicherung (DEÜV). Daten anderer Versichertengruppen, wie zum Beispiel für freiwillig Versicherte oder Handwerker, finden auf anderen Wegen unmittelbar Eingang in die Versicherungskonten. Es gilt jedoch trotz des Time-Lags von zehn Monaten zu beachten, dass die Meldungen noch nicht vollständig sind: Erfahrungsgemäß fehlen bei den Beschäftigten nach zehn Monaten noch zwei bis drei Prozent der Daten zum Berichtsjahr. Bei anderen Zeiten, wie z. B. Wehr-/Zivildienst, ist der Erfassungsgrad nach zehn Monaten sogar noch deutlich niedriger und bei Kindererziehungszeiten – mangels

Abbildung 2: Statistik der aktiv Versicherten

Termine	Jährlich zum 1.10. des Folgejahres
Datenform	Einzeldatensätze
Gegenstand	Versicherte, die durch Beiträge oder Anrechnungszeiten aktuell Rentenanwartschaften aufbauen
Quelle	DEÜV-Meldungen ⇒ Versicherungskonten
Wichtige Merkmale	Geburtsjahr, Geschlecht, Staatsangehörigkeit, Tätigkeitsschlüssel, Versicherungsbeginn (Jahr), Wohnort (Kreis), Versicherungsart am Stichtag, Arten der Versicherung (Dauern/ Beiträge/ versicherte Entgelte) im Berichtsjahr und im Vorjahr, ggf. Rentenleistungsart und Rentenbeginn
Wozu?	Aktiv Versicherte nach Art der Versicherung, Umfang und Höhe der Beiträge bzw. versicherten Entgelte, Stichtags- und Zeitraumgrößen (2-Jahres-Fenster)
Probleme	Störungen im Meldeverfahren, nie zu 100% vollständig keine Branche

Quelle: Eigene Zusammenstellung, vgl. hierzu auch VDR Statistik Versicherte: Vorwort zu Band 147.

zeitnaher Erfassung im Versicherungskonto – ist er so gering, dass sinnvolle Auswertungen nicht möglich sind.

Die Statistik der aktiv Versicherten bietet ein breites Spektrum von Merkmalen: Es reicht vom Geburtsjahr, dem Geschlecht und der Staatsangehörigkeit der Versicherten über den Tätigkeitsschlüssel³, den Versicherungsbeginn und den Wohnort (Kreis) bis hin zu einer Differenzierung nach den verschiedenen Versicherungsarten, zu denen jeweils die Dauer sowie die Beiträge bzw. Höhe der versicherten Entgelte dargestellt sind. Abgebildet wird jeweils ein Zweijahresfenster, das aus dem Berichtsjahr und aus dem Vorjahr besteht. Da in Teilbereichen erfahrungsgemäß immer wieder Störungen beim Meldeverfahren auftreten, kann die Interpretation der Ergebnisse problematisch sein. Innerhalb der Statistik der aktiv Versicherten – wie auch in allen anderen Statistiken der Rentenversicherung – ist eine Differenzierung nach Branchen nicht möglich.

Die Versicherungskontenstichprobe (**vgl. Abbildung 3**) beschäftigt sich mit den Rentenanwartschaften von Versicherten. Im Unterschied zur Statistik der aktiv Versicherten werden in dieser Statistik sämtliche in der Chronologie gespeicherten Versicherungstatbestände dargestellt. Vorhandene Merkmale sind das Geburtsjahr, das Geschlecht und die Staatsangehörigkeit des Versicherten; ein regionales Zuordnungskriterium ist nicht vorhanden. Die übrigen Merkmale beziehen sich hauptsächlich auf die Höhe und die Zusammensetzung der Rentenanwartschaften, wobei verschiedene Arten von versicherungsrechtlich relevanten Zeiten und ihr „Wert“ in Form von Entgeltpunkten bei der Rentenberechnung dargestellt sind.

³ Der Tätigkeitsschlüssel beruht auf einer Meldung im DEÜV-Verfahren und indiziert die aktuellste Tätigkeit der Versicherten: Das Merkmal enthält derzeit in den Stellen 1 bis 5 die Angaben zur Tätigkeit (vgl. Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Versicherungsnachweisen der Bundesanstalt für Arbeit): Ausgeübte Tätigkeit (Stellen 1 - 3), Stellung im Beruf (Stelle 4), Ausbildung (Stelle 5). Ist kein Tätigkeitsschlüssel zum Jahr des Leistungsfalles oder davor gespeichert, ist das Merkmal mit Nullen zu belegen. Insbesondere gilt dies für Zeiträume vor dem Jahr 2000.

Abbildung 3: Versicherungskontenstichprobe

Termine	Jährlich zum 1.10. des Folgejahres
Datenform	Einzeldatensätze
Gegenstand	Versicherte im Alter von 15 - 67 Jahren mit rentenrechtlichen Zeiten im Versicherungskonto oder einem Bonus aus Versorgungsausgleich
Quelle	Versicherungskonten
Wichtige Merkmale	<ol style="list-style-type: none"> 1. Geburtsjahr, Geschlecht, Staatsangehörigkeit, Tätigkeitsschlüssel, ggf. Rentenleistungsart und Rentenbeginn, Merkmale zur Anwendung des Fremdretenrechts, Werte zur Gesamtleistungsbewertung, Merkmale zum Umfang der vorhandenen Zeiten sowie zur Höhe und Zusammensetzung der momentanen Rentenanwartschaft 2. chronologischer Versicherungsverlauf <ul style="list-style-type: none"> • Art der versicherten Zeit • Beitragsart-Besonderheiten • Zeitpunkt und Zeitraumdauer (Monate) • Beitragshöhe / versichertes Entgelt • „Wert“ in Entgeltpunkten
Wozu?	Erworbene Rentenanwartschaften nach Höhe und Zustandekommen, Umfang der angerechneten Zeiten, Auswirkungen von rechtlichen Regelungen, Untersuchungen des chronologischen Versicherungsverlaufs
Probleme	<ul style="list-style-type: none"> • unvollständige Kontoklärung ⇒ Unterschätzung der Zeiten und Anwartschaften keine Informationen zu ausländischen Versicherungszeiten • keine Regionalangabe

Quelle: Eigene Zusammenstellung.

Dabei – und das ist das Wichtigste an dieser Statistik – ist in den Daten der Versicherungskontenstichprobe der chronologische Versicherungsverlauf dokumentiert, in dem Art, Zeitpunkt und Zeitdauer, Beitragshöhe und versichertes Entgelt sowie schließlich der Wert dieser Zeit in Form von Entgeltpunkten ausgewiesen sind. Mit dieser Statistik lassen sich daher die erworbenen Anwartschaften nach Höhe und Zustandekommen analysieren und der Umfang der angerechneten Zeiten sowie die Auswirkungen von rechtlichen Regelungen damit in Verbindung bringen. Losgelöst von rentenrechtlichen Aspekten kann der chronologische Verlauf der (versicherten) Entgelte aber ebenso von Interesse sein wie das Auftreten und die Abfolge bestimmter Arten der Versicherung, z. B. versicherungspflichtige oder geringfügige Beschäftigung, Lohnersatzleistungen, versicherungspflichtige Selbständigkeit usw. Problematisch an dieser Statistik ist, dass zuverlässige Aussagen oftmals nur zu gewinnen sind, wenn die Versicherungskonten vollständig geklärt sind. Dazu bedarf es aber der Mitwirkung der Versicherten, die in der Praxis nicht immer gegeben ist. Hinsichtlich der faktischen Anonymisierung stellen Längsschnittinformationen bekanntlich besondere Anforderungen.

Die Versorgungsausgleichsstatistik (**vgl. Abbildung 4**), die jährlich zum 1. April des Folgejahres durchgeführt wird, befasst sich mit Versicherten, für die ein Versorgungsausgleich im Versicherungskonto gespeichert ist und beinhaltet dabei jeweils sämtliche Bestandsfälle. Die Untersuchung des Anwartschaftsausgleichs in Versorgungsfällen, insbesondere die übertragenen und erworbenen Anwartschaften nach Ausgleichsformen, dürften auch die hauptsächliche Nutzungsmöglichkeit dieser Statistik bedeuten.

Abbildung 4: Versorgungsausgleichsstatistik

Termine	Jährlich zum 1.04. des Folgejahres
Datenform	Einzeldatensätze
Gegenstand	Versicherte mit einem Versorgungsausgleich (alle Bestandsfälle)
Quelle	Versorgungsausgleichsverfahren \Rightarrow Versicherungskonten
Wichtige Merkmale	Geburtsjahr / Geschlecht der Versicherten und des geschiedenen Ehepartners, Ehedauer, Art und Höhe des Versorgungsausgleichs, Beitragszahlungen zur Wiederauffüllung
Wozu?	Übertragene / erworbene Anwartschaften nach Ausgleichsformen
Probleme	Keine Informationen zu Anwartschaften insgesamt und zum Versicherungsverlauf

Quelle: Eigene Zusammenstellung.

Wichtige Merkmale, die hierbei zur Verfügung stehen, sind das Geburtsjahr und das Geschlecht – jeweils vom Versicherten und Ehepartner –, die Ehedauer, die Art und die Höhe des Versorgungsausgleichs sowie ggf. Beitragszahlungen zur Wiederauffüllung der Anwartschaften. Da aber datenmäßig kein Zusammenhang zwischen den geschiedenen Partnern besteht und jegliche Informationen über den Versicherungsverlauf und die insgesamt vorhandenen Rentenanwartschaften fehlen, sind die Nutzungsmöglichkeiten stark eingeschränkt.

2.2 Rehabilitationsstatistiken

Neben der summarisch abgewickelten Reha-Antrags- und Erledigungsstatistik sowie der Reha-Antragslaufzeiten-Statistik ist hier insbesondere die so genannte Reha-Statistik-Datenbasis (RSD) zu nennen. Diese wird als Jahrerhebung zum 15. April des Folgejahres und als Verlaufserhebung zum 15. November des Folgejahres gemeldet. Etwa 3 Monate später stehen üblicherweise Ergebnisse zur Verfügung.

Bei der Jahrerhebung der Reha-Statistik-Datenbasis (**vgl. Abbildung 5**) liegen dem VDR Individualdatensätze mit einem eingeschränkten Merkmalspektrum von sämtlichen Versicherungsträgern vor. Diese beschreiben die Struktur der abgeschlossenen Leistungen zur Teilhabe am Arbeitsleben (medizinische und berufliche Rehabilitation) im Berichtsjahr.

Im Rahmen des so genannten Reha-Kurzberichtes, zu dem von der BfA keine Individualdaten vorliegen, werden auch Anträge und Bewilligungen einbezogen. Wichtige Merkmale der Jahrerhebung sind das Geburtsjahr, das Geschlecht, der Wohn- und der Maßnahmeort (jeweils nur

Abbildung 5: Reha-Statistik-Datenbasis (RSD) Jahreserhebung

Termine	Jährlich zum 15.04. des Folgejahres
Datenform	Einzeldatensätze, von BfA zum Teil nur Summensätze
Gegenstand	1. Reha-Kurzbericht mit groben Strukturangaben zu Anträgen, Bewilligungen (und durchgeführten Leistungen im Berichtsjahr) 2. Abgeschlossene Leistungen zur Teilhabe (medizinische und berufliche Rehabilitation) im Berichtsjahr
Quelle	Reha-Antragsverfahren / Reha-Entlassungsberichte
Wichtige Merkmale	Geburtsjahr, Geschlecht, Familienstand, Wohn- und Maßnahmeort (Kreis), Antragsart, Maßnahmeart, Verfahrensart, Art der Durchführung, Bewilligte Tage und Dauer, Diagnosen und Ursache der Erkrankung, Beruf und Stellung im Beruf, Arbeit vor Antragstellung, Entlassungsform und Leistungsfähigkeit
Wozu?	Ursachen und Struktur der Leistungen zur medizinischen Rehabilitation und sonstigen Leistungen zur Teilhabe sowie der betroffenen Versicherten
Probleme	Keine einheitliche Datenstruktur (zwei verschiedene Einzeldatensatzformate und ein Summendatensatz)

Quelle: Eigene Zusammenstellung, vgl. hierzu auch VDR Statistik Rehabilitation: Vorwort zu Band 146.

der Kreis), die Antrags-, Maßnahme- und die Verfahrensart, die Entlassungsdiagnosen, der zuletzt ausgeübte Beruf sowie Angaben über Maßnahmebeginn und Maßnahmeende. Mit der Jahreserhebung der Reha-Statistik-Datenbasis lässt sich daher die Struktur der im jeweiligen Berichtsjahr durchgeführten Reha-Leistungen beschreiben.

Wesentlich mächtiger, aber auch dementsprechend schwieriger zu interpretieren, ist die Verlaufserhebung der Reha-Statistik-Datenbasis (**vgl. Abbildung 6**). Allerdings liegen dazu nicht von allen Versicherungsträgern Einzeldatensätze vor. Betrachtungsgegenstand der Verlaufserhebung sind die gestellten und die erledigten Anträge auf Leistungen zur Rehabilitation sowie vor allem die abgeschlossenen Reha-Leistungen innerhalb eines 8-Jahreszeitraumes.

Damit lassen sich sowohl die zeitliche Entwicklung auf diesem Sektor gut beobachten als auch die mehrfache Inanspruchnahme von Reha-Leistungen durch Versicherte analysieren bzw. die aus mehreren Teilleistungen bestehenden Reha-Verfahren auswerten. Darüber hinaus sind weitere Verlaufsdaten enthalten, nämlich Informationen darüber, ob und mit welcher Leistungsart ein Rentenantrag gestellt bzw. bewilligt wurde, ob und in welcher Höhe eine Rente gezahlt wird usw. Ergänzt werden die Daten mit Informationen über den Versicherten für einen elf Jahre betragenden Zeitraum. Schließlich gibt es für jeden Rehabilitanden Merkmale zu Geburtsjahr, Geschlecht, Staatsangehörigkeit und Wohnort (Kreis). Die Verlaufsdaten helfen insbesondere die Inanspruchnahme und den „Erfolg“ von Leistungen zur Rehabilitation und das Rehabilitationsgeschehen vor einer Rentengewährung zu analysieren. Die Fülle der Information bringt es jedoch mit sich, dass die Daten komplex strukturiert sind und – da sie einen 8- bzw. 11-Jahreszeitraum beschreiben – häufig von Änderungen in der Verschlüsselungsweise der zugrunde liegenden Verwaltungsdaten betroffen sind. Soweit dem VDR Individualdaten mit Verlaufscharakter vorliegen ist zu beachten, dass diese einer strengen Zweckbindung unterlie-

Abbildung 6: Reha-Statistik-Datenbasis (RSD) Verlaufserhebung

Termine	Jährlich zum 15.11. des Folgejahres
Datenform	Einzeldatensätze, von BfA zum Teil nur Summensätze
Gegenstand	<ol style="list-style-type: none"> 1. Gestellte und erledigte Reha-Anträge sowie abgeschlossene Leistungen zur Teilhabe (medizinische und berufliche Rehabilitation) im 8-Jahreszeitraum (Berichtsjahr und 7 Jahre davor) 2. Reha-Verlauf einschl. Beiträge und Rente im 8-Jahreszeitraum (Berichtsjahr und 7 Jahre davor)
Quelle	Versicherungskonto + Reha-/Rentenverfahren
Wichtige Merkmale	<ol style="list-style-type: none"> 1. Demographische Merkmale: Geburtsjahr, Geschlecht, Familienstand, Wohnort (Kreis) 2. Reha-Daten im 8-Jahresfenster: Antragsart, Maßnahmeart, Verfahrensart, Art der Durchführung, Maßnahmeort (Kreis), Antrags- und Bewilligungsdatum, ggf. Ablehnungsgründe, bewilligte Tage und Dauer, Diagnosen und Ursache der Erkrankung, Beruf und Stellung im Beruf, Arbeit vor Antragstellung, Entlassungsform, Leistungsfähigkeit, positives und negatives Leistungsbild 3. Rentendaten im 8-Jahresfenster: Antrags und Bescheidsdatum, Rentenleistungsart, Beginn- und Ende der Leistung, Rentenbetrag 4. Versicherungsdaten im 11-Jahresfenster: Beitragsart, Dauer und Höhe des/der versicherten Entgelts / Lohnersatzleistung
Wozu?	Entwicklung der Rehabilitation im 8-Jahreszeitraum, Inanspruchnahme und „Erfolg“ von Rehabilitation
Probleme	Keine einheitliche Datenstruktur (zwei verschiedene Einzeldatensatzformate und ein Summendatensatz), komplexe Datenstruktur

Quelle: Eigene Zusammenstellung.

gen und die Frage der Weitergabemöglichkeit im Rahmen des Forschungsdatenzentrums daher noch offen ist. Der Verlaufsscharakter des Datenmaterials stellt ein schwieriges Problem für eine faktische Anonymisierung dar.

2.3 Rentenstatistiken

Die als interne Arbeitsstatistik anzusehende Rentenantrags- und Erledigungsstatistik wird lediglich in summarischer Form durchgeführt. Gleiches gilt für die wenige Merkmale umfassende Rentennachbehandlungsstatistik. Drei weitere Statistiken im Rentenbereich dürften jedoch, wie die Umfrage unter potenziellen Datennutzern des FDZ-RV⁴ gezeigt hat, von besonderem Interesse sein, da sie die Hauptaufgabe der gesetzlichen Rentenversicherung beschreiben.

Die Rentenzugangs-, Rentenwegfalls- und Rentenänderungsstatistik (**vgl. Abbildung 7**) wird zum 10. Februar des auf das Berichtsjahr folgenden Jahres erstellt. Ergebnisse liegen in der Re-

⁴ Vgl. hierzu den Beitrag von Ralf Himmelreicher im vorliegenden Band.

Abbildung 7: Rentenzugangs-, Rentenwegfalls- und Rentenänderungsstatistik

Termine	Jährlich zum 10.02. des Folgejahres
Datenform	Einzeldatensätze
Gegenstand	1. Neu zugewandene Renten 2. Weggefallene Renten 3. Änderungen von Rentenleistungsart oder Teilrentenanteil
Quelle	Rentantrags-/ Rentenzahlverfahren
Wichtige Merkmale	Geburtsjahr, Geschlecht, Familienstand, Wohnort (Kreis), Diagnosen, Beruf, Rentenleistungsart, Rentenbeginn, Rentenbetrag und Zusammensetzung, Auffüllbetrag/Rentenzuschlag, Kranken-/ Pflegeversicherung (Art und Beitrag), Einkommensanrechnung, Zusammentreffen mit Einkommen, Relevanz bestimmter rentenrechtlicher Vorschriften, Anwendung von Fremdrechten und Vertragsrentenrecht, Relevanz der Arbeitsmarktlage, Versicherungsstatus und Jahresentgelt vor Leistungsfall, Werte aus der Gesamtleistungsbewertung, Werte aus der Rentenberechnung (Zeitdauer und Entgeltpunkte) Bezugsdauer/Wegfallsgrund beim Rentenwegfall, alte/neue Leistungsart bzw. Teilrentenanteil bei Rentenänderung
Wozu?	Struktur (Anzahlen, Alter, Beträge, Zeiten, Berechnungswerte, usw.) von „neuen“ Renten, Rentenwegfällen, Rentenänderungen
Probleme	Keine Rentenkumulation, kein Haushaltsbezug, keine Informationen zu anderen Alterssicherungseinkünften

Quelle: Eigene Zusammenstellung, vgl. hierzu VDR Statistik Rentenzugang: Vorwort zu Band 149.

gel etwa 2 Monate später vor. Grundsätzlich gliedert sich diese Statistik in drei Teilbereiche, nämlich die

- Rentenzugangsstatistik, in der alle im Laufe eines Berichtsjahres neu bewilligten (zugegangenen) Renten erfasst sind,
- Rentenwegfallsstatistik, in der alle im Laufe eines Berichtsjahres endgültig weggefallenen Renten enthalten sind, sowie
- Rentenänderungsstatistik, in der alle Fälle dokumentiert werden, bei denen sich im Berichtsjahr die Rentenleistungsart oder der Teilrentenanteil geändert hat oder die von einer Teilrente in eine Vollrente oder umgekehrt umgestellt wurden.

Die Daten entstammen dem Rentenbewilligungs- bzw. den Rentenzahlverfahren. Neben dem Geburtsjahr, dem Geschlecht, dem Wohnort (Kreis), der Rentenleistungsart, den Diagnosen (bei Renten wegen verminderter Erwerbsfähigkeit), dem Beruf sowie dem Rentenbetrag und dessen Zusammensetzung enthält diese Statistik eine Reihe von Merkmalen, mit denen die Auswirkungen einzelner rentenrechtlicher Vorschriften bzw. der Umfang bestimmter rentenrechtlicher Zeiten statistisch dokumentiert werden. Beim Rentenwegfall kommen etwa noch die Rentenbezugsdauer und der Wegfallsgrund hinzu. Im Falle von Rentenänderungen ist insbesondere ausgewiesen von welcher Leistungsart in welche neue Leistungsart die Umstellung erfolgt ist, bzw. welcher Teilrentenanteil vor bzw. nach der Änderung vorgelegen hat. Bei der Interpretation dieser Statistik kann problematisch sein, dass weder Informationen über eine Kumulation meh-

rerer Renten – insbesondere Witwenrente mit einer eigenen Altersrente – noch Angaben über andere Alterseinkünfte vorliegen. Auch lassen sich beispielsweise Ehepartner im Sinne eines Haushaltsbezuges nicht zusammenführen.

Die Rentenbestandsstatistik aus Konten der Rentenversicherungsträger (**vgl. Abbildung 8**) ist im Prinzip ein Spiegelbild der Rentenzugangstatistik, sie bezieht sich jedoch auf sämtliche laufenden Rentenzahlungen am 31. Dezember eines Berichtsjahres und wird ebenfalls zum 10. Februar des Folgejahres erstellt. Die Merkmale lehnen sich grundsätzlich an die Rentenzugangstatistik an – rein formal liegt sogar die gleiche Datenstruktur vor -, allerdings sind bestimmte Informationen, wie etwa die Diagnose oder der Beruf, bei Rentenbestandsfällen nicht mehr verfügbar.

Abbildung 8: Rentenbestandsstatistik aus den Konten der Rentenversicherungsträger

Termine	Jährlich zum 10.02. des Folgejahres
Datenform	Einzeldatensätze
Gegenstand	Laufende Rentenzahlungen am 31.12. des Berichtsjahres
Quelle	Rentenbestand / Versicherungskonten
Wichtige Merkmale	Geburtsjahr, Geschlecht, Familienstand, Wohnort (Kreis), Rentenleistungsart, Rentenbeginn, Rentenbetrag und Zusammensetzung, Auffüllbetrag/Rentenzuschlag, Kranken-/ Pflegeversicherung (Art und Beitrag), Einkommensanrechnung, Zusammentreffen mit Einkommen, Relevanz bestimmter rentenrechtlicher Vorschriften, Anwendung von Fremdrentenrecht und Vertragsrentenrecht, Relevanz der Arbeitsmarktlage, Werte aus der Gesamtleistungsbewertung, Werte aus der Rentenberechnung (Zeitdauern und Entgeltpunkte) bei Renten nach SGB VI
Wozu?	Struktur (Anzahlen, Alter, Beträge, Zeiten, Berechnungswerte, usw.) von laufenden Rentenzahlungen
Probleme	Keine Rentenkumulation, kein Haushaltsbezug, keine Informationen zu anderen Alterssicherungseinkünften, unterschiedliche Berechnungsarten (umgewertete RVO/AVG/RKG-Renten, umgewertete Renten aus dem Beitrittsgebiet)

Quelle: Eigene Zusammenstellung, vgl. hierzu VDR Statistik Rentenbestand: Vorwort zu Band 148 und Anhang 2.

Mit der Rentenbestandsstatistik aus den Konten der Rentenversicherungsträger lässt sich die Struktur der laufenden Rentenzahlungen beschreiben, also vor allem Mengengerüste, Rentenzahlbeträge einschließlich Teilkomponenten, Auftreten und Umfang verschiedener rentenrechtlicher Zeiten und die wertmäßigen Auswirkungen bestimmter Anrechnungsvorschriften. Als Mangel dieser Rentenbestandsstatistik ist festzuhalten, dass nur einzelne Renten ausgewiesen sind, nicht aber das Zusammentreffen mehrerer Rentenleistungen bei einer Rentenbezieherin bzw. einem Rentner. Wenn schon die Rentenkumulation bei einer Person nicht nachvollziehbar ist, so gilt dies natürlich auch für Ehepaare, d.h. die laufenden Rentenzahlungen können nicht auf Haushaltsebene aggregiert werden. Ob und in welcher Höhe Einkünfte aus anderen Alterssicherungssystemen oder u.U. aus eigener Erwerbstätigkeit vorliegen, lässt sich aus dieser Statistik nicht erkennen.

In welchem Umfang Renten aus der gesetzlichen Rentenversicherung kumulativ ausgezahlt werden, zeigt die Rentenzahlbestandsstatistik (**vgl. Abbildung 9**), die aus den laufenden Rentenzahlungen vom Renten Service der Deutschen Post und der Bundesknappschaft jeweils zum 1. Juli eines jeden Kalenderjahres gewonnen wird. D.h., in dieser Statistik können nicht nur einzelne Renten isoliert betrachtet werden, sondern auch Rentnerinnen und Rentner mit der Gesamtheit ihrer laufenden Rentenleistungen aus der gesetzlichen Rentenversicherung.

Abbildung 9: Rentenzahlbestandsstatistik (Post/Bundesknappschaft)

Termine	Jährlich zum 1.07. des Folgejahres
Datenform	Einzeldatensätze
Gegenstand	Laufende Rentenzahlungen am 01.07. des Berichtsjahres
Quelle	Postzahlbestand / Knappschaftszahlbestand
Wichtige Merkmale	Geburtsjahr, Geschlecht, Wohnort (Kreis), Rentenleistungsart, Rentenbeginn, Rentenbetrag und Zusammensetzung, Auffüllbetrag/Rentenzuschlag, Kranken-/ Pflegeversicherung (Art und Beitrag), Erstattungsbeträge
Wozu?	Struktur (Anzahlen, Alter, Beträge) von laufenden Rentenzahlungen, insbesondere Mehrfachrentenbezug
Probleme	Keine Versicherungszeiten, kein Haushaltsbezug, keine Information zu anderen Alterssicherungseinkünften

Quelle: Eigene Zusammenstellung, vgl. hierzu VDR Statistik Rentenbestand: Vorwort zu Band 148, insbesondere den Abschnitt zu „Rentenkumulationen“.

Als wichtige Merkmale liegen wiederum das Geburtsjahr, das Geschlecht, der Wohnort (Kreis), die Rentenleistungsart, der Rentenbeginn, sowie der Rentenbetrag und seine Zusammensetzung vor. Außerdem finden sich in dieser Statistik diverse Erstattungsbeträge, die der Rentenversicherung von anderen Stellen ersetzt werden. Problematisch für die Interpretation ist allerdings, dass in den Daten keine Versicherungszeiten enthalten sind, also keine Aussagen darüber gemacht werden können, in welchem zeitlichen Umfang die Versicherten Beiträge gezahlt haben. Auch die Wirkungen einzelner Berechnungsvorschriften lassen sich in der Rentenzahlbestandsstatistik nicht mehr nachvollziehen, weil entsprechende Informationen schon in den Ausgangsdaten nicht vorhanden sind. Die Statistik ist somit insbesondere dafür geeignet, die Struktur und den Umfang von Mehrfachrentenbezug auszuweisen. Darüber hinausgehende Fragestellungen, wie solche zum Haushaltskontext oder über Leistungen aus anderen Systemen der Alterssicherung, können mit dieser Statistik nicht analysiert werden.

2.4 Finanzstatistiken

Im Bereich der Finanzstatistiken existieren mehrere Berichterstattungen, von denen als wichtigste die monatlichen und jährlichen Rechnungsergebnisse zu nennen sind. Da jedoch keine Individualdaten vorliegen, wird hier auf diese Statistiken nicht näher eingegangen. Im Übrigen wird darauf hingewiesen, dass das Bundesministerium für Arbeit bzw. Gesundheit und Soziale Sicherung in seiner Reihe „Die gesetzliche Rentenversicherung in der Bundesrepublik Deutsch-

land im Jahre – statistischer und finanzieller Bericht“ detaillierte Rechnungsergebnisse der Rentenversicherungsträger veröffentlicht (vgl. BMA 1997).

2.5 Sonstige Statistiken

Anders als die Rechtsbehelfsstatistik und die Statistik über die Versendung von Versicherungsverläufen, Renteninformationen und Kontenklärungen beruht die jeweils zum 30. Juni eines Berichtsjahres erhobene Personalstatistik prinzipiell auf Einzeldatensätzen. Die Personalstatistik bezieht sich auf die bei der Rentenversicherung beschäftigten Personen. Allerdings gilt auch hier, dass nicht von allen Versicherungsträgern Einzeldatensätze vorliegen, sondern teilweise lediglich Summendaten geliefert werden. Insofern kann auf Basis der Personalstatistik keine Mikroanalyse durchgeführt werden. Ferner verfügt die amtliche Statistik mit der Personalstatistik nach § 7 Finanz- und Personalstatistikgesetz über eine Erhebung, in der vergleichbare Daten enthalten sind. Daher eignen sich weder die sonstigen Statistiken im Allgemeinen noch die Personalstatistik im Besonderen als Datensätze für das FDZ-RV.

2.6 Sondererhebungen

Zu nennen sind hier insbesondere die Sondererhebung „Vollendete Versichertenleben“ aus dem Rentenzugang, die zuletzt für das Berichtsjahr 2001 durchgeführt wurde, sowie die Sondererhebung „Altersvorsorge in Deutschland“, die derzeit für das Jahr 2002 aufbereitet wird.

Bei der Sondererhebung „Vollendete Versichertenleben“ werden die Daten aus dem Rentenzugang mit einem Versicherungsverlauf entsprechend der Versicherungskontenstichprobe kombiniert (vgl. **Abbildung 10**). D.h., es stehen zusätzlich zu den konkreten Rentenzugangsdaten die chronologischen Verlaufsdaten der rentenrechtlich relevanten Zeiten zur Verfügung.

Abbildung 10: Sondererhebung „Vollendete Versichertenleben“ aus dem Rentenzugang

Termine	Bei Bedarf einmal im Jahr, zuletzt 2001
Datenform	Einzeldatensätze wie beim Rentenzugang und zusätzlich wie bei der Versicherungskontenstichprobe
Gegenstand	Versicherungsverläufe zu aktuellen Versichertenrentenzugängen (als Stichprobe)
Quelle	Versicherungskonten
Wichtige Merkmale	Vgl. Rentenzugang und Versicherungskontenstichprobe
Wozu?	Modellrechnungen bei geplanten Rentenrechtsänderungen und für finanzielle Vorausschätzungen

Quelle: Eigene Zusammenstellung.

Diese Verbindung erlaubt es, die Einflüsse bestimmter Konstellationen im Versicherungsverlauf auf die spätere Rentenhöhe zu beurteilen. Allerdings werden solche Sondererhebungen nicht regelmäßig, sondern nur bei einem speziellen Bedarf der Rentenversicherung durchgeführt.

Bei der Sondererhebung „Altersvorsorge in Deutschland“ werden Daten aus dem chronologischen Versicherungsverlauf mit weiteren Daten aus einer Fragebogenerhebung in Verbindung gebracht, die u. a. auch andere Alterssicherungsformen betreffen (vgl. **Abbildung 11**).⁵ Aufgrund der Datenschutzklausel im Befragungsbogen besteht derzeit keine Möglichkeit, die Mikrodaten der AVID der Wissenschaft zur Verfügung zu stellen.

Abbildung 11: Sondererhebung „Altersvorsorge in Deutschland“ AVID

Termine	Bei Bedarf, zuletzt 2002
Datenform	Einzeldatensätze wie bei der Versicherungskontenstichprobe plus Befragungsdaten
Gegenstand	Befragungsdaten zu Ausbildung, Berufsleben, beabsichtigtem Rentenzeitpunkt und anderen Alterssicherungsformen sowie Versicherungsverläufe von 40 - 59-Jährigen im Ehepartnerschaftskontext
Quelle	Versicherungskonten + Befragung
Wichtige Merkmale	Vgl. Rentenzugang und Versicherungskontenstichprobe
Wozu?	Modellrechnungen für künftige Rechtsänderungen und Einordnung der Rentenversicherung ins bestehende System der Alterssicherung

Quelle: Eigene Zusammenstellung, vgl. www-altersvorsorge-in-deutschland.de.

Der Vorteil gerade dieser Sondererhebung ist vor allem, dass sie zum Einen über die Alterssicherung im Bereich der gesetzlichen Rentenversicherung hinausgeht und Anwartschaften und Zugehörigkeiten bei anderen Systemen der Altersvorsorge beleuchten kann. Zum Anderen ist es möglich, die Erwerbsbiographien von Versicherten und ihren Ehepartnern (im Haushaltskontext) unter Berücksichtigung möglicher Lücken im Versicherungsverlauf zu beschreiben. Die Datenfülle und der Längsschnittcharakter der Daten würden aber erhebliche Anonymisierungsprobleme aufwerfen.

3. Datenmengen und Verfügbarkeit

Die Daten zu fast allen Statistiken werden jeweils mit gewissem Abstand zum Erhebungsstichtag oder Berichtszeitraum bei den Versicherungsträgern erhoben, nach einer maschinellen Plausibilitätsprüfung dem VDR übermittelt und dort nochmals intensiven Plausibilitätsprüfungen unterzogen und ggf. bereinigt. Anschließend nimmt der VDR umfangreiche Auswertungen vor und sichtet diese, um die Ergebnisse abzusichern. Danach erfolgen die vorgesehenen in-

⁵ Für eine detaillierte Darstellung der AVID-Spezifika siehe den Beitrag „Erste Daten der Wissenschaft mit den Daten der GRV: Die Studie ‚Altersvorsorge in Deutschland‘ (AVID)“ von Thorsten Heien im vorliegenden Band. Für Interessenten an spezifischen Mikrodatenauswertungen der AVID zu Fragestellungen, die mit den mehreren tausend Tabellen in 7 Tabellenbänden nicht beantwortet werden können, gibt es nur folgenden Weg: Forschende können sich an das BMGS oder den VDR wenden und die Mikrodaten auswerten lassen (vgl. hierzu die Anmerkungen im Beitrag von Ralf Himmelreicher in diesem Band).

ternen und externen Berichtserstattungen und es werden teilweise tabellarisch Ergebnisse in gedruckter Form publiziert oder im Internet bereitgestellt. Grundsätzlich stehen die Datenbestände für einzelne Statistiken auch zu länger zurückliegenden Zeiträumen für Auswertungsvorhaben zur Verfügung. Ob und in welchem Umfang ältere Mikrodatensätze im Rahmen des FDZ-RV von Interesse sind, bleibt zu klären. Die vorhandenen Ressourcen sollen jedoch primär darauf verwandt werden, aktuelle Statistiken für das FDZ-RV aufzubereiten. Darüber hinaus könnten beim Rentenbezug und beim Rentenwegfall eine Aufbereitung der Daten aus einer langen Reihe von Berichtszeiträumen unter dem Gesichtspunkt des Zugangs bzw. Wegfalls von Renten für Jahrgangskohorten von besonderem Interesse sein. Es wird darauf aufmerksam gemacht, dass einzelne Datenbestände sehr umfangreich sind (**vgl. Abbildung 12**). Allerdings wird zur Anonymisierung der Daten durch Stichprobenziehung im Rahmen des FDZ-RV eine erhebliche Reduzierung des Datenumfangs erreicht werden.

Abbildung 12: Datenmengen und Verfügbarkeit

Bezeichnung der Statistik	Verfügbar seit	Anzahl der Datensätze*)	Anzahl der Merkmale *)
Stammsatzbestandsstatistik	1978	108 Mio.	18
Statistik der aktiv Versicherten	1982	42 Mio.	ca. 140
Versicherungskontenstichprobe	1992	0,5 Mio.	ca. 170 + 19*X
Versorgungsausgleichsstatistik	1981	4,6 Mio.	ca. 50
RSD Jahreserhebung (abgeschlossene Reha-Leistungen)	1971	1 Mio.	ca. 60
RSD-Verlaufserhebung	-	11 Mio.	ca. 22+140*X
Rentenzugangsstatik	1973	1,3 Mio.	ca. 210
Rentenwegfallsstatistik	1973	1,2 Mio.	ca. 150
Rentenänderungsstatistik (Rentenumwandlung)	1973	0,2 Mio.	ca. 210 + 150
Rentenbestandsstatistik aus den Konten der Rentenversicherungsträger	1992	24 Mio.	ca. 150
Rentenzahlbestandsstatistik Post/Bundesknappschaft	1981	24 Mio.	ca. 75

*) bei der aktuellen Erhebung.

Quelle: Eigene Zusammenstellung.

Die Mikrodatensätze der Rentenversicherung umfassen im Querschnitt teilweise mehr als hundert Merkmale, im Fall von Längsschnittinformationen wird ein Vielfaches an Merkmalen erreicht. Nähere Einzelheiten dazu sind in den Dokumentation der jeweiligen Statistik abgedruckt. Dazu zählen zunächst die Benutzerhinweise, die es zu den meisten Statistiken gibt. Sie enthalten neben Angaben zu den zugrundeliegenden Datenquellen sowie zum Erhebungsdesign auch weitere mögliche Besonderheiten bei der Datengewinnung und Angaben zum Meldeweg.

Die jeweiligen Merkmale mit ihren Ausprägungen sind in so genannten Datensatzbeschreibungen dokumentiert. Diese sind in ihrer derzeitigen Form für die interne Nutzung vorgesehen, sie erfordern deshalb vielfach Kenntnisse im Rentenrecht sowie über den Verfahrensablauf. Umfangreiche Verschlüsselungen, wie z.B. für Diagnosen oder Berufe ergeben sich mitunter aus separaten Schlüsselverzeichnissen. Merkmalszusammenhänge erschließen sich teilweise auch aus den Fehlerverzeichnissen, die für die maschinelle Plausibilitätsprüfung verwendet werden. Den Interessenten der verschiedenen Statistiken wird empfohlen, sich die Vorworte zu den publizierten Statistikbänden der Reihe VDR-Statistik sowie die Anhänge dieser Publikationen zu vergegenwärtigen. Diese Texte enthalten wichtige Hinweise zum Erhebungsablauf, zu den Besonderheiten in bestimmten Berichtsjahren und den verwendeten Fachbegriffen.

4. Pilotprojekt Versicherungskontenstichprobe 2001

Der VDR hat im Jahr 2003 eine projektbezogene Anonymisierung von ausgewählten Mikrodaten aus der Versicherungskontenstichprobe des Jahres 2001 vorgenommen. Der dazu entwickelte Mikrodatsatz war ein Pilotprojekt, anhand dessen der Zugang zu anonymisierten Mikrodaten der Rentenversicherung erprobt wurde. Dieser Datensatz wird im Folgenden als „Pilotprojekt Versichertenkontenstichprobe 2001“ bezeichnet.⁶ Neben demographischen Merkmalen der Versicherten enthält dieser Datensatz Verlaufsinformationen, speziell zu Beitragsarten und Beitragshöhen mit Zeitpunkten und Zeitdauern, die erreichten Entgeltpositionen sowie ggf. die Rentenleistungsart und das exakte Alter bei Rentenbeginn. Weil die Determinanten des Renteneintritts im Vordergrund der Analyse stehen, konnte die Fallauswahl auf Versicherte im Alter von 55 bis 65 Jahren beschränkt werden. Bei der Knappschaft Versicherte und Ausländer wurden ausgeschlossen. Außerdem wurden nur „geklärte“ Versicherungskonten in das Pilotprojekt einbezogen, weil ungeklärte Versicherungskonten keine zuverlässigen Rückschlüsse erlaubt hätten. Im Sinne einer Anonymisierung sind Versicherte mit selten in Anspruch genommene Leistungsarten aus dem Datenbestand selektiert worden. Nach mehreren Abstimmungsgesprächen mit den Wissenschaftlern wurden ausschließlich konventionelle Anonymisierungsmethoden eingesetzt. Dazu wurden nicht benötigte Merkmale, die zur Deanonymisierung hätten dienen können, wie etwa der Tätigkeitsschlüssel und sensible Merkmale, generell gelöscht. Metrische Merkmale wurden generell vergrößert. Bei nicht metrischen Merkmalen wurden Ausprägungen soweit als möglich zu sinnvollen Gruppen zusammengefasst, wie z.B. Rentenleistungsarten, die zu Rentenleistungsartgruppen verdichtet wurden.

Besonderes Augenmerk galt der Anonymisierung der Daten im chronologischen Teil der Versicherungskontenstichprobe, zumal die chronologische Abfolge erkennbar bleiben sollte. Zunächst wurde das Berichtsfenster, das sich prinzipiell auf das gesamte Versicherungsleben erstreckte, auf ein 10-Jahresfenster verkleinert und somit auf das Berichtsjahr und die neun davor liegenden Kalenderjahre reduziert. Die einzelnen Beitragsarten wurden wiederum zu Beitragsartgruppen zusammengefasst. Für die Untersuchung nicht benötigte Sachverhalte, wie z.B. Sachbezug, Zeiten der Kindererziehung oder Zeiten nach dem Fremdrentengesetz sowie konkrete Entgeltangaben wurden gelöscht. Sämtliche Datumsangaben wurden vergrößert und schließlich wurden alle angegebenen Entgeltpunkte gerundet.

⁶ Der Datensatz „Pilotprojekt Versichertenkontenstichprobe 2001“ ist die empirische Basis der im vorliegenden Band abgedruckten Studie „Determinanten des Renteneintritts - eine Untersuchung mit prozessproduzierten Daten“ von Stefan Kempf und Martin Kukuk.

In Sachen Pilotprojekt Versicherungskontenstichprobe 2001 kann festgehalten werden, dass der betriebene Aufwand zur Aufbereitung und Anonymisierung des Originaldatensatzes erheblich war. Bedenkt man ferner, dass sich mit dem Datensatz dieses Pilotprojektes nur noch vergleichsweise wenige Fragestellungen analysieren lassen, dann erhält man eine grobe Vorstellung davon, in welchem Umfang in Zukunft vor allem personelle Ressourcen in das FDZ-RV eingebracht werden müssen.

5. Fazit

Das Angebot an überwiegend prozessproduzierten Mikrodaten der Rentenversicherung ist sehr groß. Sie sind dafür geeignet, auf zahlreiche wissenschaftliche Fragestellungen verschiedener Disziplinen empirisch gestützte Antworten zu finden.⁷ Da das Gros der Daten vereinfacht gesprochen dafür erhoben wurde, die Höhe der jeweiligen Leistung für die Versicherten zu ermitteln, entsprechen sie nicht grundsätzlich wissenschaftlichen Anforderungen. Anstelle tiefgegliederter Differenzierung, wie sie häufig in wissenschaftsbasierten Umfragen möglich sind, haben die Daten der Rentenversicherung in der Regel die Eigenschaft, dass sie im Rahmen von Totalerhebungen meist sehr exakt erhoben wurden. D.h., wichtiges Kennzeichen der Daten der Rentenversicherung sind hohe Fallzahlen und eine hohe Validität.

Anhand des Datensatzes „Pilotprojekt Versichertenkontenstichprobe 2001“ konnte exemplarisch gezeigt werden, dass es möglich ist, selbst einen Längsschnittdatensatz der Rentenversicherung für die Wissenschaft aufzubereiten. Und zwar aufzubereiten in einer Weise, die sowohl den strengen Kriterien des Datenschutzes als auch den Anforderungen der Wissenschaftler genügt. Wenngleich dieses Beispiel lehrt, dass die zukünftige Zusammenarbeit zwischen Wissenschaft und FDZ-RV keinen finalen sondern eher einen prozesshaften Charakter haben wird, weil jede neue Erkenntnis gleichsam neue Fragen aufwerfen wird.

Die Vorgehensweise beim „Pilotprojekt Versichertenkontenstichprobe 2001“ kann sicherlich ein Modell für Gastwissenschaftler-Arbeitsplätze im FDZ-RV sein. Um im Rahmen des FDZ-RV Scientific Use Files und Public Use Files anbieten zu können, sind jedoch solche Lösungen anzustreben, die für die Aufbereitung und die Anonymisierung der Daten eine wesentlich bessere Aufwand-Nutzen Relation aufweisen.

⁷ An den Daten der Rentenversicherung sind nach den Befunden der FDZ-RV-Umfrage derzeit 40 Forscher(gruppen) aus 11 verschiedenen wissenschaftlichen Disziplinen, vor allem aus Ökonomie, Demographie, Soziologie und Medizin, interessiert (vgl. hierzu den Beitrag von Ralf Himmelreicher im vorliegenden Band).

Literatur:

- Bundesministerium für Arbeit und Sozialordnung (1997): Die gesetzliche Rentenversicherung in der Bundesrepublik Deutschland im Jahre 1996 - Statistischer und finanzieller Bericht. Bonn.
- Rehfeld, U. (2001): Die Statistiken der gesetzlichen Rentenversicherung. *Deutsche Rentenversicherung*, 3-4, 169-188.
- Verband Deutscher Rentenversicherungsträger (Hg.) (2004): *VDR Statistik Rehabilitation des Jahres 2002*, Band 146. Frankfurt am Main.
- Verband Deutscher Rentenversicherungsträger (Hg.) (2004): *VDR Statistik Rentenbestand am 31. Dezember 2003*, Band 148. Frankfurt am Main.
- Verband Deutscher Rentenversicherungsträger (Hg.) (2004): *VDR Statistik Rentenzugang des Jahre 2003 einschließlich Rentenwegfall, Rentenänderung/Änderung des Teilrentenanteil*, Band 148. Frankfurt am Main.
- Verband Deutscher Rentenversicherungsträger (Hg.) (2004): *VDR Statistik Versicherte 2001/2002*, Band 147. Frankfurt am Main.

Hilmar Luckert hat an der Julius Maximilians Universität Würzburg Mathematik studiert und sich dabei schwerpunktmäßig mit mathematischer Statistik befasst. Nach Abschluss des Studiums kam er 1981 zum VDR und war zunächst als Referent für verschiedene Statistiken zuständig. Nachdem 1987 die Abteilung „Volkswirtschaft und Statistik“ in Frankfurt eingerichtet wurde blieb Hilmar Luckert mit seinem Referat in Würzburg bis heute zuständig für die technische Durchführung sämtlicher Routinestatistiken und Berichterstattungen. Inhaltlich betreute er selbst lange Jahre die Versicherungskontenstichprobe und Sondererhebungen, entwickelte und gestaltet aber auch alle anderen Statistiken mit. Als Leiter der Abteilung „Berufliche Bildung, Statistisches Berichtswesen, Datenschutz und Datensicherheit“ sind inzwischen zum ursprünglichen Arbeitsschwerpunkt Statistik weitere hinzugekommen. Der Aufbau des Forschungsdatenzentrums der Rentenversicherung, dabei insbesondere die programm- und datentechnische Konzeption und Realisierung des Vorhabens, sieht er in diesem und im nächsten Jahr als eine seiner Kernaufgaben an.

Aspekte des Datenschutzes im Forschungsdatenzentrum der Rentenversicherung*

Claudia Heese (VDR Frankfurt)

1. Vorbemerkung

Anliegen des Datenschutzes ist es nicht, „Sand ins Getriebe“ zu streuen oder Forschungsvorhaben zu vereiteln. Datenschutz soll vielmehr widerstreitende Interessen miteinander in Einklang bringen. Bei Forschungsvorhaben muss das berechnigte Interesse des Wissenschaftlers an der Gewinnung neuer Erkenntnisse mit dem ebenfalls berechtigten Interesse des Betroffenen an der Geheimhaltung der ihn betreffenden Daten in Einklang gebracht werden. Um diese Abwägung, die sich insbesondere in § 75 SGB X spiegelt, verständlicher zu machen, sollen im Folgenden die Grundzüge des Datenschutzrechts und hier insbesondere des Sozialdatenschutzrechts kurz dargestellt werden. Es werden die Anforderungen des § 75 SGB X erläutert ebenso wie die Konsequenzen, die sich hieraus für einzelne Forschungsvorhaben und für den Betrieb des Forschungsdaten-zentrums der gesetzlichen Rentenversicherung (FDZ-RV) ergeben.

2. Grundzüge des Sozialdatenschutzrechts

2.1 Begriff der Sozialdaten

Zentraler Begriff des Datenschutzrechts ist der des Sozialgeheimnisses. Danach hat jeder einen Anspruch darauf, dass seine Sozialdaten von den Leistungsträgern nicht unbefugt erhoben, verarbeitet oder genutzt werden (vgl. § 35 Abs. 1 Satz 1 SGB I). Dieser stark ausgeprägte Schutz der Sozialdaten als Ausfluss des verfassungsrechtlich verankerten Allgemeinen Persönlichkeitsrechts ist das Gegengewicht dazu, dass die Sozialversicherung eine Pflichtversicherung ist und der Einzelne es sich nicht aussuchen kann, ob seine Sozialdaten gespeichert werden oder nicht. Sozialdaten wiederum sind gemäß § 35 Abs. 1 SGB I Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmaren natürlichen Person, die von einem Sozialleistungsträger im Hinblick auf seine Aufgaben nach dem Sozialgesetzbuch erhoben, verarbeitet oder genutzt werden. Sie setzen sich somit aus zwei Komponenten zusammen, nämlich einer persönlichen und einer sachlichen.

Die persönliche Komponente besteht darin, dass es sich bei Sozialdaten immer um Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmaren natürlichen Person handeln muss. Diese Person wird als Betroffener bezeichnet. Persönliche Verhältnisse sind beispielsweise Geburtsdatum, Versicherungsdaten, Höhe des Arbeitsentgelts, Rentendaten und Krankheiten. Wohingegen unter sachlichen Verhältnissen z.B. gestellte Anträge zu verstehen sind.

* Ich danke Romina Schneider und Boris Pantzke für die Erstellung des im Anhang des Beitrags beigefügten Prüfschemas.

Im Einzelfall kann es schwer sein, persönliche und sachliche Verhältnisse gegeneinander abzugrenzen. Da die Unterscheidung rechtlich keine Rolle spielt, sollte man nicht nach Abgrenzungskriterien suchen, sondern die beiden Begriffe als Einheit ansehen. Wichtig ist, dass alle Informationen, die etwas über eine Person aussagen, vom Schutz des Gesetzes erfasst sind. Der Gesetzgeber stellt nicht auf die persönliche Geheimnissphäre und damit auf ein individuelles Geheimhaltungsinteresse ab. Im geltenden Recht sind vielmehr alle Angaben über einen Betroffenen gleichermaßen schutzwürdig, unabhängig davon, ob es sich um elementare Daten wie Name oder Anschrift, oder ob es sich um besonders sensible Daten wie beispielsweise medizinische Daten handelt. Weiterhin unbeachtlich ist, ob die Daten bereits offenkundig sind oder nicht.

Die Einzelangaben müssen außerdem einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können. Auf eine bestimmte Person weist der Name, der Vorname und das Geburtsdatum hin. Mit Hilfe dieser Daten ist eine Person direkt erkennbar. Eine Person ist dagegen bestimmbar, wenn ihre Identität durch Zusatzabfragen herausgefunden werden kann. Beispiele hierfür sind die Versicherungsnummer, ein Aktenzeichen, eine Stammmnummer beim Arbeitsamt u.ä. Eine Person ist aber auch dann bestimmbar, wenn eine Angabe nur auf sie zu treffen kann, z.B. der erste Direktor der LVA Thüringen.

Neben dieser gerade beschriebenen persönlichen Komponente ist dem Begriff der Sozialdaten auch noch eine sachliche Komponente zueigen. Eine bestimmte Information wird nämlich nur dann als Sozialdatum geschützt, wenn sie von einer in § 35 SGB I genannten Stelle im Hinblick auf deren Aufgaben nach dem Sozialgesetzbuch erhoben, verarbeitet oder genutzt wird.

Was Aufgaben in diesem Sinne sind, wird in § 67 Abs. 2 SGB X näher umschrieben. Neben den „eigentlichen“ Aufgaben der Rentenversicherungsträger nach dem Sozialgesetzbuch - nämlich Feststellung und Erbringung von Leistungen (vgl. § 148 SGB VI) -, zählen hierzu auch Aufgaben aufgrund von Verordnungen, deren Ermächtigungsgrundlage sich im Sozialgesetzbuch befindet.

2.2 Betriebs- und Geschäftsgeheimnisse

Den Sozialdaten sind gemäß § 35 Abs. 4 SGB I Betriebs- und Geschäftsgeheimnisse gleichgestellt. Was darunter zu verstehen ist, ist in § 67 SGB X definiert. In Abs. 1 Satz 2 dieser Vorschrift heißt es, dass Betriebs- und Geschäftsgeheimnisse alle betriebs- oder geschäftsbezogenen Daten - auch die von juristischen Personen - sind, die Geheimnischarakter haben. Die Informationen dürfen nur einem eng begrenzten Personenkreis bekannt sein, und der Betriebsinhaber muss ein berechtigtes, schützenswertes Interesse an ihrer Geheimhaltung haben. Zu dieser Art von Informationen können beispielsweise die Höhe der Schulden, die Anzahl der geringfügig Beschäftigten u.ä. zählen. Wichtig ist, dass auch diese Daten dem Leistungsträger im Rahmen seiner Aufgabenerfüllung nach dem Sozialgesetzbuch bekannt geworden sein müssen. Die Gleichstellung der Betriebs- und Geschäftsgeheimnisse mit den Sozialdaten trägt dem Umstand Rechnung, dass bestimmte Firmendaten hinsichtlich ihrer Sensibilität den Sozialdaten durchaus vergleichbar sind.

2.3. Anonymisieren

Die Qualifikation von Daten als Sozialdaten entfällt, sobald sie anonymisiert werden. Nach der Legaldefinition des § 67 Abs. 8 SGB X sind Sozialdaten dann anonymisiert, wenn sie derart verändert sind, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können. Der Gesetzgeber unterscheidet folglich zwischen der echten und der faktischen Anonymisierung. Liegt eine echte Anonymisierung vor, dann kann der Personenbezug auch mit Zusatzinformationen nicht wieder hergestellt werden. Ausreichend ist jedoch bereits eine faktische Anonymisierung, die immer dann vorliegt, wenn die Wiederherstellung des Personenbezugs nicht absolut ausgeschlossen werden kann, jedoch der Aufwand für diese Deanonimisierung außerhalb eines vernünftigen Verhältnisses zu dem zu erwartenden Nutzen steht (vgl. Rombach 2002, RN120).

Folglich ist dann noch keine Anonymisierung von Daten erfolgt, wenn zwar bei einer Fragebogenaktion die Fragebogen selbst keine Angaben über die befragten Personen enthalten, wohl aber durchnummeriert sind und es noch eine Liste gibt, mittels der eine Nummer wieder einer bestimmten Person zugeordnet und damit der Personenbezug ohne großen Aufwand wieder hergestellt werden kann.¹

3. Übermittlung von Sozialdaten für Forschung und Planung gemäß § 75 SGB

Ist ein Forschungsvorhaben nicht mit anonymisierten Daten im oben beschriebenen Sinne durchführbar, dann können nur beim Vorliegen der Voraussetzungen des § 75 SGB X Sozialdaten übermittelt werden. Diese Vorschrift unterscheidet sich von den übrigen Übermittlungsvorschriften des Sozialgesetzbuches dadurch, dass sie der Übermittlung mit Einwilligung des Betroffenen den absoluten Vorrang einräumt, während bei den übrigen Übermittlungstatbeständen des SGB X Einwilligung des Betroffenen und gesetzliche Ermächtigung zur Übermittlung gleichberechtigte Rechtsgrundlagen sind.

3.1 Bestimmtes Forschungs- oder Planungsvorhaben im Sozialleistungsbereich

Erste Voraussetzung einer Übermittlung nach § 75 SGB X ist, dass das durchzuführende Forschungs- oder Planungsvorhaben hinreichend präzise ist, weil sich andernfalls die Notwendigkeit der Verwendung personenbezogener Daten nicht ausreichend sicher beurteilen lässt. Nur bei einem hinreichend konkretisierten Vorhaben kann dem Vorwurf einer Datenhaltung auf Vorrat begegnet werden.

Nach allgemeiner Auffassung ist unter Forschung jeder ernsthafte und planmäßige Versuch, in methodischer und nachprüfbarer Weise neue Erkenntnisse zu erlangen, zu verstehen.² Planung

¹ Zur Herstellung von faktischer Anonymität bei Mikrodaten vgl. Lenz et al. (2004).

² Ebenfalls für eine forschungsfreundliche Auslegung plädiert Walz (1989, RN 29).

hingegen ist die Festlegung künftigen Verhaltens auf der Basis der Abschätzung künftiger Tatsachen. Während wissenschaftliche Forschung sowohl von öffentlichen als auch von privaten Instituten durchgeführt werden kann, können Sozialdaten für ein bestimmtes Vorhaben der Planung nur dann gemäß § 75 SGB X übermittelt werden, wenn dieses von öffentlichen Stellen durchgeführt wird.

Eine zulässige Übermittlung nach § 75 Abs. 1 SGB X setzt weiterhin voraus, dass das geplante Vorhaben im Sozialleistungsbereich erfolgt, d.h., das Forschungs- oder Planungsziel muss einen erkennbaren thematischen Bezug zu den Inhalten, Trägern oder Strukturen des Systems der sozialen Sicherung aufweisen (Wiese 1980, 459). Auf dieses tatbestandliche Erfordernis sei hier noch einmal ausdrücklich hingewiesen, hat es doch zur Folge, dass beispielsweise soziologische Grundlagenforschung unzulässig ist (vgl. 2. TB/LfD für das Saarland, Ziff. 8.1.). Vielmehr hat der Gesetzgeber durch die Eingrenzung der Forschungszwecke sichergestellt, dass Daten, die von der Sozialverwaltung verarbeitet werden, auch nur für solche Forschungs- oder Planungsvorhaben zur Verfügung stehen, die sich ihrerseits wieder auf den Sozialleistungsbereich konzentrieren.

3.2 Einholen der Einwilligung nicht zumutbar

Personenbezogene Daten dürfen gemäß § 75 SGB X nur dann übermittelt werden, wenn es unzumutbar ist, die Einwilligung des Betroffenen in die Übermittlung einzuholen. Wegen des eingangs erwähnten Vorrangs der Übermittlung mit Einwilligung sind an die Unzumutbarkeit der Einholung der Einwilligung hohe Anforderungen zu stellen. Sie ist nur dann zu bejahen, wenn beispielsweise die Bekanntgabe des Vorhabens bei dem Betroffenen zu einem unverhältnismäßigen gesundheitlichen Nachteil führen würde. Dies kann insbesondere bei der Krebs- und Psychiatrieforschung der Fall sein, wenn der Betroffene selbst von seiner Erkrankung keine Kenntnis hat. Die Einholung einer Einwilligung ist auch dann unzumutbar, wenn Daten von Betroffenen übermittelt werden sollen, deren Adressen den Trägern nicht bekannt sind, wie dies beispielsweise bei einer Studie über in ihre Heimat zurückgekehrte Ausländer der Fall wäre.

Hingegen kann hoher Verwaltungs- oder Kostenaufwand ebenso wie die Eilbedürftigkeit nur in Ausnahmefällen die Unzumutbarkeit der Einholung der Einwilligung rechtfertigen. Allerdings kann die Einholung einer Einwilligung unter Umständen dann unzumutbar sein, wenn mit einem geringen Rücklauf an Einwilligungserklärungen gerechnet werden muss und es infolgedessen zu soziodemographischen Verzerrungen kommt.

Falls es zumutbar ist, eine Einwilligung des Betroffenen einzuholen, dann muss diese im Vorfeld der Übermittlung erteilt werden und kann nicht erst nachträglich eingeholt werden, was sich aus der gesetzlichen Formulierung „eingewilligt hat“ ergibt (vgl. § 183 BGB). Es handelt sich bei der Einwilligungserklärung um eine höchstpersönliche Erklärung, die vom Betroffenen selbst abgegeben werden muss. Sie muss hinsichtlich Inhalt, Zweck und Ausmaß hinreichend bestimmt sein.³

³ Ausführlich zu den Voraussetzungen einer wirksamen Einwilligungserklärung, vgl. Hesse (1997, 317 ff.).

3.3 Keine Beeinträchtigung schutzwürdiger Interessen bzw. überwiegendes öffentliches Interesse

Eine zulässige Übermittlung von Sozialdaten setzt weiterhin voraus, dass im Einzelfall positiv festgestellt worden ist, dass durch die Übermittlung keine schutzwürdigen Interessen des Betroffenen beeinträchtigt werden. Dabei sind die schutzwürdigen Interessen ausschließlich aus Sicht des Betroffenen zu beurteilen, und sie sind um so schutzwürdiger, je eher sie einen Bezug zur Intimsphäre aufweisen. Hieraus folgt beispielsweise, dass das Interesse des Betroffenen geringer wiegt, wenn für die Feststellung regionaler Besonderheiten seine Adresse übermittelt werden soll, als wenn medizinische Daten aus seinem Krankenhausentlassungsbericht übermittelt werden sollen. Ist allerdings die Adresse mit für den Betroffenen nachteiligen Zusatzinformationen verbunden (z.B. Anschrift eines Gefängnisses), kann dies wiederum zu einer anderen Gewichtung führen, als wenn es sich um eine reine Wohnanschrift handelt.

Allerdings ist auch trotz einer Beeinträchtigung von schutzwürdigen Interessen des Betroffenen eine Übermittlung von Sozialdaten dann erlaubt, wenn das öffentliche Interesse an der Forschung oder Planung das Geheimhaltungsinteresse des Betroffenen erheblich überwiegt.

3.4 Vorliegen einer Genehmigung

Weitere Voraussetzung einer zulässigen Übermittlung ist das Vorliegen einer Genehmigung durch die zuständige oberste Bundes- oder Landesbehörde, die für den Bereich, aus dem die Daten stammen, zuständig ist.

Die Genehmigung muss gemäß § 75 Abs. 2 Satz 3 SGB X den Dritten, an den die Daten übermittelt werden, die Art der zu übermittelnden Sozialdaten und den Kreis der Betroffenen, die wissenschaftliche Forschung oder die Planung, zu der die übermittelten Sozialdaten verwendet werden dürfen, und den Tag, bis zu dem die übermittelten Sozialdaten aufbewahrt werden dürfen, bezeichnen. Sie steht kraft Gesetzes unter dem Vorbehalt der nachträglichen Aufnahme, Änderung oder Ergänzung einer Auflage; allerdings darf sie nur versagt werden, wenn die oben beschriebenen Voraussetzungen des § 75 Abs. 1 SGB X nicht vorliegen.

3.5 Übermittlung an nicht-öffentliche Stellen

Sollen Sozialdaten zu Forschungszwecken an eine nicht-öffentliche Stelle übermittelt werden, dann ist die Genehmigungsbehörde verpflichtet, durch entsprechende Auflagen dafür zu sorgen, dass die Voraussetzungen des § 75 SGB X erfüllt werden und die Daten ausschließlich zu Forschungs- oder Planungszwecken gespeichert, verändert oder genutzt werden.

3.6 Einschränkung der Übermittlungsbefugnis bei besonders schutzwürdigen Sozialdaten (§ 76 SGB X)

Werden den Sozialleistungsträgern von einem Arzt oder einer anderen in § 203 StGB genannten Stelle Sozialdaten zugänglich gemacht, dann unterliegt der Träger hinsichtlich dieser Daten

den gleichen Geheimhaltungspflichten wie diese Stellen. Nach einhelliger Auffassung können diese Daten bei Vorliegen nur einer ausdrücklichen oder mutmaßlichen Einwilligung, beim Bestehen gesetzlicher Mitteilungspflichten oder aber beim Vorliegen eines übergesetzlichen Notstandes übermittelt werden.⁴

4. Forschungsgeheimnis

In der Literatur wurde immer wieder über die Schaffung eines Forschungsgeheimnisses nachgedacht (vgl. Vetter 1999, Weichert 1996, Wagner 2003). Ein solches wurde insbesondere deshalb für erforderlich gehalten, weil die medizinische Forschung durch die datenschutzrechtlichen Regelungen verhindert oder doch zumindest verteuert werde, was sich letztlich schädlich auf die Allgemeinheit aber auch auf den Wissenschaftsstandort Deutschland auswirke (vgl. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften 1995/1998).

Fraglich ist, ob sich durch die Schaffung eines Forschungsgeheimnisses die Hoffnungen seiner Befürworter auch erfüllen. Bereits jetzt ist die Forschungsfreiheit durch Art. 5 Abs. 3 GG geschützt. Da auch das Allgemeine Persönlichkeitsrecht durch Art. 2 Abs. 1 in Verbindung mit Art. 1 Abs. 1 GG verfassungsrechtlich geschützt ist, stehen sich zwei Grundrechte gegenüber, die im Wege der systematischen Auslegung durch praktische Konkordanz miteinander in Einklang gebracht werden müssen, d.h. jedes der beiden Grundrechte kann nur so weit reichen, wie es der Schutzbereich des anderen Grundrechts erlaubt (vgl. Pieroth & Schlink 2003, RN 321). Dies hätte zur Folge, dass auch nach Kodifizierung eines Forschungsgeheimnisses die Einwilligung des Betroffenen eingeholt werden müsste, wenn seine Daten für Forschungszwecke verwendet werden und dies ohne Gefährdung des Forschungszweckes möglich und auch zumutbar ist (Binne 2000).

Allerdings wäre bei der Schaffung eines Forschungsgeheimnisses das Schutzniveau der bei den forschenden Stellen gespeicherten und genutzten personenbezogenen Daten erheblich verbessert. So könnte beispielsweise bei medizinischen Forschungsdaten das Forschungsgeheimnis den Daten den gleichen Schutz angedeihen lassen, wie sie ihn beispielsweise bei einem Arzt genießen. Die Daten wären dann dem Zugriff insbesondere der Exekutive entzogen, so dass ein Beschlagnahmeverbot und ein Zeugnisverweigerungsrecht bestünden. Möglicherweise würde der erhöhte Schutz der Daten den Betroffenen dann wiederum eher in ein Forschungsvorhaben einwilligen lassen. Dies wäre aber die einzige Erleichterung, die durch Kodifizierung eines expliziten Forschungsgeheimnisses einträte.

5. Resümee

Auf den ersten Blick erscheint die Einhaltung der datenschutzrechtlichen Vorschriften die wissenschaftliche Forschung zwar erheblich zu behindern. Es gilt allerdings zu beachten, dass die Daten der Rentenversicherung vorrangig der Erbringung von Sozialleistungen dienen und des-

⁴ Vgl. Gemeinsames Rundschreiben der Spitzenverbände der Sozialversicherung zu den datenschutzrechtlichen Vorschriften des SGB I und SGB X in der Fassung des Gesetzes zur Änderung des Bundesdatenschutzgesetzes und anderer Gesetze, § 76, RN 5.

halb der Betroffene grundsätzlich nicht damit rechnen muss, dass seine Daten zu Forschungszwecken weitergegeben werden. Dieses Vertrauen gilt es, auch im Hinblick auf die Akzeptanz der Forschungsprojekte, mit dem berechtigten Interesse der Wissenschaftler an der Gewinnung neuer Erkenntnisse in Einklang zu bringen. Das IAB hat dies in einem Kurzbericht über den Start seines Forschungszentrums dahingehend auf den Punkt gebracht, dass „bei allen Bemühungen um die größtmögliche Kundenorientierung die rechtlichen Bestimmungen zum Schutz der Sozialdaten den unumstößlichen Rahmen bilden.“ (IAB 2004, 2)

Sollte sich aber während der zweijährigen Pilotierungsphase des FDZ-RV zeigen, dass der Schutz des Betroffenen durch weniger einschneidende Maßnahmen gleichermaßen sichergestellt werden kann, dann wäre dies ein Ergebnis, das unbedingt im Abschlussbericht über das Projekt Eingang finden und dem Gesetzgeber zur Kenntnis gebracht werden müsste.

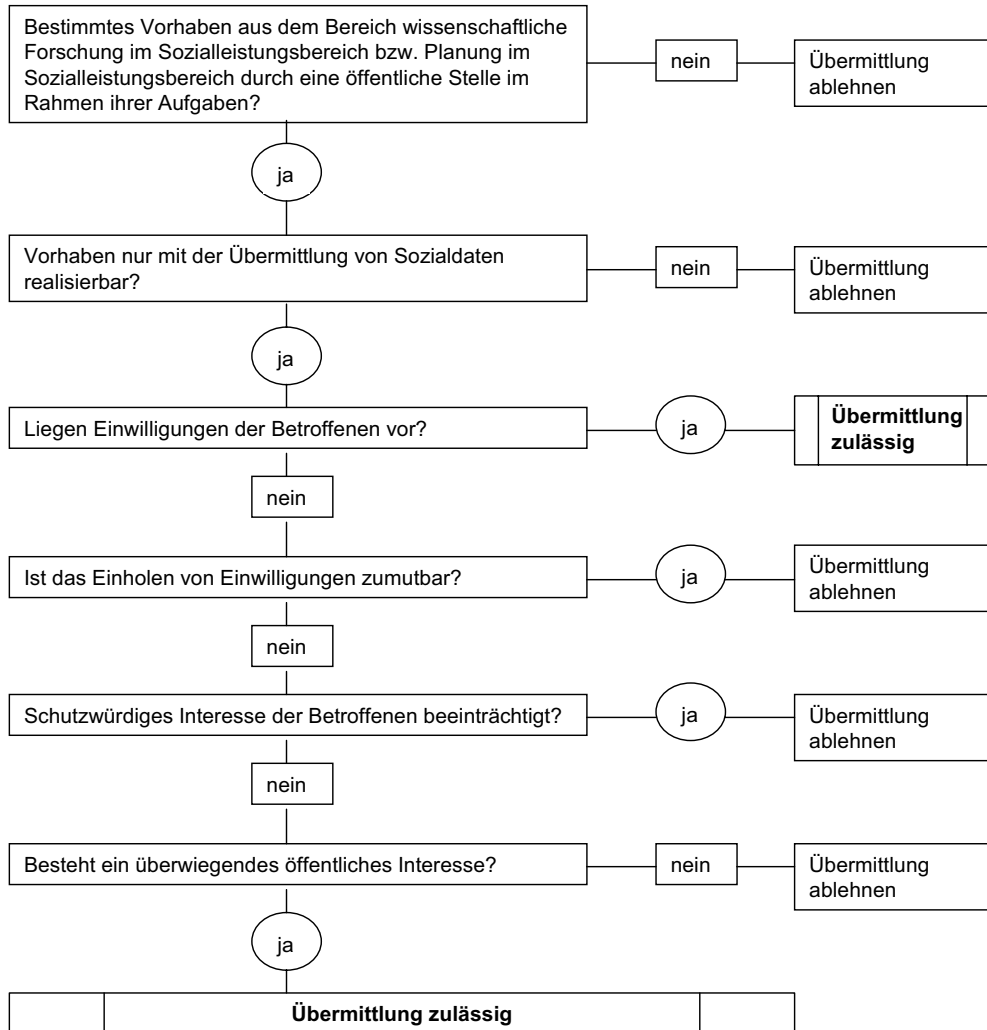
Literatur

- Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (1995 und 1998): *Resolution zum Datenschutz / medizinischen Forschungsgeheimnis*. <http://www.uni-duesseldorf.de/WWW/AWMF/awmf-frp.htm>.
- Binne, W. (2000): Datenschutz in der Rehabilitationsforschung. In: Bengel, J. & Koch, U. (Hg.): *Grundlagen der Rehabilitationswissenschaften*. Berlin: Springer, Kap. 30.
- Gemeinsames Rundschreiben der Spitzenverbände der Sozialversicherung zu den datenschutzrechtlichen Vorschriften des SGB I und SGB X in der Fassung des Gesetzes zur Änderung des Bundesdatenschutzgesetzes und anderer Gesetze.
- Heese, C. (1997): Sozialdatenschutz und wissenschaftliche Untersuchungen in der gesetzlichen Rentenversicherung. *Deutsche Rentenversicherung*, 5-6, 311-325.
- IAB (2004): Forschungsdatenzentrum der BA geht an den Start, *IAB-Kurzbericht*, 9, 1-4.
- Lenz, R., Sturm, R. & Vögrimmer, D. (2004): Maße für die faktische Anonymität von Mikrodaten. *Wirtschaft und Statistik*, 6, 621-638.
- Pieroth, B. & Schlink, B. (2003): *Grundrechte - Staatsrecht II*, 19. Aufl. Heidelberg.
- Rombach, W. (2002): In: Hauck (Hg.): *Kommentar zum Sozialgesetzbuch SGB X/1,2*.
- Vetter, R. (1999): Forschungs(daten)geheimnis. *Datenschutz und Datensicherheit*, 7, 389-391.
- Wagner, J. (2003): Arbeiten mit Einzeldaten der amtlichen Statistik am Beispiel des Monatsberichts im Verarbeitenden Gewerbe. In: Ronning, G. & Gnoss, R. (Hg.): *Anonymisierung wirtschaftsstatistischer Einzeldaten*. Forum der Bundesstatistik, Band 42, Wiesbaden: Statistisches Bundesamt, 140-146.
- Walz, S. (1989): In: Borchert, G., Hase, F. & Walz, S. (Hg.): *Gemeinschaftskommentar zum Sozialgesetzbuch - Schutz der Sozialdaten (GK-SGB X 2)*. Neuwied/Frankfurt a.M.: Luchterhand Verlag.
- Weichert, T. (1996): Datenschutz und medizinische Forschung – Was nützt ein „medizinisches Forschungsgeheimnis“? *Medizinrecht*, 14. Jg., 258-261.
- Wiese, W. (1980): Der Schutz der Sozialdaten. *Die Angestelltenversicherung*, 27. Jg., 449-461.

Claudia Heese: Juristin, Leiterin des Referates „Grundsatzfragen des Systems der sozialen Sicherheit“ (derzeit Alterseinkünftegesetz und Organisationsreform als große Themen) und Datenschutzbeauftragte des VDR.

Anhang

Abbildung 1: Prüfschema zur Feststellung der Zulässigkeit der Übermittlung von Sozialdaten bei wissenschaftlichen Forschungs- und Planungsvorhaben (§ 75 SGB X)



Besonderheiten § 76 SGB X

Sofern Sozialdaten vorliegen, die dem besonderen Berufsgeheimnis unterliegen, muss

- eine ausdrückliche oder mutmaßliche Zustimmung der Betroffenen vorliegen
oder
- die Übermittlung auf Grund besonderer Gesetze zulässig sein
oder
- die Übermittlung auf Grund eines übergesetzlichen Notstand zulässig sein

Quelle: eigene Darstellung.

Befunde einer schriftlichen Befragung unter potenziellen Nutzerinnen und Nutzern des Forschungsdatenzentrums der gesetzlichen Rentenversicherung*

Dr. Ralf K. Himmelreicher (VDR Frankfurt)

Zusammenfassung: Um Wissenschaftlern den Zugang zu Mikrodaten der gesetzlichen Rentenversicherung (gRV) zu ermöglichen, baut die gesetzliche Rentenversicherung, gefördert vom Bundesministerium für Forschung und Bildung (BMBF), ein Forschungsdatenzentrum (FDZ-RV) auf, das beim Verband Deutscher Rentenversicherungsträger (VDR) angesiedelt ist. Die prozessproduzierten Mikrodaten der Rentenversicherung sollen in Zukunft auf vier verschiedenen Zugangswegen erhältlich sein: Als anonymisierte (1) Public Use Files (PUF) und als tiefere Differenzierungen ermöglichende (2) Scientific Use Files (SUF). Daneben sollen (3) Gastwissenschaftler-Arbeitsplätze den kontrollierten Zugang zu den Datensätzen ermöglichen. Mittelfristig ist die Einrichtung eines Datenzugangs via (4) kontrollierter Datenfernverarbeitung angestrebt.

Um sich bei dem Aufbau des FDZ-RV am Datenbedarf der Wissenschaft orientieren zu können, wurde im Frühjahr 2004 eine Umfrage mit dem Titel „Daten der Rentenversicherung und Bedarf der Wissenschaft“ durchgeführt. Die wichtigsten Ergebnisse der Umfrage lassen sich wie folgt zusammenfassen: An der Umfrage haben 40 Forscher(-gruppen) aus 11 verschiedenen wissenschaftlichen Disziplinen, vor allem aus Ökonomie, Demographie, Soziologie und Medizin, teilgenommen. Sie forschen schwerpunktmäßig im Bereich der Alterssicherung sowie zu demographischen und gesundheitlichen Fragestellungen. Die Wissenschaftlerinnen und Wissenschaftler arbeiten in der Regel mit mehreren Personen in einem Forschungsprojekt zusammen; daneben gibt es auch einige Einzelforschungsvorhaben und individuelle Qualifikationsarbeiten, vor allem Promotionsvorhaben. Das Gros der Wissenschaftler möchte im Hinblick auf ihre Datenwünsche in Zukunft mit der Rentenstatistik arbeiten, vor allem mit dem Rentenzugang. Im Mittelfeld liegt die so genannte Versichertenstatistik und eine vergleichsweise geringere Nachfrage ist im Bereich der Rehabilitationsstatistik zu verzeichnen. So lange sich das FDZ-RV im Aufbau befindet und nicht alle Mikrodatensätze der Wissenschaft in anonymisierter Form zur Verfügung gestellt werden, können sich knapp 60% der befragten Forscher(-gruppen) einen temporären Aufenthalt als Gastwissenschaftler in den Räumen des VDR in Frankfurt bzw. ab Frühjahr 2005 in Berlin sowie in Würzburg vorstellen. Drei Viertel der befragten Wissenschaftler ist an absolut anonymisierten Public Use Files interessiert und möchte diese auch in der Lehre an Fachhochschulen und Universitäten einsetzen. STATA und SPSS sind die Statistikprogramme, mit denen die meisten Befragten ihre Auswertungen durchführen möchten. Die überwiegende Mehrheit der befragten Wissenschaftler zeigt großes Interesse am Aufbau des FDZ-RV und möchte ausdrücklich über weitere Aktivitäten informiert werden.

* Dieser Aufsatz erschien in leicht veränderter Fassung zuerst in Deutsche Rentenversicherung Heft 8/2004, 472-485. Siehe auch www.fdz-rv.de, Downloadbereich: "Umfrage Forschungsbedarf".

1. Einleitung

Um den Zugang empirisch forschender Wissenschaftlerinnen und Wissenschaftler zu den überwiegend prozessproduzierten Daten der gRV zu verbessern, hat die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) ausdrücklich die Einrichtung eines Forschungsdatenzentrums bei der gesetzlichen Rentenversicherung empfohlen (vgl. KVI 2001, 286). Dieser Empfehlung nachkommend wird aktuell bei der gesetzlichen Rentenversicherung als Pilotprojekt ein Forschungsdatenzentrum eingerichtet, gefördert vom Bundesministerium für Bildung und Forschung (BMBF) und angesiedelt beim VDR.

Um das Datenangebot¹ des FDZ-RV an den konkreten Bedarfen der Wissenschaft ausrichten zu können, wurde die Umfrage „Daten der Rentenversicherung und Bedarf der Wissenschaft“ durchgeführt. Vorrangiges Ziel der Umfrage war, den konkreten Datenbedarf potentieller Nutzerinnen und Nutzer von Mikrodaten der gesetzlichen Rentenversicherung zu ermitteln.² Damit in Zusammenhang stehen die Kommunikation mit der Wissenschaft und die konkreten Nutzungsbedingungen. Es wurde gefragt: Auf welchen Wegen wird der Zugang zu den Daten seitens der Wissenschaft gewünscht, welche Programme zur Analyse und damit welche Datenformate werden präferiert?

Der zweite Abschnitt beschäftigt sich mit den Ergebnissen der Umfrage, insbesondere mit präferierten Datensätzen und Zugangswegen. Im dritten Abschnitt werden schließlich die Befunde der Umfrage bezüglich des Daten- und Dienstleistungsangebotes des Forschungsdatenzentrums der gesetzlichen Rentenversicherung diskutiert.

2. Ergebnisse der Umfrage „Daten der Rentenversicherung und Bedarf der Wissenschaft“

Die Aufgabe der Umfrage besteht darin herauszufinden, welche Mikrodaten der Rentenversicherung auf welchen Wegen der Wissenschaft zur Verfügung gestellt werden sollen. Angesprochen wurde dabei nicht die Gesamtheit aller Wissenschaftlerinnen und Wissenschaftler, die Interesse an den Daten haben könnten, sondern nur jene, die bereits Kontakt zur Rentenversicherung aufgenommen haben (siehe unten). Die Umfrage „Daten der Rentenversicherung und Bedarf der Wissenschaft“ kann somit keine Repräsentativität beanspruchen. Das Formulieren von Aussagen über die Wissenschaft ist nicht angestrebt, sondern sie soll durch Erfragung von insbesondere Datenwünschen und Zugangswegen eine Orientierung beim Aufbau des Forschungsdatenzentrums der Rentenversicherung bieten.

¹ Zur Bedeutung der Daten der gesetzlichen Rentenversicherung für Sozialpolitik(-empfehlung) und wissenschaftliche Forschung siehe Rehfeld (2001/2004). Die Statistik der gesetzlichen Rentenversicherung ist auf Verbandsebene organisiert. Die rechtliche Basis der statistischen Berichterstattung ist in § 79 SGB IV, in der Verwaltungsvorschrift über die Statistik der gesetzlichen Rentenversicherung und in § 75 SGB X geregelt.

² Zur Ermittlung des Datenbedarfs, der Nutzungsmöglichkeiten sowie der Zugangswege wurde von den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder ebenfalls eine Befragung durchgeführt (siehe hierzu Zühlke u.a. 2003a/b/c).

Das FDZ-RV wird als Projekt im Rahmen des Forschungsnetzwerks Alterssicherung (FNA) beim VDR geführt (vgl. Ruland 2001, 129 ff.). Für die Umfrage „Daten der Rentenversicherung und Bedarf der Wissenschaft“ wurden zunächst Wissenschaftlerinnen und Wissenschaftler über das FNA angesprochen. Das FNA besteht seit dem Jahr 2001 und verfügt zur Vernetzung verschiedener Forschergruppen im Themenspektrum der Alterssicherung über eine Datenbank.³ Über diese Datenbank wird der so genannte FNA-Newsletter versendet, der von rund 1.400 Forscherinnen und Forschern empfangen wird. Dieser Newsletter diente als Medium, um auf unsere Umfrage aufmerksam zu machen. Der Fragebogen zur Umfrage stand zwischen Ende März und Mitte April 2004 im Internetportal unseres Forschungsdatenzentrums (www.fdz-rv.de) zum Download zur Verfügung und konnte online, per Fax oder als Brief an das FDZ-RV gesendet werden; Einsendeschluss war der 16. April 2004.

Der FNA-Newsletter dürfte überwiegend Personen erreichen, die Interesse an aktuellen Aktivitäten des Forschungsnetzwerks Alterssicherung haben. In einem engeren Kontakt mit dem FNA, z.B. über Forschungsk Kooperationen, Stipendien, Austausch von Forschungspapieren etc. stehen etwa 175 Forscher(-gruppen). Von diesem Personenkreis haben 40 Forscher(-gruppen) ihr ausdrückliches Interesse an den Daten der Rentenversicherung bekundet. Dies bedeutet überschlagsweise eine Rücklaufquote von knapp 23%.

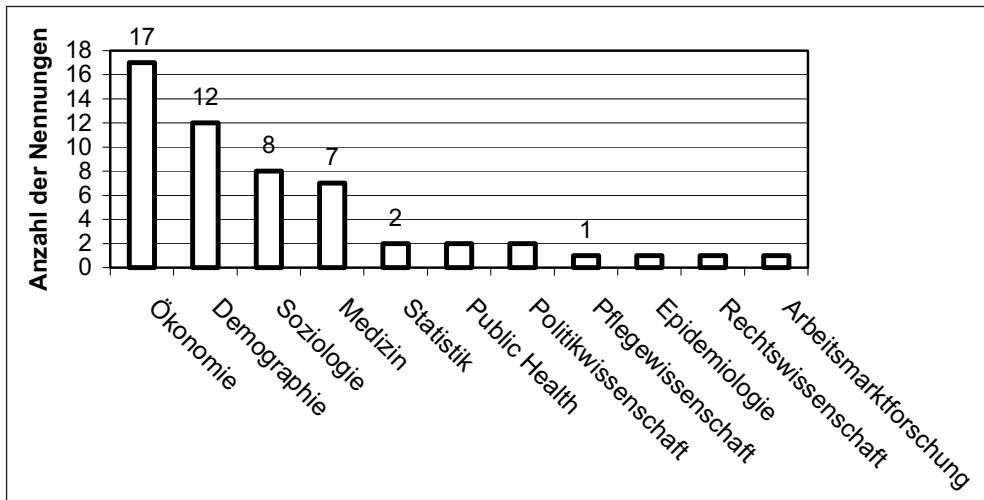
Von den 40 Forscher(-gruppen) haben immerhin 38 durch ihre Unterschrift der EDV-mäßigen Erfassung ihrer persönlichen Angaben zugestimmt; diese Personen möchten weiterhin mit dem Forschungsdatenzentrum in Kontakt bleiben und über Neuigkeiten informiert werden. Zur Kontaktpflege wurde ein Verteiler aufgebaut, über den Interessierte in Zukunft Informationen erhalten. Im Folgenden werden einige wichtige Befunde aus der Umfrage dargestellt; sie sind nach der Reihenfolge im Fragebogen geordnet.⁴ Der Aufbau des Forschungsdatenzentrums soll sich zum Einen an den Befunden der Umfrage, zum Anderen an weiteren Verfahren zum Austausch von Erfahrungen und Wünschen, wie Nutzerkonferenzen, Arbeitstagen, Berichten aus anderen Forschungsdatenzentren etc. orientieren.

2.1 Disziplinarität

Zugang zu den Individualdaten der gesetzlichen Rentenversicherung möchten Forschende haben, die sich im Hinblick auf ihre Schwerpunktdisziplin zehn verschiedenen wissenschaftlichen Fachrichtungen zuordnen. Eine vergleichsweise hohe Datennachfrage ist bei den Fachbereichen mit starker quantitativer Ausrichtung, wie Ökonomie, Demographie, Soziologie und Medizin festzustellen. Wesentlich geringer fällt die Nachfrage nach Datensätzen der Rentenversicherung der weniger stark quantitativ forschenden Disziplinen Politik- und Rechtswissenschaft sowie den eher neuen bzw. kleineren Fachbereichen Statistik, Public Health, Pflegewissenschaften und Epidemiologie aus.

³ Als offenste Form der Kooperation dient eine allen Interessierten über das Internet zugängliche Datenbank mit Informationen über Wissenschaftler, Publikationen und Projekte im Bereich der Alterssicherungsforschung. Die Datenbank spiegelt zunächst die Informationen zu den Fachgebieten und Forschungsschwerpunkten der Wissenschaftler aus dem Bereich Alterssicherung wider, die mit Hilfe eines Fragebogens erhoben werden (siehe www.vdr-fna.de).

⁴ Der Fragebogen befindet sich im Anhang.

Abbildung 1: Schwerpunktdisziplinen der Wissenschaftler

Quelle: VDR-Umfrage „Daten der Rentenversicherung und Bedarf der Wissenschaft“, März/April 2004.

Im Hinblick auf die Schwerpunktdisziplin, in der die Forscher tätig sind, konnte ferner festgestellt werden, dass Interdisziplinarität weit verbreitet ist. Den 39 Forschern, die Angaben zur Schwerpunktdisziplin gemacht haben, sind 54 Disziplinen zuzuordnen, wobei 28 Forschende ein Kernfach und 11 zwei und mehr Schwerpunktdisziplinen genannt haben: 8 Forscher fühlen sich in zwei Disziplinen zuhause, zwei in drei Fächern und ein Forscher hat vier verschiedene Disziplinen angegeben. Die beobachtbare Interdisziplinarität steht sicherlich in Zusammenhang mit dem Forschungsgegenstand Alterssicherung, der innerhalb enger Fachgrenzen kaum adäquat behandelt werden kann (vgl. Ruland 2001). Forschende mit mehr als einer Schwerpunktdisziplin haben am häufigsten die Fächerkombination Ökonomie/Soziologie genannt.

Differenziert man die häufiger genannten Disziplinen Ökonomie, Demographie, Soziologie und Medizin nach Forschungsinstituten, dann zeigt sich, dass im Fach Ökonomie immerhin 6 der 17 Nennungen aus zwei Forschungsinstituten stammen, im Fach Soziologie und Medizin gibt es lediglich ein Institut mit 2 Nennungen, während sich in der Disziplin Demographie 9 der 12 Nennungen auf ein großes Forschungsinstitut konzentrieren.

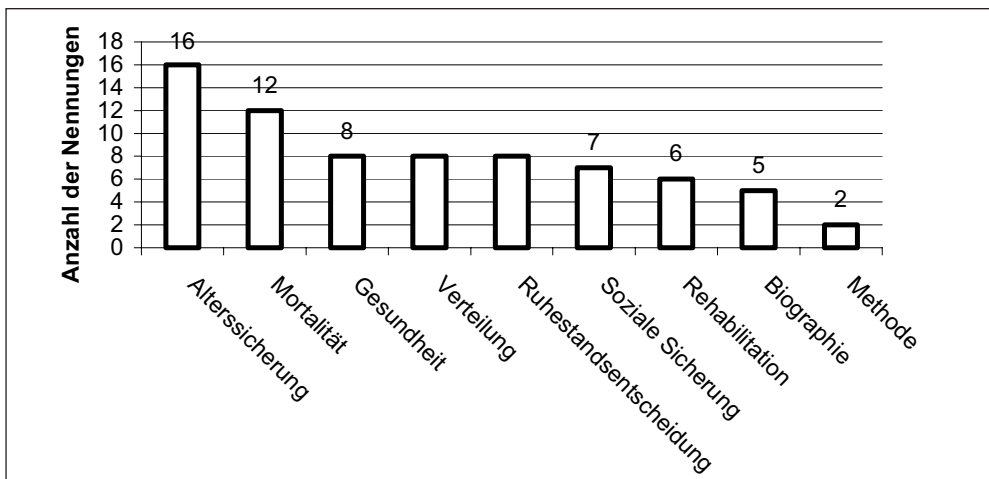
Insgesamt haben 38 Wissenschaftlerinnen und Wissenschaftler aus 23 verschiedenen Instituten an der Umfrage „Daten der Rentenversicherung und Bedarf der Wissenschaft“ teilgenommen. Die in verschiedenen Instituten forschenden Personen sind in zahlreichen wissenschaftlichen Einzeldisziplinen verankert, insbesondere in den Fächern Ökonomie, Demographie, Soziologie und Medizin, jedoch auch in einigen anderen Bereichen der empirischen Wirtschafts- und Sozialwissenschaften sowie in der Statistik und der Rechtswissenschaft. Auf Grund des hinsichtlich ihrer Einbettung in verschiedene Forschungsrichtungen breit gestreuten Interesses potenzieller Nutzer von Mikrodaten der gRV sind gleichsam verschiedene Fragestellungen zu erwarten. Insofern sind miteinander konkurrierende Ergebnisse zu erwarten, die zu Gewinn bringenden Erkenntnissen führen dürften. Diese Erwartung stützt sich nicht allein auf die verschiedenen Forschungsrichtungen, sondern auch auf die von den Wissenschaftlern genannten Forschungsfelder.

2.2 Forschungsfelder

Von den 39 Forschenden, die Angaben auf die Frage „Für welches Forschungsfeld der sozialen Sicherheit möchten Sie die Daten des FDZ-RV nutzen?“ gemacht haben, möchten in Zukunft 16 Personen im Bereich der Alterssicherung forschen. Anders formuliert: gut zwei Fünftel der potenziell RV-Daten Nutzenden möchten die Alterssicherung, den Kernbereich der gesetzlichen Rentenversicherung, erforschen.

An zweiter Stelle der am häufigsten genannten Forschungsfelder stehen Analysen zur Mortalität. Die Forschungsfelder Gesundheit, (Einkommens-)Verteilung und Ruhestandsentscheidung folgen auf Platz 3 mit je 8 Nennungen. Die Forschung im Bereich Rehabilitation und Bio-

Abbildung 2: Forschungsfelder der Wissenschaftler



Quelle: VDR-Umfrage „Daten der Rentenversicherung und Bedarf der Wissenschaft“, März/April 2004.

graphie wurde von lediglich 6 bzw. 5 Forschern genannt. Grundlegende Methodenforschung mit den Daten der gRV wollen nur wenige Wissenschaftler durchführen.

Die Forscher, die zukünftig Analysen auf Basis von Mikrodaten der gRV durchführen möchten, haben ihren Forschungsschwerpunkt zur Hälfte im Bereich der gesetzlichen Rentenversicherung, ein Drittel erforscht das Zusammenwirken von gesetzlicher, privater und betrieblicher Altersvorsorge. Einige wenige arbeiten schwerpunktmäßig im Bereich der betrieblichen Altersvorsorge und im Bereich des Zusammenwirkens von gesetzlicher und privater Altersvorsorge.

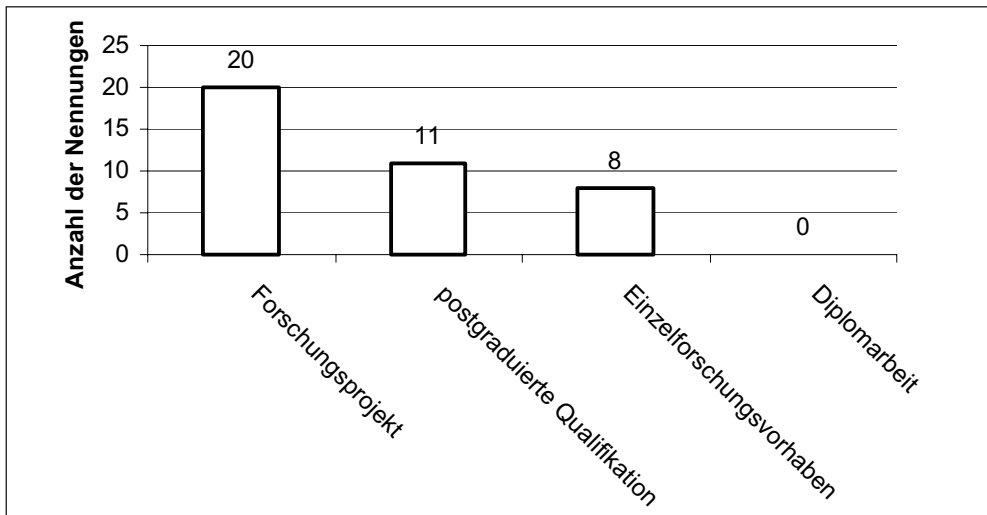
Neben den standardisiert erhobenen Forschungsfeldern im Bereich der sozialen Sicherheit wurden als weitere Forschungsgegenstände vor allem internationale Vergleiche, Fragen der Finanzierung der gesetzlichen Rentenversicherung und das Verhältnis zwischen Rentenzahlung und späterer Rentenhöhe unter Berücksichtigung der Laufzeit, also Analysen zur Rendite, angegeben.

Weitere Forschungsvorhaben außerhalb des Forschungsfeldes Soziale Sicherheit hat jeder vierte Wissenschaftler, der an unserer Umfrage teilgenommen hat. Diese liegen vor allem im Bereich der Portfolioanalyse und des Sparverhaltens, in der medizinischen Forschung und im Bereich Rehabilitation.

2.3 Forschungsprojekt

Um einordnen zu können, in welchem institutionellen Kontext in Zukunft mit den Daten der gRV gearbeitet werden soll, wurden Fragen zum institutionellen Rahmen des Forschungsvorhabens, zur Finanzierung und zur Laufzeit bereits bestehender Projekte gestellt.

Abbildung 3: Institutioneller Rahmen der Forschungsvorhaben



Quelle: VDR-Umfrage „Daten der Rentenversicherung und Bedarf der Wissenschaft“, März/April 2004.

Die schwerpunktmäßige Nutzung der prozessproduzierten Mikrodaten der gRV dürfte in Zukunft im Rahmen von Forschungsprojekten liegen. Daneben werden derzeit postgraduierte Qualifikationsvorhaben wie insbesondere Promotionen, aber auch Habilitationen durchgeführt. Ferner ist in Zukunft damit zu rechnen, dass die Daten in Einzelforschungsvorhaben zum Einsatz kommen. Bisher ist der Einsatz von Individualdaten der gRV im Rahmen von Diplomarbeiten noch nicht vorgesehen. Im Rahmen des Aufbaus des FDZ-RV wird angestrebt, so genannte stark anonymisierte Campus-Files für Studierende zur Verfügung zu stellen. Diese könnten als Strukturdatensatz fungieren, um vor der Diplomarbeit das Arbeiten mit den komplexen Daten der gRV zu üben. Insofern dürften die Daten der gRV in Zukunft bei Diplomarbeiten mit quantitativem Zuschnitt durchaus eine Rolle spielen.

Auf die Fragestellung, ob es sich bei den Forschungsvorhaben um Drittmittelprojekte handelt, antwortete ein Drittel der Befragten mit Ja. Fragt man danach, welche Institution die Drittmittelprojekte (mit-)finanziert, wurden die Deutsche Forschungsgemeinschaft (DFG), das Bundesministerium für Bildung und Forschung (BMBF), die Europäische Gemeinschaft (EU), die Hans-Böckler-Stiftung (HBS) sowie das Forschungsnetzwerk Alterssicherung (FNA) des VDR genannt. Diese Befunde verweisen darauf, dass überwiegend öffentlich-rechtliche Institutionen eine Kofinanzierung der Forschungsprojekte übernehmen.

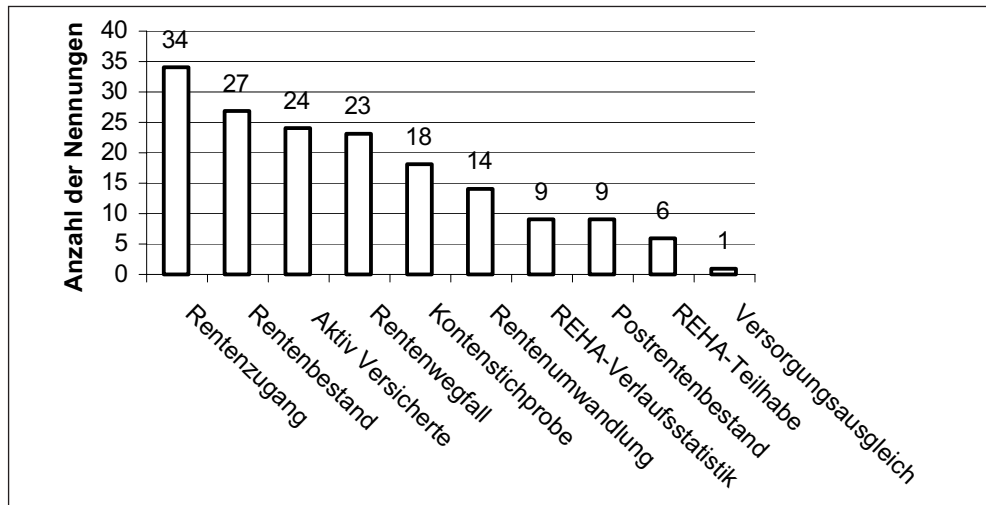
Bei jenen Befragten, die eine Kofinanzierung ihrer Forschung erhalten, läuft diese in einigen Fällen im laufenden Jahr aus oder reicht wiederum bei einigen Fällen bis maximal Ende 2005; nur bei wenigen Forscher(-gruppen) ist die Finanzierung bis zum Jahr 2006 gesichert. Diese Befunde verweisen auf zeitlich befristete projektförmige Forschungsvorhaben bei den meisten Wissenschaftlern. Die Zufriedenstellung solcher User des FDZ-RV kann nicht im Rahmen von mittel- bis langfristiger Zusammenarbeit erprobt werden, sondern bedingt eine ebenso intensive wie kurzfristige Kooperation. Das Ziel einer möglichst kurzfristigen Bereitstellung entsprechender Daten erfordert - unter Berücksichtigung forschungsspezifischer Fragestellungen und deren Auswirkungen auf die jeweiligen Datenwünsche - kurze Bearbeitungszeiten, wobei stets die Bestimmungen des Datenschutzes zu beachten sind.

2.4 Nachfrage der Wissenschaft nach Mikrodaten der Rentenversicherung

Im Fragebogen wurde unter anderem danach gefragt, welche Datensätze aus den Bereichen Versichertenstatistik, Rehabilitationsstatistik und Rentenstatistik (Mehrfachnennungen sind zulässig) nachgefragt werden. Diese Frage ist vor allem deshalb wichtig, weil sie geeignet ist herauszufinden, welche Datensätze von den Forschenden präferiert werden. Die am häufigsten genannten Datensätze sollen jene sein, die im Forschungsdatenzentrum als erste zur Verfügung gestellt werden.

Ordnet man die einzelnen Datensätze den drei großen Themenfeldern „Versichertenstatistik“, „Rehabilitationsstatistik“ und „Rentenstatistik“ zu, dann zeigt sich, dass das Gros der Wissenschaftler mit der Rentenstatistik, insbesondere mit dem Rentenzugang arbeiten möchte. Im Mittelfeld liegt die Versichertenstatistik; eine vergleichsweise geringe Nachfrage ist aus dem Bereich der Rehabilitationswissenschaft zu verzeichnen. Dieser Befund dürfte darauf zurückzuführen sein, dass Rehabilitationswissenschaft und Rentenversicherung seit vielen Jahren kooperieren.

Neben den standardisiert abgefragten prozessproduzierten Datensätzen der Rentenversicherung wurde Interesse an den REHA-Entlassungsberichten und an den Daten der AVID geäußert. Der Sonderwunsch nach REHA-Entlassungsberichten ist notiert, dürfte jedoch während des Aufbaus des FDZ-RV in der zweijährigen Projektphase mangels größerer Nachfrage nicht zu jenen gehören, die prioritär zur Verfügung gestellt werden. Hinsichtlich der Nachfrage zu Daten der AVID ist folgendes festzuhalten: Die AVID-Daten werden nicht weitergegeben, weil eine Weitergabe am Datenschutz scheitert. Diese Aussage gilt sowohl für die AVID '96 als auch für die AVID '02. In der Datenschutzerklärung bzw. in der von den Versicherten unterzeichneten Einwilligungserklärung ist zur Kontenklärung und Zusammenführung von Konten-, Befragungs-

Abbildung 4: Welche Datensätze werden präferiert?

Legende: Kontenstichprobe – Versicherungskontenstichprobe;
 Rentenumwandlung – Statistik über Änderungen der Leistungsart, des Teilrentenanteils oder von Teilrente in Vollrente;
 REHA-Verlaufstatistik – REHA-Renten-Verlaufstatistik;
 Postrentenbestand – Postrentenbestandsstatistik;
 REHA-Teilhabe – Statistik über abgeschlossene Leistungen zur Teilhabe;
 Versorgungsausgleich – Statistik über den Versorgungsausgleich.

Quelle: VDR-Umfrage „Daten der Rentenversicherung und Bedarf der Wissenschaft“, März/April 2004.

und Ehepartnerdaten festgelegt worden, dass ausschließlich die Träger der Rentenversicherung und das BMGS diese Daten auswerten dürfen.⁵

Für Interessenten an spezifischen Mikrodatenauswertungen zu Fragestellungen, die mit den mehreren tausend Tabellen in 7 Tabellenbänden nicht beantwortet werden können, gibt es nur einen Weg, um solche Informationen zu erhalten: Forschende können sich an das BMGS oder den VDR wenden und nach Genehmigung ihres Vorhabens die Mikrodaten kostenpflichtig auswerten lassen. Solche Anträge wurden in der Vergangenheit unbürokratisch und schnell bearbeitet.⁶

2.5 Zugangswege und Software

Überraschend hoch war die Bereitschaft der Forschenden so genannte Gastwissenschaftler-Arbeitsplätze beim VDR in Würzburg oder Frankfurt (ab Frühjahr 2005 in Berlin) zu besuchen. Knapp 60% der Wissenschaftler können sich eine zeitweilige Tätigkeit in den Räumen des VDR

⁵ Die Studie „Altersvorsorge in Deutschland (AVID)“ hat ein eigenes Internetportal. Unter www.altersvorsorge-in-deutschland.de sind Informationen zur AVID-Untersuchung des Verbandes Deutscher Rentenversicherungsträger und des Bundesministeriums für Gesundheit und Soziale Sicherung abgelegt. Dort sind zahlreiche Informationen, Hinweise auf Publikationen sowie die Personen zu finden, über die ein Kontakt hergestellt werden kann.

⁶ Diese Erfahrung hat der Autor als Forscher am Zentrum für Sozialpolitik der Universität Bremen selbst gemacht (vgl. Himmelreicher und Viebrok 2003 sowie Viebrok et al. 2004).

vorstellen, insbesondere in der Aufbauphase des FDZ-RV, in der zunächst noch keine faktisch oder total anonymisierten Datensätze vorliegen. Neben der persönlichen Betreuung ist ein besonderer Vorteil von Gastwissenschaftler-Arbeitsplätzen, dass das leistungsfähige Rechenzentrum des VDR genutzt werden kann. Auswertungen von Datensätzen mit einer Fallzahl jenseits der Millionengrenze erfordern selbst mit modernsten PCs einige Geduld, vor allem bei multivariaten Analysen.

Neben den Gastwissenschaftler-Arbeitsplätzen haben knapp drei Viertel der befragten Wissenschaftler angegeben, dass sie mit absolut anonymisierten Public Use Files, insbesondere zur Verwendung in der Lehre an den Universitäten, arbeiten möchten.

Hinsichtlich der Datenformate und Auswertungssoftware auf den Gastwissenschaftler-Arbeitsplätzen haben 11 Forschende angegeben, dass sie ihre Auswertungen mit STATA durchführen möchten, 8 Personen präferieren SPSS, mit SAS möchte lediglich ein Forscher arbeiten. Weitere Auswertungsprogramme, die jeweils nur einfach genannt wurden, sind Oracle und Matlab.

Für den Aufbau des FDZ-RV hinsichtlich der Zugangswege bedeuten die Befunde, dass man in der Gründungsphase mit Gastwissenschaftler-Arbeitsplätzen einen Großteil der Interessenten mit Daten der gRV bedienen kann. Nach der Gründungsphase sollte verstärkt daran gearbeitet werden, Scientific und Public Use Files zur Verfügung zu stellen. Hierzu ist die Beschreibung der Erhebungsmethode sowie die Dokumentation der Daten zu verbessern.

Hinsichtlich der Statistik-Software wird vor allem STATA aber auch SPSS präferiert. Wahrscheinlich werden die Datensätze des FDZ-RV im CSV-Format zur Verfügung gestellt, so dass sie problemlos in gängige Software-Pakete eingelesen werden können. Die Gastwissenschaftler-Arbeitsplätze sollten sowohl mit STATA als auch mit SPSS ausgestattet sein, im Idealfall auch mit SAS.⁷

3. Zusammenfassung und Ausblick

Das sich im Aufbau befindende Forschungsdatenzentrum der gesetzlichen Rentenversicherung (FDZ-RV) verfolgt das Ziel, der Wissenschaft prozessproduzierte Mikrodaten der Rentenversicherung zur Verfügung zu stellen. Vorgesehen ist, die Daten in Zukunft auf vier verschiedenen Wegen anzubieten: als absolut anonymisierte Public Use Files (PUF) und als tiefere Differenzierungen ermöglichende Scientific Use Files (SUF). Vor allem während der Aufbauphase sollen Gastwissenschaftler-Arbeitsplätze den kontrollierten Zugang zu den Datensätzen ermöglichen. Die Option der kontrollierten Datenfernverarbeitung ist im Hinblick auf inhaltliche und technische Voraussetzungen sowie auf den Datenschutz als besonders anspruchsvoll einzuschätzen und kann deshalb erst mittelfristig realisiert werden.

Da die gesetzliche Rentenversicherung über ein umfassendes wie langjähriges statistisches Berichtswesen mit mehreren verschiedenen Mikrodatensätzen verfügt, die beim Aufbau des FDZ nicht auf einmal angeboten werden können, diene die Umfrage „Daten der Rentenversi-

⁷ Eine Ausstattung der Gastwissenschaftler-Arbeitsplätze im FDZ-RV mit STATA, SPSS und SAS hätte den Vorteil, dass im Hinblick auf die Software-Bestückung vergleichbare Arbeitsbedingungen bestehen würden, wie in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (siehe den Beitrag von Zühlke et al. (2003) bzw. im vorliegenden Band).

cherung und Bedarf der Wissenschaft“ dazu herauszufinden, welche Datensätze auf welchen Zugangswegen von der Wissenschaft gewünscht werden. Ein weiteres Ziel der Umfrage war, einen Adressverteiler aufzubauen, um über die weitere Entwicklung des FDZ-RV, Tagungen etc. informieren zu können.

Die wichtigsten Befunde der Umfrage sowie die sich für den Aufbau des FDZ-RV ergebenden Implikationen werden hier zusammengefasst:

- An der Umfrage haben 40 Forscher(-gruppen) aus 11 verschiedenen wissenschaftlichen Disziplinen, vor allem aus Ökonomie, Demographie, Soziologie und Medizin, teilgenommen. Dies bedeutet für das FDZ-RV, dass eine interdisziplinäre Ausrichtung erstrebenswert ist. Bezüglich z.B. der Datendokumentation bzw. sonstiger Beschreibungen sollte auf spezifische Fachtermini zugunsten allgemein verständlicher Interpretationen verzichtet werden.
- Die Forschungsfelder liegen schwerpunktmäßig im Bereich der Alterssicherung, Demographie und Gesundheit. Die Wissenschaftler planen mithin Forschungsgegenstände zu analysieren, die im Zentrum des Interesses der gRV liegen, so dass diesbezüglich ein hoher Erkenntniszuwachs zu erwarten ist.
- Die Wissenschaftlerinnen und Wissenschaftler arbeiten in der Regel in einem Forschungsprojekt zusammen mit mehreren Personen; daneben gibt es auch einige Einzelforschungsvorhaben und individuelle Qualifikationsarbeiten, vor allem Promotionsvorhaben. Insbesondere im Hinblick auf Forschergruppen ist daran zu denken, dass die Datennutzungsverträge so ausgestaltet sein sollten, dass sie von definierten Personengruppen und nicht nur von Einzelpersonen genutzt werden können.
- Das Gros der Wissenschaftler möchte im Hinblick auf ihre Datenwünsche in Zukunft mit der Rentenversicherungsstatistik arbeiten, vor allem mit dem Rentenzugang. Im Mittelfeld liegt die so genannte Versichertenstatistik und eine vergleichsweise geringere Nachfrage ist im Bereich der Rehabilitationsstatistik zu verzeichnen. Dieser Befund bedeutet für das FDZ-RV, dass zunächst die Rentenstatistik inklusive Dokumentation aufbereitet und für die Wissenschaft zur Verfügung gestellt wird, dann folgt die Versicherten- und schließlich die Rehabilitationsstatistik.
- Solange sich das FDZ-RV im Aufbau befindet und nicht alle Mikrodatensätze der Wissenschaft in anonymisierter Form zur Verfügung gestellt werden, können sich knapp 60% der Befragten einen temporären Aufenthalt als Gastwissenschaftler in den Räumen des VDR in Frankfurt bzw. ab Frühjahr 2005 in Berlin sowie in Würzburg vorstellen. An diesem Befund kann man ablesen, dass die hohe Bereitschaft der Wissenschaftler in den Räumen des VDR arbeiten zu wollen, schon bald zur Eröffnung der ersten Gastwissenschaftler-Arbeitsplätze führen wird. Es ist jedoch zu erwarten, dass verfügbare Scientific Use Files einen Teil der Gastwissenschaftler-Arbeitsplätze verdrängen wird. Diese Einschätzung dürfte insbesondere dann zutreffen, wenn es sich um mit Mikrodaten der Rentenversicherung erfahrene Forscher handelt.
- Drei Viertel der befragten Wissenschaftler ist an absolut anonymisierten Public Use Files interessiert und möchte diese auch in der Lehre an Fachhochschulen und Universitäten einsetzen. D.h. auf längere Sicht ist zu erwarten, dass Mikrodaten der gRV in Diplom- und Examensarbeiten verwendet werden. Hieraus könnte ein Bindungseffekt entstehen, wodurch Forschung auf Basis von Daten der gRV verbreitet wird.

- STATA und SPSS sind die Statistikprogramme, mit denen die meisten Befragten ihre Auswertungen durchführen möchten.

Schließlich ist festzuhalten, dass die überwiegende Mehrheit der befragten Wissenschaftler großes Interesse am Aufbau des FDZ-RV hat und ausdrücklich über weitere Aktivitäten informiert werden möchte. Insofern besteht bereits heute seitens der Wissenschaft der Wunsch nach einer Dauereinrichtung des FDZ-RV. Ferner dürfte es seinen Beitrag dazu leisten, die Erforschung des Systems der Alterssicherung in Deutschland voranzutreiben, letztlich zum Nutzen der Versicherten und Rentner.⁸

Literatur

- Himmelreicher, R.* (2004): Daten der Rentenversicherung und Bedarf der Wissenschaft. Befunde einer Befragung unter potenziellen Nutzern des Forschungsdatenzentrums der gesetzlichen Rentenversicherung. *Deutsche Rentenversicherung*, 8, 472-485.
- Himmelreicher, R. & Viebrok, H.* (2003): Die „Riester-Rente“ und einige Folgen für Alterseinkünfte. Befunde einer Simulationsstudie auf der Grundlage typischer Erwerbs- und Familienbiographien. *ZeS-Arbeitspapier 4/2003*, Zentrum für Sozialpolitik der Universität Bremen, Bremen: Eigenverlag.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (2001): *Wege zu einer besseren informationellen Infrastruktur*. Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. Nomos: Baden-Baden.
- Krupp, H.-J.* (2004): Mikroanalysen und amtliche Statistik - gestern, heute, morgen. In: Merz, J., Zwick, M. u.a.: *MIKAS - Mikroanalysen und amtliche Statistik*. Hrsg. vom Statistischen Bundesamt: Statistik und Wissenschaft, 27-35, Bd. 1. Wiesbaden.
- Rehfeld, U.* (2001): Die Statistiken der gesetzlichen Rentenversicherung. *Deutsche Rentenversicherung*, 3-4, 169-188.
- Rehfeld, U.* (2004): Datenangebot und Informationsbedarf im Bereich der Alterssicherung – Forschungsdatenzentrum der Rentenversicherung im Aufbau. *Deutsche Rentenversicherung*, 1-2, 63-75.
- Ruland, F.* (2001): Forschungsnetzwerk Alterssicherung – verstärkte Kooperation zwischen Rentenversicherung und Wissenschaft im Bereich der Alterssicherung. *Deutsche Rentenversicherung*, 3-4, 129 –136.
- Viebrok, H., Himmelreicher, R. & Schmähl, W.* (2004): *Private Altersvorsorge statt gesetzlicher Rente: Wer gewinnt, wer verliert?* Herausgegeben von W. Schmähl. Beiträge zur Sozial- und Verteilungspolitik, Band 3, Lit Verlag: Münster.
- Zühlke, S. & Hetke, U.* (2003): Datenbedarf und Datenzugang: Ergebnisse der ersten Nutzerbefragung des Forschungsdatenzentrums der Statistischen Landesämter. In: *Allgemeines Statistisches Archiv*, Band 87, Heidelberg: Physika, 321 – 334.

⁸ Auf den Aspekt, „(...) dass die unabhängige wissenschaftliche Analyse dazu beitragen kann, die Effizienz und Zielgerechtigkeit des politischen Handelns in einem derartigen Bereich zu beurteilen.“ hat Krupp (2004, 31 f.) hingewiesen.

Zühlke, S. & Zwick, M. (2003): Datenbedarf und Datennutzungsmöglichkeiten. Das Dienstleistungsangebot der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder. In: Statistisches Bundesamt (Hg.): *Wirtschaft und Statistik, Sonderausgabe ISI-Weltkongress 2003*, Stuttgart: Metzler-Poeschel, 71 – 73.

Zühlke, S. & Zwick, M., Scharnhorst, S. & Wende, T. (2003): Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder. *Wirtschaft und Statistik*, 10, 906 – 911.

Dr. Ralf K. Himmelreicher ist Elektriker und Dipl.-Soziologe, studierte Soziologie, VWL und Statistik an der FU in Berlin und war wissenschaftlicher Mitarbeiter an der Arbeitnehmerkammer Bremen. Promotion (Dr. rer. pol.) am Institut für Konjunktur- und Strukturforchung (IKSF) der Universität Bremen bei Prof. Thomas von der Vring, Wissenschaftlicher Mitarbeiter am Zentrum für Sozialpolitik an der Universität Bremen bei Prof. Dr. Winfried Schmähl, seit 2003 ist er Referent beim VDR in Frankfurt (Abteilung Volkswirtschaft und Statistik) und seit Anfang des Jahres 2004 schwerpunktmäßig mit dem Aufbau des FDZ-RV befasst.

Anlage

Umfrage des Forschungsdatenzentrums der gesetzlichen Rentenversicherung

FDZ-RV zum Thema

Daten der Rentenversicherung und Bedarf der Wissenschaft

Um Wissenschaftlerinnen und Wissenschaftlern den Zugang zu den überwiegend prozessproduzierten Statistikdaten der Gesetzlichen Rentenversicherung zu erleichtern und deren Auswertungsmöglichkeiten generell zu verbessern, hat die Kommission zur Verbesserung der wissenschaftlichen Infrastruktur zwischen Wissenschaft und Statistik (KVI) die Einrichtung eines Forschungsdatenzentrums bei der gesetzlichen Rentenversicherung empfohlen. Die gesetzliche Rentenversicherung richtet dieser Empfehlung entsprechend zurzeit ein Forschungsdatenzentrum ein, das beim VDR angesiedelt sein wird. Das Dienstleistungsangebot dieses Forschungsdatenzentrums soll sich in seinem Aufbau und seiner Entwicklung am gegenwärtigen und zukünftigen wissenschaftlichen Bedarf orientieren. In einem ersten Schritt ist die Frage zu klären, zu welchem Zweck welche Daten rund um den Forschungsgegenstand „Alterssicherung“ benötigt werden. Zur Ermittlung des Datenbedarfs wird diese Umfrage durchgeführt, an deren reger Teilnahme wir Sie - im Sinne der Wissenschaft – nachdrücklich bitten möchten. Die weitere Arbeitsplanung bezüglich des Forschungsdatenzentrums wird an den Ergebnissen dieser Umfrage ausgerichtet sein.

Im Rahmen dieser Umfrage soll außerdem ein Verteiler aufgebaut werden, um interessierte Personen in Zukunft über Neuigkeiten aus dem FDZ-RV informieren zu können. Sofern Sie weitere Informationen zum Forschungsdatenzentrum der Gesetzlichen Rentenversicherung erhalten möchten, bitten wir Sie, die Einverständniserklärung am Ende des Fragebogens mit „Ja“ zu beantworten, Ihre Adresse anzugeben, die wir zu diesem Zwecke speichern werden und den Fragebogen zu unterschreiben.

**Zu weiteren Informationen zum FDZ-RV siehe
(auch zum Download unter www.fdz-rv.de):**

Rehfeld, Uwe G.: „Datenangebot und Informationsbedarf im Bereich der Alterssicherung - Forschungsdatenzentrum der Rentenversicherung im Aufbau“, in: DRV 1-2/2004, S. 63 – 75.

Rehfeld, Uwe G.: „Die Statistiken der gesetzlichen Rentenversicherung“, in: DRV 3-4/2001, S.169 – 188.

1. In welcher Disziplin möchten Sie mit den Daten des FDZ-RV arbeiten?

(bitte geben Sie Ihre Schwerpunktdisziplin an)

- | | | |
|---|--------------------------------------|---|
| <input type="checkbox"/> Ökonomie | <input type="checkbox"/> Demographie | <input type="checkbox"/> Pflegewissenschaft |
| <input type="checkbox"/> Rechtswissenschaften | <input type="checkbox"/> Statistik | <input type="checkbox"/> sonstige Disziplin |
| <input type="checkbox"/> Soziologie | <input type="checkbox"/> Medizin | _____ |

2.1 Für welches Forschungsfeld der sozialen Sicherheit möchten Sie die Daten des FDZ-RV nutzen?

(bitte geben Sie Ihr derzeit wichtigstes Forschungsfeld an)

- | | |
|---|--|
| <input type="checkbox"/> Soziale Sicherung | <input type="checkbox"/> Methodenforschung |
| <input type="checkbox"/> Verteilungsanalysen | <input type="checkbox"/> Gesundheit |
| <input type="checkbox"/> Alterssicherung, darunter vor allem: | <input type="checkbox"/> Mortalität |
| <input type="checkbox"/> Gesetzliche Rentenversicherung | <input type="checkbox"/> Rehabilitation |
| <input type="checkbox"/> Betriebliche Altersvorsorge | <input type="checkbox"/> Ruhestandsentscheidung |
| <input type="checkbox"/> Private Altersvorsorge | <input type="checkbox"/> weitere Forschungsfelder: |
| <input type="checkbox"/> Lebensverlaufsforschung | |

2.2 Haben Sie weitere Forschungsvorhaben außerhalb der Thematik der sozialen Sicherung?

- ja nein

wenn ja, welche: _____

3.1 In welchem Rahmen ist Ihr Forschungsvorhaben angesiedelt?

- Habilitation Forschungsprojekt
- Promotion - Ihre Funktion: _____
- Diplomarbeit - Projektleitung: _____
- Einzelforschungsvorhaben - in Kooperation mit wem: _____

3.2 Handelt es sich um ein Drittmittelprojekt?

- ja nein

wenn ja, von wem wird dieses (mit-)finanziert: _____

3.3 Läuft dieses Projekt bereits?

- ja nein

wenn ja, bis wann läuft es: _____

4. An welchen Statistikdaten der gesetzlichen Rentenversicherung besteht vor allem Interesse? (Mehrfachnennungen sind zulässig)

Versichertenstatistik

- Statistik der aktiv Versicherten
- Versicherungskontenstichprobe
- Statistik über den Versorgungsausgleich

Rehabilitationsstatistik

- Statistik über abgeschlossene Leistungen zur Teilhabe
- Reha-Renten-Verlaufsstatistik

Rentenstatistik

- Rentenzugangsstatisik
- Rentenwegfallsstatistik
- Statistik über Änderungen der Leistungsart, des Teilrentenanteils oder von Teilrente in Vollrente
- Rentenbestandsstatistik
- Postrentenbestandsstatistik

Sonstige Statistik (ggf. benennen)

5.1 Sofern (noch) keine faktisch oder absolut anonymisierten Daten vorliegen: Besteht grundsätzlich Interesse an der Nutzung eines Gastwissenschaftler-Arbeitsplatzes in Frankfurt am Main oder Würzburg?

- ja (weiter mit Frage 5.2) nein (weiter mit Frage 6)

5.2 Bei Interesse an der Nutzung von Gastwissenschaftler-Arbeitsplätzen: Welche Software bevorzugen Sie?

- SPSS
- SAS
- STATA
- sonstige: _____

6. Besteht über Ihr Forschungsvorhaben hinaus ein Interesse an Public Use Files (absolut anonymisierte Mikrodaten), insbesondere zur Verwendung in der Lehre?

ja nein

7. Haben Sie Anregungen und Hinweise, die Sie gerne im Rahmen der Service-Angebote des FDZ-RV berücksichtigt haben möchten?

Einverständniserklärung:

- Nein, ich möchte routinemäßig keine weitere Information erhalten
- Ja, ich bin an den RV-Daten interessiert und mit der EDV-mäßigen Registrierung meiner persönlichen Angaben einverstanden.

Name und akademischer Grad: _____

Funktion: _____

Institution: _____

Straße, Hausnummer: _____

PLZ, Ort: _____

Tel. / FAX: _____

E-mail / WWW: _____

Datum und Unterschrift: _____

Vielen Dank für Ihre Teilnahme an der Umfrage. Bitte senden Sie den ausgefüllten Fragebogen bis zum **16.04.2004** an **Verband Deutscher Rentenversicherungsträger, Ref. 5.2.2 – FDZ-RV - Eysseneckstr. 55, 60322 Frankfurt am Main** oder faxen sie an **069 / 1522 320**. Weitere Informationen zur Statistik und den Daten der gesetzlichen Rentenversicherung finden Sie unter www.vdr.de/statistik bzw. unter www.fdz-rv.de. Bitte informieren Sie ihre Kolleginnen und Kollegen.

1.2 Forschungsdatenzentren der Statistischen Ämter und der BA

Die Perspektive der Datenproduzenten wird im folgenden Abschnitt um die Erfahrungen anderer öffentlicher Forschungsdatenzentren erweitert. Die Einrichtung von Forschungsdatenzentren war neben der Schaffung von Servicezentren und der Gründung eines Rates für Sozial- und Wirtschaftsdaten einer der zentralen Vorschläge der Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI). In einem ersten Schritt wurde der Aufbau von Forschungsdatenzentren beim Statistischen Bundesamt und bei den Landesämtern, der Bundesagentur für Arbeit und bei der Rentenversicherung empfohlen (KVI 2001, 285 ff.).

Die Gründung des FDZ des Statistischen Bundesamtes erfolgte noch im Jahr 2001 und im März 2002 wurde ein erstes Forschungsdatenzentrum der Statistischen Landesämter installiert. Die amtliche Statistik hat somit einen ersten Zugang der Wissenschaft zu Mikrodaten geschaffen.

Das Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (IAB) wurde im April 2004 eingerichtet, wobei der Zugang zu Daten der BA z.B. im Rahmen des Betriebspanels bereits seit 1999 über die Schalterstelle des IAB möglich ist. Wie bei der Rentenversicherung handelt es sich bei den Daten der BA überwiegend um Sozialdaten.

Die Bereitstellung von Mikrodaten im Rahmen der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder sowie der Bundesagentur für Arbeit erfolgt unter ähnlichen Bedingungen und liefert hilfreiche Hinweise für einen erfolgreichen Aufbau des FDZ-RV. Zur Koordination der Maßnahmen und zum gegenseitigen Erfahrungsaustausch der öffentlichen Forschungsdatenzentren wurde zudem der Arbeitskreis öffentliche Forschungsdatenzentren (AKöffFDZ) eingerichtet.

Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder*

*Dr. Sylvia Zühlke, Markus Zwick, Sebastian Scharnhorst und Thomas Wende
(FDZ der Statistischen Ämter des Bundes und der Länder)*

Einleitung

Die Komplexität des wirtschaftlichen und sozialen Wandels sowie die Fortschritte in Wissenschaft und Informationstechnik haben den Datenbedarf moderner Gesellschaften grundlegend verändert. Die zur Analyse und Gestaltung moderner Gesellschaften erforderlichen Daten müssen insbesondere Informationen über Teilgruppen der Gesellschaft liefern sowie die Analysen des wirtschaftlichen und sozialen Wandels auf der Basis von Längsschnittdaten erlauben. Aufgrund des geänderten Informationsbedarfs ist die Veröffentlichung von Ergebnissen in Form von Tabellen heute nicht mehr ausreichend. Vielmehr entspricht es den methodischen und inhaltlichen Erfordernissen, statistische Daten entsprechend dem wissenschaftlichen Datenbedarf bereitzustellen. Dazu gehört es, Zugangsmöglichkeiten zu anonymisierten und nicht anonymisierten Mikrodaten zu schaffen, mit denen differenziertere Analysen durchgeführt werden können.

Vor diesem Hintergrund wurde in den letzten Jahren in Deutschland intensiv über den Zugang der Wissenschaft zu den Mikrodaten der amtlichen Statistik diskutiert. Durch die „Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik“ (KVI) wurden dafür im Auftrag des Bundesministeriums für Bildung und Forschung verschiedene Vorschläge zur Verbesserung der Wechselwirkungen zwischen Wissenschaft und Statistik erarbeitet (vgl. KVI 2001). Eine der zentralen institutionellen Empfehlungen dieser Kommission bestand in der möglichst raschen Einrichtung von Forschungsdatenzentren bei den Datenproduzenten.

Diese Empfehlung wurde von der amtlichen Statistik aufgegriffen. Im Jahr 2001 hat das Statistische Bundesamt ein solches Forschungsdatenzentrum etabliert. Ein weiteres Forschungsdatenzentrum der Statistischen Landesämter wurde im März 2002 eingerichtet, als eine gemeinsame Einrichtung aller Statistischen Landesämter mit 16 regionalen Standorten. Mit diesen Forschungsdatenzentren intensiviert die deutsche amtliche Statistik ihre bisherigen Bemühungen, die Mikrodaten der amtlichen Statistik für wissenschaftliche Analysen zugänglich zu machen.

Ziel dieses Aufsatzes ist es, die neuen Formen und Wege der Nutzung amtlicher Mikrodaten vorzustellen, die sich aus der Einrichtung der Forschungsdatenzentren ergeben haben. Hierfür folgt zunächst ein Überblick über die Entwicklung der Rahmenbedingungen für die Nutzung von Mikrodaten der amtlichen Statistik in Deutschland. Anschließend werden die Ziele und Aufgaben der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder beschrieben sowie die unterschiedlichen Möglichkeiten der Datennutzung präsentiert, die mit den Forschungsdatenzentren angeboten werden.

* Dieser Beitrag erschien unter gleichem Titel u.a. in *Wirtschaft und Statistik*, 10/2003, 906-911.

Die Nutzung von Mikrodaten der amtlichen Statistik in Deutschland

Die Nutzung von Mikrodaten der amtlichen Statistik durch die Wissenschaft wurde in Deutschland sehr stark von der Entwicklung des Bundesstatistikgesetzes (BStatG) beeinflusst, das 1951 erstmals verabschiedet und in den Jahren 1981 und 1987 novelliert wurde.

Die Übermittlung von Mikrodaten an Dritte wurde bei Entstehung des Bundesstatistikgesetzes kaum diskutiert, so dass in der Fassung von 1951 hierzu noch keine explizite Regelung enthalten war. Aufgrund der fehlenden Möglichkeiten umfangreiche Mikrodaten zu verarbeiten, war die Nachfrage nach Mikrodaten in den 50er, 60er und frühen 70er Jahren relativ gering. Sofern die Übermittlung von Einzelangaben nicht explizit in anderen Gesetzen geregelt war, wurden für einzelne Projekte formal anonymisierte Mikrodaten zur Verfügung gestellt. So wurden amtliche Mikrodaten seitens der Wissenschaft erstmals in dem Projekt „Sozialpolitisches Entscheidungs- und Indikatorensystem für die Bundesrepublik Deutschland (SPES)“ analysiert, das 1972 bis 1978 durchgeführt wurde (Krupp 1973). Diesem Projekt stellte die amtliche Statistik formal anonymisierte Mikrodaten des Mikrozensus und der Einkommens- und Verbrauchsstichprobe zur Verfügung. Für das gleiche Projekt wurde auch eine Stichprobe der Volkszählung 1970 (VZ70) auf der Grundlage des Volkszählungsgesetzes bereit gestellt.

Als die Nachfrage nach Mikrodaten mit der Fortentwicklung der Informationstechnik schnell anstieg, wurde die Übermittlung von Mikrodaten an Dritte erstmals explizit mit der Verabschiedung des Bundesdatenschutzgesetzes im Jahr 1977 und mit der Novellierung des BStatG im Jahr 1981 geregelt. Mit der Einführung der sogenannten Weiterleitungsvorschrift des § 11 BStatG 1981 wurde die Absicht verfolgt, eine bessere Versorgung der Wissenschaft mit Mikrodaten zu ermöglichen. Diese Vorschrift sah die Weitergabe von Mikrodaten in absolut anonymisierter Form an die Nutzer vor. Mit der Anwendung dieser Möglichkeit innerhalb konkreter Projekte zeigte sich allerdings, dass die Anforderung an ein solchermaßen absolut anonymisiertes Datenmaterial so restriktiv war, dass es in der Folgezeit nur noch in Ausnahmefällen zu Datenlieferungen kam. Aufgrund der gesetzlichen Änderung des BStatG konnten für nachfolgende Projekte wie „Vergleichende Analysen der Sozialstruktur mit Massendaten“ (VASMA) oder dem Sonderforschungsbereich 3 „Mikroanalytische Grundlagen der Gesellschaftspolitik“ der Deutschen Forschungsgemeinschaft nur noch absolut anonymisierte Mikrodaten zu relativ hohen Kosten zur Verfügung gestellt werden.¹

Das Volkszählungsurteil von 1983 und die damit einhergehende Diskussion zeigten, dass sowohl die informationelle Selbstbestimmung als auch die Wissenschaftsfreiheit, beide in Artikel 5 des Grundgesetzes verankert, als wichtige Grundrechte gleichrangig zu behandeln sind. Somit war der Gesetzgeber aufgerufen, für einen adäquaten Datenzugang Sorge zu tragen. Mit der weiteren Modifikation des BStatG im Jahre 1987 und der Einführung einer faktischen Anonymität für die Weitergabe von Mikrodaten an die Wissenschaft kam er diesen Anforderungen nach. Die faktische Anonymität erlaubte es nun, im Rahmen des sogenannten Wissenschaftsprivilegs, unter bestimmten Voraussetzungen Mikrodaten an die Wissenschaft zu liefern, die ein Restrisiko der Deanonymisierung aufwiesen. In der Folgezeit konkretisierten verschie-

¹ Zu den Projekten vgl. Hauser (1994). Die Ergebnisse des Projektes VASMA sind unter <http://www.gesis.org/Dauerbeobachtung/Mikrodaten/Daten/brd/literatur.pdf> dokumentiert.

dene Projekte die Ausgestaltung faktisch anonymisierter Mikrodatensätze. Insbesondere die Ergebnisse des Projektes „Die faktische Anonymisierung von Mikrodaten“ erlaubten es, ab Mitte der neunziger Jahre erste standardisierte faktisch anonymisierte Mikrodatensätze für den Bereich der Haushalts- und Personenerhebungen zur Verfügung zu stellen (vgl. Müller 1991).

Während hierdurch und durch weitere Aktivitäten außerhalb der amtlichen Statistik² insbesondere für sozialwissenschaftliche Fragestellungen eine neue Datenbasis geschaffen wurde, konnten im Bereich der Wirtschaftswissenschaften aufgrund der größeren Schwierigkeiten bei der Anonymisierung von Betriebs- und Unternehmensdaten vergleichbare Fortschritte zunächst nicht erzielt werden. Der fehlende Zugang insbesondere zu den Unternehmens- und Betriebsdaten wurde seitens der Wissenschaft in dem Memorandum „Erfolgsbedingungen empirischer Wirtschaftsforschung und empirisch gestützter wirtschafts- und sozialpolitischer Beratung“ aufgegriffen (Hauser et al. 1998). Hierin wurde gefordert, den Zugang zu schwer anonymisierbaren Mikrodaten innerhalb der Räumlichkeiten der Datenproduzenten zu ermöglichen. Die Diskussion über den Mikrodatenzugang der Wissenschaft gewann mit dem Memorandum sowie dem Symposium „Kooperation zwischen Wissenschaft und amtlicher Statistik - Praxis und Perspektiven“ (Müller et al. 1999) im Jahre 1999 eine neue Dynamik, die auch von der Politik aufgegriffen wurde.

In der Folge erarbeitete die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) eine Reihe von Empfehlungen zur Verbesserung der Kooperation von Wissenschaft und Statistik (KVI 2001). Diese reichen von der Mitwirkung der Datennutzer bei der Aufstellung der Erhebungs- und Aufbereitungsprogramme über die Perspektiven einer modernen Aus- und Weiterbildung in der Statistik bis zu den verschiedenen Möglichkeiten für den Zugang der Wissenschaft zu den Mikrodaten der öffentlichen Datenproduzenten. Eine wesentliche institutionelle Forderung bezieht sich auf die Einrichtung von Forschungsdatenzentren bei den Datenproduzenten und auf die Einrichtung von Servicezentren. In der Zwischenzeit werden die Empfehlungen der Kommission zur Verbesserung der informationellen Infrastruktur von dem Gründungsausschuss des Rates für Sozial- und Wirtschaftsdaten umgesetzt, so dass eine Reihe von Forschungsdatenzentren und Servicezentren bereits ihre Arbeit aufnehmen konnten (vgl. Lüttinger et al. 2003).

Hierzu zählen auch die beiden Forschungsdatenzentren der amtlichen Statistik. Zwar sind diese zwei voneinander unabhängige Einrichtungen, gleichwohl stimmen sie ihre Arbeit eng untereinander ab, um der Wissenschaft ein gemeinsames Angebot für einen verbesserten Datenzugang zu unterbreiten. Den Empfehlungen der KVI folgend, nach denen die Einrichtung einer leistungsfähigen Dateninfrastruktur eine Aufgabe der Forschungsförderung ist, haben sowohl das Statistische Bundesamt als auch die Statistischen Landesämter einen Förderantrag beim Bundesministerium für Bildung und Forschung gestellt. Mit Hilfe dieser Mittel soll ein umfangreiches Dienstleistungs- und Datenangebot geschaffen werden.

² Außerhalb der amtlichen Statistik haben sich insbesondere mit dem „Sozio-oekonomischen Panel“ (SOEP) und der „Allgemeine Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) Erhebungen etabliert, die regelmäßig durchgeführt werden und für wissenschaftliche Analysezwecke zur Verfügung stehen.

Ziele und Aufgaben der Forschungsdatenzentren

Das wesentliche Ziel der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder besteht darin, den Zugang der Wissenschaft zu den Mikrodaten der amtlichen Statistik durch die Einrichtung unterschiedlicher Datennutzungswege zu erleichtern. Eine wesentliche Voraussetzung für die Realisierung dieses Ziels besteht in der grundsätzlichen Verbesserung der Dateninfrastruktur durch den Aufbau einer fachlich zentralisierten Datenhaltung für ausgewählte Statistiken und durch die Einrichtung eines Metadateninformationssystems.

In Deutschland wird der überwiegende Teil der Statistikerhebungen dezentral in den Statistischen Landesämtern durchgeführt, so dass dort über 90 Prozent aller Mikrodaten der amtlichen Statistik erhoben, aufbereitet und gespeichert werden. Da sich wissenschaftliche Analysen in der Regel jedoch auf mehrere Bundesländer oder das gesamte Bundesgebiet beziehen, planen die Statistischen Landesämter den Aufbau einer fachlich zentralisierten Datenhaltung. Dadurch wird es möglich sein, die Mikrodaten der amtlichen Statistik länderübergreifend an allen regionalen Standorten der Forschungsdatenzentren zu nutzen.

Damit wissenschaftliche Nutzer die Mikrodaten der amtlichen Statistik analysieren und interpretieren können, benötigen sie zudem umfassende Informationen über die Datensätze sowie über die Erhebung, Aufbereitung und Qualität der Daten. Daher werden die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder ein internetbasiertes Metadateninformationssystem entwickeln, in dem sich die Nutzer über die Erhebungen der amtlichen Statistik informieren können.

Datennutzungswege

Um der Wissenschaft den Zugang zu dem gesamten Informationspotential der amtlichen Statistik zu eröffnen, richten die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder unterschiedliche Zugangswege zur Nutzung ihrer Mikrodaten ein. Damit werden den Nutzern zusätzliche und weitergehende Möglichkeiten zur Analysen von Mikrodaten der amtlichen Statistik eröffnet, als dies bislang möglich war.

Ansatzpunkt dieser zusätzlichen Nutzungswege ist es, die Vermeidung einer Reidentifikation der Auskunftgebenden nicht mehr allein durch Veränderungen im Datenmaterial sicherzustellen, sondern auch durch die Regulierung des Datenzugriffs. Die einzelnen Nutzungswege resultieren daher aus verschiedenen Kombinationen von Datenanonymisierung und Zugriffsregulierung.

a) *Absolut anonymisierte Mikrodatensätze*

Absolut anonymisierte Daten werden durch Aggregation oder durch die Entfernung einzelner Merkmale soweit verändert, dass eine Identifizierung der Auskunftgebenden nach menschlichem Ermessen unmöglich gemacht wird. Die amtliche Statistik bietet absolut anonymisierte Mikrodaten in Form sogenannter Public Use Files (PUF) an. Diese können allen interessierten Personen zur Verfügung gestellt werden.

Bislang wurden solche Datensätze für die Sozialhilfestatistik sowie für die Zeitbudgeterhebung erstellt. Auch in diesem Segment werden die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder zukünftig verstärkt an einer Ausweitung ihres Angebotes arbeiten. Eine weitere Hauptzielrichtung der Public Use Files liegt im Bereich der Hochschullehre. Die Forschungsdatenzentren entwickeln derzeit sogenannte Campus-Files, die an Hochschulen zu Lehrzwecken eingesetzt werden können. Es ist geplant, einen ersten Campus-File des Mikrozensus 1998 noch im laufenden Semester zur Verfügung zu stellen. Dieser Datensatz soll Studentinnen und Studenten schon frühzeitig die Möglichkeit geben, die Besonderheiten der Analyse von Mikrodaten kennen zu lernen.

b) Faktisch anonymisierte Mikrodatensätze

Eine absolute Datenanonymisierung birgt den Nachteil, dass damit auch ein erheblicher Teil der statistischen Information verloren geht. Dagegen werden Mikrodaten als faktisch anonymisiert bezeichnet, wenn die Deanonymisierung zwar nicht gänzlich ausgeschlossen werden kann, die Angaben jedoch nur mit einem unverhältnismäßig hohen Aufwand an Zeit, Kosten und Arbeitskraft dem jeweiligen Merkmalsträger zugeordnet werden können.³ Die Hauptzielrichtung der faktischen Anonymisierung besteht folglich darin, durch behutsame Informationsreduktion und Informationsveränderungen die Zuordnungsmöglichkeiten von Merkmalsausprägungen zu den entsprechenden Merkmalsträgern zu verringern, dabei jedoch den statistischen Informationsgehalt zu schonen. Hierfür müssen für jede einzelne Erhebung der Aufwand und der Nutzen einer Deanonymisierung analysiert werden. Dabei können unterschiedliche Anonymisierungsverfahren zur Anwendung kommen.⁴ Nach Maßgabe des Bundesstatistikgesetzes können faktisch anonymisierte Daten allerdings nur wissenschaftlichen Einrichtungen und nur zur Durchführung wissenschaftlicher Vorhaben zugänglich gemacht werden.

Faktische Anonymität resultiert also nicht allein aus dem realen Informationsgehalt der Daten, sondern auch aus den bestehenden Möglichkeiten zur Deanonymisierung. Wann ein Mikrodatensatz als faktisch anonym bezeichnet werden kann, hängt daher insbesondere davon ab, unter welchen Rahmenbedingungen die Daten verarbeitet werden. So ist von entscheidender Bedeutung, welches Zusatzwissen vorliegt und wo die Datennutzung stattfindet. Abhängig davon, ob die Mikrodaten extern oder in den Statistischen Ämtern genutzt werden, kann die faktische Anonymität mit mehr oder minder starken Informationsveränderungen erreicht werden.

Ein immer wieder deutlich vorgetragener Wunsch der Wissenschaft ist es, Mikrodaten in anonymisierter Form am eigenen Arbeitsplatz zu nutzen. Die faktische Anonymisierung der Mikrodaten ermöglicht diese Übermittlung nicht vollständig anonymisierter Mikrodaten zur externen (Off-Site-) Nutzung in wissenschaftlichen Einrichtungen. Da jedoch allein durch die Herausgabe dieser Daten ein höheres Deanonymisierungsrisiko besteht als bei der Nutzung in einem Statistischen Amt, ist die Datenanonymisierung relativ ausgeprägt. Die für diese Nutzungsform erzeugten Datensätze werden als Scientific Use Files (SUF) bezeichnet.

Die amtliche Statistik bietet im Bereich der personenbezogenen Daten mit dem Mikrozensus, der Einkommens- und Verbrauchsstichprobe und der Zeitbudgeterhebung bereits ein breites Datenspektrum als Scientific Use File an. Die Forschungsdatenzentren des Bundes und der

³ Diese Regelung geht zurück auf den § 16 Abs. 6 BstatG.

⁴ Ein Überblick über Anonymisierungsmethoden ist zu finden in Köhler (1999).

Länder sind bestrebt, dieses Angebot sukzessive zu erweitern. So laufen zur Zeit Projekte zur Anonymisierung der Lohn- und Einkommensteuerstatistik und der Diagnosestatistik. Im kommenden Jahr soll darüber hinaus mit der Gehalts- und Lohnstrukturerhebung erstmals auch ein sogenannter Employer-Employee-Datensatz zur standardmäßigen Anonymisierung herangezogen werden. Mit dem Projekt „Faktische Anonymisierung wirtschaftsstatistischer Einzeldaten“ versucht die amtliche Statistik, gemeinsam mit den wissenschaftlichen Nutzern auch in diesem Bereich anonymisierte Standardfiles zu entwickeln (Sturm 2002).

c) Projektbezogene faktische Anonymisierung zur On-Site-Nutzung

Bei geringer Nachfrage nach einzelnen Statistiken und bei schwer zu anonymisierenden Mikrodaten wird es in vielen Fällen nicht sinnvoll sein, in einem aufwändigen Verfahren standardisierte Scientific Use Files zu erstellen. Hier ist eine projektbezogene Anonymisierung der Daten sinnvoller. Dies hat den praktischen Vorteil, dass dabei nicht die gesamte Statistik anonymisiert wird, sondern lediglich die aus ihr benötigten Merkmale.

Die projektbezogene Anonymisierung erzeugt ebenfalls eine faktische Datenanonymität. Diese Daten können jedoch nur in den Räumlichkeiten der Forschungsdatenzentren des Bundes und der Länder an sogenannten Gastwissenschaftler-Arbeitsplätzen ausgewertet werden. Da die Mikrodaten dabei die Räumlichkeiten der amtlichen Statistik nicht verlassen und die Nutzer nicht beliebig über die Daten verfügen und sie kaum mit Zusatzinformationen kombinieren können, hat diese Nutzungsform einen weiteren wesentlichen Vorteil. Weil die faktische Anonymität hier bereits bei wesentlich geringeren Veränderungen im Datenmaterial erreicht wird als bei der Erstellung von Scientific Use Files zur Off-Site Nutzung, verbleiben folglich auch mehr Informationen im Datenmaterial.

Um die Attraktivität der On-Site-Nutzung von Mikrodaten zusätzlich zu erhöhen und die regionale Erreichbarkeit dieser Nutzungsform zu gewährleisten, richten die Forschungsdatenzentren derzeit neben den bereits bestehenden Gastwissenschaftler-Arbeitsplätzen in Berlin, Bonn und Wiesbaden weitere Arbeitsplätze in allen Statistischen Landesämtern ein. An diesen können die Daten mit den gängigen statistischen Analyseprogrammen ausgewertet werden.

d) Nutzung amtlicher Mikrodaten durch kontrollierte Datenfernverarbeitung

Die Nutzung von geschützten Datenbeständen via Datenfernverarbeitung ist eine relativ neue Entwicklung, deren Bedeutung in der Zukunft zunehmen wird.⁵ Mittels dieses Verfahrens ist es der Wissenschaft möglich, das Informationspotential von nicht oder nur formal anonymisiertem Mikrodatenmaterial zu nutzen, ohne jedoch selbst direkten Zugriff auf diese Daten zu haben. Die Wissenschaftler entwickeln dabei Auswertungsprogramme (Syntax-Skripte), die dann von den Mitarbeitern der Forschungsdatenzentren an den Originaldaten angewendet werden. Diese Datendienstleistung wird zur Zeit in den Programmen SPSS, SAS und STATA angeboten. Die Datenfernverarbeitung ist im Gegensatz zu den Scientific Use Files nicht auf einen bestimmten Personenkreis beschränkt und bietet daher auch für ausländische Wissenschaftler und für den nichtwissenschaftlichen Interessentenkreis eine Möglichkeit, Mikrodaten der amtlichen Statistik zu nutzen.

⁵ Erste Erfahrungen wurden hier im Bereich der Steuerstatistiken gesammelt (vgl. Zwick 1999).

Zur praktikablen Anwendung der Datenfernverarbeitung stellen die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder Datenstrukturfiles zur Verfügung, die es den Nutzern ermöglichen, ihre Analyseprogramme auf die Struktur der Originaldaten abzustimmen. Diese Datenstrukturfiles geben die Datenstruktur des originären Datensatzes wieder, ohne inhaltliche Informationen zu transportieren. Das Material ist im Aufbau also in der Merkmalsstruktur, Anzahl der Satzstellen, Datensatzlänge identisch zum Originalmaterial. Durch ein technisches Verfahren sind die Daten aber so verfälscht, dass nur noch synthetische inhaltsleere Datensätze zur Verfügung stehen. Zurzeit stehen Datenstrukturfiles für die Mikrodaten des Mikrozensus und der Lohn- und Einkommensteuerstatistik zur Verfügung.

Derzeit ist die Datenfernverarbeitung ein relativ arbeitsaufwändiges Verfahren, da die Programmsyntax vorab auf Deanonymisierungsstrategien und der Datenoutput anschließend auf Geheimhaltungsfälle überprüft werden müssen. Bislang müssen diese Schritte manuell durchgeführt werden. Zwar liegen mittlerweile auch erste automatisierte Verfahren für diese Prüfungen vor, jedoch ist auch mit diesen Lösungsansätzen derzeit noch keine vollständige Automatisierung der Datenfernverarbeitung möglich.

Ein wesentliches Ziel der Forschungsdatenzentren besteht daher darin, Verfahren zu entwickeln, die eine stärkere Automatisierung der Datenfernverarbeitung erlauben. Ansatzpunkte hierfür bieten insbesondere die LIS/LES-Datenbank sowie das in Dänemark praktizierte Verfahren einer On-Line-Datenverarbeitung.⁶ Die LIS/LES-Datenbank erlaubt den direkten Datenzugriff auf die Mikrodaten der Luxembourg Income Study bzw. Luxembourg Employment Study über das Internet (vgl. Smeeding et al. 2002). Über die Angabe eines projektbezogenen Passworts können durch die Übermittlung von SAS-, SPSS- oder STATA-Files automatisch Datenauswertungen gestartet werden. LIS/LES ist so ausgestaltet, dass auch eine begrenzte Syntax- und Ergebnisüberprüfung erfolgt.

Das dänische Modell sieht hingegen vor, dass Wissenschaftler direkt von ihrem Arbeitsplatz aus auf einen Server zugreifen können, der von dem Statistischen Amt für Forschungszwecke eingerichtet und gepflegt wird.⁷ Die Wissenschaftler erhalten hierzu Zugriff auf ein Verzeichnis, in dem nur die Daten abgelegt sind, die sie für ihr Forschungsvorhaben benötigen. Die Daten können von dem Wissenschaftler in ein für sie angelegtes Arbeitsverzeichnis kopiert werden; von dort werden auch die Analysen durchgeführt. Die Ergebnisse werden automatisch per E-Mail versendet. Der gesamte Prozess der Datenverarbeitung kann von dem Statistischen Amt jederzeit daraufhin überprüft werden, ob die vertraglich vereinbarten Regeln eingehalten werden.

Die hier im Einzelnen dargestellten Datennutzungswege der Forschungsdatenzentren können auch miteinander kombiniert werden. Unabhängig davon, welche Möglichkeit des Datenzugangs gewählt wird, unterliegt die Datenbereitstellung allerdings einer Zweckbindung. Dies bedeutet, dass die Mikrodaten nicht für allgemeine Forschungszwecke zur Verfügung gestellt werden können, sondern nur für ein definiertes und zeitlich begrenztes Forschungsvorhaben.

⁶ Darüber hinaus bietet das Programm μ -ARGUS die Möglichkeit, Ergebnisse, die in Form von Tabellen vorliegen, maschinell geheim zu halten. Das Programm wurde vom Statistischen Zentralamt der Niederlande für Eurostat entwickelt und für die tabellen-übergreifende Geheimhaltung vom Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen erweitert. Es wird zur Zeit in Deutschland ausführlich getestet.

⁷ Das Modell wird in dem Bericht „Access to Microdata in the Nordic Countries“ beschrieben, der vom Statistischen Amt Schwedens 2003 herausgegeben wurde.

Sonderaufbereitungen

Neben den dargestellten Datennutzungswegen besteht für alle Nutzergruppen weiterhin die Möglichkeit, Sonderaufbereitungen in Auftrag zu geben. Hierbei handelt es sich um Datenauswertungen, die auf den speziellen Informationsbedarf eines bestimmten Nutzers zugeschnitten sind und nicht oder nur unzureichend mit den sonstigen Datennutzungswegen umgesetzt werden können. Im Gegensatz zur kontrollierten Datenfernverarbeitung werden die Auswertungsprogramme nicht von den Nutzern, sondern von der amtlichen Statistik erstellt. Hierzu wird der Datenbedarf im Dialog zwischen amtlicher Statistik und Nutzer soweit konkretisiert, dass ein Auswertungsprogramm erstellt werden kann. Anschließend wird das nicht oder nur formal anonymisierte Datenmaterial mit Hilfe des entwickelten Programms ausgewertet. Die Ergebnisse werden nach einer Geheimhaltungsprüfung übermittelt. Der Nutzer hat demnach keinen direkten Kontakt mit den nicht oder nur formal anonymisierten Mikrodaten.

Datenbedarf der Wissenschaft und Nutzungspräferenzen

Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder streben an, ihr Dienstleistungsangebot dem wissenschaftlichen Bedarf entsprechend weiterzuentwickeln und befragen daher regelmäßig potenzielle Nutzer nach ihren Präferenzen. Um den konkreten Datenbedarf der Wissenschaft bei der Entwicklung des Dienstleistungsangebotes zu berücksichtigen, hat das Forschungsdatenzentrum der Statistischen Landesämter im Sommer 2002 eine Nutzerbefragung durchgeführt (Zühlke & Hetke 2003). Ziel der Umfrage war es, die potenziellen Nutzerinnen und Nutzer des künftigen Datenangebotes der Forschungsdatenzentren anzusprechen und ihren konkreten Datenbedarf zu ermitteln. Dabei konnten sich die Befragten auch zu den unterschiedlichen Nutzungsmöglichkeiten, den von ihnen eingesetzten Analyseprogrammen und ihrem Interesse an geplanten Veranstaltungen der beiden Forschungsdatenzentren äußern.

Von den siebenhundert befragten Wissenschaftlern gaben immerhin knapp 600 an, dass sie im Rahmen ihrer wissenschaftlichen Tätigkeit Mikrodaten nutzen oder zukünftig benötigen werden. Insgesamt weisen die Ergebnisse der Nutzerbefragung somit auf ein großes Interesse der Wissenschaft an der Nutzung von Mikrodaten der amtlichen Statistik hin. Da der in der Umfrage angegebene Datenbedarf fachlich sehr breit gestreut ist, werden sich die Forschungsdatenzentren nicht auf die Bereitstellung einzelner weniger Statistiken konzentrieren können, sondern müssen mittelfristig ein breites Datenangebot bereithalten.

Bezüglich der angebotenen Datennutzungswege weist die Befragung auf eine eindeutige Präferenz für die Nutzung von faktisch und total anonymisierten Daten am eigenen Arbeitsplatz hin, während die Arbeit an den Gastwissenschaftler-Arbeitsplätzen und die kontrollierte Datenfernverarbeitung zum Befragungszeitpunkt kaum auf das Interesse der Wissenschaft stießen. Es wird jedoch nicht möglich sein, alle relevanten Erhebungen der amtlichen Statistik als anonymisierte Datensätze anzubieten. Die Forschungsdatenzentren planen daher, die Attraktivität der alternativen Nutzungswege durch deren nutzungsgerechte Ausgestaltung deutlich zu steigern. Insbesondere die Einrichtung von Gastwissenschaftler-Arbeitsplätzen in allen regionalen Standorten der Forschungsdatenzentren wird die regionale Erreichbarkeit des Dienstleistungsangebot der amtlichen Statistik deutlich verbessern, so dass der Zugang zu den Mikrodaten wesentlich erleichtert wird.

Ausblick

Durch die beschriebene Weiterentwicklung der Dateninfrastruktur und die Einrichtung unterschiedlicher Datennutzungswege wird die Mikrodatenbasis für wissenschaftliche Analyse-zwecke entscheidend verbessert werden. Die Diskussion um den Zugang der Wissenschaft zu dem Informationspotenzial der öffentlichen Datenproduzenten ist jedoch mit der Einrichtung von Forschungsdatenzentren bei den unterschiedlichen öffentlichen Datenproduzenten nicht abgeschlossen. Zukünftige Fragen des Datenzugangs betreffen neben grundsätzlichen Regelungen insbesondere auch die Bereitstellung international vergleichbarer Mikrodaten.

Die Diskussion um eine grundsätzliche Regelung des Datenzugangs wird vom Gründungsausschuss des Rates für Sozial- und Wirtschaftsdaten unter dem Begriff des „Forschungsdatengeheimnisses“ geführt. Der Vorschlag des Gründungsausschusses sieht vor, dass Wissenschaftlerinnen und Wissenschaftler den Mitarbeiterinnen und Mitarbeitern der Datenproduzenten gleich gestellt werden und somit weitergehende Rechte zur Nutzung der Mikrodaten als bisher erhalten sollen. Gleichzeitig sollen ein Zeugnisverweigerungsrecht sowie ein Beschlagnahmeverbot den unbefugten Zugriff auf Mikrodaten, die bei den Wissenschaftlern gespeichert sind, verhindern.

Ein wesentliches Problem, Mikrodaten für die international vergleichende wissenschaftliche Forschung zu nutzen, besteht darin, dass in den einzelnen Ländern der Mikrodatenzugang sehr unterschiedlich geregelt ist.⁸ Dadurch ist die Beschaffung verschiedener nationaler Mikrodaten mit einem recht hohen Aufwand verbunden. Erste Ansätze, den Datenzugang zu den Erhebungen der Europäischen Union zu vereinheitlichen, wurden mit der Verordnung Nr. 322/97 und darauf aufbauend 831/2002 für Gemeinschaftsstatistiken unternommen.⁹ Mit diesen Verordnungen wird die länderübergreifende Bereitstellung von Mikrodaten der Arbeitskräftestichprobe, des Europäischen Haushaltspanels, der Erhebung über die berufliche Weiterbildung und der Innovationserhebung der Gemeinschaft angestrebt. Die Umsetzung dieser Verordnungen wird jedoch zurzeit dadurch erschwert, dass die dort vorgesehenen Regelungen teilweise den nationalen Regelungen in den Ländern der Europäischen Union widersprechen.

Aufgrund der dargestellten Entwicklungen ist davon auszugehen, dass sich die Datenzugangsmöglichkeiten der Wissenschaft auch zukünftig weiter entwickeln werden. Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder werden sich an diesem Prozess durch eigene Ausgestaltungsvorschläge weiterhin intensiv beteiligen.

⁸ Eine Darstellung der verschiedenen nationalen Ansätze zum Datenzugang findet sich in den Tagungsbeiträgen zum internationalen „Workshop on Microdata“ vom 21./22. August 2003 unter <http://www.micro2122.scb.se>.

⁹ Verordnung (EG) Nr. 831/2002 der Kommission vom 17. Mai 2002 zur Durchführung der Verordnung (EG) Nr. 322/97 des Rates über die Gemeinschaftsstatistiken - Regelung des Zugangs zu vertraulichen Daten für wissenschaftliche Zwecke (Amtsbl. der EG Nr. L 133 vom 18. Mai 2002, 7 ff.).

Literatur

- Hauser, R. (Hg.) (1994): *Mikroanalytische Grundlagen der Gesellschaftspolitik: Ergebnisse aus dem gleichnamigen Sonderforschungsbereich*, Band 1 und 2. Berlin.
- Hauser, R., Wagner, G. & Zimmermann, K. (1998): Erfolgsbedingungen empirischer Wirtschaftsforschung und empirisch gestützter wirtschafts- und sozialpolitischer Beratung. Ein Memorandum, *Allgemeines Statistisches Archiv*, 82, 369-379.
- Krupp, H.-J. (1973): Sozialpolitisches Entscheidungs- und Indikatorensystem für die Bundesrepublik Deutschland. *Allgemeines Statistisches Archiv*, Band 57, 380-387.
- Köhler, S. (1999): Anonymisierung von Mikrodaten in der Bundesrepublik und ihre Nutzung – Ein Überblick. In: Statistisches Bundesamt (Hg.): *Forum der Bundesstatistik*, Band 31. Stuttgart: Metzler-Poeschel, 133-144.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (2001): *Wege zu einer besseren informationellen Infrastruktur*. Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. Baden-Baden: Nomos.
- Lüttinger, P., Schimpl-Neimanns, B., Wirth H. & Papastefanou G. (2003): Mikrodaten (German Microdata Lab): Das Servicezentrum für amtliche Mikrodaten bei ZUMA. *ZUMA-Nachrichten*, 52, 153-172.
- Müller, W., Blien, U., Knoche, P. & Wirth, H. (1991): Die faktische Anonymität von Mikrodaten. In: Statistisches Bundesamt (Hg.): *Forum der Bundesstatistik*, Band 19. Stuttgart: Metzler-Poeschel.
- Smeeding, T., Jesuit, D. & Alkemade, P. (2002): The LIS/LES Project Databank: Introduction and Overview. *Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 122, 3, 497- 517.
- Statistisches Bundesamt (Hg.) (1999): Kooperation zwischen Wissenschaft und amtlicher Statistik - Praxis und Perspektiven - Beiträge zum Symposium am 31. Mai/1. Juni 1999 in Wiesbaden. *Forum der Bundesstatistik*, Band 34. Stuttgart: Metzler-Poeschel.
- Sturm, R. (2002): Wirtschaftsstatistische Einzeldaten für die Wissenschaft. *Wirtschaft und Statistik*, 2, 101-109.
- Zühlke, S. & Hetke, U. (2003): Datenbedarf und Datenzugang: Ergebnisse der ersten Nutzerbefragung des Forschungsdatenzentrums der Statistischen Landesämter. *Allgemeines Statistisches Archiv*, Band 87, 321-334.
- Zwick, M. (1999): Steuerstatistische Einzeldaten und ihre Auswertungsmöglichkeiten für die Wissenschaft. *Allgemeines Statistisches Archiv*, 83, 2, 248-253.

Dr. Sylvia Zühlke leitet die Geschäftsstelle des Forschungsdatenzentrums der Statistischen Landesämter. Arbeitsschwerpunkte: Projektkoordination, Beratung beim Zugang zu Mikrodaten der amtlichen Statistik, Weiterentwicklung des Datenzugangs. Forschungsschwerpunkte: Arbeitsmarktforschung und Längsschnittanalysen.

Diplom-Volkswirt Markus Zwick leitet das Forschungsdatenzentrum des Statistischen Bundesamtes. Seine Arbeitsschwerpunkte sind: Koordination von Projekten des Forschungsdatenzentrums, Rechtliche Fragen des Zugangs zu Mikrodaten der amtlichen Statistik, Weiter-

entwicklung des Einzeldatenzugangs für die Wissenschaft. Seine Forschungsschwerpunkte sind: Mikrosimulationsmodelle der Steuerstatistik, Weiterentwicklung der Anonymisierungsmethodik.

Sebastian Scharnhorst ist wissenschaftlicher Mitarbeiter im Forschungsdatenzentrum der Statistischen Landesämter. Arbeitsschwerpunkte: Anonymisierungsmethoden, Entwicklung von Metadaten, Rückerschließung von Mikrodaten.

Diplom-Soziologe Thomas Wende ist wissenschaftlicher Mitarbeiter im Forschungsdatenzentrum des Statistischen Bundesamtes. Seine Arbeitsschwerpunkte sind: Weiterentwicklung des automatisierten Fernrechnens, Entwicklung von Public Use Files für die Lehre (sog. Campus Files) und internationale Fragen des Mikrodatenzugangs.

Das **Forschungsdatenzentrum der Statistischen Landesämter** arbeitet eng mit dem **Forschungsdatenzentrum des Statistischen Bundesamtes** zusammen, um für die Wissenschaft ein abgestimmtes Dienstleistungsangebot der amtlichen Statistik zu entwickeln. Ausführliche Informationen zu den beiden Forschungsdatenzentren, zum Datenangebot und zu den Nutzungsmöglichkeiten sind im gemeinsamen Internetangebot unter www.forschungsdatenzentrum.de zu finden.

Datenzugang und Datenverfügbarkeit im Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung

Dr. Annette Kohlmann (IAB der BA – FDZ, Nürnberg)

1. Der Datenzugang in der Vergangenheit

Der Aufbau eines Forschungsdatenzentrums (FDZ), das sich vornehmlich um die Belange der Arbeitsmarktforschung kümmert, ist eine Aufgabe, die in der Bundesagentur für Arbeit (BA) und dem Institut für Arbeitsmarkt- und Berufsforschung (IAB) schon sehr früh als wichtig erkannt wurde. So gab es im Rahmen von einzelnen drittmittelgeförderten Projekten und unter Einsatz von Eigenmitteln der BA seit Jahren bereits Vorarbeiten, die sich auf die Erstellung und Bereitstellung von Scientific Use Files und die Datenfernverarbeitung konzentrierten.

Der bisherige Zugang zu den Daten der BA bzw. des IAB wurde von Forscherseite wiederholt als vergleichsweise gut, aber dennoch verbesserungsfähig angesehen. Restriktionen bestanden vor allem im Hinblick auf die Zugangsmodalitäten der Datenübermittlung von Originalmaterial (vgl. Kapitel 3), aber auch in Bezug auf inhaltliche Aspekte bei faktisch anonymisierten Scientific Use Files, die aus Datenschutzgründen zum Teil stark vergrößerte Merkmale aufwiesen und deren Analysierbarkeit dadurch eingeschränkt war.

Generell ist festzustellen, dass der Zugang zu den Daten der BA bzw. des IAB in der Vergangenheit dadurch gekennzeichnet war, dass keine institutionalisierte Anlaufstelle existierte und daraus einerseits für die Dateninteressenten ein erhöhter Aufwand und andererseits für die MitarbeiterInnen des IAB eine erhöhte Arbeitsbelastung außerhalb ihres originären Tätigkeitsfelds resultierte. Angesichts der Vielfalt und der Komplexität der Datenquellen, der vor Einführung des FDZ geringen Bündelung von Beratungskompetenz und der eher für interne Belange konzipierten Dokumentationsmaterialien entstand ein erheblicher Nachholbedarf an Service für die Wissenschaft.

2. Aufbau eines Forschungsdatenzentrums der BA im IAB

Diese Lücke soll das seit Anfang April 2004 neu geschaffene FDZ der BA im IAB schließen, das sich gegenwärtig noch in der Aufbauphase befindet. Das Ziel dieses Pilotprojekts, das aus Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) und des Bundesministeriums für Wirtschaft und Arbeit (BMWA) bis Ende November 2006 gefördert wird, besteht darin, eine funktionsfähige Organisation zu errichten, die unter Einhaltung rechtlicher Normen (insbesondere des Datenschutzes) die Aufbereitung, Dokumentation und Bereitstellung von Echtdaten der BA und des IAB übernimmt. Die Bereitstellung von Originalmaterial bezieht sich vor allem auf Daten, die nicht ohne erheblichen Informationsverlust anonymisiert werden können.

Damit wird entsprechend den Empfehlungen der Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) eine Erleichterung des Zugangs zu den amtlichen Mikrodaten der BA und des IAB für die wissenschaftliche Forschung angestrebt. Daher ist gerade in der ersten Phase des FDZ die Entwicklung von *transparenten und standardisierten Zugangsregelungen*, bei denen eine *Gleichbehandlung von Anfragen externer ForscherInnen* gewährleistet ist, von zentraler Bedeutung. Dies impliziert, dass Anfragen ohne Ansehen der antragstellenden Person im FDZ bearbeitet werden, sondern ausschließlich nach ihrem Eingangstermin. Ebenso gilt, dass Anfragen und die Art der Datennutzung durch externe Nutzer vertraulich behandelt werden.¹

Um den Datenzugang für Externe zu ermöglichen, klärt das FDZ unter *Einhaltung von Datenschutzregelungen* die Datennutzungsmöglichkeit für Externe ab. Dies geschieht in einem standardisierten Verfahren, das – anders als in der Vergangenheit – verfahrenstechnisch vereinfacht wird. Ziel ist es, nicht mehr individuelle Datenzugangsberechtigungen beantragen zu müssen, sondern über die Bildung von Fall- und Nutzergruppen ein standardisiertes Genehmigungsverfahren zu erlangen, das vor allem im Hinblick auf Gastwissenschaftler-Arbeitsplätze (vgl. Kapitel 3) von großer Bedeutung ist. Das FDZ definiert hierbei im Vorfeld – in Kooperation mit der Wissenschaft – Datensätze, die möglichst vielseitig statistisch auswertbar sind. Darüber hinaus fällt in den Aufgabenbereich des FDZ auch die Kontrolle von Analyseergebnissen hinsichtlich datenschutzrechtlicher Sachverhalte. Dies trifft vor allem auf die Nutzung des Originalmaterials der Daten der BA und des IAB durch Externe zu. Damit wird für ForscherInnen Rechtssicherheit hergestellt und der Aufwand für die Beteiligten minimiert.

Die Erleichterung des Datenzugangs, die durch das FDZ geschaffen wird, betrifft jedoch nicht nur technische und datenschutzrechtliche Aspekte. Da es sich bei den Daten der BA und des IAB in der Regel um Sozialdaten handelt, d.h. um im Verwaltungsablauf anfallende (prozessproduzierte) Daten, liegt ein besonderes Augenmerk auf einer *genauen und umfassenden Dokumentation* unter Berücksichtigung datensicherheitsrechtlicher Aspekte. Amtliche Daten sind nicht ohne weiteres mit den üblicherweise in sozial- und wirtschaftswissenschaftlichen Surveys gewonnenen Daten zu vergleichen. Gerade im Hinblick auf die Datengenese sind bei diesen Daten die Dokumentationen und Arbeitshilfen von besonderer Bedeutung, um ein Arbeiten mit den Daten für Personen, die bislang geringe oder keine Erfahrungen mit amtlichen Daten gesammelt haben, zu ermöglichen.²

3. Zugang zu den Daten am FDZ der BA im IAB

Bei den Daten³ der BA bzw. des IAB, die im FDZ zur Verfügung gestellt werden, handelt es sich in der Regel um Sozialdaten nach § 67 Sozialgesetzbuch (SGB) X bzw. um Sozialdaten gleichgestellte Daten. Sozialdaten sind Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbarer natürlichen Person, die von der BA im Hinblick auf ihre Aufgaben nach dem SGB III erhoben und verarbeitet werden.

¹ Daten von NutzerInnen des Datenmaterials der BA/des IAB werden im FDZ ausschließlich zu Dokumentations- und Evaluationszwecken gesammelt. Eine darüber hinausgehende Veröffentlichung von Nutzerdaten erfolgt ausschließlich mit Zustimmung der NutzerInnen.

² Aspekte der Datendokumentation und der Beratung werden in diesem Beitrag nicht weiter behandelt.

³ ‚Daten‘ bezeichnet hier ausschließlich Mikrodatenfiles. Analyseergebnisse basierend auf Aggregatdaten sind dadurch nicht betroffen und können beim Statistik-Datenzentrum der BA (s.u.) beantragt werden.

Die Sensibilität dieser Daten erfordert es, dass für den Zugang von Externen zu den Daten der BA und des IAB spezifische Anforderungen gelten. So wird der Zugang zu nicht anonymisierten Sozialdaten prinzipiell durch § 75 SGB X reglementiert, der besagt, dass diese Daten nur für Forschung im Bereich der Sozialversicherung zugänglich sind.⁴ Die Datennutzung für kommerzielle Forschung und Interessen ist damit explizit ausgeschlossen.

Datenübermittlung nach § 75 SGB X

Nicht anonymisierte Sozialdaten können in einem spezifischen Verfahren an Externe übermittelt werden, Ansprechpartner ist dabei bis auf weiteres das Datenzentrum der Statistik der BA⁵, bei dem die Übermittlung von Individualdaten für die Forschung und Planung nach § 75 SGB X beantragt werden kann.

Diese Übermittlung von Sozialdaten muss durch das Bundesministerium für Wirtschaft und Arbeit genehmigt werden. Die Bedingungen für eine Genehmigung sind folgendermaßen definiert:

- Es muss sich um ein Projekt aus dem Sozialleistungsbereich handeln,
- das öffentliche Interesse muss das Geheimhaltungsinteresse des Betroffenen erheblich überwiegen,
- das Einholen der Einwilligung des Betroffenen muss unzumutbar sein und
- der Zweck des Vorhabens darf nicht auf andere Weise (z.B. durch ein vorhandenes Scientific Use File oder über Ferndatenanalyse) erreichbar sein.

Die Kosten, die durch dieses Verfahren entstehen, werden in Rechnung gestellt.

Dieses Verfahren der Datenübermittlung gehörte in der Vergangenheit zu den üblichen Zugangswegen für Externe zu amtlichen Daten über den Arbeitsmarkt und die wirtschaftliche Entwicklung in Deutschland.

Weiterhin bestand in der Vergangenheit die Möglichkeit, im Rahmen von Kooperationsprojekten mit MitarbeiterInnen des IAB Datenzugang zu erhalten oder die Daten im Rahmen von Informationsaufenthalten im IAB auszuwerten, wie dies z.B. im Rahmen des neu aufgelegten Graduiertenprogramms des IAB möglich ist.

Diese nicht-standardisierten Datenzugangswege der Vergangenheit bestehen nach wie vor. Zudem ist auch die Möglichkeit gegeben, Sonderauswertungen durch das Datenzentrum der Statistik der BA anfertigen zu lassen. Diese Sonderauswertungen sind kostenpflichtig und beziehen sich auf Auswertungen aus den amtlichen Statistiken der BA zur Lage und Entwicklung auf dem Arbeits- und Ausbildungsmarkt, zur Beschäftigung, zur Arbeitsförderung inkl. Eingliederungsbilanz sowie zu den Lohnersatz- und sonstigen Leistungen.

Da all diese Verfahren in der Regel ressourcenintensiv, langwierig oder abhängig von subjektiven Interessen der Beteiligten sind, resultierten in der Vergangenheit zwangsläufig ungleiche

⁴ Für anonymisierte Daten, wie z.B. Scientific Use Files, gilt die Bedingung, dass diese für die Arbeitsmarkt- und Berufsforschung nutzbar sind (SGB III, § 282, 7).

⁵ Erreichbar unter der E-mail Adresse: Service-Haus.Statistik-Datenzentrum@arbeitsagentur.de

Chancen beim Datenzugang, da die Ressourcenausstattung der antragstellenden Forschungsinstitute sehr unterschiedlich war.

Daher füllt das FDZ hier eine wichtige Lücke, indem standardisierte und transparente Zugangswege die oben genannten weiterhin bestehenden Möglichkeiten ergänzen.

Generell lassen sich drei prinzipielle Zugangswege⁶ zu den Daten der BA und des IAB im FDZ unterscheiden:

1. *Scientific Use Files* sind faktisch anonymisierte Auszüge aus Originaldatensätzen, die der Forschung unter bestimmten Auflagen zeitlich befristet zur Verfügung gestellt werden. Seit 1995 steht eine 1%ige Stichprobe der Beschäftigten- und Leistungsempfängerhistorik zur Verfügung. Diese IAB-Beschäftigtenstichprobe (IABS) wird als Basis- und Regionalfile über das Zentralarchiv für empirische Sozialforschung in Köln (ZA) vermittelt. Die vorhandenen Scientific Use Files werden möglichst regelmäßig durch das FDZ aktualisiert.
2. Bei der *Datenfernverarbeitung (Schalterstelle)* erhalten externe ForscherInnen aufgrund der Sensibilität der Daten keinen direkten Zugang. Stattdessen werden von ihnen Analysesyntaxen eingesendet, die anhand des Originalmaterials prozessiert werden. Die im Hinblick auf datenschutzrechtliche Aspekte geprüften Ergebnistabellen werden an die Nutzer zurückgesendet. Dieser Weg wird bei Daten gewählt, die nicht ohne großen Informationsverlust anonymisiert werden können.
3. *Gastwissenschaftler-Arbeitsplätze* sind der im Hinblick auf den Datenzugang restriktivste, gleichzeitig im Hinblick auf die Analysemöglichkeiten jedoch auch ergiebigste Zugangsweg für externe ForscherInnen. Gastwissenschaftler-Arbeitsplätze sind dann sinnvoll, wenn das Datenmaterial nicht ohne erheblichen Informationsverlust (faktisch) anonymisierbar ist, gleichermaßen aber so sensibel und die Bearbeitung so aufwändig ist, dass Auswertungen nur am Ort der Datenhaltung durchgeführt werden können. Im FDZ der BA im IAB liegen diese Daten schwach anonymisiert (d.h. ohne Identifikatoren, aber mit vollem Merkmalsumfang) vor.

4. Datenbestand der BA bzw. des IAB

Der Datenbestand der BA und des IAB ist heterogen im Hinblick auf die Genese, den Inhalt und die Zeitdimension der Daten.⁷

Im Hinblick auf die *Datengenese* kann generell unterschieden werden zwischen Daten, die über das Meldeverfahren der Sozialversicherung erfasst bzw. durch BA-interne Verfahren gesammelt und Daten, die durch Befragungen des IAB erhoben werden.

Die unterschiedlichen *Zeitdimensionen* beziehen sich auf stichtagsbezogene Statistiken, Längsschnittdaten auf Stichtagsbasis und zeitraumbezogene Datensätze (Historiken).

⁶ Die Erstellung und Bereitstellung von Public Use Files zu Daten der BA und des IAB wird gegenwärtig noch nicht praktiziert. Eine Prüfung ihrer Machbarkeit erfolgt jedoch im Rahmen des Pilotprojekts FDZ.

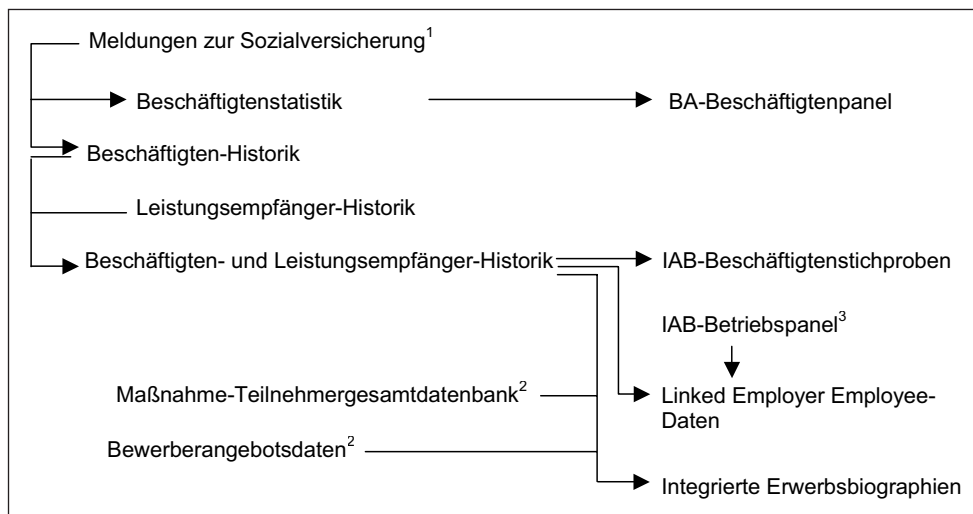
⁷ Im Folgenden wird ausschließlich die Aufbereitung von Daten angesprochen, die relevant ist für die im FDZ zur Verfügung gestellten Daten. Es werden nicht alle vorhandenen Quelldaten beschrieben. Ein Überblick über alle Daten, die bis Ende der Projektlaufzeit zur Verfügung gestellt werden sollen – soweit datenschutzrechtlich möglich – wird in Oertel et al. (2006) gegeben.

Inhaltlich umfassen die Daten der BA Informationen zu Beschäftigten, Betrieben, Arbeit-suchenden, arbeitsmarktpolitischen Maßnahmen und Leistungsempfängern.

Aus verschiedenen Einzeldaten werden im IAB integrierte Datenbasen erstellt, in denen meh-rere Teilaspekte des Arbeitsmarktes miteinander verknüpft werden. Das FDZ strebt die sukzes-sive Bereitstellung dieser Daten unter Berücksichtigung der geltenden Datenschutzregelungen (vgl. Kapitel 3) an.

Die folgende Darstellung verdeutlicht den Zusammenhang zwischen Datenquellen, Datensät-zen und deren Bereitstellung im FDZ:

Abbildung 1: Datenquellen und Datenwege



1 = Datenquelle: Meldungen an die Sozialversicherungsträger
 2 = Datenquelle: BA-interne Verfahren
 3 = Datenquelle: IAB-Survey
 Quelle: eigene Darstellung.

Grundlage der meisten Daten der BA und des IAB sind die Meldungen der Arbeitgeber über ihre sozialversicherungspflichtig beschäftigten Arbeitnehmer an die Krankenkassen. Die gesetzliche Grundlage ist das mit Wirkung vom 1. Januar 1973 eingeführte integrierte Meldeverfahren zur Kranken-, Renten- und Arbeitslosenversicherung (DEVO/DÜVO), das seit dem 1.1.1999 in einer aktualisierten Version (DEÜV) gilt.⁸ Damit liegt eine Vollerhebung aller sozialversicherungspflichtig Beschäftigten (seit 1999 auch der geringfügig Beschäftigten) in Deutschland vor, die gegenwärtig ca. 32 Mio. Beschäftigungsverhältnisse umfasst.

8 DEVO (Datenerfassungsverordnung) - Verordnung über die Erfassung von Daten für die Träger der Sozialversicherung und für die Bundesanstalt für Arbeit vom 24.11.1972 (BGBl. I, 2159 ff.).
 DÜVO (Datenübertragungsverordnung) - Verordnung über die Datenermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der Bundesanstalt für Arbeit vom 18.12.1972 (BGBl. I, 2482 ff.).
 DEÜV (Datenerfassungs- und übermittlungsverordnung) - Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung vom 10.2.1998 (BGBl. I, 343 ff.).

Die Krankenkassen leiten die Meldungen an die Rentenversicherungsträger weiter, von wo aus sie dann der BA übermittelt werden, die diese Daten gemäß ihres gesetzlichen Auftrags (SGB III, § 281) registriert und in der *Beschäftigtenstatistik* verwaltet. Der Eingang der Daten erfolgt verzögert, da Arbeitgebern Abgabefristen für die Meldungen eingeräumt werden. Diese Meldungen sind die Basis für die Ziehung des BA-Beschäftigtenpanels, das als Scientific Use File externen Nutzern zur Verfügung gestellt wird (vgl. Kapitel 5).

Im IAB werden die Meldungen zur Sozialversicherung archiviert und als Historikdatensatz aufbereitet, der *Beschäftigten-Historik* (BeH). Die Beschäftigtenhistorik umfasst personen-, betriebs- und beschäftigungsbezogene Merkmale sowie Merkmale zum Jahreskonto des/der Versicherten.

Die *Leistungsempfänger-Historik* (LeH) speist sich aus Daten aus dem Verfahren zur Auszahlung von Leistungen (coLei), einem internen Verfahren der BA. Diese Daten umfassen neben personenbezogenen Informationen auch Informationen über den Leistungsbezug (Arbeitslosengeld, Arbeitslosenhilfe, Unterhaltsgeld, Kranken-/Pflegeversicherungsbeiträge) und Merkmale, die sich auf das letzte Beschäftigungsverhältnis beziehen. Die Daten werden aus den Arbeitsagenturen täglich an die Krankenkassen im Rahmen des integrierten Meldeverfahrens übermittelt.

Aus der BeH und der LeH wird im IAB die *Beschäftigten- und Leistungsempfänger-Historik* (BLH) auf der Basis der Sozialversicherungsnummern aufgebaut. Dieser neu entstandene Datensatz ist zentral für fast alle weiteren Datenmengen, die im FDZ der BA im IAB der Fachöffentlichkeit zur Verfügung gestellt werden. Die IAB-Beschäftigtenstichproben werden als (bislang 1%ige) Samples aus der BLH gezogen (vgl. Kapitel 5).

Die personenbezogenen Daten der BLH können in einem weiteren Schritt mit Betriebsinformationen verknüpft werden, z.B. mit dem IAB-Betriebspanel (vgl. Kapitel 5), einer Panelbefragung des Forschungsbereichs „Betriebe und Beschäftigung“ des IAB. Daraus entstehen verschiedene Varianten eines *Linked Employer Employee Datensatzes* (LIAB), die Angebots- und Nachfrageseite auf dem Arbeitsmarkt gleichermaßen analysierbar machen (vgl. Kapitel 6).

Zur Generierung einer anderen integrierten Datenbasis, den *Integrierten Erwerbsbiographien* (IEB, vgl. Kapitel 6) werden neben der BLH zusätzliche interne BA-interne Datenbasen herangezogen. Dazu gehört die *Maßnahme-Teilnehmergesamtdatenbank* (MTG), die im IAB auf der Basis der BA-Verfahren zur Arbeitsvermittlung (coArb) und zur Verwaltung von Maßnahmeteilnahmen (coSach) sowie auf Basis der zentralen Kundendatenbank der Arbeitsagenturen (CORA) erstellt wird. Sie umfasst personenbezogene und erwerbsbiographische Merkmale, Merkmale zum Leistungsbezug sowie Informationen zu Maßnahmen und Maßnahmeteilnahme.⁹

Eine weitere Datenquelle der IEB sind die *Bewerberangebotsdaten* (BewA); diese umfasst alle Arbeitssuchenden, die sich in den Arbeitsagenturen als solche melden und gibt damit einen umfassenderen Einblick in das Arbeitskräftereservoir als die Leistungsempfänger-Historik, die

⁹ Die Daten zu Maßnahmen und Maßnahmeteilnahme umfassen u.a. den Bereich der beruflichen Fortbildung, Weiterbildung, Trainingsmaßnahmen, Deutschlehrgänge, die Freie Förderung, Arbeitsbeschaffungsmaßnahmen, Strukturanpassungsmaßnahmen, Lohnkostenzuschuss-Altfälle, Eingliederungszuschüsse, Eingliederungsvertrag, Mobilitätshilfen, Überbrückungsgeld für ExistenzgründerInnen, Beschäftigungshilfen für Langzeitarbeitslose etc.

ausschließlich Personen erfasst, die Leistungen beziehen. Sie enthält neben personenbezogenen Merkmalen erwerbsbiographische Variablen und BA-dienstleistungsbezogene Merkmale (wie z.B. Beratung, Leistungsbezug, Förderung) und stammt aus dem BA-internen Verfahren der Arbeitsvermittlung (coArb). Der aus BeH, LeH, MTG und BewA generierte Datensatz der IEB ermöglicht damit eine sehr umfassende Analyse von Erwerbsbiographien.

5. Gegenwärtig verfügbare Daten im FDZ der BA im IAB

Gegenwärtig¹⁰ stehen als Scientific Use Files die IAB-Beschäftigtenstichproben (Basisfile und Regionalfile) und das BA-Beschäftigtenpanel zur Verfügung. Mittels Ferndatenanalyse ist das IAB-Betriebspanel für externe Wissenschaftlerinnen zugänglich. Die Nutzung von Daten, die ausschließlich im Rahmen von Gastaufenthalten im FDZ ausgewertet werden können, wird vorbereitet (vgl. Kapitel 6).

IAB-Beschäftigtenstichproben

Die IAB-Beschäftigtenstichproben liegen mittlerweile schon seit neun Jahren als Scientific Use Files der Fachöffentlichkeit zur Nutzung vor. Hierbei handelte es sich bislang um 1%-Stichproben aller sozialversicherungspflichtig Beschäftigten, die aus der im IAB als Längsschnitt aufbereiteten BLH gezogen wurden.

Zwei Varianten sind zu unterscheiden:

- ein Basisfile, der die Jahre 1975 bis 1995 umfasst und aus Gründen der Anonymisierung nur eine geringe regionale Differenzierung aufweist, sowie
- ein Regionalfile, das die Jahre 1975 bis 1997 umfasst, eine differenzierte regionale Gliederung besitzt, auf der inhaltlichen Ebene allerdings durch stärkere Zusammenfassungen als das Basisfile gekennzeichnet ist.

Die Erstellung beider Varianten der IAB-Beschäftigtenstichprobe wurde durch die Deutsche Forschungsgemeinschaft im Rahmen von Forschungsprojekten mitfinanziert. Die Scientific Use Files sind über das Zentralarchiv für empirische Sozialforschung (ZA) in Köln verfügbar.¹¹

Gegenwärtig wird im Rahmen eines Verbundprojekts zwischen der Universität Siegen und dem IAB eine Aktualisierung der Regionalstichprobe erarbeitet. Dieses neue Regionalfile umfasst den Zeitraum 1975 bis 2001 und eine größere Population (2%-Stichprobe), was ein geringeres Ausmaß von Merkmalsvergrößerungen ermöglicht. In diesem neuen File sind ab April 1999 aufgrund der Neuregelung der geringfügigen Beschäftigung auch Informationen zu geringfügigen Beschäftigungsverhältnissen verfügbar, darüber hinaus können zudem auch die Abgabegründe von Meldungen stärker differenziert werden. Dieses Scientific Use File wird ebenso wie seine Vorgänger über das ZA interessierten wissenschaftlichen Nutzern zur Verfügung gestellt werden.

¹⁰ Die folgende Beschreibung bezieht sich auf den Stand im August 2004.

¹¹ ZA-Studiennr. 3136 (IAB-Beschäftigtenstichprobe, Basisfile) und ZA-Studiennr. 3348 (IAB-Beschäftigtenstichprobe, Regionalfile).

BA-Beschäftigtenpanel

Neben der IAB-Beschäftigtenstichprobe existiert seit kurzem eine weitere Datenquelle für Analysen über Beschäftigte, das Beschäftigtenpanel der BA. Hierbei handelt es sich um eine knapp 2%-Stichprobe aus den Quartalsdaten der Beschäftigtenstatistik der BA, die mit Informationen zu Betrieben angereichert wurden. Im Unterschied zur Beschäftigtenstichprobe handelt es sich also nicht um vollständige Biographien, sondern um quartalsweise abgelegte Informationen zu Beschäftigten.

Gegenwärtig liegen 17 Wellen vor, vom 1. Quartal 1998 bis zum 4. Quartal 2002. Eine jährliche Aktualisierung wird durch das FDZ vorgenommen werden.

Jede Welle umfasst ca. 500.000-600.000 sozialversicherungspflichtig Beschäftigte (seit Juni 1999 auch geringfügig Beschäftigte). Im Beschäftigtenpanel sind verschiedene soziodemographische Variablen, Individualmerkmale und betriebliche Merkmale (v.a. Merkmale zur Zusammensetzung der Belegschaft) enthalten.

Die Stichprobenauswahl auf der Basis von Geburtstagen bewirkt einen Ausgleich der Panelmortalität. Ein weiterer Vorteil des Datensatzes ist die strukturtreue Abbildung von Saison- und Konjunkturmustern und der zeitnahe Rand (gegenwärtig bis Ende 2002) sowie ein vereinfachtes Datenhandling von komplexen prozessproduzierten Daten.

Das BA-Beschäftigtenpanel kann gegenwärtig über die BA bezogen werden und wird demnächst als Scientific Use File über das ZA beziehbar sein. Darüberhinaus soll es für differenzierte (regionale) Analysen auch über die Datenfernverarbeitung im FDZ zugänglich gemacht werden.¹²

IAB-Betriebspanel

Das IAB-Betriebspanel ist die größte Betriebsbefragung der Bundesrepublik Deutschland, bei der jährlich ca. 16.000 Betriebe aller Wirtschaftszweige und Betriebsgrößenklassen Auskunft zu verschiedenen, in der Hauptsache personalpolitischen Fragestellungen, geben.¹³ Mit dem IAB-Betriebspanel sind repräsentative Auswertungen auch auf der Bundeslandebene durch Hochrechnung innerhalb der Zellen einer Matrix von zehn Betriebsgrößenklassen und 20 Wirtschaftszweigen möglich. Das IAB-Betriebspanel wird vom Forschungsbereich „Betriebe und Beschäftigung“¹⁴ des IAB durchgeführt. Der Datenzugang für externe ForscherInnen erfolgt seit 1999 über die Schalterstelle des IAB-Betriebspanels, die vom BMBF gefördert und im Mai 2004 in das FDZ integriert wurde.

Gegenüber Individualdaten ist es bei Betriebsdaten weitaus schwieriger, eine faktische Anonymisierung ohne einen unverhältnismäßigen Informationsverlust zu gewährleisten. Vor diesem Hintergrund wurde die Möglichkeit der Erstellung eines Scientific Use Files zum IAB-Betriebspanel im IAB untersucht. Gegenwärtig ist das IAB an einem Projekt des Statistischen Bundesamtes beteiligt, in dem u.a. auch die Möglichkeiten einer adäquaten Anonymisierung des Betriebspanels erforscht werden. Mittelfristig ist geplant, dass Forscher im Rahmen von Gastaufenthalten im FDZ mit den Originaldaten arbeiten können.

¹² Nähere Informationen unter <http://www.pub.arbeitsamt.de/hst/services/statistik/000200/html/panel/panel.html>.

¹³ Weitere Informationen zum Betriebspanel sind erhältlich unter <http://betriebspanel.iab.de>.

¹⁴ Vgl. <http://www.iab.de/asp/mitarbeiterDB/nodeBereichDetails.asp?pkyNode=9>.

6. Zukünftig verfügbare Daten im FDZ der BA im IAB

Für die weitere Entwicklung des FDZ sind gegenwärtig zwei Arbeitsschwerpunkte auszumachen: Zum einen die Öffnung des FDZ für Gastwissenschaftler und zum anderen die Bereitstellung von integrierten Daten.

Die Öffnung für Gastwissenschaftler bedeutet, dass alle bislang verfügbaren Daten (IAB-Beschäftigtenstichproben, BA-Beschäftigtenpanel und IAB-Betriebspanel) ebenso wie auch die zukünftig verfügbaren Datensätze (s.u.) in schwach anonymisierten Versionen im FDZ vorgehalten und interessierten Nutzern nach datenschutzrechtlicher Prüfung zugänglich gemacht werden. Schwach anonymisiert heißt, dass ausschließlich Identifikatoren (z.B. Sozialversicherungsnummern) gelöscht bzw. verfremdet werden und ansonsten der vollständige Merkmalsumfang der Daten angeboten wird. Damit besteht für Forscher die Möglichkeit, Analysen mit stärker differenzierten Merkmalen vorzunehmen. Die Analyseergebnisse, die in Gastaufenthalten gewonnen werden, werden im FDZ datenschutzrechtlich geprüft.

Wie in Kapitel deutlich wurde, wird im FDZ der BA im IAB die Bereitstellung von integrierten Daten der Bereitstellung einzelner, unverknüpfter Datenbestände vorgezogen. Diese Priorisierung von Daten, die durch die Verknüpfung mehrerer inhaltlicher Aspekte charakterisiert sind, begründet sich in den größeren Analysemöglichkeiten. So ist es in der Regel für die Arbeitsmarktforschung von besonderem Interesse, nicht Informationen z.B. zu Beschäftigung und Maßnahmeteilnahme getrennt zu analysieren, sondern gerade die Zusammenhänge und ihre Verschränkung zwischen den beiden Aspekten des Arbeitsmarkts zu untersuchen. Ähnliches gilt auch für die Angebots- und Nachfrageseite auf dem Arbeitsmarkt.

Mit der Verknüpfung ist jedoch in der Regel die datenschutzrechtliche Anforderung verbunden, dass diese Daten ausschließlich an abgeschotteten Arbeitsplätzen nutzbar sind, damit das aufgrund des größeren Merkmalsumfangs erhöhte Reidentifikationsrisiko von Betrieben bzw. Individuen möglichst gering gehalten wird.

Gegenwärtig sind zwei integrierte Datensätze im FDZ geplant, Linked Employer Employee Daten (LIAB) und die integrierten Erwerbsbiographien (IEB).

Linked Employer Employee Daten (LIAB)

Die Verknüpfung zwischen Beschäftigteninformationen und Betriebsinformationen ist vor allem dann hilfreich, wenn Interaktionen zwischen Angebots- und Nachfrageseite auf dem Arbeitsmarkt untersucht werden. Aufgrund der Komplexität der Daten und datenschutzrechtlicher Restriktionen lagen im IAB bislang ausschließlich Querschnitte auf der Personenebene zum 30.6. eines Jahres (Erhebungszeitpunkte der Betriebspanel-Wellen) bzw. individuelle Datensätze einzelner IAB-Projekte vor. „Den“ LIAB gab es bislang nicht. Die Nutzung dieser Daten war in der Regel ausschließlich über Kooperationsprojekte mit IAB-WissenschaftlerInnen möglich.

Es ist geplant, im Forschungsdatenzentrums der BA im IAB Varianten der Linked Employer Employee Daten möglichst bald im Rahmen von Gastaufenthalten zur Verfügung zu stellen. Da unterschiedliche Forschungsfragen jeweils unterschiedliche Variablenkataloge und Datenmodelle (Querschnitt/Längsschnitt) erfordern, werden gegenwärtig Konzepte von LIAB-Varianten erarbeitet, die dann standardisiert zugänglich gemacht werden sollen. Zu diesen gegenwärtig

diskutierten Varianten gehören auf betrieblicher Ebene aggregierte Individualinformationen als Zusatzvariablen zum IAB-Betriebspanel, Jahresscheiben zum 30.6. für die Jahre 1993 bis 2001 inkl. generierter Variablen zur Erwerbsbiographie, Mobilitätsdatensätze mit jeweils drei Jahresscheiben (d.h. Verknüpfung mit dem Vor- und dem Nachfolgejahr), integrierte Beschäftigtenhistoriken für eine Auswahl von je 2000 Betrieben in Ost- und Westdeutschland zwischen 1993 und 2002 sowie integrierte Historiken für alle Betriebe, die zwischen 1993 und 2002 durchgängig am IAB-Betriebspanel teilgenommen haben.

Da die Verknüpfung dieser beiden Datenmengen sehr speicherintensiv ist, müssen sparsame Datenmodelle gewählt werden, die eine maximale Auswertbarkeit der Daten erlauben. Um diese Datenmodelle möglichst effektiv gestalten zu können, plant das Forschungsdatenzentrum in Kooperation mit der Arbeitsgruppe „LIAB“ des IAB und dem Forschungsbereich „Betriebe und Beschäftigung“ des IAB einen Workshop zu den Linked Employer Employee Daten des IAB im November 2004. Ziel dieses Workshops ist es, in Kooperation mit interessierten Wissenschaftlern die vorbereiteten Datenkonzepte zu diskutieren und - soweit technisch und datenschutzrechtlich möglich - Nutzerwünsche mit aufzunehmen.

Gastaufenthalte sollen dann (zunächst im Testbetrieb) sobald wie möglich im FDZ realisiert werden.

Integrierte Erwerbsbiographien (IEB)¹⁵

Wie aus der Abbildung in Kapitel 4 hervorgeht, stellen die integrierten Erwerbsbiographien eine Verknüpfung zwischen einem Datensatz, der aus dem amtlichen Meldeverfahren resultiert (BeH) und drei Datensätzen, die aus BA-internen Verfahren entstehen (LeH, MTG und BewA) dar.

Diese vier Datenquellen werden über die Sozialversicherungsnummer bzw. die Kundennummer im IAB miteinander verknüpft. Daraus entsteht ein Datensatz, der für den Zeitraum 1990 bis 2004 ca. 637 Mio Beobachtungen von 60 Mio Individuen umfasst. Hierbei ist jedoch darauf hinzuweisen, dass sich nicht sämtliche Basisdaten über den kompletten Zeitraum zwischen 1990 und 2004 erstrecken; gegenwärtig wird der komplette Zeitraum (1990 bis Juni 2004) lediglich von Merkmalen aus der LeH abgedeckt. Bei den anderen drei Datenbasen wurden alle zum Erstzeitpunkt der IEB verfügbaren Meldungen berücksichtigt. Informationen aus der BeH umfassen so den Zeitraum 1990 bis Ende 2002 (Grund der Verzögerungen sind mögliche Nachmeldungen, vgl. Kapitel 4). Informationen aus der MTG liegen vor für den Zeitraum Januar 2000 bis Juni 2004, Bewerberangebote von Januar 1997 bis Juni 2004.

Analysen zu Erwerbsbiographien sind in diesem zeitlichen Rahmen vor allem im Hinblick auf Effekte von Beschäftigungen und Leistungsempfang auf spätere Maßnahmeteilnahme bzw. Bewerberangebote möglich.

Aufgrund der unterschiedlichen Herkunft und des unterschiedlichen verfahrenstechnischen Hintergrunds der Daten stellt die Aufbereitung und Überprüfung dieser einheitlichen Datenbasis eine große Herausforderung dar, die gegenwärtig im IAB geleistet wird.

¹⁵ Diese Planung steht unter der Einschränkung, dass das bislang prototypisch entwickelte Generierungsverfahren datenschutzrechtlich genehmigt wurde. Dieser Prozessschritt ist derzeit noch nicht abgeschlossen.

Es ist geplant, diese Daten ebenso wie die LIAB-Varianten - sobald datentechnisch und datenschutzrechtlich möglich – der Fachöffentlichkeit in Form von Gastaufenthalten schwach anonymisiert zur Verfügung zu stellen.

Aus Gründen der Verarbeitbarkeit der Daten und des Speicherproblems wird dies im Rahmen einer 2%-Stichprobe auf der Basis des Geburtstagskonzept erfolgen. Im Hinblick auf die Möglichkeit der faktischen Anonymisierung der Daten für die Erstellung eines Scientific Use Files fanden erste Beratungen in einer Arbeitsgruppe aus IAB, externen Wissenschaftlern¹⁶ und dem Bundesbeauftragten für Datenschutz statt.

7. Fazit

Der hier vorgelegte Überblick über die Datenverfügbarkeit und den Datenzugang im FDZ der BA im IAB spiegelt den gegenwärtigen Stand (August 2004) der Arbeiten im FDZ wider. Sicherlich kann zum gegenwärtigen Zeitpunkt nur ein vorsichtiger Blick auf die Zukunft des FDZ geworfen werden.

Abschließend kann jedoch festgehalten werden, dass das FDZ der BA im IAB dann der richtige Ansprechpartner für Externe ist, wenn diese

- empirische wissenschaftliche nicht-kommerzielle Forschung betreiben,
- den Zugang zu Mikrodaten über den Arbeitsmarkt in Deutschland suchen,
- bereit sind, sich in komplexe Datensätze einzuarbeiten und
- bereit sind, diese als Scientific Use Files, per Datenfernverarbeitung oder im Rahmen von Gastaufenthalten zu nutzen.

Für Interessenten aus der kommerziellen Forschung, Personen, die Auftrags- oder Sonderauswertungen durchführen lassen möchten bzw. Aggregatdaten oder Sozialdaten nach § 75 SGB X übermittelt bekommen möchten bietet sich das Serviceangebot des Datenzentrums der Statistik der BA an (s.o.).

Literatur

Bellmann, L. (2002): Das IAB-Betriebspanel. Konzeption und Anwendungsbereiche. *Allgemeines Statistisches Archiv*, 86, 2, 177-188.

Bellmann, L., Bender, S. & Kölling, A. (2002): Der Linked Employer-Employee-Datensatz aus IAB-Betriebspanel und Beschäftigtenstatistik der Bundesanstalt für Arbeit (LIAB). In: Kleinhenz, G. (Hg.): *IAB-Kompodium Arbeitsmarkt- und Berufsforschung. Beiträge zur Arbeitsmarkt- und Berufsforschung*, BeitrAB 250, 21-29. http://doku.iab.de/beitrab/2002/beitr250_103.pdf.

¹⁶ Beteiligte Institutionen sind IWH, IZA, RWI, ZUMA und die Universität St. Gallen.

- Bellmann, L., Kohaut, S. & Lahner, M. (2002): Das IAB-Betriebspanel – Ansatz und Analysepotenziale. In: G. Kleinhenz (Hg.): *IAB-Kompendium Arbeitsmarkt- und Berufsforschung. Beiträge zur Arbeitsmarkt- und Berufsforschung*, BeitrAB 250, 13-20. http://doku.iab.de/beitrab/2002/beitr250_102.pdf.
- Bender, S., Haas, A. (2002): Die IAB-Beschäftigtenstichprobe. In: G. Kleinhenz (Hg.): *IAB-Kompendium Arbeitsmarkt- und Berufsforschung. Beiträge zur Arbeitsmarkt- und Berufsforschung*, BeitrAB 250, 3-12. http://doku.iab.de/beitrab/2002/beitr250_101.pdf.
- Brand, R., Bender, S. & Kohaut, S. (1999): Möglichkeiten der Erstellung eines Scientific Use Files aus dem IAB-Betriebspanel. In: Statistisches Bundesamt (Hg.): *Möglichkeiten einer wissenschaftlichen Nutzung von Unternehmensdaten aus der amtlichen Statistik*, Band 14 der Schriftenreihe Spektrum Bundesstatistik. Stuttgart: Metzler-Poeschel, 148-165.
- Brinkmann, C. (2002): Pilotprojekt "Beschleunigte Datenbereitstellung für externe Institute zum Zwecke der Evaluation arbeitsmarktpolitischer Instrumente". *IAB-Werkstattbericht* Nr. 2/20.3.2002, 13-15. <http://doku.iab.de/werkber/2002/wb0202.pdf>.
- Koch, I. & Meinken, H. (2004): The Employment Panel of the German Federal Employment Service. *Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften*, Im Druck.
- Koch, I. & Meinken, H. (2003): Das Beschäftigtenpanel der Bundesanstalt für Arbeit. Ein neues Scientific Use File. *ZA-Information* 52, 159-170 http://www.za.uni-koeln.de/publications/pdf/za_info/ZA-Info-52.pdf.
- Kölling, A. (2000): The IAB-Establishment Panel. *Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 120, 2, 291-300.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (2001): *Wege zu einer besseren informationellen Infrastruktur*. Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. Baden-Baden: Nomos.
- Oertel, M., Passenberger, J. & Janser, M. (2004): Forschungsdatenzentrum der BA geht an den Start. IAB-Kurzbericht 9/2004. <http://doku.iab.de/kurzber/2004/kb0904.pdf>.
- Sturm, R. (2002): Wirtschaftsstatistische Einzeldaten für die Wissenschaft. *Wirtschaft und Statistik*, 2, 101-109.

Dr. Annette Kohlmann: Studium der Soziologie und Politischen Wissenschaften an der Universität Mannheim; 1995 Diplom; wiss. Mitarbeiterin an der TU Chemnitz (2000 Promotion zum Dr. phil.), am Max-Planck-Institut für Demographische Forschung, Rostock (2000-2002), und am Zentrum für Umfragen, Methoden und Analysen (ZUMA), Mannheim (2002-2004), seitdem Schwerpunkte im Bereich der Datenprüfung, -aufbereitung und -dokumentation sowie der Beratung zu Daten, ab April 2004 am IAB als Leiterin des Forschungsdatenzentrums (gegenwärtige Hauptaufgabe: Aufbau des FDZ der BA, sonstige Forschungsschwerpunkte: Erwerbsverläufe von Müttern bzw. Wiedereintritt von Frauen in den Arbeitsmarkt nach der Geburt von Kindern).

Das **IAB** ist das Forschungsinstitut der **Bundesagentur für Arbeit** und vom Gesetzgeber mit der Aufgabe betraut, Wirkungen der Arbeitsförderung zu analysieren und interdisziplinäre Arbeitsmarktforschung zu betreiben (siehe <http://www.iab.de>).

2. FDZ-RV und Wissenschaft

In diesem zweiten Kapitel werden einige Forschungsvorhaben skizziert, die nur auf der Grundlage von Mikrodaten der Rentenversicherung möglich sind. Zunächst, in Abschnitt 2.1 werden erste Erfahrungen aus zwei laufenden Projekten mit Mikrodaten der Rentenversicherung dokumentiert. In Abschnitt 2.2 steht der zukünftige Datenbedarf der Wissenschaft im Vordergrund.

2.1 Erfahrungen der Wissenschaft mit Daten der RV

Erfahrungen mit Daten der Rentenversicherung wurden aus Kooperationen des VDR einerseits im Rahmen der AVID-Erhebung und andererseits mit dem Pilotprojekt Versichertenkontenstichprobe 2001 gewonnen:

Die Studie Altersvorsorge in Deutschland (AVID) wird im Auftrag von VDR und Bundesministerium für Gesundheit und Soziale Sicherung (BMGS) durchgeführt. Obwohl die Daten der AVID nicht im FDZ-RV verfügbar sein werden, sind die hier gewonnenen Erfahrungswerte im Umgang mit Versichertenkontendaten ein wichtiger Bezugspunkt für die Datenaufbereitung im FDZ-RV.

Das Pilotprojekt Versichertenkontenstichprobe 2001 wurde in Zusammenarbeit mit Prof. Martin Kukuk und Stefan Kempf von der Universität Würzburg gestartet. Hierbei wurden unter Einhaltung der datenschutzrechtlichen Regelungen Verfahren zum Zuschnitt von Datensätzen erprobt.

Erste Erfahrungen der Wissenschaft mit Daten der RV: Die Studie „Altersvorsorge in Deutschland“ (AVID)*

Dr. Thorsten Heien (TNS Infratest Sozialforschung, München)

1. Einleitung

Um dem steigenden Bedarf der Wissenschaft nach Daten im Bereich der Alterssicherung gerecht zu werden, baut die gesetzliche Rentenversicherung (GRV) derzeit ein Forschungsdatenzentrum (FDZ-RV) auf, das einen Zugang zu Mikrodaten der GRV auf verschiedenen Wegen – Public Use Files, Scientific Use Files, Gastwissenschaftler-Arbeitsplätze – ermöglichen soll (Rehfeld 2004, 73). Die in diesem Kontext interessierenden Erfahrungen von TNS Infratest Sozialforschung mit GRV-Daten werden im Folgenden am Beispiel der im Auftrag des Verbandes Deutscher Rentenversicherungsträger (VDR) und des Bundesministeriums für Gesundheit und Soziale Sicherung (BMGS) durchgeführten Untersuchung „Altersvorsorge in Deutschland“ (AVID) geschildert.¹ Die Darstellung beschränkt sich im Wesentlichen auf Erfahrungen mit „originären“ Rentenkontendaten im Rahmen der AVID (1996 und 2002) und lässt Erfahrungen mit der Auswertung der AVID-Daten außer Acht, da die hier primär interessierenden Daten der aktuellen AVID 2002 voraussichtlich erst im Sommer 2006 vorliegen werden (Roth et al. 2002, 640).

Zunächst wird jedoch ein Rückblick auf die „historischen“ Erfahrungen von TNS Infratest Sozialforschung mit GRV-Daten gegeben (Kapitel 2), dem sich eine kurze inhaltliche und methodische Vorstellung der AVID 2002 anschließt (Kapitel 3). Es folgt ein Überblick über den grundsätzlichen Umgang mit GRV-Versichertenkontendaten bei der AVID 2002 (Kapitel 4), ehe die konkreten Fragen der Kontenklärung durch die Rentenversicherungsträger (Kapitel 5) und der Klärung von Biographielücken durch TNS Infratest Sozialforschung im Rahmen der Datenprüfung (Kapitel 6) diskutiert werden. Abschließend werden ein kurzes Fazit zur Bedeutung der GRV-Daten für die AVID gezogen und die sich daraus ergebenden Schlussfolgerungen erörtert (Kapitel 7).

2. Erfahrungen von TNS Infratest Sozialforschung mit GRV-Daten

Die Erfahrungen von TNS Infratest Sozialforschung mit Daten der GRV beschränken sich nicht allein auf die AVID, sondern reichen historisch viel weiter zurück: Bereits vor 25 Jahren wurde im Rahmen der im Auftrag des damaligen Bundesministeriums für Arbeit und Sozialordnung (BMA) und in Kooperation mit dem Sonderforschungsbereich 3 „Mikroanalytische Grundlagen der Gesellschaftspolitik“ an den Universitäten Frankfurt am Main und Mannheim durchgeführ-

* Redaktionell überarbeitete Version eines Vortrags im Rahmen des Workshops „Forschungsrelevante Daten der Rentenversicherung: Aufbau des Forschungsdatenzentrums der gesetzlichen Rentenversicherung (FDZ-RV)“ des Verbandes Deutscher Rentenversicherungsträger, Würzburg, 28./29. Juni 2004. Die Angaben im Text beziehen sich auf den Stand im August 2004.

¹ Die Unternehmensbezeichnung „TNS Infratest Sozialforschung“ existiert erst seit Ende Januar 2004. Obwohl sich die vorliegenden Ausführungen auch auf weiter zurück liegende Zeitpunkte bzw. -räume beziehen, wird im Folgenden ausschließlich die aktuelle Bezeichnung verwendet.

ten Lebenslagenenerhebung („Grunddaten zur Entwicklung von Lebenslagen“) zur Ermittlung der Alterssicherungsanwartschaften der deutschen Wohnbevölkerung der Geburtsjahrgänge 1920 bis 1955 erstmals auf Rentenkontendaten der GRV zurückgegriffen (Infratest Sozialforschung 1986; vgl. **Abbildung 1**).

Dazu wurden die Befragten um die Teilnahme an einem Kontenklärungsverfahren gebeten, das zwischen März 1981 und Februar 1983 durchgeführt wurde. TNS Infratest Sozialforschung erhielt im Anschluss – mit Zustimmung der Befragungsperson – vom jeweiligen Rentenversicherungsträger eine Kopie der so genannten „Rentenauskunft“, die vercodet und ausgewertet wurde. Ergänzt wurden diese Informationen durch eine persönlich-mündliche Befragung von über 9.000 Personen sowie Arbeitgeberauskünften zur betrieblichen Altersversorgung (BAV). Im Rahmen der zwischen November 1980 und April 1981 durchgeführten Befragung wurden der Lebenslauf, berufliche und soziodemographische Merkmale der Person und des Haushalts sowie Anwartschaften in Sicherungssystemen außerhalb der GRV erhoben. Befragte, die als Arbeiter oder Angestellte in der Privatwirtschaft beschäftigt waren, erhielten nach Abschluss der mündlichen Befragung einen Zusatzfragebogen, der von ihrem Arbeitgeber beantwortet werden sollte. Diese Zusatzbefragung fand zwischen März und Oktober 1981 statt (vgl. Kiel 1984, Infratest 1986).

Abbildung 1: Erfahrungen von TNS Infratest Sozialforschung mit GRV-Daten

1) Lebenslagenuntersuchung (1979-1985; BMA)

- Aktuelle Alterssicherungsanwartschaften der deutschen Wohnbevölkerung der Geburtsjahrgänge 1920-1955
- Datenquellen:
 - a) Persönlich-mündliche Befragung von 9.097 Personen
 - b) GRV-Rentenkontendaten
 - c) Arbeitgeberauskünfte zur BAV

2) Alterssicherung in Deutschland 1986 (1986-1988; BMA)

- Einkommenssituation der Bevölkerung der Geburtsjahrgänge 1931 und früher
- Datenquellen:
 - a) Schriftliche Befragung von 29.364 Personen;
 - b) GRV-Rentendaten

3) Altersvorsorge in Deutschland 1996 (1995-2000; VDR/BMA)

- Alterssicherungsanwartschaften der deutschen GRV-Rentenversicherten der Geburtsjahrgänge 1936-1955 und ihrer Ehepartner (im 65. Lebensjahr)
- Datenquellen:
 - a) Persönlich-mündliche Befragung von 14.434 Personen
 - b) GRV-Rentenkontendaten
 - d) Fortschreibung der (Erwerbs-)Biographien

4) Altersvorsorge in Deutschland 2002 (2001-2006; VDR/BMGs)

- Vgl. Kapitel 3

Von 1986 an hat TNS Infratest Sozialforschung im Auftrag des BMA die erste Untersuchung zur „Alterssicherung in Deutschland“ (ASID '86) der Bevölkerung ab 55 Jahren (Geburtsjahrgänge 1931 und früher) durchgeführt,² die als Datenquelle für Informationen zur Art und Höhe von Einkommen auf der Personen- und Ehepartnerebene sowie zu den Determinanten der Alterseinkommen konzipiert war. Den Kern der Untersuchung bildete die Erfassung von 17 Einkommensarten, bei Ehepaaren für beide Ehepartner und bei Witwen differenziert nach eigenen und abgeleiteten Leistungen (Infratest Sozialforschung 1988).

Die Untersuchung basierte auf einer schriftlichen Befragung, in der Angaben zu insgesamt 29.364 Personen erhoben wurden.³ Die Befragten der ASID '86 wurden u.a. um ihr Einverständnis zur Übermittlung von GRV-Rentendaten durch den jeweiligen Versicherungsträger gebeten, die die Rentenart, das Jahr des Rentenbeginns, den monatlichen Zahlbetrag, den Krankenversicherungszuschuss sowie verschiedene Versicherungszeiten enthielten. Die von den Versicherungsträgern auf einem Formular übermittelten Informationen wurden von TNS Infratest Sozialforschung vercodet und ausgewertet (Infratest Sozialforschung 1988, 84).

Knapp zehn Jahre nach der ASID '86 führte TNS Infratest Sozialforschung im Auftrag von VDR und BMA die erste AVID durch. Ziel der AVID 1996 war die Verbesserung der Informationsbasis der GRV im Hinblick auf a) soziodemographische Merkmale der Versicherten, b) Lücken in den Versicherungsbiographien und Gründe für fehlende Beitragszeiten, c) die zeitliche Entwicklung der Biographien, d) künftige Anwartschaften von Frauen und Männern (auch auf Ehepaarebene) und e) die Kumulation von GRV-Anwartschaften mit Anwartschaften aus anderen Alterssicherungssystemen (vgl. Kortmann & Schatz 1999, Roth et al. 2002).

Die AVID 1996 ging aus der Lebenslagenerhebung hervor, unterschied sich aber von ihr zum einen durch eine andere Grundgesamtheit: Einbezogen wurden die damals 40- bis unter 60-jährigen (Geburtsjahrgänge 1936 bis 1955) deutschen GRV-Versicherten und – unabhängig von Alter, Nationalität und GRV-Konto – ihre Ehepartner (Kortmann und Schatz 1999; vgl. **Abbildung 1**). Zum anderen wurden bei der AVID 1996 nicht nur aktuelle, sondern auch zukünftige Alterssicherungsanwartschaften betrachtet. Dies machte eine neue Datenquelle nötig: Neben der persönlich-mündlichen Befragung von 14.434 Personen und der Klärung der GRV-Rentenkonto wurden deshalb die Lebens- und Erwerbsverläufe der Befragten mittels eines eigens entwickelten Mikrosimulationsmodells bis zum jeweiligen 65. Lebensjahr (als Regelaltersgrenze der GRV) fortgeschrieben (vgl. Schatz et al. 2002).⁴

Als vierte und (bisher) letzte Erfahrung von TNS Infratest Sozialforschung mit „originären“ GRV-Daten ist schließlich die aktuelle AVID 2002 zu nennen, die derzeit im Auftrag von VDR und BMGS durchgeführt und deshalb im Folgenden ausführlicher vorgestellt wird.

2 Auch in den Jahren 1992, 1995, 1999 und 2003 wurde TNS Infratest Sozialforschung mit der Durchführung der ASID betraut.

3 Darunter befanden sich auch Angaben zu 2.169 bereits verstorbenen Ehemännern (Infratest Sozialforschung 1988, 10).

4 Informationen zur betrieblichen Altersversorgung wurden bei der AVID 1996 im Rahmen der persönlich-mündlichen Befragung erhoben.

3. „Altersvorsorge in Deutschland 2002“ (AVID 2002)

Gegenüber der Vorgängeruntersuchung wurde bei der AVID 2002 die Grundgesamtheit ausgeweitet und zwar auf die deutsche Wohnbevölkerung (der Geburtsjahrgänge 1942 bis 1961) und – unabhängig von Alter, Nationalität und GRV-Konto – ihre Ehepartner. Dies geschah nicht zuletzt deshalb, weil die mangelnde Bevölkerungsrepräsentativität ein Kritikpunkt an der AVID 1996 gewesen war. Durch die veränderte Grundgesamtheit war für die AVID 2002 eine Stichprobe auf Basis der Versicherungskontenstichprobe des VDR (wie noch 1996) nicht mehr möglich. Stattdessen wurde zwischen Oktober und Dezember 2002 eine repräsentative Stichprobe aus dem Access-Panel von TNS Infratest TPI schriftlich befragt.⁵ Dabei handelt es sich um einen großen Pool befragungsbereiter Haushalte,⁶ für die aufgrund von Vorbefragungen und gezielten Screening-Befragungen detaillierte und aktuelle soziodemographische Informationen zur Verfügung stehen. Access-Panels finden in der empirischen Sozialforschung bisher eher selten Verwendung, mittlerweile plant allerdings z.B. das Statistische Bundesamt angesichts kontinuierlich sinkender Ausschöpfungsquoten bei Befragungen auf der Grundlage anderer Stichprobenarten den Aufbau eines solchen Panels auf Basis des Mikrozensus (vgl. Bechthold et al. 2002, Lotze und Breiholz 2002).

Neben der hohen Befragungsbereitschaft sprachen im Rahmen der AVID 2002 aus methodischer Sicht – im Vergleich zur Alternative einer Einwohnermelderegisterstichprobe (Roth et al. 2002, 635) – für das Access-Panel vor allem:

- die höhere Qualität der ausgefüllten Fragebögen,
- die im Hinblick auf die für 2004 geplante Nachbefragung wichtige Vertrautheit der Panel-Teilnehmer mit Mehrfachbefragungen,
- die für die Nacherhebung von Angaben im Rahmen der Datenprüfung wichtige Verfügbarkeit von Telefonnummern von annähernd 95% der Befragungsteilnehmer sowie
- die bessere Erreichbarkeit von in Bevölkerungsumfragen unterrepräsentierten Gruppen durch die Ziehung einer Stichprobe unter bestimmten Quotierungsvorgaben.

Im Übrigen hat eine kritische Evaluation der soziodemographischen Strukturen des Access-Panels bzw. der daraus gezogenen Bruttostichprobe in der Regel eine gute Übereinstimmung mit den Strukturen in anderen Erhebungen (Mikrozensus, Sozio-oekonomisches Panel etc.) gezeigt. In Verbindung mit den Kostenvorteilen gegenüber einer Befragung auf Basis einer Einwohnermelderegisterstichprobe wurde deshalb für die AVID 2002 der methodische Nachteil einer nicht echten Zufallsstichprobe in Kauf genommen.⁷

⁵ Um den Fragebogen und den Ablauf der Befragung ausgiebig zu testen, wurde zwischen Mai und Juli 2002 ein Pretest mit 534 Befragten durchgeführt.

⁶ Im Oktober 2002 gehörten dem Access-Panel von TNS Infratest TPI 71.865 Haushalte mit insgesamt 164.986 Personen an (NFO TPI Test-Panel-Institut 2002, 1).

⁷ Allerdings darf in diesem Zusammenhang nicht unerwähnt bleiben, dass auch Nettostichproben von Befragungen auf Basis von Einwohnermelderegisterstichproben aufgrund vollständiger („unit non-response“) oder teilweiser Verweigerungen („item non-response“) Verzerrungen unterliegen. Angesichts der Komplexität des Fragenprogramms der AVID 2002 war davon auszugehen, dass sich derartige Verzerrungen nicht zufällig über alle Befragten verteilen würden.

Im Mittelpunkt der schriftlichen Befragung der AVID 2002 standen Informationen zur Soziodemographie (Familienstand, Beruf, Bildung, Ausbildung, Einkommen etc.), zur Erwerbsbiographie sowie zur Absicherung außerhalb der GRV in Regel- und Zusatzversicherungssystemen (betriebliche Altersversorgung, Zusatzversorgung des öffentlichen Dienstes, Beamtenversorgung, landwirtschaftliche Alterssicherung, berufsständische Versorgung) und durch private Vorsorge (private Lebens- und Rentenversicherungen, Wohneigentum, nicht jedoch weitere Vermögensanlagen). Da die Befragung des Jahres 2002 aufgrund des kurzen Zeitraums seit Inkrafttreten der neuen Fördermöglichkeiten der privaten Vorsorge gemäß Altersvermögensgesetz (AvmG) und Altersvermögensergänzungsgesetz (AvmEG) diesbezüglich nur einen vorläufigen Eindruck liefern konnte, findet im Sommer 2004 eine Nachbefragung zu Art und Höhe der seit Herbst 2002 vereinbarten zusätzlichen betrieblichen und privaten Vorsorge sowie sonstiger privater Vorsorgemaßnahmen statt. Die Klärung der GRV-Konten wurde bereits im Mai 2004 abgeschlossen (vgl. Kapitel 4).

Im Rahmen der seit Juli 2003 laufenden Einzelfallprüfung erfolgen durch ein Team von neun Editoren pro befragter Person über 1.000 Prüfungen innerhalb der Befragungsdaten und im Abgleich von Befragungs- und Versicherungskontendaten. So werden zum Beispiel die Filterführungen, zulässige Codes, zulässige Wertebereiche, Inkonsistenzen zwischen demographischen Angaben sowie Inkonsistenzen zwischen Einkommensarten und demographischen Angaben geprüft. Als besonders prüfintensiv haben sich der Lebenslauf der Befragten und die verschiedenen Versicherungsverträge (neue staatlich geförderte Altersvorsorge, Kapitallebensversicherungen, private Rentenversicherungen) erwiesen. Bei letzteren bereiten den Befragten vor allem die detaillierten Angaben Probleme. Zudem werden häufig Monats- mit Jahresangaben und Euro- mit DM-Angaben verwechselt sowie Verträge an der falschen Stelle im Fragebogen eingetragen.

Im Anschluss an die Datenprüfung werden die Lebens- und Erwerbsverläufe mit Hilfe des bereits bei der AVID 1996 verwendeten, mittlerweile aber weiterentwickelten Mikrosimulationsmodells bis zum 65. Lebensjahr der Befragten statistisch fortgeschrieben und auf der Basis des jeweils geltenden Sozial- und Steuerrechts die Nettogesamteinkommen berechnet.⁸ Abgeschlossen wird die Untersuchung durch eine ausführliche Dokumentation der Ergebnisse in Form von Mikrodaten, Tabellen und Berichten (Roth et al. 2002: 629 ff.).⁹

4. Umgang mit GRV-Versichertenkontendaten in der AVID

Nun zum primär interessierenden Umgang mit den GRV-Versichertenkontendaten im Rahmen der AVID 2002, der in Abbildung 2 einschließlich der Aufgabenverteilung auf die verschiedenen beteiligten Institutionen – TNS Infratest TPI, TNS Infratest Sozialforschung, Datenstelle der Rentenversicherungsträger (DSRV), 26 Rentenversicherungsträger – im Überblick dargestellt ist. Demnach wurden im Rahmen der schriftlichen Befragung durch TNS Infratest TPI die Befrag-

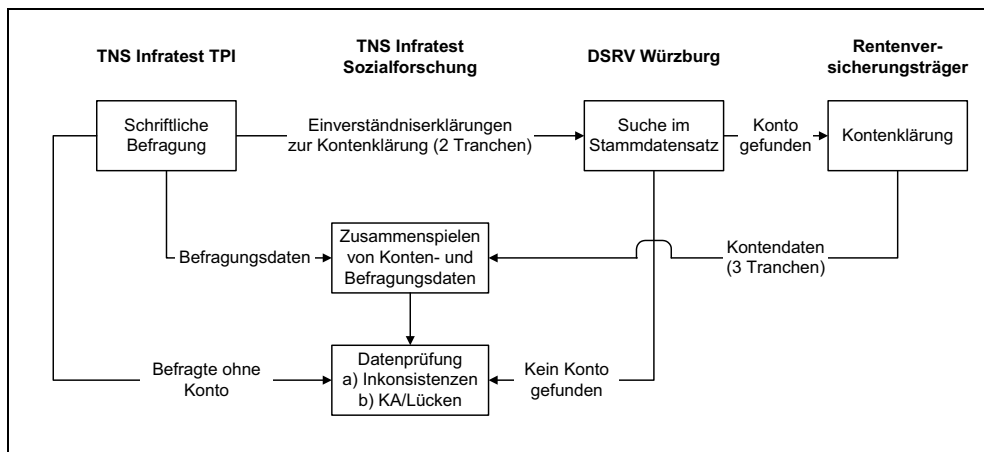
⁸ Als wesentliche Neuerungen des Modells sind zu nennen: a) die Berücksichtigung von Einkommen oberhalb der Beitragsbemessungsgrenze, b) der Längsschnittbezug des Arbeitszeitmodells, c) die detailliertere Schätzung von Leistungen der betrieblichen Zusatzversorgung und d) der längere empirische Stützzeitraum (2002: 1992 bis 2002; 1996: 1992 bis 1996) für die Fortschreibung der Biographien (Roth et al. 2002, 638).

⁹ Weitere Informationen zur AVID (1996 und 2002) finden sich im Internet unter: www.altersvorsorge-in-deutschland.de.

ten zunächst um ihre „Einverständniserklärung zur Kontenklärung und zum Zusammenspielen von Konten- und Befragungsdaten“ gebeten. Das entsprechende Formular enthielt Angaben zu Namen, Geburtstag und -ort, Anschrift und die Rentenversicherungsnummer und wurde im Rücklauf sofort von den Fragebogen getrennt.¹⁰

Die Daten der erfassten Einverständniserklärungen wurden in zwei zeitversetzten Tranchen – bei der ersten Tranche im November 2002 7.396 Fälle und bei der zweiten Tranche im Februar 2003 7.278 Fälle – an die DSRV geschickt, wo die dazugehörigen Konten im Stammsatzbestand gesucht wurden, zunächst über die Rentenversicherungsnummer und dann über die weiteren Angaben.¹¹ Die insgesamt 13.746 Fälle, in denen ein Versicherungskonto gefunden werden konnte, wurden schließlich von der DSRV zwecks Einleitung eines Kontenklärungsverfahrens an die jeweiligen Rentenversicherungsträger übergeben. Dabei handelt es sich um dasselbe Verfahren, das von den Versicherten vor dem Renteneintritt durchgeführt werden muss.

Abbildung 2: Umgang mit GRV-Versichertenkontendaten in der AVID 2002



Quelle: eigene Darstellung.

Die Daten der geklärten Konten wurden von den Rentenversicherungsträgern in drei zeitversetzten Tranchen (März 2003, September 2003, Mai 2004) an TNS Infratest Sozialforschung geliefert,¹² wo sie mit den Befragungsdaten zusammengespielt wurden. In den Gesamtdatensatz, der derzeit immer noch intensiv geprüft wird (vgl. Kapitel 3), gehen darüber hinaus die Fälle ein, in denen kein Konto gefunden wurde (obwohl sie den Befragungsangaben nach eines haben müssten),¹³ und die Fälle ohne Konto.

10 Für Details des Datenschutzes im Rahmen der AVID vgl. Roth et al. (2002, 619).

11 Bei der ersten Tranche gelang dies in 94,9% aller Fälle (N = 6.682). Die restlichen 5,1% aller Fälle, für die kein Konto gefunden wurde (N = 358), gingen zur erneuten Überarbeitung an TNS Infratest Sozialforschung zurück. Im Rahmen der zweiten Tranche konnte 97,1% aller Fälle ein Versicherungskonto zugewiesen werden (N = 7.064), in den restlichen 2,9% der Fälle gelang dies nicht (N = 214).

12 Die Bundesversicherungsanstalt für Angestellte (BfA) hat als einziger Träger im März 2004 eine zusätzliche Tranche geliefert.

13 Bei diesen Fällen handelt es sich um potenzielle Ausfälle, deren genauer Status im Rahmen der Datenprüfung geklärt werden muss.

Im Anschluss an Schilderung des grundsätzlichen Umgangs mit den GRV-Versichertenkonten im Rahmen der AVID 2002 sollen nun noch einige Zahlen zur Bilanz a) der Kontenklärung durch die Rentenversicherungsträger (Kapitel 5) und b) der Klärung von Biographielücken durch TNS Infratest Sozialforschung im Rahmen der Datenprüfung (Kapitel 6) präsentiert werden.

5. Klärung von GRV-Versichertenkonten in der AVID

In **Tabelle 1** werden die Erfolge der Kontenklärung durch die Rentenversicherungsträger im Rahmen der AVID 1996 und 2002 (soweit wie möglich) miteinander verglichen. Dabei ist zu berücksichtigen, dass sich die Fallzahl für 2002 auf den Stand vor der (derzeit noch nicht abgeschlossenen) Datenprüfung bezieht, die Fallzahl für 1996 dagegen auf den Stand nach Abschluss der Datenprüfung. Deshalb wird die endgültige Fallzahl der AVID 2002 geringer sein als die in **Tabelle 1** ausgewiesenen 14.545 Fälle.

Der Anteil der Fälle ohne Konto lag zunächst einmal 1996 mit 2,3% deutlich niedriger als 2002 (5,5%), weil sie nur als Ehepartner (eines GRV-Versicherten) in die Stichprobe gelangen konnten, während dies 2002 auch als Zielperson möglich war.¹⁴ Von den verbleibenden 14.100 Konten konnten bei der AVID 1996 fast 90%, nämlich 12.406 von den Trägern geklärt werden. Dies ist ein außerordentlich hoher Wert, berücksichtigt man a) die „Umstellungsarbeiten“ im Zuge der deutsch-deutschen Vereinigung und b) den großen Aufwand einer Kontenklärung für die Betroffenen. Letzteres gilt vor allem für die neuen Bundesländer, wo zum Zeitpunkt der Kontenklärung der AVID 1996 zahlreiche Versichertenkonten noch nicht maschinell geführt worden waren (Infratest Burke Sozialforschung 2000, 46). Dementsprechend entfiel ein überproportionaler Anteil der 11% nicht (abschließend) geklärten Konten auf die neuen Bundesländer (1.090

Tabelle 1: Bilanz der Klärung von GRV-Versichertenkonten 1996 und 2002

	AVID 1996			AVID 2002		
	abs.	%	%	abs.	%	%
Befragte*	14.434	100,0		14.545	100,0	
Kein Konto	334	2,3		799	5,5	
Konto	14.100	97,7	100,0	13.746	94,5	100,0
Geklärt	12.406		88,0	...**		...**
Nicht geklärt	1.609		11,4	...**		...**
Leer	85		0,6	...**		...**

Quelle: eigene Darstellung

Anmerkung: * 1996: nach Datenprüfung; 2002: vor Datenprüfung; ** steht noch nicht fest.

¹⁴ Allerdings kann sich durch Ausfälle im Rahmen der Datenprüfung die Relation von Fällen mit Konto zu Fällen ohne Konto für die AVID 2002 noch ändern.

von 1.609; nicht in Tabelle 1 ausgewiesen). Diese quantitativen Unterschiede gingen einher mit qualitativen Unterschieden: Während in den neuen Bundesländern vor allem Beitragszeiten nicht geklärt werden konnten, handelte es sich in den alten Bundesländern im Wesentlichen um nicht geklärt Anrechnungszeiten (Infratest Burke Sozialforschung 2000, 48). 85 Konten (0,6%) erwiesen sich schließlich im Rahmen der Kontenklärung der AVID 1996 als leer.

Obwohl die Kontendaten für die AVID 2002 mittlerweile vollständig vorliegen (**vgl. Tabelle 1**), kann aufgrund der noch laufenden Datenprüfung der endgültige Anteil geklärter Konten noch nicht bestimmt werden. Derzeitigen Prognosen zufolge wird dieser Anteil allerdings über 90% liegen, was eine Steigerung gegenüber dem bereits sehr guten Ergebnis von 1996 und dementsprechend eine noch höhere Datenqualität bedeuten würde.

6. Klärung von Biographielücken in der AVID

Abschließend wollen wir die Bilanz der Klärung der Lücken in den GRV-Konten durch TNS Infratest Sozialforschung bei der AVID 1996 betrachten. Betrachtet wurde im Rahmen der Lückenklärung der Zeitraum vom 15. Geburtstag der Befragten bis zum Dezember 1995 (als Befragungszeitpunkt; **vgl. Tabelle 2**).¹⁵ Bei einem unterstellten Durchschnittsalter von 50 Jahren (im Dezember 1995) lagen für jeden Befragten im Durchschnitt 35 (= 50-15) Biographiejahre vor. Multipliziert man diese Zahl mit 12 (Monaten) und der Anzahl der Befragten, resultiert die Anzahl der vorliegenden Biographiemonate bis zum Befragungszeitpunkt. In den alten Bundesländern waren dies bei der AVID 1996 ca. 4 Mio. Monate, in den neuen Bundesländern ca. 2 Mio. Monate (**vgl. Tabelle 2**).

Im Anschluss an die Kontenklärung durch die Rentenversicherungsträger waren in den alten Bundesländern bei der AVID 1996 etwas mehr als drei Viertel (75,6%) dieser Monate durch rentenrechtliche Zeiten abgedeckt, in den neuen Bundesländern sogar etwas mehr als vier Fünf-

Tabelle 2: Bilanz der Klärung von Biographielücken 1996

	Alte Bundesländer		Neue Bundesländer	
	Tsd.	%	Tsd.	%
Monate insg.	3.927,6	100,1	1.933,5	100,0
GRV geklärt (= Kontenklärung)	2.968,7	75,6	1.554,3	80,4
AVID geklärt (= Datenprüfung)	834,7	21,3	71,1	3,7
Nicht geklärt	124,2	3,2	308,1	15,9

Quelle: eigene Darstellung.

Anmerkung: * ab dem 16. Lebensjahr bis Dezember 1995.

¹⁵ Es sollten nur Lücken von einem Jahr und länger geklärt werden.

tel (80,4%). Letzteres ist im Wesentlichen auf folgende Charakteristika von Erwerbstätigkeiten bzw. -biographien in der ehemaligen DDR zurückzuführen: a) die höheren Erwerbsquoten (vor allem von Frauen), b) die fast durchgehende Sozialversicherungspflicht von Tätigkeiten und c) die schnellere Möglichkeit zur Rückkehr von Frauen an ihren Arbeitsplatz nach der Geburt von Kindern (Infratest Burke Sozialforschung 2000: 58).¹⁶

In den alten Bundesländern konnten weitere 21,3% der Biographiemonate im Rahmen der Datenprüfung durch TNS Infratest Sozialforschung geklärt werden, in den neuen Bundesländern dagegen nur 3,7%. Letzteres ist im Wesentlichen auf den hohen Anteil nicht erfasster Beitragszeiten zurückzuführen, die sich in Tabelle 2 als nicht geklärte Monate wieder finden (neue Bundesländer: 15,9%).¹⁷ In den alten Bundesländern konnten letztendlich nur 3,2% der Biographiemonate nicht geklärt werden. Ursächlich waren hier zum einen Erinnerungslücken bei den befragten Personen, zum anderen die Tatsache, dass die betreffenden Personen im Rahmen der telefonischen Nachbefragung nicht erreicht werden konnten bzw. über keinen Telefonschluss verfügten.

Für die AVID 2002 ist, so viel lässt sich derzeit bereits sagen, mit einem noch niedrigeren Anteil nicht geklärter Biographiemonate zu rechnen. Dies ist vor allem auf a) die im vornherein große Anzahl verfügbarer Telefonnummern von Befragungsteilnehmern (vgl. Kapitel 3) und b) die – im Vergleich zu erstmals für eine derartige Befragung ausgewählten Personen – größere Kooperationsbereitschaft der Teilnehmer des Access-Panels zurückzuführen. Das Problem eines deutlich höheren Anteils nicht geklärter Biographiemonate in den neuen Bundesländern besteht weiterhin, lässt sich aufgrund der nicht abgeschlossenen Datenprüfung aber noch nicht endgültig quantifizieren.

7. Fazit und Schlussfolgerungen

Als Fazit der Schilderung der Erfahrungen von TNS Infratest Sozialforschung mit GRV-Daten im Rahmen der Untersuchung „Altersvorsorge in Deutschland“ ist zum einen die zentrale Bedeutung der Versichertenkontendaten für die AVID zu konstatieren. Dies schlägt sich einerseits in dem hohen Verbreitungsgrad von GRV-Anwartschaften in der Bevölkerung nieder (vgl. Tabelle 1). Andererseits zeigen sich im Rahmen der Datenprüfung bisweilen deutliche Abweichungen zwischen den Versichertenkontendaten und den Angaben der Befragten im Rahmen der schriftlichen Befragung und hier vor allem zu ihrem Lebenslauf. Diese Abweichungen sind wohl im Wesentlichen auf Erinnerungslücken bei den befragten Personen zurückzuführen. Daraus folgt, dass im Interesse valider Ergebnisse eine AVID ohne die Integration von Versichertenkontendaten undenkbar ist.

¹⁶ Der Anteil der in den neuen Bundesländern in der AVID 1996 durch rentenrechtliche Zeiten abgedeckten Monate wäre allerdings bei einer erfolgreicherer Kontenklärung durch die Rentenversicherungsträger (vgl. Kapitel 5) noch höher ausgefallen.

¹⁷ Da Lücken in den Erwerbsbiographien (vor allem der Männer) in der ehemaligen DDR sehr selten waren, wurden längere Lücken (30 Monate und mehr) in den neuen Bundesländern vor 1992 im Rahmen der Analysen als sozialversicherungspflichtige Erwerbstätigkeitsepisoden behandelt (Infratest Burke Sozialforschung 2000: 131).

Zum anderen bleibt aus der Schilderung der Erfahrungen von TNS Infratest Sozialforschung festzuhalten, dass die aktuelle AVID 2002 im Hinblick auf die Konten- (vgl. Kapitel 5) und die Lückenklärung (vgl. Kapitel 6) noch erfolgreicher sein wird als ihre Vorgängerin. Dies bedeutet eine noch höhere Datenqualität als bei der AVID 1996, von der nicht zuletzt auch die mit den Daten arbeitende Wissenschaft profitiert.

Literatur

- Bechthold, S., Müller, A. & Pötsch, O. (2002): Ein Access-Panel als Auswahlgrundlage für Haushalts- und Personenerhebungen ohne Auskunftspflicht. In: *Wirtschaft und Statistik*, 5/2002, 345-358.
- Infratest Burke Sozialforschung (2000): Altersvorsorge in Deutschland 1996 (AVID '96) - Methodenbericht: Phasen I und II: Datenerhebung und -aufbereitung, Fortschreibung der Biographien und Berechnung der Alterseinkommen. In: Verband Deutscher Rentenversicherungsträger und Bundesministerium für Arbeit und Sozialordnung (Hg.): *CD-ROM zur Deutschen Rentenversicherung* – Schriftenreihe Band 19 und BMA-Forschungsbericht Nr. 277, Frankfurt a. M./Bonn.
- Infratest Sozialforschung (1986): Zukünftige Rentnergeneration - Anwartschaften in der Alterssicherung der Geburtsjahrgänge 1920-1955. Untersuchungsbericht zum Forschungsprojekt „Grunddaten zur Entwicklung von Lebenslagen“. *Forschungsbericht 142 des Bundesministeriums für Arbeit und Sozialordnung*.
- Infratest Sozialforschung (1988): Alterssicherung in Deutschland 1986 – Methodenbericht. In: Bundesministerium für Arbeit und Sozialordnung (Hg.): *BMA-Forschungsbericht 200-M*, Bonn.
- Kiel, W. (1984): *Datenhandbuch zur Lebenslagen-Studie. Haupterhebung und Rentenbescheide*. Arbeitspapier des Sonderforschungsbereichs 3, Frankfurt am Main.
- Kortmann, K. & Schatz, C. (1999): Altersvorsorge in Deutschland 1996 (AVID '96) – Lebensverläufe und künftige Einkommen im Alter. Verband Deutscher Rentenversicherungsträger (Hg.): *DRV-Schriften* Band 19; und Bundesministerium für Arbeit und Sozialordnung (Hg.): *BMA-Forschungsbericht* Band 277, Frankfurt/München.
- Lotze, S. & Breiholz, H. (2002): Zum neuen Erhebungsdesign des Mikrozensus. In: *Wirtschaft und Statistik*, 5/2002, 359-366.
- NFO TPI TestPanel-Institut (2002): *Access Panel Statistics* Oktober 2002, Wetzlar.
- Rehfeld, U. (2004): Datenangebot und Informationsbedarf im Bereich der Alterssicherung – Forschungsdatenzentrum der Rentenversicherung im Aufbau. In: *Deutsche Rentenversicherung*, 1-2, 63-75.
- Roth, M., Stegmann, M. & Bieber, U. (2002): Die Aktualisierung der Studie Altersvorsorge in Deutschland. Inhaltliche und methodische Neuerungen der AVID 2002. In: *Deutsche Rentenversicherung*, 11, 612-641.
- Schatz, C., Merz, J. & Kortmann, K. (2002): Künftige Alterseinkommen. Eine Mikrosimulationsstudie zur Entwicklung der Renten und der Altersvorsorge in Deutschland (AVID '96). In: *Schmollers Jahrbuch*, 122, 227-260.

Dr. Thorsten Heien, Dipl.-Soz.: Von 1998 bis 2001 Wissenschaftlicher Mitarbeiter in zwei Forschungsprojekten zur sozialpolitischen Einstellungsforschung an der Fakultät für Soziologie der Universität Bielefeld. 2002 Promotion bei Prof. Dr. Hans-Jürgen Andreß („Wohlfahrtsansprüche in Europa“). Seit 2001 Projektleiter (Senior Research Manager) bei TNS Infratest Sozialforschung im Bereich „Sozialpolitische Forschung“. Dort verantwortlich für Untersuchungen zur Altersvorsorge und -sicherung (u.a. „Altersvorsorge in Deutschland 2002“, „Spar- und Finanzanlageverhalten privater Haushalte 2001, 2003“) sowie zur Armuts- und Sozialhilfeforschung (u.a. „Niedrigeinkommens-Panel“).

TNS Infratest Sozialforschung (München) gehört zu TNS Infratest innerhalb des weltweit tätigen Marktforschungsunternehmens TNS (Taylor Nelson Sofres, London). Neben der Markt- und Medienforschung betreibt TNS Infratest bereits seit den 1950er Jahren Sozial- und Politikforschung. 1975 wurde dafür eine eigene Gesellschaft gegründet, die Infratest Sozialforschung GmbH, die heute auf dem Gebiet der sozialwissenschaftlichen Umfrageforschung das führende privatwirtschaftliche Institut in Deutschland ist (siehe www.tns-infratest-sofo.com).

Determinanten des Renteneintritts – eine Untersuchung mit prozessproduzierten Daten

Stefan Kempf und Prof. Dr. Martin Kukuk (Universität Würzburg)

1. Einleitung

Die Frage, warum ein Versicherter zu einem bestimmten Zeitpunkt aus dem Erwerbsleben ausscheidet, ist insbesondere bei Rentensystemen interessant, die ein Wahlrecht des Renteneintritts einräumen. Das deutsche System gewährt solche Wahlrechte, die in verschiedenen Reformen der letzten 20 Jahre zum Teil gravierende Änderungen erfahren haben. Ein Ziel dieser Untersuchung ist es, die Anreizwirkungen der Wahlrechte empirisch zu untersuchen. Im Mittelpunkt steht die gesetzliche Rentenversicherung, die nach wie vor eine Hauptsäule der finanziellen Absicherung im Rentenalter darstellt, auch wenn die private und betriebliche Altersvorsorge in der Zukunft zunehmende Bedeutung gewinnen werden.

Die gesetzlichen Bestimmungen für einen Renteneintritt sehen neuerdings vor, dass ein Arbeitnehmer im Alter zwischen 62 und 65 Jahren in Rente gehen kann, wenn bestimmte Bedingungen erfüllt sind. Ab 62 Jahren können Versicherte mit einer sehr langen Erwerbsbiographie und ab 65 Jahren alle Erwerbstätigen nach einer Wartezeit von mindestens 5 Jahren in Rente gehen. Für das Zugangsjahr 2001, für das im Rahmen dieser Arbeit die Determinanten der Verrentung untersucht werden, kommt hinzu, dass auch Versicherte ab 60 Jahren nach einer Phase der Arbeitslosigkeit oder Altersteilzeit in Rente gehen konnten. Die genauen gesetzlichen Bestimmungen sind im 6. Sozialgesetzbuch festgelegt.

Innerhalb dieses gesetzlichen Rahmens können Arbeitnehmer über ihren Renteneintritt relativ frei entscheiden: Nach den alten Regelungen konnten Versicherte mit einer langen Erwerbsbiographie ab 60 über Altersteilzeit das Verlassen des Erwerbslebens frühzeitig planen oder mit 62 in Rente gehen. Dagegen sind Dinge wie Arbeitslosigkeit nicht auf konkrete Überlegungen vom Versicherten zurückzuführen, sondern mehr auf ein unfreiwilliges Ereignis wie etwa den Abbau von Arbeitskräften durch den Arbeitgeber. Im Krankheitsfall kann ein Versicherter optional auch eine Rente wegen Erwerbsminderung beantragen. Krankheiten können zwar nicht als echte Wahlmöglichkeiten angesehen werden, dennoch steht es den Betroffenen eventuell offen, sich eingeschränkt weiter am Arbeitsleben zu beteiligen oder aber aus dem Erwerbsleben auszuscheiden. Die Regelaltersrente eröffnete schließlich allen Versicherten, die sich einen Anspruch auf eine Rente erworben haben, den Weg in die Altersrente.

Neben den gesetzlichen Rahmenbedingungen und den damit verbundenen Wahlmöglichkeiten gibt es Entscheidungskriterien, nach denen sich ein Ausstieg aus dem Erwerbsleben richtet. Gesundheitliche Probleme oder Episoden von Arbeitslosigkeit im höheren Alter müssen nicht unbedingt gleich zur Beendigung des Erwerbslebens führen, können aber ein Beweggrund dafür sein, den Verrentungszeitpunkt früher zu wählen. Finanzielle Überlegungen spielen sicherlich eine gewichtige Rolle. Wird ein früher Zeitpunkt für die Verrentung gewählt, erhält man die über die Beiträge erworbene Rentenzahlung abzüglich eventueller Abschläge. Man gibt zusätzlich künftiges Arbeitseinkommen und damit verbunden eine höhere Rentenanwartschaft

auf. Dies wird mit dem Konzept des Optionswertes (Stock & Wise 1990) beurteilt. Darüber hinaus kann die Situation des Ehepartners ebenso mit ein Grund für das Verlassen des Erwerbslebens sein wie die finanzielle private und/oder betriebliche Absicherung.

Der Optionswert wurde u.a. bereits in Schnabel (2003), Börsch-Supan (1992) und Schmidt (1994) verwendet. Datengrundlage für diese Schätzungen ist das Sozio-Ökonomische Panel. Der Nachteil dieser Daten ist, dass präzise Angaben über die Einkommenshistorie nicht vorhanden sind und Informationen zu den erworbenen Rentenansprüchen gänzlich fehlen, die behelfsweise abgeschätzt werden müssen. In diesem Beitrag verwenden wir prozessproduzierte Daten des Verbands Deutscher Rentenversicherungsträger (VDR), die präzise Informationen über das sozialversicherungspflichtige Einkommen und die erworbenen Rentenansprüche enthalten. Damit lassen sich potenzielle Arbeitseinkommen in der Zukunft relativ genau schätzen, so dass sich die finanziellen Aspekte der Rentenentscheidung in enger Analogie zu Stock und Wise (1990) für das deutsche Rentenversicherungssystem empirisch abschätzen lassen. Der Verwendung prozessproduzierter Informationen sind in Deutschland allgemein enge Grenzen durch den gesetzlich geregelten Schutz individueller Daten gesetzt. Die Erhebung und Speicherung der Prozessdaten beispielsweise bei Sozialversicherungsanstalten und den Finanzverwaltungen verfolgen einen eigenständigen Zweck, weshalb der Zugang und die Verwertung dieser Daten äußerst restriktiv gehandhabt wird. Für die wissenschaftliche Forschung sind diese Informationen aber von immenser Bedeutung, um wirtschaftspolitisch relevante Antworten geben zu können. Unsere Arbeit soll hierzu einen Beitrag liefern und aufzeigen, welche Analysemöglichkeiten mit bisher für die Wissenschaft unzugänglichen Individualdaten bestehen.

Im weiteren Verlauf wird das zugrunde liegende einperiodige Optionswertmodell dargestellt. Anschließend wird im Abschnitt 3 der für die Schätzung verwendete Datensatz erläutert. Im Abschnitt 4 werden die untersuchten Schätzmodelle vorgestellt sowie die daraus resultierenden Ergebnisse im Abschnitt 5 präsentiert. Das Fazit im Abschnitt 6 schließt den Beitrag ab.

2. Das Optionswertmodell

Um die Anreizwirkungen der gesetzlichen Rentenversicherung zu beurteilen, können die zukünftig zu erwartenden Einkommensströme auf einen bestimmten Zeitpunkt abdiskontiert werden. Bei einer kurzfristigen Betrachtung werden die Kapitalwerte zweier aufeinander folgenden Renteneintrittsjahre miteinander verglichen. Der Renteneintritt wird aufgeschoben, wenn der zweite Wert den ersten übersteigt. Der generelle intertemporale Ansatz wählt den Zeitpunkt, der den maximalen Kapitalwert aufweist.

Der Mangel dieses Konzepts besteht darin, dass die Annahme zugrunde liegt, dass ein Euro in Rente denselben Wert besitzt, wie ein Euro in der Erwerbsphase. Nicht berücksichtigt ist beispielsweise der größere Umfang an Freizeit, der einen Euro in Rente wertvoller erscheinen lässt. In direktem Zusammenhang steht dabei auch, dass der Konsum von Gütern eine lineare Funktion vom Einkommen ist, nicht aber der Nutzen, den die Konsumenten daraus ziehen. Außerdem unterstellt das Kapitalwertkonzept in ihrem Verhalten risikoneutrale Versicherte. Zu beobachten ist aber, dass die Versicherten eine Glättung ihrer Einkommen während der Renten- und Erwerbsphase anstreben. Aus diesen Gründen werden nun keine Einkommens- und Renten-

niveaus betrachtet, sondern verschiedene Nutzenfunktionen in der Erwerbs- und Rentenphase.

Daraus resultieren aufgrund einer einperiodigen oder intertemporalen Betrachtung zwei weitere Berechnungsmöglichkeiten für einen Optionswert, die das Wahlverhalten besser beschreiben sollten als bei einer reinen Kapitalwertbetrachtung.

Der Nutzen zum Betrachtungszeitpunkt s , den ein Individuum bei einem Renteneintritt im Alter r bis zu seinem Todeszeitpunkt T realisieren kann, ergibt sich durch die Summe der auf den Zeitpunkt s mit dem Kapitalmarktzins i abdiskontierten Nutzenniveaus:

$$V_s(r) = \sum_{t=s}^{r-1} \left(\frac{1}{1+i} \right)^{t-s} U_E(Y_t) + \sum_{t=r}^T \left(\frac{1}{1+i} \right)^{t-s} U_R(B_t(r)) \quad \forall \quad r \geq s$$

$B_t(r)$ entspricht dem Renten- und Y_t dem Erwerbseinkommen in den jeweiligen Lebensjahren t . Die Höhe des Renteneinkommens richtet sich unter anderem nach dem Renteneintrittsalter r . $U_E(\cdot)$ und $U_R(\cdot)$ bezeichnen die Nutzenfunktionen in Abhängigkeit des Einkommens in der jeweiligen Lebensphase.

Ein Versicherter wird nun das Renteneintrittsalter r derart wählen, dass eine Verschiebung des Renteneintritts um ein Jahr zu einem geringeren Nutzenniveau führen würde. Ist $V_s(r+1)$ hingegen größer als $V_s(r)$, wird der Versicherte mit der Verrentung mindestens bis zum Zeitpunkt $r+1$ warten. Deshalb ist es sinnvoll, folgende Differenz zu betrachten:

$$G_s(r+1) = V_s(r) - V_s(r+1)$$

Nimmt der Ausdruck positive Werte an, wird der Renteneintritt in r gewählt, ansonsten wird er aufgeschoben. Die Höhe der zukünftigen Einkommen Y_t und damit auch $B_t(r)$ sind jedoch mit Unsicherheit behaftet, so dass das Individuum darüber Erwartungen zum Zeitpunkt s bilden muss. Dies führt zur neuen Definition der Funktion $G_s(\cdot)$:

$$G_s(r+1) = E(V_s(r)) - E(V_s(r+1))$$

Die Nutzenfunktionen werden analog zu Stock und Wise (1990) als Potenzfunktion des Einkommens und einem additiven stochastischen Störterm modelliert. In der Erwerbsphase erzielt das Individuum einen Nutzen U_E von:

$$U_E(Y_t) = Y_t^\gamma + \varepsilon_E$$

Aufgrund der Annahme abnehmender Grenznutzen sollte γ zwischen 0 und 1 liegen. In der Rentenphase lautet die Nutzenfunktion U_R aufgrund der unterstellten Bewertung der Freizeit:

$$U_R(B_t(r)) = (k \cdot B_t(r))^\gamma + \varepsilon_R$$

ε_E und ε_R stellen dabei zwei voneinander unabhängige Störterme dar. Aus Vereinfachungsgründen betrachten wir in dieser Arbeit nur die kurzfristige Entscheidung modelliert (einperiodiger Fall), wofür sich der Optionswert zu folgendem Ausdruck ergibt:

$$G_r(r+1) = E((k \cdot B_r(r))^\gamma) - E(Y_r^\gamma) - \sum_{t=r+1}^T \left(\frac{1}{1+i} \right)^{t-r} [E((k \cdot B_r(r+1))^\gamma) - E((k \cdot B_r(r))^\gamma)]$$

Bei einer intertemporalen Betrachtung des Optionswerts wird im Vergleich zum vorherigen Modell berücksichtigt, dass der Versicherte mehr als zwei Möglichkeiten für einen Renteneintritt hat. Die Alternative, heute im Alter r in Rente zu gehen, wird damit verglichen, wann $V_S(r)$ maximal wird. Intertemporal lautet das Optimierungskalkül:

$$\min G_s(r^*) = E(V_S(r)) - E(V_S(r^*))$$

$V_S(r^*)$ entspricht dem maximalen Nutzenniveau, das durch Renteneintritt zum Zeitpunkt r^* erzielt werden kann.

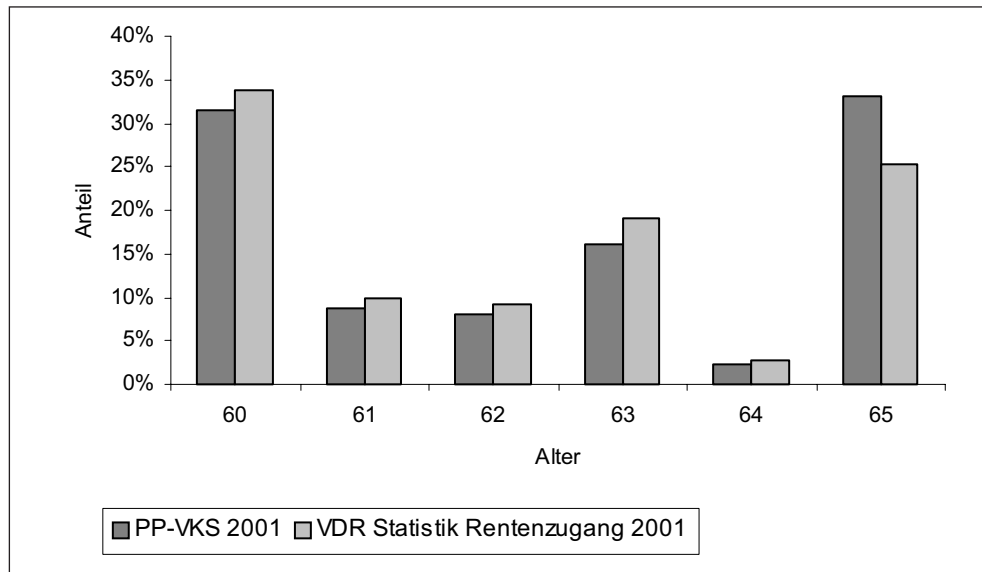
3. Datensatz (PP-VKS 2001)

Die empirische Relevanz des Modells wird nun mit prozessproduzierten Daten des VDR überprüft. Dabei handelt sich um die Versichertenkontenstichprobe aus dem Jahre 2001, die für dieses Forschungsvorhaben unter dem Namen „Pilotprojekt Versichertenkontenstichprobe 2001 (PP-VKS 2001)“ anonymisiert und aufbereitet wurde.¹ Weil sich die Erhebungsperiode auf das Jahr 2001 bezieht, sind die zu diesem Zeitpunkt gültigen Regeln zu berücksichtigen, nach denen sich ein Versicherter mit 60 Jahren für eine Rente wegen Arbeitslosigkeit oder für Altersrente entscheiden konnte. Im Datensatz PP-VKS 2001 sind versicherte Personen im Alter zwischen 55 und 65 Jahren enthalten, die entweder bereits verrentet sind oder noch im Berufsleben stehen.

Der Datensatz „Pilotprojekt Versichertenkontenstichprobe 2001“ enthält personenbezogene Daten wie Geschlecht, Geburtsdatum und die Anzahl der Kinder. Außerdem geht daraus hervor, ob ein Versicherter zuletzt Angestellter oder Arbeiter war sowie die Art und Höhe der Rente, die nun bezogen wird. Darüber hinaus liefern die Daten Aufschluss über den Status des Versicherten in den Jahren 1991 bis 2001. Dies umfasst Informationen über Arbeitslosigkeit, Krankheit ohne Lohnfortzahlung, Altersteilzeit, ein bestehendes Arbeitsverhältnis sowie die in diesen Zeiten erwirtschafteten Entgeltpunkte und über den Bezug von Altersrente. Außerdem kann danach unterschieden werden, ob die Entgeltpunkte in Ost- oder Westdeutschland erworben wurden.

¹ Weitere Anmerkungen zur Anonymisierung und Aufbereitung des Datensatzes "Pilotprojekt Versichertenkontenstichprobe 2001" sind in diesem Band im Beitrag von Hilmar Luckert enthalten.

Abbildung 1: Rentenzugänge 2001 in Deutschland, Renten wegen Alters, Männer, Verteilung nach Alter in Prozent



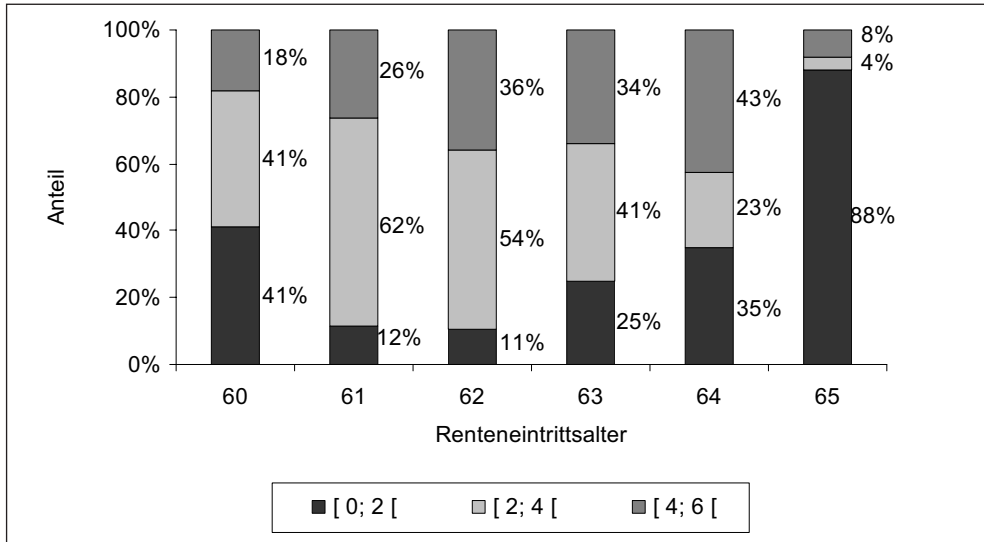
Quelle: PP-VKS 2001 und VDR Rentenzugangstatistik 2001, eigene Berechnungen.

In dieser Untersuchung werden nur männliche Versicherte betrachtet, die zu Beginn des Jahres 2001 noch nicht in Rente waren. Außerdem werden für die Schätzungen Personen nicht berücksichtigt, die in 2000 keine Entgeltpunkte verdient haben oder bereits in Altersteilzeit waren. In **Abbildung 1** wird der Renteneintrittsjahrgang 2001 der in der Schätzung verwendeten Individuen der Grundgesamtheit gegenübergestellt. Bis auf die Altersgruppe 65 sind die Abweichungen nicht allzu groß. Die Abweichungen werden unter anderem durch unterschiedliche Abgrenzungen von Alter und Rentenzugangszeitpunkt sowie durch die Abgrenzung der Stichprobe verursacht. Für die Aussagekraft der Schätzungen sind diese Abweichungen jedoch nicht von allzu großer Bedeutung.

Reil-Held (2000) findet einen positiven Zusammenhang zwischen Verdienst und einem späteren Verrentungszeitpunkt. Dies wollen wir zunächst explorativ überprüfen, indem für das Untersuchungsjahr 2001 die Einkommensverteilung der Rentenzugänge nach Alter betrachtet wird. Aus **Abbildung 2** erkennt man, dass der Anteil der Individuen mit höherem Einkommen bis zum Alter von 64 Jahren wächst, d.h., dass sie länger im Erwerbsleben verbleiben. Auf der anderen Seite zeigt sich für die unterste Einkommensklasse, dass ab der Altersgruppe 61 der Anteil stetig zunimmt, was eventuell durch die Notwendigkeit, das Rentenniveau durch zusätzliche Versicherungsjahre noch zu erhöhen, erklärt werden kann.

Betrachtet man anstelle der Summe der Entgeltpunkte der letzten 36 Monate die Summe aller Entgeltpunkte, die über den aktuellen Rentenwert und den Zugangsfaktor die monatliche Rentenhöhe bestimmt, so ergibt sich ein zur **Abbildung 2** ähnliches Bild. An dieser Stelle sei nochmals auf den klaren Vorteil der prozessproduzierten Daten hingewiesen. Die Erwerbs-

Abbildung 2: Rentenzugänge 2001 in Deutschland, Renten wegen Alters, Männer, Summe der in 36 Monaten vor Rentenbeginn erwirtschafteten Entgeltpunkte, Verteilung in Prozent der jeweiligen Altersgruppe



Quelle: PP-VKS 2001, eigene Berechnungen.

historie der sozialversicherungspflichtigen Einkünfte und damit verbunden die erworbenen Rentenansprüche sind exakt vorhanden und müssen nicht durch Schätzungen aus anderen Angaben gewonnen werden, was immer mit zusätzlicher Unsicherheit über die gewonnenen Werte verbunden ist.

Da die Zielsetzung des Datengewinnungsprozesses die individuelle Anspruchsberechnung ist, sind lediglich spärliche Angaben zur sonstigen Lebenssituation vorhanden. Für die Entscheidung über den Zeitpunkt der Verrentung eines Versicherten sind aber nicht nur die eigenen Einkünfte, sondern auch die Einkommenssituation und der Erwerbsstatus des Lebenspartners von entscheidender Bedeutung. Da im Datensatz keinerlei Verknüpfungen zum Partner möglich sind, können die Analysen nur auf der partiellen individuellen Ebene untersucht werden.

4. Schätzmodell

Wir beschränken unsere Analyse auf den einperiodigen Fall des Optionswertes. Es soll also nur die Entscheidung untersucht werden, im Jahr 2001 in Rente zu gehen oder die Verrentung bis zum Jahr 2002 aufzuschieben. Dabei nehmen wir vereinfachend an, dass zu Beginn eines Jahres die Entscheidung über den Renteneintritt getroffen wird. Diese binäre Entscheidung wird mit Hilfe des Probitansatzes modelliert. Als Referenzpunkt dient das Modell unter Einbeziehung von Scheinvariablen für Alter sowie für Arbeitslosigkeit in den letzten 12 Monaten und für Erhalt von Krankengeld ebenso in den letzten 12 Monaten. Dies soll die im Rentensystem angelegten institutionellen Rahmenbedingungen abbilden. Darüber hinaus wird untersucht, inwiefern die aus-

schließliche Verwendung des Optionswertes das Renteneintrittsverhalten erklären kann. Hierbei muss einschränkend eingeräumt werden, dass im Abschnitt 2 der Optionswert beispielsweise den Erwartungswert vom potenzierten zukünftigen Einkommen enthält. Selbst wenn der Erwartungswert des Einkommens $E(Y_r^y)$ bekannt wäre, ließe sich $E(Y_r^y)$ nicht einfach ermitteln. Da wir sehr genaue Einkommensangaben der Vergangenheit besitzen, gehen wir von einem deterministischen zukünftigen Einkommen aus, so dass sich der Optionswert im einperiodigen Fall reduziert zu:

Die Entscheidung zu Beginn des Jahres 2001 beruht auf Schätzungen für das zu erzielende Arbeitseinkommen in diesem Jahr. Als Schätzung dafür wird das Arbeitseinkommen im Jahr 2000 verwendet. Da der Zuwachs an Entgeltpunkten und damit auch an zukünftigen Rentenzahlun-

$$G_r(r+1) = (k \cdot B_r(r))^y - Y_r^y - \frac{1 - (1+i)^{r+1-T}}{1 - (1+i)^{-1}} \left[(k \cdot B_r(r+1))^y - (k \cdot B_r(r))^y \right]$$

gen proportional zum Arbeitseinkommen im Betrachtungszeitpunkt ist, sind die verwendeten Größen für Y und die Veränderung von B hochgradig korreliert. Zusätzlich stellt sich die Frage, ob alle Parameter im nicht-linearen Optionswertausdruck im Rahmen des Probitmodells identifiziert und damit schätzbar sind. Über Simulationen unter Berücksichtigung der im Datensatz vorhandenen Werte für Y und B kann gezeigt werden, dass im Fall der Festlegung des Zinssatzes i sowie einer Annahme über ein fixes T alle anderen Parameter identifiziert sind, weshalb die Schätzungen für verschiedene Zinssätze durchgeführt und dargestellt werden.

5. Ergebnisse

Zunächst wird die Verrentungsentscheidung im Jahr 2001 mit Hilfe der für die Auswertungen selektierten $n=3.452$ Beobachtungen unter Berücksichtigung von Altersvariablen untersucht. Dies führt zu einem Pseudo- R^2 von 8,5% nach McFadden (bzw. 12,5% nach Veall und Zimmermann, vgl. Ronning 1990, Abschnitt 2.5.1). Eine Erweiterung um die Scheinvariablen für Arbeitslosigkeit sowie für Krankheit erhöht diese Maße auf 24% bzw. 30%, was die Bedeutung der Interaktion der anderen Sozialversicherungssysteme für die Verrentungsentscheidung hervorhebt (Schnabel 2003). Die Ergebnisse sind in **Tabelle 1** zusammengefasst:

Die Altersvariablen besitzen allesamt einen signifikanten Einfluss auf das Renteneintrittsverhalten. Im Vergleich zum 65. Lebensjahr, bei dem die Renteneintrittswahrscheinlichkeit bei ca. 93% liegt, sinkt die Wahrscheinlichkeit in Rente zu gehen bei geringerem Alter. Darüber hinaus lassen sich die Spitzen beim Rentenzugang im Alter von 60 und 63 Jahren sehr gut wieder finden. Ab 60, als dem ersten möglichen Zeitpunkt für den Renteneintritt, sinkt die Renteneintrittswahrscheinlichkeit zunächst. Mit 63, wenn die Versicherten die Wahl haben, eine Altersrente für langjährige Versicherte in Anspruch zu nehmen, steigt die Wahrscheinlichkeit wieder sprunghaft an, um mit 64 Jahren wieder leicht abzunehmen.

Die Aufnahme von Scheinvariablen für Arbeitslosigkeit und Krankheit im Jahr 2000 in Spezifikation 3 bewirkt eine signifikante Verbesserung. Die Tatsache, im Jahr 2000 krank ohne Lohnfortzahlung bzw. arbeitslos gewesen zu sein, unabhängig von der Länge dieser Phasen, hat ei-

Tabelle 1: Binäres Probitmodell

Variable	Spezifikation 1		Spezifikation 2		Spezifikation 3	
	Parameter	t-Wert	Parameter	t-Wert	Parameter	t-Wert
c	-0,5025	-22,49	1,4753	9,73	1,3940	9,16
d(60)			-2,1049	-13,44	-2,6616	-16,59
d(61)			-2,2056	-13,84	-2,6395	-16,17
d(62)			-2,2821	-14,05	-2,6719	-16,03
d(63)			-1,7011	-10,49	-1,8429	-11,28
d(64)			-1,7503	-10,15	-1,8845	-10,84
d(AL)					1,3566	24,19
d(KA)					0,5854	5,88
ln L	-2.130,594		-1.948,461		-1.619,070	

Quelle: PP-VKS 2001, eigene Berechnungen.

nen starken positiven Einfluss auf die Wahrscheinlichkeit, in Rente zu gehen: Eine Arbeitslosigkeit erhöht die Wahrscheinlichkeit für einen Renteneintritt um 48,78 % und eine Krankheit um 21,75%.

Als nächstes wenden wir uns der Frage zu, ob das Optionswertmodell das Renteneintrittsverhalten ebenfalls so gut erklären kann wie die bisherigen Modelle. Wie oben erwähnt, müssen Werte für den Abdiskontierungszinssatz i restringiert werden, um die anderen Parameter des Optionswertes schätzen zu können. Die Ergebnisse sind in **Tabelle 2** zusammengestellt:

Tabelle 2: Schätzung des Optionswerts im binären Probitmodell

Variable	Spezifikation 4 $i = 1,5 \%$		Spezifikation 5 $i = 3,0 \%$		Spezifikation 6 $i = 4,5 \%$	
	Parameter	t-Wert	Parameter	t-Wert	Parameter	t-Wert
c	-0,7552	-6,46	-0,7325	-7,00	-0,7144	-7,55
α	0,000270	1,20	0,000100	1,17	0,000036	1,44
k^γ	6,1002	7,27	6,4078	7,78	6,7475	8,18
γ	0,7171	10,00	0,8153	10,96	0,9151	11,90
ln L	-1925,480	-1920,837	-1916,278			

Quelle: PP-VKS 2001, eigene Berechnungen.

Man erkennt, dass für verschiedene Zinssätze die resultierenden Log-Likelihoodwerte etwas besser als in der Spezifikation 2 sind, die ausschließlich auf die Altersdummys zurückgreift. Das Konzept kann demnach den Renteneintritt recht gut erklären (Pseudo- R^2 nach McFadden: 10%). Die Schätzung des Parameters α vor dem Optionswert hat das erwartete positive Vorzeichen. Eine Erhöhung des Optionswerts lässt die Wahrscheinlichkeit für einen Renteneintritt in 2001 steigen. Einschränkend muss gesagt werden, dass der geschätzte Koeffizient gemessen am t-Wert nicht signifikant von Null verschieden ist, was eigentlich gegen den Optionswert

als erklärende Variable sprechen würde. Allerdings zeigt der Likelihoodquotiententest für das Modell mit Optionswert einen signifikanten Erklärungsgehalt an. In den im Abschnitt 4 erwähnten Simulationsexperimenten fiel bereits die hohe Streuung der Schätzungen für α auf, was durch hoch korrelierte Beobachtungen für Y und die Veränderung von B verursacht sein könnte.

Die Schätzungen für den Parameter γ weisen allesamt eine erwartete Größenordnung auf. In den Spezifikationen 4 und 5 sind die Koeffizienten signifikant von eins verschieden, so dass die Schätzungen abnehmende Grenznutzen bestätigen. Werden höhere Kalkulationszinssätze verwendet, resultieren Koeffizientenschätzungen größer als eins. Jedoch steht die Wahl der Zinssätze in Einklang mit sonst verwendeten Größenordnungen (Börsch-Supan 1992). Der Koeffizient k^2 deutet darauf hin, dass die Versicherten eine extreme Präferenz haben, in Rente zu gehen. Nach den geschätzten Nutzenfunktionen für die Erwerbs- und Rentenphase ist nach Spezifikation 6 ein Euro in Rente genauso viel wert wie 8,05 € in der Erwerbsphase. Im Vergleich zu den Schätzungen von Stock und Wise (1990), die γ mit 0,632 schätzen, ist der hier geschätzte Wert für k wesentlich größer. Neben unbeobachtbaren Größen wie Haushaltseinkünfte kann eine Überschätzung von k auch darauf zurückgeführt werden, dass in der Schätzung das sozialversicherungspflichtige Einkommen ohne Abzüge verwendet wird. Die Differenz zwischen Nettoeinkommen und potenzieller Nettorente dürfte demnach geringer ausfallen, so dass ein betragsmäßig kleineres k zu erwarten wäre.

Wir integrieren nun beide Ansätze, um die finanziellen Anreize ebenso wie die Interaktionen zu den anderen Sozialversicherungssystemen beurteilen zu können. Dabei wählen wir einen Zinssatz in Höhe von 3%. Die Ergebnisse dieser Schätzung finden sich in **Tabelle 3**.

Tabelle 3: Probit Schätzung des kombinierten Modells

Variable	Parameter	t-Wert
c	0,0426	0,19
d(AL)	1,2606	20,69
d(KA)	0,6192	5,69
d(60)	-1,6339	-7,98
d(61)	-1,7589	-8,65
d(62)	-1,9801	-9,68
d(63)	-2,0669	-9,90
d(64)	-2,0948	-9,67
a	0,00010	0,83
kg	5,8317	2,94
g	0,7644	8,22
ln L	-1529,061	

Quelle: PP-VKS 2001, eigene Berechnungen.

Durch die Hinzunahme des Optionswertes gegenüber der Spezifikation 3 steigt der Erklärungsgehalt signifikant an. Das Pseudo- R^2 nach McFadden erhöht sich auf 28,23%. Wiederum haben die Koeffizienten die erwarteten Vorzeichen und Größenordnungen, auch wenn die Standardabweichung für die Schätzung von α wiederum sehr groß ist. Darüber hinaus bleiben neben dem Optionswert die Scheinvariablen zu Alter, Krankheit und Arbeitslosigkeit weiterhin signifikant, die letzten beiden verändern kaum ihre Größenordnung im Vergleich zu Spezifikation 3. Dies zeigt an, dass neben rein finanziellen Überlegungen auch andere Einflussfaktoren wie Gesundheit und Beschäftigungsmöglichkeiten weiterhin wichtige Determinanten zur Erklärung des Renteneintrittsverhaltens bleiben.

6. Fazit

Der Optionswert spiegelt die finanziellen Anreize der gesetzlichen Rentenversicherung für den Übertritt in die Rente gegenüber einem Verbleib in der Erwerbstätigkeit wider. Diese Anreize bleiben auch noch nach der Einführung der Abschläge in Höhe von 0,3 % pro Rentenmonat, die ein Versicherter vor Erreichen des 65. Lebensjahres in Rente geht, bestehen. Die Schätzungen deuten auf extreme Präferenzen für Freizeit hin, so dass die Abschläge weit höher ausfallen müssten, um Anreize für einen späteren Renteneintritt zu setzen.

Die geschätzte starke Präferenz für Renteneinkommen, die sich durch den hohen Wert von k^γ ergibt, kann durch nicht beobachtbare Einflussfaktoren überschätzt sein. Da keinerlei Informationen über den Erwerbsstatus und das Einkommen des Lebenspartners sowie über sonstige private und betriebliche Altersversorgung vorliegen, wird der Wert eines Euro aus der gesetzlichen Rentenversicherung des Individuums gegenüber einem Euro aus Arbeitseinkommen höchstwahrscheinlich überschätzt. Weitere Einflussfaktoren wie Abfindungen oder drohender Arbeitsplatzverlust, die einen Renteneintritt begünstigen, können ebenfalls nicht beobachtet werden und können zu Verzerrungen in den Schätzungen führen.

Die Interaktionen zu den anderen sozialen Sicherungssystemen haben für die Verrentungsentscheidung eine enorme Bedeutung. Arbeitslosigkeit oder Krankheit sind bedeutende Faktoren, die zu einem früheren Eintritt in die Altersrente führen. In unsere Analyse geht nur die Tatsache ein, ob ein Individuum im letzten Jahr Krankengeld oder Arbeitslosengeld bezogen hat, nicht aber die Tatsache, wie lange dies im Einzelfall war. In zukünftigen Analysen wird dies berücksichtigt werden.

Literatur

- Börsch-Supan, A. (1992): Population Aging, Social Security Design and Early Retirement. *Journal of Institutional and Theoretical Economics*, 148, 533-557.
- Reil-Held, A. (2000): Einkommen und Sterblichkeit in Deutschland: Leben Reiche länger? Universität Mannheim, SFB 504 *Diskussionspapier*, 00-14.
- Ronning, G. (1991): *Mikroökometrie*. Springer-Verlag: Berlin.
- Schmidt, P. (1994): Die Wahl des Rentenalters. Peterlang-Verlag: Frankfurt.
- Schnabel, R. (2003): *The Joint Effects of Unemployment Insurance and Public Pension in Germany*. Arbeitspapier. Universität Essen.
- Stock, J. & Wise, D. (1990): Pensions, The Option Value of Work and Retirement. *Econometrica*, 58, 5, 1151-1180.

Dipl.-Vw. Stefan Kempf ist wissenschaftlicher Mitarbeiter bei Prof. Dr. Martin Kukuk am Lehrstuhl für Ökonometrie der Universität Würzburg. Forschungsschwerpunkte sind die Ökonometrie im Allgemeinen sowie Determinanten für die Wahl des Renteneintrittalters.

Prof. Dr. Martin Kukuk: Studium der VWL in Konstanz. Diplom 1988, Promotion 1991 (Uni Konstanz). Feodor-Lynen Stipendiat der A.v.Humboldt-Stiftung und Visiting Assistant Professor am Economics Department der McMaster University, Hamilton/Ontario, Kanada (1992). Nach Stationen im Sonderforschungsbereich 178 an der Uni Konstanz, dem Zentrum für Europäische Wirtschaftsforschung, Mannheim, und der Uni Tübingen, Habilitation dort im Fach Ökonometrie und Statistik (2000). Rufannahme auf eine Professur für BWL, insb. Empirische Wirtschaftsforschung an der Uni Magdeburg (2002), Rufannahme auf den Lehrstuhl für Ökonometrie an der Uni Würzburg (Sept 2002).

Der **Lehrstuhl für Ökonometrie der Wirtschaftswissenschaftliche Fakultät (Universität Würzburg)** beschäftigt sich mit der Überprüfung von Hypothesen über ökonomische Zusammenhänge auf Grundlage relevanter empirischer Informationen und mittels geeigneter Methoden der mathematischen Statistik. Ziel unserer Arbeit ist es, Entscheidungen in der Wirtschaftspolitik oder im Management auf eine gesicherte Basis zu stellen (siehe www.wifak.uni-wuerzburg.de/qwf).

2.2 Projektskizzen und Datenbedarf der Wissenschaft

Eines der wichtigsten Anliegen, das mit der Veranstaltung des Workshops verfolgt wurde, ist die differenzierte Ermittlung des Datenbedarfs der Wissenschaftlerinnen und Wissenschaftler. Über die vorangegangene quantitative Erhebung der Datenwünsche hinaus galt es, die aktuellen Fragestellungen der Forschung genauer kennen zu lernen, um die künftige Ausrichtung des Datenangebots des FDZ-RV an konkreten Forschungsvorhaben orientieren zu können.

Zum Abschluss des ersten Veranstaltungstages wurden deshalb laufende und geplante Projekte der teilnehmenden Forscherinnen und Forscher vorgestellt. In kompakten Folienvorträgen schilderten die Wissenschaftler ihre Forschungsinteressen und ihre wichtigsten Fragestellungen.¹ Mitglieder nahezu aller teilnehmenden Forschergruppen präsentierten ihre jeweiligen Forschungsprojekte, deren theoretische Einbettung sowie abgeleitete Forschungshypothesen. Im Mittelpunkt standen die sich daraus ergebenden Anforderungen an die Daten bzw. benötigte Variablen aus den verschiedenen Datensätzen der Rentenversicherung.

Gemeinsam ist den Projekten aus unterschiedlichen Disziplinen, dass sie aufbauend auf den Mikrodaten der Rentenversicherung nach neuen Erkenntnissen streben. Die Publikation der Projektskizzen an dieser Stelle ist insbesondere durch den Wunsch motiviert, den Kommunikationsprozess zwischen Wissenschaft und FDZ-RV von Anfang an transparent zu gestalten und künftige Entscheidungen bezüglich des Zuschnitts der bereitgestellten Datensätze nachvollziehbar zu machen. Zu diesem Zweck wurden die teilnehmenden Forscherteams gebeten, ihre Fragestellungen und die Konzepte ihrer Forschungsvorhabens schriftlich zu charakterisieren und diese Projektskizzen dem VDR zukommen zu lassen.

Für die Anfertigung der Projektbeschreibungen wurden den Forschergruppen während des Workshops Hinweise für die inhaltliche Gliederung zur Erstellung ihrer Beiträge mitgegeben.² Erfreulicherweise haben sich die Wissenschaftlerinnen und Wissenschaftler weitgehend an das vorgegebene Schema gehalten, das es den Lesern erleichtern soll, sich einen schnellen Überblick zu verschaffen und einzelne Projektskizzen zu vergleichen.

Die in diesem Abschnitt abgedruckten Projektskizzen sind folgendermaßen gegliedert: Zunächst wird eine Beschreibung des organisatorischen Rahmens eines jeden Projektes gegeben und in die Forschungsthematik eingeführt. Nachdem der angestrebte Erkenntnisgewinn des Forschungsvorhabens herausgestellt wurde, wird eine theoretische Einbettung der Untersuchung vorgenommen und das methodische Vorgehen dargelegt.

Im Mittelpunkt der Beiträge steht die Frage nach dem Datenbedarf: Welche speziellen Datensätze werden für das Forschungsvorhaben benötigt? Welche Daten stehen bereits zur Verfügung und woran scheitern weitergehende Untersuchungen? Inwieweit können Mikrodaten der Rentenversicherung solche Lücken füllen und zu neuen Erkenntnissen beitragen?

¹ Die am Workshop teilnehmenden Wissenschaftler waren im Vorhinein seitens der Tagungsleitung darum gebeten worden, ihre Forschungsprojekte anhand einer mitgebrachten Folie zu skizzieren.

² Die Gliederungsvorgabe enthielt folgende acht Punkte: (1) Einleitung mit Kurzcharakterisierung des Forschungsprojektes, (2) Fragestellung, (3) Ziel/möglicher Erkenntnisgewinn, (4) theoretische Einbettung, (5) methodisches Vorgehen, (6) Datenbedarf mit Bezug auf Daten des FDZ-RV mit min-max-Abwägung, (7) gewünschte und akzeptierbare Datenzugangswege, (8) Zusammenfassung - welche Schritte sind realisierbar und woran Scheitern bestimmte Dinge?

Schließlich werden die Datenanforderungen im Hinblick auf Minimal- und Maximalwünsche abgewogen und mit Bezug auf die verschiedenen diskutierten Wege des Datenzugangs Präferenzen geäußert.

Die Zusammenführung der Projektskizzen in diesem Tagungsband veranschaulicht die Bandbreite der Einsatzmöglichkeiten von Mikrodaten der Rentenversicherung, die im Rahmen des FDZ-RV angeboten werden sollen. Gleichzeitig wird bereits in der Aufbauphase der enge Kontakt zu den Forschungsinstituten gepflegt und eine transparente wie kooperative Arbeitsweise zwischen dem Forschungsdatenzentrum und seinen zukünftigen Nutzern kultiviert.

Die Projektskizzen gliedern sich nach dem konkreten Datenbedarf, wobei sich deren Reihenfolge an den Umfrageergebnissen³ orientiert: Zuerst werden die Projektskizzen abgedruckt, die schwerpunktmäßig auf Analysen der Rentenstatistik basieren; dann folgen jene, die auf Datengrundlage der Versichertenstatistik Berechnungen durchführen möchten. Zuletzt schließen sich die Projektskizzen an, die sich auf die Reha-Statistik konzentrieren. Innerhalb dieser Kategorien sind die Projektskizzen jeweils in alphabetischer Ordnung des zuerst genannten Autors aufgeführt.

³ Siehe hierzu den Beitrag von Ralf Himmelreicher im vorliegenden Band.

Übergänge in den Ruhestand

Sabine Böttcher (Zentrum für Sozialforschung Halle)

Einleitung

Der ostdeutsche Arbeitsmarkt ist gekennzeichnet durch einige spezifische Problemlagen, die besonders die Erwerbssituation der jüngeren und älteren Erwerbsfähigen betreffen: Zum einen sind durch die Regelungen zum Vorruhestand und zur Altersteilzeit in den Nachwendejahren ein Großteil der über 55-Jährigen vorzeitig und vor allem (fast) gleichzeitig aus dem Erwerbsleben aus. Damit kam es zum „Abbruch“ natürlicher Verrentungsprozesse, da in den Folgejahren kaum noch Erwerbstätige das Rentenalter erreichten. Erst jetzt, nachdem die damals jüngeren Alterskohorten „durchgealtert“ sind, setzt sich der Prozess der natürlichen Verrentung wieder langsam in Gang. Zum anderen charakterisiert sich der ostdeutsche Arbeitsmarkt durch eine hohe Arbeitslosigkeit der jüngeren und älteren Erwerbsfähigen. Aufgrund der Gleichzeitigkeit des Wiedereinsatzes des Generationenaustausches einerseits und hoher Arbeitslosigkeit gerade unter Älteren andererseits ist zu fragen, ob und inwieweit sich Arbeitslosigkeit als Übergangsstadium in Rente etabliert hat.

Im Rahmen des Projektes „ostmobil – Analysen zur Mobilität auf dem ostdeutschen Arbeitsmarkt“ beschäftigt sich der Berichtsteil 3 „Übergänge in den Ruhestand“ neben der Erfassung und quantitativen Beschreibung der Entwicklung des Arbeitsmarktes durch Übergänge in den Ruhestand mit der Analyse des Austrittsverhaltens ostdeutscher Arbeitnehmer aus dem Arbeitsmarkt.

Derzeit läuft im Rahmen dieses Projektes eine methodische Vorstudie zur „Eignung des Mikrozensuspanels für Analysen des Übergangs von der Erwerbstätigkeit in den Ruhestand“ in Zusammenarbeit mit dem Statistischen Landesamt Sachsen-Anhalt und dem Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen.

Fragestellung

In der Analyse von Erwerbsverläufen älterer Arbeitnehmer mit dem besonderen Augenmerk auf der Organisation von Ausstiegen aus dem Erwerbsleben stehen folgende Fragen im Vordergrund:

- Wie stellt sich die Arbeitsmarktsituation der über 50-jährigen Männer und Frauen in Ost- und Westdeutschland dar?
- In welchem Lebensalter verlassen ältere Arbeitnehmer endgültig den Arbeitsmarkt?
- Aus welchen Erwerbs- bzw. Nichterwerbsverhältnissen heraus erfolgt der Übergang in den Ruhestand? Das besondere Augenmerk liegt dabei auf den Status Erwerbstätigkeit, Arbeitslosigkeit und sonstige Nichterwerbstätigkeit (z.B. Unterhalt).

- Lassen sich spezifische Beschäftigungsverhältnisse darstellen, aus denen heraus der direkte Übergang von Erwerbstätigkeit in Ruhestand erfolgt?
- Können berufsgruppen- oder branchenspezifische Austrittswahrscheinlichkeiten abgebildet werden?

Ziel und möglicher Erkenntnisgewinn

Ziel des Projektes ist die Analyse der Erwerbsbeteiligung, der Erwerbsverläufe und der Austrittsynamik älterer Arbeitnehmer (ab 50 Jahren) auf dem ostdeutschen Arbeitsmarkt. Im Vordergrund steht dabei nach der Darstellung der Erwerbsbeteiligung dieser Altersgruppe die Abbildung des Anteils derjenigen, die aus der Erwerbstätigkeit heraus direkt in den Ruhestand wechseln im Vergleich zu denjenigen, die den Übergang in den Ruhestand über den „Umweg“ Arbeitslosigkeit realisieren müssen. Dabei ist u.a. von besonderem Interesse, ob bei denjenigen, die bis zum Eintritt in den Ruhestand erwerbstätig sind, berufs- oder branchenspezifische Merkmale differenziert werden können.

Ein weiteres Ziel ist die Nutzbarmachung amtlicher Statistiken für ein weites Praxisfeld, in dem Routinen entwickelt werden, mit denen diese amtlichen Statistiken im weiteren fortgeschrieben werden können.

Methodisches Vorgehen

Bisher:

- Analyse der aggregierten Arbeitsmarktdaten für die Jahre 1995 – 2002 für Männer und Frauen in Ost- und Westdeutschland
- Analyse der aggregierten Berufsdaten aus den Meldungen der Arbeitgeber zur Sozialversicherung (Beschäftigtenstatistik)
- Analyse der aggregierten Rentenzugangs- und Rentenbestandsdaten für die Jahre 1995 – 2002 nach ausgewählten Rentenarten für Männer und Frauen in Ost- und Westdeutschland

Derzeit:

- Methodische Vorstudie zur Eignung des Mikrozensuspanels für Analysen des Übergangs von Erwerbstätigkeit in den Ruhestand

Geplant (bei Eignung des Mikrozensuspanels):

- Aggregation personenbezogener Erwerbsverläufe über 55-jähriger Männer und Frauen mit Statusveränderungen (Status: erwerbstätig, arbeitslos, in Rente, Unterhalt, Sonstiges)
- Aggregation personenbezogener Rentenübergangsdaten (u.a. auch zur Erhöhung der Validität der mittels des Mikrozensuspanels gewonnenen Ergebnisse)

Datenbedarf mit Bezug zu FDZ-RV-Daten und Zugangswege

Zur Aggregation personenbezogener Rentenübergangsdaten werden aus dem Datenbestand des Forschungsdatenzentrums der Rentenversicherung folgende Daten benötigt:

- Geschlecht
- Region (Ost-/Westdeutschland, evtl. Bundesländerebene)
- Datum des Rentenüberganges oder Alter beim Übergang in Rente
- Rentenart
- Versicherungsverhältnis vor dem Leistungsfall
- Beruf/Tätigkeit bei Rentenübergang bzw. Antragstellung

Als möglicher Zugangsweg zu den Daten scheinen Gastwissenschaftler-Arbeitsplätze am günstigsten. Als Statistikprogramm wird STATA präferiert.

Sabine Böttcher (geb. 1969), Studium der Soziologie an der Martin-Luther-Universität Halle-Wittenberg (1988-1995), seit 1996 Projektmitarbeit in der Forschungsgemeinschaft für Konflikt- und Sozialstudien Halle e.V. und seit 2000 wissenschaftliche Mitarbeiterin am Zentrum für Sozialforschung Halle (zsh). Arbeitsschwerpunkte: Arbeitsmarkt, Erwerbsverläufe.

Zentrum für Sozialforschung Halle e.V. an der Martin-Luther-Universität Halle-Wittenberg, gegründet im Winter 1996/1997. Ein Team von ca. 15 Mitarbeitern bearbeitet mit ausgeprägter Konzentration auf ostdeutsche Entwicklungen und Problemlagen Forschungsprojekte zu folgenden zwei Schwerpunkten: 1) Arbeitsmarkt, Ausbildungssysteme, Einmündungs- und Berufsverläufe und 2) Unternehmensstruktur und betriebliche Strategien (siehe <http://www.zsh-online.de>).

Berichtssystem zum Altersübergang aus der Sicht von Beschäftigten und Betrieben (Altersübergangs-Monitor)

Renate Büttner und PD Dr. Matthias Knuth (Institut Arbeit und Technik, Gelsenkirchen)

Einleitung

Wie in vielen anderen entwickelten Gesellschaften gibt es in Deutschland einen Trend zum frühen Ausstieg aus der Erwerbstätigkeit. Die Alterung der Gesellschaft und die daraus resultierende Belastung der sozialen Sicherungssysteme zwingen dazu, diese Entwicklung umzukehren. Dazu brauchen Sozialpartner und politisch Verantwortliche Informationen über das aktuelle Geschehen sowie die Optionen der Beschäftigten und der Betriebe. Die öffentliche Diskussion wird oft so geführt, als ginge es nur um die Heraufsetzung des gesetzlichen Rentenalters. Die Renteneintritte allein sind jedoch ein unzureichender Indikator, weil es zwischen Erwerbstätigkeit und Verrentung verschiedene, durch die jeweiligen gesetzlichen Rahmenbedingungen forcierte Übergangsstadien gibt. Dazu gibt es zwar eine Fülle verschiedener Datenquellen, was jedoch fehlt, ist die systematische Zusammenschau der verschiedenen Mechanismen und der über sie Auskunft gebenden Daten.

Vor diesem Hintergrund hat die Hans-Böckler-Stiftung das Institut Arbeit und Technik beauftragt, auf der Grundlage vorhandener, aber bisher nicht systematisch aufeinander bezogener Daten ein Berichtssystem zum Altersübergangs-Geschehen zu entwickeln.

Zielsetzung

Unter „Altersübergang“ verstehen wir:

1. den Prozess, der von der Erwerbstätigkeit in die Altersrente führt,
2. die Lebensphase zwischen dem Austritt aus Erwerbstätigkeit und dem „Ankommen“ in der Altersrente,
3. die unterschiedlichen sozialrechtlichen Status, die im Prozess des Altersübergangs durchlaufen werden, und ihre konsekutive sozialrechtliche Bedingtheit.

Angesichts der Tatsache, dass die Mehrheit der Zugänge in Altersrenten nicht aus dem Beschäftigungsstatus erfolgt, kann man davon sprechen, dass der Altersübergang sich als quasi vierte Phase des Lebenslaufes (im Sinne des ursprünglich dreigliedrigen Schemas von Kohli 1985 und 1994) institutionalisiert hat.

Um sowohl betrieblichen als auch gesellschaftlichen Akteuren ein möglichst zeitnahes sowie repräsentatives Bild des Altersübergangs-Geschehens und der – aufgrund der in Betrieben und bei Beschäftigten vorherrschenden Einstellungen, Erwartungen und Optionen – mittelfristig zu erwartenden Entwicklung zu vermitteln, wird innerhalb einer zweijährigen Pilotphase ein Altersübergangs-Monitor entwickelt. Durch dieses Berichtssystem erhalten die Entscheidungsträger der Gesetzgebung und der Tarifpolitik eine Rückmeldung, wie sich die bereits vollzogenen so-

wie weitere noch ausstehende Veränderungen der institutionellen Rahmenbedingungen des Altersübergangs auf das Verhalten und die Orientierungen von Betrieben und Beschäftigten auswirken.

Der Regelkreis zwischen der Umgestaltung der Rahmenbedingungen, den dadurch bewirkten (oder auch nicht bewirkten) Verhaltensänderungen und der Wahrnehmung dieser Veränderungen durch die Entscheidungsträger ist derzeit aufgrund der Zersplitterung der Datenlage und des zu großen Zeitverzuges bei der Aufbereitung dieser Daten unvollständig und unwirksam. In amtlichen Statistiken öffentlich zugängliche Daten werden nicht handlungsrelevant, weil niemand sie zu interpretierbaren Informationen verdichtet, sie graphisch veranschaulicht und als Orientierungswissen in die öffentliche Debatte einspeist. Im Rahmen des Projektes werden verschiedene ohnehin existierende Prozessstatistiken und Panelerhebungen, die bereits relevante Informationen zum Altersübergangsgeschehen beinhalten, systematisch und kontinuierlich ausgewertet, die entsprechenden Ergebnisse zusammengeführt und vor dem Hintergrund sich verändernder institutioneller Rahmenbedingungen interpretiert.

Die Ergebnisse des Projekts werden in unregelmäßigen Abständen in elektronischer Form als „Altersübergangs-Report“ auf den Internet-Seiten der Hans Böckler Stiftung und des Instituts Arbeit und Technik publiziert und zum download angeboten.

Fragestellungen

Angesichts der vielfältigen Bemühungen von Gesetzgeber und Tarifparteien, den Altersübergang umzugestalten und eine zumindest teilzeitige Erwerbstätigkeit bis in ein etwas höheres Lebensalter zu erreichen, ist es eine für die Gestaltung künftiger Altersübergänge entscheidende Frage, ob und in welcher Weise die bisherigen Reformen der Rahmenbedingungen Veränderungen des realen Altersübergangsgeschehens hervorgerufen haben. Wie Koller (2001) zeigt, ist diese scheinbar einfache, deskriptiv zu beantwortende Frage alles andere als trivial. Zeitreihen von altersübergangs-relevanten Phänomenen sind für sich allein nicht aussagefähig, sondern müssen im Kontext demographischer Veränderungen und des Konjunkturzyklus gewürdigt werden. Außerdem bilden verschiedene Datenquellen (s.u.) jeweils nur einen Ausschnitt des Altersübergangs-Geschehens ab; sie müssen mosaikartig zusammengefügt werden, um zu einer verlässlichen Einschätzung zu kommen. Hierfür genügen keine Momentaufnahmen, sondern es ist eine kontinuierliche Beobachtung erforderlich.

Im einzelnen geht es bei dem zu entwickelnden Monitoring von Altersübergängen um folgende Fragen:

- Wie verändern sich die Altersstrukturen unter den Beschäftigten im Vergleich zu denen der Bevölkerung im Erwerbsalter?
- Mit welchen Maßnahmen reagieren die Betriebe auf den demographischen Wandel?
- Gibt es Anzeichen für eine Umkehr des Trends zum immer früheren Erwerbsausstieg und Renteneintritt?
- Wie verändern sich die Muster der Überbrückung zwischen Arbeit und Rente unter veränderten sozialrechtlichen Rahmenbedingungen?

- Wie schlägt sich das nieder in der Inanspruchnahme von Lohnersatzleistungen der Bundesagentur für Arbeit oder Sozialhilfe, von Altersteilzeit oder von Instrumenten der aktiven Arbeitsförderung?
- Wie verändern sich die Optionen der Betriebe sowie die Lebensplanung der Beschäftigten bezüglich der Altersarbeit und des Altersübergangs? Gibt es bei den Beschäftigten geschlechtsspezifische Unterschiede? Welchen Einfluss hat der Haushaltskontext?

Gesellschaftspolitischer Begründungszusammenhang

Angesichts der langfristigen Vorhersehbarkeit demographischer Veränderungen wird die Alterung der deutschen (und westeuropäischen) Gesellschaft zwar spät, jedoch in Deutschland immerhin seit etwa Anfang der neunziger Jahre des vergangenen Jahrhunderts verstärkt diskutiert.¹ Dabei dominierten lange Zeit Fragen der Ergänzung und damit Entlastung der beitragsfinanzierten Alterssicherung und der Heraufsetzung der Altersgrenzen, ab denen volle Altersrenten beansprucht werden können. Inzwischen kommen auch Fragen der Familien- und Bildungspolitik als Bedingungsfaktoren des Reproduktionsverhaltens auf die Tagesordnung, haben aber angesichts leerer öffentlicher Kassen einen schweren Stand.

Während die Geburtenraten nur langfristig und – über sie – der Altersaufbau der Bevölkerung nur in noch sehr viel längerfristiger Perspektive beeinflussbar sind, bildete die vergleichsweise kurzfristig beeinflussbare Stellgröße, nämlich das faktische Alter des Austritts aus Erwerbstätigkeit und Beitragszahlung und des Eintritts in den Bezug irgendeiner Form von Transferleistung, lange Zeit einen „blinden Fleck“ der gesellschaftspolitischen Debatte. Das dürfte im Wesentlichen auf folgende Gründe zurückzuführen sein:

1. So sehr die Sozialpartner auf gesamtgesellschaftlicher Ebene das Ansteigen der Lohnnebenkosten durch den frühen Altersübergang beklagen mögen, so sehr haben sie in konkreten Situationen der Personalanpassung ein Interesse an der Erhaltung dieses Ventils.
2. Bei der Bewältigung des Strukturbruchs und der massiven Arbeitsplatzverluste, die die deutsche Wirtschafts- und Währungsunion und die anschließende Vereinigung auslösten, spielte der Altersübergang mit einem Höchststand von fast 900.000 Beziehern von Altersübergangsgeld im Jahre 1993 eine prominente Rolle. Die Alternativlosigkeit dieser Praxis hat die kritische Diskussion des Altersübergangs in Deutschland um mehr als eine halbe Dekade verzögert.
3. Frühe Altersübergänge beinhalten verschiedene sozialrechtlich definierte Status, die öffentlich unter anderen Paradigmata wie Langzeitarbeitslosigkeit oder Arbeitslosigkeit Geringqualifizierter diskutiert werden. Inwieweit es sich hierbei um Altersübergänge handelt, stellt sich definitiv erst im Nachhinein bei Erreichen einer Altersrente heraus. Obwohl den meisten Menschen aus der Erfahrung im persönlichen Umfeld durchaus geläufig, fehlt doch das öffentliche Bewusstsein davon, welch großer Anteil von Sozialtransfers der verschiedensten Kategorien faktisch die Funktion hat, den Austritt aus Erwerbsarbeit vorzuverlegen und eine Brücke zur Altersrente herzustellen.

¹ Die Enquête-Kommission „Demographischer Wandel“ wurde vom 12. Deutschen Bundestag im Oktober 1992 erstmalig eingesetzt (Enquête-Kommission 2002).

4. Da der öffentliche Diskurs vom Denken in statischen Quantitäten statt in Prozessen und Übergängen geprägt ist, ist die Befürchtung verbreitet, dass die Jugendarbeitslosigkeit in dem Maße steigen müsse, wie der Erwerbsaustritt der Älteren hinausgezögert wird. Inzwischen haben jedoch verschiedene Studien erhebliche Zweifel am Bestehen eines so einfachen Zusammenhangs geweckt (Sackmann 1997, Sörensen 1998, Bender, Preißler & Wübke 2000, Knuth 1999, Peters & Steiner 2000).
5. Es wird in der Regel übersehen, dass angesichts der großen Zahl von Betroffenen und der vergleichsweise kurzen Lebenszeitspanne, in der die Mehrzahl der Austritte erfolgt, eine Verlängerung des durchschnittlichen Alters des Erwerbsaustritts um wenige Monate oder gar um ein Jahr eine entscheidende Trendwende bedeuten und für die sozialen Sicherungssysteme einen relevanten Entlastungseffekt bringen würde. Das gilt insbesondere auch unter dem Gesichtspunkt, dass der historisch stärkste Geburtsjahrgang 1964, der im Jahre 2019 das 55. Lebensjahr vollenden wird, fast doppelt so stark besetzt ist wie der Geburtsjahrgang 1945, der sich derzeit im Altersübergang befindet und zudem eine höhere Erwerbsbeteiligung der Frauen aufweist. Anstatt die notwendigen kleinen Schritte einer wirklichen Trendwende anzugehen, tauchen periodisch politische Vorschläge zur Heraufsetzung des offiziellen Rentenalters auf, die sehr weit vom realen Altersübergangsgeschehen entfernt sind.

Aus dem Abstand internationaler Beobachter wurde deshalb die Situation des Altersübergangs in Deutschland als Politikblockade analysiert und mit der viel zitierten „Politikverdrossenheit“ in Zusammenhang gebracht (Mares 1999). Es bedurfte eines wirtschaftlichen Aufschwungs mit punktueller und vorübergehender Knappheit von Fachkräften, damit die Partner des „Bündnisses für Arbeit, Ausbildung und Wettbewerbsfähigkeit“ sich am 4. März 2001 auf einen Paradigmenwechsel bei der Beschäftigung älterer Arbeitnehmer² verständigen konnten. Die Wirksamkeit dieses hochrangig deklarierten gesellschaftlichen Konsens über die Notwendigkeit einer wirksamen Umsteuerung beim Altersübergang ist angesichts der in Deutschland anhaltenden und von einer weiteren Welle des Personalabbaus in Großunternehmen gekennzeichneten Rezession jedoch weiterhin fraglich.

Mit den seit 1992 durchgeführten Rentenreformen versucht die Gesetzgebung, dem Trend zu einem immer früheren Erwerbsausstieg entgegen zu wirken. Die schrittweise Anhebung des Eintrittsalters für die verschiedenen Arten vorgezogener Altersrenten sowie die Einführung von Rentenabschlägen für deren vorzeitige Inanspruchnahme sollten bei den nachwachsenden Geburtsjahrgängen eine grundsätzliche Verhaltensänderung hin zu einem späteren Renteneintritt bewirken. Durch verlängerte Beitragszeiten und eine verkürzte Rentenlaufzeit erhoffte man sich eine zweifache Entlastung der Rentenkassen. Soweit eine solche Verhaltensänderung nicht eintrat, sollten zumindest die Abschläge für Kostenneutralität der Frühverrentung sorgen.

Parallel zu den Rentenreformen wurde mit der Ablösung des Arbeitsförderungsgesetzes durch das SGB III im Jahre 1998 die Altersstaffel für verlängerte Bezugszeiten des Arbeitslosengeldes heraufgesetzt und die Ausnahme von ratiellen Sozialplanleistungen bei der Bedürftigkeitsprüfung für die Arbeitslosenhilfe aufgehoben, so dass die bisherigen Muster des „Vorruhestandes“ entweder für die Betriebe deutlich teurer oder für die Beschäftigten deutlich ungünstiger wurden. Die mit dem Namen „Hartz“ verbundenen Reformen setzen diese Entwicklung fort. Diese Regelungen werden aber erst in der Zukunft wirksam.

² Vgl. Deutscher Bundestag (2001) - Vorausgegangen war eine entsprechende Positionierung der Arbeitgeberverbände (vgl. BDA 2000).

Methodisches Vorgehen

Die Motive der Betriebe und Beschäftigten beim Altersübergang und die Mechanismen, unter denen er bisher gestaltet wird, sind weitgehend erforscht. Zielsetzung des Projektes ist es deshalb nicht, neue theoretische Erklärungen des Altersübergangsverhaltens zu liefern. Es geht vielmehr um die Entwicklung und Exploration neuer Möglichkeiten eines zeitnahen Monitorings des Altersübergangsgeschehens.

Das Berichtssystem Altersübergang wird aus der Kombination einschlägiger Datenquellen entwickelt und im Hinblick auf aktuelle Prozesse des Altersübergangs interpretiert. Im Vordergrund steht somit nicht der statistische Nachweis von Korrelationen, sondern die aktuelle, anschauliche und politisch relevante Beschreibung des Altersübergangsgeschehens und der Altersübergangsoptionen. Diesbezügliche, von Betrieben und Beschäftigten zwischen 1996 und 2003 erhobene Daten werden innerhalb der zweijährigen Pilotphase sekundäranalytisch untersucht. Zur Gewinnung von Vergleichsmaßstäben werden zudem Zeitreihen ab Anfang der neunziger Jahre zurückverfolgt, soweit die benutzten Datensätze das zulassen. Der Vergleich zwischen alten und neuen Bundesländern sowie zwischen Frauen und Männern wird sich als Standard durch sämtliche Auswertungen ziehen.

Zur Entwicklung eines „Altersübergangs-Reports“ werden folgende Datenquellen und Datensätze herangezogen:

- diverse Arbeitsmarktstatistiken der Bundesagentur für Arbeit
- Mikrozensus
- Rentenzugangsstatisik des VDR
- deutscher Alters-Survey
- IAB-Betriebspanel.

Im Hinblick auf eine notwendige Weiterentwicklung des Altersübergangs-Monitors werden die Nutzung weiterer Datenquellen sowie Möglichkeiten der direkten Verknüpfung von Datensätzen geprüft und konzeptionell vorbereitet. Das Projekt hat somit Pilotcharakter zur Entwicklung eines neuen Aspektes kontinuierlicher Sozialberichterstattung, für die es nach Ablauf der von der Hans-Böckler-Stiftung finanzierten zweijährigen Pilotphase andere Förderer bzw. institutionelle Träger zu gewinnen gilt.

Datenbedarf und Zugangswege

Die Rentenzugangsstatisik des Verbandes Deutscher Rentenversicherungsträger (VDR) gibt u.a. Auskunft darüber, in welchem Alter welche Anzahl von Versicherten in welche Art von Rente zugegangen ist. Im Kontext unserer Fragestellung beschränken wir uns auf Versichertenrenten, da die Hinterbliebenenrenten nicht unmittelbar mit dem Übergang aus der Erwerbstätigkeit in den Ruhestand verknüpft sind. Andererseits würde eine Beschränkung allein auf die Altersrenten dazu führen, ein eventuelles „Ausweichen“ auf Erwerbsminderungsrenten zu übersehen.

Zugänge in Erwerbsminderungsrenten in einem Alter unter 50 Jahren bleiben jedoch außer Betracht, weil sie nicht zur Problematik des Altersübergangs gehören.

Um Verhaltensänderungen beim Renteneintritt untersuchen zu können, bedarf es kohortenspezifischer Analysen.³ Die Analyse des Renteneintritts nach Geburtskohorten ist beim VDR bislang jedoch nur für abgeschlossene, d.h. vollständig dem Erwerbsalter entwachsene Kohorten entwickelt (vgl. Kruse 2001); für Analysen am aktuellen Rand braucht man dagegen ein Verfahren zur Analyse un abgeschlossener Geburtskohorten, das beim VDR zwar in Entwicklung, während der Projektlaufzeit jedoch noch nicht verfügbar ist. Die im Hinblick auf den Renteneintritt von uns zu behandelnden Fragestellungen werden deshalb anhand der uns vom Verband Deutscher Rentenversicherungsträger (VDR) per Datenträger zur Verfügung gestellten aggregierten Rentenzugangsstatisiken untersucht. Gegenüber den vom VDR im Internet öffentlich zugänglich gemachten Statistiken ermöglichen diese differenziertere altersspezifische Analysen. Auf eine dennoch bestehende Beschränkung dieser Daten muss hingewiesen werden:

- Es handelt sich um nach zahlreichen Merkmalen aggregierte Tabellen von Rentenzugängen, nicht um individuelle Daten des Altersübergangsverlaufs von Personen. Insofern sind unterschiedliche rentenrechtliche Ereignisse individuell nicht zuzuordnen.
- Diese Daten geben auch keine Auskunft darüber, ob die Personen, die in die Zählung der Rentenzugänge eingegangen sind, unmittelbar vorher beschäftigt waren oder nicht. Lediglich bei den vorzeitigen Altersrenten wegen Arbeitslosigkeit oder nach Altersteilzeit ergibt sich der vorausgegangene Status aus der Definition der Zugangsvoraussetzungen.
- Die für Insider verfassten Datendokumentationen und -erläuterungen sind für externe Nutzer zudem kaum verständlich und müssten für die Zukunft besser lesbar werden.
- Die Bereitstellung der VDR-Daten in der aktuellen Form (Software/Bedienungsoberfläche) ist wenig nutzerfreundlich; die Zusammenstellung der für das Projekt erforderlichen Rentenzugangsdaten war sehr zeitaufwändig.

Für die Zukunft wäre deshalb die Bereitstellung von Individualdaten in Form von Scientific Use Files wünschenswert, die neben persönlichen, berufsbiographischen und rentenrelevanten Merkmalen auch Informationen zum Haushaltskontext bzw. Erwerbsstatus des Ehepartners beinhalten sollten.

- Die Rentenversicherungsdaten sind durch die Träger der Rentenversicherung jedoch bereits soweit anonymisiert, dass Individualdaten aus verschiedenen Zeitwellen und Ereignissen, z.B. Reha- und Rentenereignissen, nicht zusammengebracht werden können.
- Da es sich bei den Daten der Rentenversicherungsträger um prozessproduzierte Daten handelt, werden die künftig vom Forschungsdatenzentrum bereitgestellten Datensätze die Bedarfe der Wissenschaft wohl auch zukünftig nicht hinreichend abdecken und die Behandlung von bestimmten Untersuchungsfragen, wie etwa den Haushaltskontext, nicht erlauben.

³ Ein Beispiel: Wenn die Geburtskohorte X zu Y Prozent bereits mit Erreichen des 60. Lebensjahres in den Bezug einer Altersrente eintrat, die Geburtskohorte 1942 dagegen, die 2002 das 60. Lebensjahr erreichte, nur zu Z Prozent, dann hat sich am Altersübergangverhalten offensichtlich etwas verändert.

- Eine für die Wissenschaft wünschenswerte Datenintegration durch den VDR, die umfassende Informationen zu individuellen Versicherungsverläufen liefern könnte, wird durch interne datenschutzrechtliche Bestimmungen und die unterschiedliche Interessenlage der 26 Verbandsmitglieder bzw. Sozialversicherungsträger erschwert.

Literatur

- BDA, Bundesvereinigung der Deutschen Arbeitgeberverbände (2000): Erhöhung der Erwerbsbeteiligung Älterer Arbeitnehmer. Ein Diskussionspapier der BDA. Berlin: BDA.
- Bender, S., Preißler, J. & Wübbecke, C. (2000): Betriebliche Determinanten des Generationenaustausches in westdeutschen Betrieben. Eine Untersuchung auf der Basis des IAB-Employer-Employee-Datensatzes für die Jahre 1994 und 1995. In: George, R. & Struck, O. (Hg.): *Generationenaustausch im Unternehmen*. München: Hampp.
- Deutscher Bundestag, Ausschuss für Arbeit und Sozialordnung (Hg.) (2001): *Ergebnisse des 7. Spitzengesprächs des Bündnisses für Arbeit, Ausbildung und Wettbewerbsfähigkeit am 4. März 2001*. Berlin: Ausschussdrucksache 14/1345.
- Enquête-Kommission „Demographischer Wandel“ (2002): *Herausforderungen unserer älter werdenden Gesellschaft an den einzelnen und die Politik. Schlussbericht*. Bundestags-Drucksache 14/8800 vom 28.03.2002.
- Knuth, M. (1999): Senkung der Arbeitslosigkeit durch Ausstieg aus dem Vorruhestand. Gesellschaftliche und betriebliche Innovationserfordernisse im Umgang mit dem strukturellen und demographischen Wandel. In: Brödner, P., Helmstädter, E. & Widmaier, B. (Hg.): *Wissens- teilung*, München: Hampp, 107-144.
- Kohli, M. (1985): Die Institutionalisierung des Lebenslaufs. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 1, 1-29.
- Kohli, M. (1994): Institutionalisierung und Individualisierung der Erwerbsbiographie. In: Beck, U. & Beck-Gernsheim, E. (Hg.): *Risikante Freiheiten*. Frankfurt a. Main: Suhrkamp, 219-244.
- Koller, B. (2001): Das Rentenalter wurde angehoben - zieht der Arbeitsmarkt mit? Eine Analyse zum Übergang in Rente, zu Erwerbsbeteiligung und Arbeitslosigkeit Älterer. *IAB-Werkstattbericht* Nr. 7/01 v. 29.6.
- Kruse, E. (2001): Kohortenspezifische Analyse des Rentenzugangs von Versichertenrenten. In: Barkholdt, C. (Hg.): *Prekärer Übergang in den Ruhestand*. Opladen: Westdeutscher Verlag.
- Mares, I. (1999): The evolution of early retirement: explaining policy blockage and policy frustration. Washington: *American Institute for Contemporary German Studies Policy Papers* no. 11.
- Peters, R.-H. & Steiner, V. (2000): *Auswirkungen einer Verkürzung der Lebensarbeitszeit auf Beschäftigung, Wachstum und Wettbewerbsfähigkeit der Metall- und Elektroindustrie*. Kurzfassung zum Gutachten für den Arbeitgeberverband Gesamtmetall. Mannheim: ZEW.
- Sackmann, R. (1997): Der Einfluß von Verrentungsprozessen und Mobilitätsprozessen auf die Arbeitsmarktrisiken von Berufseinsteigern. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 3, 675-687.
- Sørensen, O.B. (1998): Unterschiedliche Praktiken in bezug auf das Rentenalter: Angemessene Reaktion auf die Entwicklung des Arbeitsmarktes? *Deutsche Rentenversicherung*, 3-4, 179-208.

Renate Büttner ist Wissenschaftliche Mitarbeiterin im Forschungsschwerpunkt „Entwicklungstrends des Erwerbssystems“ im Institut Arbeit und Technik und Projektmitarbeiterin.

PD Dr. Matthias Knuth ist Forschungsdirektor im Forschungsschwerpunkt „Entwicklungstrends des Erwerbssystems“, wissenschaftlicher Geschäftsführer im Institut Arbeit und Technik und Projektleiter.

Der **Forschungsschwerpunkt „Entwicklungstrends des Erwerbssystems“** stellt empirisch fundiertes Orientierungswissen über die mittel- und langfristige Entwicklung des Erwerbssystems bereit, schwerpunktmäßig zu folgenden Themenfeldern: 1) Dynamik und Stabilität von Beschäftigung, Mobilität der Arbeitskräfte, Förderung des Beschäftigtentransfers bei Personalanpassung, 2) Entwicklung von Humankapital und Innovationsfähigkeit, 3) Steigerung der Erwerbsbeteiligung von Älteren, Frauen und MigrantInnen, 4) Entwicklung des Übergangsverhaltens von der Erwerbs- in die Ruhestandsphase und ihre Auswirkung auf die sozialen Sicherungssysteme, 5) Umgestaltung des Arbeitsmarktregimes.

Das **Institut Arbeit und Technik** ist Teil des Wissenschaftszentrums Nordrhein-Westfalen, unter dessen Dach das Wuppertaler Institut für Klima, Umwelt und Energie sowie das Kulturwissenschaftliche Institut in Essen ebenfalls angesiedelt sind. Gemeinsam mit den drei Instituten organisiert das Wissenschaftszentrum NRW den Dialog zwischen Wissenschaft, Wirtschaft, Kultur und Politik. Das Institut Arbeit und Technik ist ein anwendungsorientiertes Forschungs- und Entwicklungsinstitut des Landes Nordrhein-Westfalen. Sein Auftrag ist, innovative Lösungswege für die Bewältigung des Strukturwandels zu entwickeln und zu erproben. Maßstäbe des Erfolges sind dabei vor allem die Verbesserung der Beschäftigungslage und die Nachhaltigkeit der Wertschöpfung. (siehe <http://iat-info.iatge.de>).

Intragenerationale Umverteilungswirkungen in der gesetzlichen Rentenversicherung

Hans-Martin von Gaudecker

(Mannheimer Forschungsinstitut Ökonomie und demographischer Wandel)

Einleitung

Ein positiver Zusammenhang zwischen Einkommen und Sterblichkeit ist hinlänglich bekannt. Dieses Projekt wird versuchen, diese Beziehung für Rentner der deutschen gesetzlichen Rentenversicherung (GRV) nachzuweisen und zu quantifizieren. Ein positiver Zusammenhang führte dazu, dass Gutverdiener aufgrund ihrer längeren Lebenserwartung einen höheren Barwert der Rentenzahlungen hätten. Diese regressive Umverteilung würde die sonstigen Effekte des deutschen Steuer- und Sozialsystems zumindest teilweise konterkarieren. Auch diese Umverteilungswirkung soll quantifiziert werden.

Fragestellung und Methodik

Der Zusammenhang von Einkommen und Sterblichkeit stößt sowohl in der theoretischen als auch in der empirischen internationalen Literatur auf großes Interesse.¹ Die meisten Studien beschränken sich wegen des Simultanitätsproblems auf eine Korrelation der beiden Größen und gehen nicht auf Kausalitätsmechanismen ein. Ein Problem dieser Arbeiten ist, dass sie auf Umfragedaten beruhen, die lediglich das momentane Einkommen erfassen. Datensätze des Verbands Deutscher Rentenversicherungsträger e.V. (VDR) sind demgegenüber besonders aufschlussreich, da sie eine Messgröße für das Lebenseinkommen beinhalten. Dieses sollte aus theoretischer Sicht die höchste Relevanz für die Sterblichkeit haben. Im Gegensatz dazu beinhalten Umfragedaten lediglich punktuelle Angaben zum momentanen Einkommen. Diese können durch kurzfristige Phänomene beeinflusst werden (z.B. Arbeitslosigkeit, Börsenschwankungen, etc.), welche vermutlich wenig Einfluss auf die Sterblichkeit haben. Gesundheitsvorsorge und -versorgung, Lebensgewohnheiten und ähnliche Faktoren hängen vor allem mit dem langfristigen Einkommen zusammen. Die in VDR-Daten vorhandenen Renteneinkommen bzw. pro Jahr gesammelten Entgeltpunkte bieten eine ideale Möglichkeit, von kurzfristigen Einflüssen zu abstrahieren.

Von Interesse in diesem Projekt ist der Zusammenhang von Einkommen und Sterblichkeit innerhalb von Kohorten, d.h. Personen eines bestimmten Geburtsjahrgangs. Die Daten des VDR sind jedoch periodenweise aufgebaut. Mit anderen Worten, sie beinhalten mehrere reale Kohorten in bestimmten Jahren. Da die Rente in Deutschland nicht kohortenspezifisch ist, lassen sich Renteneinkommen und Entgeltpunkte über alle Personen dieser Geburtsjahrgänge hinweg vergleichen. So wird es unter der Annahme eines zeitkonstanten Zusammenhangs zwischen Einkommen und Sterblichkeit möglich, daraus eine fiktive Kohorte zu konstruieren und Lebenserwartungen in Abhängigkeit vom Einkommen zu berechnen.

¹ Für einen exzellenten Überblick vgl. Smith (1999).

In einer Vorstudie zu diesem Projekt (v. Gaudecker 2004), für die auf Kreisebene aggregierte Daten für Baden-Württemberg zur Verfügung standen, wurde die prinzipielle Machbarkeit dieses Ansatzes gezeigt. Für jede Altersklasse ab 65 Jahre wurde die Überlebenswahrscheinlichkeit in Abhängigkeit vom Einkommen berechnet (es wurden 35 separate Probit-Modelle geschätzt). Daraus konnten Lebenserwartungen als Funktion der Renteneinkünfte bestimmt werden. Die Ergebnisse erscheinen plausibel und ermutigen, mit diesem Ansatz weiter zu arbeiten. Sie sind jedoch aufgrund der Kombination von nichtlinearen Modellen und aggregierten Daten zum derzeitigen Zeitpunkt nicht statistisch belastbar.

Wird die Lebenserwartung mit Individualdaten bestimmt, so könnten für verschiedene Einkommensklassen interne Zinsfüße (d.h. implizite Renditen des Umlageverfahrens) berechnet werden. Dies vermittelt einen Eindruck von den Umverteilungswirkungen des Rentensystems, die aus der differentiellen Mortalität entstehen. Die Hypothese ist hier, dass Personen mit einem hohen Renteneinkommen tendenziell länger leben und damit auch länger in den Genuss einer Rente kommen. Dies würde die generelle Umverteilung durch das deutsche Steuer- und Sozialsystem zumindest teilweise wieder kompensieren, mit derzeit kontrovers diskutierten sozialpolitischen Schlussfolgerungen (vgl. z.B. BMGS 2003). Um die Umverteilungswirkungen zu bestimmen, würden Barwertberechnungen für Annuitäten herangezogen (Mitchell et al. 1999).

Möglicher Erkenntnisgewinn und Einordnung in die Literatur

Diese Studie reiht sich in diverse Publikationen zur Mortalität der Versicherten in der Gesetzlichen Rentenversicherung ein. Müller und Rehfeld analysieren die Gesamtsterblichkeit von Rentnern und differenzieren nach Dauer der Beitragszahlung und Erhalt von Erwerbsunfähigkeitsrenten (Müller & Rehfeld 1985). Klosterhuis und Müller-Fahrnow finden eine deutlich geringere Sterblichkeit von sozialversicherungspflichtig Beschäftigten mit hohem Einkommen, betrachten jedoch nicht die Rentenbezieher (Klosterhuis & Müller-Fahrnow 1993). Kruse untersucht diese Personengruppe, er findet allerdings keine Korrelation zwischen Einkommen und Sterblichkeit (Kruse 2000). Diesen Studien ist gemein, dass sie im Wesentlichen auf statistisch-deskriptiven Ansätzen beruhen. In diesem Projekt sollen die Möglichkeiten ausgelotet werden, ökonometrische Methoden auf die oben angesprochene Thematik anzuwenden. Ferner geht es über die eben genannten Studien dahingehend hinaus, dass explizit differenzierte Renditen berechnet werden sollen.

Datenbedarf

Der Datenbedarf bezieht sich auf die Rentenwegfalls- und Rentenbestandsstatistiken. Aufgrund von starken Kohorteneffekten bei Frauen wird sich die Analyse nur für Männer durchführen lassen, dieser Punkt hat sich in der Vorstudie bereits herauskristallisiert. Konkret sehen Minimal- und Maximalvorstellungen wie folgt aus:

Minimum:

- Nach Alter geschichtete Stichprobe für Rentner ab 65 Jahren aus Rentenwegfalls- und Rentenbestandsstatistik für ein Jahr. Besonders hohe, dünn besetzte Alter könnten abgeschnitten oder zusammengefasst werden.
- Merkmale: Geburtsjahr, Rentenzahlbetrag, Zahl der Beitragsjahre, durchschnittliche Entgeltpunkte pro Beitragsjahr (wie in dem Tabellensatz für Baden-Württemberg).

Maximum:

- Vollerhebung der Rentner ab 65 Jahren aus Rentenwegfalls- und Rentenbestandsstatistik für ein Jahr. Besonders hohe, dünn besetzte Alter könnten abgeschnitten oder zusammengefasst werden.
- Merkmale: Geburtsjahr, Rentenzahlbetrag und detaillierte Erwerbsbiographie. Die genaue Variablenausprägung bezüglich des aktiven Versichertenlebens könnte geklärt werden, wenn es aktuell wird.

Zugangswege und Realisierbarkeit

- Gewünscht: Scientific Use File.
- Akzeptierbar: Gastwissenschaftler-Arbeitsplatz, kontrollierte Datenfernverarbeitung.

Nach den Diskussionen beim Workshop halte ich den Minimalwunsch auf jeden Fall für realisierbar. Die Maximalvorstellungen könnten an Datenschutzgründen scheitern, ein Kompromiss wird sicherlich möglich sein. Generell ließe sich bereits mit dem Minimalprogramm eine Menge aussagen. Die genaueren Erwerbsbiographieangaben dienen vor allem dazu, sich auf Personen zu beschränken, die einen angemessen hohen Anteil ihres Erwerbslebens sozialversicherungspflichtig beschäftigt waren.

Literatur

- Bundesministerium für Gesundheit und Soziale Sicherung (Hg.) (2003): *Nachhaltigkeit in der Finanzierung der Sozialen Sicherungssysteme: Bericht der Kommission*. Berlin.
- von Gaudecker, Hans-Martin (2004): *Zum Zusammenhang von Einkommen und Sterblichkeit bei Rentnern der gesetzlichen Rentenversicherung*. Unveröffentlichtes Manuskript, Universität Mannheim.
- Klosterhuis, H. & W. Müller-Fahrnow (1993): Sozialschicht und Sterblichkeit bei männlichen Angestellten aus den alten Bundesländern. In: Mielck, A. (Hg.): *Krankheit und soziale Ungleichheit*. Opladen: Leske und Budrich.
- Kruse, E. (2000): Sterblichkeit und fernere Lebenserwartung von Versichertenrentnern der gesetzlichen Rentenversicherung - Aktuelle Ergebnisse einer Sterbetafelberechnung für 1996/1998. *Deutsche Rentenversicherung*, 3-4, 121-138.

Mitchell, O., Poterba, J., Warshawsky, M. & Brown, J. (1999): New Evidence on the Money's Worth of Individual Annuities. American Economic Review, December, 1299-1318.

Müller, H.-W. & U. Rehfeld (1985): Zur Rentnersterblichkeit unter besonderer Berücksichtigung langjährig berufstätiger Frauen und Männer. Blätter der Deutschen Gesellschaft für Versicherungsmathematik, 17(2), 141-162.

Smith, J. (1999): Healthy Bodies and Thick Wallets: The Dual Relation between Health and Socioeconomic Status. Journal of Economic Perspectives, 13 (2), 145-166.

Hans-Martin von Gaudecker, Diplom-Volkswirt, MEA und Zentrum für wirtschaftswissenschaftliche Doktorandenstudien (CDSEM), Universität Mannheim. Forschungsinteressen: Mikroökonomie, Renten- und Gesundheitsökonomie, Versicherungsökonomie, Mortalität und Langlebigkeit.

Mannheimer Forschungsinstitut Ökonomie und demographischer Wandel (MEA), Direktor Axel Börsch-Supan, Ph.D. Das MEA ist ein Forschungsinstitut in der Fakultät für Volkswirtschaftslehre der Universität Mannheim. Aufgabe des MEA ist es, mikro- und makroökonomische Aspekte des demographischen Wandels zu antizipieren und zu begleiten. Durch den Aufbau von empirischen Modellen und daraus folgenden Prognosen leitet MEA Handlungsempfehlungen für Wirtschaft und Politik ab (siehe www.mea.uni-mannheim.de).

Rentenversicherung und heterogene Lebenserwartung

Stefan Hupfeld (Universität Konstanz)

Einleitung

Mit den erstmals vom FDZ-RV zur Verfügung gestellten Daten wollen wir Phänomene wie Umverteilungsströme und Frühverrentungsverhalten im Rahmen der gesetzlichen Rentenversicherung analysieren. Ausgangspunkt ist dabei die internationale empirische Evidenz für einen Gleichlauf von sozioökonomischen Größen wie Einkommen und Vermögen mit Lebenserwartung oder Gesundheitszustand. Neueste Untersuchungen bestätigen diesen Zusammenhang für Deutschland (siehe z.B. Reil-Held (2002) für eine Untersuchung anhand des Sozioökonomischen Panels). Theoretische und methodische Überlegungen legen allerdings nahe, dass die kausale Richtung dieses Zusammenhangs uneindeutig ist und sich die verschiedenen Größen gegenseitig beeinflussen.

Wenn wir zunächst die Existenz eines solchen Zusammenhangs unterstellen, führt die Auszahlung der gesetzlichen Rente als Annuität zu Umverteilung von Beziehern niedriger Renten zu Beziehern von hohen Renten. Berücksichtigt man jedoch den individuellen Zeitpunkt des Renteneintritts, der aus gesundheitlichen Gründen oder (in Grenzen) auf Grund freier Entscheidung erfolgen kann, ist die Richtung des Effektes nicht mehr eindeutig.

Anhand von Daten des VDR soll untersucht werden, ob und in welcher Größenordnung sich diese Argumentation bestätigen lässt.

Fragestellung und Ziel

Im deutschen System der gesetzlichen Rentenversicherung kann es insbesondere aus zwei Gründen zu Umverteilung zwischen verschiedenen Einkommensgruppen kommen. Erstens hängt schon vor Renteneintritt ein höheres Einkommen mit besserer Gesundheit und höherer Lebenserwartung zusammen. Die Höhe der später ausgezahlten Renten bemisst sich nach Dauer und Höhe der vorangegangenen einkommensabhängigen Einzahlungen; nimmt man den Zeitpunkt des Renteneintritts einmal als gegeben, so erhält ein Bezieher hoher Einkommen nicht nur monatlich eine höhere Rente, sondern auch die Rentenbezugsdauer ist erwartungsgemäß länger als bei Rentnern mit zuvor niedrigen Einkommen. Zweitens hat ein Bezieher hoher Renten zusätzlich die Möglichkeit, in seine Gesundheit zu investieren und damit seine Lebenserwartung weiter zu erhöhen. Beide Mechanismen benachteiligen die unteren Einkommensgruppen. Anhand der VDR-Daten wollen wir untersuchen, ob und in welcher Größe diese Umverteilung in der Realität auftritt.

Da in gewissem Rahmen der Zeitpunkt des Renteneintritts aber vom Versicherten bestimmt werden kann, ist es möglich, dass diese Umverteilungswirkungen durch individuelle Entscheidungen ganz oder teilweise kompensiert werden. Ein Bezieher eines niedrigen Einkommens kann sich entweder in Antizipation dieser für ihn nachteiligen Wirkungen oder aufgrund des

durchschnittlich schlechteren Gesundheitszustandes für einen früheren Renteneintritt entscheiden, wodurch sich seine Rentenbezugsdauer erhöht. Hier ist wiederum die Frage, ob sich ein solcher Zusammenhang zwischen Renteneintrittsalter und Einkommens- bzw. Rentenhöhe auch statistisch zeigen lässt.

Voraussetzung für die Vermutungen ist die Korrelation von Wohlstand und Lebenserwartung. Dieser Zusammenhang wurde für Deutschland bislang kaum untersucht. Deshalb soll diese Fragestellung soweit wie möglich mit den VDR-Daten untersucht werden.

Der gesamte Fragenkomplex wird durch theoretische Arbeiten begleitet, aus denen empirisch überprüfbare Aussagen abgeleitet werden sollen. Wir gehen davon aus, dass Individuen in verschiedenen Eigenschaften heterogen sind. Diese Heterogenität äußert sich in unterschiedlichen Bildungsniveaus, unterschiedlicher Akkumulation von Vermögen und unterschiedlichen Investitionen in die eigene Gesundheit und Lebenserwartung. Vor diesem Hintergrund trifft der Versicherte eine rationale Entscheidung über den Zeitpunkt seines Renteneintritts.

Theoretische Einbettung

Eine erste theoretische Erklärung für den Gleichlauf von Wohlstand und Gesundheit findet sich im Gesundheitskapitalmodell nach Grossman (1972). Hier hat ein Individuum in Abhängigkeit seiner Präferenzen und seiner finanziellen Ausstattung die Möglichkeit, Zeit oder Geld in seine Gesundheit zu investieren. Ein besserer Gesundheitszustand führt zunächst sofort zu höherem Wohlbefinden. In Folge erlaubt dieser verbesserte Gesundheitszustand die Erzielung eines höheren Einkommens, welches für Konsum oder weitere Investitionen in Gesundheit genutzt werden kann. Gesundheit hat hier also zwei Eigenschaften, nämlich die eines Konsumguts und die einer Investition.

Eine treibende Größe sowohl für Einkommen als auch für das Wissen um die Wirksamkeit von Gesundheitsinvestitionen ist der Bildungsstand. In empirischen Untersuchungen erweist er sich in der Regel als signifikant positiv (z.B. Adams et al. 2003) für die Höhe des Einkommens und den Gesundheitsstatus. Auch in den theoretischen Grundlagen, schon beginnend mit Grossman (1972), wird die Rolle der Bildung in diesem Rahmen erkannt. Es ist aus empirischer Sicht noch unsicher, welche der Größen tatsächlich die treibende Kraft ist und in welche Richtung die kausalen Beziehungen verlaufen. So finden Meer et al. (2002) und Adams et al. (2003) trotz vorhandener Korrelation nach Berücksichtigung von Endogenität zwischen den wichtigsten Variablen nur noch schwache kausale Beziehungen zwischen Wohlstand und Gesundheit.

Um schließlich einen theoretischen Hintergrund für Art und Größe der Umverteilung zu finden, bieten sich Modelle aus dem Bereich der impliziten Besteuerung an. Hier werden Einzahlungen und erwartete Auszahlungen aus dem Rentensystem miteinander in Beziehung gesetzt, um so die implizite Steuerlast eines Individuums oder einer Kohorte zu bestimmen. Die Arbeiten von Diamond (2003) bieten hierfür einen Ausgangspunkt. Mögliche Folgerungen für die Gestaltung des Rentensystems finden sich bei Breyer (1997, 2004).

Methodik

Die Endogenität (also die gegenseitige Beeinflussung) der interessierenden Größen wie Einkommen, Rentenhöhe und Lebenserwartung bzw. Rentenbezugsdauer verlangt nach Verfahren, die dieses Problem berücksichtigen. Je nach Verfügbarkeit in den Daten müssen geeignete Instrumentenvariablen gefunden werden, um unverzerrte Schätzungen zu erhalten. Da es sich um Individualdaten handelt, sollen Verfahren aus der Mikroökonomie angewendet werden. Die Fragestellung verlangt eigentlich nach den abgeschlossenen Versichertenbiographien einer Kohorte; da die Daten aber zum deutlich größeren Teil aus nicht abgeschlossenen Versichertenleben bestehen, muss diese Zensur auch methodisch berücksichtigt werden.

Datenbedarf

Die ideale Datenbasis besteht aus vollständigen (individuellen) Versichertenbiographien einer bzw. mehrerer Alterskohorten. Um die Effekte genau zu identifizieren, sind Kovariablen nötig, die über soziodemographische Charakteristika Auskunft geben, unter denen der individuelle Bildungsstand eine prominente Rolle spielt. Sehr nützlich wären auch Angaben über weitere Einkommensarten und Vermögen der Versicherten. Um die Frage nach der kausalen Richtung der Beziehungen zu behandeln, wären des Weiteren Paneldaten wünschenswert.

Die verschiedenen Rentenstatistiken (Zugangs-, Änderungs- und Wegfallstatistik), die Versicherungskontenstichprobe und die Sondererhebung „Vollendete Versichertenleben“ entsprechen dem Datenbedarf am ehesten. Langfristig ist es wünschenswert, eine Zusammenführung der verschiedenen Rentenstatistiken und zu verschiedenen Zeitpunkten zu ermöglichen.

Alle Datensätze sollten aus methodischen Gründen so vollständig wie möglich sein, besonders mit Blick auf die Anzahl der Kovariablen. Genauso ist es wünschenswert, dass die Anonymisierungsverfahren die Aussagekraft der Daten so wenig wie möglich vermindern.

Zugangswege

Der gewünschte Zugangsweg besteht in Scientific Use Files für die Analyse am eigenen Arbeitsplatz. Alle anderen Zugangswege, wie ein Gastwissenschaftler-Arbeitsplatz oder kontrollierte Datenfernverarbeitung, sind akzeptabel und werden angestrebt, falls die Veröffentlichung der Scientific Use Files scheitern sollte. Besonders in den letztgenannten Fällen wäre die frühe Veröffentlichung von simulierten oder stark anonymisierten Daten hilfreich, um Programme entsprechend vorzubereiten und zu testen.

Zusammenfassung

Die Veröffentlichung der Statistiken des VDR durch das FDZ-RV bietet die erste Möglichkeit, Fragestellungen mit Individualdaten der gesetzlichen Rentenversicherung zu untersuchen. Besonders die Tatsache, dass es sich bei einigen Datensätzen um Vollerhebungen handelt, kommt der potentiellen Aussagekraft der Ergebnisse zu Gute. Aus den Ergebnissen lassen sich mög-

licherweise Erkenntnisse über Phänomene wie Umverteilungsströme und Frühverrentungsverhalten gewinnen.

Leider fehlen in den Daten Angaben zu Vermögen und Einkünften aus anderen Quellen als abhängiger Erwerbsarbeit. Für die Zukunft ist zu hoffen, dass sich die verschiedenen Rentenstatistiken zusammenfügen lassen, um einen möglichst genauen Einblick in die Versichertenbiographien zu erhalten und diese über die Zeit zu verfolgen.

Literatur

- Adams, P., McFadden, D. & Ribeiro, T. (2003):* Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status. *Journal of Econometrics*, 112, 3-56.
- Breyer, F. (1997):* Sind äquivalente Renten fair? In: Hauser, R. (Hg.): *Reform des Sozialstaats*. Berlin, 169-180.
- Breyer, F. (2004):* Auf Leben und Tod – steigende Lebenserwartung und Sozialversicherung. *Perspektiven der Wirtschaftspolitik*, 5, 227-241.
- Diamond, P.A. (2003):* *Taxation, Incomplete Markets, and Social Security*. Cambridge MA: MIT Press.
- Grossman, M. (1972):* *The Demand for Health: A Theoretical and Empirical Investigation*. New York: Columbia University Press of the National Bureau of Economic Research.
- Reil-Held, A. (2002):* *Die Rolle intergenerationaler Transfers in Einkommen und Vermögen der älteren Menschen in Deutschland*. Dissertation. Universität Mannheim.

Stefan Hupfeld, Dipl.-Volkswirt. Wissenschaftlicher Angestellter am Lehrstuhl Prof. Dr. Friedrich Breyer, Universität Konstanz, Teilnehmer am Doktorandenprogramm „Quantitative Economics and Finance“ und Assoziation mit der DFG Forschergruppe „Heterogene Arbeit“. Forschungsinteresse: Optimale Gestaltung der Rentenversicherung und heterogene Lebenserwartung (siehe <http://wipol.wiwi.uni-konstanz.de/deutsch/hupfeld.htm>).

Prof. Dr. Friedrich Breyer ist Inhaber des Lehrstuhls für Volkswirtschaftslehre, insbesondere Wirtschafts- und Sozialpolitik, an der Universität Konstanz. Das Projekt „Lebensarbeitszeit, Humankapitalbildung und Transfersystem“ ist Teil der von der DFG geförderten Forschergruppe „Heterogene Arbeit: Positive und normative Aspekte der Qualifikationsstruktur“.

Erwerbstätigkeit und Altersmortalität

Prof. Dr. Barbara Riedmüller und Michaela Willert (Freie Universität Berlin)

Einleitung

Der sich vollziehende demographische Wandel und der sich vergrößernde Altenquotient sowie die damit verbundenen künftigen finanziellen Belastungen des Sozialbudgets bestimmen die Diskussion über die gegenwärtigen Reformen im System der deutschen Alterssicherung. Als Gründe für die weiterhin steigende Lebenserwartung der Bevölkerung in entwickelten Ländern werden die besseren sozio-ökonomischen Rahmenbedingungen und Fortschritte im Gesundheitswesen genannt. Dieser Trend wurde in der Vergangenheit besonders von der Langlebigkeit der Frauen getragen, deren Lebenserwartung die der Männer um mehrere Jahre überstieg. Zugleich sind die vergangenen Jahrzehnte durch gesellschaftliche Veränderungen, insbesondere bezüglich der Rolle der Frauen auf individueller, familialer und sozialer Ebene, geprägt. Das führt in zunehmendem Maße zu einer Annäherung von weiblichen und männlichen Erwerbsverläufen. Welchen Einfluss langjährige Erwerbstätigkeit auf die Altersmortalität von Männern und Frauen hat, ist derzeit aber nur unzureichend wissenschaftlich beleuchtet.¹ Im hier skizzierten Projekt soll auf Basis von Daten der Deutschen Rentenversicherung die These untersucht werden, ob langjährige Erwerbsarbeit von Frauen positiv mit höherer Altersmortalität korreliert ist. Ein solcher Zusammenhang zwischen Erwerbstätigkeit und Altersmortalität hätte bei steigenden Erwerbsquoten von Frauen maßgebliche Auswirkungen auf die Prognose der ferneren Lebenserwartung und damit die Größe der Alten-Population und Rentenbezieher in einer Gesellschaft.

Forschungsstand und Herleitung der These

Die Unterschiede in der Mortalität von Frauen und Männern sind Gegenstand einiger Untersuchungen.² Zum einen wurden Belege für biologische bzw. genetische Unterschiede gefunden, wie die unterschiedliche Anfälligkeit der Geschlechter für ischämische Herzerkrankungen. Diese wurden jedoch um unterschiedliche geschlechtsspezifische Verhaltensdispositionen erweitert, die in jungen Erwachsenenjahren zu einer höheren Sterblichkeit bei Männern führen (z.B. risikoreicheres Verhalten durch das männliche Sexualhormon Testosteron). Weiterhin spielen sozio-ökonomische Faktoren eine Rolle. So wird angenommen, dass ein gefährlicheres Arbeitsumfeld bzw. körperlich schwere Tätigkeiten die Mortalitätsraten männlicher Erwachsener erhöhen. Innerhalb dieser Gruppe ist die Erkenntnis relativ gut abgesichert, dass soziale Unterschiede zu Mortalitätsdifferenzen führen (vgl. Liang et al. 2002, Reil-Held 2000). Ein geringerer sozialer Status führt durch schlechtere Ernährungsgewohnheiten, schwierigere Arbeitsbedingungen und ein höheres Risikoverhalten zu höheren Mortalitätsraten. Inwieweit Gleiches für Frauen zutrifft, wurde bislang nicht eindeutig belegt. Allerdings wurde in verschiedenen Ländstudien³ gezeigt, dass bestimmte Krankheiten als Todesursache mit dem sozio-ökonomi-

¹ Eine Untersuchung von Müller und Rehfeld aus dem Jahr 1985 bezieht sich auf Geburtsjahrgänge vor 1917, deren Lebensbedingungen nicht vergleichbar sind (Müller & Rehfeld 1985).

² Für einen Überblick vgl. Gjonca et al. (1999).

³ Für einen Literaturüberblick vgl. Martikainen (1995).

schen Status negativ korreliert sind. Der Grund für die Schwierigkeiten bei der Analyse liegt vor allem darin, dass Konzepte sozialer Stratifikation, die sich auf den beruflichen Hintergrund beziehen, auf Frauen zum Teil nur schwer übertragen werden können. So ist für auf dem Arbeitsmarkt inaktive Frauen der potenzielle Berufsstatus nicht zu bestimmen, oder Frauen sind aufgrund von Familienpausen in einer niedrigeren beruflichen Position, als es ihrer Ausbildung oder einer vorhergehenden Tätigkeit entspricht. Der Versuch, den Status des Ehepartners oder das Haushaltseinkommen als Indikator für den sozialen Status der Frau zu verwenden, scheitert häufig an der Datenlage. Zudem wird angenommen, dass z.B. im Gegensatz zu Männern der Ehestand ein weiteres soziales Distinktionsmerkmal ist,⁴ ein Indikator, der ebenfalls nicht immer erhoben wurde. Schließlich treten häufig Kohorteneffekte als Störfaktoren für die Analyse von Mortalitätsunterschieden sowohl im Vergleich mit Männern als auch innerhalb der weiblichen Population auf.

Gerade diese Effekte sollen im hier skizzierten Projekt näher beleuchtet werden. Die Lebens- und Arbeitsbedingungen jüngerer und älterer Rentnerinnenkohorten unterschieden sich stark voneinander. Die Erwerbsquoten von Frauen in der Bundesrepublik stiegen, das Familienmodell des männlichen Ernährers wurde durch das der zumindest hinzuverdienenden Ehefrau ergänzt. Aus dem 2-Phasenmodell weiblicher Erwerbsverläufe (Erwerbstätigkeit und Ausstieg mit Familiengründung) wurde das 3-Phasenmodell mit erneuter Erwerbstätigkeit nach der Familienphase. In den Daten der ASID 1999 ist dies an einem sprunghaften Anstieg der Versicherungsjahre der 55 bis unter 65-Jährigen im Vergleich zu den älteren westdeutschen Frauen ersichtlich: Waren die über 65-Jährigen durchschnittlich 24 Jahre in der GRV versichert, waren es bei den jüngeren Frauen bereits 32 Jahre (BMA 1999, Tab. 1205). Die AVID 1996 weist für zwischen 1936 und 1940 geborene sogar durchschnittlich 34 Versicherungsjahre aus, davon knapp 23 Erwerbsjahre (vgl. VDR & BMA 1999, Tab. 1004).⁵

Dieser Wandel der Erwerbsverläufe fand in den europäischen Ländern je nach Wohlfahrtstypus unterschiedlich statt. So zeichneten sich vor allem die skandinavischen und in geringerem Maße auch die liberalen Wohlfahrtsstaaten wie bspw. Großbritannien schon in der Mitte des 20. Jahrhunderts durch höhere Frauenerwerbsquoten als die konservativen, auf traditionellen Sozialversicherungssystemen basierenden Wohlfahrtsstaaten aus. Vergleicht man dies mit der Entwicklung der Differenzen der Lebenserwartung zwischen Männern und Frauen, ist der stärkere Rückgang des gender gap in den Ländern mit hoher Frauenerwerbsbeteiligung augenfällig (**Tabelle 1**).

Die Erkenntnisse über soziale Differenzen von Mortalität und Morbidität, Ansätze zur Erklärung des gender gap in der Mortalität und internationale Vergleichsdaten legen die These nahe, dass eine Verlängerung der Erwerbsphasen im Lebensverlauf von Frauen zur Verringerung ihrer Lebenserwartung beitragen.

Der bisherige Forschungsstand auf diesem Gebiet bezieht sich überwiegend auf die Mortalität über die gesamte Lebensspanne hinweg. Ergebnisse zur Altersmortalität liegen hingegen kaum vor. So weisen Liang et al. (2002) darauf hin, dass in Studien über den Zusammenhang von so-

⁴ Zum Einfluss der Wahl unterschiedlicher Indikatoren für den sozio-ökonomischen Status auf die Mortalität vgl. Kunst et al. (2004).

⁵ Für die jüngste Untersuchungskohorte der 1951-1955 Geborenen steigt die prognostizierte Zahl der Erwerbsjahre auf 28 Jahre, die der Versicherungsjahre auf insgesamt 38,7 Jahre (VDR & BMA 1999).

Tabelle 1: Frauenerwerbsquoten und Differenz zwischen Lebenserwartung bei der Geburt von Frauen und Männern im internationalen Vergleich

	Labour force participation rate 1970 (in %)*	Labour force participation rate 1990 (in %)*	1990-1970 (in %-points)	Female-Male Diff. 1980 (in years)**	Female-Male Diff. 1996 (in years)**	1996-1980 (in years)
Germany	48	57	9	6,8	6,5	-0,3
France	48	58	10	8,2	7,8	-0,4
Sweden	59	80	21	6,0	5,0	-1,0
Denmark	58	78	20	6,1	5,2	-0,9
United Kingdom	51	65	14	6,0	4,9	-1,1

* Quelle: OECD, Statistical Compendium; ** Quelle: Gjonça et al. 1999, Table 1.

zialem Status und Mortalität die Alten meist als homogene Gruppe erfasst sind und nicht zwischen jüngeren und älteren Alten unterschieden wird. Zudem ist ihre Anzahl in derartigen Studien zumeist gering. Dies ist auch eine Folge des Fehlens von hinreichend großen Individual-Datensätzen, die Informationen über Erwerbsverläufe und Altersmortalität verknüpfen. Hier eröffnet die Rentenbestands- und Rentenwegfallsstatistik der Rentenversicherung neue Analysemöglichkeiten, da sie große Fallzahlen und detaillierte Informationen zu Erwerbsdauer und individuell erzielten Einkommen beinhaltet.⁶

Für Deutschland existieren aktuell widersprüchliche Ergebnisse: Eine Studie des IAB auf Basis des SOEP (1996) und qualitativer Interviews mit Frauen über 70 belegen bereits den vermuteten Zusammenhang zwischen Erwerbstätigkeit und Mortalität bzw. Morbidität (vgl. Meyer et al. 1996); Kruse (2000), der Daten des VDR nutzte, kommt hingegen nicht zu eindeutigen Ergebnissen.

Ziele und möglicher Erkenntnisgewinn

Der Erkenntnisgewinn einer solchen Studie ist immanent: Wenn ein positiver Zusammenhang zwischen Erwerbsdauer und Altersmortalität besteht, nimmt bei einem steigenden Anteil lange erwerbstätiger Frauen in der Bevölkerung die Lebenserwartung in geringerem Maße zu als bisher angenommen. Damit ist auch der Anstieg der Rentenbezugszeiten geringer. Dies hat sowohl Auswirkungen auf das Verhältnis der Erwerbsbevölkerung zu Rentenempfängern als auch auf die Tarife privater Rentenversicherer. Der Nachweis eines positiven Zusammenhangs zwischen langjähriger Erwerbstätigkeit und Mortalität würde zu einer Verbesserung der Prognosefähigkeit über die künftige Rentnerpopulation führen und Ausmaß und Zeitpunkt notwendiger politischer Maßnahmen genauer determinieren. Aus politikwissenschaftlicher Perspektive lässt sich im Rahmen des hier skizzierten Projektes die Validität von Begründungszusammenhängen und politischen Diskursen zu Reformvorhaben eruieren.

⁶ Eine Berufsbiographie ist zwar auch im SOEP enthalten, doch stehen bei Rentnern und Rentnerinnen keine Einkommensverläufe aus der Zeit vor der Erhebung zur Verfügung, und die Fallzahlen sind für Kohortenanalysen zu gering.

Bennett und Olshansky (1996) zeigen, dass bereits geringe Korrekturen der Mortalitätsraten große Wirkungen haben. Ihre Anpassung der von der US-amerikanischen Social Security Administration (SSA) verwendeten Mortalitätsraten ab 65 Jahren an die Raten aus Ländern mit einer verlässlicheren Erfassung von Todesfällen führte für das Jahr 1990 zu einer Verringerung der ferneren Lebenserwartung der 65-jährigen Frauen von 0,5 Jahren, im Jahr 2050 von 1,2 Jahren (Männer: -0,3 und -4,0 Jahre) (Bennett & Olshansky 1996, 712). Die Projektion der Rentnerpopulation im Jahr 2050 mit den angepassten Werten ergab eine Verringerung um 4,8% gegenüber der Vorausberechnung der SSA (Bennett & Olshansky 1996, 714). Eine durch die Behörde durchgeführte Sensitivitätsanalyse der Prognosefaktoren für die Finanzierung der Sozialfonds ergab, dass die Annahmen über die künftige Mortalität die Ergebnisse am stärksten beeinflussen (Bennett & Olshansky 1996, 720 f.).

Für Deutschland zeigt Wilke (2004) anhand der 10. koordinierten Bevölkerungsvorausberechnung des Statistischen Bundesamtes (Statistisches Bundesamt 2003) den Einfluss der Annahmen zur Lebenserwartung auf den Beitragssatz und das Rentenniveau der GRV. Ihrer Berechnung zufolge führt die Verwendung einer Variante mit einer im Vergleich zum Referenzszenario um 1,5 Jahre höheren Lebenserwartung zum Anstieg des Beitragssatzes um 0,2 Prozentpunkte und einer Absenkung des Rentenniveaus um 0,4 Prozentpunkte (Wilke 2004, 22 f.).

In der Berechnung des Statistischen Bundesamtes wird der Anstieg weiblicher Erwerbstätigkeit in jüngeren Kohorten bei den Prognosen bislang nicht explizit berücksichtigt. So geht die Prognosevariante mit dem geringsten Anstieg der ferneren Lebenserwartung von den heute international erreichten niedrigsten Sterbewahrscheinlichkeiten aus (Statistisches Bundesamt 2003, 19). Für Variante L2 wurde die Abnahme der Sterblichkeit in der Bundesrepublik seit 1970 verwendet. Die von der „Kommission für Nachhaltigkeit in der Finanzierung der sozialen Sicherungssysteme“ (BMGS 2003, 52 f.) getroffenen Annahmen bezüglich der Mortalität bis 2030 wurden nicht expliziert. Das skizzierte Projekt könnte einen wichtigen Beitrag zur Verbesserung der Prognosefähigkeit leisten.

Methodisches Vorgehen

In der Rentenzugangstatistik des VDR waren im Jahr 2002 die Geburtskohorten von 1904 bis 1937 erfasst (VDR 2003). Diese dürften folglich auch in der Rentenbestands- und in der Rentenwegfallstatistik zu finden sein.⁷ Mit Hilfe der Rentenbestands- und Rentenwegfallstatistik soll eine Kohortenanalyse erfolgen. Für jede reale Geburtskohorte werden Untergruppen mit unterschiedlich langen Erwerbs- bzw. Versicherungsdauern gebildet. Feinere Differenzierungen könnten anhand der Summe der Entgeltpunkte (EP) oder der durchschnittlich je Versicherungsjahr erzielten EP gebildet werden. Eine Hypothese wäre, dass Frauen mit Teilzeiterwerbstätigkeit weniger belastet sind als jene mit Vollzeiterwerb, was sich annäherungsweise mit Hilfe der Entgeltpunkte ermitteln ließe. Zudem soll der Einfluss sozialer Stratifizierung durch die Unterscheidung nach Versicherungszweigen berücksichtigt werden.⁸ Für diese Untergruppen werden die Wahrscheinlichkeiten errechnet, das Erhebungsjahr zu erleben. Im Wesentlichen handelt es sich um einen deskriptiv-statistischen Ansatz.

⁷ Insbesondere bei den jüngsten Rentenzugangskohorten kann jedoch davon ausgegangen werden, dass diese den Rentenzugang eine Weile überleben werden.

⁸ Zur Legitimität dieser Differenzierung vgl. Rehfeld & Scheitl (1991, 295 f.).

Datenbedarf (mit Bezug zu FDZ-RV-Daten und min-max-Abwägung)

Benötigt werden Daten aus der Rentenzugangs- und Wegfallstatistik.

Minimum:

- Nach Geburtsjahr geschichtete Stichprobe für Männer und Frauen aus Rentenwegfalls- und Rentenbestandsstatistik für ein Jahr. Besonders hohe, dünn besetzte Alter könnten abgeschnitten oder zusammengefasst werden.
- Merkmale: Geburtsjahr, Summe der Entgeltpunkte, Versicherungszeit / Zahl der Beitragsjahre, durchschnittliche Entgeltpunkte pro Beitragsjahr, Versicherungszweig, Wegfallsgrund, Wohnort Ost/West.

Maximum:

- Vollerhebung aus Rentenwegfalls- und Rentenbestandsstatistik für ein Jahr. Besonders hohe, dünn besetzte Alter könnten abgeschnitten oder zusammengefasst werden.
- Zusätzliche Merkmale: chronologische Erwerbsbiographie, Art der Entgeltpunkte, Rentenbeginn, Beruf.

Zugangswege

- Gewünscht: Scientific Use File
- Akzeptierbar: Gastwissenschaftler-Arbeitsplatz (soweit finanzierbar), kontrollierte Datenfernverarbeitung.

Zusammenfassung

Die Bereitstellung von Daten der gesetzlichen Rentenversicherung eröffnet vielfältige neue multi- und interdisziplinäre Analysemöglichkeiten. Ein Beispiel dafür ist das hier skizzierte Projekt, auf dessen Basis weitere politikwissenschaftlich relevante Fragestellungen bearbeitet werden können. Auch lassen sich Projekte über den Wandel der Erwerbsverläufe oder über den Einfluss des Alters auf die soziale Stratifizierung innerhalb einer Gesellschaft vorstellen. Zu bedauern ist, dass Datensätze aus der AVID-Studie der Wissenschaft nicht zugänglich gemacht werden. So fehlen in der Rentenbestands- und Wegfallsstatistik wichtige Indikatoren wie Haushaltzusammenhang oder kumulierte Alterseinkommen aus verschiedenen Quellen. Auch der Beruf und Bildungsstand werden unzureichend bzw. gar nicht erhoben.

Bei der Schaffung von Gastwissenschaftler-Arbeitsplätzen wäre es schön, wenn auch Berlin als Standort berücksichtigt werden könnte.

Literatur

- Bennett, N. & Olshansky, S. (1996): Forecasting US Age Structure and the Future of Social Security: The Impact of Adjustments to Official Mortality Schedules. *Population and Development Review*, 22, 4, 703-727.
- Bundesministerium für Arbeit und Sozialordnung (Hg.) (2001): ASID 1999. *Tabellenband*. München.
- Bundesministerium für Gesundheit und Soziale Sicherung (2003): *Nachhaltigkeit in der Finanzierung der sozialen Sicherungssysteme. Endbericht der Kommission*. Berlin.
- Gjonça, A., Tomassini, C. & Vaupel, J. (1999): Male-female Differences in Mortality in the Developed World. *Max-Planck-Institut für demografische Forschung Rostock, Working Paper* 1999-009.
- Kruse, E. (2000): Sterblichkeit und fernere Lebenserwartung von Versichertenrentnern der gesetzlichen Rentenversicherung - Aktuelle Ergebnisse einer Sterbetafelberechnung für 1996/1998. *Deutsche Rentenversicherung*, 3-4, 121-138.
- Kunst, A., Andersen, O., Bos, V., Cardano, M., Costa, G., Harding, S., Hemström, Ö., Layte, R., Mackenbach, J., Regidor, E., Reid, A., Santana, P. & Valkonen, T. (2004): Monitoring of trends in socioeconomic inequalities in mortality: Experiences from a European project. *Demographic Research*, Special Collection 2, Article 9, Max-Planck-Institut für demografische Forschung Rostock, 229-253.
- Liang, J., Bennett, J., Krause, N., Kobayashi, E., Kim, H.-K., Brown, J., Akiyama, H., Sugisawa, H. & Jain, A. (2002): Old Age Mortality in Japan: Does the Socioeconomic Gradient Interact With Gender and Age? *Journal of Gerontology: Social Sciences*, 57 B, 5, 294-307.
- Martikainen, Pekka (1995): Mortality and Socio-economic Status Among Finnish Women. *Population Studies*, 49, 71-90.
- Meyer, S., Müller, P., Schulze, E., Häußler, B.; Reschke, P. & Staffeldt, T. (1996): Lebenserwartung erwerbstätiger Frauen, Bundesministerium für Arbeit und Sozialforschung. *Forschungsbericht Sozialforschung* Nr. 252, Bonn.
- Müller, H.-W. & Rehfeld U. (1985): Die Sterblichkeit von Altersrentnern in der gesetzlichen Rentenversicherung in den Jahren 1982-1984 - zur Lebenserwartung berufstätiger Frauen und Männer, *Deutsche Rentenversicherung*, 6-7-8, 382-403.
- Rehfeld, U. & Scheitl, O. (1991): Sterblichkeit und fernere Lebenserwartung von Rentnern der gesetzlichen Rentenversicherung - aktuelle Ergebnisse für 1986/1988 und Bilanz zum bisherigen Untersuchungsstand. *Deutsche Rentenversicherung*, 4-5, 289-320.
- Reil-Held, A. (2000): Einkommen und Sterblichkeit in Deutschland: Leben Reiche länger? Universität Mannheim, SFB 504 *Diskussionspapier* 00-14.
- Statistisches Bundesamt (2003): *Bevölkerung Deutschlands bis 2050. 10. koordinierte Bevölkerungsvorausberechnung*. Wiesbaden.
- Verband Deutscher Rentenversicherungsträger (Hg.) (2003): *Rentenversicherung in Zahlen, DRV-Schriften*, Band. 22, Frankfurt a.M.
- Verband Deutscher Rentenversicherungsträger und Bundesministerium für Arbeit und Sozialordnung (Hg.) (1999): AVID 1996. *Tabellenband II*. München.
- Wilke, Christina Benita (2004): Ein Simulationsmodell des Rentenversicherungssystems. Konzeption und ausgewählte Anwendungen von MEA-PENSIM, *MEA Discussion Paper* 48-04, Mannheim.

Prof. Dr. Barbara Riedmüller lehrt und forscht auf dem Gebiet der Komparatistik mit Schwerpunkten in der vergleichenden Wohlfahrtsforschung an der Freien Universität Berlin, Fachbereich Politikwissenschaft. Schwerpunkte in der Forschung sind zurzeit zum einen die private Alterssicherung im europäischen Vergleich, ein Forschungsprojekt gefördert von der Europäischen Union, und zum anderen ein von der Böckler-Stiftung gefördertes Projekt über die Arbeitsmarktintegration allein erziehender Frauen. Sie ist unter anderem Mitglied des Sozialbeirats für die Rentenversicherung der Bundesregierung und Vorsitzende des Beirats des Zentrums für Sozialpolitik der Universität Bremen. Für weitere Informationen konsultieren Sie bitte unsere Website: www.polwiss.fu-berlin.de/people/ried/index.html.

Dipl.-Soz. Michaela Willert ist wissenschaftliche Mitarbeiterin im EU-geförderten und von Prof. Riedmüller geleiteten Projekt „Private Pensions and Social Inclusion in Europe: A study of Germany, Italy, the Netherlands, Poland, Switzerland and the United Kingdom“.

Datenbedarf in der Mortalitätsforschung in Deutschland

Dr. Rembrandt Scholz (Max-Planck-Institut für demografische Forschung Rostock)

Einleitung

Die Abteilung „Alterung und Langlebigkeit“ und das „Datenlabor für demographische Daten“ des MPI für demografische Forschung Rostock bearbeiten Forschungsprojekte zur Darstellung der Mortalität. Hierzu werden spezifische Datensätze analysiert und es werden Referenzen für die Datenqualität benötigt, die repräsentative Aussagen für Deutschland zulassen.

Die Verfügbarkeit der amtlichen Mortalitätsdaten im hohen Alter ist eingeschränkt. Die Zunahme des zeitlichen Abstandes zur Volkszählung 1987 hat Einfluss auf die Datenqualität durch Fortschreibung von Fehlern. Diese Einschränkung ist bei den VDR-Daten nicht gegeben. Die Validität der Daten des VDR ist durch die Zahlung von finanziellen Leistungen und die dadurch bedingten Rückkopplungen besonders hoch.

Mit den ca. 12 Mio. Personen mit Versichertenrenten im Alter von 65 Jahren und älter wird etwa 93 % der Bevölkerung abgebildet. Die Altersverteilung ist bis ins höchste Alter stark besetzt. Damit sind Aussagen im hohen Altersbereich auch zuverlässig.

Die Erklärung von Phänomenen der Mortalitätsentwicklung kommt sehr schnell an Grenzen, die durch die Datenverfügbarkeit verursacht sind. Die Möglichkeit der Analyse der VDR-Daten eröffnet ein neues Feld zu bearbeitender Themen. Uns interessieren Aussagen zur differentiellen Sterblichkeit nach sozialen Schichten (Einkommen). Wir suchen nach Aussagen zum Grenzalter und dessen wahrscheinlicher Entwicklung. Weitere Merkmale, die als Differenzierungsmerkmale geeignet sind, sollen einbezogen werden.

Die Datensätze des FDZ-RV sollten ständig für die Analyse von Entwicklungen der Mortalität in Deutschland genutzt werden (Monitoring).

Fragestellung

Auf der Hand liegen die Fragestellungen, die unmittelbar eine Differenzierung nach Merkmalen zulassen, welche sonst nicht zu erhalten sind. Dazu gehören alle Merkmale, die in Deutschland nicht mit dem Sterbefall statistisch erfasst werden können.

Die Daten des VDR sind hinsichtlich der Mortalität mit Hilfe der Konstruktion von Sterbetafeln und der Berechnung der Lebenserwartung der Rentner bereits analysiert worden (Kruse 2000, Rehfeld & Scheitl 1991). Diese Analysen sollten vertieft und mit weiteren Merkmalen verknüpft und als Längsschnitte angelegt werden. Es werden in diesem Stadium der Analyse im Gegensatz zu den anderen Forschungsprojekten nicht die Individualdatensätze sondern Merkmalstabellen für Summensätze benötigt. Das betrifft die mehrdimensionale Summation der Bestandsdaten und entsprechende Summensätze für Personen, die verstorben sind.

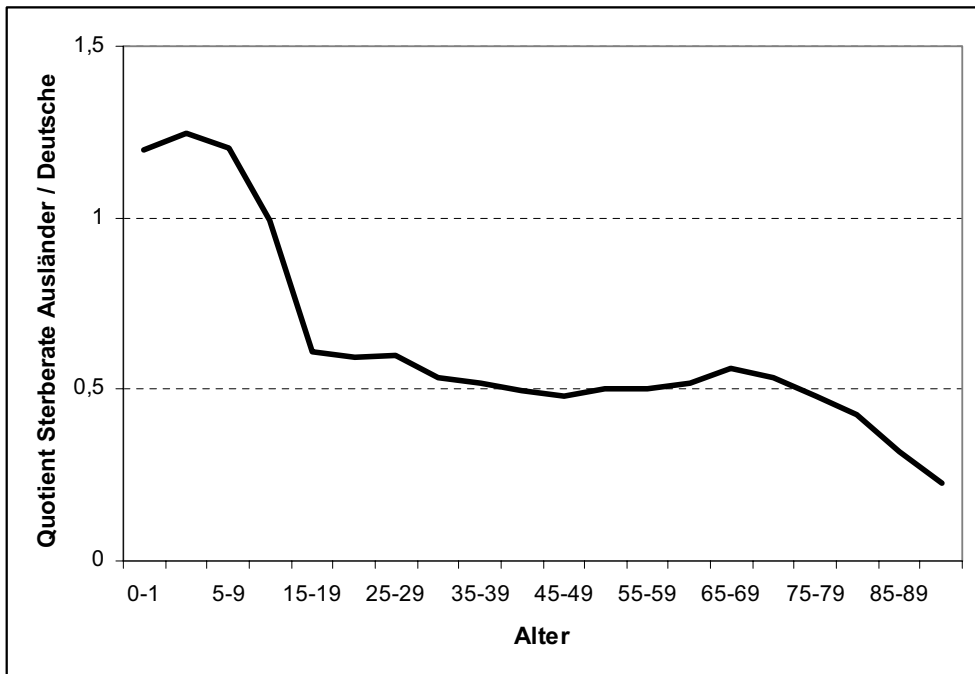
a) *Vergleichende Studie über die Entwicklung der Sterblichkeit Ost West*

Die nach der Wende einsetzende starke Angleichung der Lebensverhältnisse in Ost und West ist ein einmaliger ‚gesellschaftspolitischer Modellversuch‘: die Veränderung der gesamten gesellschaftlichen Verhältnisse, des Rentensystems, der Infrastruktur bis zum Versorgungssystem von medizinischen Leistungen. Bei der Mortalitätsanalyse zeigt sich, dass die Geschlechter unterschiedlich von den Veränderungen profitieren können. Die Vertiefung der Analysen mit weiteren Merkmalen ist angezeigt, um eine bessere Bewertung vorzunehmen. Für die sehr alten Geburtsjahrgänge konnte die Übersterblichkeit in Ostdeutschland abgebaut werden (Scholz & Maier 2003). Offene Fragen ergeben sich zu dem genauen Wirkmechanismus.

b) *Sterblichkeit der Ausländer in Deutschland*

Die amtliche Statistik hat für das Merkmal Staatsbürgerschaft die Schwierigkeit, den Bestand in der richtigen Größenordnung abzubilden - damit kann die Sterblichkeit der Ausländer nicht richtig eingeschätzt werden.

Abbildung 1: Verteilung der altersspezifischen Quotienten der Sterberaten Ausländer / Deutsche, Deutschland alte Länder 1980-95



Quelle: berechnet nach Angaben vom Statistischen Bundesamt.

Durch den Ausländeranteil von 9 % ist ein Einfluss auf die Gesamtsterblichkeit in Deutschland gegeben, insbesondere durch die regionale Ungleichverteilung.

In der amtlichen Statistik ist die Sterblichkeit der Ausländer wesentlich geringer als die der Deutschen. Damit wird die Lebenserwartung in Deutschland erhöht. Die Sterblichkeit der Ausländer wird durch eine Vielzahl von Faktoren beeinflusst, unter anderem durch selektive Zuwanderung und die selektive Abwanderung. Sind diese Faktoren allerdings ausreichend zur Erklärung der Unterschiede und haben wir nicht auch ein Problem in der statistischen Erfassung der korrekten Bestände im Ausländerzentralregister? Während die Erfassung der Mortalität der Ausländer in Deutschland eine hohe Validität hat, ist diese bei den Beständen der Ausländer wahrscheinlich nicht gegeben. Wie hoch ist eine eventuelle Übererfassung und welche Korrekturen müssen vorgenommen werden?

c) Differentielle Sterblichkeit in Deutschland

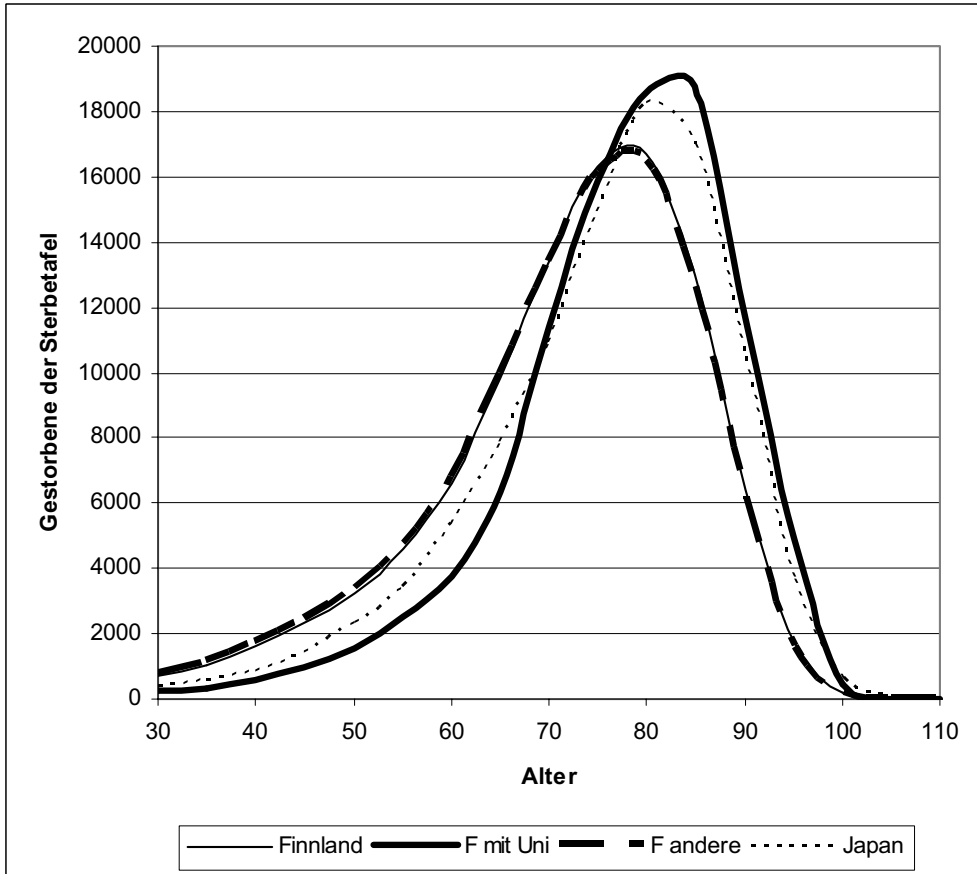
Die Entwicklung der Sterblichkeit hat sich in der Vergangenheit immer an dem Verlauf der Entwicklung der Sterblichkeit der günstigen sozialen Schichten orientiert. Die günstigere soziale Schicht realisiert zur gleichen Zeit immer bessere Werte der Mortalität als der Durchschnitt der Bevölkerung. Somit wird eine mögliche Entwicklung der Sterblichkeit für einen bestimmten Stand des medizinischen Fortschrittes von günstigeren sozialen Gruppen angezeigt – illustriert an der Abbildung der Sterbefälle nach Alter für Finnland und Japan (**Abbildung 2**).

Die Sterblichkeit in Finnland ist nach dem Bildungsabschluss (Hochschule ja/ nein) differenziert und zeigt auf dem Niveau von Sterbetafeln für Finnland 1995 einen Vergleich mit Japan 2000. Die realisierte Mortalität der Personen mit höheren Bildungsabschlüssen in Finnland ist teilweise günstiger als die Mortalität in Japan.

Hypothese

Die soziale Differenzierung nach Einkommen (Gesamtrentenzahlbetrag) ist ein Merkmal, welches die Population in soziale Gruppen teilt, die mit unterschiedlicher Mortalität verbunden sind. Der Trend der Sterblichkeitsentwicklung folgt dem Gradienten der Mortalität nach sozialen Schichten (Otterblad 1991).

Abbildung 2: Verteilung der gestorbenen Männer auf die Sterbetafel standardisiert nach Alter in Finnland 1995 und Japan 2000, Finnland nach Bildung, Universitätsabschluss und andere



Quelle: berechnet nach HMD und Valkonen, nicht publiziert.

Methode

Bestimmung von altersspezifischen Mortalitätsraten, aus denen Sterbetafeln abgeleitet werden. Die Aussagen sollten möglichst viele Merkmale berücksichtigen, um eventuelle Wechselwirkungen rechnerisch auszuschalten. Als Minimum einer solchen Analyse werden Klassen des Gesamtrentenzahlbetrages gebildet, welche die Population und die Sterbefälle in Gruppen einteilt (z.B. Perzentile um den Mittelwert). Für die Aussagefähigkeit sind auch die Dimensionen der Datenausfälle wichtig. Als Minimum werden für jede Einkommensklasse mehrdimensionale Tabellen für die Merkmale benötigt:

z.B. 3 Tabellen für jeweils die Population und die Sterbefälle:

- Gesamt-Population, Gesamt-Sterbefälle nach Alter, Geschlecht, Ost / West,
- 30% Perzentil des Einkommens der Population und der dazugehörigen Sterbefälle nach Alter, Geschlecht, Ost / West,
- 70% Perzentil des Einkommens der Population und der dazugehörigen Sterbefälle nach Alter, Geschlecht, Ost / West,

Weitere Merkmale können eine Analyse vertiefen und auch für weitergehende Aussagen wichtig sein: Familienstand, Bildung, Kinderzahl, Informationen zur Erwerbsbiographie, Beruf, Migration, Region und Haushaltssituation. Dieses Projekt wäre Teil der internationalen Studie Vanguard Project.

Es ist ein Trend der Zunahme der Lebenserwartung in vielen Ländern zu beobachten. Bei der Untersuchung im internationalen Vergleich stellt sich für die Gesamtentwicklung bisher ein linearer Trend der Zunahme der Lebenserwartung dar (Oeppen & Vaupel 2002). Eine todesursachenspezifische Einschränkung des Forschungsansatzes führt zu einer noch stärkeren Zunahme (Vallin & Mesle 2004). Es ist zu vermuten, dass die Entwicklung der Mortalität am stärksten mit sozialen Merkmalen wie Bildung und Einkommen zusammenhängt.

Bei der Vorhersage der weiteren Entwicklung ist die Erreichung von Grenzwerten des menschlichen Lebens bisher noch nicht sichtbar. Die Zunahme der Lebenserwartung wird für die sozial besser ausgestatteten Personen mit höherer Bildung und besseren finanziellen Möglichkeiten dazu führen, dass die eventuell vorhandenen Grenzen der Mortalitätsentwicklung früher sichtbar werden müssen. Die Vorhut (engl. Vanguard) der Entwicklung kann die zukünftigen Möglichkeiten bei dem aktuellen medizinischen Niveau darstellen. Das MPI demografische Forschung organisiert ein internationales Forschungsprojekt und möchte auch einen deutschen Datensatz zu dieser Zielstellung beifügen.

Datenbedarf

Aus der Rentenbestandsstatistik und der Rentenwegfallsstatistik Auszahlungen als mehrdimensionale Summenstatistiken möglichst detailliert und nach den Merkmalen:

- Alter (nach Einzelalter)
- Geschlecht
- Familienstand
- Staatsbürgerschaft
- Bildungsabschluss / Schulabschluss
- Rentenzahlbetrag
- Region: Ost / West.

Referenzen

- Kruse, E. (2000): *Sterblichkeit und fernere Lebenserwartung von Versichertenrentnern der gesetzlichen Rentenversicherung: Aktuelle Ergebnisse einer Sterbetafelberechnung für 1996/1998. Deutsche Rentenversicherung*, 3-4, 121-138.
- Oeppen, J. & Vaupel, J.W. (2002): Broken limits to life expectancy. *Science*, 296, 1029-1030.
- Otterblad Olausson, P. (1991): Mortality among the elderly in Sweden by social class. *Social Science & Medicine*, 32, 437-440.
- Rehfeld, U. & Scheitl, O. (1991): Sterblichkeit und fernere Lebenserwartung von Rentnern der gesetzlichen Rentenversicherung – aktuelle Ergebnisse für 1986/1988 und Bilanz zum bisherigen Untersuchungsstand. *Deutsche Rentenversicherung*, 4-5, 289-320.
- Scholz, R.D. & Maier, H. (2003): *German unification and the plasticity of mortality at older ages*. WP-2003-31, <http://www.demogr.mpg.de/papers/working/wp-2003-031.pdf>.
- Vallin, J. & Meslè, F. (2004): *The lowest mortalities*. Vortrag, 1. Symposium Human Mortality Database, Rostock 17./ 18.06.04.

Rembrandt Scholz ist wissenschaftlicher Mitarbeiter im Datenlabor für demografische Daten des MPI. Der Schwerpunkt seiner Arbeit liegt auf der Erstellung und Pflege von Mortalitätsdatenbanken und der Durchführung von Forschungsprojekten zur Sterblichkeit in Deutschland.

Am **MPI für demografische Forschung** bestehen zwei wissenschaftliche Abteilungen mit den Forschungsprogrammen „Altern“ (Prof. Vaupel) und „Fertilität und Familiendynamik“ (Prof. Hoem). Das Forschungsprogramm Altern widmet sich der Erforschung jener Faktoren, die Altern, Langlebigkeit und Sterblichkeit bestimmen. Ein Schwerpunkt liegt dabei auf der Entwicklung im hohen Alter. Vor dem Hintergrund steigender Lebenserwartung und einer zunehmenden Zahl älterer Menschen richtet sich das Interesse darauf, allgemeine Trends wie Unterschiede von Alterungsprozessen und Entwicklungen von Langlebigkeit und Mortalität in verschiedenen Ländern zu erfassen. Die von Prof. Hoem geleitete Abteilung untersucht die Geburten und Familienentwicklung im heutigen Europa. Eine ausführliche Vorstellung des Institutsprofils findet sich unter www.demogr.mpg.de.

Zielvorstellungen in der Alterssicherung: Konkretisierung, Operationalisierung und Messung als Grundlage für normative Aussagen

Dr. Rainer Unger (Zentrum für Sozialpolitik der Universität Bremen)

Einleitung

In den letzten Jahren sind in Deutschland im Rahmen der Alterssicherungspolitik gravierende Änderungen der gesetzlichen Grundlagen, insbesondere in der gesetzlichen Rentenversicherung, vorgenommen worden. Mit der Verabschiedung des Altersvermögensgesetzes (AVmG) und des Altersvermögensergänzungsgesetzes (AVmEG) erfolgte ein Richtungswandel in der Alterssicherungspolitik, in der die private Altersvorsorge durch steuerliche Maßnahmen massiv unterstützt wird, während das Leistungsniveau der gesetzlichen Rentenversicherung (GRV) reduziert wird. Die jüngste Minderung des Rentenniveaus ist mit dem Rentenversicherungsnachhaltigkeitsgesetz verbunden. Hier deutet sich eine Entwicklung an, die langfristig dazu führen könnte, dass die Leistungen der GRV auch für langjährig Versicherte nur noch eine Mindestsicherung darstellen. Die - insbesondere auch längerfristigen - individuellen Auswirkungen durch Änderungen auf der Leistungsseite sind allerdings kaum Gegenstand der politischen Erörterungen. Die Diskussion konzentriert sich vielmehr auf den Beitragssatz und dessen Entwicklung und auf die Möglichkeiten der Leistungskürzung, um diesen Zielwert gewährleisten zu können.

Die Diskussion verdeutlicht einmal mehr, dass kaum Kenntnisse über das tatsächliche Sicherungsniveau der Rentergeneration vorhanden sind und in der Regel das Eckrentenniveau - eine sehr leicht zu berechnende Größe - zur Beurteilung herangezogen wird. Durch die Änderung des Absicherungsziels, in dem die angestrebte private Vorsorge in das angestrebte Absicherungsniveau mit eingeht, ist dieser Indikator allerdings noch weniger als bisher geeignet, adäquate Informationen über das Absicherungsniveau zu liefern. Es ist daher erforderlich, Indikatoren zur Messung des Leistungsniveaus der Alterssicherungssysteme zu haben, um u.a. Hinweise auf die bewusst geschaffene „Sicherungslücke“ zwischen den durch die gesetzliche Rentenversicherung zur Verfügung gestellten Leistungen und dem vom Individuum bzw. Haushalt angestrebten Gesamtsicherungsniveau zu liefern, und so erforderlichen Handlungsbedarf aufzuzeigen.

Fragestellung

Die Fragestellung bezieht sich auf die adäquate Ermittlung des Leistungsniveaus von Alterssicherungssystemen. Dabei geht es darum zu analysieren, wie das Sicherungsniveau charakterisiert werden kann.

Zielsetzung

Zielsetzung ist die Entwicklung von Indikatoren. Dazu sollen Einkommensersatzraten beim altersbedingten Ausscheiden aus dem Erwerbsleben, aber auch im Falle von Invalidität und beim Tod eines Ehepartners ermittelt werden. Diese beziehen sich nicht nur auf den unmittelbaren Zeitpunkt vor und nach dem Eintritt des entsprechenden Ereignisses, sondern es sollen auch längere Phasen aus der Versicherten- und der Haushaltsbiographie berücksichtigt werden. Dieses Vorgehen ermöglicht es, Aussagen über die Verstetigungsfunktion der gesetzlichen Rentenversicherung und anderer Alterssicherungsinstitutionen abzuleiten.

Einerseits wird dazu das Absicherungsniveau der älteren Bevölkerung nach dem Übergang von der Erwerbstätigkeit in die Nacherwerbsphase einer empirischen Analyse unterzogen. Dabei soll die Betrachtung des Einkommens, das sich oftmals nur auf eine Einkunftsart (zum Beispiel die Rente aus der GRV) stützt, erweitert werden, um ein Bild der tatsächlichen Einkommenslage im Alter zu vermitteln. Im Vordergrund steht das Zusammentreffen von verschiedenen eigenen Leistungen aus den Alterssicherungssystemen. Dazu werden neben den Einkommen aus der gesetzlichen Rentenversicherung (der ersten Säule der Alterssicherung) die betriebliche Altersversorgung und die Zusatzversorgung des öffentlichen Dienstes (die zweite Säule der Alterssicherung) sowie Vermögenseinkünfte (die dritte Säule der Alterssicherung) in die Analysen mit einbezogen. Andererseits soll die Verstetigungsfunktion der Alterssicherung über den gesamten Lebenslauf einer empirischen Analyse unterzogen werden. Dabei soll zusätzlich zur individuenbezogenen Analyse auch der Haushaltszusammenhang berücksichtigt werden. Dadurch können Veränderungen in der Entwicklung der Nacherwerbsphase, etwa durch veränderte Haushaltszusammensetzungen, und damit Kenntnisse über das erreichte Absicherungsniveau im Haushaltskontext vertieft werden.

Methodisches Vorgehen

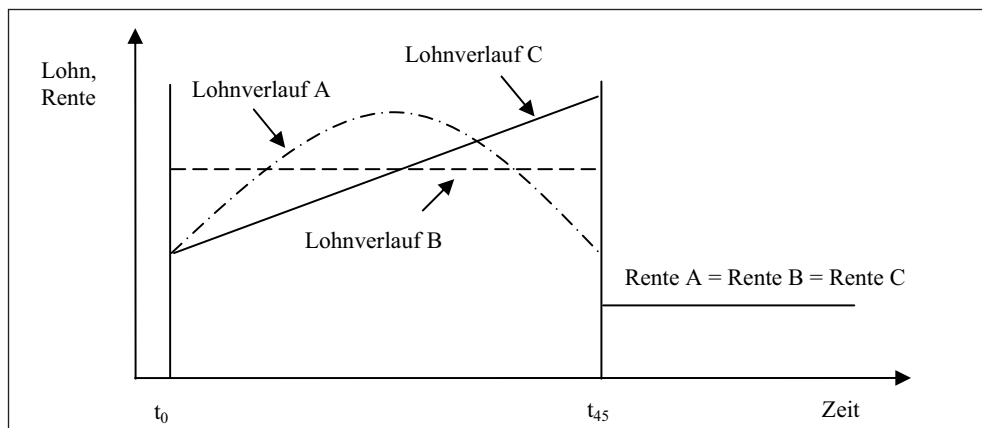
Bei der Analyse des Absicherungsniveaus wird auf den Lebensverlauf Bezug genommen und die Einkommenssituation über sogenannte Ersatzraten beschrieben, die vergleichend die materielle Situation vor und nach dem Ausscheiden aus der Erwerbstätigkeit charakterisieren. Es soll – im Prinzip ausgehend von der Definition der Ersatzrate als dem Verhältnis von Rentenzahlbetrag zu sozialversicherungspflichtigem Bruttomonatsarbeitsentgelt – untersucht werden, welche Konsequenzen der Lohnverlauf über verschiedene näher zu spezifizierende Zeiträume vor der Verrentung für diese Ersatzrate haben kann.

Hierbei geht es um den Zeitraum, für den die jeweiligen Einkommen ermittelt werden. So ist beispielsweise zu bedenken, dass eine Teilzeittätigkeit oder Arbeitslosigkeit vor dem Ausscheiden aus der Erwerbstätigkeit zwangsläufig zu einem niedrigeren Entgelt führen würde und damit zu einem höheren Rentenniveau. Um derartige Aspekte berücksichtigen zu können, sollte die bisherige eher als „zeitpunktbezogen“ zu charakterisierende Betrachtungsweise durch eine „zeitraumbezogene“ ergänzt werden. Abbildung 1 illustriert diesen Zusammenhang. In der **Abbildung** sind drei verschiedene Lohnverläufe schematisch dargestellt, die alle zu einer gleich hohen Rente führen. So wird ersichtlich, dass das letzte Entgelt nur bedingt als Vergleichsgröße zur Ermittlung eines Indikators für die Beurteilung des Absicherungsniveaus herangezogen

werden kann. Betrachtet man nur das letzte Arbeitsentgelt und die Zugangsrente, so würde sich folgende Ordnung (bei gleicher Entgeltpunktsomme) ergeben:

$$\text{Rentenniveau: } \frac{\text{Lohnverlauf A}_{t_{45}}}{\text{Rente}} \rangle \frac{\text{Lohnverlauf B}_{t_{45}}}{\text{Rente}} \rangle \frac{\text{Lohnverlauf C}_{t_{45}}}{\text{Rente}}$$

Abbildung 1: Lohnersatzrate bei unterschiedlichen Einkommensverläufen



Quelle: eigene Darstellung.

In einem zweiten Schritt wird das Zusammentreffen von verschiedenen eigenen Leistungen aus den Alterssicherungssystemen analysiert. Als nächstes ist daher das Einkommenskonzept zu erweitern, in dem andere Einkommenskomponenten bei der Ermittlung von Ersatzraten berücksichtigt werden. In Abhängigkeit von den Einkunftsarten ist dann auch ein Übergang von der individuellen zur haushaltsspezifischen Betrachtungsweise notwendig. Dies wird sich auf Ein- und Zweipersonenhaushalte beschränken können. Ausgehend von den im ersten Schritt erarbeiteten Ergebnissen soll anhand verschiedener Einkommens- und Verbrauchsstichproben (EVS) des Statistischen Bundesamtes die Betrachtung des Einkommens erweitert werden. Hier steht die Ermittlung von Einkommenspositionen im Vordergrund. Anhand der EVS können für unterschiedliche Einkommenskonzepte – beginnend mit dem sozialversicherungspflichtigen Bruttomonatsarbeitsentgelt (dem Einkommensbegriff, der den Analysen anhand der prozessproduzierten Daten zugrunde liegt) bis hin zu dem verfügbaren Einkommen – Einkommensrelationen ermittelt werden, d. h. die Positionierung der Haushalte, die Rentenzahlungen beziehen, zu anderen Haushalten, deren Haupteinkommensbezieher/in sich beispielsweise in der letzten Phase der Erwerbstätigkeit befindet. Hier geht es u. a. darum, Alternativen zur bisherigen Ermittlung des Brutto- bzw. Nettorentenniveaus aufzuzeigen und die Konsequenzen, die mit der Wahl eines bestimmten Indikators verbunden sind, zu verdeutlichen.

Die Erweiterung der Nenner- und Zählergröße um weitere Einkunftsarten sowie die alternative Wahl der Untersuchungseinheit zwischen „Individuum“ und „Haushalt“ ist jedoch auch auf dieser Analyseebene noch zu ergänzen durch den Aspekt der Verstetigungsfunktion der Alterssicherung. Der dritte Schritt beinhaltet daher Analysen anhand des sozio-ökonomischen Panels

(SOEP). Dieser Datensatz stellt umfassende Informationen über die Einkommenslage von Haushalten bereit und ermöglicht Longitudinalanalysen über den Zeitraum von 1984 bis 2001. Hier sollen dieselben Einkommenskonzepte und Haushaltsdefinitionen wie bei den Analysen auf der Basis der Einkommens- und Verbrauchsstichproben verwendet werden, ergänzt um den Verlaufsaspekt. Anhand dieser Datenbasis ist es somit möglich, die anhand der prozessproduzierten Daten der Rentenversicherung und der EVS untersuchte Verstetigungsfunktion der Alterssicherung einer ergänzenden Analyse auf Haushaltsebene und für Einkommensaggregate zu unterziehen – wenn auch der Zeitraum nur einen Teil der Versicherten- und Haushaltsbiographie erfasst.

Datenbedarf

Den zentralen Bezugspunkt der Analysen bildet das individuelle Rentenzugangsniveau als Relation von individuellen sozialversicherungspflichtigen Bruttomonatseinkommen und dem Rentenzahlbetrag. Die dafür am besten geeignete Datenbasis sind die beim VDR verfügbaren Rentenzugangsdaten sowie die Sondererhebung „Vollendete Versichertenleben“ (VVL). **Tabelle 1** gibt einen Überblick über die zur Berechnung des individuellen Rentenniveaus benötigten Variablen. Derartige prozessproduzierte Daten ermöglichen umfangreiche Analysen auf der individuellen Ebene, wobei die Daten in hohem Maße valide sind. Ein weiterer Vorteil dieser Daten ist, dass sie Längsschnittanalysen ermöglichen. Diese sind erforderlich, da auch längere Phasen aus den Versichertenbiographien berücksichtigt werden sollen. Die Rentenzugangsstichproben haben dabei einerseits den Vorteil großer Fallzahlen (Vollerhebung), wodurch sich die Analysen auch für die Subgruppen der Versicherungsarten und Versicherungszweige durchführen lassen. Andererseits decken die Rentenzugangsdaten jedoch nicht die Jahresentgelte während des gesamten Beitragszeitraums, sondern lediglich die der letzten drei Jahre ab.

Tabelle 1: Variablenübersicht

Demographische Merkmale	Merkmale zur Versicherung (für jede versicherungsrelevante Zeit)	Merkmale zur Leistungsbewertung
Geburtsjahr, Familienstand, Geschlecht	Versicherungszweig, Jahresentgelt, Entgeltpunkte, Beitragszeiten	Rentenart, rentenrechtliche Zeiten, Summe der Entgeltpunkte, Rentenbetrag

Neben der Rentenzugangsstatistik bietet die Sondererhebung „Vollendete Versichertenleben“ Analysemöglichkeiten über den gesamten versicherungsrelevanten Erwerbsverlauf, wobei die Erhebung allerdings letztmalig 2001 durchgeführt wurde und sich nur auf einen Teil der Rentenzugänge bezieht. Eine jährliche Neuerhebung hätte dem gegenüber den Vorteil, in einer fortzuführenden Verteilungsberichterstattung das individuelle Absicherungsniveau aus der GRV einer empirischen Analyse zu unterziehen.

Auf Grund der Vor- und Nachteile, die mit den Datensätzen verbunden sind, wird daher angestrebt, die Analysen sowohl mit der Rentenzugangsstichprobe als auch mit der Sondererhebung „Vollendete Versichertenleben“ durchzuführen.

Zugangswege

Als Zugangswege zu den Daten der Rentenversicherung sind Aufbereitungen in Form von Public Use Files und Scientific Use Files, wie z.B. beim SOEP, der EVS oder dem Mikrozensus, wünschenswert. Ein akzeptierbarer Weg wäre ein Gastwissenschaftler-Arbeitsplatz im Forschungsdatenzentrum, wobei dieser mit erheblichen Kosten- und Zeitbeschränkungen verbunden wäre.

Zusammenfassung

Das Projekt hat zum Ziel, das Absicherungsniveau der älteren Bevölkerung anhand des individuellen Rentenniveaus, das auf einem erweiterten Einkommenskonzept basiert, einer empirischen Analyse zu unterziehen, um die Kenntnisse über das erreichte Sicherungsniveau zu vertiefen. Auf dieser Basis lässt sich die gegenwärtig schrittweise eingeleitete Reduzierung des Absicherungsniveaus der Alterssicherung der GRV vor dem Hintergrund der Entwicklung der Förderung der betrieblichen Alterssicherung und der privaten Altersvorsorge empirisch beschreiben. Von hohem Interesse sind daher besonders Analysen zum individuellen Rentenniveau, die alle „Drei Säulen der Alterssicherung“ abdecken.

Neben der guten Eignung der Rentenzugangsdaten sowie der Sondererhebung VWL sind auch Defizite in Bezug auf das Forschungsvorhaben zu sehen. Anhand der prozessproduzierten Daten des VDR ist es z.B. nicht möglich, weitere Einkommenskomponenten oder die Haushaltssituation zu berücksichtigen. Des Weiteren sind in den Daten des VDR zwar Angaben zu der Höhe des Bruttomonatsarbeitsentgeltes für die versicherungsrelevante Zeit vorhanden, aus welcher Quelle (z.B. Arbeitslosigkeit) die Einkommen resultieren ist dabei aber offen. Da alternative Konzepte zur Ermittlung von Einkommensersatzraten verwendet und bewertet werden sollen, wird im Projekt daher auf anderes Datenmaterial (EVS, SOEP) zurückgegriffen. Ein erster Schritt, um diese Defizite zu beheben, wäre eine Zusammenführung von Versicherteninformationen von Ehepaaren innerhalb der Rentenstatistik aber auch eine Verknüpfung mit anderen Datensätzen wie z.B. dem Mikrozensus oder dem SOEP.

Dr. phil. Rainer Unger, Wissenschaftlicher Mitarbeiter am Zentrum für Sozialpolitik (Wirtschaftswissenschaftliche Abteilung) der Universität Bremen. Studium der Soziologie, Erziehungswissenschaft und des Öffentlichen Rechts an der Universität Heidelberg von 1991 bis 1998. Von 1998 bis 2002 Wissenschaftlicher Mitarbeiter am Institut für Soziologie der Universität Heidelberg.

Das **Zentrum für Sozialpolitik** (ZeS), ein Forschungsinstitut der Universität Bremen, wurde Ende 1988 als erstes interdisziplinäres Sozialpolitikinstitut der Bundesrepublik mit fünf Abteilungen (Theorie und Verfassung des Wohlfahrtsstaates, Institutionen und Geschichte des Wohlfahrtsstaates, Wirtschaftswissenschaftliche Abteilung, Gesundheitspolitik, Arbeits- und Sozialmedizin, Geschlechterpolitik im Wohlfahrtsstaat) gegründet (siehe www.zes.uni-bremen.de).

Analysesystem Alterssicherung

Martin Grub (Fraunhofer-Institut für Angewandte Informationstechnik, Sankt Augustin)

Einleitung

Die Struktur der Alterssicherung wird sich in den nächsten Jahrzehnten tiefgreifend wandeln: Veränderte Erwerbsbiographien, steigende Lebenserwartung, niedrige Geburtenraten, Risikomischung zwischen kapital- und umlagefinanzierten Sicherungsinstrumenten und nicht zuletzt sich wandelnde Werte einer alternden Gesellschaft stellen enorme Herausforderungen an ein jedes Alterssicherungssystem dar. In jedweder Form werden Leistungen dabei angewiesen sein auf einen konsensualen Interessensausgleich zwischen den Generationen und Bevölkerungsgruppen.

Unverändert bleibt lediglich das Sicherungsziel: Menschen in verschiedensten Lebenslagen zu ermöglichen, Risiken zu versichern und Vorsorge zu treffen für ein würdiges Leben im Alter. Das macht die Wirkungsanalyse aktueller Reformpolitik existentiell wichtig und gleichzeitig schwierig: Das komplexe Zusammenwirken des Steuer-, Sozial-, Arbeits- und Versicherungsrechts mit einer sich wandelnden Arbeitswelt unter Unsicherheit will im Modell genauso verstanden werden wie Anpassungsstrategien der Versicherten an die sich ändernde Umwelt. Im Simulationsexperiment lassen sich politische Parameter dann so justieren, dass die daraus resultierenden Anreize Vorsorgeverhalten unterstützen und einen sozialen Ausgleich wahren, der nachhaltig akzeptiert zu werden verspricht.

Fragestellung

Genau in diesem Erkenntnisinteresse besteht das Anliegen des Analysesystems Alterssicherung, AsA. Seit Ende der 80er Jahre entwickelt und betreibt das Fraunhofer Institut für Angewandte Informationstechnik (FIT, vormals GMD) Simulationsmodelle zur Gesetzesabschätzung im Bereich der Alterssicherung. Traditionell stehen dabei distributive und fiskalische Fragen im Vordergrund: Wer gewinnt, wer verliert durch eine Reform? Wie viel kostet eine Reform?

Seitdem die Alterssicherungspolitik etwa mit dem Altersvermögensgesetz 2004 verstärkt auf Verhaltensanreize zur Vorsorgesteuerung setzt, sind neue Fragen hinzu gekommen: Wie verändert sich die primäre Umverteilung unter Berücksichtigung langfristiger Verhaltensanpassungen? Wie zielgenau wirken Anreize? Aber auch die Frage nach Vorsorgerisiken stellt sich neu, seitdem der kapitalgedeckten Vorsorge in Ergänzung umlagefinanzierter Sicherungsinstrumente eine größere Bedeutung zukommt: In welchem Umfang lassen sich durch Mischstrategien systemspezifische Risiken diversifizieren? Wie wirkt das Beschäftigungsrisiko auf die unterschiedlichen Vorsorgesäulen?

Methodik

In erster Näherung erlauben statische Einzelfallmodelle, Nettozahllasten einer Reform für einen spezifischen Versicherten- oder Haushaltstyp zu quantifizieren. Dazu werden Gesetzesvorschriften in numerisch lösbare Gleichungssysteme der Form $v_i = G(x_i, \theta)$ übersetzt und unter Rückgriff auf den Gesetzesparametervektor θ auf den Fall i (Versicherter, Haushalt, Steuerpflichtiger, ...), beschrieben durch einen Vektor von Zustandsvariablen $x_i \in X$, angewendet. G bezeichnet dabei den Gesetzesalgorithmus. Das in der Regel in Geld gemessene Ergebnis v_i (Bruttorente, Haushaltseinkommen, Steuerschuld, ...) wird dabei direkt als Maß für das Wohlergehen von i interpretiert. Die Auswirkung einer Gesetzesänderung von $G(X, \theta_1)$ zu $G(X, \theta_2)$ auf i ergeben sich daher als

$$\Delta v_i = G(x_i, \theta_1) - G(x_i, \theta_2). \quad (1)$$

(1) Unter Rückgriff auf eine populationsrepräsentative Stichprobe $i=1..N$ lassen sich Aussagen zu aggregierten Verteilungswirkungen und Lasteninzidenz treffen sowie fiskalische Kosten abschätzen. Dazu dienen Mikrosimulationsmodelle. AsA verfügt über beide Modelltypen (vgl. Abbildung 1). Die Stärke dieser Modelle liegt darin, die Wirkungen einer Reform „am Tag danach“ sehr detailliert abbilden zu können. Ihre Schwäche liegt darin, nicht langfristig über diesen Tag hinaus zu blicken und Verhaltensreaktionen und Risiken nicht mit zu berücksichtigen.

Im Auftrag des Forschungsnetzwerks Alterssicherung (FNA) des VDR entwickelt das FIT zur Zeit ein ergänzendes, stochastisch dynamisches Lebenszyklusmodell, AsA_{DT}. Auch dieses Modell soll helfen, Verteilungswirkungen und Anreize politischer Reformalternativen im Längs- und im Querschnitt zu verstehen. Sein Schwerpunkt liegt aber gerade in der „Problemzone“ klassischer Mikromodelle: Die Projektionen berücksichtigen Verhaltensreaktionen unter mehrdimensionaler Unsicherheit. Angesichts der Komplexität menschlicher Entscheidungen erweist es sich dazu als sinnvoll, nicht eine Gelddifferenz einer Reform nach (1), sondern eine gewissen Bedingungen genügende sog. Nutzenfunktion u als Maß des Wohlergehens zu unterstellen. Aus deren intertemporaler Maximierung können sämtliche Entscheidungen γ des Akteurs i zu allen Zeitpunkten t unter den Randbedingungen x_{it} (Vektor von Zustandsvariablen von i in t), θ_t (Vektor von Politikvariablen) konsistent abgeleitet werden. Die bedingten Verhaltensreaktionen eines Akteurs sind demnach eindeutig beschrieben durch einen Maximierungsplan des Erwartungsnutzens U_i über den erwarteten Lebenshorizont:

$$\max_{y_i \in Y} U_i = E_0 \left[\sum_{t=t_{\min}}^{T_{\max}} \beta^t \cdot u(x_{i,t}, y_{i,t}, \theta_t) \right] \quad \text{s.t.} \quad x_{i,t+1} = \Gamma(x_{i,t}, y_{i,t}) \quad (2)$$

mit β als Diskontfaktor, E als Erwartungsoperator zum Planungszeitpunkt $t=0$ und Γ als rekursives Entwicklungsgesetz des Zustands $x_i \in X$. Erfüllt die intratemporale Nutzenfunktion u bestimmte formale Bedingungen, lässt sich dieser Maximierungsplan über den erwarteten, maximalen Lebenshorizont $t = t_{\min} \dots T_{\max}$ in $T-1$ rekursive Teilmaximierungsprobleme zerlegen:

$$v_{i,t} = \max_{y_{i,t} \in Y} \left[u(x_{i,t}, y_{i,t}, \theta_t) + \beta \cdot E_t \left[v_{i,t+1}(x_{i,t+1}, y_{i,t}, \theta_{t+1}) \right] \right]. \quad (3)$$

Diese Gleichung beschreibt ein rekursives Problem dynamischer Optimierung, das numerisch approximativ lösbar ist. Sie bildet den Optimierungskern von AsA_{d1} . Drei Innovationen stehen im Mittelpunkt des Projekts:

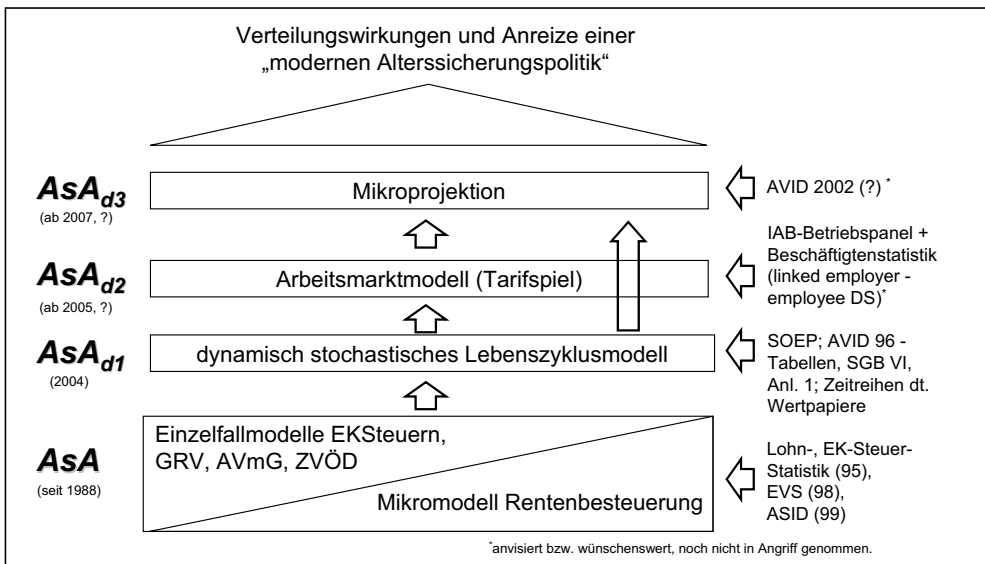
Mit AsA_{d1} liegt erstmalig ein Instrument vor, das eine detaillierte Modellierung des deutschen Steuer- und Sozialrechts in ein Entscheidungsmodell rationaler Erwartungen über mehrdimensionale Unsicherheit integriert.

Zweitens ist AsA_{d1} im Unterschied zum Gros existierender Gleichgewichtsmodelle so konstruiert, dass sukzessive Reformpfade und Übergangsregelungen konsistent analysiert werden können. Diesen Weg weiter zu entwickeln, ist insbesondere für die Alterssicherungspolitik mit ihren langen Planungshorizonten bedeutsam.

Schließlich geht mit AsA_{d1} der Versuch einher, ein ökonomisch und numerisch anspruchsvolles Analyseinstrument in einer modernen Softwareumgebung anwendungsfreundlich umzusetzen. Mit Projektabschluss wird der VDR das Simulationsinstrument zu eigenen Experimenten erhalten.

In weiteren Schritten soll AsA_{d1} um ein Arbeitsmarktmodell und eine mikroökonomische Projektion von Rentenanwartschaften ergänzt werden. Erwartungsgemäß erweist sich das Beschäftigungsrisiko in AsA_{d1} als quantitativ dominierend und kaum zu diversifizieren. Aus diesem Grund soll die Arbeitsnachfrageseite in die Analyse des Beschäftigungsrisikos mit einbezogen werden. Dazu soll ein Arbeitsmarktmodell, AsA_{d2} , entwickelt werden. Das Ziel besteht darin, in Kombination der beiden Instrumente die indirekten Auswirkungen arbeitsmarktpolitischer Reformen auf die Alterssicherung abzuschätzen. Auf diese Erkenntnisse aufbauend soll in einem

Abbildung 1: Forschungsplan (dynamisches) Analysesystem Alterssicherung



Quelle: eigene Darstellung.

dritten Schritt versucht werden, ein Entscheidungsmodell Alterssicherung in seinen Strukturparametern mikroökonomisch zu schätzen. Ziel dieses abschließenden Vorhabens, AsA_{d3} , ist die Projektion existierender Anwartschaften gegenüber den unterschiedlichen Vorsorgesäulen auf den Zeitpunkt des Renteneintritts unter Berücksichtigung von Risiken und Verhaltensanpassungen. **Abbildung 1** skizziert diese Entwicklungsperspektive und die (wünschenswerten) Hauptdatenquellen.

Datenbedarf, Zugangswege und Zusammenfassung

Die existierenden Einzelfall- und Mikromodelle aus AsA und AsA_{d1} sind gegenwärtig aus Sekundärquellen versorgt (SOEP, EVS, ASID, Lohn- und EK-Steuer-Statistik). Das Arbeitsmarktmodell AsA_{d2} wird angebotsseitig auf der empirischen Grundlage von AsA_{d1} aufbauen. Nachfrageseitig soll eine dynamische Arbeitsnachfragefunktion aus Mikrodaten des IAB-Betriebspanels (linked employer-employee Datensätze) geschätzt werden.

Problematisch ist die Datenlage in Hinblick auf die angestrebte Mikroprojektion in AsA_{d3} . Zwar wird ab 2006 mit der AVID 2002 ein geradezu idealer Datensatz für dieses Anliegen existieren. Leider sieht es gegenwärtig aber so aus, als dass die Wissenschaft erneut (wie bereits im Fall der AVID '96) keinen Zugang zu den Mikrodaten erhalten wird. Nach einer genaueren Projektspezifikation von AsA_{d3} wird zu diskutieren sein, ob sich irgendwelche Alternativen ergeben oder aber das Projekt an der Datenverfügbarkeit zu scheitern droht. Angesichts dieser Lage existieren keine Präferenzen über Zugangswege.

Martin Grub (31): Volkswirt, Projektleiter dynamisches Analysesystem Alterssicherung, Forschungsgruppe Mikroanalytische Simulationsmodelle, MikMod, am Fraunhofer Institut für Angewandte Informationstechnik, FIT.

Das **Fraunhofer-Institut für Angewandte Informationstechnik (FIT)** erforscht und entwickelt nutzerorientierte Informations- und Kooperationssysteme. Ziel ist die Optimierung der Nutzungs- und Anwendungstauglichkeit von Informationstechnik im Wechselspiel mit Organisation, Geschäftsprozess und Handlungspraxis. Das Institut adressiert sechs Anwendungsfelder: Kooperationssysteme, Teamarbeit in Erweiterter Realität, Informatik in den Lebenswissenschaften, Mobiles Wissen, Benutzerorientierte Softwaretechnik sowie Entscheidungs- und Prozessunterstützung. Hier beschäftigt sich das Institut mit Mikromodellierung und simulationsbasierter Gesetzesfolgenabschätzung, speziell für Steuer- und Transfergesetze (siehe <http://www.fit.fraunhofer.de>).

Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland

Dr. Holger Seibert (Max-Planck-Institut für Bildungsforschung, Berlin)

In diesem Beitrag wird das Projekt „Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland“ vorgestellt, welches am Max-Planck-Institut für Bildungsforschung in Berlin im Forschungsbereich „Bildung, Arbeit und gesellschaftliche Entwicklung“ durchgeführt wird. An diesem Projekt wird exemplarisch der Bedarf für die Daten der Rentenversicherung sowie die Anforderungen an das Forschungsdatenzentrum der Rentenversicherung aus der Sicht dieser Forschungsgruppe des Max-Planck-Instituts erläutert.

Das Projekt beschäftigt sich mit den Auswirkungen ökonomischer und sozialer Entwicklungen der 1980er und 1990er Jahre auf Bildung, Ausbildung und Erwerbstätigkeit junger Erwachsener. Dabei werden die Folgen von Bildungsexpansion, berufsstrukturellem Wandel, demographischen Schwankungen und sich tendenziell verschlechternden Arbeitsmarktbedingungen in ihren Auswirkungen auf soziale, geschlechtsspezifische und ethnische Unterschiede untersucht.

Es gilt herauszufinden, in welchem Ausmaß der „Normallebensverlauf“ noch immer durch relativ standardisierte Übergänge zwischen Bildung, Ausbildung und der Arbeitswelt geprägt ist. Dabei interessiert auch, welche Rolle bestehende soziale Institutionen und sozialpolitische Steuerungsinstrumente für die Sicherung eines hohen Ausbildungsniveaus und die Platzierung der Menschen auf dem Arbeitsmarkt spielen.

Theoretisch stützt sich das Projekt auf die Vorleistungen, die im Rahmen vorangegangener Forschungsprojekte am Max-Planck-Institut entstanden sind (vgl. hierzu Mayer & Carrol 1987, Blossfeld & Mayer 1988, Blossfeld 1990, Mayer 1996, 2001). Unter dem Begriff Lebensverlauf wird nicht nur die „Abfolge von Aktivitäten und Ereignissen in verschiedenen Lebensbereichen bzw. Handlungsfeldern“ (Mayer 2001, 446) verstanden. Über die „Teilhabe an gesellschaftlicher Positionen“ und damit die „Mitgliedschaft in institutionellen Ordnungen“ zielt der Lebensverlaufsbegriff damit gleichzeitig auf die „soziostrukturelle Einbettung von Individuen im Verlauf ihres gesamten Lebens“ (ebd.).

Mit der Lebensverlaufsstudie, die am Max-Planck-Institut seit den 1980er Jahren erhoben wird, steht eine umfangreiche Datenbasis für die empirische Forschung zur Verfügung. In der Lebensverlaufsstudie werden Personen retrospektiv zu verschiedenen Aspekten ihres Lebensverlaufs von der Geburt bis zum Erhebungszeitpunkt befragt. Im Vordergrund der Befragung stehen die individuellen Bildungs-, Ausbildungs- und Erwerbsverläufe. Daneben werden die Wohngeschichte sowie Informationen zu Eltern, Partnern und Kindern rekonstruiert.

Indem die Daten der Geburtskohorten 1964 und 1971 in den Kontext vorangegangener Lebensverlaufsstudien gestellt werden, ist es möglich, die kollektiven Erfahrungen im Übergang zwischen Bildung, Ausbildung und Arbeitswelt über Jahrzehnte hinweg bis in die jüngste Zeit zu rekonstruieren. Dadurch lässt sich empirisch untersuchen, wie sich die institutionellen Verknüpfungen zwischen Schule und Ausbildungssystem sowie zwischen Arbeitsmarkt und Tätig-

keitsstrukturen verändert haben. Über Dauer und Fragmentierung dieser Übergangsphasen und über die verschiedenen Wege zum Erwerb von Fertigkeiten oder einen frühen Einstieg in das Arbeitsleben wurden konkurrierende Hypothesen gebildet, die im aktuellen Projekt eingehend untersucht werden. Ein Überblick über die Bandbreite der mit den Daten untersuchten Fragestellungen geben Mayer & Hillmert (im Erscheinen).

Die vom Forschungsbereich gesammelten Lebensverlaufsdaten werden mit Längsschnittdaten der amtlichen Statistik verbunden. Der Vergleich der Daten des Projektes mit vom Institut für Arbeitsmarkt- und Berufsforschung aufbereiteten Daten aus der Beschäftigtenstatistik gibt Aufschluss über die Beziehungen zwischen Entwicklung und Eigenschaften von Unternehmen und individuellen Erwerbskarrieren. Die Kombination beider Datenquellen ermöglicht es außerdem, der Frage nach Qualität und Genauigkeit von retrospektiv erhobenen Daten auf den Grund zu gehen (vgl. Reimer 2003).

Im Zuge der jüngsten Kohortenstudie wurden 1998/99 ca. 2.900 Personen befragt, die in den Jahren 1964 bzw. 1971 geboren wurden (vgl. Hillmert im Erscheinen). Zum ersten Mal in einer Teilstudie des Lebensverlaufsprojekts sind auch deutschsprachige Ausländer eingeschlossen. Es ist darüber hinaus anhand der Elterndaten möglich, auch Migranten bzw. Personen mit Migrationshintergrund zu identifizieren, die bereits über die deutsche Staatsangehörigkeit verfügen (vgl. Seibert im Erscheinen).

1964 war der geburtenstärkste Jahrgang seit Bestehen der Bundesrepublik. Aus diesem Grund ist anzunehmen, dass die Mitglieder dieser Kohorte hinsichtlich ihrer Chancen auf dem Ausbildungs- und Arbeitsmarkt in ungleich höherem Maße Konkurrenzdruck ausgesetzt gewesen sind. In Westdeutschland ist die Anzahl der Mitglieder beider Geburtskohorten über die Jahre auf Grund von Immigration stetig angewachsen (Hillmert 2001).

Wie sich der Eintritt in den Arbeitsmarkt in Ost- und Westdeutschland gestaltet, kann an der Kohorte 1971 gut untersucht werden, da die erhobenen Daten hinsichtlich ihrer Struktur mit den Daten des Teilprojektes „Ostdeutsche Lebensverläufe nach der Wiedervereinigung“ (1996) übereinstimmen (vgl. z.B. Diewald et al. 2000, Solga & Diewald 2001, Goedicke 2002).

Zwar bieten die erhobenen Lebensverlaufsdaten, deren Qualität in einem aufwändigen Editingsverfahren nach der Erhebung noch einmal deutlich verbessert wurde (Hillmert 2002), ein großes Analysepotenzial insbesondere im Hinblick auf Längsschnittauswertungen und damit auf die Analyse der zeitlichen Binnenstruktur von Bildungs- und Ausbildungsverläufen (Hillmert im Erscheinen), dennoch sind auch einige Einschränkungen hinzunehmen. Dabei handelt es sich im Wesentlichen um:

- die Beschränkung auf nur zwei Geburtsjahre (1964 und 1971)
- Ungenauigkeiten und mögliche Erinnerungsfehler bei der retrospektiven Befragung
- geringe Fallzahlen für Ausländer und Personen mit Migrationshintergrund.

Die Beschränkungen auf die Geburtskohorten 1964 und 1971 macht lediglich einen Vergleich dieser beiden Kohorten bzw. einen Vergleich mit den älteren Kohorten aus der Lebensverlaufs-

studie möglich. Analysen im historischen Längsschnitt über mehrere Kohorten hinweg sind so nicht möglich.

Ungenauigkeiten und Erinnerungsfehler in der retrospektiven Rekonstruktion von Lebensverläufen reduzieren die Validität der erhobenen Daten. Es wurde daher ein Erhebungsinstrument gewählt, das das Erinnern bestimmter Ereignisse bzw. Episoden positiv fördert. So haben Tests ergeben, dass zeitliche Abweichungen in der erinnerten Lebensgeschichte meist nur einige wenige Monate betragen und symmetrisch ausfallen, d.h. es wird in der Erinnerung sowohl länger als auch kürzer datiert, sodass hier kein systematischer Fehldatier-Bias entsteht. Es kommt jedoch mitunter vor, dass kurze Episoden, die in der subjektiven Lebensgeschichte der Personen keinen hohen Stellwert einnehmen (z.B. Überbrückungsarbeitslosigkeiten), nicht oder nur unzuverlässig und eher verkürzt berichtet werden (Reimer 2003).

Schließlich stehen für kleinere Bevölkerungsgruppen, z.B. Ausländer, nur relativ geringe Fallzahlen zur Verfügung. Zwar liegt die Gesamtfallzahl der Erhebung mit 2.900 Fällen bereits über dem Durchschnitt, der bei Befragungen mit standardisierten Interviews in der Regel realisiert wird, dennoch sind die Fallzahlen von Minderheitengruppen oft zu klein für eingehende statistische Analysen. Der Umstand, dass in der Lebensverlaufsstudie nur deutschsprechende Ausländer befragt wurden, führt zudem zu einer deutlichen Unterrepräsentanz im Vergleich zur ausländischen Gesamtbevölkerung.

Zusätzlicher Datenbedarf

Um den oben beschriebenen Einschränkungen zu begegnen, bietet es sich an, zusätzliche Datenquellen heranzuziehen, um die Qualität der eigenen Daten zu überprüfen bzw. Analysen durchzuführen, die die Daten aufgrund der dargestellten Probleme nicht zulassen. Dies ist zum Teil bereits geschehen bzw. in Angriff genommen worden (Hillmert & Kröhnert 2001, Matthes & Wehner 2002, Reimer & Künster 2004, Seibert 2004 im Erscheinen).

Die Daten der Rentenversicherung bieten sich hierbei als besonders geeignete Ergänzung an, sofern es mit ihnen möglich ist, individuelle Ausbildungs- und Erwerbsverläufe nachzuzeichnen. Dies sollte mit Rückgriff auf die *Statistik der aktiv und latent Versicherten* prinzipiell möglich sein. Der besondere Reiz solcher Daten liegt aus unserer Sicht in der besonderen Genauigkeit der Daten in Bezug auf die zeitliche Dauer und Abfolge von Lebensepisoden.

Datierungsprobleme gibt es hier nicht, da es sich um Meldungen der Arbeitgeber bzw. der Versicherten selbst handelt, die bereits zum Zeitpunkt der Versicherungsnahme gemacht wurden. Auch dürften Ausländer und andere Minderheiten in umfangreicher Fallzahl repräsentiert sein, da der Datenbestand der Rentenversicherung einer Vollerhebung gleichkommt. Darüber hinaus sind mit den Rentendaten keine Beschränkungen auf bestimmte Geburtsjahrgänge hinzunehmen.

Neben der *Statistik der aktiv und latent Versicherten* ist für uns auch die *Versicherungskontenstichprobe* ein interessanter Datensatz, der je nach Fallzahl ebenfalls die zuvor beschriebenen Potenziale für empirische Analysen besitzen sollte.

Wünschenswerte Datenzugangswege

Für unseren Forschungsbereich wäre es wünschenswert, die Daten der Rentenversicherung in unserem Haus selbst auswerten zu können. Sollte dies aus datenschutztechnischen Gründen nicht möglich sein, würden wir eine Art der kontrollierten Datenfernverarbeitung präferieren, die uns einen Zugang auf die Rechner des Forschungsdatenzentrums ermöglicht (z.B. über CITRIX), sodass wir die Analysen selbständig über das Internet erledigen können.

Das Angebot, als Gastwissenschaftler am Forschungsdatenzentrum tätig zu sein, wäre für uns aufgrund der anfallenden Übernachtungskosten vermutlich nur für kurze Zeiten von maximal einer Woche denkbar. Ein solcher Aufenthalt würde sich auf die tatsächlichen Rechenarbeiten beschränken und sollte auf jeden Fall von einem wissenschaftlichen Mitarbeiter bzw. wissenschaftlich-technischen Assistenten des Forschungsdatenzentrums begleitet werden. Insgesamt ist es aber fraglich, ob empirische Analysen innerhalb einer Woche zufriedenstellend abgeschlossen werden können. Sollte später die Notwendigkeit der Modifikation der Berechnungen festgestellt werden, wäre ein neuerlicher Aufenthalt im Forschungsdatenzentrum nötig. Es bleibt die Vermutung, dass diese Art des Datenzugriffs eher umständlich ist und ein relativ hohes Risiko birgt, am Ende zu nur halb fertigen Berechnungen zu gelangen.

Für die Vorbereitung eines Forschungsaufenthalts im Forschungsdatenzentrum aber auch für eine Datenfernverarbeitung wäre es hilfreich, zuvor auf anonymisierte Testdaten (wie z.B. Campus-Files) Zugriff zu haben, um so das Analysepotenzial der Daten einschätzen zu können und ein „Gefühl für die Daten“ zu bekommen. Mindestens aber sollten Variablenlisten und Codepläne für die Vorbereitung empirischer Analysen zur Verfügung stehen.

Zusammenfassung

In diesem Beitrag wurde das Projekt „Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland“ vorgestellt, welches am Max-Planck-Institut für Bildungsforschung im Fachbereich „Bildung, Arbeit und gesellschaftliche Entwicklung“ angesiedelt ist. Im Rahmen des beschriebenen Projekts wurden retrospektiv erfragte, individuelle Lebensverlaufsdaten von ca. 2900 Personen (Lebensverlaufsstudie West 1964/1971) erhoben, mit denen bereits umfangreiche empirische Analysen durchgeführt wurden und noch immer werden.

Als Ergänzung zum Analysepotenzial dieser Daten sowie zur externen Validierung würden wir gern auf die Daten der Rentenversicherung zurückgreifen. Hierzu bieten sich insbesondere die Statistik der aktiv und latent Versicherten sowie die Versicherungskontenstichprobe an. Um die Arbeit mit den Rentendaten für uns möglichst effektiv zu gestalten, plädieren wir für eine Nutzung der Daten in unserem Haus. Wenn dies mit dem Datenschutz nicht vereinbar sein sollte, wäre eine Datenfernverarbeitung (z.B. über CITRIX) eine mögliche Alternative. Das Angebot von Gastwissenschaftler-Arbeitsplätzen im Forschungsdatenzentrum stellt für uns lediglich eine nachrangige Möglichkeit des Datenzugangs dar.

Im Vorfeld der Arbeit mit den Daten der Rentenversicherung sollten Testdatensätze sowie Variablenlisten und Codepläne über das Internet zur Verfügung gestellt werden, um so das Analysepotenzial der Daten möglichst genau einschätzen zu können.

Literatur

- Blossfeld, H.-P. (1990): Berufsverläufe und Arbeitsmarktprozesse. Ergebnisse sozialstruktureller Längsschnittuntersuchungen, in: Mayer, Karl Ulrich (Hg.): *Lebensverläufe und sozialer Wandel. Sonderheft 31 der Kölner Zeitschrift für Soziologie und Sozialpsychologie*, Opladen: Westdeutscher Verlag, 118-145.
- Blossfeld, H.-P. & Mayer, K.U. 1988: Arbeitsmarktsegmentation in der Bundesrepublik Deutschland. Eine empirische Überprüfung von Segmentierungstheorien aus der Perspektive des Lebenslaufs, *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 40, 2, 262-283.
- Diewald, M., Goedicke, A, & Solga, H. (2000): Arbeitsmarktcompetenzen in Ostdeutschland - nicht vorhanden oder nicht gefragt? *Aus Politik und Zeitgeschichte*, B 40, 39-46.
- Goedicke, A. (2002): *Beschäftigungschancen und Betriebszugehörigkeit. Die Folgen betrieblichen Wandels für ostdeutsche Erwerbstätige nach 1989*. Wiesbaden: Westdeutscher Verlag.
- Hillmert, S. (2001): Kohortendynamik und Konkurrenz an den zwei Schwellen des dualen Ausbildungssystems: Übergänge zwischen Schule und Arbeitsmarkt im Kontext ökonomischen und demographischen Wandels. *Arbeitspapier Nr. 2 des Projekts Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland*, Berlin: Max-Planck-Institut für Bildungsforschung.
- Hillmert, S. (2002): Edition von Lebensverlaufsdaten: Zur Relevanz einer systematischen Einzelfallbearbeitung bei standardisierten Befragungen. *ZUMA-Nachrichten*, 51, 120-140.
- Hillmert, S. (im Erscheinen): Die Lebensverlaufsstudie 64/71: Projekt, Datenerhebung und Edition, in: Mayer, K.U. & Hillmert, S. (Hg.): Geboren 1964 und 1971. *Untersuchungen zum Wandel von Ausbildungs- und Berufschancen*, Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hillmert, S. & Kröhnert, S. (2001): *Vergleich der Lebensverlaufsstudie 64/71-West mit Erhebungen des Mikrozensus auf Basis ausgewählter Randverteilungen*. Technischer Arbeitsbericht, Berlin: Max-Planck-Institut für Bildungsforschung.
- Matthes, B. & Wehner, S. (2002): Ein Vergleich ausgewählter Randverteilungen der Lebensverlaufsstudie Ost (Geburtskohorte 1971) mit den Erhebungen des Mikrozensus. *Arbeitsbericht Nr. 1 aus dem Projekt Ostdeutsche Lebensverläufe im Transformationsprozess*, Berlin: Max-Planck-Institut für Bildungsforschung.
- Mayer, K.U. (1996): Ausbildungswege und Berufskarrieren. In: Bundesinstitut für Berufsbildung (Hg.): *Forschung im Dienst von Praxis und Politik. Dokumentation der Festveranstaltung zum 25-jährigen Bestehen des Bundesinstituts für Berufsbildung*, Bielefeld: Bertelsmann, 113-145.
- Mayer, K.U. (2001): Lebensverlauf. In: Schäfers, Bernhard / Zapf, Wolfgang (Hg.): *Handwörterbuch zur Gesellschaft Deutschlands*, 2. Auflage, Opladen: Leske + Budrich, 446-460.
- Mayer, K.U. & Carrol, G.R. (1987): Jobs and Classes: Structural Constraints on Career Mobility. *European Sociological Review*, 3, 1, 14-38.
- Mayer, K.U. & Hillmert, S. (Hg.) (im Erscheinen): Geboren 1964 und 1971. *Untersuchungen zum Wandel von Ausbildungs- und Berufschancen*, Wiesbaden: VS Verlag für Sozialwissenschaften.
- Reimer, M. (2003): Autobiografisches Erinnern und retrospektive Längsschnittdatenerhebung: Was wissen wir und was würden wir gerne wissen? *BIOS - Zeitschrift für Biographieforschung, Oral History und Lebensverlaufsanalysen*, 16, 1.

- Reimer, M. & Künster, R. (2004): *Linking job episodes from retrospective surveys and social security data: Specific challenges, feasibility and quality of outcome*. Paper to be presented at the RC 33 Sixth International Conference on Social Science Methodology.
- Seibert, H. (2004): *Integration durch Ausbildung? Berufliche Platzierung ausländischer Ausbildungsabsolventen der Geburtsjahrgänge 1960 bis 1971*. Dissertation am Fachbereich Politik- und Sozialwissenschaften der Freien Universität Berlin.
- Seibert, H. (im Erscheinen): *Wer zu spät kommt... Schulausbildung und der Erwerbseinstieg von Ausbildungsabsolventen ausländischer Herkunft in Deutschland*, in: Mayer, K.U. & Hillmert, S. (Hg.): *Geboren 1964 und 1971*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Solga, H. & Diwald, M. (2001): *The East German Labour Market after German Unification: A Study of Structural Change and Occupational Matching*. *Work, Employment & Society*, 15, 1, 95-126.

Dr. Holger Seibert ist wissenschaftlicher Mitarbeiter am Max-Planck-Institut für Bildungsforschung im Forschungsbereich „Bildung, Arbeit und gesellschaftliche Entwicklung“. Seine Forschungsinteressen konzentrieren sich auf Bildung, Berufsausbildung und Arbeitsmarkteinstieg, Migranten im Bildungs- und Ausbildungssystem sowie die berufliche Platzierung von Ausländern in Deutschland.

Der **Forschungsbereich „Bildung, Arbeit und gesellschaftliche Entwicklung“ des Max-Planck-Instituts für Bildungsforschung** analysiert die Zusammenhänge und Wechselwirkungen zwischen Familie, Bildungssystem und Arbeitsmarkt im Kontext sich wandelnder Wohlfahrtsstaaten. Im Kern des Forschungsprogramms steht die Analyse von Lebensverläufen als Schnittstelle zwischen institutionellen Umwelten, menschlicher Entwicklung und sozio-ökonomischem Wandel (siehe www.mpib-berlin.mpg.de).

Berentungen aufgrund verminderter Erwerbsfähigkeit bei Arbeitnehmern in der Bauwirtschaft

*Dr. Volker Arndt, PD Dr. Dietrich Rothenbacher und Prof. Dr. Hermann Brenner
(Deutsches Zentrum für Altersforschung, Heidelberg)*

Einleitung

Berentungen aufgrund verminderter Erwerbsfähigkeit sind sowohl unter sozialmedizinischen Gesichtspunkten als auch aus der Perspektive der Alterssicherung nach wie vor sehr bedeutsam. Derzeit werden in der Bundesrepublik Deutschland pro Jahr etwa 180.000 Männer und Frauen aus gesundheitlichen Gründen frühberentet. Durch ein Screening im Rahmen der arbeitsmedizinischen Betreuung könnten frühberentungsgefährdete Arbeitnehmer identifiziert und gezielten Interventionsmaßnahmen zugewiesen werden, um deren weitere Arbeitsfähigkeit zu sichern und eine drohende Frühberentung abzuwenden. Ziel des geplanten Forschungsvorhabens ist, die prognostische Bedeutung beruflicher, medizinischer und im persönlichen Verhalten begründeter Risikofaktoren bzw. Risikoindikatoren für die Frühinvalidität aufgrund sozialmedizinisch besonders relevanter Diagnosen (Erkrankungen des Stütz- und Bewegungsapparates, Herz-Kreislaufkrankungen, bösartige Neubildungen etc.) detailliert zu untersuchen. Hierzu wurde von der Projektgruppe eine für Deutschland bislang einzigartige Kohorte von 20.000 Beschäftigten der Bauwirtschaft, einer Branche mit überdurchschnittlich hohen Frühberentungsraten, aufgebaut. Im Rahmen des geplanten Forschungsvorhabens soll nun diese Kohorte nach einer durchschnittlichen Follow-up-Dauer von 15 Jahren über einen Abgleich mit Angaben aus dem Datenbestand des VDR im Hinblick auf das Frühberentungsgeschehen nachbeobachtet werden.

Fragestellung

Die Studie soll dabei gezielt folgende Fragestellungen untersuchen:

- Welche gesundheitlichen Risikoschwerpunkte (Diagnosen) lassen sich beim Frühberentungsgeschehen von Bauarbeitern identifizieren?
- Welche beruflichen, individuellen und medizinischen Faktoren weisen prospektiv auf ein erhöhtes Frühberentungsrisiko hin und können langfristig für ein Screening auf eine drohende Frühberentung eingesetzt werden?

Ziel und möglicher Erkenntnisgewinn

Ziel des geplanten Forschungsvorhabens ist, die prognostische Bedeutung beruflicher, medizinischer und im persönlichen Verhalten begründeter Risikofaktoren bzw. Risikoindikatoren für die Frühinvalidität bzw. Sterblichkeit aufgrund sozialmedizinisch besonders relevanter Diagnosen (Erkrankungen des Stütz- und Bewegungsapparates, Herz-Kreislaufkrankungen, bösartige Neubildungen etc.) detailliert zu untersuchen. Darüber hinaus soll das Forschungsvorhaben

neue Erkenntnisse im Hinblick auf mögliche neue Berufskrankheiten sowie die Verhütung sonstiger arbeitsbedingter Gesundheitsrisiken erbringen.

Die Kenntnis von Risikofaktoren einer Frühberentung oder einer erhöhten Mortalität ist Voraussetzung für gezielte Präventions- bzw. Interventionsmaßnahmen zum Schutz der Arbeitnehmer und daher für den Arbeitgeber im Rahmen seiner Fürsorgepflicht gegenüber den Beschäftigten sowie für die Träger der gesetzlichen Unfallversicherung im Rahmen ihres erweiterten Präventionsauftrags von besonderem Interesse. Die zu erwartenden Ergebnisse tragen erheblich zu einer Verbesserung des Frühwarnsystems für Beschäftigte der Bauwirtschaft bei und unterstützen damit einen effizienten Einsatz von Präventions- und Interventionsmaßnahmen zur Verringerung von Morbidität, Frühberentung und Mortalität einer besonders belasteten Berufsgruppe.

Theoretische Einbettung und Vorarbeiten

Es ist bekannt, dass nur rund die Hälfte aller aufgrund einer verminderten Erwerbsfähigkeit Frühberenteten in den letzten 5 Jahren vor Rentenbeginn an einer Rehabilitationsmaßnahme teilgenommen hat (Schuntermann et al. 1990). Bisherige Untersuchungen anderer (vgl. Biefang et al. 1990/1996/1998) als auch unserer Arbeitsgruppe (vgl. Arndt et al. 1995/1996/1998/1999/2004 sowie Rothenbacher et al. 1997a/1997b/1998) zum Screening rehabilitationsbedürftiger Versicherter konnten erste Anhaltspunkte für Risikoindikatoren, die mit einer erhöhten Frühberentungswahrscheinlichkeit einhergehen, identifizieren. Allerdings ist der Kenntnisstand über die langfristige prognostische Bedeutung von Belastungsfaktoren und körperlichen Befunden noch unzureichend, so dass eine gezielte Identifikation von Risikogruppen und Risikotätigkeiten als Voraussetzung für eine effektive Prävention im Rahmen der arbeitsmedizinischen Betreuung bislang nur eingeschränkt möglich war. Hier fehlen insbesondere für Deutschland die Ergebnisse groß angelegter Kohortenstudien, um Risikofaktoren und Risikoindikatoren einer gesundheitsbedingten Frühberentung zu identifizieren und daraus konkrete Ansatzpunkte für die Prävention abzuleiten.

Angesichts der demographischen Entwicklung, der hohen Belastung und Beanspruchung von Arbeitnehmern in der Bauwirtschaft sowie den hohen Frühberentungsraten in der Bauwirtschaft wurde vor ungefähr zehn Jahren die Studie „Risikogruppe ältere Arbeitnehmer in der Bauwirtschaft“ von der Arbeitsgemeinschaft der Bau-Berufsgenossenschaften initiiert und von der Projektgruppe, bestehend aus Vertretern des Arbeitsmedizinischen Dienstes der Württembergischen Bau-Berufsgenossenschaft (Leiter: Dr. med. U. Daniel), dem Institut für Arbeits-, Sozial- und Umweltmedizin an der Universität Ulm (damaliger Leiter Prof. Dr. med. Dr. h.c. T.M. Fliedner) sowie der Abteilung Epidemiologie am Deutschen Zentrum für Altersforschung (vormals Universität Ulm; Leiter Prof. Dr. med. H. Brenner, M.P.H.), konzipiert und erfolgreich durchgeführt. In enger Zusammenarbeit mit der LVA Württemberg, dem Sozialministerium Baden-Württemberg und dem Landesbeauftragten für Datenschutz wurde dabei ein effizientes Verfahren zur Nachbeobachtung großer Kohorten entwickelt und eingesetzt.

In dieser Studie konnte gezeigt werden, wie man mit epidemiologischen Methoden die Bedeutung der bei den arbeitsmedizinischen Vorsorgeuntersuchungen erhobenen Befunde und Diagnosen für die Morbidität, Frühberentung und Mortalität quantifizieren kann. Damit wurde eine

Basis für eine gezielte Vorsorge einschließlich der erforderlichen Präventions- und Rehabilitationsmaßnahmen geschaffen. Die zahlreichen präventiv-medizinisch relevanten Ergebnisse haben Eingang in die weltweit führenden arbeitsmedizinischen und epidemiologischen Fachzeitschriften und große Beachtung gefunden (vgl. Arndt et al. 1995/1996/1998/1999/2004 sowie Rothenbacher et al. 1997a/1997b/1998).

Es zeigte sich aber, dass innerhalb der bisherigen Follow-up-Periode von durchschnittlich zehn Jahren annähernd ein Viertel (23%) aller Versicherten zu einem anderen Versicherungsträger gewechselt ist und für das Follow-up nicht weiter berücksichtigt werden konnte, da nur die Daten aus dem Bestand der Landesversicherungsanstalt Württemberg zur Verfügung standen. Ob die hohe berufliche Mobilität Auswirkungen auf die Ergebnisse zum Frühberentungsgeschehen hat, ist bislang noch nicht mit Sicherheit zu beurteilen. Durch die Einrichtung des FDZ-RV mit der Möglichkeit, das Frühberentungsgeschehen in der Ausgangskohorte vollständig zu erfassen, könnte die Aussagekraft der Studie entscheidend erhöht werden.

Methodisches Vorgehen

Studiendesign und Studienpopulation

Im Rahmen einer retrospektiven Kohortenstudie werden ca. 20.000 Beschäftigte der Bauindustrie, bei denen in den Jahren 1986-1992 eine routinemäßige, arbeitsmedizinische Vorsorgeuntersuchung in den arbeitsmedizinischen Zentren der Württembergischen Bau-Berufsgenossenschaft durchgeführt wurde, auf in der Zwischenzeit eingetretene gesundheitsbedingte Frühberentungen nachbeobachtet. Ausgangspunkt stellen die im Rahmen dieser arbeitsmedizinischen Vorsorgeuntersuchungen standardisiert erhobenen und auf EDV erfassten Angaben zu Anamnese, Arbeitsplatzbedingungen sowie die Befunde und Diagnosen der eingehenden körperlichen Untersuchung (einschließlich Laborbefunde und Funktionsuntersuchungen) dar. Bei einem Großteil der Beschäftigten wurden zwischenzeitlich wiederholte Untersuchungen durchgeführt (durchschnittlich ca. drei pro Teilnehmer), die ebenfalls auf EDV erfasst wurden.

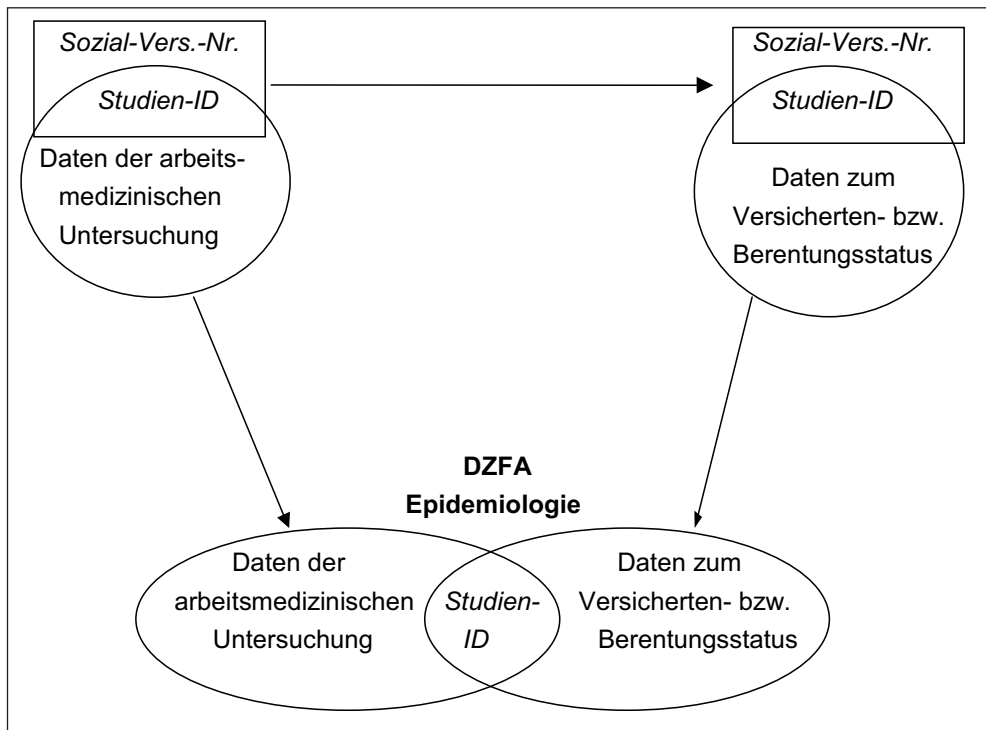
Ermittlung des Berentungsstatus und Übermittlung der Daten unter Wahrung datenschutzrechtlicher Belange

Für das Forschungsprojekt sollen die im Rahmen der arbeitsmedizinischen Vorsorgeuntersuchungen erhobenen Daten der Studienteilnehmer um Angaben zum Frühberentungsgeschehen, zu stattgefundenen Rehabilitationsmaßnahmen und zum aktuellen Vitalstatus über den Datenbestand des VDR (Routinestatistiken) ergänzt werden. Das Vorgehen und der Inhalt der Abfrage entsprechen dabei weitgehend dem Verfahren, das in der Vergangenheit bereits erfolgreich für (und beschränkt auf) Daten der LVA Württemberg durchgeführt werden konnte und für diese beschränkte Anwendung vom zuständigen Landesbeauftragten für Datenschutz sowie Sozialministerium geprüft und genehmigt wurde.

Hierzu übermittelt der Arbeitsmedizinische Dienst der Württembergischen Bau-Berufsgenossenschaft die zu einer eindeutigen Personenzuordnung erforderlichen Daten (Rentenversicherungs-Nummer, ggf. Name/Geburtsdatum/Adresse) sowie einen Umsteiger („Studien-ID“), jedoch ausdrücklich keine medizinischen Daten an das Forschungsdatenzentrum des VDR

(siehe **Abbildung 1**). Die in diesem Schritt übertragenen Daten sind für den VDR ohne jeglichen Informationswert. Dort erfolgt eine automatisierte Ermittlung aller relevanten „Follow-up“-Daten der Studienteilnehmer aus dem Datenbestand der Routinestatistiken (aktueller Versichertenstatus, Datum und Hauptdiagnose einer eventuellen Frühberentung sowie etwaiger Rehabilitationsmaßnahmen) anhand der Sozialversicherungsnummer (oder eines seitens der Rentenversicherungsträger generierten Umsteigers).

Abbildung 1: Zusammenführung der Daten unter Wahrung datenschutzrechtlicher Belange



Quelle: eigene Darstellung.

Die im Rahmen der Nachbeobachtung erhobenen Daten zum Berentungsgeschehen bzw. zur Mortalität werden dann in faktisch anonymisierter Form (nur mit der Studien-ID versehen, also ohne Sozialversicherungsnummer, Name, Geburtsdatum und Adresse) an die Abteilung Epidemiologie am Deutschen Zentrum für Altersforschung in Heidelberg übermittelt, dort über die Studien-ID mit den ebenfalls faktisch anonymisiert vorliegenden Daten der arbeitsmedizinischen Untersuchungen zusammengeführt und epidemiologisch ausgewertet. Es werden im DZFA keine Daten vorgehalten oder analysiert, über die indirekt oder direkt ein Personenbezug herstellbar wäre.

Datenbedarf

Im Einzelnen sollen folgende Angaben aus dem „Versichertenkonto“ übermittelt werden:

- Aktueller Versichertenstatus (noch versichert bzw. Versicherungsverhältnis beendet)
- Aktuelle bzw. letzte Tätigkeit
- Falls Versicherungsverhältnis beendet:
 - Datum der Beendigung des Versicherungsverhältnisses
 - Grund der Beendigung des Versicherungsverhältnisses (Altersrente, BU/EU-Rente bzw. Berentung aufgrund verminderter Erwerbsfähigkeit, Verstorben, Sonstiges)
- Im Falle einer EU/BU-Rente bzw. Berentung aufgrund verminderter Erwerbsfähigkeit:
 - Leistungsart der Rente (EU, BU, Erwerbsfähigkeit teilw. bzw. voll gemindert)
 - Befristung: ja/nein
 - Hauptdiagnose (nach ICD kodiert)
 - Prüfung der Arbeitsmarktlage
- Stationäre Rehabilitationsmaßnahmen (im Laufe des Follow-up): ja/nein
 - Beginn (Monat/Jahr) der angetretenen Rehabilitationsmaßnahmen
- Sterbedatum bei Todesfall nach Beendigung des Versicherungsverhältnisses

Zugangswege

Die für die Durchführung der skizzierten Studie notwendigen Daten sind allesamt in den Routinestatistiken des VDR als Individualdaten verfügbar:

- Stammsatzbestandsstatistik (Vitalstatus)
- Aktiv Versicherte (Tätigkeitsschlüssel)
- Jahresehebung der Reha-Statistik-Datenbasis (Datum und Diagnose abgeschlossener Reha- Maßnahmen)
- Rentenzugangstatistik (Diagnose, Leistungsart, Wegfallsgrund, Datum, ...)

Lediglich zur Ermittlung der entsprechenden Follow-up-Daten aus dem Bestand des VDR und zur Zusammenführung mit den Untersuchungsdaten auf individueller Ebene sind Modifikationen der Routinestatistiken erforderlich.

Zusammenfassung

Durch die Etablierung des FDZ-RV könnte erstmals für Deutschland ein effizientes Verfahren für die Durchführung von länderübergreifenden Kohortenstudien zum Thema Frühberentung und

Rehabilitation realisiert werden. Die bisherige Konzeption für das Forschungsdatenzentrum sieht bislang keine personenbezogene Verknüpfung aus den unterschiedlichen Routinestatistiken vor. Es ist an die Entscheidungsträger zu appellieren, die für einen personenbezogenen Abgleich unter Wahrung der datenschutzrechtlichen Belange notwendigen Voraussetzungen zu schaffen.

Literatur

- Arndt, V., Brenner, H., Daniel, U., Fliedner, T.M., Fraise, E., Rothenbacher, D., Schuberth, S. & Zschenderlein, B. (1995): Risikogruppe Ältere Arbeitnehmer in der Bauwirtschaft - Stufe III Retrospektive Kohortenstudie. *Schriftenreihe Arbeitssicherheit und Arbeitsmedizin in der Bauwirtschaft*, Band 10.2, Berufsgenossenschaften der Bauwirtschaft, Böblingen, Ulm.
- Arndt, V., Rothenbacher, D., Brenner, H., Fraise, E., Zschenderlein, B., Daniel, U., Schuberth, S. & Fliedner, T.M. (1996): Older workers in the construction industry: results of a routine health examination and a five year follow up. *Occupational and environmental medicine*, 53, 686-691.
- Arndt, V., Brenner, H., Rothenbacher, D., Zschenderlein, B., Fraise, E. & Fliedner, T.M. (1998): Elevated liver enzyme activity in construction workers: prevalence and impact on early retirement and all-cause mortality. *International archives of occupational and environmental health*, 71, 405-412.
- Arndt, V. & Brenner, H. (1999): *Prädiktoren der Frühberentung bei Arbeitnehmern in der Bauwirtschaft*, Abschlussbericht an das Forschungsinstitut für Rehabilitationsmedizin an der Universität Ulm, Ulm.
- Arndt, V., Rothenbacher, D., Daniel, U., Zschenderlein, B., Schuberth, S. & Brenner, H. (2004): *Construction work and disability pension - a 10-year follow-up of 14 000 male workers*, b, (submitted).
- Biefang, S., Gerdes, N., Hoeltz, J. & Potthoff, F. (1990): Prädiktoren der Frühberentung und gezielte Zuweisung zur medizinischen Rehabilitation. *Sozial- und Präventivmedizin*, 3, 129-137.
- Biefang, S., Potthoff, F., Bellach, B., Ziese, T. & Buschmann-Steinhage, R. (1996): Prädiktoren des Rehabilitations- und Berentungsgeschehens - Ergebnisse einer Längsschnittuntersuchung. *Deutsche Rentenversicherung*, 1-2, 84-108.
- Biefang, S., Potthoff, P., Bellach, B.M. & Buschmann-Steinhage, R. (1998): Predictors of early retirement and rehabilitation for use in a screening to detect workers in need of rehabilitation. *International Journal of Rehabilitation Research*, 21, 13-28.
- Rothenbacher, D., Arndt, V., Fraise, E., Daniel, U., Fliedner, T.M. & Brenner, H. (1997a): Chronic respiratory disease morbidity in construction workers: patterns and prognostic significance for permanent disability and overall mortality. *The European respiratory journal*, a, 10, 1093-1099.
- Rothenbacher, D., Brenner, H., Arndt, V., Fraise, E., Zschenderlein, B., Fliedner, T.M. (1997b): Disorders of the back and spine in construction workers. Prevalence and prognostic value for disability. *Spine*, b, 22, 1481-1486.
- Rothenbacher, D., Arndt, V., Fraise, E., Zschenderlein, B., Fliedner, T.M. & Brenner, H. (1998): Early retirement due to permanent disability in relation to smoking in workers of the construction industry. *Journal of occupational and environmental medicine*, 40, 63-68.
- Schuntermann, M.F., Löffler, H.E., Müller-Fahrnow, W. & Braun, R. (1990): Die Rehabilitations-Verlaufsstatistik - Ergebnisse eines Forschungsprojekts zur Epidemiologie in der medizinischen Rehabilitation. *Deutsche Rentenversicherung*, 2, 66-100.

Dr. med. Volker Arndt (M.P.H.): Wissenschaftlicher Mitarbeiter in der Abteilung Epidemiologie des DZFA (seit 2002), davor Referent der Abteilung Epidemiologie im Berufsgenossenschaftlichen Institut für Arbeitssicherheit in Sankt Augustin (1999-2001) sowie wissenschaftlicher Mitarbeiter der Universität Ulm (Abt. Epidemiologie; Institut für Arbeits-, Sozial- und Umweltmedizin; 1993-1999).

Privatdozent Dr. med. Dietrich Rothenbacher (M.P.H.): Stellvertretender Leiter der Abteilung Epidemiologie des DZFA (seit 2000), davor wissenschaftlicher Mitarbeiter der Universität Ulm (Abt. Epidemiologie; Medizinische Universitätsklinik; Institut für Arbeits-, Sozial- und Umweltmedizin; 1993-2000), Arzt für Arbeitsmedizin.

Prof. Dr. med. Hermann Brenner (M.P.H.): Wissenschaftlicher Stiftungsvorstand und Leiter der Abteilung Epidemiologie des DZFA (seit 2000), davor Leiter der Abteilung Epidemiologie an der Universität Ulm (1995-2000), Professur für Öffentliche Gesundheit und Epidemiologie am Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie der Ludwig-Maximilians-Universität München (1993-1994), Leiter der Forschungs- und Geschäftsstelle Epidemiologie an der Universität Ulm (1990-1993).

Die **Abteilung für Epidemiologie am DZFA: (Leitung Prof. Dr. Hermann Brenner)** untersucht, überwiegend im Rahmen großer Längsschnittstudien, die Verbreitung, Risikofaktoren und Chancen der Verhütung von Erkrankungen und funktionellen Einschränkungen im höheren Lebensalter. Zu den krankheitsübergreifenden Schwerpunkten der Abteilung gehören versorgungsepidemiologische und rehabilitationswissenschaftliche Fragestellungen, sowie die Weiterentwicklung epidemiologischer Methoden sowohl im Bereich der deskriptiven als auch der analytischen Epidemiologie.

Das **Deutsche Zentrum für Altersforschung:** (DZFA) an der Ruprecht-Karls-Universität Heidelberg wurde 1995 vom Bund und dem Land Baden-Württemberg als Stiftung des öffentlichen Rechts gegründet und untersucht in einem interdisziplinären Umfeld (Epidemiologie, Entwicklungsforschung, soziale und ökologische Gerontologie, molekulare Altersforschung (im Aufbau)) die vielschichtigen Aspekte der Altersprozesse im Erwachsenenalter (siehe www.dzfa.uni-heidelberg.de).

Nutzung von Reha-Routinedaten für die Forschung

Rainer Kaluscha

(Forschungsinstitut für Rehabilitationsmedizin an der Universität Ulm)

Einleitung

Im Projekt „Reha-Net“, das im gemeinsamen Förderschwerpunkt „Rehabilitationswissenschaften“ des Verbandes Deutscher Rentenversicherer (VDR) und des Bundesministeriums für Bildung und Forschung (BMBF) gefördert wird, hat der rehabilitationswissenschaftliche Forschungsverbund Ulm ein Konzept entwickelt, das die Zusammenarbeit von Kostenträgern, Kliniken und Forschungseinrichtungen durch eine gemeinsame Datenbasis, das sogenannte „Patientenkonto“, ermöglicht (Kaluscha & Jacobi 2000). Dabei wird zunächst der vereinheitlichte Entlassungsbericht des VDR aus der Rehaklinik an den Kostenträger (hier: LVA Baden-Württemberg) übermittelt. Dieser ergänzt sozio-ökonomische Daten aus dem Versichertenkonto, z.B. Informationen über Rentenansprüche und –zahlungen, anonymisiert die Daten und stellt sie dem Forschungsverbund zur Verfügung.

Dabei werden die Blätter 1, 1a, 1b und 2 des vereinheitlichten Entlassungsberichts des VDR sowie der am Forschungsinstitut entwickelte Qualitätssicherungsbogen der LVA Baden-Württemberg mit strukturierten Daten und Freitexten in einer Forschungsdatenbank abgelegt. Diese erlaubt mittels einer Erweiterung des Datenbankmanagementsystems um eine Textauswertungskomponente (Oracle Intermedia) die integrierte Auswertung von strukturierten Daten und Freitexten.

Fragestellung und Ziel

Anhand dieses Datenpools lassen sich vielfältige Fragestellungen untersuchen, z.B. welche Rehabilitanden von einer Reha-Maßnahme oder einer bestimmten Therapie über- bzw. unterdurchschnittlich profitieren oder ob umgekehrt ein Therapiesetting einem anderen vorzuziehen ist. Dies könnte zu einer verbesserten Zuweisung der Rehabilitanden zu unterschiedlichen Rehabilitationsangeboten beitragen.

Ein weiterer Schwerpunkt liegt in der Entwicklung von Prognosemodellen zur Vorhersage des Reha-Outcomes oder von Frühberentung. Diese Prognosemodelle können bei der frühzeitigen Identifikation von Risikogruppen, die einer besonderen Behandlung bedürfen, beitragen.

Angesichts längerer Lebensarbeitszeiten, eines nach hinten verschobenen Renteneintrittsalters sowie des jetzt schon absehbaren Rückgangs an jungen Nachwuchskräften mit dem Eintritt der geburtenschwachen Jahrgänge ins Berufsleben wird die Rehabilitation künftig bei der (Re-)Integration älterer Arbeitnehmer („Reha vor Rente“) sowie der Erhaltung der Selbständigkeit älterer Menschen („Reha vor Pflege“) ein immer wichtigerer Baustein sein, um den Herausforderungen einer alternden Gesellschaft gerecht zu werden. Deshalb muss die Rehabilitationsforschung rechtzeitig zu den richtigen Weichenstellungen in diesem Bereich beitragen.

Datenbedarf

Das oben vorgestellte Modell der Zusammenarbeit zwischen Kostenträgern und Forschungseinrichtungen ließe sich prinzipiell auch bundesweit umsetzen, zumal gemäß § 301 SGB V die Entlassungsberichte vermehrt elektronisch bei den Kostenträgern eingehen. Um bundesweit repräsentative Daten zu erhalten, wäre als Fernziel hier an den Aufbau einer Rehabilitationsdatenbank nach dem Muster des amerikanischen Uniform Data System for Medical Rehabilitation (UDSMR) (UDSMR 2004) zu denken.

Derzeit ist unklar, inwieweit die von der LVA Baden-Württemberg zur Verfügung gestellten Daten für die gesamte Bundesrepublik repräsentativ sind. Da es sich um Daten der Arbeiterrentenversicherung handelt, ist zu vermuten, dass bezüglich sozialer Schicht oder Ausbildung ein Bias besteht. Außerdem mögen auch regionale Besonderheiten, z.B. die unterschiedliche Arbeitsmarktsituation in verschiedenen Bundesländern, einen Einfluss auf die Reintegration von Rehabilitanden in das Berufsleben haben. Daher wäre für unser Projekt in einem ersten Schritt der Zugang zu Vergleichsdaten des VDR eine große Hilfe. Datenbedarf besteht in den Bereichen Frühberentung sowie Zeiten der Arbeitsunfähigkeit vor und nach Reha, wobei etwa nach Diagnosen und Regionen zu unterscheiden wäre. Der Datenbedarf lässt sich jedoch nicht a priori exakt definieren, sondern ist u.a. davon abhängig, welche Parameter sich bei der Modellierung als relevant erweisen. Für diese Parameter wäre dann zu prüfen, ob ihre Verteilung auffällig von den Vergleichswerten abweicht oder ob die Modelle übertragbar sind.

Für einige Fragestellungen wäre auch die Zusammenführung mit Daten aus anderen Quellen interessant. Hier wäre z.B. an mögliche Zusammenhänge zwischen Frühberentungen und Arbeitsmarkt zu denken, bei der eine Zusammenführung entsprechender Daten der Bundesagentur für Arbeit anhand relevanter Merkmale wie Berufsklassenschlüssel und Landkreis bzw. Postleitzahlbezirk mit Daten der Rentenversicherung erforderlich wäre.

Zugangswege

Aus unserer Sicht wären Scientific Use Files der bevorzugte Zugangsweg. Erfahrungsgemäß ist die wissenschaftliche Auswertung ein iterativer Prozess; oft tauchen nach den ersten Auswertungsschritten Detailfragen auf, die weitere Analysen erfordern. Bei Gastwissenschaftler-Arbeitsplätzen wären dann mehrere Reisen erforderlich. Alternativ wäre auch kontrollierte Datenfernverarbeitung denkbar.

Literatur

Kaluscha R. & Jacobi E (2000): Eine Datenbank zur Effektivitätsbeurteilung: Das Datenkonzept des rehabilitationswissenschaftlichen Forschungsverbundes Ulm. In: Schuntermann F. & Schliehe F. (Hg.): 9. Rehabilitationswissenschaftliches Kolloquium. DRV-Schriften 20. Verband Deutscher Rentenversicherungsträger. Frankfurt a.M., 218–219.

Uniform Data System for Medical Rehabilitation: *Mission and Vision*. Online: <http://www.udsmr.org/vision.htm> (21.05.2004).

Rainer Kaluscha hat an der Technischen Universität Braunschweig Informatik mit Anwendungsfach Psychologie studiert. Nach einer Tätigkeit als Software-Entwickler in einem mittelständischen Software-Haus und als wissenschaftlicher Mitarbeiter am Institut für Psychologie der Karl-Franzens-Universität Graz kam er 1999 an das Forschungsinstitut für Rehabilitationsmedizin.

Das **Forschungsinstitut für Rehabilitationsmedizin** ist eine Einrichtung einer Stiftung der Landesversicherungsanstalt Baden-Württemberg, der Waldburg-Zeil-Kliniken GmbH sowie des Universitätsklinikums Ulm. Als An-Institut der Universität Ulm ist es per Kooperationsvertrag universitäre Einrichtung mit allen Rechten und Pflichten in Forschung und Lehre. Das Institut steht einem der bundesweit acht Rehaforschungsverbände im gemeinsamen Förderschwerpunkt „Rehabilitationswissenschaften“ des Verbandes Deutscher Rentenversicherer (VDR) und des Bundesministeriums für Bildung und Forschung (BMBF) vor (siehe www.uni-ulm.de/institute/reha-forschung).

3. Zusammenfassung

Forschungsdatenzentren stehen für eine moderne Dateninfrastruktur wissensbasierter Gesellschaften. Sie eröffnen die Möglichkeit, vorhandene Informationspotenziale durch die Vernetzung bestehender Institutionen besser auszuschöpfen. Der Gedanke, durch die Kooperation zwischen öffentlichen Datenproduzenten und wissenschaftlicher Forschung das Verständnis insbesondere sozialer und ökonomischer Prozesse zu verbessern, lag auch der Gründung der Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) zugrunde:

„Die verschiedenen öffentlichen Datenproduzenten (u.a. amtliche Statistik, Arbeitsverwaltung und Sozialversicherungsträger) und die Disziplinen der Wirtschafts- und Sozialwissenschaften mit ihren großen Dateninstrumenten müssen die wechselseitige Stärkung durch Zusammenarbeit erkennen und sich als gemeinsame Träger der Entwicklung neuer Formen informationeller Infrastruktur zusammenfinden.“ (KVI 2001, 289)

Im Fall der Rentenversicherung ist der Verband Deutscher Rentenversicherungsträger (VDR) gleichsam Träger der statistischen Informationen für die gesamte Rentenversicherung. Dort ist das Forschungsdatenzentrum bis zur Organisationsreform angesiedelt.¹ Für den Aufbau der Service-Einrichtung FDZ-RV ist eine enge Zusammenarbeit mit der Wissenschaft Voraussetzung für eine wechselseitige Stärkung und Verbesserung des Daten- wie Wissenstransfers. Die Ziele des ersten gemeinsam mit der Wissenschaft ausgerichteten Workshops des FDZ-RV standen somit fest: *Information* und *Kooperation*. Diese beiden Attribute spielten und spielen eine zentrale Rolle.

Das Datenangebot der Rentenversicherung

Durch die statistische Auswertung prozessproduzierter Mikrodaten lässt sich der Einfluss von institutionellen Rahmenbedingungen bzw. von politisch motivierten Eingriffen in komplexe gesellschaftliche Entwicklungen adäquat analysieren. Die Gründung des FDZ-RV ist ein wichtiger Meilenstein im Prozess einer sukzessiven Fortentwicklung des Informationswesens der Rentenversicherung. Bislang wurden die Daten überwiegend in diversen amtlichen Meldungen, Broschüren und Statistikbänden gedruckt oder digital verarbeitet, meist in aggregierter Form als Tabellen oder Graphiken. Der VDR führt zusätzlich auf Anfrage Sonderauswertungen durch, die er wiederum in der Regel in Tabellenform bereitstellt.

Aus dem umfangreichen Datenfundus der Rentenversicherung sind die Mikrodaten, die zukünftig anonymisiert und nutzerfreundlich im FDZ-RV angeboten werden, noch zu entwickeln. Die Charakterisierung einzelner Datensätze und Sondererhebungen mit ihrem jeweiligen Merkmalskatalog veranschaulicht, dass umfangreiche statistische Informationen über Renten, Rentner und Versicherte sowie damit in Zusammenhang stehende Bereiche existieren. Das Gros der Datensätze der Rentenversicherung ist prozessproduziert, d.h., sie bauen primär auf den Notwendigkeiten der Verwaltung auf, um insbesondere Anwartschaften der Versicherten berechnen zu können.

¹ Die Organisationsreform in der gesetzlichen Rentenversicherung sieht u.a. mit Wirkung zum 1.10.2005 die Zusammenfassung von BfA und VDR zur „Deutschen Rentenversicherung Bund“ vor. Für einen Überblick über die Organisationsreform vgl. Standfest (2004).

Die Aspekte des Datenschutzes markieren im Prozess der Datenbereitstellung den unumstößlichen Rahmen. Sozialdaten sind besonders sensible Informationen, deren Übermittlung nur zulässig ist, wenn die Voraussetzungen des § 75 SGB X erfüllt sind. Eine Übermittlung der Sozialdaten ist in der Regel nicht ohne die Einwilligung der Betroffenen möglich. Diese im Falle großer Fallzahlen nur schwer zu erfüllende Auflage entfällt, wenn die Daten nicht ohne unverhältnismäßig großen Aufwand bestimmten Personen zugeordnet werden können, d.h. „faktisch anonymisiert“ worden sind. Diesbezüglich darf gemäß § 67 Abs. 8 SGB X der erwartbare Nutzen aus der Missbrauchshandlung nicht den dafür notwendigen Aufwand an Zeit, Kosten und Arbeitskraft übersteigen.

Die Regeln des Datenschutzes erfordern, dass die jeweiligen Mikrodaten der Rentenversicherung für die Weitergabe an die Wissenschaft in ihrem Informationsgehalt reduziert werden. Auch in faktisch anonymisierter Form sind mit Mikrodaten der Rentenversicherung große Fortschritte in der Forschung realisierbar, wenn die hergestellten Einzeldatensätze an die Spezifika der wissenschaftlichen Forschungsgegenstände angepasst werden.² In Bezug auf den Datenschutz ist darauf hinzuweisen, dass es sich bei den prozessproduzierten Mikrodaten der Rentenversicherung in der Regel um Totalerhebungen handelt, deren Fallzahlen regelmäßig die Millionengrenze übersteigen.³ Das bedeutet, dass selbst mit der Ziehung kleiner Stichproben, die ein Instrument zur Anonymisierung darstellen, noch immer Fallzahlen zu erreichen sind, die von repräsentativen Bevölkerungsumfragen nicht annähernd erreicht werden.

Die internen Mikrodatensätze des VDR enthalten keine persönlichen Identifikatoren, wie die Rentenversicherungsnummern der Versicherten; diese sind ausschließlich beim zuständigen Rentenversicherungsträger gespeichert. Dennoch ist durch die Verknüpfung verschiedener im Datensatz vorhandener Informationen oder durch eine Kombination mit anderen Datenquellen – durch Hinzufügung von Zusatzwissen – unter Umständen eine Reidentifizierung einzelner Personen möglich, insbesondere im Fall von Angehörigen zahlenmäßig kleiner Bevölkerungsgruppen.

Diverse Wortbeiträge auf dem Workshop wie auch schriftliche Beiträge im vorliegenden Band verweisen an einigen Stellen auf ein Spannungsverhältnis, das zwischen datenschutzrechtlichen Regelungen, die die Datenweitergabe einschränken, und den häufig umfassenden und differenzierten Datenwünschen der Wissenschaft besteht. Um mögliche Konflikte konstruktiv zu lösen, ist es seitens des FDZ-RV sehr wichtig, den Datenbedarf der verschiedenen Wissenschaftler(-gruppen) sehr genau zu kennen.

Im Vorfeld des Workshops wurde eine Umfrage durchgeführt, um den Datenbedarf unter potenziellen Nutzerinnen und Nutzern des FDZ-RV zu erheben. Die Befunde vermittelten ein klares Bild des Bedarfes der Wissenschaft an Daten der Rentenversicherung. Die größte Nachfrage der interdisziplinär zusammengesetzten Forschergruppen besteht nach Daten der Rentenstatistik (Rentenzugangs-, Rentenwegfalls-, Änderungs- und Rentenbestandsstatistik), insbesondere nach denen der Rentenzugangsstatik. Die Versicherten- und Rehabilitationssta-

² Stefan Kempf und Martin Kukuk zeigen mit ihrer im vorliegenden Band abgedruckten Arbeit auf: „(...), welche Analysemöglichkeiten mit bisher für die Wissenschaft unzugänglichen Individualdaten bestehen.“ (S. 102).

³ Zu den jeweiligen Fallzahlen der Mikrodatensätze der Rentenversicherung siehe Abbildung 12 im Beitrag von Hilmar Luckert (S. 37).

tistiken werden aufgrund ihrer geringeren Nachfrage zu späteren Zeitpunkten vom Forschungsdatenzentrum aufbereitet.

Die vier wichtigsten Wege, über die das Forschungsdatenzentrum der gesetzlichen Rentenversicherung Daten nach entsprechender Beantragung und Erfüllung der datenschutzrechtlichen Voraussetzungen zur Verfügung stellen wird, sind:

- Die Weitergabe von total anonymisierten *Public Use Files*; das sind auf kleinen Stichproben basierende absolut anonymisierte Datensätze, die ohne großen Aufwand zu beziehen sein werden.
- Faktisch anonymisierte *Scientific Use Files* werden der Forschung mit enger vertraglicher Bindung angeboten; diese Daten ermöglichen eine differenzierte Analyse im Hinblick auf bestimmte thematische Schwerpunkte.
- Zudem sollen zunächst in Frankfurt, ab April 2005 in Berlin und in Würzburg einige *Gastwissenschaftler-Arbeitsplätze* eingerichtet werden, über die ein kontrollierter Zugang zu schwach anonymisierten Datensätzen der Rentenversicherung ermöglicht wird.
- Schließlich ist mittelfristig die Option der *kontrollierten Datenfernverarbeitung* anvisiert, mithilfe derer ein nach vertraglicher Regelung kontrolliertes Fernrechnen mit anonymisierten Daten ermöglicht werden soll.

Öffentliche Forschungsdatenzentren

Die vier genannten Zugangswege zu Einzeldatensätzen bilden ebenfalls die Matrix des Serviceangebots der beiden anderen öffentlichen und vom BMBF geförderten Forschungsdatenzentren. Je nach Struktur der Datenbestände und der existierenden Dateninfrastruktur werden die öffentlichen Datenproduzenten in ihren Forschungsdatenzentren jeweils adäquate Schwerpunkte zwischen den genannten Zugriffswegen setzen. Die jeweiligen Erfahrungen, die in den öffentlichen Forschungsdatenzentren gemacht werden, werden untereinander im *Arbeitskreis öffentliche Forschungsdatenzentren (AKöFFFDZ)* ausgetauscht. Diese sind für den Aufbau des vergleichsweise jungen FDZ-RV von besonders hohem Wert.

Die Forschungsdatenzentren der *Statistischen Ämter des Bundes und der Länder* wurden bereits im Jahr 2001 gegründet und haben deshalb Pionierarbeit auf dem Gebiet des geregelten Datenzugangs geleistet. Diese Forschungsdatenzentren haben in ihren verschiedenen Kanälen des Datenzugangs bereits eine rege Nutzung zu verzeichnen: *Public Use Files* beispielsweise des Mikrozensus mit ca. einem Drittel der Variablen stehen zum Download bereit und diverse *Scientific Use Files* sind in Planung.

Die datenschutzrechtlichen Grundlagen zwischen den öffentlichen Forschungsdatenzentren unterscheiden sich: Für die Weitergabe von Daten der amtlichen Statistik müssen sie gemäß Bundesstatistikgesetz zwar ebenso faktisch anonymisiert werden wie die Daten der Sozialversicherung; die Weitergabemodalitäten sind allerdings weniger restriktiv, weil in den Stichproben der Statistischen Ämter weniger personenbezogene Merkmale enthalten sind.

Das FDZ der BA im *Institut für Arbeitsmarkt- und Berufsforschung (IAB)*, das seit April 2004 besteht, verfügt sowohl über Sozialdaten, die dem Sozialdatenschutz unterstehen, als auch über Daten von Betrieben. Insbesondere Betriebsdaten, wie beispielsweise jene aus dem IAB Betriebspanel, müssen sehr streng geschützt, d.h. aufwändig anonymisiert werden. Ein im Rahmen der Weitergabe von Daten des IAB Betriebspanels seit Jahren erprobtes Prozedere sieht vor, dass nach Eingang der Auswertungssyntax bei der Schalterstelle, die sich neuerdings im FDZ befindet, eine Analyse der Originaldaten durchgeführt wird. Bevor die Befunde dieser Analyse an den Nutzer gehen, werden sie einer Datenschutzprüfung unterzogen. Um von vornherein zielgerichtete und fehlerfreie Anfragen zu erhalten, legt das FDZ der BA viel Gewicht auf eine ausführliche Dokumentation der Daten. Außerdem stehen den Interessierten Testdaten (Strukturdatensätze) und Codepläne für das Schreiben der Syntax zur Verfügung.

Betrachtet man die Aufgaben, die alle drei vom BMBF geförderten öffentlichen Forschungsdatenzentren zu bewerkstelligen haben, sind trotz unterschiedlicher Datenlage zahlreiche Gemeinsamkeiten festzuhalten. Bezug nehmend auf den Datenschutz sind dies Verfahren zur Anonymisierung (insbesondere Stichprobenziehung, Vergrößerung von Merkmalen), um die Daten an die Wissenschaft weitergeben zu können. Parallel hierzu ist die Datennutzung zu vereinfachen, indem die jeweilige Erhebungsmethode, Variablenlisten und Merkmalsausprägungen möglichst unter Einbezug des jeweiligen Rechtsstands nutzerfreundlich zu überarbeiten sind. Diese Aufgaben werden im AKöFFFDZ und mit den neuen Servicezentren abgesprochen und aufeinander abgestimmt.

Erfahrungen der Wissenschaft mit Daten der Rentenversicherung

Die Erstellung eines *Scientific Use Files* aus den Daten der Rentenversicherung nach den Anforderungen der Wissenschaft wurde im „Pilotprojekt Versichertenkontenstichprobe 2001 (PP-VKS 2001)“ erprobt. Die Wissenschaftler konnten ihr ökonomisches Modell im Unterschied zu bisherigen Schätzungen, die mit Befragungsdaten durchgeführt wurden, auf Basis der besonders validen, weil prozessproduzierten Daten der Rentenversicherung berechnen.⁴ Kennzeichen der Daten der Rentenversicherung ist insbesondere die Kombination aus hoher Fallzahl mit genauen Datums- und Einkommensangaben, vor allem im Hinblick auf weit zurückliegende Ereignisse. Diese Vorteile gegenüber aus Befragungen gewonnenen Daten mit höherer Fehlerhaftigkeit veranschaulichen den besonderen Nutzen, der durch das Forschungsdatenzentrum der gesetzlichen Rentenversicherung erzielt werden kann.

Eine prominente Sondererhebung, in der Daten der GRV mit Befragungsdaten kombiniert werden, stellt die Studie *Altersvorsorge in Deutschland (AVID)* dar. Die AVID ist für den Bereich der Alterssicherung von großer Bedeutung. Da die Einverständniserklärung jedes Befragten unter der Bedingung die Informationen nicht weiterzugeben eingeholt wurde, ist eine Weitergabe der AVID-Daten an externe Wissenschaftler nicht möglich.

Die Vorerfahrungen aus dem Pilotprojekt Versichertenkontenstichprobe und der langjährigen Zusammenarbeit mit den Projektpartnern der AVID verweisen auf einen zentralen Befund: Der Datenfundus der Rentenversicherung ist einzigartig und es ist ein wichtiges Projekt anonymisierte Mikrodaten im Rahmen des FDZ-RV der Wissenschaft zur Verfügung zu stellen.

⁴ Siehe den Beitrag von Stefan Kempf und Martin Kukuk in diesem Band.

Ermittlung des Datenbedarfs der Wissenschaft durch ein dreistufiges Verfahren

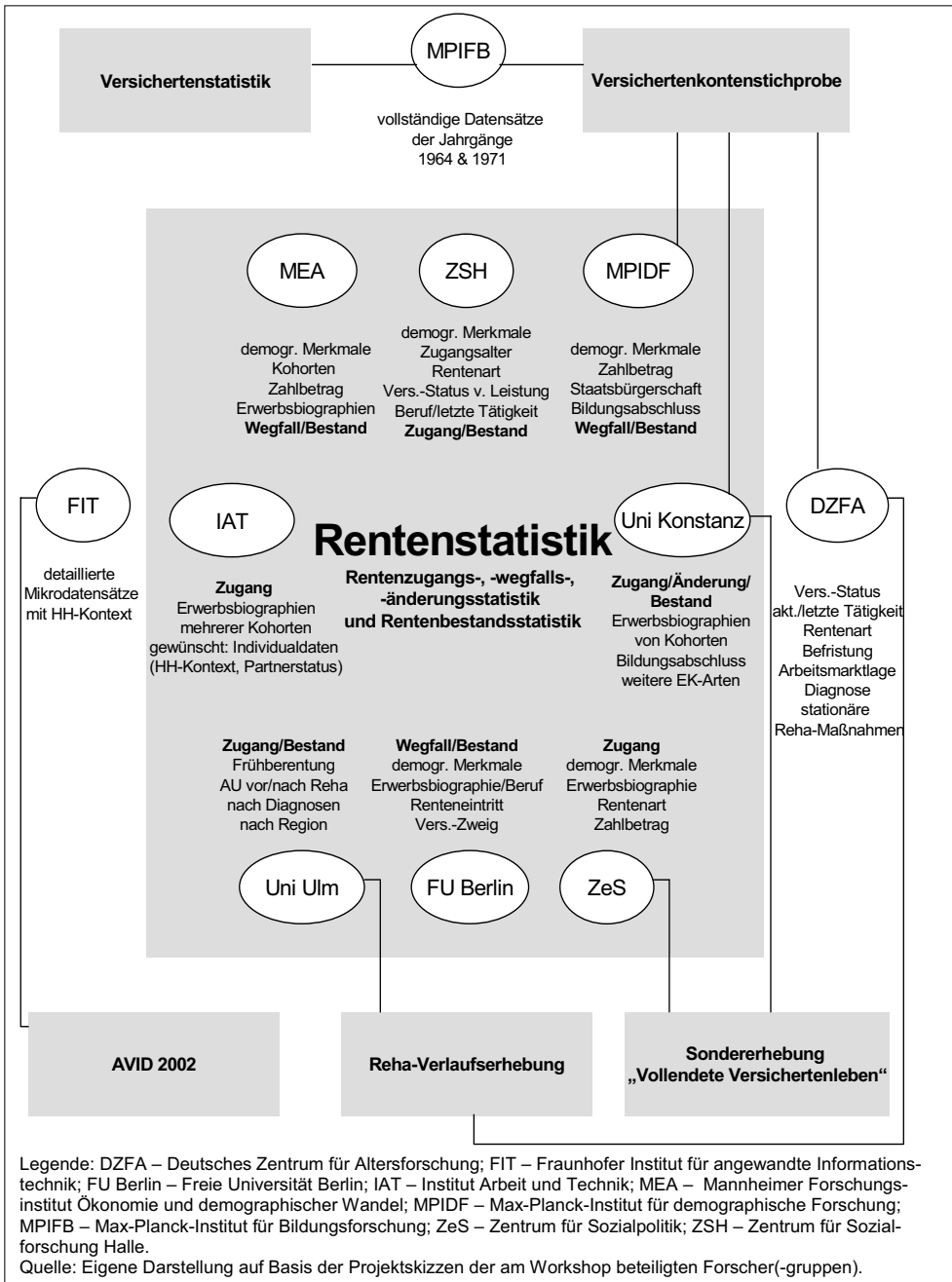
Um sich beim Aufbau des FDZ-RV an den Datenwünschen von interessierten Forscher(-gruppen) orientieren zu können, wurde den Wissenschaftlerinnen und Wissenschaftlern im Rahmen eines dreistufigen Verfahrens die Möglichkeit gegeben, ihre Vorstellungen sowohl über die aufzubereitenden Daten als auch über die präferierten Zugangswege zu diesen darzustellen.

Zunächst wurde in einem ersten Schritt den Wissenschaftlerinnen und Wissenschaftlern mittels eines Fragebogens Gelegenheit gegeben, ihre Forschungsprojekte und dafür benötigte Daten der Rentenversicherung und gewünschte Zugangswege zu diesen darzustellen.⁵ Die anhand des Rücklaufs der Fragebögen ermittelten Forschenden wurden daraufhin zum Workshop „Forschungsrelevante Daten der Rentenversicherung: Aufbau des Forschungsdatenzentrums der gesetzlichen Rentenversicherung (FDZ-RV)“ im Juni 2004 nach Würzburg eingeladen. Dort präsentierten sie auf einer einseitigen Folie kurz und prägnant ihre aktuellen bzw. in Zukunft angestrebten Projekte, insbesondere im Hinblick auf den benötigten Datenbedarf. Alle am Workshop Teilnehmenden wurden so in einem kooperativen Verfahren über die Projekte der Forscherinnen und Forscher informiert, und die spezifischen Anforderungen, die an das FDZ-RV gestellt werden, näher erläutert.

Im Zuge der Herausgabe des vorliegenden Tagungsbandes durch den VDR wurden die verschiedenen Wissenschaftler um die Übermittlung von Projektskizzen gebeten. In diesen sollte durch die Abwägung von Minimal- und Maximalanforderungen aus der Sicht des FDZ-RV eine Einschätzung über den für die Forschungsvorhaben existentiellen Datenbedarf gewonnen werden. Das dreistufige Verfahren hatte zum Ziel, die Forschenden über das Datenangebot der Rentenversicherung und Modalitäten der Datenweitergabe zu informieren, und seitens des FDZ-RV die Aufbereitung der gewünschten Daten im Hinblick auf forschungsrelevante Fragestellungen zu organisieren und zu spezifizieren. Die Beiträge finden sich im vorliegenden Band in Abschnitt 2.2 *Projektskizzen und Datenbedarf der Wissenschaft* und werden in **Abbildung 1** graphisch zusammengefasst.

⁵ Vgl. hierzu den Beitrag von Ralf Himmelreicher im vorliegenden Band und www.fdz-rv.de.

Abbildung 1: Datenbedarf der Wissenschaft



Im Zentrum des Datenbedarfs der Wissenschaft steht die Rentenstatistik, insbesondere die Rentenzugangs- und Rentenwegfallsstatistik sowie die Rentenbestandsstatistik (**vgl. Abbildung 1**). Insofern wurden die durch die Umfrage „Datenbedarf der Wissenschaft“ ermittelten Befunde durch die Projektskizzen bestätigt und insoweit erweitert, als dass es nun möglich ist, sich ein differenziertes Bild über forschungsrelevante Merkmale und deren Ausprägungen der jeweiligen Forschergruppen machen zu können. Die Vielfalt der Datenwünsche aus den verschiedenen wissenschaftlichen Instituten verweist zudem auf die sich interdisziplinär zusammensetzende Wissenschaftsgemeinde, die in Zukunft mit den Daten der Rentenversicherung arbeiten möchte. Ökonomie, Soziologie und Demographie sind die Disziplinen, denen sich die meisten Forschenden nach ihrer Selbsteinschätzung zuordnen. Das starke Interesse verschiedener Forschungseinrichtungen an den Daten der Rentenversicherung sowie disziplinübergreifende Forschungsansätze versprechen dem komplexen Feld der Alterssicherung gerecht zu werden (vgl. Ruland 2001, 130 f.).

Die sich im Mittelpunkt von **Abbildung 1** befindende Rentenzugangsstatistik wird zunächst in der Version 2003 als erster Mikrodatsatz über das FDZ-RV nach vorheriger vertraglicher Regelung für die Wissenschaft in anonymisierter Form beziehbar sein. An der Aufbereitung dieses Datensatzes und der Verbesserung der Dokumentation wird derzeit mit Hochdruck gearbeitet. In einem nächsten Schritt folgen weitere Querschnittdatensätze der Rentenstatistik sowie, um in Zukunft Zeitreihen erstellen zu können, entsprechende Vor(vor)jahresdatensätze. Im Anschluss daran ist geplant, die Mikrodaten aus den Versicherten- und Reha-Statistiken aufzubereiten. Vergleichsweise einfach gestaltet sich die Datenaufbereitung bei den Querschnittsdatsätzen. Wird jedoch die Perspektive des Längsschnitts - wie bei der Versicherungskontenstichprobe, der Reha-Verlaufserhebung und der Sondererhebung Vollendete Versichertenleben möglich - eingenommen, sind noch einige Probleme zu lösen, die im Bereich Datenschutz und Komplexität der sequentiellen Datensätze liegen. Im Rahmen des „Pilotprojektes Versichertenkontenstichprobe 2001“ konnte beispielhaft geklärt werden, dass selbst ein Längsschnittdatsatz der Rentenversicherung im Rahmen der datenschutzrechtlichen Regelungen weitergegeben werden kann. Dieser Datensatz wurde in einer Weise aufbereitet, die sowohl den strengen Kriterien des Datenschutzes als auch den Anforderungen der Wissenschaft nach hinreichend differenziertem Datenmaterial genügt.

Zugangswege

Die Herausforderung, die Ziele ‚hinreichende Differenzierung der Daten‘ und ‚Einhaltung der Bestimmungen des Sozialdatenschutzes‘ zu erreichen, wurde von allen Beteiligten erkannt; es wurde ein gemeinsames Vorgehen bei der Lösung dieser Probleme zugesagt. Anhand der Diskussionen zwischen Wissenschaftlern und Mitarbeitern der Rentenversicherung stellte sich heraus, dass ein möglicher Weg zur Konzipierung forschungsrelevanter Datensätze darin liegt, themenbezogene Scientific Use Files zu erzeugen. Diese enthalten wenige Merkmale, welche jedoch eine tiefere Differenzierung zulassen. Ein solches diskursives Vorgehen, bei dem sich Wissenschaftlergruppen auf die notwendigen Variablen verständigen, kann sowohl eine hinreichend differenzierte Erforschung der jeweiligen Aspekte der Alterssicherung als auch eine Einhaltung der Datenschutzmaßgaben gewährleisten. Gleichzeitig ist ein solches Vorgehen wenig routinisierbar und dadurch sehr arbeitsintensiv. Um statistische Prozeduren, die zukünftig mit *Scientific Use Files* zu berechnen sind, zuvor erproben zu können, sollen den Forschenden sogenannte Strukturdatensätze für Übungszwecke und Testläufe angeboten werden.

Im FDZ-RV wird für alle Wissenschaftler ein transparentes wie gleichberechtigtes Zugangsverfahren entwickelt. Die Diskussion über die Ausgestaltung des Datenzugangs konzentriert sich hierbei auf *Gastwissenschaftler-Arbeitsplätze* und *Scientific Use Files*. Da *Public Use Files* absolut anonymisiert werden, sind sie aufgrund des geringeren Informationsgehaltes für Forschungszwecke weniger geeignet, während solche Daten in Form von z.B. *Campus Files* in der Lehre an Schulen und Hochschulen erkenntnissteigernd eingesetzt werden können. *Kontrollierte Datenfernverarbeitung*, die für Wissenschaftler sehr attraktiv ist, kann wegen technischer und datenschutzrechtlicher Voraussetzungen im FDZ-RV erst mittelfristig als Datenzugangsweg erschlossen werden.

Für einen *Gastwissenschaftler-Arbeitsplatz* im FDZ-RV zeigten sich in der Befragung unter den potenziellen FDZ-RV-Daten Nutzenden knapp 60 % interessiert. Die ersten Gastwissenschaftler werden voraussichtlich noch im laufenden Jahr 2004 an technisch abgeschotteten Arbeitsplätzen mit so genannten stand alone PCs mit der Auswertung von anonymisierten Daten des Rentenzugangs 2003 beginnen können.

Einige Wissenschaftler äußerten sich skeptisch gegenüber dem Modell des *Gastwissenschaftler-Arbeitsplatzes*, weil hierdurch nur Wenigen der Zugriff auf die Mikrodaten der Rentenversicherung ermöglicht würde. Ein anderer Vorbehalt, der in den Projektskizzen der Wissenschaftler auftaucht, betrifft den Aufwand an Zeit und die damit verbundenen Reisekosten. Von Seiten der Rentenversicherung wird keine flächendeckende Versorgung mit *Gastwissenschaftler-Arbeitsplätzen*, wie von den Forschungsdatenzentren der Statistischen Landesämtern geleistet werden. Trotzdem wird z.B. dem Wunsch, auch in Berlin *Gastwissenschaftler-Arbeitsplätze* einzurichten, schon ab Frühjahr 2005 durch den Umzug der Geschäftsstelle des VDR von Frankfurt am Main in die Bundeshauptstadt entsprochen werden. Damit ist der Zugang an einem Ort mit einer hohen Dichte an wissenschaftlichen Forschungsinstitutionen gewährleistet. Der *Gastwissenschaftler-Arbeitsplatz* Standort in Würzburg bleibt erhalten.

Die Funktion der *Gastwissenschaftler-Arbeitsplätze* liegt neben der direkten Unterstützung von Forschungsvorhaben durch die Verfügbarmachung von Mikrodaten der Rentenversicherung und der Betreuung der Forscher unter anderem im Hinblick auf die Spezifika der Datensätze auch darin, die *Scientific Use Files* zu verbessern. Durch die verschiedenen Fragestellungen und Herangehensweisen der externen Forscher sollen die *Scientific Use Files* disziplinäre Besonderheiten möglichst ausgewogen berücksichtigen; zugleich soll die Dokumentation der Datensätze fachübergreifend verbessert werden.

Dokumentation und Datentiefe

Die Datenbestände der Rentenversicherung sind gleichsam das akkumulierte Resultat eines historischen Prozesses. Trotz großer Kontinuität in der Rentenversicherung wurden durch die Politik vielfältige gesetzliche Änderungen vorgenommen, welche die Rentenansprüche verschiedener Geburtskohorten und Versichertengruppen in unterschiedlicher Weise betreffen – all diese Effekte spiegeln sich in den Daten wider. Die komplizierten Feinheiten einzelner Sachverhalte erschweren die statistische Datenverarbeitung durch die Wissenschaft. In diesem Punkt unterscheiden sich Mikrodaten öffentlicher Datenproduzenten stark von solchen Daten, die unmittelbar für Forschungszwecke erhoben wurden. Zwar sind die detaillierten Datenmerkmale in den methodischen Anhängen der Statistikbände dokumentiert; eine differenzierte Analyse,

insbesondere von Längsschnittdaten, erfordert allerdings eine ausgiebige Auseinandersetzung mit den Einzelbestimmungen und ihrer Entstehung. Für eine adäquate Betreuung von Forschungsprojekten, die Analysen auf Basis von Daten der Rentenversicherung erstellen, ist für Wissenschaftler der enge Kontakt zum Forschungsdatenzentrum der Rentenversicherung und später auch zu den neu entstehenden Servicezentren (KVI 2001, 291 f.) zu empfehlen.

Datenbedarf und Datenschutz

Auf Forderungen der Wissenschaft nach möglichst kompletten Mikrodatensätzen, die z.B. eine Zusammenführung von Personendaten zu Haushaltsdaten ermöglichen würde, wurde seitens des FDZ-RV entgegnet, dass sich das Kriterium der faktischen Anonymisierung auch auf denkbare Versuche erstreckt, mehrere veröffentlichte Themendatenfiles zu vereinen (*merging*). In der sich auf dem Workshop entfaltenden Diskussion wurde von Seiten der Forscher insistiert, dass die Daten der Rentenversicherung in anonymisierter Form wichtige Informationen, wie beispielsweise über die Einkommenssituation von Ehepartnern, systematisch im Dunkeln ließen. Tatsächlich sind derartige Daten zum Haushaltskontext und zu den gesamten Haushaltseinkünften nicht in den individuellen Versicherungskonten, die auf sozialversicherungspflichtige Einkünfte abstellen, gespeichert und können somit auch nicht in den Rentenstatistiken enthalten sein.⁶ Es ist die gesetzliche Pflicht der Halter der Sozialdaten, die persönliche Integrität der Versicherten zu wahren. Der Gesetzgeber hat diese Vorkehrungen zum Schutze seiner Bürger getroffen und das FDZ-RV wird sich strikt an die Regeln des Datenschutzes halten.

Gleichzeitig sehen die Mitarbeiter und Mitarbeiterinnen des FDZ-RV ihre Aufgabe in der Unterstützung der wissenschaftlichen Forschung und halten es für eine der zentralen Funktionen, den Datenbedarf der Wissenschaft an den Stellen, an denen eine Weitergabe der Daten nicht möglich ist, entsprechend zu dokumentieren. In diesem Zusammenhang signalisierte die Datenschutzbeauftragte des VDR, dass eine Abmilderung des rigiden Charakters des § 75 SGB X durch den Gesetzgeber angezeigt wäre, wenn sich erweise, „dass der Schutz der Betroffenen durch weniger einschneidende Maßnahmen gleichermaßen sichergestellt werden kann“ (siehe den Beitrag von Claudia Heese im vorliegenden Band). Eine solche Initiative könnte aufgrund der Erfahrungen des Forschungsdatenzentrums der Rentenversicherung und deren Evaluation angestoßen werden.

In der abschließenden Diskussion des Workshops wurden die zentralen Aspekte bei der Ausgestaltung der Datenwege des Forschungsdatenzentrums kontrovers zur Sprache gebracht. Ein wichtiges Thema war die Frage nach möglichen vertraglichen Alternativen zum aufwändigen und mitunter langwierigen Anonymisieren der Sozialdaten. Für die Wissenschaft wäre eine solche Lösung mit rein kontraktuellen Sanktionsmechanismen deshalb attraktiv, weil sie einen schnelleren Datenzugang ermöglichen würde. Es wurde von Forscherseite argumentiert, dass wertvolle Informationen der Forschung zugänglich seien, die durch die Informationsreduktion des Anonymisierens verloren gingen. Die momentane Rechtslage lässt eine rein vertragliche Übereinkunft zur Weitergabe von Sozialdaten nicht zu, weil der Missbrauch hier zwar zu ahnden, nicht aber zu verhindern ist. Vielmehr ist ein Nutzervertrag in jedem Fall notwendig, um die Weitergabe faktisch anonymisierter Einzeldatensätze an Bedingungen knüpfen zu können. Eine Auflage ist beispielsweise die Löschung der Daten nach dem Projektende.

⁶ Zu den in den Mikrodatensätzen der Rentenversicherung enthaltenen wichtigen Merkmalen siehe den Beitrag von Hilmar Luckert in diesem Band.

Abschließende Bemerkungen

Insgesamt betrachtet zeigt das breite Interesse verschiedenster Forschungsinstitute an den Daten der Rentenversicherung und die Teilnahme der jeweiligen Forscher(-gruppen) am ersten Workshop des FDZ-RV, dass der Aufbau eines Forschungsdatenzentrums beim VDR eine wichtige Erweiterung der wissenschaftlichen Forschungsmöglichkeiten wie des statistischen Projektwesens der Rentenversicherung selbst darstellt. Die in diesem Band abgedruckten Projektskizzen verweisen auf Forschungsfelder, auf denen gestützt auf Daten der Rentenversicherung neue Erkenntnisse zu erwarten sind. Eine Verbesserung der empirischen Forschungsmöglichkeiten im Feld der Altersvorsorge erscheint vor dem Hintergrund einer zunehmenden Individualisierung derselben zwingend erforderlich (vgl. Himmelreicher 2003). Bei den ersten Schritten des Aufbaus des FDZ-RV hat sich jedoch auch gezeigt, dass noch zahlreiche weitere Schritte zu gehen sind, bis der Übergang von einem Pilotprojekt hin zu einem benutzerfreundlichen Datenanbieter für wissenschaftliche Forschung geschafft ist.

Seitens des FDZ-RV soll der Aufbau und die Datenvergabe nachvollziehbar und transparent gestaltet werden. Deshalb wird ein zweiter Workshop zum Aufbau des Forschungsdatenzentrums der Rentenversicherung vom 27. bis 29. Juni 2005 beim VDR in Würzburg ausgerichtet. Getreu den Attributen *Information* und *Kooperation* soll auch über diesen Workshop in einem Tagungsband berichtet werden.

Literatur

- Himmelreicher, R.* (2003): Eigenverantwortung und Eigenvorsorge in der Alterssicherung: Das Beispiel Riester Rente. In: J. Allmendinger (Hrsg.): Entstaatlichung und soziale Sicherheit, Verhandlungen des 31. Kongresses der Deutschen Gesellschaft für Soziologie in Leipzig 2002 (Band II CD), Opladen: Leske + Budrich.
- KVI, Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (2001): Wege zu einer besseren informationellen Infrastruktur. Gutachten der vom Bundesministerium für Bildung und Forschung eingesetzten Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. Nomos: Baden-Baden.
- Ruland, F.* (2001): Forschungsnetzwerk Alterssicherung – verstärkte Kooperation zwischen Rentenversicherung und Wissenschaft im Bereich der Alterssicherung. Deutsche Rentenversicherung, 3-4, 129-136.
- Standfest, E.* (2004): Die Organisationsreform in der gesetzlichen Rentenversicherung. Deutsche Rentenversicherung, 8, 459-467.

Teilnehmerliste

Workshop „Das Forschungsdatenzentrum der gesetzlichen Rentenversicherung (FDZ-RV) im Aufbau“ vom 28. bis 29. Juni 2004 in Würzburg

1.	Arndt, Dr.	Volker	Deutsches Zentrum für Altersforschung Heidelberg
2.	Böttcher	Sabine	Zentrum für Sozialforschung Halle Halle
3.	Bütefisch	Thomas	Verband Deutscher Rentenversicherungsträger Frankfurt am Main
4.	Büttner	Renate	Institut Arbeit und Technik Gelsenkirchen
5.	Faik, Dr.	Jürgen	Verband Deutscher Rentenversicherungsträger Frankfurt am Main
6.	Grub	Martin	Fraunhofer Institut für Angewandte Informationstechnik , St. Augustin
7.	Habla	Heike	Statistisches Bundesamt Wiesbaden
8.	Heese	Claudia	Verband Deutscher Rentenversicherungsträger Frankfurt am Main
9.	Heien, Dr.	Thorsten	tns Infratest Sozialforschung München
10.	Heineck	Martin	Universität Konstanz Konstanz
11.	Himmelreicher, Dr.	Ralf	Verband Deutscher Rentenversicherungsträger Frankfurt am Main
12.	Hupfeld	Stefan	Universität Konstanz Konstanz
13.	Kaldybajewa, Dr.	Kalamkas	Bundesversicherungsanstalt für Angestellte Berlin
14.	Kaluscha	Rainer	Forschungsinstitut für Rehabilitationsmedizin der Universität Ulm , Ulm
15.	Kempf	Stefan	Universität Würzburg Würzburg
16.	Kirschner, Dr.	Walther	Kurpark-Klinik Bad Nauheim
17.	Klosterhuis, Dr.	Here	Bundesversicherungsanstalt für Angestellte Berlin
18.	Kohlmann, Dr.	Annette	Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit – Forschungsdatenzentrum , Nürnberg
19.	Konold	Michael	Forschungsdatenzentrum der Statistischen Landesämter , Düsseldorf

20.	Kukuk, Prof. Dr.	Martin	Universität Würzburg Würzburg
21.	Luckert	Hilmar	Verband Deutscher Rentenversicherungsträger Würzburg
22.	Padberg, Dr.	Frank	Hessisches Statistisches Landesamt Wiesbaden
23.	Pantzke	Boris	Verband Deutscher Rentenversicherungsträger Frankfurt am Main
24.	Poulev	Kim	Verband Deutscher Rentenversicherungsträger Frankfurt am Main
25.	Radl	Jonas	Verband Deutscher Rentenversicherungsträger Frankfurt am Main
26.	Rasner	Anika	Max-Planck Institut für demografische Forschung, Rostock
27.	Rehfeld	Uwe	Verband Deutscher Rentenversicherungsträger Frankfurt am Main
28.	Roß	Elisabeth	Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit – Forschungsda- tenzentrum, Nürnberg
29.	Rütters	Christine	Verband Deutscher Rentenversicherungsträger Büro, Berlin
30.	Scholz, Dr.	Rembrandt	Max-Planck Institut für demografische Forschung, Rostock
31.	Seibert	Holger	Max-Planck Institut für Bildungsforschung Berlin
32.	Sommer	Mathias	MEA Universität Mannheim Mannheim
33.	Unger, Dr.	Rainer	ZeS Universität Bremen Bremen
34.	Viebrok, Dr.	Holger	ZeS Universität Bremen Bremen
35.	von Gaudecker	Hans-Martin	MEA Universität Mannheim Mannheim
36.	Willert	Michaela	Freie Universität Berlin Berlin
38.	Zwick	Markus	Forschungsdatenzentrum des Statistischen Bundesamtes, Wiesbaden