

SHARE-RV: Eine Datengrundlage für Analysen zu Alterssicherung, Gesundheit und Familie auf der Basis des Survey of Health, Ageing and Retirement in Europe und der Daten der Deutschen Rentenversicherung

Tatjana Mika, Christin Czaplicki

In einem Kooperationsprojekt der Deutschen Rentenversicherung Bund mit dem Mannheimer Institut Ökonomie und Demographischer Wandel (Universität Mannheim) wird ein Forschungsdatensatz erstellt, in dem biographische Informationen aus dem Versicherungskonto mit Interviews des wissenschaftlich erprobten Survey of Health, Ageing and Retirement in Europe (SHARE) verbunden werden. Die Daten bieten damit die Möglichkeit, detaillierte Informationen zu den Erwerbsbiographien im Zusammenhang mit der Familien- und Haushaltssituation sowie dem Gesundheitszustand der Befragten zu betrachten.

1. Ziele des Projektes SHARE-RV¹

Das Forschungsdatenzentrum der Deutschen Rentenversicherung (FDZ-RV) stellt bereits seit 2004 statistisch aufbereitete Informationen aus den Verwaltungsdaten zur wissenschaftlichen Nutzung zur Verfügung. Mit dem Projekt SHARE-RV soll das Datenangebot erweitert werden, indem Befragungsdaten aus SHARE² mit Informationen aus dem Versicherungsverlauf kombiniert werden.

Durch diese Verknüpfung entstehen wesentliche Vorteile. Während die Prozessdaten der Rentenversicherung (RV) um sozioökonomische und gesundheitliche Hintergrundinformationen der Versicherten erweitert werden, können die Befragungsdaten des SHARE von der Qualität der genau erfassten sozialversicherungspflichtigen Erwerbsbiographien und Rentenanwartschaften profitieren. Wie im Datenverknüpfungsjahr BASiD³ des FDZ-RV werden beide Datenquellen unter Verwendung der Sozialversicherungsnummer verknüpft.

2. SHARE

SHARE ist ein multidisziplinärer, international vergleichender Paneldatensatz, der die Lebenssituation älterer Menschen in Europa abbildet. Dabei enthält die repräsentativ angelegte Datenbasis Informationen zur gesundheitlichen, familiären und wirtschaftlichen Situation von Über-50-Jährigen in Privathaushalten⁴, wobei die Inhalte im Folgenden näher erläutert werden sollen. Die erste Erhebung von SHARE erfolgte 2004/2005 in elf europäischen Ländern (Belgien, Dänemark, Deutschland, Frankreich, Griechenland,

Niederlande, Italien, Österreich, Schweden, Schweiz und Spanien). Die zweite Erhebungswelle wurde 2006/2007 mit über 30 000 Befragten in insgesamt 14 europäischen Ländern und Israel durchgeführt. Die aktuelle dritte Befragungswelle von 2008/2009 enthält schließlich Informationen zu den retrospektiven Lebensgeschichten der Befragten (SHARELIFE), wobei hierfür ein neu entworfener lebensverlaufsorientierter Fragenkatalog zum Einsatz kam. Durch die regelmäßige Wiederholung der Befragung der über 30 000 Studienteilnehmer können Veränderungen der wirtschaftlichen, gesundheitlichen und sozialen Lage älterer Menschen in Europa beobachtet

Tatjana Mika ist Mitarbeiterin im Forschungsdatenzentrum der Rentenversicherung, Christin Czaplicki ist Mitarbeiterin im Projekt SHARE-RV – Geschäftsbereich Forschung und Entwicklung der Deutschen Rentenversicherung Bund.

¹ Die Verknüpfung der Befragungsdaten mit den administrativen Daten der Rentenversicherung wird gefördert durch die Volkswagen Stiftung (Az.: II/83 176). Die Erhebung von SHARE wird durch das 5. Rahmenprogramm der Europäischen Kommission (Projekt QLK6-CT-2001-00360 in dem Themenbereich Lebensqualität), durch das 6. Rahmenprogramm (SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5-CT-2005-028857 und SHARELIFE, CIT4-CT-2006-028812) und durch das 7. Rahmenprogramm (SHARE-PREP, 211909 und SHARE-LEAP, 227822) gefördert. Weitere Förderungen erfolgen durch das U.S. National Institute on Ageing (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4533-01 und OGHA 04-064, IAG BSR06-11, OGHA04-064) sowie durch weitere nationale Stellen (vgl. www.share-project.org für eine vollständige Liste der fördernden Institutionen).

² Weitere Informationen auf www.share-project.org.

³ Vgl. Budzak, Steppich (2010): BASiD: Biographiedaten ausgewählter Sozialversicherungsträger in Deutschland. In: RVaktuell 3/2010, Jahrgang 57: S.111-116.

⁴ Vgl. Börsch-Supan, Hank, Jürges, Schröder (Hrsg.) (2009): 50plus in Deutschland und Europa – Ergebnisse des Survey of Health, Ageing and Retirement in Europe. Wiesbaden: VS Verlag.

werden. Auf diese Weise wird es möglich, Einflussfaktoren einschneidender Ereignisse in Verbindung mit den institutionellen Rahmenbedingungen (z. B. Gesundheits- und Rentensystem) über den Lebensverlauf zu identifizieren und zu untersuchen.

2.1 Gesundheit

Gesundheit hat einen wesentlichen Einfluss auf die Lebensqualität einer Person und bestimmt zudem die Teilhabe an vielen Lebensbereichen. SHARE bietet in diesem Zusammenhang umfangreiche Möglichkeiten zur Untersuchung altersspezifischer Gesundheitsprobleme und der daraus entstehenden Konsequenzen für einzelne Lebensbereiche wie Beruf oder Familie. So stehen z. B. subjektive und objektive Indikatoren zur Erfassung der physischen Gesundheit⁵ zur Verfügung: subjektiv erfragtes Gesundheitsempfinden⁶, berichtete Krankheiten und Symptome, Einschränkungen in physischen Funktionen und Einschränkungen in Aktivitäten des alltäglichen Lebens. Während die subjektiven Maße auf Berichten der Befragten beruhen, werden die objektiven Maße durch verschiedene Messungen von Körperfunktionen wie Gehgeschwindigkeit, Greifkraft, Lungenvolumen und den Stuhl-Aufsteh-Test - erhoben.

Mit diesen detaillierten Informationen zum Gesundheitsstatus einer Person bietet SHARE die Chance, die Messgenauigkeit subjektiver und objektiver Indikatoren miteinander zu vergleichen⁷. Neben den

Indikatoren zur Bestimmung der physischen Gesundheit wird in SHARE auch die mentale Verfassung der Studienteilnehmer erhoben. Hierzu wird die Euro-D-Skala für Altersdepression verwendet, die dazu dient, mithilfe von Symptomen das Risiko einer Depression einzuschätzen. Darüber hinaus wird aus den Ergebnissen verschiedener Konzentrationsübungen und rechnerischer Aufgaben ein numerischer Wert errechnet, der die kognitive Leistungsfähigkeit der Befragten widerspiegelt.

Während mit den ersten beiden Erhebungswellen überwiegend Faktoren wie Bildung und Einkommen im Zusammenhang mit sozialen Unterschieden im Gesundheitszustand thematisiert wurden⁸, wird der Fokus der dritten Erhebung in SHARELIFE zusätzlich auf die gesundheitliche Versorgung während der Kindheit sowie Belastungen, denen Studienteilnehmer während der Kindheit ausgesetzt waren, gelegt. Darüber hinaus werden die Arbeitsbedingungen für den gegenwärtigen und die vergangenen Arbeitsplätze erfragt. Damit wird der Umfang der verfügbaren Indikatoren, die sich auf die Gesundheit auswirken könnten, erheblich erweitert.

2.2 Familie und soziale Netzwerke

Die Familie ist im gesamten Lebensverlauf ein zentrales Element der emotionalen Bedürfnisbefriedigung sowie der sozialen und ökonomischen Absicherung. Auch bei älteren Personen finden die Interaktionen innerhalb der Familie dabei nicht nur auf der Paarebene statt, sondern über die Grenzen einer Generation hinweg. Um diese komplexen Beziehungsmuster untersuchen zu können, werden Informationen benötigt, die den Austausch zwischen den Generationen wiedergeben. SHARE bietet hierfür einen Fragenkatalog, der zahlreiche Analysemöglichkeiten familiärer Beziehungen ermöglicht. So werden durch die Befragung beider Partner eines Haushaltes zunächst Interaktionsprozesse untersuchbar, die zwischen ihnen stattfinden⁹. Dabei sind insbesondere die Partnerschaftshistorien, die ab der dritten Welle des SHARE erhoben werden, von wissenschaftlicher Bedeutung, da diese u. a. Hinweise auf die Stabilität und die Dauer von Beziehungen geben.

Darüber hinaus werden die Austauschbeziehungen zwischen Eltern und Kindern detailliert erfasst, wodurch die familiäre Solidarität auch zwischen zwei Generationen betrachtet werden kann¹⁰. Insbesondere die Wohnortnähe, die Kontakthäufigkeit oder soziale Transfers, z. B. in Form von Unterstützung bei Haushalts- oder Pflgetätigkeiten¹¹, sind gute Indikatoren zur Charakterisierung dieser Beziehungsmuster. Die Erhebung von Informationen zur Beziehung von Großeltern und Enkelkindern ermöglicht schließlich die Analyse der Familienbeziehungen über drei Generationen, wobei insbesondere die Betreuung der Enkel im Vordergrund steht¹².

Neben den familiären Beziehungen enthält SHARE Informationen zu den sozialen Netzwerken der Stu-

⁵ Mackenbach, Avendano, Andersen-Ranberg, Aro (2005): Physical Health. In: Börsch-Supan, Brügl, Jürges, Mackenbach, Siegrist, Weber (2005): Health, Ageing and Retirement in Europe – First Results from SHARE. Mannheim: MEA.

⁶ Um Unterschiede zwischen verschiedenen Messinstrumenten erfassen zu können, wird im Interviewverlauf der ersten Erhebungswelle neben der europäischen Skala der subjektiv empfundenen Gesundheit zusätzlich die amerikanische Skala verwendet, wobei sich die Skalen in der Anzahl und der Formulierung der Antwortkategorien unterscheiden.

⁷ Jürges (2007): True health vs. response styles: Exploring cross-country differences in self-reported health. Health Economics 16 (2): S.163–178.

⁸ Jürges (2008): Health inequalities by education, income and wealth: A comparison of 11 European countries and the US. Applied Economics Letters 1 (5).

⁹ Hank, Jürges (2007): Gender and the division of household labor in older couples: A European perspective, Journal of Family Issues, 28 (3), S.399–421.

¹⁰ Brandt, Deindl, Haberkern, Szydlik (2008): Reziprozität zwischen erwachsenen Generationen: Familiäre Transfers im Lebensverlauf. Zeitschrift für Gerontologie und Geriatrie 41 (5): S.374–381.

¹¹ Brandt, Szydlik (2008): Soziale Dienste und Hilfe zwischen Generationen in Europa. Social services and help between generations in Europe. Zeitschrift für Soziologie 37 (4): S.301–320.

¹² Hank, Buber (2010): Die Betreuung der Enkelkinder durch ihre Großeltern: Wohlfahrtsstaatliche Einflüsse auf soziale Transfers zwischen den Generationen. In: Ette, Ruckdeschel, Unger (eds.): Potentiale intergenerationaler Beziehungen: Chancen und Herausforderungen für die Gestaltung des demografischen Wandels, S.321–371. Würzburg: Ergon Verlag.

dienteilnehmer. So werden neben ehrenamtlichen Tätigkeiten in Vereinen auch Betreuungs- oder Unterstützungstätigkeiten, die für bzw. von Freunde(n) oder Nachbarn ausgeübt wurden, erfragt¹³. Dabei wird nicht nur die reine Tätigkeit erfasst, sondern auch die Häufigkeit der aktiven Ausübung. Mit diesen Informationen können u. a. zusätzlich zu den familiären Beziehungsmustern Analysen zum produktiven Altern und zu dessen Einflussfaktoren auch im internationalen Vergleich durchgeführt werden.

2.3 Wirtschaftliche Situation, Einkommen und Vermögen

Traditionell wurde die individuelle ökonomische Absicherung im Alter der Familie zugeschrieben. Im Zuge der Ausdifferenzierung der Familienaufgaben und der Einführung der Sozialversicherungssysteme gewann jedoch die individuelle finanzielle Unabhängigkeit an Bedeutung. SHARE bietet diesbezüglich ein breites Spektrum an Indikatoren, die die wirtschaftliche Lage der Studienteilnehmer erfassen.

Das Einkommen wird hierbei nicht nur auf der Individualebene erfragt, sondern für alle Haushaltsmitglieder. Dazu gehören Einkommen aus Erwerbstätigkeit, Renten oder auch Mieteinnahmen¹⁴. Auf diese Weise kann ein Gesamteinkommen auf Haushaltsebene erstellt werden, das alle Einkünfte aus verschiedenen Quellen beinhaltet. Durch die differenzierte Erfassung der verschiedenen Einkommensarten wird die Vielfalt finanzieller Ressourcen der älteren Bevölkerung wiedergegeben. Neben der Einkommenssituation der Befragten und ihrer Haushaltsmitglieder liefert SHARE außerdem Informationen zum Beschäftigungsstatus, wodurch die Daten es erlauben, nicht nur Personen zu betrachten, die bereits das Rentenalter erreicht haben, sondern auch Personen, denen der Übergang ins Rentenalter noch bevorsteht.

Durch die zusätzliche Erfassung von Haushaltsausgaben, z. B. für Lebensmittel oder Telefon, werden Analysen möglich, die die Konsumfähigkeit im Alter betreffen, d. h., inwiefern die Ressourcen zur Aufrechterhaltung des Lebensstandards beitragen¹⁵. Über die Einkommensinformationen hinaus enthält SHARE Informationen zu finanziellen Transfers¹⁶, die zur Unterstützung von Familienmitgliedern dienen oder in Form von Schenkungen oder Erbschaften erfolgen¹⁷.

3. Forschungsdaten aus den Versicherungskonten

Die Befragten von SHARE repräsentieren die ältere Bevölkerung Deutschlands ab 50 Jahren. Es sind sowohl aktiv oder passiv Versicherte als auch Rentner befragt worden. Daher stellte sich die Frage, welches erprobte Format aus der Statistik der RV als Ergänzung für diese Befragung sinnvoll ist.

Im Hinblick auf die Versicherungsbiographie erscheint die Versicherungskontenstichprobe als passendes

Datensatzformat, weil die Längsschnittdaten der Versicherungsbiographie eine besonders gelungene Ergänzung zur retrospektiven Befragung von SHARELIFE sind. Im Test erwies sich, dass selbst für bereits mehrere Jahre vor der Befragung verrentete Befragte noch eine Meldung im Format der Versicherungskontenstichprobe erstellt werden konnte.

3.1 Konzept der Versicherungskontenstichprobe

Die Versicherungskontenstichprobe ist ein jährlich erstellter Datensatz, der Auskunft über die erreichten Anwartschaften der Versicherten gibt¹⁸. Der Datensatz besteht aus einem fixen Teil, der demographische Angaben wie Geschlecht, Geburtsjahr und Bundesland enthält. Die rentenbezogenen Variablen des fixen Teils weisen die erreichte Anwartschaft in der RV zum Zeitpunkt des Stichtags, also des letzten Tages des Berichtsjahres, aus. Weil die Mehrzahl der Versicherten zum Zeitpunkt des Stichtags noch keine Rente bezieht, wird für den Zweck der Statistikmeldung der Wert einer hypothetischen Erwerbsminderungsrente zu diesem Zeitpunkt errechnet. Hierfür wird die Sachbearbeitung beim Rentenversicherungsträger aufgefordert, das Konto zu klären und eine Anwartschaft unter der Annahme zu berechnen, dass die Person ab dem Tag nach dem Stichtag erwerbsgemindert wäre. Damit wird der zum Stichtag geltende Rechtsstand zugrunde gelegt, so dass alle Biographien der unterschiedlichen Geburtskohorten unter dem gleichen Recht vergleichbar werden.

Zu bedenken ist allerdings, dass damit auch bereits vor längerer Zeit unter einem anderen Rechtsstand verrentete Personen ihre Biographie nach aktuell geltendem Recht bewertet bekommen. Die ausgewiesenen Entgeltpunktbeträge können damit unter Umständen von der tatsächlich gezahlten Rente abweichen. Je länger die tatsächliche Verrentung zurückliegt, umso häufiger ist aufgrund der in der

¹³ Hank, Stuck (2008): Volunteer work, informal help and care among the 50+ in Europe: Further evidence for „linked“ productive activities at older ages. *Social Science Research* 37 (4): S. 1280–1291.

¹⁴ Paccagnella, Weber (2005): Household Income. In: Börsch-Supan et al., a. a. O., S. 296–301.

¹⁵ Browning, Madsen (2005): Consumption. In: Börsch-Supan, et al., a. a. O., S. 318–324.

¹⁶ Attias-Donfut, Ogg, Wolff (2005): Financial transfers. In: Börsch-Supan et al., a. a. O., S. 179–185.

¹⁷ Brandt, Deindl, Haberkern, Szydlik (2009): Geben und Nehmen zwischen Generationen: Der Austausch von Geld und Zeit in Europa. In: Börsch-Supan, et al., 50+ in Deutschland und Europa – Ergebnisse des Survey of Health, Ageing and Retirement in Europe [Series „Alter(n) und Gesellschaft“], S. 95–114. Wiesbaden: VS Verlag für Sozialwissenschaften.

¹⁸ S. Tabellenband „Anwartschaften“, der jährlich von der Deutschen Rentenversicherung herausgegeben wird. Die Ergebnisse sind auch im Internet unter www.forschung-rentenversicherung.de/einsehbar.

Zwischenzeit wirksam gewordenen Rechtsreformen eine Abweichung zu erwarten.

3.2 Biographiedaten aus den Versicherungskonten

Der für Forschungszwecke besonders interessante Teil der Versicherungskontenstichprobe ist die Meldung der gesamten versicherungspflichtigen Biographie in einem abgestimmten Konzept von sog. Beitragszeiten (BYAT) und Beitragsbesonderheiten (BYATSO), die unterschiedliche Beitrags- und sonstige Versicherungszeiten wiedergeben. Für sozialwissenschaftliche Zwecke werden diese vielen unterschiedlichen Biographiezeiten in den Meldungen im FDZ-RV in einen übersichtlichen Katalog von sozialen Erwerbssituationen (SES) übersetzt¹⁹. Dieses Konzept hat den Vorteil, dass es für längere Zeiträume gültige Zustände umfasst und daher auch die Vergleichbarkeit unterschiedlicher Geburtskohorten gewährleistet ist.

Der besondere Reiz der Längsschnittdaten, die aus dem Versicherungskonto generiert werden, ist die Vollständigkeit der Erfassung. Mehrfach durchgeführte Panel-Befragungen, die ebenfalls zu Längsschnitt-Biographiedaten führen, kämpfen mit der zunehmenden Nichtteilnahme der ursprünglich kontaktierten Personen im Laufe der Zeit. Diese sog. Panel-Mortalität hat innerhalb weniger Jahre nicht nur eine starke Reduktion der Fälle zur Folge, die für empirische Analysen zur Verfügung stehen, sondern ist zudem nicht zufällig, sondern häuft sich in bestimmten Lebenssituationen wie z.B. Scheidung, Umzug aufgrund von Arbeitsplatzwechsel, langwieriger Erkrankungen oder aufgrund von Veränderungen der allgemeinen Rahmenbedingungen. Wird dagegen das Leben retrospektiv z.B. im Rentenalter erfragt, treten Erinnerungslücken auf, die vor allem länger zurück-

liegende Zeiträume betreffen. Im Gegensatz hierzu sind die aus den Konten gewonnenen Informationen über das zurückliegende Erwerbsleben verzerrungsfrei²⁰.

Die wichtigsten Informationen aus den Versicherungskonten sind die Zeiten der sozialversicherungspflichtigen Beschäftigung und die in diesen Zeiträumen erworbenen Entgeltpunkte aus Erwerbstätigkeit. Entgeltpunkte haben den Vorzug, dass sie ohne Berücksichtigung von Inflation über die Zeit vergleichbar sind, weil sie das jeweilige individuelle Einkommen mit dem Durchschnittsentgelt des gleichen Jahres vergleichen. Damit lässt sich ohne größere Mühe der Einkommensverlauf über die gesamte Erwerbskarriere darstellen²¹. Von den weiteren sozialen Erwerbssituationen ist u. a. die Arbeitslosigkeit besonders relevant für die Forschung mit Daten aus Versicherungskonten²², da diese Zeiten in Befragungen häufig fehlerhaft, nämlich nur verkürzt, in Erinnerung bleiben.

4. Konzept der Verknüpfung der Daten

Die Verknüpfung von Befragungs- und prozessproduzierten Daten dient der Erstellung von Forschungsdaten, indem Daten aus administrativen Prozessen, z.B. den Verlauf der sozialversicherungspflichtigen Beschäftigung in den Konten der gesetzlichen RV, und die nur von den Personen selbst ermittelbaren Informationen wie z.B. das vorhandene Vermögen im Haushalt, verbinden.

Dabei sind zwei Wege gangbar: entweder kann eine Stichprobe aus den Verwaltungsdaten oder eine bevölkerungsrepräsentative Befragung die Ausgangsbasis für die Verbindung sein. Welcher Weg sinnvoller ist, hängt von den beabsichtigten Forschungsfragen ab, die mit dem Datensatz bearbeitet werden sollen. Soll vorrangig eine ausreichende Anzahl von Personen im Datensatz enthalten sein, die eine bestimmte Maßnahme oder Geldleistung bekommen haben, dann sollte der Datensatz von einer Stichprobe der administrativen Daten ausgehen²³. Ist beabsichtigt, die soziale und gesundheitliche Lage der gesamten Bevölkerung zu untersuchen und dabei auch diejenigen, die in der gesetzlichen RV abgesichert sind, zu vergleichen, dann sollte von einer bevölkerungsrepräsentativen Stichprobe ausgegangen werden.

Für das SHARE-RV-Projekt wird die bevölkerungsrepräsentative Befragung der international vergleichenden SHARE-Studie zu Grunde gelegt. Damit ist es möglich, die Versicherten der gesetzlichen RV im Vergleich mit anderen Gruppen der Bevölkerung wie den Beamten und Selbständigen zu untersuchen. Allerdings ist in diesem Studiendesign nicht für alle Befragten die Kopplung mit Daten der RV möglich, weil nicht alle ein Konto bei der RV besitzen.

4.1 Ermittlung der Sozialversicherungsnummer

Die Befragung für die SHARE-Studie wird von einem Befragungsinstitut durchgeführt. Die Interviewer sind

¹⁹ Stegmann, Aufbereitung der Sondererhebung „Versicherungskontenstichprobe (VSKT)“ als Scientific Use File für das FDZ-RV in: Die Versicherungskontenstichprobe als Scientific Use File, DRV Schriften Band 79, 2008, S. 17–33.

²⁰ Allerdings kann die Schaffung neuer sozialversicherungsrechtlicher Zustände wie die Anerkennung von Zeiten ehrenamtlicher Pflege oder die Ausweitung der Versicherungspflicht für Arbeitslose zu Periodensprüngen in den Daten führen (Mika, The effects of social and institutional change on data production. The case of welfare state reforms on the rise and decline of unemployment and care-giving in german pension fund data, in: Historical social research, Vol. 34 [2009] 3, S. 115–137).

²¹ Z.B. Fachinger/Künemund, Grenzen und Verbesserungsmöglichkeiten des Scientific Use Files der Versicherungskontenstichprobe 2005, S. 105.

²² Protsch/Groß, Die Bedeutung des Scientific Use Files der Versicherungskontenstichprobe 2005 aus der Sicht der sozialen Ungleichheitsforschung, in: Die Versicherungskontenstichprobe als Scientific Use File, DRV Schriften, Band 79, 2008, S. 125–146.

²³ Ein Beispiel hierfür ist der PASS-Datensatz, der die soziale Lage von Beziehern des Arbeitslosengelds II untersuchen sollte. Hierfür wurde vom Institut für Arbeitsmarkt und Berufsforschung (IAB) eine Stichprobe von Beziehern dieser Geldleistung als Ausgangsbasis gezogen (www.iab.de/de/befragungen/iab-haushaltspanel-pass.aspx).

dabei angewiesen, die Befragten um die Verknüpfung der Befragungsdaten mit Informationen aus dem Konto der RV zu bitten, wobei ausdrücklich darauf hingewiesen wird, dass die Teilnahme freiwillig ist. Im Falle der Zustimmung übergibt der Interviewer einen Bogen mit der Bitte, Name, Vorname und eventuell Geburtsname (Anfangsbuchstabe) sowie Geburtsdatum und Versicherungsnummer anzugeben. Die Zustimmung zur Datenverknüpfung wird durch die Unterschrift wirksam. Mit einem bereits frankierten Briefumschlag werden diese Bögen von den Befragten selbständig an das FDZ-RV geschickt. Durch diesen Ablauf ist sichergestellt, dass weder das Befragungsinstitut noch die Partner von SHARE an der Universität Mannheim die Versicherungsnummer erfahren. Das Forschungsdatenzentrum überprüft die eingesandten Versicherungsnummern auf Plausibilität²⁴.

4.2 Kooperation aller Rentenversicherungsträger bei der Datenerhebung

Die Befragten der SHARE-Studie verteilen sich über alle Rentenversicherungsträger (RV-Träger). Es gibt jedoch keinen Datenbestand von Statistikdaten für die Versicherten und Rentner, der die Versicherungsnummer enthält und aus dem für Befragungsdaten in kurzer Zeit entsprechende prozessproduzierte Daten hinzugefügt werden könnten. Daher müssen diese Daten speziell für das entsprechende Projekt aus den Konten der RV-Träger gezogen und gemeldet werden. Die Teilnahme an Projekten, die mit der Erstellung derartiger Datensätze einhergehen, muss daher von allen RV-Trägern gemeinsam gebilligt werden. Die Gremien der RV dienen einer solchen gemeinsamen Beschlussfassung. Das Projekt SHARE-RV wurde von den zuständigen Gremien der RV befürwortet, wodurch die Kooperation aller RV-Träger bei der Datenerstellung sichergestellt ist. Die gemeinsame Datenstelle der Deutschen Rentenversicherung in Würzburg koordinierte die Datenanforderung und prüfte die eingehenden Datenlieferungen.

Um die Arbeit der Datenstelle und der RV-Träger zu erleichtern, wurde einerseits auf das Konzept des bewährten Statistikdatensatzes „Versicherungskontenstichprobe“ zurückgegriffen und andererseits der Zeitpunkt der Datenmeldung auf den für die reguläre Statistikmeldung gelegt. Zudem wurde auf die Anforderung einer speziellen Kontenklärung für die Daten verzichtet, weil bei Personen ab dem Alter 50 die Hoffnung bestand, dass die Konten schon überwiegend geklärt wären.

Das Projekt war mit seiner Datenanforderung insofern innovativ, als Daten für die Versicherungskontenstichprobe sonst nur für Personen bis zum Alter 67 für die Statistik gemeldet werden, die Befragten des SHARE-Projektes teilweise aber älter sind. Daher wurden in diesem Fall erstmals Längsschnitt-Biographiedaten von über 67-Jährigen aus den Konten der RV-Träger generiert. Das Ergebnis war beeindruckend: alle RV-Träger konnten für die angefragten Versiche-

rungsnummern Längsschnittdaten für alle Altersgruppen der SHARE-Befragung liefern.

4.3 Anonymisierung

Die Daten der RV-Träger werden bereits ohne Sozialversicherungsnummer an die Datenstelle zurückgemeldet. Die Daten enthalten zu diesem Zeitpunkt allerdings noch die Originalnummer des Fragebogens. Weil auch diese an einem anderen Ort, nämlich dem Befragungsinstitut, mit den persönlichen Angaben zusammengeführt werden könnte, durchlaufen die Daten einen weiteren Anonymisierungsschritt. Sie werden mit der Fragebogennummer an die zentrale Datenkoordinationsstelle des internationalen SHARE-Projekts gesandt (centeERdata in den Niederlanden)²⁵. Von dort werden die Statistikdaten mit einer neuen Personenkennziffer an den Projektpartner MEA in Mannheim geschickt. Nun haben die Befragungsdaten und die Daten aus den Versicherungskonten der RV die gleiche Personenkennziffer und können zusammengeführt werden. Die im Statistikformat erhobenen Daten werden unmittelbar in das Format des Scientific Use Files „Versicherungskontenstichprobe“ umgesetzt, wodurch Statistikdaten in faktisch anonymisierter Form vorliegen.

5. Veröffentlichung der Forschungsdaten SHARE-RV

Ziel des Projektes SHARE-RV ist die Herstellung eines Forschungsdatensatzes, der nicht nur den Projektpartnern, sondern auch weiteren Forschern an anderen Forschungseinrichtungen und Universitäten zur Verfügung stehen kann und damit von einer Vielzahl von Sozialforschern genutzt werden kann. Beide Projektpartner, die Deutsche Rentenversicherung Bund und SHARE, stellen hierfür jeweils ihren eigenen Datensatz nach ihren Regeln zur Verfügung. Die Weitergabe der Forschungsdaten aus den Konten der RV erfordert einen Vertrag mit der Deutschen Rentenversicherung Bund. Die Daten werden grundsätzlich nur dann weitergegeben, wenn sie für ein Projekt mit einer Fragestellung zum Themenbereich der sozialen Sicherung benötigt werden, das an einer öffentlichen Forschungseinrichtung oder Universität durchgeführt werden soll. Über das FDZ-RV wird dann ein Datennutzungsvertrag geschlossen, in dem sich die Besteller zur Einhaltung der Datenschutzvorschriften durch alle Mitarbeiter verpflichten. Der Zugang zu den Daten des SHARE erfolgt über das Forschungsdatenzentrum SHARE²⁶, indem ein Nutzungsantrag gestellt wird, durch den Zugangsdaten für das Internetportal bereitgestellt werden und die Daten anschließend frei zum Download zur Verfügung stehen.

²⁴ Das stellt insofern eine Schwierigkeit dar, als die Versicherungsnummer der Rentner länger ist und daher auf dem Zustimmungsbogen Raum für diese sehr lange Nummer geboten werden muss.

²⁵ www.centerdata.nl/en/TopMenu/Over_CenteERdata/.

²⁶ www.share-project.org.