

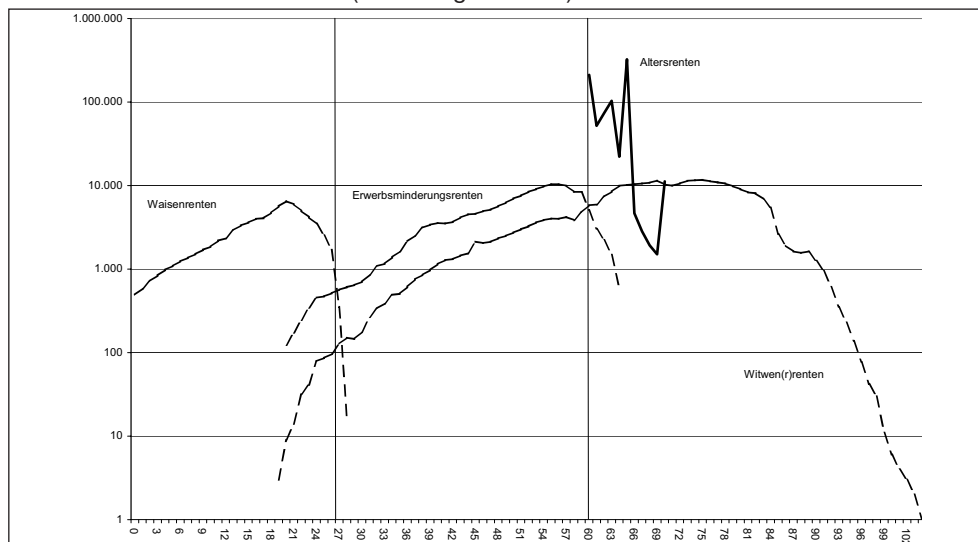
Zusammenfassung und Ausblick auf die weitere Entwicklung des FDZ-RV

Ralf K. Himmelreicher und Jonas Radl
Deutsche Rentenversicherung Bund, FDZ-RV Berlin

Die Daten der Deutschen Rentenversicherung stellen für die Wissenschaft einen großen Fundus an Mikrodaten dar, die nahezu den gesamten Lebensverlauf der Versicherten umschließen. **Abbildung 1** zeigt, in welchem Lebensalter Personen in Deutschland erstmalig eine Rente beziehen und in welchen Altersgruppen diese besonders häufig in Anspruch genommen werden (können); das Vorkommen der verschiedenen Rentenarten weist darauf hin, welche unterschiedlichen Risiken von der Deutschen Rentenversicherung abgesichert werden. In die Rentenversicherung ist nahezu die gesamte volljährige Bevölkerung integriert, weil eigene Beiträge abgeführt werden oder etwa Kinder erzogen oder Angehörige gepflegt werden.

Es gibt sowohl eigene als auch abgeleitete Renten. Eigene, auf selbst geleisteten Beiträgen beruhende Renten, sind so genannte Versichertenrenten; sie lassen sich in Renten wegen verminderter Erwerbsfähigkeit und wegen Alters differenzieren. Abgeleitete Renten werden aus den Konten der Versicherten abgeleitet, bei dieser Rentenart handelt es sich um Witwen(r)- und Waisenrenten – also Renten wegen Todes. Letztere vergrößern den in der Rentenversicherung einbezogenen Personenkreis. Aus **Abbildung 1** geht hervor, dass im Jahre 2004 die Neuzugänge in Altersrenten mit einem Alter von 60, 63 und 65 Jahren mit großem Abstand die häufigste Verrentungsursache darstellte. Renten wegen verminderter Erwerbsfähigkeit kommen bei 55-Jährigen besonders häufig vor, während Waisenrenten sich überwiegend bei jungen Erwachsenen, Witwen- und Witwerrenten bei Personen im Alter zwischen 65 und 85 Jahren kumulieren.

Abbildung 1: Anzahl der Rentenzugänge nach verschiedenen Rentenarten und Einzelalter in Deutschland 2004 (Anzahl logarithmiert)



Quelle: VDR Rentenzugangsstatisik 2004; für Renten wegen verminderter Erwerbsfähigkeit Tabelle 201.00 Z RV, für Renten wegen Alters Tabelle 301.00 Z RV, für Witwen(r)renten Tabelle 402.00 Z RV, für Waisenrenten Tabelle 601.00 Z RV; kumulierte Altersklassen wurden geglättet.

Die in Abbildung 1 dargestellten Rentenzugänge repräsentieren den Ausschnitt aus den Statistiken der Deutschen Rentenversicherung, der von der Wissenschaft besonders häufig nachgefragt wird. Neben der Rentenzugangsstatistik existieren weitere prozessproduzierte Datensätze, die in Zukunft im FDZ-RV insbesondere als Scientific Use Files (SUFs) angeboten werden sollen. Einen Überblick über das verfügbare und potenzielle Datenangebot des FDZ-RV zeigt **Abbildung 2**.

Abbildung 2: Potenziell und aktuell verfügbare prozessproduzierte Mikrodaten im FDZ-RV

Rentenzugangsstatistik	Rentenbestandsstatistik	Rentenzahlbestandsstatistik	Versichertenstatistik	Rehastatistik	Versorgungsausgleichsstatistik
Querschnittsdatsätze					
Rentenzugang Rentenwegfall und Rentenänderung	Rentenzahlungen aus den Trägerkonten	Mehrfachbezug von Rentenzahlungen	Statistik der aktiv Versicherten	RSD-Jahreserhebung abgeschlossener Reha-Leistungen	Bestandsfälle aller Versicherten mit Versorgungsausgleich
Längsschnittsdatsätze					
Vollendete Versichertenleben			Versicherungskontenstichprobe	RSD-Verlaufserhebung	

Quelle: eigene Abbildung; = aktuell verfügbar im FDZ-RV.

Als Scientific Use Files stehen der Wissenschaft aktuelle Querschnittsdatsätze der Rentenstatistik (Rentenzugang, Rentenwegfall und Rentenbestand) für die Berichtsjahre 2003 und 2004 zur Verfügung. Diese Daten treffen – wie die in diesem Band zusammengestellten Beiträge zeigen – vor allem auf das Interesse von Forscherinnen und Forschern der Sozial- und Wirtschaftswissenschaften sowie der Demografie. Mit der für die Zukunft geplanten Erweiterung des Datenangebots auf die Versicherten- und Rehabilitationsstatistik wird einerseits der interdisziplinäre Charakter der Nutzerstruktur des FDZ-RV erweitert werden. Andererseits sollen durch die angestrebte Aufbereitung der Längsschnittdaten „Versicherungskontenstichprobe“ und „Vollendete Versichertenleben“ die an biografischen Verlaufsangaben interessierte Forschungsgemeinde angesprochen werden. Im Hinblick auf den Einbezug weiterer Datensätze in das Angebot des FDZ-RV ist jedoch darauf hinzuweisen, dass eine Ausweitung des Datenangebots ressourcenintensiv ist.

Die im vorliegenden Band dokumentierten Analysen basieren überwiegend auf SUFs des FDZ-RV, die aus der Rentenzugangs-, Rentenbestands- und Rentenwegfallstatistik der Deutschen Rentenversicherung erstellt wurden. Bei diesen Datensätzen handelt es sich um erstmals für die nichtkommerzielle wissenschaftliche Nutzung erstellte faktisch anonymisierte Mikrodaten. Ziel der vorgeschalteten Umfrage und des ersten Workshops war es, die Datenprioritäten der Wis-

senschaftler zu erheben und bei der Auswahl und Aufbereitung der Datensätze einfließen zu lassen. Datenschutzrechtliche Prüfverfahren und die Besonderheiten der prozessproduzierten Daten zogen eine Suche nach Kompromisslösungen nach sich. Inzwischen sind die ersten SUFs aufbereitet, und erste Analysen sind erstellt.

Die konstruktiven Ergebnisse der Diskussionen auf dem zweiten Workshop des FDZ-RV sowie der vielen persönlichen Gespräche sind in diesen Tagungsband eingeflossen. Die Anregungen und Hinweise, die sich auch in den Beiträgen der Tagungsdokumentation wiederfinden, beziehen sich auf drei verschiedene Aspekte: (1) die Organisation des Forschungsdatenzentrums, (2) die Struktur der Datensätze und den Datenzugang und (3) die Merkmale der Datensätze.

(1) Organisation des Forschungsdatenzentrums:

Die Bewertungen der bisherigen Arbeit des Forschungsdatenzentrums durch die versammelte scientific community fielen positiv aus. Besonders wurde die gute und verbindliche Betreuung der Nutzer durch die Mitarbeiter des FDZ-RV und anderer Experten in der Deutschen Rentenversicherung Bund gewürdigt.

Das Internetportal (www.fdz-rv.de) wurde wegen des reichhaltigen Informationsangebotes und der benutzerfreundlich zu beziehenden Anträge gelobt. Darüber hinaus können über das Internetportal detaillierte Datensatzbeschreibungen und Codepläne heruntergeladen werden. Die Tagungsteilnehmer äußerten den Wunsch, eine Auflistung der auf Daten des FDZ-RV basierenden Publikationen im Internetportal bereitzustellen.

Prozessproduzierte Daten sind häufig schwer zu interpretieren, weil sie eine spezifische rechtliche Grundlage haben. Im Fall der Daten der Rentenversicherung wurden diese insbesondere deshalb erhoben, um die Höhe der Ansprüche der Versicherten berechnen zu können. Die wichtigste Anforderung an die Dokumentationen der Daten des FDZ-RV ist darum, dass sie die einzelnen Rechtsgrundlagen benennen und ihre Umsetzung in die Statistiken transparent machen. In diesem Zusammenhang ist einzuräumen, dass eine vollständige Aufklärung der Bedeutungsgehalte einzelner Variablen den Forscherinnen und Forschern eine Auseinandersetzung mit dem Sozialgesetzbuch, insbesondere mit dem SGB VI, abverlangt. Dass sich die Dokumentation der Daten verbessert, belegt der vorliegende Tagungsband. Die methodisch orientierten Beiträge im ersten Kapitel fungieren als Operationalisierungs- und Interpretationshilfe und zeigen Analysepotenziale auf.

Öffentlich geförderte Forschungsdatenzentren sind im Arbeitskreis öffentlicher Forschungsdatenzentren (AKöffFDZ) zusammengeschlossen. In diesem Kreis werden Fragen zur konzeptionellen Ausrichtung der Forschungsdatenzentren mehrmals im Jahr diskutiert; dabei wird angestrebt, die Maßnahmen und Termine der verschiedenen Forschungsdatenzentren zu koordinieren. Die Praxis der gegenseitigen Teilnahme an Veranstaltungen der anderen Forschungsdatenzentren wurde auf dem zweiten Workshop des FDZ-RV fortgesetzt und trägt dazu bei, diese Kooperation zu verstetigen (vgl. Wübbeke sowie Philipp & Schoffer im vorliegenden Band). Ferner werden durch den Rat für Sozial- und Wirtschaftsdaten (www.RatSWD.de) Kooperationsvorhaben zwischen Wissenschaft und Datenproduzenten gefördert.

(2) Struktur der Datensätze und Datenzugang

Die Anonymisierungsverfahren, mithilfe derer die Anforderungen des Datenschutzes im FDZ-RV erfüllt werden, wurden von den Teilnehmern des Workshops für gut befunden. Indem auf verfremdende Verfahren wie den Variablenringtausch verzichtet wurde, werden die Auswertungsmöglichkeiten kaum eingeschränkt. Trotzdem wird durch die Restriktion der minimalen, univariaten Zellenbesetzungen mittels Zusammenfassung von seltenen Ausprägungen sowie mittels top- und bottom-coding den Ansprüchen des Datenschutzes entsprochen. Durch Ziehung einer zufälligen Stichprobe aus den Totalerhebungen gelangt man zu repräsentativen Datensätzen. Bei diesen handelt es sich überwiegend um Fall- und nicht um Personenstatistiken (vgl. Himmelreicher im vorliegenden Band).

Inzwischen wurden erste Erfahrungen mit Gastwissenschaftlern gesammelt. Die angestrebte tiefere Datenstruktur von Datensätzen für die Nutzung der Daten in den Räumen des FDZ-RV erwies sich unter den Bedingungen der hohen Kontrolle an einem Gastwissenschaftler-Arbeitsplatz als datenschutzrechtlich unbedenklich. Der persönliche Kontakt zwischen Gastwissenschaftlern und den Mitarbeitern und Mitarbeiterinnen des Forschungsdatenzentrums stellte sich dabei für beide Seiten als konstruktiv heraus, da einerseits Unterstützung im Bereich der Interpretation der Daten angeboten wurde und andererseits Unklarheiten in den Datensatzdokumentationen gefunden und behoben werden konnten.

Die großen Fallzahlen in den SUFs des FDZ-RV sind ein herausragendes Charakteristikum der Datensätze, die im Gegensatz zu surveybasierten Daten die Möglichkeit bieten, differenzierte Analysen selbst zu kleinen Bevölkerungsgruppen zu unternehmen. Hoch- und Höchstaltrigen kommen in der Regel in hinreichender Zahl vor (vgl. Scholz im vorliegenden Band). Migranten lassen sich danach unterscheiden, ob es sich um Personen mit ausländischer oder deutscher Nationalität (z.B. Spätaussiedler) handelt (vgl. Mika im vorliegenden Band). Eine allgemeine Stärke der Daten des FDZ-RV ist ihre hohe Validität, die durch umfangreiche routinemäßige Fehler- und Plausibilitätsprüfungen gesichert ist (vgl. Luckert im vorliegenden Band).

(3) Merkmale der Datensätze

Zur Aufbereitung einiger Merkmale im SUF Versichertenrentenzugang 2003 wurden verschiedene Anregungen gegeben: Kritisiert wurde das Fehlen des Merkmals Leistungsart mit einer Differenzierung nach den verschiedenen Altersrentenarten. Des Weiteren wurde darauf hingewiesen, die drei Merkmale zum Bruttojahresverdienst möglichst nicht mit der Anzahl der dazu aufgewendeten Kalendertage zu gewichten, sondern beide Merkmale getrennt auszuweisen. Beide Vorschläge wurden im SUF Rentenzugang 2004 umgesetzt. Außerdem finden sich seit dem Berichtsjahr 2004 sowohl der Tätigkeitenschlüssel als auch die Bildungsangabe im Datensatz, sodass sozialwissenschaftliche Skalen zu Beruf und Prestige operationalisiert werden können (vgl. Stegmann im vorliegenden Band).

Bei Angaben zum Erwerbseinkommen und der Höhe der gesetzlichen Rente bewähren sich prozessproduzierte Daten besonders, weil keine verzerrenden Falschantworten möglich sind (vgl. Berkel & Börsch-Supan sowie Unger im vorliegenden Band). Hervorgehoben wird auch die Verlässlichkeit der Informationen zu Beschäftigungszeiten und Entgeltbestandteilen. Mithilfe der verfügbaren SUFs lassen sich viele Forschungsfragen adäquat operationalisieren, weil die Höhe der Entgeltpunkte näherungsweise die Höhe des Erwerbseinkommens über den Le-

benszyklus widerspiegelt. Die Auswertungsmöglichkeiten werden allerdings durch die Rechtszensierung des Einkommens, die aufgrund der Beitragsbemessungsgrenze zu Stande kommt, begrenzt (vgl. von Gaudecker im vorliegenden Band).

Interessant sind Vergleiche zwischen Mikrodaten der Rentenversicherung und Daten der amtlichen Statistik. In Bezug auf die Mortalitätsforschung erwiesen sich die Daten der Rentenversicherung als wertvolle Datenquelle, die bestehende Lücken in der Abdeckung des Mikrozensus zu schließen hilft (vgl. Scholz im vorliegenden Band). Bei der Untersuchung von Übergängen in den Ruhestand erschweren die unterschiedlichen Definitionen des Erwerbsstatus in amtlicher und Rentenstatistik den Vergleich (vgl. Böttcher im vorliegenden Band).

Den SUF Versichertenrentenzugang zeichnet das Beobachtungsfenster der letzten drei Jahre vor Beginn des erstmaligen Rentenbezuges aus, innerhalb derer differenzierte Daten zu Einkommen und erwerbsbiografischen Status der Versicherten vorliegen (vgl. Böttcher im vorliegenden Band). Aufgrund dieser Längsschnittkomponenten kann der SUF Versichertenrentenzugang als *retrospektiver Querschnitt* gelten. Das Analysepotenzial des Datensatzes wird dadurch – gerade für die Erforschung von Altersübergängen – gesteigert (vgl. Büttner & Brussig im vorliegenden Band). Der hohe Anteil von Versicherten mit unbekanntem Status vermindert die Güte der diesbezüglichen Befunde.

Der beklagte Mangel an Vermögensangaben in den Datensätzen des FDZ-RV lässt sich nicht beseitigen. In dieser Hinsicht sind jedoch auch Surveydaten wenig verlässlich. Gleichfalls sind die in einigen Forschungszusammenhängen wünschenswerten Daten zu subjektiven Erwartungen oder der Zufriedenheit mit der Rentenversicherung nicht erhoben worden. Der Haushaltskontext der Rentenversicherten und -bezieher lässt sich nicht ermitteln. Diese Möglichkeit wäre lediglich mit verknüpften Daten gegeben.

Fazit und Ausblick

Das Forschungsdatenzentrum der Rentenversicherung besteht seit Januar 2004 (vgl. Rehfeld im vorliegenden Band). In den zwei Jahren seit seiner Gründung haben annähernd fünfzig Personen einen Vertrag zur Datennutzung mit dem Forschungsdatenzentrum abgeschlossen, und nahezu ebenso viele Personen sind als potenzielle Nutzer registriert. Die Infrastruktur des Forschungsdatenzentrums mit seinen beiden Standorten in Berlin und Würzburg ist eingerichtet. Eine vom Rat für Sozial- und Wirtschaftsdaten eingesetzte Expertengruppe hat das FDZ-RV im Juli 2005 evaluiert und die bisherigen Schritte des Aufbaus für gut befunden. Derzeit werden die SUFs aktualisiert und weitere Daten, insbesondere die Längsschnittdaten der Rentenversicherung, sollen in das Datenangebot des FDZ-RV integriert werden.

In Zukunft wird von den Mitarbeiterinnen und Mitarbeitern angestrebt, sofern es im Rahmen des vorhandenen Arbeitsvolumens möglich ist, verstärkt selbst Analysen mit Daten des FDZ-RV durchzuführen, um die sozialpolitische Diskussion zu bereichern und die SUFs des FDZ-RV in der scientific community bekannter zu machen; hierzu sollen weiterhin Publikationen beitragen – einige befinden sich bereits in Vorbereitung.

Mit einem großen Teil der Nutzerinnen und Nutzer der Daten des Forschungsdatenzentrums konnte – beginnend mit der Umfrage zu den Nutzerpräferenzen, gefolgt vom ersten bis hin zum zweiten Workshop des Forschungsdatenzentrums – eine längere Beziehung aufgebaut werden. Ein weiterer Nutzerkreis konnte erst in den letzten Monaten, u.a. mit dem Umzug des FDZ-RV nach Berlin erreicht werden. Mit alten wie neuen Nutzern des FDZ-RV streben wir eine intensive Kooperation an und hoffen, dazu beitragen zu können, dass mit den prozessproduzierten Daten des Forschungsdatenzentrums gerne geforscht wird. Zu einem verstärkten Erfahrungsaustausch soll der dritte Workshop des Forschungsdatenzentrums der Rentenversicherung beitragen, der vom 26. bis 28. Juni 2006 in Bensheim stattfinden wird. Ein call for papers wird im Frühjahr 2006 versendet, und wie gewohnt sollen die dort gehaltenen Vorträge in einem Tagungsband publiziert werden.