

FDZ-Biografiedatensatz
für die
Biografiedaten zu Vollendeten Versichertenleben (VVL) 2007
(SUFVVL2007_Fix und SUFVVL2007_“*Verlaufsmerkmal*“)

Umsetzung des SK79
in einen anonymisierten Datensatz fester Satzlänge:
Sequentielle Biografiedaten

Autor:
Dr. Michael Stegmann
(FDZ-RV, Würzburg)

Stand: 04.05.2010

Inhalt

1	DIE SONDERERHEBUNG „VOLLENDETE VERSICHERTENLEBEN“ (VVL)	3
1.1	GRUNDSÄTZLICHES	3
1.2	AUFBAU DER DATENSÄTZE	4
1.2.1	<i>Datensatz SK 79</i>	4
1.2.2	<i>Datensatz SK 90</i>	4
1.2.3	<i>Selektion der Dateninformation</i>	4
2	SK 79 ALS BIOGRAFIEDATENSATZ	6
2.1	SEQUENTIELLE BIOGRAFIEDATEN	6
2.2	ZEITLICH PARALLELE UND ÜBERLAPPENDE SOWIE IM MONAT AUF EINANDER FOLGENDE BLÖCKE....	8
2.2.1	<i>Zuordnungshierarchie</i>	9
2.2.2	<i>Prioritätenregelung bei gleicher Hierarchie</i>	9
2.2.2.1	Parallele, überlappende oder im Monat aufeinander folgende Pflichtbeiträge	9
2.2.2.2	Zeitliche Parallelität ohne Beiträge	9
2.2.2.3	Zeitliche Überlappung zweier Blöcke.....	9
2.2.2.4	Sonderfall Arbeitgeberwechsel im Monat und parallele Beschäftigung.....	10
2.2.3	<i>Wann gilt ein Monat als belegt</i>	10
2.2.4	<i>Ausnahme Kindererziehungszeit</i>	10
2.3	UMSETZUNG DER RENTENRECHTLICHEN BIOGRAFIEANGABEN	11
2.3.1	<i>Einführung von Sozialen Erwerbssituationen (SES)</i>	12
2.3.2	<i>Sondermerkmal Geringfügigkeit</i>	15
2.3.3	<i>Belegte Tage im Biografie monat</i>	15
2.3.4	<i>Monatsbezogene Entgeltpunktinformationen</i>	15
2.3.5	<i>Sondermerkmal Kumulation von Beschäftigungen</i>	16
2.3.6	<i>Sonderfall Zurechnungszeit und Ende der Biografie</i>	16
3	UMSETZUNG IN FESTE SATZLÄNGE	17
3.1	MERKMALE.....	17
3.2	DATENSATZSTRUKTUR.....	18
4	DATENWEITERGABE VON MERKMALEN DES FIXEN TEILS	18
5	EINGRENZUNG UND HOCHRECHNUNG DER STICHPROBE	18
6	ABLAUFSHEMA DER DATENGENERIERUNG	19
7	ANONYMISIERUNG UND DATENSCHUTZ	21

1 Die Sondererhebung „Vollendete Versichertenleben“ (VVL)

1.1 Grundsätzliches

Das Konzept der VVL ist darauf angelegt Rechtsänderungen in ihren Auswirkungen möglichst empirisch zu überprüfen. Die vorhandenen Routinestatistiken reichen dafür nicht aus.

Es wird für ein Rentenzugangsjahr eine Stichprobe auf maschinellm Wege aus dem Versicherungskonto in Form des Statistikdatensatzes SK79, wie er bei den Versicherungskontenstichprobe gemäß § 1 Abs. 2 RSVwV verwendet wird, erstellt.

Für die Rentenzugangsfälle, für welche die Konten vollständig aktuell aufbereitet sind, wird von der Deutschen Rentenversicherung Bund, Geschäftsbereich 0500, eine Stichprobe gezogen. Für diese Fälle werden maschinell Datensätze über den Versicherungsverlauf (SK 79) von den kontenführenden Versicherungsträgern bereitgestellt. Es wurde für die Erhebung 2007 eine Brutto-Stichprobe in der Größenordnung von rund 145.000 Fällen aus dem Rentenzugang (ca. 1/5 der Gesamtfälle) angestrebt, wobei insbesondere nur Neuzugänge, Versichertenrenten und Nicht-Vertragsrenten in die Erhebung einbezogen wurden.

Die VVL lässt sich demnach beschreiben als eine systematische Zufallsauswahl aus dem Rentenzugang mit Meldegrund 10 (Festsetzung ohne unmittelbar vorhergehenden Rentenbezug aus einer gesetzlichen Rentenversicherung). Es werden nur ausgewählte Leistungsarten berücksichtigt (Erwerbsminderungs- und Altersrenten). Außerdem darf es sich nur um Nichtvertragsrenten handeln. Die VVL ist brutto eine 20%-Stichprobe dieser Fälle des Rentenzugangs.

Kurz beschrieben, wird bei der Gewinnung der Fälle folgendermaßen verfahren:

- Ziehen einer geeigneten Stichprobe aus den Rentenzugangsdaten (SK90) des Berichtsjahres (Deutsche Rentenversicherung Bund, Geschäftsbereich 0500).
- Übermittlung der Statistikdatensätze an die kontoführenden Anstalten,
- Reanonymisierung durch die Versicherungsträger,
- die Versicherungsträger erstellen die Versicherungskontendatensätze SK79 und übermitteln diese unter der gleichen anonymen Zählnummer wie der zugehörige Zugangsdatensatz an die Deutsche Rentenversicherung Bund, Geschäftsbereich 0500.

Bei den Fällen der VVL handelt es sich um Personen, die im Berichtsjahr dem Rentenzugang der genannten Rentenarten zuzurechnen sind. Es handelt sich also um Renten die rechtskräftig beschieden sind und damit auf geklärten Versicherungskonten basieren.

1.2 Aufbau der Datensätze

Für die Untersuchungspersonen liegen Informationen nach dem Datensatz SK 79 und Daten in Form des SK 90 vor.

1.2.1 Datensatz SK 79

Der Datensatz SK 79 zur Versicherungskontenstichprobe gemäß § 1 Abs. 2 RSVwV gliedert sich in einen festen Teil mit datentechnischen und demographischen Merkmalen sowie Ergebnissen aus der Rentenberechnung und in einen variablen Teil mit Datenblöcken für jede versicherungsrechtlich relevante Zeit. Der feste Satzteil mit Werten aus der Rentenberechnung entspricht weitestgehend dem Rentendatensatz SK 90.

1.2.2 Datensatz SK 90

Der SK 90 ist der einheitliche Statistikdatensatz zur Rentenzugangsstatisik nach § 6 RSVwV und zur Rentenbestandsstatistik nach § 7 RSVwV aus den Konten der Rentenversicherungsträger. Für die Fälle des Rentenzugangs werden u.a. soziodemografische Angaben, Angaben zum Rentenbeginn, Merkmale zur Rehabilitation und zur Gesamtleistungsbewertung sowie zur Rentenberechnung gemeldet. Die Informationen beziehen sich auf die Sachverhalte zum Zeitpunkt der Rentengewährung.

1.2.3 Selektion der Dateninformation

Aus dem SK 79 werden die zusammengefassten Verlaufsinformationen weitergegeben. Dazu wird der variable Teil in der nachfolgend beschriebenen Art und Weise aufbereitet. Die summarischen Merkmale zur Rentenberechnung und die soziodemografischen Angaben stammen aus dem SK 90 und dem festen Teil des SK 79. Aus dem SK 79 werden die Informationen zur Soziodemografie weitergegeben, die sich im SK 90 nicht finden. Das Gros der Merkmale – insbesondere die Werte zur Rentenberechnung – sind dem SK 90 entnommen.

Es sei erwähnt, dass es Unterschiede zwischen dem SK 79 und dem SK 90 geben kann. Dies resultiert daraus, dass der SK 79 auf den Informationen des Versicherungskontos beruht, das zum Erhebungszeitpunkt vorliegt. Die Rente wird in der VVL also zum 1.1. des Berichtsfolgejahres berechnet. Es findet also der Rechtsstand ab 1.1. des Berichtsfolgejahres Anwendung, dieser kann sich vom Rechtsstand zum Zeitpunkt der tatsächlichen Rentengewährung unterscheiden. Haben sich zwischen dem Zeitpunkt des Rentenzugangs (Inhalt des SK 90) und der Erhebung des SK 79 Sachverhalte geändert, kann daraus außerdem eine Abweichung in der Bewertung der Versicherungsbiografie entstehen. Vor allem kann daraus eine abweichende Zurechnungszeit bei EM-Renten resultieren. Es ist aber auch möglich, dass nach Antritt der Rente in einigen Fällen noch weitere Anwartschaften erworben wurden. Ferner finden sich keine ehgattenbezogenen Sachverhalte, die im Rahmen der Rentengewährung beachtet werden, als *verlaufsbezogene Informationen im Rentenversicherungskonto*. Als Beispiel sind Zeiten nach dem Fremdrentengesetz zu benennen, in vielen Fällen fehlt im SK 79 die Begrenzung der Entgeltpunkte nach §22b FRG. Da der feste Teil des SK 79 aber ausschließlich auf den Informationen beruht, die im Konto geführt werden, sind diese Zeiten nicht ausgewiesen.

Die Werte der Rentenberechnung, die aus dem SK 90 entnommen werden, spiegeln die Situation wider, die für die Höhe der Rente maßgeblich ist. Diese Informationen werden weitergegeben. Würde man die Rente auf Basis der Informationen des variablen Teils ermitteln, kann es - wie beschrieben - zu Abweichungen zu diesem realen Wert kommen. Dies wird im weitergegebenen Datensatz nicht deutlich, da die Zusammenfassung der Verlaufsmerkmale keine exakte Berechnung der Rentenanwartschaften zulässt, was auch nicht die Intention des Biografiedatensatzes ist.

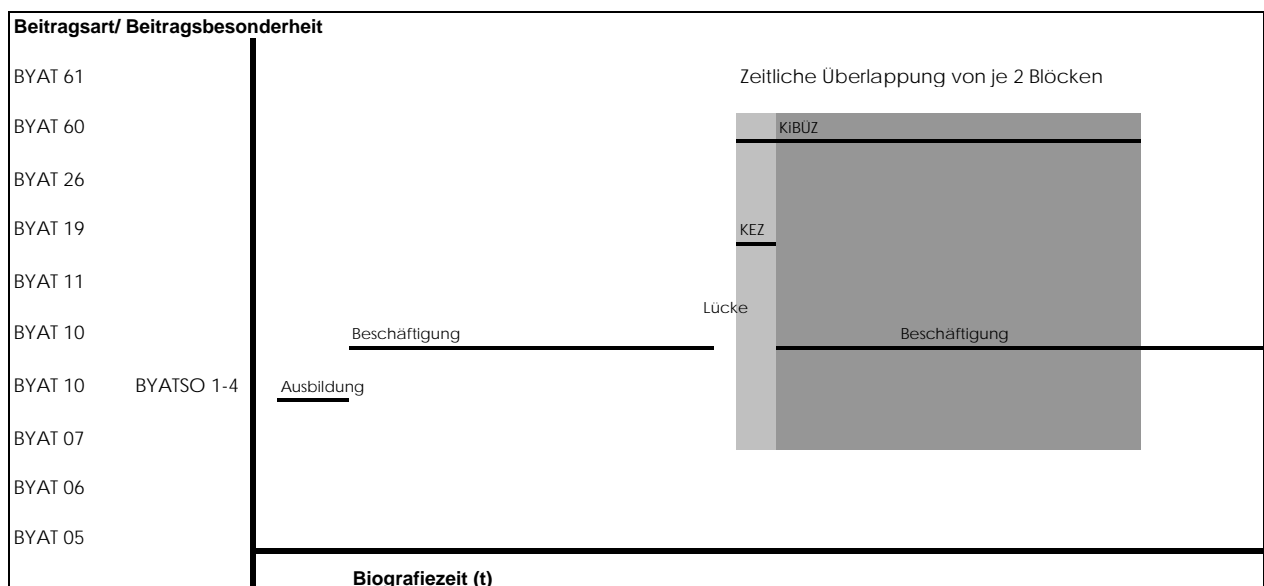
Die Weitergabe der Informationen zur Rentenberechnung aus dem SK 90 bürgt dafür, dass verlässlich Auskunft gegeben wird über Höhe und Zusammensetzung der Rentenanwartschaften. Die Längsschnittinformationen aus dem variablen Teil des SK 79 machen es u.a. möglich, verschiedene Phasen der Biografie zu untersuchen und unter Beachtung der Summeninformationen auch Rückschlüsse auf die damit in Beziehung stehende Rente zu ziehen.

2 SK 79 als Biografiedatensatz

2.1 Sequentielle Biografiedaten

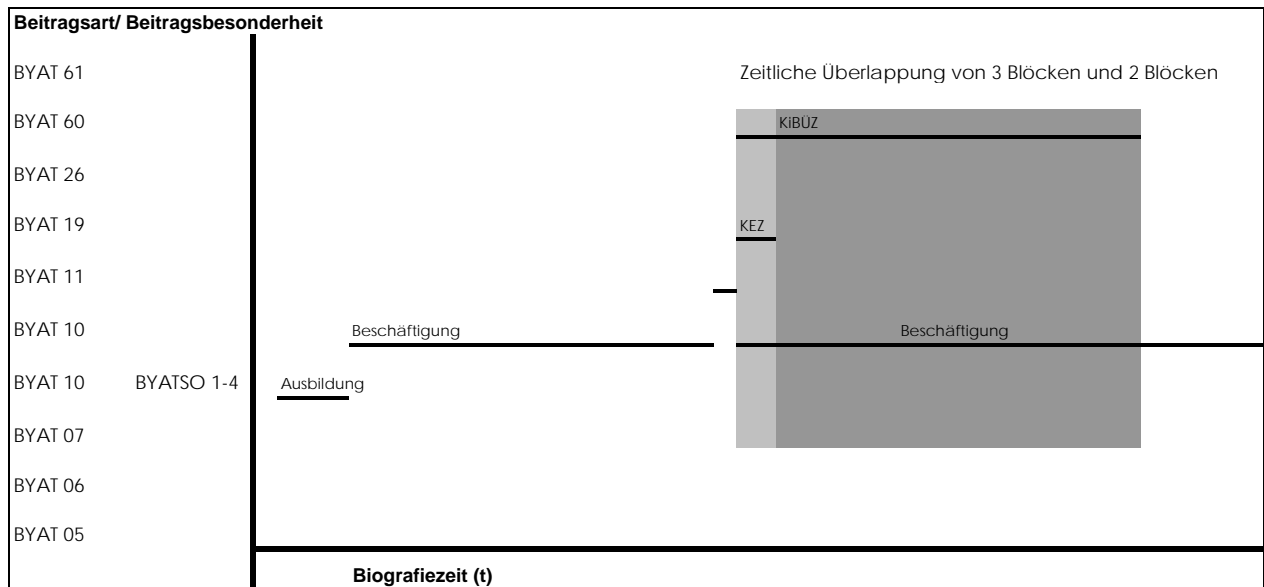
Im variablen Teil des SK 79 werden in zeitlichen Blöcken die Informationen aus dem Rentenversicherungskonto aufgeführt. Die Blöcke sind zeitlich festgelegt durch die „von/bis“-Variablen (VNZR/BSZR). Dabei ist es wichtig zu beachten, dass die einzelnen zeitlichen Blöcke auf der Zeitachse nicht chronologisch sein müssen: Zwischen zwei aufeinanderfolgenden Blöcken kann „Biografiezeit“ fehlen. Es ist aber auch möglich, dass zwei oder sogar mehrere Blöcke zeitlich parallel liegen oder sich zeitlich überlappen. Im Beispiel 1 ist eine Biografie dargestellt, die relativ chronologisch verläuft: Vor der Geburt eines Kindes ergibt sich eine zeitliche Lücke ohne Information, durch die Kinderberücksichtigungszeit kommt es dann zu paralleler Blockbildung Kindererziehungszeit (KEZ)/Berücksichtigungszeit (KiBÜZ) und Beschäftigung/KiBÜZ.

Schaubild1: BEISPIEL 1



Im Beispiel 2 kommt es sogar zu einer zeitlichen Überlappung von drei Blöcken. Beschäftigung, Kindererziehungszeit (KEZ) und Berücksichtigungszeit (KiBÜZ) fallen zeitlich zusammen. Im weiteren Verlauf liegen Beschäftigung und KiBÜZ parallel.

Schaubild 2: BEISPIEL 2

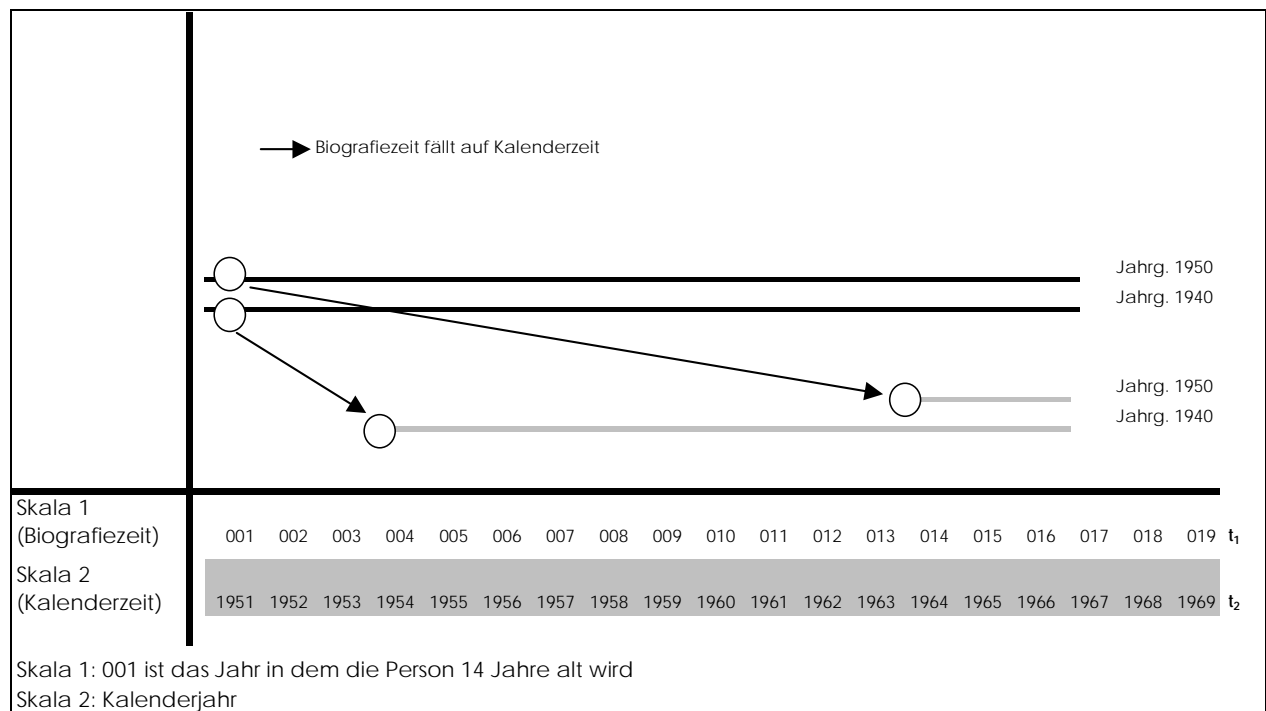


Ein sequentieller Biografiedatensatz zeichnet sich dadurch aus, dass die Datenlogik sich an Biografiemonaten oder -jahren orientiert. Zur besseren Verständlichkeit kann man die Herangehensweise als Frage formulieren, die lautet: Welchen Status hat die Person im Monat X ihres Lebens inne?

Dabei wirft sich die nächste Frage auf: Auf welche Zeit bezieht sich Monat X. Hierzu wählen einige Ansätze (Blossfeld 1985) als einheitliche Bezugszeit die Kalenderjahre. Das bedeutet, der Datensatz beginnt zum Beispiel im Januar des Jahres 1950. D.h., jeder Datensatz beginnt mit Monat 1, dem Januar 1950, gleich ob die Biografie hier bereits begonnen hat (Kalenderzeit). Der Datensatz endet im Monat mit dem jüngsten Ereignis oder einem anderen festgelegten Datum. Damit ergeben sich für „junge“ Fälle zahlreiche Leerstellen am Beginn des Datensatzes und für „alte“ Fälle viele Leermerkmale am Ende.

Eine andere Herangehensweise ist es, an das Lebensalter (Biografiezeit) anzuknüpfen (Untersuchung AVID 1996 und 2005). Der Datensatz beginnt immer im Januar des Jahres, in dem die Person z.B. 14 Jahre alt wird und endet im Dezember des Jahres, in dem die Person 65 Jahre alt wird (AVID 1996). Dies hat den Vorteil, dass der Datensatz grundsätzlich kürzer wird. Es ergibt sich jedoch auch der Nachteil, dass für die Biografien kein direkter Bezug zu einem Kalenderjahr hergestellt ist (die gleichen Biografiemonate finden zu unterschiedlichen Zeitpunkten statt). Das nachfolgende *Schaubild 3* illustriert den Unterschied auf Jahresebene.

Schaubild 3: Kalenderzeit und Biografiezeit



Vorgehen:

Für eine Umsetzung des SK79 in sequentielle Biografiedaten wurde analog der „AVID-Logik“ ebenfalls an das Biografiezeitkonzept angeknüpft. Die Untersuchung AVID 1996 hat dazu Erfahrungen erbracht, die zeigen, dass sich dieses Konzept für die RV-Daten bewährt hat. Im FDZ-Biografiedatensatz für die VVL ist der Biografieverlauf monatsgenau wiedergegeben. Das heißt, die zeitliche Einheit ist jeweils ein Biografiemonat. Der erste Monat ist dabei stets der Januar des Jahres, in dem die Person 14 Jahre alt geworden ist. Der letzte Monat ist der Dezember des Jahres, in dem die Person 65 Jahre alt wird bzw. geworden ist.

2.2 Zeitlich parallele und überlappende sowie im Monat aufeinander folgende Blöcke

Folgt man der Systematik der AVID 1996, dann wird für jeden Monat nur ein Ereignis bzw. Status abgelegt; es gibt also nur eine Statusinformation für jeden Verlaufsmonat. Bei diesem Vorgehen muss bei parallelen bzw. zeitlich überlappenden Blöcken eine Prioritätenregel gefunden werden, die festlegt, welcher variable Block herangezogen wird.

Die Monatsbelegung wird nach folgender Zuordnungshierarchie und Prioritätenregelung vorgenommen.

2.2.1 Zuordnungshierarchie

Ergeben sich für einen Monat mehr als eine Statusinformation aus der Kombination der Merkmale BYAT und BYATSO, wird der Monat aus dem Block beschickt, der in der Zuordnungshierarchie nach dem Merkmal BYAT zuerst kommt:

1. Pflichtbeitrag außer Kindererziehungszeit (BYAT 4 – 18),
2. freiwilliger Beitrag (20, 21)
3. Anrechnungszeit (40-51)
4. Ersatzzeit (30, 31)
5. FZR (25, 26)
6. Rentenbezug (70, 71, 72, 90)
7. Kindererziehungszeit und Erziehung mehrerer Kinder (19, 29)
8. Berücksichtigungszeit und Gutschrift Kinder (60, 61)

2.2.2 Prioritätenregelung bei gleicher Hierarchie

2.2.2.1 Parallele, überlappende oder im Monat aufeinander folgende Pflichtbeiträge
Ergeben sich für einen Monat parallele, überlappende oder im Monat aufeinander folgende Pflichtbeiträge der Stufe 1 der Zuordnungshierarchie, hat der Pflichtbeitrag Priorität, für den sich der auf den Monat bezogene höhere Beitrag ergibt. Der Monatswert ergibt sich durch die Variable EGPT für den Block, sie wird auf den Tag umgerechnet und mit der Anzahl der belegten Tage im Monat multipliziert.

2.2.2.2 Zeitliche Parallelität ohne Beiträge

Liegen Blöcke zeitlich parallel und sind zeitlich deckungsgleich wobei keine Angaben zur Variable EGPT (z.B. Deckungsgleichheit von zwei Anrechnungszeiten, 1. Zeit vom 1-31.12; Zeit 2. 1-31.12) existieren, wird der Block genommen, der im Datensatz zuerst kommt.

2.2.2.3 Zeitliche Überlappung zweier Blöcke

Überlappen Blöcke zeitlich, wobei keine Angaben zur Variable EGPT (z.B. Überlappung von zwei Anrechnungszeiten,) existieren, wird der Block genommen, auf den mehr Tage des Monats entfallen. Ergeben sich gleich viele Tage im Monat, wird der Block genommen, der zeitlich früher beginnt (z.B. Beginn am 1.1 und Beginn am 3.1; es zieht Block mit Beginn am 1.1).

2.2.2.4 Sonderfall Arbeitgeberwechsel im Monat und parallele Beschäftigung

Ein Sonderfall bildet der Sachverhalt, wenn es innerhalb eines Monats zu einem Arbeitgeberwechsel kommt (BYAT = 10). In diesem Fall liegen zwei Blöcke für den Monat vor, die sich (mit Zeitabstand) aneinander anschließen. Diese werden als **eine** Information gewertet. In diesem Fall werden die monatsbezogenen Entgeltpunkte aus beiden Blöcken (mEGPT, mEGPTAN) addiert und als mEGPT bzw mEGPTAN abgespeichert. Besonders behandelt wird auch die Variable MANZ: Sie ergibt sich aus der Summe der Kalendertage, die durch Beschäftigung im Monat belegt sind (überlappende Tage zählen nur einfach). Die übrigen Verlaufsmerkmale beziehen sich auf den Block, der nach der voranstehenden Prioritätenregelung zieht.

Liegen zwei oder mehrere Beschäftigungen im Monat parallel vor, z.B. durch zwei Teilzeitbeschäftigungen, wird ebenfalls die Summe aus allen Beschäftigungen in den Merkmalen mEGPT und mEGPTAN eingestellt, für MANZ wird genauso, wie oben beschrieben, verfahren (Im Falle vollkommener Parallelität wird also nicht summiert = komplette Überlappung). Die übrigen Verlaufsmerkmale beziehen sich ebenfalls auf den Block, der nach der Prioritätenregelung zieht.

Es wird folgendermaßen verfahren, um die monatsbezogenen Entgeltpunkte zu ermitteln:

1. Errechnen eines Tagesentgelts ($tEGPT_k$, $tEGPTAN_k$) für alle k Blöcke, die im Monat X liegen.
2. Errechnen der belegten Tage im Monat ($MANZ_k$) X für alle k Blöcke, die im Monat X liegen.
3. Errechnen von monatsbezogenen EGPT-Werten ($mEGPT_k$, $mEGPTAN_k$) für alle k Blöcke im Monat X .
4. mEGPT als Summe aller $mEGPT_k$
5. Ermittlung vom MANZ aus allen $MANZ_k$ (Berücksichtigung der Überlappungen)

2.2.3 Wann gilt ein Monat als belegt

Hierbei bleibt noch die Frage offen, ab wann ein Monat als belegt gilt, wenn der Zeitraum auf Monatsangaben umgerechnet wird: Es wird davon ausgegangen, dass ein Tag im Monat genügt, um diesen als belegt anzusehen.

2.2.4 Ausnahme Kindererziehungszeit

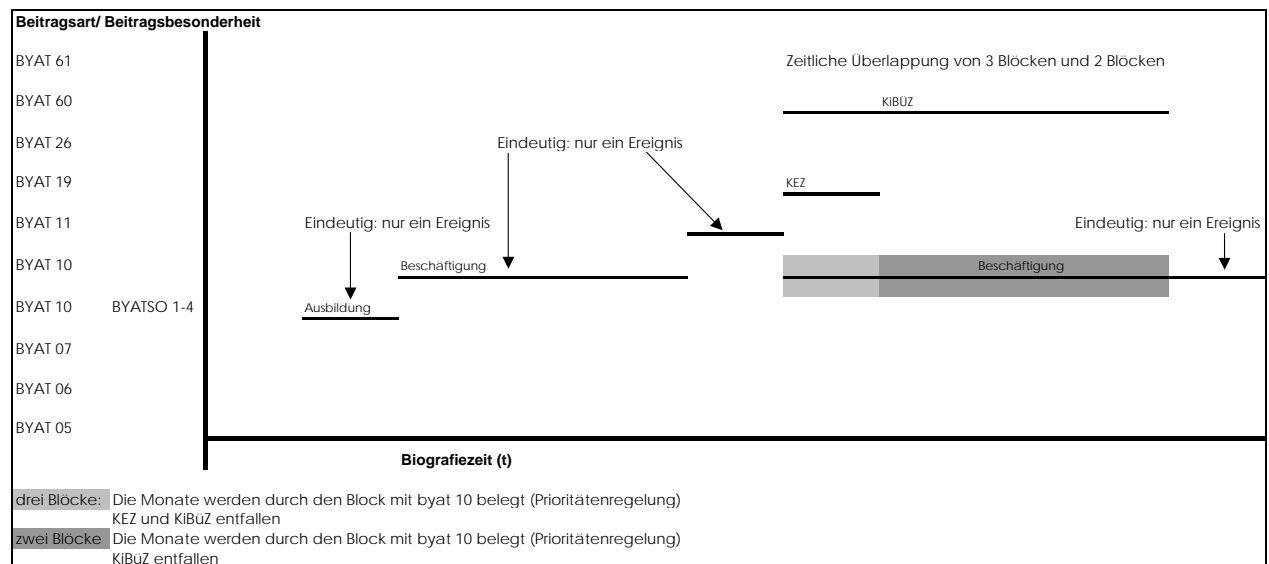
Eine Ausnahme bilden die Beiträge für Kindererziehungszeiten (BYAT 19). Sie werden grundsätzlich nachrangig behandelt. Liegt eine andere Belegung des Monats parallel, ist immer dieser Sachverhalt für die Belegung relevant. Auch die gleichzeitige Erziehung mehre-

rer Kinder (BYAT 61) ist nachrangig zu behandeln. Dies geschieht aber automatisch, da parallel keine anderen Pflichtbeiträge liegen können.

Schaubild 4 zeigt ein vereinfachtes Beispiel für die Umsetzung dieser Regelung.

Schaubild 4

Beispiel für Belegung der Monate nach der Prioritätenregel



Die gewählte Form der Daten hat den Vorteil, dass gängige Statistikpakete (SPSS, SAS u.a.) Module für die Analyse vorgeben. Das für Ereignisdaten genutzte Programm TDA ist in der Lage solche Sequenzdaten mit einem entsprechenden Command-File in Episodendaten umzusetzen und zu analysieren.

2.3 Umsetzung der rentenrechtlichen Biografieangaben

Über die Variablenkombination Beitragsart und Beitragsbesonderheit (BYAT/BYATSO) lässt sich die persönliche Situation im entsprechenden Zeitraum nachzeichnen. Dabei ist aber zu berücksichtigen, dass es sich um Informationen handelt, die für die Rentenberechnung von Belang sind. Die Informationen beschränken sich deshalb sachgemäß auf diese Tatbestände. Zudem ist zu berücksichtigen, dass einzelne Merkmalskombinationen sich durch einen unterschiedlichen zeitlichen Bezug ergeben, bezogen auf die Lebenssituation aber als deckungsgleich zu bezeichnen sind. Die Unterscheidung ergibt sich durch den unterschiedlichen Rechtsstand, der Anwendung findet. Für diesen zeitlichen Bezug lässt sich ein Beispiel anführen: Eine Person, die pflegt (und die gesetzlichen und sozialversicherungs-

rechtlichen Anforderungen erfüllt), hat bis 1.4.1995 BYAT 60 + BYATSO 2, also eine Berücksichtigungszeit für Pflege. Ab dem 1.4.2005 ist sie unter BYAT 7, Pflichtbeitrag für Pflege, zu finden. Die soziale Situation dieser Person ist in beiden Fällen identisch. Eine weitere soziale Situation, die sich durch diesen zeitlich-rechtlichen Aspekt in verschiedenen BYAT/BYATSO-Kombinationen findet, ist Arbeitslosigkeit.

2.3.1 Einführung von Sozialen Erwerbsituationen (SES)

Zur Datenweitergabe eignen sich die rentenrechtlichen Tatbestände in der Regel nicht. Es ist sinnvoll, eine Rückführung auf soziale Situationen durchzuführen. In den allermeisten Fällen dürfte hier das Hauptinteresse liegen. Aus Sicht des Datenschutzes ist damit bereits ein wesentlicher Schritt zur Anonymisierung unternommen. Es wird die unten stehenden Umschlüsselung der BYAT/BYATSO-Kombinationen in Soziale Erwerbssituationen (SES) verwendet. Diese wurde im Rahmen der Sonderuntersuchung Altersvorsorge in Deutschland (AVID 1996 und 2005) entwickelt (Tabelle 1). Die folgende Systematik lehnt sich an die Vorgehensweise der AVID an und entwickelt diese fort.

Tabelle 1: Zuordnung BYAT/BYATSO zu SES
(ist BYATSO nicht spezifiziert, muss es nicht beachtet werden)

BYAT	BYATSO	SES	Bedeutung	SES
Missing	missing	0		keine Information
40	4,6,7,8,C	1	AZ (Schul)ausbildung	Schulische
41	4,6,7,8,C		AZ (Schul)ausbildung	Ausbildung ¹⁾
42			AZ (Fach)hochschule	
43			AZ (Fach)hochschule	
48	4,6,7,8,C		AZ Schulausbildung	
20	3		Nachzahlung für Schulausbildung	
21	3		nach §207 SGB VI	
10	1,6	2	Pflichtbeitrag berufl. Ausbildung	Berufliche
10	2,7		Pflichtbeitrag berufl. Ausbildung	Ausbildung ²⁾
7		3	Pflege Pflichtbeitrag	Nicht-
20	8		freiwilliger Beitrag Pflege	erwerbsmäßige
21	8		freiwilliger Beitrag Pflege	Pflege
60	2		Berücksichtigungszeit Pflege	
11		4	Pflichtbeitrag Mutterschaft	Kindererziehung
19			Kindererziehungszeit	und
29			Beitrag gleichzeitig mehrere Kin-	Haushalt ³⁾
40	2		Schwangerschaft	
41	2		Schwangerschaft	
48	2		Schwangerschaft	
60	1		Berücksichtigungszeit	
61			Berücksichtigungszeit und Gut-	schrift
12		5	Pflichtbeitrag Reha, bei Kranken-	Arbeitsunfähigkeit
40	1,A		geld u.a.	und
41	1,A		AZ Reha	Krankheit ⁴⁾
48	1,A (evtl. 1a)		AZ Reha	
4		6	Pflichtbeitrag Arbeitslosenhil-	Arbeitslos: Arbeits-
			fe/ALG II	losenhilfe/ALG II
13		7	Pflichtbeitrag Arbeitslosengeld	Arbeitslos: Arbeits-
13/8+			Arbeitslos ⁷⁾ 1.3.90 bis 30.06.92	losengeld
10+40/3	Osten ⁷⁾			
40	3,B	8	AZ Arbeitslos	Arbeitslos: Anrech-
				nungszeit
41	3,B		AZ Arbeitslos	
48	3,B		AZ Arbeitslos	
14		9	Pflichtbeitrag Wehr- od. Zivildienst	Wehr- und
				Zivildienst
5		10	Arbeitgeber Geringfügig	Geringfügig
6			Pflichtbeitrag Geringfügig bei Auf-	beschäftigt
			stockung	
17		11	Pflichtbeitrag Selbständige	Selbständig
60	6,7		Berücksichtigungszeit und Selb-	ständig

25		12	Ausschließlich FZR	
20	0,2,5-7		freiwillige Beiträge	Sonstiges
21	0,2,5-7		freiwillige Beiträge	
40	9		AZ sonstig	
41	9		AZ sonstig	
48	9		AZ sonstig	
30			Ersatzzeit	
31			Ersatzzeit	
9		13	Altersteilzeit Pflichtbeitrag	sozialversiche- rungspflichtige Erwerbstätig
10	0,3,4,5,8,9		Pflichtbeitrag	
15			Pflichtbeitrag Nachversicherung echt	
16			Pflichtbeitrag Nachversicherung fiktiv	
26			Datenaustausch Erhöhung fiktive FZR	
8			Pflichtbeitrag für EU im Beitrittsge- biet	
49			Nachversicherung	
20	1		Freiw. Versicherung Überschrei- ten d. Jahresverdienstgrenze	
21	1			
20	4		Nachzahlung bei	
21	4		Heiraterstattung ⁵⁾	
50, 51		14	Zurechnungszeit	Zurechnungszeit
70,71,72		15	Rentenbezug	Rentenbezug
40,41,48	5		AZ Rentenbezug (vor Umwandlung in Altersrente war diese Zeit Zurechnungszeit)	
90			Bergmannsprämie	(aus eigener Versicherung)
70			Rentenbezugszeit	
18			Vorruhestandsgeldbezug	

1) gelöscht

2)Zeiten einer tatsächlichen Ausbildung, maximal 36 Monate !!

Als Zeiten einer beruflichen Ausbildung sind auch solche erfasst, die als Zweitbeitrag entrichtet wurden. Nach der Prioritätenregelung für Pflichtbeiträge aber nur dann, wenn die erzielten Entgeltpunkte höher sind als die aus dem Erstbeitrag. Fiktive Zeiten einer beruflichen Ausbildung sind nicht als berufliche Ausbildung erfasst. Auch nicht erfasst sind Zeiten einer tatsächlichen beruflichen Ausbildung, die nicht als solche anerkannt werden, da sie unter 10/3 bzw. 10/4 (SK79) gemeldet werden und nicht von den fiktiven Ausbildungszeiten zu unterscheiden sind. In der VSKT 2005 sind dies alle Monate mit tatsächlicher Ausbildung, die die ersten 36 Monate mit beruflicher Ausbildung übersteigen. In den kommenden Jahren sind es dann auch die Monate, die durch die Neuregelung der Anerkennung von Schul- und Berufsausbildung nicht mehr als beitragsgemindert anerkannt werden (werden z.B. 36 Monate Schulausbildung an Schulen mit berufsbildendem Charakter bereits anerkannt, dann sind dies alle Monate mit Berufsausbildung). Diese Zeiten fallen dann auch unter 10/3 und 10/4. Ab 2009 entfällt die Bewertung fiktiver Ausbildung gänzlich, so dass ab diesem Zeitpunkt die Summe aus 10/1 bis 10/4 alle beruflichen Ausbildungszeiten erfasst (10/1+10/2 die zu bewertenden und 10/3 + 10/4 die nicht zu bewertenden, Analog 10/6-10/9).

3) Dies sind Zeiten der Kindererziehung oder der Berücksichtigung von Kindererziehung. Sie bilden jedoch die Zeiten, die im rentenrechtlichen Sinne als Kindererziehung anerkannt werden, nur dann komplett ab, wenn es in dieser Zeit nicht auch zu anderen rentenrechtlichen Zeiten kommt (siehe 2.2).

4) Zeiten, die rückwirkend als Erwerbsminderungsrente anerkannt wurden (Zeitraum zwischen Antragstellung und Rentengewährung) und für die Pflichtbeiträge und eine Anrechnungszeit wegen Rentenbezug abgelegt sind, sind nach der Prioritätenregelung Zeiten der Krankheit oder Arbeitslosigkeit (Vorrang des Pflichtbeitrags).

5) Nachzahlungen für Heiratsersatzung finden sich in der Regel am Beginn der Versicherungsbiografie. In der VVL 2004 betrifft dies 12% der Frauen. In der Regel sind dies Zeiten der Erwerbstätigkeit, deswegen werden die Nachzahlungen auch als solche behandelt. **Achtung:** Nicht quantifizierbar ist der Anteil der Frauen, die eine Heiratsersatzung durchgeführt haben (bis 1967 möglich) und die Beiträge nicht nachgezahlt haben. In diesen Fällen ergibt sich eine Lücke zu Beginn der **Versicherungsbiografie**.

6) entfallen

7) Zeiten des Bezugs einer Lohnersatzleistung nach dem Recht der Arbeitsförderung, worunter nach der Einführung der Arbeitsämter im Beitrittsgebiet das Arbeitslosengeld, die Arbeitslosenhilfe und das Unterhaltsgeld fielen (ab 3.10.1990 auch Altersübergangsgeld und Eingliederungsgeld), sind Anrechnungszeiten (§252a SGB VI). Während des Bezugs dieser Leistungen bestand Versicherungspflicht zur Rentenversicherung. Deshalb finden sich im Versicherungskonto parallel Pflichtbeitragszeiten (byat 10) und Anrechnungszeiten bzw. Zweitbeitrag wegen Arbeitslosigkeit (byat 40/byatso 3 und byat 13/byatso 8). Gesetzlich war dies zunächst bis zum 31.12.1991 beschränkt, wurde aber bis zum 30.06.1992 verlängert. Im SUFVSKT finden sich bis zum Jahr 2006 in den Verlaufsmerkmalen MEGPT nur die Anwartschaften, die für den Erstbetrag gespeichert wurden. Die Anwartschaften für diese Arbeitslosigkeitsphasen insgesamt finden sich in MEGPTAN und GMEGPT sowie GMEGPTAN. AB dem Jahr 2007 finden sich auch unter MEGPT die Gesamtanwartschaften.

2.3.2 Sondermerkmal Geringfügigkeit

Es wird ein Sondermerkmal zur Erfassung der geringfügigen Beschäftigung eingeführt (NJOB). Liegt im Monat ausschließlich oder ergänzend eine geringfügige Beschäftigung vor, dann wird die Variable auf 1 gesetzt. Geringfügige Beschäftigung liegt vor, wenn BYAT mit 5 oder 6 belegt ist.

2.3.3 Belegte Tage im Biografienmonat

Das Verlaufsmerkmal „Belegte Tage im Monat“ (MANZ) dokumentiert, wie viele Tage im jeweiligen Biografienmonat durch die dokumentierte SES belegt sind. Dabei ist der Sonderfall bei Arbeitgeberwechsel innerhalb eines Monats und mehrere parallele sozialversicherungspflichtige Beschäftigungen berücksichtigt (siehe 3.2).

2.3.4 Monatsbezogene Entgeltpunktinformationen

Die Variablen EGPT und EGPTAN des für die SES maßgeblichen Blocks werden auf den Monat bezogen (mEGPT, mEGPTAN). Dazu wird ein tagesbezogener Entgeltpunktwert ermittelt: der jeweilige Wert wird durch die Anzahl der Tage im Gesamtzeitraum des entsprechenden Blocks dividiert. Der tagesbezogene Entgeltpunktwert wird mit der Variablen MANZ (Anzahl der belegten Tage im Monat) multipliziert. Zu beachten sind die Sonderregelungen, wenn es sich um einen Monat handelt, in dem mehrere sozialversicherungspflichtige Beschäftigungen vorliegen.

In der technischen Umsetzung wird der Tageswert als „Double“-Variable gehalten, der Monatswert wird auf die vierte Nachkommastelle gerundet. Im Ergebnis resultieren daraus geringe Rundungsfehler. Durch die unterschiedliche Dauer der Monate (30/31 Tage) ergeben

sich bei voll belegten Monaten geringfügige Schwankungen (im Kalendermonat Februar fallen diese etwas größer aus).

Beispiel: Belegter Zeitraum 01.01 bis 31.03; 0,25 Entgeltpunkte

- Tage insgesamt: 90
- Tagesentgeltpunktwert: 0,002777...
- Jan=0,086087...; 0,0861 (auf vier Stellen gerundet)
- Feb=0,077756..., 0,0778 ""
- Mrz=0,086087...; 0,0861 ""
- Zeitraumwert= 0,2500

Zusätzlich werden die Merkmale gmEGPT und gmEGPTAN ermittelt und weitergegeben. In diesen Merkmalen wird die Summe aller monatsbezogenen Entgeltpunkte abgelegt, die sich für diesen Monat ergeben. Es werden damit auch die Entgeltpunkte berücksichtigt, die sich für die Blöcke ergeben, die nicht die SES bestimmen. Damit werden alle aus rentenrechtlicher Sicht für diesen Monat relevanten Entgeltpunktinformationen weitergegeben.

Zur Berechnung werden dabei für jeden Block, der auf den Monat fällt, die monatsbezogenen Entgeltpunkte ermittelt. Diese Werte werden als Double-Wert gehalten, also nicht gerundet, und aufsummiert. Die abschließende Summe wird auf vier Nachkommastellen gerundet.

2.3.5 Sondermerkmal Kumulation von Beschäftigungen

Wie unter 3.2.2.4 beschrieben werden die Entgeltpunktinformationen addiert, wenn mehrere Beschäftigungen in einem Monat zusammentreffen. Die übrigen Verlaufsmerkmale richten sich nach dem Block, welcher der Prioritätenregelung folgend zu belegen ist. Um solche Blöcke identifizieren zu können, wird ein Verlaufsmerkmal (JKUM) eingeführt, das mit eins belegt wird, wenn im entsprechenden Monat mindestens zwei Blöcke mit sozialversicherungspflichtiger Beschäftigung (BYAT 10) liegen.

2.3.6 Sonderfall Zurechnungszeit und Ende der Biografie

Eine Zurechnungszeit ergibt sich bei einer Erwerbsminderungsrente. Dabei handelt sich um keine empirisch tatsächlich erlebte Zeit, soweit diese nach dem Berichtsjahr liegt. Sie ist dann ein rechtliches Konstrukt zur Ermittlung der Erwerbsminderungsrente. Wenn die Zurechnungszeit in den Zeitraum vor der Erhebung fällt, ist dieser Teil aber als SES „Erwerbsgemindert“ zu qualifizieren, denn dann entspricht sie dem sozialen Status in dem bereits er-

lebten Zeitraum. Das Ende des Beobachtungszeitraums kann über das Merkmal BJDEZ eingegrenzt werden.

3 Umsetzung in feste Satzlänge

3.1 Merkmale

Grundsätzlich werden folgende Merkmale aus dem variablen Teil des SK79 als Monatsangaben weitergeben:

VSGR	(lässt Unterscheidung pflichtversicherter Handwerker zu)
SES	BYAT+BYATSO als Umschlüsselung aus Tabelle 1
PFLEGE	Flag: Nichterwerbsmäßige Pflege im Monat X
KRANK	Flag: Arbeitsunfähigkeit oder Rehabilitation im Monat X
ALOS	Flag: Arbeitslos im Monat X
KI	(Rückschlüsse auf parallele Kindererziehung)
GM	(Rückschlüsse auf beitragsgeminderte Zeiten)
RCEG	(als Hilfsmerkmal für BHBR)
RTVS/KZOST	(Paralleler Rentenbezug / Ost-Entgeltpunkte)
BHBR	(Beschäftigungsbereich)
mEGPT	(Entgeltpunkte für Monat X bezogen auf die SES, !Besonderheit Doppel- oder Mehrfachbeschäftigung im Monat)
mEGPTAN	(Angerechnete Entgeltpunkte für Monat X bezogen auf die SES, !Besonderheit Doppel- oder Mehrfachbeschäftigung im Monat)
gmEGPT	(Entgeltpunkte insgesamt für Monat X)
gmEGPTAN	(Angerechnete Entgeltpunkte insgesamt für Monat X)
NJOB	Dokumentation Geringfügige Beschäftigung
MANZ	Anzahl der durch die SES belegten Tage im Monat; !Besonderheit Doppel- oder Mehrfachbeschäftigung im Monat
JKUM	Vorliegen von mehreren sozialversicherungspflichtigen Beschäftigungen (BYAT 10) im Monat
Kind3	Anzahl der Kinder 36 Monate und jünger im Monat X
Kind12	Anzahl der Kinder 144 Monate und jünger im Monat X
FZR	Flag: Versicherung in der freiw. Zusatzverorgung (DDR)
SDDR	Flag: Beiträge nach §256a Abs. 3 SGB VI

3.2 Datensatzstruktur

Die Daten werden analog der Biografiezeit als Monatsangaben abgespeichert. Dabei werden Die Merkmale geblockt von 001 bis 624 abgelegt:

VSGR001	VSGR002	...	VSGR624	SES001	SES002	...	SES624	KI001	KI002	...	KI624	...	GM001	GM002

... fortlaufend bis 624

Zusätzlich wird die Fallnummer abgespeichert. Der fixe Teil des Datensatzes SK79 wird abgespalten und separat gespeichert. Ein Bezug ist über die Fallnummer möglich. Welche Daten des fixen Teils weitergegeben werden, wird gesondert geregelt.

4 Datenweitergabe von Merkmalen des fixen Teils

Bei der VVL handelt es sich bei den Altersrenten um rentenrechtlich abgeschlossene Biografien und bei den Erwerbsminderungsrenten um tatsächliche Rentenzugänge aufgrund Erwerbsminderung. Deshalb können auch Variablen aus der Rentenberechnung weitergegeben werden. Die Merkmale zur Rentenberechnung und einige ergänzende soziodemografische Merkmale werden aus dem Datensatz entnommen, der für die Fälle jeweils als Bestandteil der Rentenzugangstatistik gemeldet wird (SK90). Zusätzlich werden aus dem fixen Datenteil des Sk79 Merkmale hinzugefügt, die im Rentenzugangsdatsatz nicht enthalten sind.

Die genaue Beschreibung ist dem Codeplan (**Anlage**) zu entnehmen.

5 Eingrenzung und Hochrechnung der Stichprobe

Die VVL ist eine systematische Zufallsauswahl aus dem Rentenzugang mit Meldegrund 10 (Festsetzung ohne unmittelbar vorhergehenden Rentenbezug aus einer gesetzlichen Rentenversicherung). Es werden nur ausgewählte Leistungsarten berücksichtigt (Erwerbsminderungs- und Altersrenten). Außerdem darf es sich nur um Nichtvertragsrenten handeln. Die VVL ist brutto eine 20%-Stichprobe dieser Fälle des Rentenzugangs gezogen.

Es wird aus der vorhandenen Netto-Stichprobe eine zufällige Substichprobe von 25% gezogen. Dabei werden nur Fälle der Jahrgänge berücksichtigt die im Berichtsjahr zwischen 30 und 65 Jahre alt werden. Daraus ergibt sich eine 5%-Stichprobe für die genannten Jahrgänge.. Da es sich um eine proportionale Zufallsstichprobe ohne besondere Schichtung handelt, kann von einem Hochrechnungsfaktor abgesehen werden.

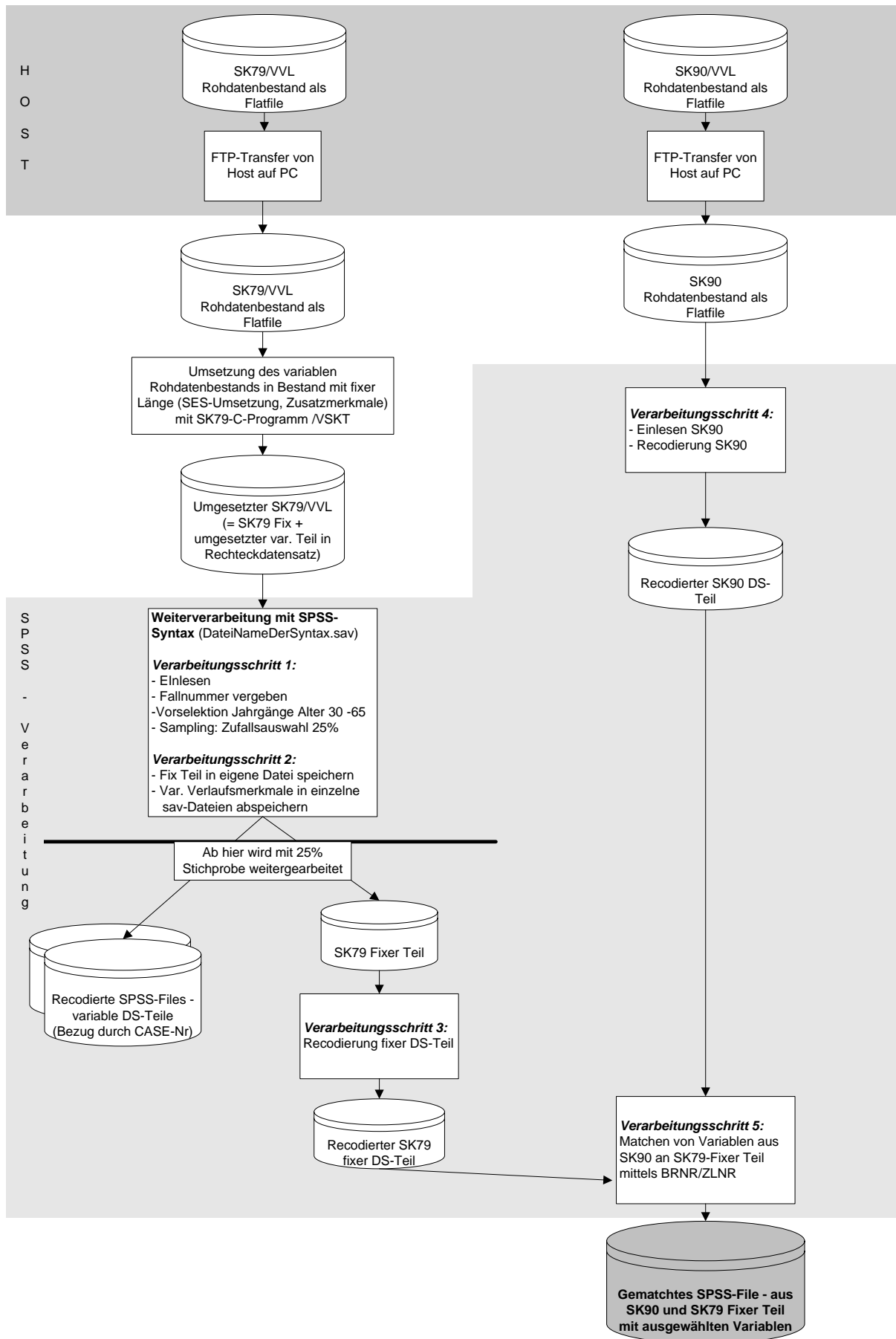
6 Ablaufschema der Datengenerierung

Die beiden folgenden Schemata illustrieren die Produktion der Stichproben der VVL 2004 und der VSKT für das FDZ. Als Endprodukt werden folgende Dateien erzeugt:

Fixer Datenteil (Strukturmerkmale, Soziodemografie, Summenmerkmale der Rentenberechnung)	Biografiebezogene Verlaufsmerkmale
SUFVVVl“Jahr“_fix	SUFVVVl“Jahr“_“Merkmalsname“ (z.B. SUFVVVl2007_SES)

Im ersten Schritt werden die Daten zur VVL, Verlaufsdaten mit fixem Rentenberechnungsteil (SK79) und die Daten zum Rentenzugang (SK90) via FTP vom Großrechner auf den Verarbeitungsrechner gespielt. Mit einem C-Programm wird anschließend der variable Teil des SK79 in ein Datenformat mit fixer Länge transferiert, dabei findet auch die SES-Umsetzung statt. Im Anschluss wird eine eindeutige Fallnummer für jede Untersuchungsperson vergeben und die Stichprobenziehung durchgeführt. Da der Endnutzer in der Regel nur einen Teil der biografiebezogenen Verlaufsmerkmale für seine Analysen benötigt, werden die einzelnen Verlaufsmerkmale in separate Datenfiles aufgeteilt (die eindeutige Fallnummer wird weitergeben). Im nächsten Schritt wird der fixe Teil des SK79 in SPSS eingelesen und dort bearbeitet (Rekodierungen). Danach werden die Daten zum Rentenzugangsgeschehen (SK90) in SPSS eingelesen und rekodiert (einschließlich Zusammenfassung der Werte zur Rentenberechnung). Anhand der eindeutigen Identifikation durch Bereichsnummer und Zählnummer werden ausgewählte Merkmale des SK90 (Werte zur Rentenberechnung) an den modifizierten Datenteil Sk79 fix gespielt.

Ablaufschema VVL



7 Anonymisierung und Datenschutz

Die vorgeschlagene Aufbereitung der Längsschnittdaten ergibt einen Datensatz, der faktisch anonymisiert ist. Die faktische Anonymisierung der Verlaufsdaten wird im Wesentlichen durch folgende Schritte erreicht:

1. Die detaillierten rentenrechtlichen Längsschnittinformationen, die sich aus den Variablen BYAT und BYATSO ergeben, werden reduziert auf wenige Zustände (so genannte SES). Selbst die Kenntnis über rentenrechtliche Sachverhalte mit zeitlichem Bezug kann damit nicht mehr zur Re-Identifikation herangezogen werden.
2. Aus dem fixen Teil des SK79 werden nur Merkmale zur soziodemografischen Profilierung weitergegeben. Diese werden in ihren Ausprägungen den allgemeinen Anonymisierungsregeln des FDZ folgend reduziert. Bei der VVL werden zusätzlich Merkmale der Rentenberechnung aus dem Zugangsdatensatz SK90 weitergegeben.
3. Es wird eine 25%-Substichprobe weitergegeben. Zusätzlich findet eine Altersbegrenzung bei den Untersuchungspersonen von 30 bis 65 Jahren statt. Bei der VVL 2004 ergibt sich z.B. eine Substichprobengröße von ca. 41.000 Fällen, Bereits die Weitergabe als Substichproben und deren Umfang macht eindeutige Re-Identifikation nahezu unmöglich.