

FDZ-Biografiedatensatz
aus der
Versicherungskontenstichprobe
(SUFVSKT2007_Fix und SUFVSKT2007_“*Verlaufsmerkmal*“)

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

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

Stand: 30.04.2010

Inhalt

1	DIE ERHEBUNG DER VERSICHERUNGSKONTENSTICHPROBE (VSKT)	3
1.1	ALLGEMEINES.....	3
1.2	STICHPROBE	3
1.3	PANELPFLEGE	6
1.3.1	<i>Nachziehen von Panelfällen</i>	6
1.3.2	<i>Ausscheiden von Panelfällen</i>	6
1.4	MELDEWEG UND MELDEZEITPUNKT	7
1.5	ERLÄUTERUNGEN ZUM ORIGINALDATENSATZ ZUR VSKT	9
1.5.1	<i>Satzaufbau</i>	9
1.5.2	<i>Erläuterungen zum Datensatz</i>	9
1.6	HOCHRECHNUNGSRAHMEN	10
1.7	HOCHRECHNUNG DER STICHPROBE DER VSKT	10
1.8	SUBSTICHPROBE	12
2	SK 79 ALS BIOGRAFIEDATENSATZ	12
2.1	SEQUENTIELLE BIOGRAFIEDATEN	12
2.2	ZEITLICH PARALLELE UND ÜBERLAPPENDE SOWIE IM MONAT AUF EINANDER FOLGENDE BLÖCKE..	15
2.2.1	<i>Zuordnungshierarchie</i>	15
2.2.2	<i>Prioritätenregelung bei gleicher Hierarchie</i>	16
2.2.2.1	Parallele, überlappende oder im Monat aufeinander folgende Pflichtbeiträge	16
2.2.2.2	Zeitliche Parallelität ohne Beiträge	16
2.2.2.3	Zeitliche Überlappung zweier Blöcke.....	16
2.2.2.4	Sonderfall Arbeitgeberwechsel im Monat und parallele Beschäftigung.....	16
2.2.3	<i>Wann gilt ein Monat als belegt</i>	17
2.2.4	<i>Ausnahme Kindererziehungszeit</i>	17
2.3	UMSETZUNG DER RENTENRECHTLICHEN BIOGRAFIEANGABEN	18
2.3.1	<i>Einführung von Sozialen Erwerbssituationen (SES)</i>	19
2.3.2	<i>Sondermerkmal Geringfügigkeit</i>	21
2.3.3	<i>Belegte Tage im Biografiemonat</i>	21
2.3.4	<i>Monatsbezogene Entgeltpunktinformationen</i>	21
2.3.5	<i>Sondermerkmal Kumulation von Beschäftigungen</i>	22
2.3.6	<i>Zurechnungszeit und Ende der Biografie</i>	23
2.3.6.1	Fälle mit rein fiktiven Zurechnungszeiten.....	23
2.3.6.2	Fälle mit echten und erlebten Zurechnungszeiten.....	23
2.3.6.3	Zensurierung des Datensatzes	23
3	RENTENBERECHNUNG UND SUMMENMERKMALE ZUR GRV-RENTE	24
3.1	RENTENRECHTLICHE BEWERTUNG DER BIOGRAFIE	24
3.2	HINTERGRÜNDE UND KONSEQUENZEN FÜR DIE INTERPRETATION	25
4	UMSETZUNG IN FESTE SATZLÄNGE	27
4.1	MERKMALE.....	27
4.2	DATENSATZSTRUKTUR	27
5	DATENWEITERGABE VON MERKMALEN DES FIXEN TEILS	28
6	EINGRENZUNG UND HOCHRECHNUNG DER STICHPROBE	28
7	ABLAUFSCHHEMA DER DATENGENERIERUNG	28
8	ANONYMISIERUNG UND DATENSCHUTZ	29

1 Die Erhebung der Versicherungskontenstichprobe (VSKT)

1.1 Allgemeines

Die VSKT informiert über die Versicherten der deutschen gesetzlichen Rentenversicherung und den Stand ihrer Rentenanwartschaften. Sie liefert Informationen über sämtliche (gespeicherten) rentenrelevanten Tatbestände. Die Ergebnisse dienen als Grundlage für Planungsaufgaben des Gesetzgebers, interne Planungen der Versicherungsträger und Information von Presse und Öffentlichkeit. Insbesondere soll die Statistik Informationen zum Zwecke der Vorausschätzung der finanziellen Entwicklung in der Rentenversicherung bereitstellen.

Die Berichterstattung wurde in Form einer geschichteten Zufallsstichprobe aus den Versicherungskonten erhoben und in den Folgejahren als Panel fortgeführt und gepflegt. Eine Panel-Erhebung wurde gewählt, um einerseits auf geklärte Konten zurückgreifen zu können, andererseits aber den Aufwand der Kontenklärung möglichst gering zu halten.

Als Versicherte zählen in der Statistik alle diejenigen Personen, für die bei einem Träger der deutschen Rentenversicherung ein Versicherungskonto geführt wird, das

- am Auswertungstag nicht still- oder totgelegt ist,
- für Zeiten bis zum Stichtag (31.12. des Berichtsjahres) Beitragszeiten enthält (d. h. das Konto ist nicht "leer") oder für das ein Bonus aus einem Versorgungsausgleich gespeichert ist,
- keinen Eintrag enthält, dass der Versicherte am Stichtag (31.12. des Berichtsjahres) bereits verstorben ist,
- eine Person betrifft, die am Stichtag (31.12. des Berichtsjahres) mindestens 15 und höchstens 67 Jahre alt ist.

Die so abgegrenzten Versicherten stellen die **Grundgesamtheit der Erhebung** dar.

Ein Versicherungskonto gilt für diese Statistik als geklärt, wenn es mit oder ohne Mitwirkung des Versicherten geklärt worden ist und das BIS-Datum der Kontenklärung vom Stichtag (31.12. des Berichtsjahres) an höchstens 6 Jahre zurückliegt.

1.2 Stichprobe

Aus dem Stammsatzbestand wurde im Jahr 1983 eine geschichtete Zufallsstichprobe mit nicht-proportionalen Auswahlwahrscheinlichkeiten gezogen und seitdem in Form eines Panels laufend gepflegt. Zielstellung der Stichprobenzusammensetzung war in den einzelnen Schichten relativ gleiche Fallzahlen zu erhalten. Schichtungsmerkmale sind dabei die Zielgruppe (Geschlecht, Staatsangehörigkeit, Versicherungszweig des aktuellen Kontoführers)

sowie der Altersjahrgang. Die Ziehung erfolgt beim GB0500¹ aus dem Stammsatzbestand der DSRV, indem in jeder Zielgruppe und jedem Altersjahrgang von einem zufällig gewählten Anfangswert an jede x-te Versicherungsnummer ausgewählt wird. Die Stammsätze sind wie folgt sortiert, wobei obenstehende Begriffe Vorrang vor weiter unten stehenden Begriffen haben:

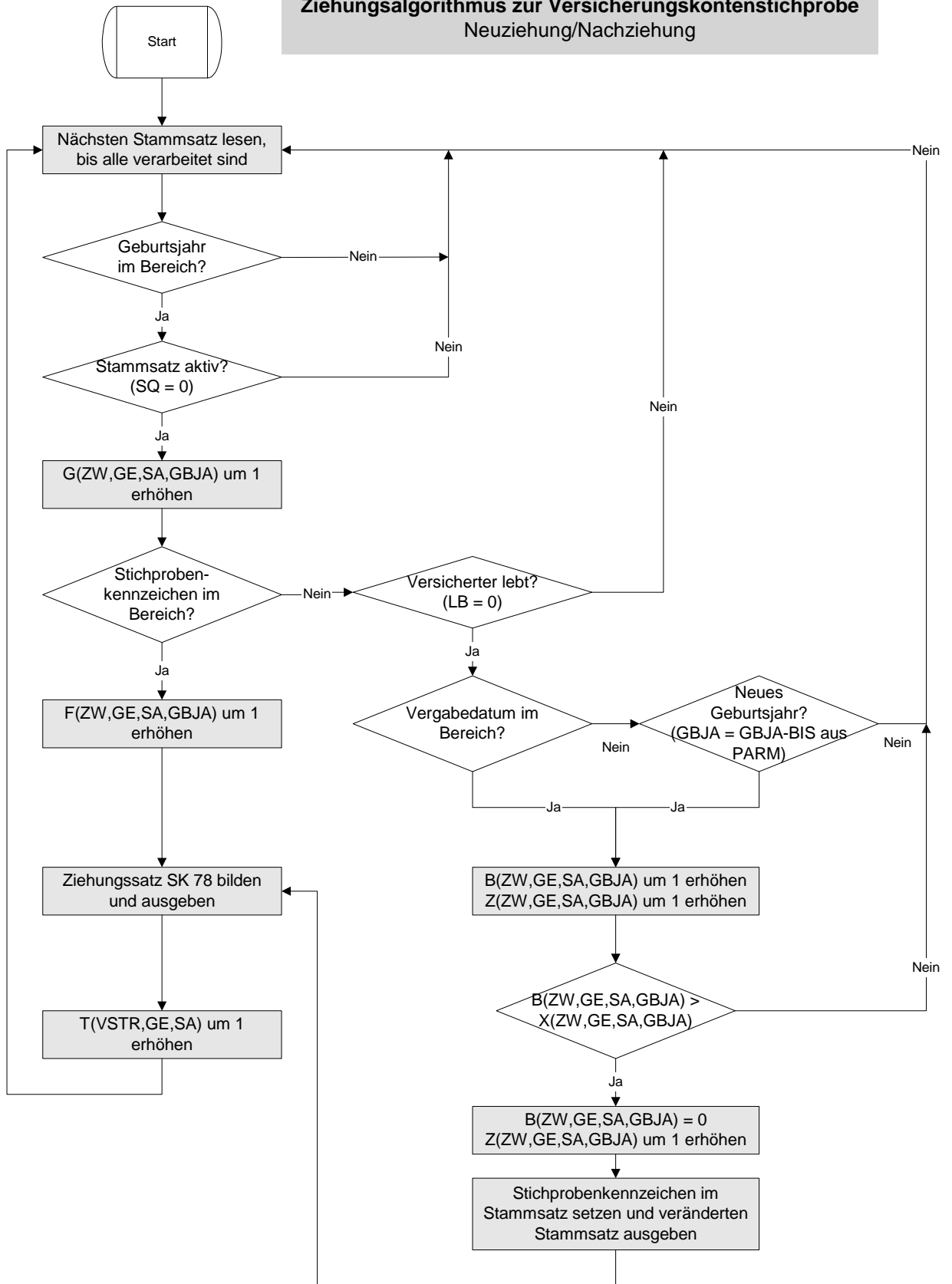
- Geburtsjahr
- Geburtsmonat
- Geburtstag
- Alpha-Zeichen der Versicherungsnummer
- Seriennummer der Versicherungsnummer
- Bereichsnummer der Versicherungsnummer (Vergabeanstalt)

Die Ziehungssystematik lässt bei dieser Anordnung keine Verletzung der Zufälligkeit befürchten. Die für die Stichprobe ausgewählten Fälle werden im Stammsatz mit dem Jahr des Erhebungsstichtages, zu dem sie gezogen wurden, gekennzeichnet. Für die Hochrechnung wird im Stammsatz außerdem abgelegt, welchem Versicherungszweig der Versicherte bei der Ziehung angehört. Bei späteren Erhebungen können die bereits gezogenen Panelfälle an der Kennzeichnung erkannt werden.

Die Ziehung erfolgt dann anhand des nachstehenden Ablaufschemas.

¹ Geschäftsbereich 0500 der Deutschen Rentenversicherung Bund (ehemals VDR Würzburg); nachfolgend GB0500 genannt.

Ziehungsalgorithmus zur Versicherungskontenstichprobe Neuziehung/Nachziehung



Es gelten folgende Abkürzungen:

B (ZW, GE, SA, GBJA):	Ziehungszähler Zählt die seit dem letzten Stichprobenfall nicht ausgewählten Fälle der Ziehungsgrundgesamtheit.
F (ZW, GE, SA, GBJA):	Altfallzähler Zählt die bereits in Stichprobenfälle gekennzeichneten Sätze.
G (ZW, GE, SA, GBJA):	Stichprobengrundgesamtheitszähler Zählt die Fälle, die zur betrachteten Grundgesamtheit gehören.
X (ZW, GE, SA, GBJA):	Schrittweiten Gibt an, jeder wievielte Satz ausgewählt wird; wird von außen vorgegeben.
Y (ZW, GE, SA, GBJA):	Neufallzähler Zählt die neu gezogenen Fälle der Stichprobe.
Z (ZW, GE, SA, GBJA):	Ziehungsgrundgesamtheitszähler Zählt die für die aktuelle Ziehung in Frage kommenden Sätze.

1.3 Panelpflege

1.3.1 Nachziehen von Panelfällen

Damit das Panel auch in den auf die erstmalige Erhebung zum Jahr 1983 folgenden Jahren eine Zufallsauswahl aller Versicherten darstellt, müssen neu hinzugekommene Versicherte anteilmäßig nachgezogen werden. Dies geschieht dadurch, dass die Auswahl bei der Nachziehung auf die neu hinzugekommenen Versicherungsnummern beschränkt wird. Damit die Stichprobe nicht verzerrt wird, sind bei Nachziehungen innerhalb bereits besetzter Gruppen dieselben Auswahlwahrscheinlichkeiten wie bei der ursprünglichen (erstmaligen) Panelziehung zu verwenden. Die nachgezogenen Fälle werden gekennzeichnet, indem das Jahr des Stichtages, zu dem die Fälle nachgezogen wurden, im Stammsatz abgelegt wird. Die Nachziehung wird jeweils zu Jahresbeginn vorgenommen.

1.3.2 Ausscheiden von Panelfällen

Wie aus der Grundgesamtheit können auch aus dem Panel Fälle wegfallen. Dazu kommt es,

- wenn das Konto still- oder totgelegt wird,
- wenn die Beiträge erstattet werden und nicht wieder neue Beiträge eingehen,
- wenn der Versicherte die Altersgrenze übersteigt, d. h. am Stichtag (31.12. des Berichtsjahres) 68 Jahre oder älter ist,
- wenn der Versicherte bis einschließlich des Stichtages (31.12. des Berichtsjahres) verstorben ist.

Daneben können Panelfälle auch vorübergehend, d. h. für eine Erhebung ausfallen,

- wenn die Kontoführung nach der Datenanforderung wechselt,
- wenn das Konto am Auswertungstag nicht auswertbar ist. (Eine Kontosperrung ist kein genereller Hinderungsgrund)

Ein Wechsel der Kontoführung hat nicht das Ausscheiden aus dem Panel zur Folge. Eine Verzerrung der Stichprobe durch Wanderungsfälle wird mittels einer geeigneten Hochrechnung vermieden (nach 2005 im Prinzip nicht mehr nötig, da die Einteilung nach Versicherungszweigen entfällt).

Es kann davon ausgegangen werden, dass sich das Ausscheiden von Versicherten aus der Grundgesamtheit im Panel entsprechend bemerkbar macht. Die Zufälligkeit der Stichprobe bleibt daher erhalten.

Im Stammsatzbestand wird die Kennzeichnung als Stichprobenfall gelöscht, wenn der Stammsatz still- oder tot gelegt wird. Außerdem werden die Kennzeichnungen der Fälle gelöscht, die die obere Altersgrenze überschritten haben.

1.4 Meldeweg und Meldezeitpunkt

Für die Stichprobe zum 31.12. eines Jahres X erhalten die Versicherungsträger jeweils zu Beginn des Jahres X + 1 die Panelfälle (einschl. der nachgezogenen Versicherungsnummern) in Form von Anforderungsdatensätzen (SK 78). Die Aufteilung auf die Versicherungsträger erfolgt nach dem am Auswertungstag im Stammsatz eingetragenen aktuellen Kontoführer. Für die Rentenversicherungsträger in den neuen Bundesländern ist eine Beteiligung erst seit der Erhebung zum 31.12.1995 im Oktober 1996 verbindlich vorgesehen.

Da im Stammsatz nicht sämtliche notwendigen Informationen zur Auswahl der Stichprobenfälle vorhanden sind oder sich diese zwischenzeitlich verändert haben können, prüfen die Versicherungsträger, inwieweit es sich bei den gemeldeten Versicherungsnummern um Fälle der Grundgesamtheit handelt. Versicherungskonten, auf die mindestens einer der nachstehenden Ausschlussgründe zutrifft, werden aus den Stichprobenfällen ausgesondert und nicht weiter betrachtet. Die nachstehende Übersicht nennt die Ausschlussgründe und zeigt, wann sie geprüft werden.

Ausschlussgrund	Prüfung erfolgt durch			
	GB0500		Versicherungsträger	
	bei Ziehung (Jahresbeginn)	bei Datenanforderung (Mitte September)	bei der Meldung zur Kontenklärung (Jahresbeginn)	bei der Auswertung (01.10.)
Konto ist am Verarbeitungstag still- oder totgelegt	ja	ja	ja	ja
Anstalt ist am Verarbeitungstag nicht aktueller Kontoführer	nein	nein	ja	ja
Versicherter ist am Stichtag (31.12.) bereits verstorben	ja	nein	ja	ja
Im Konto sind weder anrechenbare Beitragszeiten gespeichert, noch liegt ein Bonus aus einem Versorgungsausgleich vor	nein	nein	ja	ja
Konto ist nicht auswertbar (gesperrt)	nein	nein	nein	ja

Soweit Konten nicht ausgeschlossen werden, ist zu prüfen, ob sie geklärt sind. Für ungeklärte Konten von Deutschen mit Wohnort in Deutschland, die am Stichtag (31.12. des Berichtsjahres) mindestens 30 Jahre alt sind, ist ein Klärungsverfahren einzuleiten, falls nicht schon eines läuft. Für Versicherte mit Wohnort in den neuen Bundesländern einschließlich des Ostteils Berlins ist dabei eine Beschränkung auf die Alter ab 45 Jahre zulässig. Nicht zu klären sind Konten, die unter einen Ausschlussgrund fallen, sowie Konten von Versicherten mit Wohnort im Ausland.

Mitte September des Jahres X + 1 ermittelt der GB 0500 aus dem Stammsatzbestand die Panelfälle samt aktuellem Kontoführer und fordert die Daten an, indem die Versicherungsträger nochmals die Stichprobenfälle in Form von Anforderungsdatensätzen erhalten.

Die Versicherungsträger prüfen wiederum die Ausschlussgründe (s. o.) und erstellen für die verbleibenden Fälle die Datensätze SK 79 unter der Annahme des Versicherungsfalles "Erwerbsunfähigkeit am Stichtag (31.12. des Jahres X)" einschließlich der Übergangsregelungen (auch bezogen auf die FRG-Rechtsanwendung). Zurechnungszeiten sind dabei anzusetzen.

Vor der Übermittlung an den GB500 sind die Datensätze von den Versicherungsträgern einer Plausibilitätsprüfung zu unterziehen. Der GB0500 stellt hierzu ein Fehlerprüfprogramm als Unterprogramm bereit. Treten in bedeutendem Umfang Fehler auf (mehr als 0,5 %), sind die Ursachen zu ergründen und zu beseitigen. Grundsätzlich werden die Fehlerfälle mit gemeldet und ggf. vom GB0500 ausgesondert. Die Daten werden dem GB0500 mit dem für die Übersendung von Statistikdaten vereinbarten Verfahren zur Verfügung gestellt.

Mit der Lieferung der Datensätze SK 79 stellen die Versicherungsträger dem GB0500 eine Übersicht über die angeforderten, gelieferten und ausgeschiedenen Fälle zur Verfügung.

1.5 Erläuterungen zum Originaldatensatz zur VSKT

1.5.1 Satzaufbau

Der Datensatz SK 79 ist variabel lang und besteht aus einem festen Teil mit Angaben für alle Versicherungszweige und versicherungszweigspezifischen Angaben sowie aus einer individuell verschiedenen Anzahl von Blöcken, die über die einzelnen Zeiten berichten.

1.5.2 Erläuterungen zum Datensatz

Grundsätzlich sind alle nicht beschickten Felder mit Blanks zu füllen. In den nicht beschickten Feldern sind aber auch Nullen zugelassen, sofern damit keine Aussage verbunden ist. Das Feld Reserve ist immer mit "Nullen" zu belegen. Lediglich bei der Sondererhebung "Vollendete Versichertenleben" kann dort die Angabe einer Betriebsnummer erforderlich sein.

Entgeltdaten nach dem Stichtag (31.12. des Berichtsjahres) sind im variablen Teil des Datensatzes, wie auch im festen Teil nicht auszuweisen. Dies gilt auch für unmittelbar entrichtete Beiträge.

1.6 Hochrechnungsrahmen

Für Abschätzungen zur Genauigkeit der Schätzwerte aus der Stichprobe und für die Hochrechnung der Ergebnisse auf die Grundgesamtheit werden die Anzahlen der auf die einzelnen Ziel- und Altersgruppen entfallenden Versicherten benötigt. Diese Informationen werden aus der "Statistik der aktiv Versicherten" und der "Statistik der latent Versicherten" vom GB0500 gewonnen.

1.7 Hochrechnung der Stichprobe der VSKT

Die Stichprobe wurde bei Eröffnung des Panels so gestaltet, dass in den zugrunde gelegten Schichten möglichst gleiche Fallzahlen entstehen.

Um die Versichertengrundgesamtheit zu beschreiben, müssen die Auswertungsergebnisse der Stichprobe deshalb hochgerechnet werden. Da die Fallauswahl in der geschichteten Stichprobe mit unterschiedlichen Auswahlwahrscheinlichkeiten erfolgt und außerdem unplausible Stichprobenfälle von der Auswertung ausgeschlossen werden, müssen die Fälle der Stichprobe bei Auswertungen mit unterschiedlichen Faktoren berücksichtigt werden. Die Hochrechnungsfaktoren für die einzelnen Teilgesamtheiten sind bestimmt durch den Versicherungszweig (bis 2005 relevant), zu dem der letzte Beitrag entrichtet wurde, das Alter, das Geschlecht, die Staatsangehörigkeit (Inländer/Ausländer) und das Versicherungsverhältnis am Stichtag (31.12. des Berichtsjahres).

Bei der Ziehung der Versicherten sowie bei Auswertungen werden bis zum Erhebungsjahr 2005 12 Zielgruppen unterschieden:

Zielgruppe 1:	Männer, Deutsche,	DRV Regionalträger, bis 2004 AR
Zielgruppe 2:	Frauen, Deutsche,	DRV Regionalträger, bis 2004 AR
Zielgruppe 3:	Männer, Ausländer,	DRV Regionalträger, bis 2004 AR
Zielgruppe 4:	Frauen, Ausländer,	DRV Regionalträger, bis 2004 AR
Zielgruppe 5:	Männer, Deutsche,	DRV Bund, bis 2004 AV
Zielgruppe 6:	Frauen, Deutsche,	DRV Bund, bis 2004 AV
Zielgruppe 7:	Männer, Ausländer,	DRV Bund, bis 2004 AV
Zielgruppe 8:	Frauen, Ausländer,	DRV Bund, bis 2004 AV
Zielgruppe 9:	Männer, Deutsche,	DRV KBS, bis 2004 KN

Zielgruppe 10: Frauen, Deutsche, DRV KBS, bis 2004 KN
 Zielgruppe 11: Männer, Ausländer, DRV KBS, bis 2004 KN
 Zielgruppe 12: Frauen, Ausländer, DRV KBS, bis 2004 KN

Für eine durch Versicherungszweig, Geschlecht, Staatsangehörigkeit und Geburtsjahr bestimmte Teilgesamtheit mögen folgende Bezeichnungen gelten:

Y = Zahl der Versicherten der Teilgesamtheit

Bei den Stichprobenfällen einer Teilgesamtheit sind drei Gruppen zu unterscheiden, da Fälle z.B. als AR-Fälle gezogen wurden, in der Zwischenzeit aber den Zweig gewechselt haben:

Gruppe 1: Fälle, die bei der Ziehung zur DRV Regionalträger/AR bis 2004 gehörten

Gruppe 2: Fälle, die bei der Ziehung zur DRV Bund/AV bis 2004 gehörten

Gruppe 3: Fälle, die bei der Ziehung zur DRV KNS/KN bis 2004 gehörten

X_i = Zahl der Stichprobenfälle aus Gruppe i

S_i = Ziehungsschrittweite für Fälle der Gruppe i = $\frac{1}{S_i}$

P_i = Auswahlwahrscheinlichkeit für Fälle der Gruppe i

Ein Fall der Gruppe i wird dann mit nachstehendem Faktor hochgerechnet:

$$P_i \times \frac{Y}{\left[\frac{X_1}{P_1} + \frac{X_2}{P_2} + \frac{X_3}{P_3} \right]} = \frac{Y \cdot S_i}{X_1 S_1 + X_2 S_2 + X_3 S_3}$$

Beispiel:

Idealfall (es kommen keine Wanderungen der Stichprobenfälle zwischen den Zweigen vor):

Y= DRV Bund/Deutsch / Weiblich/ Jahrgang 1940, insgesamt 1000 Fälle

G1= 100 Fälle; P1=10%, Schrittweite= 10

G2= 0 Fälle, P2= 10%, Schrittweite= 10

G3= 0 Fälle, P3= 50%, Schrittweite=10.

$$1000 \cdot 0,1 / 100 \cdot 0,1 = 10$$

Es kommen Wanderungen der Stichprobenfälle zwischen den Zweigen vor):

Y= DRV Bund/Deutsch / Weiblich/ Jahrgang 1940, insgesamt 1000 Fälle

G1= 80 Fälle; P1=10%, Schrittweite= 10

G2= 15 Fälle, P2= 10%, Schrittweite= 10

G3= 5 Fälle, P3= 50%, Schrittweite=10.

$$1000 \cdot 0,1 / 80 \cdot 0,1 + 15 \cdot 0,1 + 5 \cdot 0,5 = 8,33$$

Soweit möglich, wird bei der Hochrechnung von Versicherten ohne Rentenbezug nach folgenden Versichertengruppen unterschieden:

- aktiv Versicherte
- latent Versicherte und
- Übergangsfälle.
- Rentenbezieher

Ist in der Stichprobe eine Fallgruppe nicht besetzt, obwohl es in der Grundgesamtheit Versicherte diesen Typs gibt, so wird die nächst übergeordnete Versichertengruppe für die Hochrechnung herangezogen.

1.8 Substichprobe

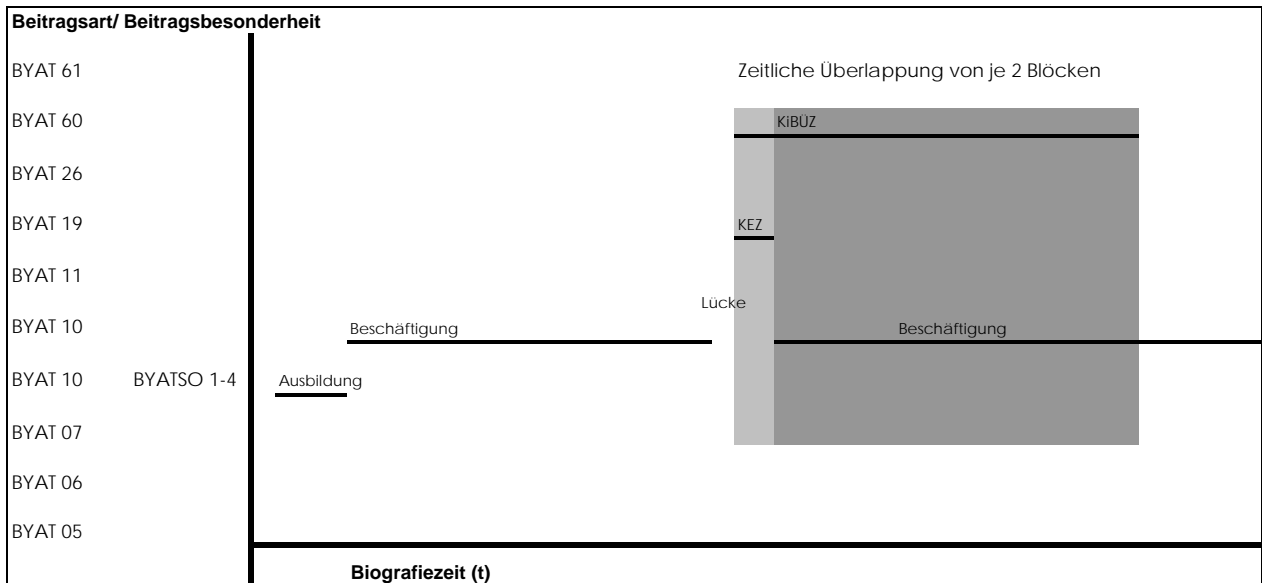
Für den SUF wird aus der beschriebenen Stichprobe eine geschichtete Substichprobe von 25% gezogen. Schichtungskriterien sind der Versicherungszeitpunkt, der Geburtsjahrgang, das Geschlecht und die Nationalität. Diese Kriterien zur Ziehung der Substichprobe entsprechen damit weitgehend denen der Originalziehung. Da die VSKT eine disproportionale Stichprobe darstellt, müssen die Hochrechnungsfaktoren angepasst werden. Dies geschieht durch eine erneute Hochrechnung. Für die Datenweitergabe wird danach die Auswahl anschließend auf die im Inland lebenden Deutschen im Alter 30 bis 67 Jahre begrenzt.

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 ohne chronologische Brüche 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 kaum Informationslücken enthält: 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.

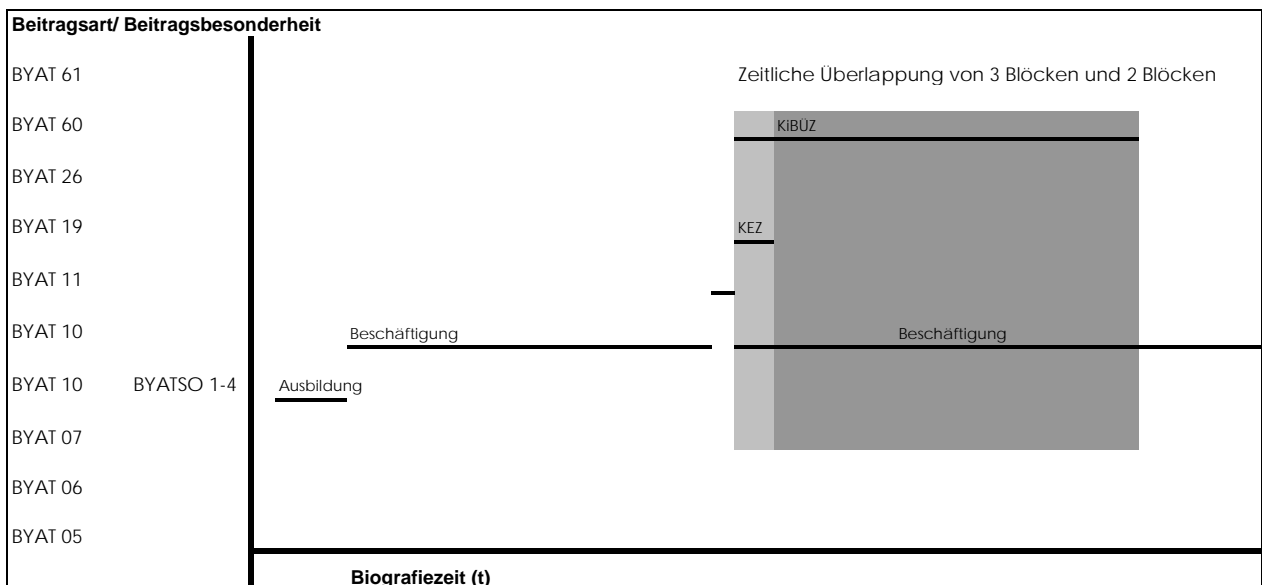
Schaubild1: BEISPIEL 1



(BYAT= Beitragsart)

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



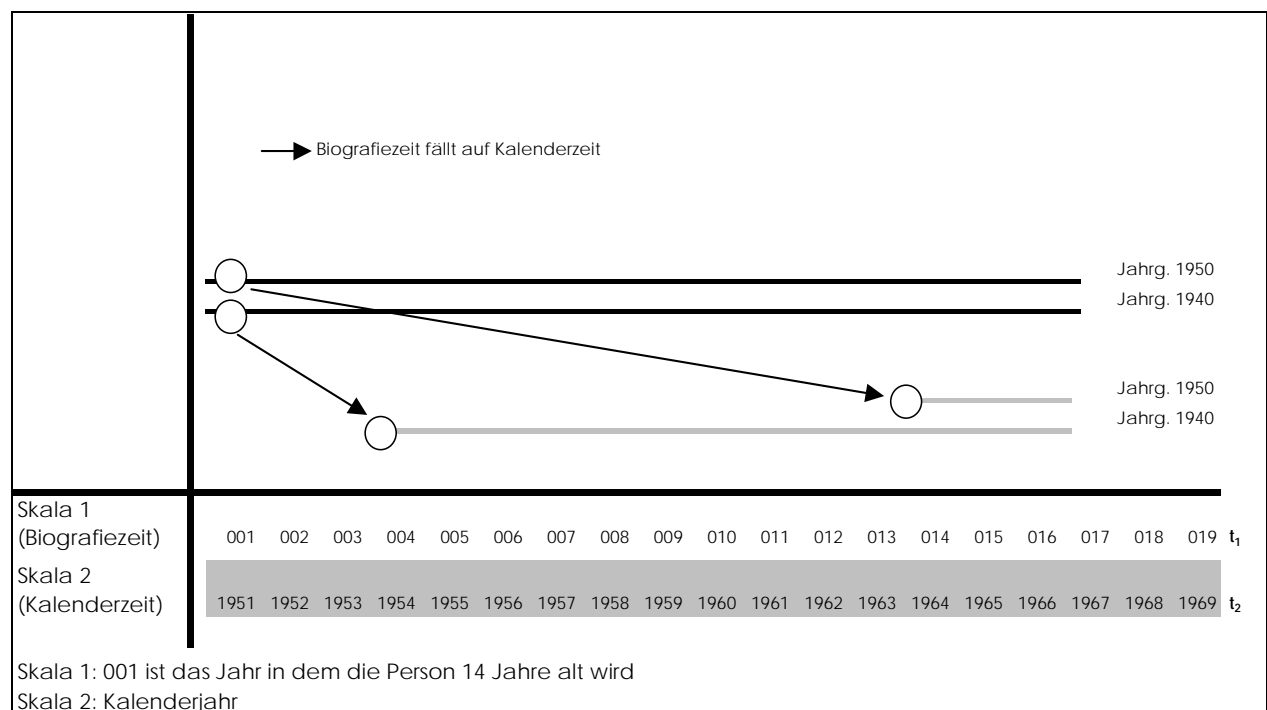
(BYAT= Beitragsart)

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). 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 VSKT 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 und im Monat aufeinander folgenden 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 Beitragsart (BYAT) und Beitragsbesonderheit (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 auf den Monat bezogene Variable EGPT für den Block. 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, z.B. beide vom 01.-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 der Block mit Beginn am 1.1). Analog wird verfahren, wenn die Blöcke im Monat aufeinander folgen.

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 für MANZ nicht summiert = komplette Überlap-

pung). Die übrigen Verlaufsmerkmale beziehen sich ebenfalls auf den Block, der nach der Prioritätenregelung zieht.

Zusammengefasst 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 von $MANZ$ 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.

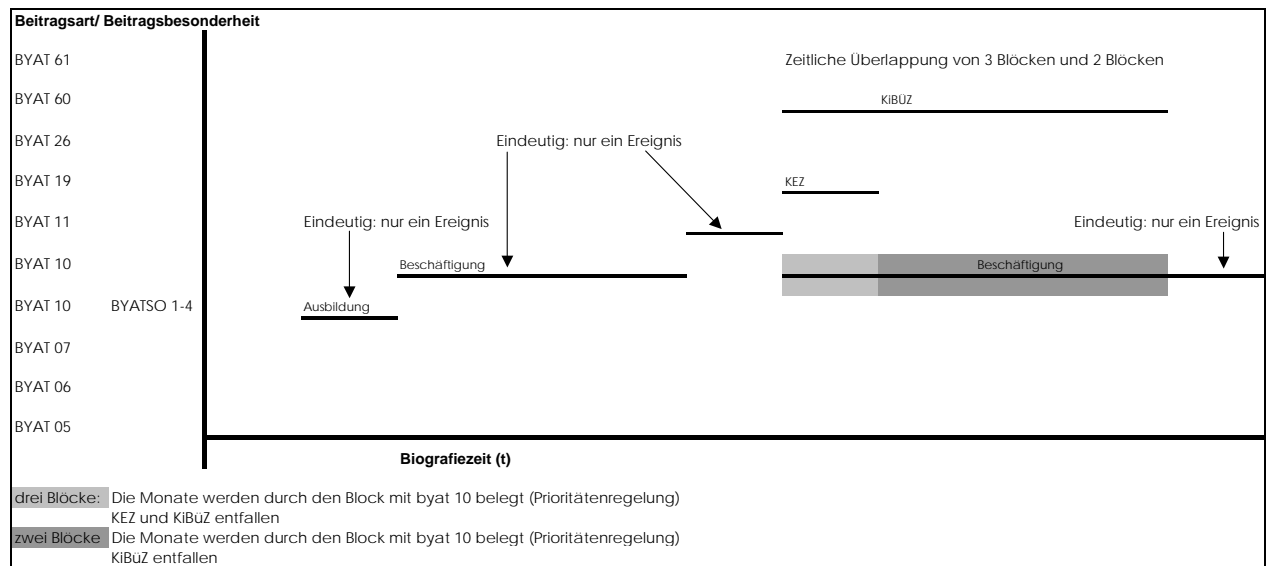
2.2.4 Ausnahme Kindererziehungszeit

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 mehrerer 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 für Ereignisdaten viel genutzten Programme TDA und STATA sind in der Lage solche Sequenzdaten mit einem entsprechenden Command-File oder Do-Files in Episodendaten umzusetzen und zu analysieren.

2.3 Umsetzung der rentenrechtlichen Biografieangaben

Über die Kombination der Merkmale 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 sozialversicherungsrechtlichen 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 in Tabelle 1 stehenden Umschlüsselung der BYAT/BYATSO-Kombinationen in Soziale Erwerbssituationen (SES) zu verwendet. Diese Systematik lehnt sich an die Vorgehensweise der AVID 1996 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+				losengeld
10+40/3	Osten ⁷⁾		Arbeitslos ⁷⁾ 1.3.90 bis 30.06.92	
40	3,B	8	AZ Arbeitslos	Arbeitslos: Anrech-
41	3,B		AZ Arbeitslos	nungszeit
48	3,B		AZ Arbeitslos	

14		9	Pflichtbeitrag Wehr- od. Zivildienst	Wehr- und Zivildienst
5 6		10	Arbeitgeber Geringfügig Pflichtbeitrag Geringfügig bei Aufstockung	Geringfügig beschäftigt
17 60	6,7	11	Pflichtbeitrag Selbständige Berücksichtigungszeit und Selbständig	Selbständig
25 20 21 40 41 48 30 31	0,2,5-7 0,2,5-7 9 9 9	12	Ausschließlich FZR freiwillige Beiträge freiwillige Beiträge AZ sonstig AZ sonstig AZ sonstig Ersatzzeit Ersatzzeit	Sonstiges
9 10 15 16 26 8 49 20 21 20 21	0,3,4,5,8,9 1 1 4 4	13	Altersteilzeit Pflichtbeitrag Pflichtbeitrag Pflichtbeitrag Nachversicherung echt Pflichtbeitrag Nachversicherung fiktiv Datenaustausch Erhöhung fiktive FZR Pflichtbeitrag für EU im Beitrittsgebiet Nachversicherung Freiw. Versicherung Überschreiten d. Jahresverdienstgrenze Nachzahlung bei Heiratserstattung ⁵⁾	sozialversicherungspflichtige Erwerbstätig
50, 51		14	Zurechnungszeit	Zurechnungszeit
70,71,72 40,41,48 90 70 18	5	15	Rentenbezug AZ Rentenbezug (vor Umwandlung in Altersrente war diese Zeit Zurechnungszeit) Bergmannsprämie Rentenbezugszeit Vorruhestandsgeldbezug	Rentenbezug (aus eigener Versicherung)

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.

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 er-

mittelt: 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 diesem 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 2.2.2 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 Zurechnungszeit und Ende der Biografie

Zurechnungszeiten ergeben sich als rentenrechtliche Zeiten bei einer Erwerbsminderungsrente. Zum Zeitpunkt des Rentenbeginns stellen diese rentenrechtlichen Zeiten keine tatsächlich erlebten Biografiepassagen dar. Sie sind vielmehr ein rechtliches Konstrukt: Die Versicherungsbiografie wird dabei mit Zurechnungszeiten vom Zeitpunkt der Erwerbsminderung bis zum 60. Lebensjahr aufgefüllt. Der Wert der Zurechnungszeit im Rahmen der Rentenberechnung ergibt sich dabei aus der Gesamtleistungsbewertung. Es wird quasi so getan, als hätte der Betroffene ab dem Zeitpunkt der Erwerbsminderung bis zum 60. Lebensjahr Anwartschaften erworben, die dem Durchschnitt seiner bisherigen Biografie entsprechen.

Im Rahmen der VSKT sind diese Zeiten am Ende der Biografie zu unterscheiden nach solchen, die zum Erhebungszeitpunkt tatsächlich verlebt wurden und solchen, die in die Zukunft reichen. Dies wird im Folgenden noch näher ausgeführt.

2.3.6.1 Fälle mit rein fiktiven Zurechnungszeiten

Personen, die zum Erhebungszeitpunkt keine Versichertenrente beziehen (PSGR=99), werden in der VSKT so behandelt, als würde zum Zeitpunkt der Erhebung eine (fiktive) Erwerbsminderung eintreten. Dies wird so gehandhabt, um eine Rentenberechnung durchführen zu können. Die Zurechnungszeit ist damit ein fiktives Konstrukt. Die Monate mit SES 14 am Ende der Biografien dieser Personen sind damit keine empirischen Informationen.

2.3.6.2 Fälle mit echten und erlebten Zurechnungszeiten

Für Personen, bei denen eine Erwerbsminderungsrente vorliegt, verhält sich die Situation anders. Zurechnungszeiten in Form von SES 14 bis zum Erhebungsstichtag sind echte und erlebte Zeiten. Die Information ist bis zum Erhebungsstichtag eine empirische Tatsache (Erwerbsminderung). Die Zurechnungszeiten, die über diesen Zeitpunkt hinausgehen sind tatsächliche Zurechnungszeiten, die allerdings in die Zukunft reichen. Man kann vermuten, dass die Person auch weiterhin erwerbsgemindert sein wird; dies ist jedoch nicht zwingend.

Eine Besonderheit bilden in diesem Kontext Personen, deren Zurechnungszeit (SES 14) zum Erhebungszeitpunkt beginnt. Hier liegen echte, aber keine tatsächlich erlebten Zurechnungszeiten vor.

2.3.6.3 Zensurierung des Datensatzes

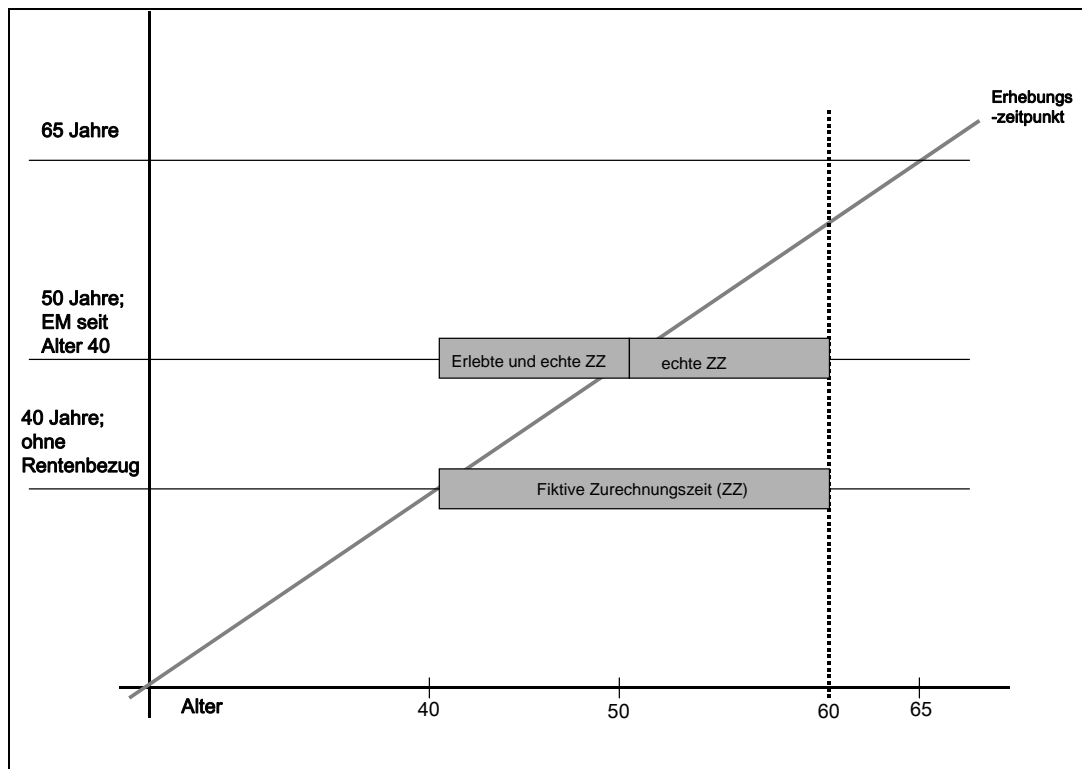
Die Biografie endet mit der letzten rentenrechtlichen Information:

- Ist die Biografie im Monat 624 noch nicht in Altersrente, wird die Biografie auf den Monat 624 begrenzt.

- Ist die Person in EM-Rente oder Altersrente, dann wird die Rentenbezugszeit bzw. Anrechnungszeit wegen Rentenbezug als letzte Information (SES 15) weitergegeben.
- Ist die Person weder in EM-Rente noch in Altersrente (PSGR =99), dann liegt bei ihr eine fiktive Zurechnungszeit als letzte Information (SES 14) vor.
- Sind die Voraussetzungen für eine fiktive EM-Rente nicht erfüllt (PSGR =99), endet die Biografie mit „missings“.

Im folgenden Schaubild 5 ist die Aufteilung der Zurechnungszeiten schematisch dargestellt.

Schaubild 5



3 Rentenberechnung und Summenmerkmale zur GRV-Rente

3.1 Rentenrechtliche Bewertung der Biografie

Grundlage der Rentenberechnung für jeden Fall ist eine Rentenauskunft zum 31.12. des Berichtsjahres. Für jeden Fall wird damit unterstellt, dass er zum 01.01 des folgenden Jahres in EM-Rente geht. Zur Ermittlung wird dabei der Rechtsstand zum 31.12 des Berichtsjahres herangezogen. Die zu diesem Zeitpunkt gültigen rentenrechtlichen Regelungen finden Anwendung. Das gilt damit auch für Übergangsregelungen.

Beispiele:

1.) Die Person im Alter von 63 Jahren ist zum Stichtag 31.12.2005 (VSKT 2005) tatsächlich in Altersrente (Altersrente für Schwerbehinderte). Es wird dieser Person nun eine Rentenauskunft ausgestellt für eine EM-Rente ab 01.01.2006. Bei der rentenrechtlichen Bewertung dieser Biografie werden z.B. auch die neuen Regelungen zur Bewertung der Schulausbildung abgebildet, für den Fall, dass ein Renteneintritt zum 01.01.2006 geschieht. Die Zeiten des Rentenbezugs fließen als Anrechnungszeiten (BYAT 71) nicht in die Bewertung ein; sie beeinflussen den Gesamtleistungswert nicht und sind nicht in der Summenvariable AZ enthalten.

Die tatsächliche Bewertung der Biografie (insbesondere Rechtsstand, auf dessen Basis die Rente ermittelt wurde) und die tatsächliche GRV-Rente sind ohne Bedeutung.

2.) Eine Person im Alter 63 ist zum Stichtag 31.12.2005 tatsächlich in Altersrente (Altersrente für Schwerbehinderte). Sie ist aber bereits seit ihrem 50. Lebensjahr in Rente. Die EM-Rente (bzw. EU-Rente) wurde in eine Altersrente umgewandelt. Da die Altersrente durch Veränderungen im Leistungsrecht geringer ausgefallen wäre, hat bei der Rentenumwandlung der Besitzschutz gegriffen, der eine unverminderte Rentenhöhe garantiert.

Zur Bewertung dieser Biografie gilt dasselbe Vorgehen wie unter 1. beschrieben. Zusätzlich ist zu erwähnen, dass die Neuberechnung auch keinen Besitzschutz berücksichtigt. Auch hier wird so getan, als ob die Person zum 01.01.2006 in die EM-Rente einträte.

3.) Eine Person im Alter 45 ist zum Stichtag 31.12.2005 erwerbstätig. Ihr wird unterstellt, dass sie zum 01.01.2006 in EM-Rente geht. Die Bewertung der Biografie basiert ebenfalls auf diesem Rechtsstand.

Hinsichtlich der rentenrechtlichen Zeiten ist weiterhin zu erwähnen, dass in die Rentenberechnung nur Zeiten in Deutschland und nach Fremdrentenrecht gutgeschriebene Zeiten einfließen. Rentenrechtliche Zeiten aus Vertragsländern bleiben unberücksichtigt.

3.2 Hintergründe und Konsequenzen für die Interpretation

Das Verfahren, jeder Person eine EM-Rente zum 01.01. des Jahres, das auf das Berichtsjahr folgt, zu unterstellen, wurde gewählt, um auch für jüngere Jahrgänge eine Rente ermitteln zu können, die die beitragsfreien Zeiten berücksichtigt. Diese werden durch die Gesamtleistungsbewertung mit Entgeltpunkten versehen.

Vor allen bei den Altersgruppen unter 60 Jahren ist die Systematik der Gewährung einer EM-Rente zu beachten. Die EM-Rente enthält eine so genannte Zurechnungszeit vom Eintritt der Erwerbsminderung bis zum Zeitpunkt der Vollendung des 60. Lebensjahres. Diese Zurechnungszeit wird über die Gesamtleistungsbewertung bewertet. Der Gesamtleistungswert ist der Grundwert oder der Vergleichswert (der höhere von beiden). Diese werden auf die folgende Art berechnet.

Grundwert:

(Entgeltpunkte Beitragszeiten + beitragsgeminderte Zeiten + Berücksichtigungszeiten)

Belegungsfähige Monate im Gesamtzeitraum

Nenner: Vollendung des 17. Lebensjahrs bis Monat Erwerbsminderung (hier) abzüglich beitragsfreier Zeiten soweit nicht gleichzeitig Berücksichtigungszeiten. Monate mit rentenrechtlichen Zeiten vor dem 17. Lebensjahr werden hinzuaddiert.

Vergleichswert:

(Summe Entgeltpunkte Grundbewertung – Abzug besonderer Zeiten*)

Belegungsfähige Monate im Gesamtzeitraum

- * - beitragsgeminderte Zeiten,
- für Berücksichtigungszeiten, die gleichzeitig beitragsfreie Zeiten sind und
- für Beitrags- und Berücksichtigungszeiten während des Bezuges einer Rente aus eigener Versicherung.

Nenner: Entsprechend wird die Summe der belegungsfähigen Monate im Gesamtzeitraum um die Kalendermonate der oben genannten Zeiten vermindert.

Die Gesamtleistungsbewertung ermittelt für die Zurechnungszeiten also Rentenanwartschaften, die die durchschnittliche Belegung der bislang durchlebten Biografie widerspiegeln. Man kann dieses Vorgehen aus Sicht der Bewertung für die Rente auch als einfache Simulation der Biografie bis zum 60. Lebensjahr bezeichnen. Es wird quasi simuliert, dass der Durchschnitt des bisherigen Versicherungslebens den zukünftigen Verlauf der Biografie ergibt.

Dieses Vorgehen hat Konsequenzen für die Interpretation der Ergebnisse. Bei Kohortenvergleichen finden sich in der Höhe und Zusammensetzung von Rentenansprüchen die Auswirkungen der (fortgeschriebenen) Biografien. Ein Einfluss des Leistungsrechts auf unterschiedliche Rentenhöhen, wird mit der fiktiven und damit auf dem gleichen Rechtsstand basierenden Rentenbewertung bis auf die gesetzlichen Übergangsregelungen ausgeschaltet.

4 Umsetzung in feste Satzlänge

4.1 Merkmale

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

VSGR	(lässt Unterscheidung pflichtversicherte Handwerker zu)
SES	BYAT+BYATSO als Umschlüsselung aus Tabelle 1
KI	(Rückschlüsse auf parallele Kindererziehung)
GM	(Rückschlüsse auf beitragsgeminderte Zeiten)
RCEG	(als Hilfsmerkmal für BHBR)
RTVS	(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
PFLEGE	Flag: Nichterwerbsmäßige Pflege im Monat X
KRANK	Flag: Arbeitsunfähigkeit oder Rehabilitation im Monat X
ALOS	Flag: Arbeitslos im Monat X

4.2 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 (siehe Codeplan).

5 Datenweitergabe von Merkmalen des fixen Teils

Bei der VSKT handelt es sich zum kleineren Teil um rentenrechtlich abgeschlossene Biografien und in der Mehrzahl um solche, die noch nicht abgeschlossen sind. Die Merkmale des fixen Teils, die Summenmerkmale und die Gesamtanwartschaften enthalten, basieren damit zum Teil auf fiktiven Werten. Bei allen Fällen wird eine Erwerbsminderung unterstellt, um die Biografie rentenrechtlich bewerten zu können. Bezugsjahr für den Eintritt der Erwerbsminderung ist das Erhebungsjahr.

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

6 Eingrenzung und Hochrechnung der Stichprobe

Es wird aus der vorhandenen Netto-Stichprobe eine zufällige, geschichtete Substichprobe von 25% gezogen. Für die Datenweitergabe werden nur im Inland lebende Deutsche im Alter von 30 bis 67 Jahren berücksichtigt. Es ergibt sich damit z.B. eine Stichprobengröße in der VSKT 2005, die **59.457 Fälle** umfasst. Da es sich um eine disproportionale Zufallsstichprobe handelt, werden die Hochrechnungsfaktoren hierauf angepasst (vgl. Abschnitt 1.8).

7 Ablaufschema der Datengenerierung

Das folgende Schema illustriert die Produktion der Stichproben der VSKT für das FDZ-RV. Als Endprodukt werden folgende Dateien erzeugt:

Fixer Datenteil (Strukturmerkmale, Soziodemografie, Summenmerkmale der Rentenberechnung)	Biografiebezogene Verlaufsmerkmale
SUFVSKT2005_fix	SUFVSKT2005_“Merkmalsname“ (z.B. SUFVSKT2005_SES)

Im ersten Schritt werden die Daten zur VSKT-Verlaufsdaten mit fixem Rentenberechnungsteil (SK79) 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 Dateien aufgeteilt (die eindeutige Fallnummer wird weitergeben). Im nächsten Schritt wird der fixe Teil des SK79 in SPSS eingelesen und dort bearbeitet (Rekodierungen).

8 Anonymisierung und Datenschutz

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

1. Die detaillierten rentenrechtlichen Längsschnittinformationen, die sich aus den Variablen BYAT und BYATSO ergeben, werden reduziert auf 15 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 Merkmale zur soziodemografischen Profilierung und zu rentenrechtlichen Zeiten sowie zur Rentenberechnung weitergegeben. Diese werden in ihren Ausprägungen den allgemeinen Anonymisierungsregeln des FDZ folgend reduziert.
3. Es wird eine Substichprobe weitergegeben. Zusätzlich wird eine Altersbegrenzung von 30 bis 67 Jahren und eine Eingrenzung auf im Inland lebende Deutsche bei den Untersuchungspersonen durchgeführt. Bereits die Weitergabe als Substichproben und deren Umfang macht die eindeutige Re-Identifikation nahezu unmöglich.

Ablaufschema VSKT

