

**Einfluss der demografischen Alterung  
auf die Beitragseinnahmen der  
Gesetzlichen Rentenversicherung  
aus sozialversicherungspflichtiger  
Beschäftigung**

# **Einfluss der demografischen Alterung auf die Beitragseinnahmen der Gesetzlichen Renten- versicherung (GRV) aus sozialversicherungs- pflichtiger Beschäftigung**

Doris Söhnlein, Johann Fuchs\*

Der vorliegende Bericht veröffentlicht die Ergebnisse des mit Mitteln des Forschungsnetzwerkes Alterssicherung der Deutschen Rentenversicherung Bund finanzierten Projekts „Beitragsentwicklung der Gesetzlichen Rentenversicherung in Abhängigkeit von der Erwerbsstruktur“. Für die Finanzierung des Projekts danken wir der Deutschen Rentenversicherung Bund.

\* Doris Söhnlein und Johann Fuchs sind Mitarbeiter am Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesagentur für Arbeit, Nürnberg.

## Inhaltsverzeichnis

Zusammenfassung: .....	3
1 Problemstellung .....	5
2 Forschungsansatz .....	6
3 Beschreibung der Daten .....	9
4 Schätzergebnisse für den Einfluss der Altersstruktur der SV- Beschäftigten .....	15
4.1 Modelle mit ersten Differenzen .....	17
4.2 Modelle mit Bestandsgrößen: ein Kointegrationsansatz .....	30
5 Modelle mit der Altersstruktur der Bevölkerung bzw. des Erwerbspersonenpotenzials .....	34
6 Fazit .....	38
Literatur .....	40
Anhang .....	41
Anhang A1: Variablenbeschreibung .....	41
Tabelle A2: Tests auf Stationarität .....	42
Anhang A3: Alternative Modellspezifikationen .....	43
Anhang A4: Korrelationsmatrix Beschäftigte, Bevölkerung, Erwerbspersonenpotenzial .....	47

## **Zusammenfassung:**

Die vorliegende Studie untersuchte die Frage, ob sich die demografische Alterung auf Einnahmen der Gesetzlichen Rentenversicherung (GRV) auswirken könnte. Dazu wurde der Einfluss der Altersstruktur der sozialversicherungspflichtig Beschäftigten auf die Beitragseinnahmen der GRV unter Berücksichtigung der Entwicklung weiterer Faktoren analysiert.

Als Daten standen – neben den Beitragseinnahmen der GRV und den bei der Bundesagentur für Arbeit gemeldeten sozialversicherungspflichtig Beschäftigten – u.a. das Bruttoinlandsprodukt (BIP), der Beitragssatz zur GRV und einige Strukturmerkmale der Beschäftigung, wie Frauenanteil, Teilzeitquote, Ausländeranteil, Qualifikationsstruktur, zur Verfügung. Der Stützzeitraum reichte von 1976 bis 2003. Die Modellbildung beschränkte sich auf die alten Länder.

Mit den zur Verfügung stehenden Variablen wurden unterschiedliche zeitreihenanalytische Modelle gebildet. Weil die meisten Zeitreihen nicht stationär sind, erfolgten die ökonometrischen Schätzungen mit den jährlichen Veränderungen der Variablen.

Für die Berechnungen wurde der Alterseinfluss mit zwei unterschiedlichen Strategien modelliert: Zum ersten wurden Altersklassen als erklärende Variable verwendet. In einem zweiten Ansatz gingen die statistischen Parameter der Altersverteilung (wie Mittelwert, Streuung, Schiefe, Wölbung) als Variable ein.

Die Berechnungen zeigten, dass beide Typen von Altersvariablen statistisch signifikant sind. Das Modell, das den Alterseinfluss mit Altersklassen schätzte, war dem alternativ berechneten Ansatz - mit den Parametern der Altersverteilung als erklärende Variable - allerdings geringfügig überlegen. Insbesondere konnte das flexiblere Alterklassenmodell die Variabilität der Einnahmenentwicklung genauer nachvollziehen.

Von den übrigen Zeitreihen erwiesen sich die Wachstumsrate des BIP, der Beitragssatz zur GRV und die Qualifikationsstruktur (Anteil der Beschäftigten mit Hochschulabschluss) als die einflussreichsten Variablen.

Weil sich insbesondere für eine langfristige Betrachtung Bestände bzw. Niveaus besser eignen als Veränderungen, wurde ansatzweise untersucht, ob im Rahmen einer Kointegrationsmethodologie eine Modellierung mit den Niveaus der Zeitreihen möglich ist. Die ersten Ergebnisse kann man nur bedingt als erfolgversprechend bezeichnen. Trotzdem sollte der Ansatz nicht aus den Augen verloren werden. Wenn man mit einem Fehlerkorrekturmodell Bestände und Veränderungen zusammen in einem Modell schätzt, würden sich sowohl Aussagen über die langfristige Beziehung als auch über die kurzfristige Dynamik treffen lassen.

Ein Projektschritt befasste sich mit der Möglichkeit, ob die ermittelten Befunde auf die Altersstruktur der Bevölkerung bzw. des Erwerbspotenzials übertragbar sind. Auf der Grundlage erster Schätzungen kann auch diese Frage nur eingeschränkt bejaht werden. Es sind insbesondere noch weitere Untersuchungen hinsichtlich der Modellstabilität erforderlich, damit sich der geschätzte Modelltypus für längerfristige „Prognosen“ eignet, bei denen man eine Projektion der Bevölkerung oder des Erwerbspotenzials zugrunde legt.

Darüber hinaus haben sich im Laufe der Projektarbeit eine Reihe offener Fragen ergeben:

Die Beitragseinnahmen enthalten auch Beiträge von nicht-sozialversicherungspflichtig Beschäftigten. Vermutlich wäre eine Modellbildung mit den Beitragseinnahmen aus sozialversicherungspflichtiger Beschäftigung vorzuziehen. Alternativ müssten in den Modellen eine Reihe weiterer Faktoren als erklärende Variable einbezogen werden.

In den hier geschätzten Modellen war der Beitragssatz zur GRV eine exogene Variable. Näher untersucht werden sollte aber vielleicht die wechselseitige Abhängigkeit von Einnahmen und Ausgaben.

Die Analysen basierten auf westdeutschen Zeitreihen. Zu prüfen wäre, wie sehr sich die Modelle ändern, wenn man bis 1990 westdeutsche Werte und ab 1991 gesamtdeutsche Werte zugrunde legt.

Ein methodisches Problem verursacht der Integrationsgrad der Zeitreihen, da nicht alle Zeitreihen vom selben Grad integriert sind.

## 1 Problemstellung

Die weit überwiegende Mehrheit der Demografen geht davon aus, dass die Bevölkerung und das Erwerbspersonenpotenzial (Arbeitskräfteangebot) Deutschlands auf lange Sicht gesehen weniger wird und „altert“. Wahrscheinlich werden den Betrieben hierzulande immer weniger junge und stattdessen immer mehr ältere Arbeitskräfte zur Verfügung stehen.

Neben den direkten Arbeitsmarktwirkungen der Demografie, z.B. auf den Umfang und die Zusammensetzung des Erwerbspersonenpotenzials, sind indirekte Effekte zu erwarten. Die demografischen Veränderungen werden voraussichtlich die wirtschaftliche Entwicklung beeinflussen, die Produktivität, die sozialen Sicherungssysteme usw. Besonders betroffen ist die gesetzliche Rentenversicherung. Ein Aspekt ist die Ausgabenentwicklung der Gesetzlichen Rentenversicherung (GRV), weil voraussichtlich die Zahl der Neurentner und – bei steigender Lebenserwartung - ihre durchschnittliche Rentenbezugsdauer zunehmen werden.

Möglicherweise beeinflusst der demografische Wandel auch die Einnahmen der GRV, obwohl diese sicherlich größtenteils von der Gesamtbeschäftigung und den Einkommen der (sozialversicherungspflichtig) Beschäftigten determiniert werden. Ein unseres Wissens bislang nicht untersuchter Faktor ist der Einfluss der *Altersstruktur* der Beschäftigten auf die Einnahmen der GRV. Denkbar wäre es beispielsweise, dass - bei gleichbleibender Gesamtbeschäftigung - ein Anstieg der Zahl der älteren Arbeitnehmer in den Betrieben zu mehr Einnahmen der GRV führt. Welche Richtung und vor allem welche Stärke der Zusammenhang hat, hängt jedoch von vielen Faktoren ab. Die Senioritätsentlohnung wird wohl genau so eine Rolle spielen wie die Qualifikationsstruktur der Beschäftigten, der Anteil der Teilzeitbeschäftigten und die altersspezifische Arbeitslosigkeit.

Man kann wohl annehmen, dass die Alterung nur eine Determinante neben anderen, wahrscheinlich wichtigeren Einflussgrößen (z.B. die Beschäftigungslage, das Wirtschaftswachstum) ist. Außerdem sind bislang noch keine größeren demografischen Veränderungen in der Bevölkerung und auch nicht bei den Beschäftigten feststellbar. So stieg das Durchschnittsalter der bei der Bundesagentur für Arbeit gemeldeten sozialversicherungspflichtig Beschäftigten zwischen 1976 und 2003 von 36,7 auf 38,95 Jahre. Doch zeigen sich erste Veränderungen im Altersaufbau. Insbesondere

sank der Anteil der Jüngeren seit Anfang der 90er Jahre deutlich. Mit dem älter werden der Baby-Boomer nahm gleichzeitig der Anteil der „mittleren“ Altersgruppe (30-49 Jahre) zu.

Die folgende Analyse untersucht, wie sich diese noch schwachen Veränderungen auf die Einnahmen der GRV ausgewirkt haben, wobei die Haupteinflussfaktoren (wie das Wirtschaftswachstum) in die Analyse als Kovariate eingehen. Unsere Arbeit versteht sich als Vorstudie, mit der die grundsätzliche Machbarkeit einer solchen Analyse aufgezeigt werden soll. Weil die Verhältnisse in den neuen Ländern möglicherweise völlig andere sind, beschränkt sich die Untersuchung in diesem Stadium auf die alten Länder. Damit wird soweit wie möglich der Bruch in den Daten umgangen, den die Wiedervereinigung erzeugte.

Im folgenden Kapitel wird dazu unser Forschungsansatz vorgestellt. Kapitel 3 behandelt die Daten. Die ökonometrischen Analysen enthält das vierte Kapitel. Im fünften Kapitel wird die Frage gestellt, ob auf der Basis des Forschungsansatzes „Prognosen“ möglich sind. Dazu wird in dieser Arbeit zunächst der Zusammenhang der Alterung von Bevölkerung und Erwerbspersonenpotenzial und der Einnahmenseite der GRV analysiert. Im abschließenden Fazit werden denkbare Weiterungen und Verbesserungen des Ansatzes diskutiert.

## 2 Forschungsansatz

Im Rentenversicherungsbericht 2006 werden die Einnahmen der Gesetzlichen Rentenversicherung (GRV) für das Jahr 2005 auf 231,7 Mrd. € beziffert (BMAS 2006). Fast drei Viertel der Einnahmen entfielen 2005 auf Beitragseinnahmen (davon 85 % Pflichtbeiträge), der Rest überwiegend auf den Bundeszuschuss. Seit 1976 sind die Einnahmen aus den alten Bundesländern jährlich fast linear gestiegen (Abb. 1). Vereinfacht ausgedrückt, ergeben sich die Beitragseinnahmen aus dem Beitragssatz zur gesetzlichen Rentenversicherung, dem durchschnittlichen Jahreseinkommen (Bruttolohn) und der Anzahl der Beitragszahler ab.<sup>1</sup> Der Beitragssatz ist die Stellgröße, mit der Einnahmen und Ausgaben - unter Berücksichtigung

---

<sup>1</sup> Zum Verfahren der Vorausberechnung der Einnahmen der GRV siehe das Kap. 3.3 im Rentenversicherungsbericht (BMAS 2006)

der übrigen Einnahmen - zum Ausgleich gebracht werden. Die Lohnsumme aus durchschnittlichem Jahreseinkommen und Beitragszahlern hängt ihrerseits von anderen Faktoren ab, insbesondere natürlich vom Wirtschaftswachstum, der Lohnentwicklung und möglicherweise von Veränderungen in der Zusammensetzung der Beitragszahler. Letztere sind ihrerseits wieder teilweise ein Spiegel der wirtschaftlichen Dynamik (z.B. Qualifikationsanforderungen) - teilweise beeinflussen sie diese Dynamik (Alterung und Produktivität) - oder auch gesetzlicher Rahmenbedingungen (z.B. Möglichkeiten der Frühverrentung).

Es gibt eine nicht mehr überschaubare Literatur zu der Frage, wie sich die Ausgaben für Renten und Pensionen angesichts der demografischen Alterung entwickeln könnten. Ob und welchen Effekt die demografische Entwicklung auf die Beitragseinnahmen der GRV hat, wird – wenn überhaupt - eher indirekt behandelt.<sup>2</sup> Meist geschieht dies im Rahmen der Berechnung von „Belastungsquoten“ wie dem Altenquotienten, der die Zahl der älteren Einwohner auf die Bevölkerung im Erwerbsalter bezieht. Sollte bei einer zurückgehenden Erwerbsbevölkerung auch die Zahl der Beschäftigten sinken, dann würde dies die Beiträge aus sozialversicherungspflichtiger Beschäftigung reduzieren.

Die sog. Senioritätsentlohnung und das Altersprofil der Arbeitsproduktivität dürften die wichtigsten Ansatzpunkte zur Erklärung geben, warum sich die Alterung der Beschäftigten auf die Einnahmen der GRV auswirken könnte. Es ist allerdings nicht unumstritten, ob die Arbeitsproduktivität altersabhängig ist (z.B. Börsch-Supan 2003) und es ist ebenso wenig geklärt, wie verbreitet in Deutschland die Senioritätsentlohnung ist (z.B. OECD 2005: 117ff.). Beides einmal angenommen, könnte eine demografisch bedingte Alterung der Arbeitskräfte die Lohnsumme beeinflussen und damit auch die Einnahmenseite der GRV.

Ein weiterer, indirekter Effekt könnte sich über die Qualifikationstendenzen einstellen. Wenn auf große Kohorten eine schwächer besetzte folgt und beide Kohorten unterschiedlich gut qualifiziert sind, sind Einflüsse auf die

---

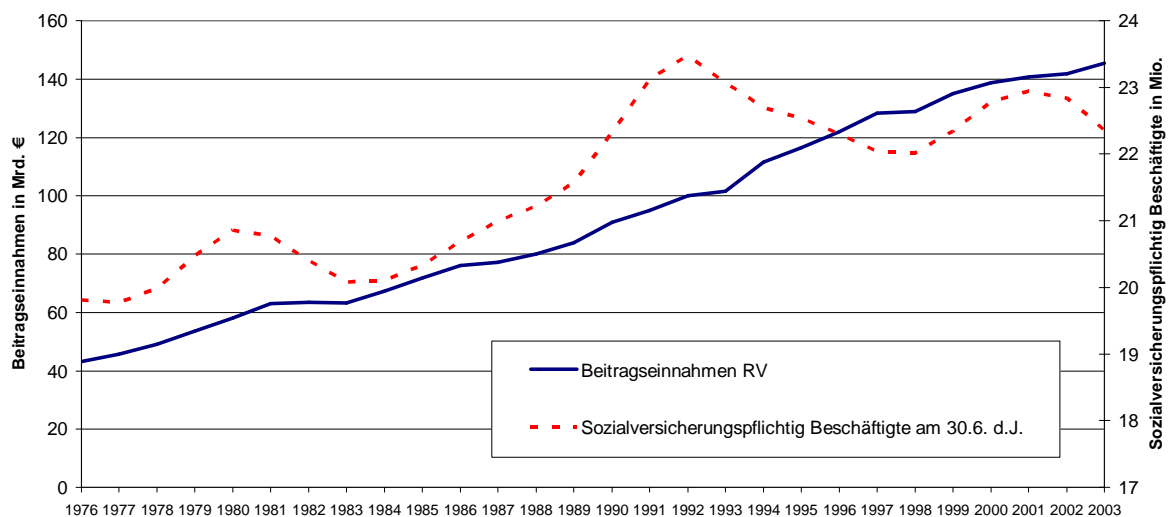
<sup>2</sup> So stellen Ulrich & Erbsland (1997) einen positiven Zusammenhang zwischen demografischer Alterung und den Steuereinnahmen fest. Als Grund dafür nennen sie die in der Vergangenheit gestiegene Bevölkerungszahl.



Einnahmen denkbar, weil besser qualifizierte Beschäftigte mehr verdienen. Der Alterseffekt ergibt sich dabei aus der unterschiedlichen Größe der Kohorten und nicht weil das Alters-Verdienstprofil auf einer Qualifikationsebene besonders ausgeprägt wäre (OECD 2005: 116 ff.)

Ob die Alterung der Beschäftigten die Beitragseinnahmen der GRV tangiert, ist demnach unklar. Diese Frage soll deshalb hier im Rahmen einer explorativen Studie untersucht werden.

**Abb. 1: Einnahmen der Rentenversicherung und sozialversicherungspflichtig Beschäftigte von 1976 bis 2003 (alte Bundesländer)**



Quelle: DRV, BA

Zunächst wird untersucht, ob es einen Zusammenhang gibt zwischen der Altersstruktur der sozialversicherungspflichtig Beschäftigten und den gesamten, aus der Beschäftigung resultierenden Einnahmen der GRV. Hierbei werden eine Reihe von anderen Einflussfaktoren, wie die Qualifikationsstruktur, die Teilzeitquote und das BIP, als Kovariate berücksichtigt. Es wird eine Regressionsbeziehung bestimmt, mit den Einnahmen als abhängige Variable (Regressand) auf der „linken Seite“ der Gleichung. Die rechte Seite der Gleichung enthält neben den genannten Kovariaten ein Maß für die Altersstruktur. Welches sich dafür am besten eignet, ist eine empirische Frage. Gedacht wird an die typischen statistischen Parameter mit denen man Lage, Streuung, Schiefe und Wölbung einer Verteilung charakterisiert. Die Alterung müsste sich auch in diesen Parametern niederschlagen. Neben diesem parametrischen Ansatz soll die in Klassen eingeteilte

Altersverteilung der Beschäftigten als Gruppe von Regressoren getestet werden.

Es ist anzunehmen, dass die Altersstruktur der Beschäftigten mit der der Bevölkerung und des Erwerbspotenzials korreliert. Aufgrund dieser Korrelation könnte sich der Zusammenhang zwischen Beitragseinnahmen und Alterung der Beschäftigten auf die Alterung der Bevölkerung bzw. des Erwerbspotenzials übertragen lassen. Würde sich dies bestätigen, dann wäre man vielleicht sogar in der Lage – mit der gebotenen Vorsicht – aus der fast schon als gesichert geltenden Alterung des Erwerbspotenzials bzw. der Bevölkerung künftige Folgen für die Einnahmenseite der GRV abzuleiten.

Die vorliegende Arbeit beschränkt sich auf die alten Länder, für die im Gegensatz zu den neuen Ländern relativ lange Zeitreihen vorliegen. Eine tiefere regionale Differenzierung würde den Rahmen der Vorstudie sprengen. Allerdings werden sowohl der Frauenanteil als auch der Ausländeranteil (jeweils an den sozialversicherungspflichtig Beschäftigten) als Kovariate in die Modellgleichungen eingehen.

### **3 Beschreibung der Daten**

Für die erwähnten Variablen liegen Zeitreihen aus dem Zeitraum 1976 bis 2003 vor.

Die Einnahmen der GRV für die alten Bundesländer entstammen der Statistik der DRV und beziehen sich auf die Beiträge zur „Gesetzlichen Rentenversicherung“ (Allgemeine und Knappschaftliche Rentenversicherung). Wie in der obigen Abbildung 1 deutlich zu sehen ist, haben die Beitragseinnahmen seit 1976 stetig und fast linear von Jahr zu Jahr zugenommen.

Den weitaus größten Teil der Beitragseinnahmen machen Pflichtbeiträge von abhängig Beschäftigten aus. Für Gesamtdeutschland waren dies im Jahr 2003 rund 84 % (nach Bloss 2006: 29). Die größten Posten, die nicht aus Beschäftigung stammen, waren Pflichtbeiträge für Leistungsempfänger der Bundesagentur für Arbeit (5,9 %) und für Kindererziehungszeiten (7 %).

Die Daten für die sozialversicherungspflichtig Beschäftigten (siehe Abb. 1) stammen von der Bundesagentur für Arbeit (BA). Für das Jahr 2003 gibt die BA eine Zahl von rund 22,36 Mio. sozialversicherungspflichtigen Beschäftigten an. Das sind weit weniger als die 33,36 Mio. „Aktiv Versicherten“ der GRV, wo 30,05 Mio. als Pflichtversicherte, 509 Tsd. als freiwillig Versicherte, fast 2,5 Mio. als geringfügig Beschäftigte und etwas mehr als 300 Tsd. als Anrechnungszeitversicherte erfasst sind.

Bei den Einnahmen werden für einige Gruppen (Pflicht-)Beiträge erhoben, die keiner sozialversicherungspflichtigen Beschäftigung nach den Kriterien des SGB III/der BA nachgehen. Angesichts der Größenordnung der Beitragseinnahmen spielt dies vor allem bei den Leistungsempfängern der BA und den Pflichtbeiträgen für Kindererziehungszeiten eine Rolle.

Da die Untersuchung den Charakter einer Vorstudie hat, wurde entschieden, trotzdem mit den vorliegenden BA-Daten für die Beschäftigten zu rechnen.

Die Beschäftigtendaten der BA sind nach den Merkmalen Einzelalter, Geschlecht, Deutsche/Ausländer, Teilzeitbeschäftigung und Qualifikation untergliedert, allerdings nicht in dieser Kombination, sondern gesondert nach den einzelnen Merkmalen. Die Daten beziehen sich auf den 30.6. eines Jahres und sind für die alten Länder verfügbar.

Für das Merkmal Alter wurden zwei Variablentypen gebildet: Altersklassen und die Parameter der Altersverteilung als Variable.

Die Aufteilung der Beschäftigten und die Bildung von Anteilen erfolgte einmal für die drei Hauptaltersklassen (14 bis unter 30, 30 bis unter 50, 50 bis 65) und zum zweiten für 5er Altersgruppen (14-19, 20-24 usw.)<sup>3</sup>

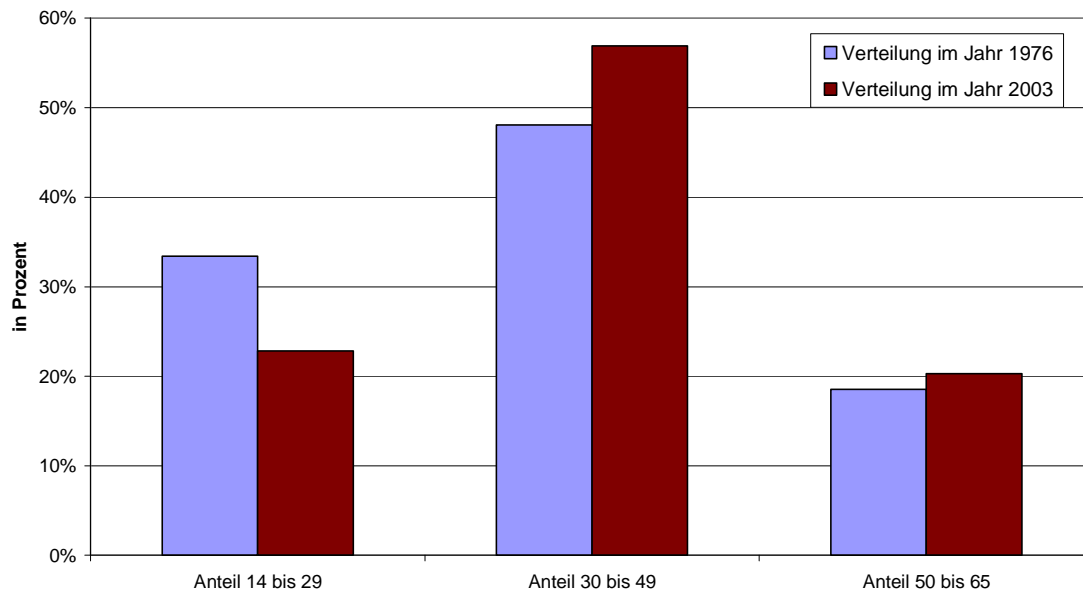
Eine gröbere Einteilung der Altersklassen ist statistisch günstig, weil damit Freiheitsgrade gespart werden, denn mit jeder Klasse verliert man einen

---

<sup>3</sup> Die BA-Statistik weist in der Tat vereinzelt 0 bis 14-Jährige und 65- bis 99-Jährige (!) aus. Offensichtlich liegen hier Fehler vor. Aus diesem Grund wurden die unter 14-Jährigen und die über 65-Jährigen aus der Analyse ausgeschlossen. Die 14-Jährigen wurden zur jüngsten Altersgruppe und die 65-Jährigen zur höchsten Altersgruppe gerechnet, weil das Alter in der BA-Statistik unseres Wissens nach früher über das Geburtsjahr errechnet wurde.

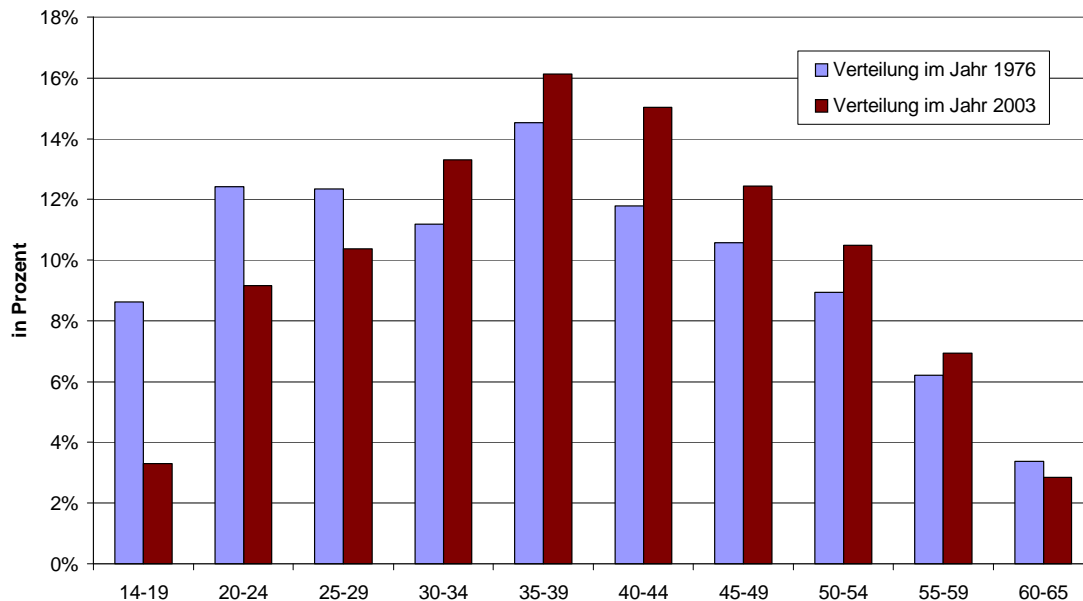
Freiheitsgrad. Andererseits gehen mit einer groben Einteilung Informationen verloren. Deshalb wurden beide Variable alternativ in die Analysen einbezogen.

**Abb. 2: Altersverteilung der sozialversicherungspflichtig Beschäftigten**  
(Summe über alle Altersklassen = 100 %)



Quelle: Eigene Berechnungen mit Daten der BA.

**Abb. 3: Altersverteilung der sozialversicherungspflichtig Beschäftigten**  
(Summe über alle Altersklassen = 100 %)



Quelle: Eigene Berechnungen mit Daten der BA.

Eine sparsamere Modellierung erlaubt die Verwendung von Verteilungsparametern. Für das Merkmal Alter der Beschäftigten wurden die typischen

Parameter einer Verteilung berechnet: Mittelwert, Median, Standardabweichung, Schiefe und Wölbung.

Die Parameter der Altersverteilung wurden wie folgt berechnet:

$$p(i) = \frac{\text{Anzahl sozialversicherungspflichtig Beschäftigter (Anz. sozpf. Besch.) im Alter}(i)}{\text{Gesamtzahl sozpf. Beschäftigter}}$$

$$ME = \text{Median} := \begin{cases} \frac{\text{Gesamtzahl sozpf. Beschäftigter}}{2} \leq \text{Anz. sozpf. Besch. im Alter } [14, \dots, ME] \\ \frac{\text{Gesamtzahl sozpf. Beschäftigter}}{2} > \text{Anz. sozpf. Besch. im Alter } ]ME, \dots, 65] \end{cases}$$

$$MW = \text{Mittelwert} = \sum_{i=14}^{65} i * p(i)$$

$$SD = \text{Standardabweichung} = \sqrt{\sum_{i=14}^{65} p(i) * (i - MW)^2}$$

$$SK = \text{Schiefe} = \frac{\sum_{i=14}^{65} p(i) * (i - MW)^3}{\text{Varianz}^{3/2}} = \frac{\sum_{i=14}^{65} p(i) * (i - MW)^3}{SD^3}$$

$$KURT = \text{Wölbung} = \frac{\sum_{i=14}^{65} p(i) * (i - MW)^4}{SD^4} - 3$$

Abbildung 4 zeigt, wie sich die Parameter im Laufe der Zeit entwickelt haben. Das Durchschnittsalter der Beschäftigten steigt seit längerem, wobei der Median erst in der jüngeren Vergangenheit nachzog. Die Streuung der Verteilung – gemessen mit der Standardabweichung – ist seit Mitte der 80er Jahre deutlich zurückgegangen.

Auffällig sind auch die Veränderungen der Schiefe. Ausgehend von einer stärkeren Linkssteilheit tendierte die Verteilung zunächst Richtung Symmetrie, d.h. die Schiefe näherte sich dem Wert Null. Danach wurde die Verteilung wieder linkssteiler. In der jüngeren Vergangenheit wurde die Verteilung zunehmend symmetrischer. Diese Maßzahl beschreibt, was sich in Abb. 3 andeutete: weniger jüngere Beschäftigte, mehr im mittleren Alter.

Der Parameter Wölbung zeigt, dass die Altersverteilung der Beschäftigten in den letzten Jahren steiler wurde. (Bei einem Wert von Null hätte die Verteilung dieselbe Wölbung wie die Normalverteilung.) Damit kommt zum Ausdruck, dass die jüngeren Arbeitskräfte weniger wurden, während der Anteil der mittleren Klassen zunahm.

Abb. 4a: Parameter der Altersverteilung der SV-Beschäftigten

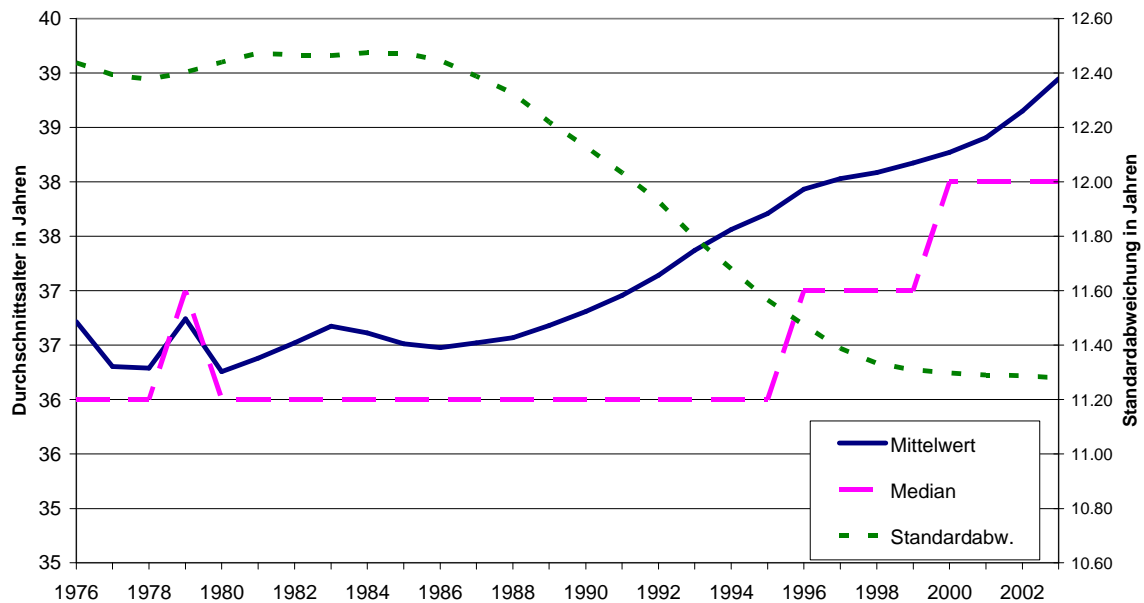
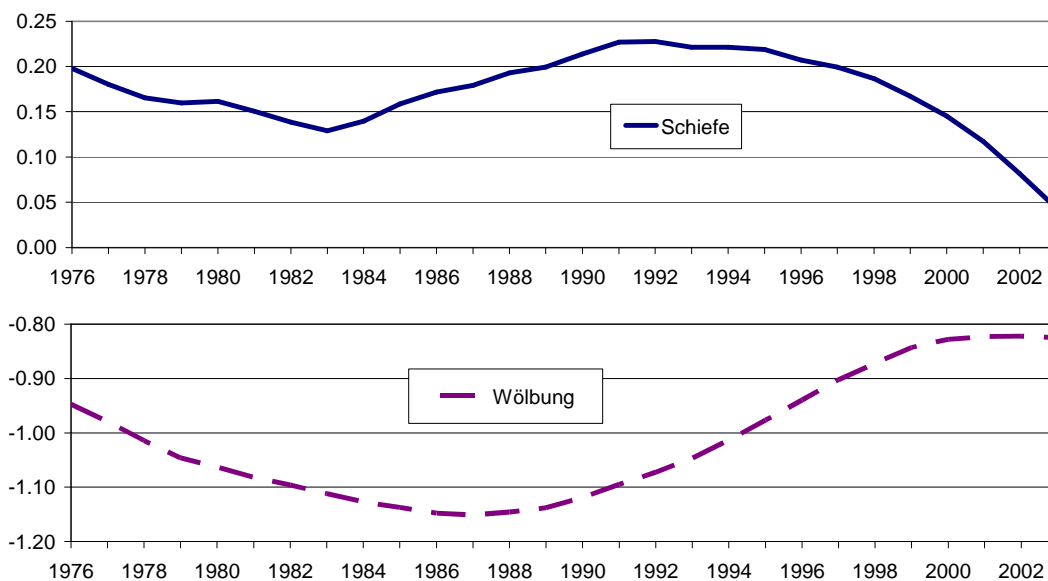


Abb. 4b: Parameter der Altersverteilung der SV-Beschäftigten



Quelle: Eigene Berechnungen mit Daten der BA

Weiter wurden mit den Angaben der BA folgende Variable gebildet (jeweils zum 30.6. eines Jahres, sozialversicherungspflichtig Beschäftigte (SVB) in den alten Ländern):

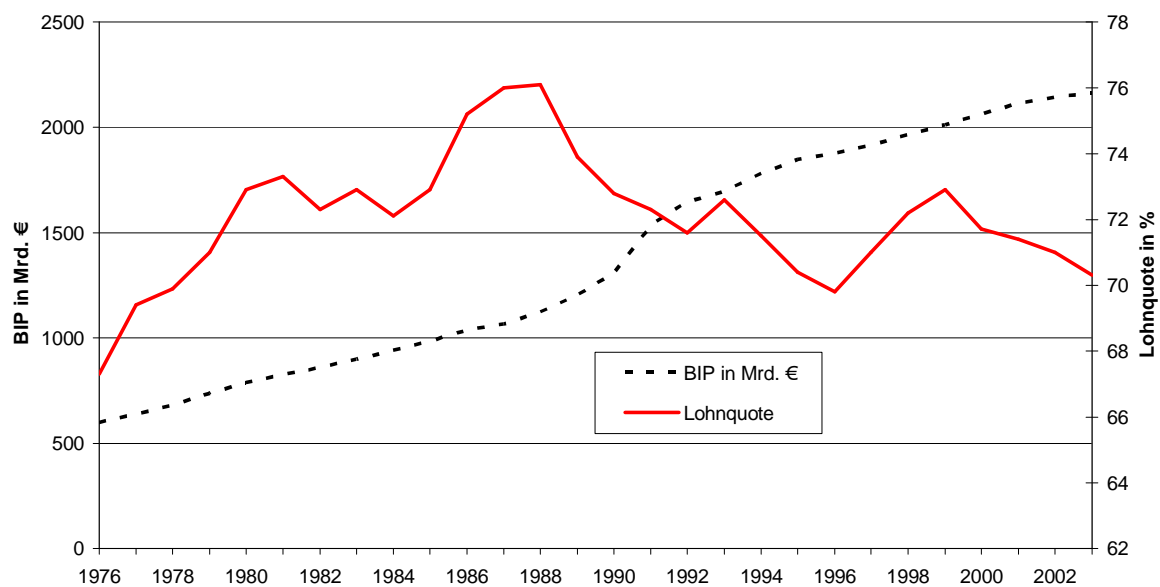
- Teilzeitquote der Frauen/der Männer

- Teilzeitquote nach Umfang (Quote für die Teilzeitbeschäftigten mit einer Wochenarbeitszeit unter 20 Stunden und für die Teilzeitbeschäftigten mit mehr 20 Wochenstunden)
- Frauenanteil
- Ausländeranteil
- Quote der SV-Beschäftigten nach Qualifikationen: ohne Berufsausbildung, mit Berufsausbildung oder Fachschule, mit (Fach-)Hochschulabschluss

Als weitere Variable verwendeten wir den Beitragssatz zur GRV, der offenkundig eine sehr wichtige Stellschraube ist, um Einnahmen und Ausgaben auszugleichen.

Mit dem Bruttoinlandsprodukt (BIP) berücksichtigte die Modellbildung eine zentrale ökonomische Variable als Indikator für die Einkommensentwicklung. Das BIP wird ab 1992 allerdings nur noch für Gesamtdeutschland ausgewiesen.

**Abb. 5: Gesamtwirtschaftliche Größen - BIP und Lohnquote**  
(Westdeutschland, ab 1991 Gesamtdeutschland)



Datenquelle: Statistisches Bundesamt

Außerdem haben wir die Lohnquote (Bruttoeinkommen aus unselbständiger Arbeit in % des Volkseinkommens) herangezogen. Auch von dieser Größe gibt es ab 1991 nur noch gesamtdeutsche Werte.

## 4 Schätzergebnisse für den Einfluss der Altersstruktur der SV-Beschäftigten

Alle Variablen liegen als Zeitreihen vor. Das statistische Modell erklärt die Entwicklung der Rentenversicherungsbeiträge mit den gesamtwirtschaftlichen Größen, dem Beitragssatz, der Altersverteilung der Beschäftigten und den anderen Strukturmerkmalen der Beschäftigung. Die zeitreihenanalytische Vorgehensweise geht von einem stochastischen Prozess aus. Voraussetzung für diese ökonometrische Modellierung ist Stationarität aller Variablen. Sind die Variablen nichtstationär, gelten die Signifikanzniveaus der üblichen Tests nicht mehr („spurious results“).

Für die Niveaus der Variablen deuten die gängigsten Unit Roots Tests (z.B. der Augmented Dickey-Fuller Test (ADF)) - bis auf einige Altersvariablen<sup>4</sup> - auf Nichtstationarität hin. Stationär waren jedoch in den meisten Fällen die ersten Differenzen (siehe Tabelle A2 im Anhang). Es erschien deshalb sinnvoll, bei der anschließenden Modellbildung auf Variablendifferenzen aufzubauen.<sup>5</sup>

Das Modell nimmt eine bestimmte Kausalität an: Die zu erklärende Variable sind die Beitragseinnahmen. Die übrigen Variablen sind erklärende Einflussgrößen (Regressoren). Die Altersverteilung und die übrigen Strukturmerkmale der Beschäftigung (z.B. Qualifikationsstruktur, Teilzeitanteil) dürften nicht von den gesamten Beitragseinnahmen abhängen. Insofern ist die Kausalrichtung eindeutig.

Der Beitragssatz ist jedoch eine Größe, für die man eine anders gerichtete Kausalität annehmen könnte. Wenn die Beitragseinnahmen nicht reichen, um die Ausgaben zu decken, wäre eine denkbare Reaktion eine Anhebung des Beitragssatzes. Tatsächlich führt der Granger-Kausalitätstest zu dem Ergebnis, dass die Einnahmen den Beitragssatz determinieren und nicht umgekehrt. Aber der Beitragssatz kann kaum als stochastische Variable angesehen werden. Es ist mehr ein politisches bzw. aus Sicht der GRV

---

<sup>4</sup> Bei den Altersvariablen, insbesondere den Altersklassen, müsste die Frage der Stationarität allerdings noch einmal überdacht werden (gemeinsame Unit Root). Dies war wegen der Kürze der Bearbeitungsdauer des Projekts nicht abschließend zu klären.

<sup>5</sup> Es wurden auch Tests mit einer als Quote definierten Einnahmenvariable durchgeführt. Dafür wurden die Beitragseinnahmen in Relation zum BIP gesetzt. Als Ergebnis kam heraus, dass die Quote weniger als die Differenz geeignet ist.



„betriebswirtschaftliches“ Instrument, das gesetzt wird. Für die vorliegende Arbeit betrachten wir den Beitragssatz deshalb als eine die Einnahmenseite beeinflussende Größe. Es wäre aber nicht uninteressant, den Zusammenhang von Beitragssatz und Einnahmen im Kontext der Ausgaben zu untersuchen.<sup>6</sup> Dies wurde in der vorliegenden Arbeit noch nicht versucht.

Auch beim Zusammenhang von Beitragseinnahmen und BIP sehen wir die Möglichkeit einer wechselseitigen Beziehung. Immerhin machen die Leistungen der GRV zu einem nicht geringen Teil die Wirtschaftskraft aus.<sup>7</sup> Die Richtung des Zusammenhangs wurde deshalb mit dem Granger-Kausalitätstest geprüft.<sup>8</sup> Der Test spricht für eine Abhängigkeit der Einnahmen vom BIP und nicht umgekehrt.

Derselbe Test für die Lohnquote und die Beitragseinnahmen ergab statistisch überhaupt keinen wechselseitigen Zusammenhang zwischen diesen Variablen. Das Ergebnis spricht also auch nicht für eine Abhängigkeit der Lohnquote von den Einnahmen.

Für eine genaue Spezifikation des vermuteten Zusammenhangs fehlen die theoretische und die empirische Grundlage. Getestet wurde deshalb eine Reihe von durchaus unterschiedlichen Modellen. Auf der rechten Seite der Gleichung standen in allen Fällen eine (oder mehrere) Altersvariable sowie die übrigen Einflussgrößen als Kovariate. Nachdem die Zeitreihen den Tests zufolge wahrscheinlich nicht stationär sind, wurden zunächst Modelle in ersten Differenzen geschätzt.

Im Weiteren werden nur die wichtigsten Ergebnisse kommentiert. Von den umfangreichen Rechnungen stehen ausgewählte Ergebnistabellen im Anhang. Zusätzliche Tabellen werden auf Anfrage gerne zur Verfügung gestellt.

---

<sup>6</sup> Langfristige Zusammenhänge zwischen demografischen Variablen und den Sozialabgaben finden Ulrich & Erbsland im Rahmen einer Kointegrationsanalyse (1997).

<sup>7</sup> Siehe Bloss (2006), die die Bedeutung der Sozialversicherungssysteme auf regionaler Ebene untersuchte.

<sup>8</sup> Für den Test von Granger & Sims auf Kausalität siehe Pindyck/Rubinfeld 1998: 242f.

## 4.1 Modelle mit ersten Differenzen

Ausgangspunkt der Modellbildung zum Alterseinfluss ist die Abhängigkeit der Beitragseinnahmen der GRV vom BIP, der Zahl der sozialversicherungspflichtig Beschäftigten (SVB) und dem Beitragssatz zur GRV. Mit diesen Variablen lässt sich die Veränderung der Einnahmen nur zu rund 10 % erklären (gemessen mit dem  $R^2$ , siehe Tabelle A30 im Anhang). Diese Grundgleichung wird im Folgenden um den Alterseinfluss erweitert.

Eine interessante Variable wäre an sich die Lohnquote, da sie zusammen mit BIP und Beitragssatz die Beitragseinnahmen der GRV fast vollständig (wenn auch multiplikativ, nicht additiv) determiniert.<sup>9</sup> Wie bereits im zweiten Kapitel diskutiert wurde, dürfte der Effekt der Altersvariablen im Wesentlichen über die Lohnquote vermittelt werden. Tatsächlich verändern sich bei Aufnahme der Lohnquote die Koeffizienten der Altersvariablen. Wenn sowohl die Altersvariablen als auch die Lohnquote in ein Modell eingehen, dann ist die Lohnquote zwar hochsignifikant, aber das Vorzeichen ihres Koeffizienten ist verkehrt: mit steigender Lohnquote würden danach die Beitragseinnahmen sinken. Aus diesem Grund haben wir sie aus den weiteren Analysen ausgeschlossen.

Zu unterscheiden sind zwei prinzipielle Modellansätze. Zum einen wurde mit Altersklassen gerechnet, zum anderen mit den Parametern der Altersverteilung.

### Modellierung des Einflusses des Alters mit Altersklassen

Es zeigt sich zunächst, dass es günstiger ist, die Anzahl der sozialversicherungspflichtig Beschäftigten aus dem Modell auszuschließen. Mit ihr verliert die Altersklassenvariable etwas an Einfluss, ohne dass die Variable „Anzahl der SV-Beschäftigten“ signifikant wäre.

Einen statistisch signifikanten Einfluss des Alters fanden wir ausschließlich für die 5-Jahres-Altersgruppen. Modelle mit einer gröberen Altersstruktur (mit den drei Hauptaltersklassen unter 30, 30 bis unter 50, 50 bis 65) ergaben keine signifikanten Resultate. Die Berücksichtigung weiterer Variab-

---

<sup>9</sup> Ein Modell mit der Lohnquote und dem BIP als alleinige exogene Variable (jeweils logarithmiert) „erklärt“ die Einnahmen der GRV zu gut 99 %.

len – wie die Teilzeitquote, die Qualifikationsstruktur, das BIP, der Beitragssatz usw. – änderte nichts an diesem Grundmuster.

Von den vielen geschätzten Gleichungen für die Entwicklung der Beitragseinnahmen haben wir das folgende Modell 1 ausgewählt, weil sich ihre statistischen Eigenschaften alles in allem am überzeugendsten darstellten. Im Anhang stehen unkommentiert die Ergebnisse für weitere Modelle, die unserer Auffassung nach keine entscheidende Verbesserung gegenüber Gleichung 1 bringen.

Modell 1:

$$\begin{aligned} D(\text{EINNAHMEN}) = & -195049.1545 * D(\text{SVB\_20\_24}) - 441798.6385 * D(\text{SVB\_25\_29}) - \\ & 244738.3457 * D(\text{SVB\_30\_34}) - 297876.8519 * D(\text{SVB\_35\_39}) - \\ & 406733.6303 * D(\text{SVB\_40\_44}) - 565627.1607 * D(\text{SVB\_45\_49}) - \\ & 377226.4648 * D(\text{SVB\_50\_54}) + 21621.35634 * D(\text{SVB\_55\_59}) + \\ & 501759.1886 * D(\text{SVB\_60\_65}) + 2086.416621 * D(\text{BS\_RV}) + \\ & 50921.91591 * W\_BIP + 1114929.893 * D(\text{QUAL\_UNI}) \end{aligned}$$

Erläuterung: siehe Tabelle 1

Die Beitragseinnahmen der GRV sind nach dieser Schätzung eine Funktion der Altersverteilung (5er-Altersklassen), des Beitragssatzes zur GRV, des BIP und der Qualifikationsstruktur der Beschäftigten, und zwar modelliert mit den ersten Differenzen bzw. der Wachstumsrate.

Erwartungsgemäß wirken sich ein höheres Wirtschaftswachstum, ein höherer Beitragssatz und eine bessere Qualifikationsstruktur (höherer Verdienst!) positiv auf die Beitragseinnahmen der GRV aus.

Etwas schwierig zu interpretieren ist der Einfluss der Altersvariablen. Die einzelnen Koeffizienten der Altersklassen können nicht unabhängig voneinander gesehen werden. Der Grund dafür ist, dass die Summe der 10 Altersklassen immer 100 % ergibt. Erhöht sich der Anteil in einer Altersklasse muss der der übrigen entsprechend sinken.<sup>10</sup>

---

<sup>10</sup> Aus mathematischen Gründen dürfen in die Regressionsgleichung nicht alle Altersklassen eingehen, andernfalls läge strenge Kollinearität vor und das Modell ließe sich nicht mehr schätzen. Es spielt statistisch keine Rolle, welche Klasse dies ist, weil immer das Gesamtsystem der Altersklassen zu betrachten ist. Für alle hier gerechneten Modelle wurde als Basis die Altersklasse der 14- bis 19-jährigen Beschäftigten gewählt. Erhöht man im Rahmen einer Simulation den Anteil einer Altersklasse, ohne dass man die Anteile der übrigen im Modell aufgenommenen Klassen verändert, geht dies automatisch

Um ein Beispiel zu geben: Der Anteil der Altersklasse 25 bis 29 sinkt um 1 %-Punkt und der Anteil der 40- bis 44-Jährigen steigt um denselben Wert. Der negative Regressionskoeffizient bewirkt steigende Beitragseinnahmen aufgrund der Altersklasse 25 bis 29 und sinkende aufgrund der Klasse der 40- bis 44-Jährigen. Die gesamte Nettoveränderung der Einnahmen ergäbe sich dann aus dem Plus und Minus dieser beiden Klassen.

Nimmt allerdings der Anteil der Älteren, z.B. der 55-59-Jährigen, zu Lasten einer jüngeren Altersklasse zu, dann erhöhen sich die Einnahmen, weil das Modell für die Älteren positive Koeffizienten ausweist. Dieses Ergebnis steht im Einklang mit der Senioritätsentlohnung.

Tabelle 1 enthält das Schätzergebnis für Gleichung 1.<sup>11</sup>

Die neun Koeffizienten der Altersvariablen wurden mit dem Wald Test auf gemeinsame Signifikanz getestet. Die Teststatistik (F-Wert 3,305, df (9,15)) war hochsignifikant ( $p = 0,02$ ).

Das Modell dürfte frei von Heteroskedastizität sein, da der in Tabelle 1 angegebene White Test nicht signifikant ist.

Im Modell sind deutlich Spuren von Autokorrelation feststellbar (siehe Durbin-Watson-Statistik und Breusch-Godfrey Serial Correlation LM Test). Deshalb wurde es mit einer Varianz-Kovarianzmatrix geschätzt, die auch in diesem Fall konsistente Schätzungen ergibt (Heteroskedasticity and Autokorrelation Consistent Covariances (HAC) von Newey-West.).<sup>12</sup> Da die Autokorrelationsfunktion auf einen AR-Prozess 2ter Ordnung hindeutete, wurden die HAC mit einem Lag von 2 geschätzt.

Darüber hinaus modellierten wir einen ARMA-Ansatz. Signifikant waren autoregressive Terme 1. und 2. Ordnung (AR(1) und AR(2)). Eine grundlegende Änderung der Ergebnisse ergibt sich dadurch nicht (siehe Anhang A31).

---

zu „Lasten“ dieser Basisvariablen. Rechnerisch zeigt sich der Zusammenhang zwischen den Altersvariablen in der Kovarianz der Koeffizientenschätzungen.

<sup>11</sup> Alle ökonometrischen Berechnungen wurden mit dem Programm EViews durchgeführt.

<sup>12</sup> Siehe Manual des Statistik- und Ökonometrie-Programms EViews 5 (2004: 457).

**Tabelle 1: Modell mit Altersklassen (Modell 1) – Schätzung mit 1. Differenzen**

Dependent Variable: D(EINNAHMEN)

Method: Least Squares

Sample (adjusted): 1977 2003

Newey-West HAC Standard Errors &amp; Covariance (lag truncation=2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SVB_20_24)	-195049.2	130859.9	-1.490519	0.1568
D(SVB_25_29)	-441798.6	214400.5	-2.060623	0.0571
D(SVB_30_34)	-244738.3	107282.8	-2.281245	0.0376
D(SVB_35_39)	-297876.9	116923.8	-2.547615	0.0223
D(SVB_40_44)	-406733.6	167176.7	-2.432956	0.0280
D(SVB_45_49)	-565627.2	136639.9	-4.139547	0.0009
D(SVB_50_54)	-377226.5	176274.6	-2.139994	0.0492
D(SVB_55_59)	21621.36	139723.8	0.154744	0.8791
D(SVB_60_65)	501759.2	102222.7	4.908493	0.0002
D(BS_RV)	2086.417	561.3664	3.716676	0.0021
W_BIP	50921.92	11075.04	4.597898	0.0003
D(QUAL_UNI)	1114930.	195120.3	5.714063	0.0000
R-squared	0.815178	Mean dependent var		3787.464
Adjusted R-squared	0.679643	S.D. dependent var		2233.930
S.E. of regression	1264.407	Akaike info criterion		17.42370
Sum squared resid	23980879	Schwarz criterion		17.99962
Log likelihood	-223.2199	Durbin-Watson stat		2.530778
<b>White Heteroskedasticity Test:</b>				
F-statistic	4.144895	Prob. F(24,2)		0.212481
Obs*R-squared	26.46786	Prob. Chi-Square(24)		0.329861
<b>Breusch-Godfrey Serial Correlation LM Test:</b>				
F-statistic	5.156216	Prob. F(2,13)		0.022455
Obs*R-squared	11.92893	Prob. Chi-Square(2)		0.002568
<b>Ramsey RESET Test:</b>				
F-statistic	1.150019	Prob. F(2,13)		0.346841
Log likelihood ratio	4.398459	Prob. Chi-Square(2)		0.110889

*Erläuterungen:**D(..): Veränderung der Variable gegenüber dem Vorjahr (1. Differenz)**SVB\_20\_24 ist der Anteil der 20- bis 24jährigen an allen SV-Beschäftigten; entsprechendes gilt für die übrigen Altersgruppen**BS\_RV: Beitragssatz zur GRV**W\_BIP: Wachstumsrate des BIP**Qual\_Uni : Anteil der SV-Beschäftigten mit (Fach-)Hochschulabschluss**Quelle: Eigene Berechnungen.*

Nach dem Ergebnis des Jarque-Bera-Tests dürften die Residuen normalverteilt sein.

Die funktionale Form (lineare Gleichung) wurde mit dem „regression specification error test“ (RESET) geprüft. Der Test war nicht signifikant, d.h. die Annahme einer linearen Beziehung ist statistisch haltbar.

Nun lag der Likelihood Ratio Test des RESET nur knapp über dem 10 %-Signifikanzniveau und das Modell enthält ein gewisses Maß an Autokorrelation. Beides kommt auch vor, wenn in einem Modell wichtige Einflussgrößen fehlen. Deshalb wurde für die obige Gleichung die Signifikanz weiterer Variablen getestet.

Eine Konstante führt zu keiner signifikanten Verbesserung des Modells.

Die Hinzunahme einzelner parametrischer Altersvariablen schwächt den Einfluss der 5er Altersklassen etwas ab, aber die parametrische Altersvariable bleibt insignifikant. Nur der Regressionskoeffizient der Standardabweichung ist signifikant, allerdings verbessert sich dadurch Gleichung 1 nicht signifikant (siehe Anhang A32).

Der Frauenanteil, der Ausländeranteil und die Teilzeitquote (an den SV-Beschäftigten) waren ebenfalls nicht signifikant.

Nachdem die Einnahmen auch Versicherungsbeiträge für bei der Bundesagentur für Arbeit registrierten Arbeitslosen enthalten, wurde der Einfluss der Arbeitslosigkeit geprüft. Die Variable ist signifikant und kann Modell 1 geringfügig verbessern (siehe Anhang, Tabelle A33).

In den Stützzeitraum der Schätzungen fällt die Wiedervereinigung mit ihren erheblichen Verwerfungen in ökonomischer und sozialer Hinsicht. Deshalb prüften wir, ob sich dies in den untersuchten Beziehungen zeigt.

Zunächst wurde der Chow Test durchgeführt. Dieser teilt den Stützzeitraum auf und schätzt für den Zeitraum vor und den nach dem Strukturbruch je eine – identische – Gleichung. Die Schätzungen für die beiden Zeiträume vor und nach 1990/1991 unterschieden sich nicht signifikant, d.h. der Test deutet auf keinen Strukturbruch hin. Für den Zeitraum nach 1990/91 liegen aber nur 12 Werte vor, also gerade so viele wie Regresso-

ren. Insofern ist der Chow Test hier nur sehr eingeschränkt interpretierbar.

Weiter haben wir mit einer 0-1-Dummyvariablen einen einfachen Sprung in den Beziehungen ab 1991 modelliert; dies entspricht einer Änderung in der Regressionskonstanten. Auch die Dummyvariable brachte keinen Hinweis auf einen Strukturbruch.

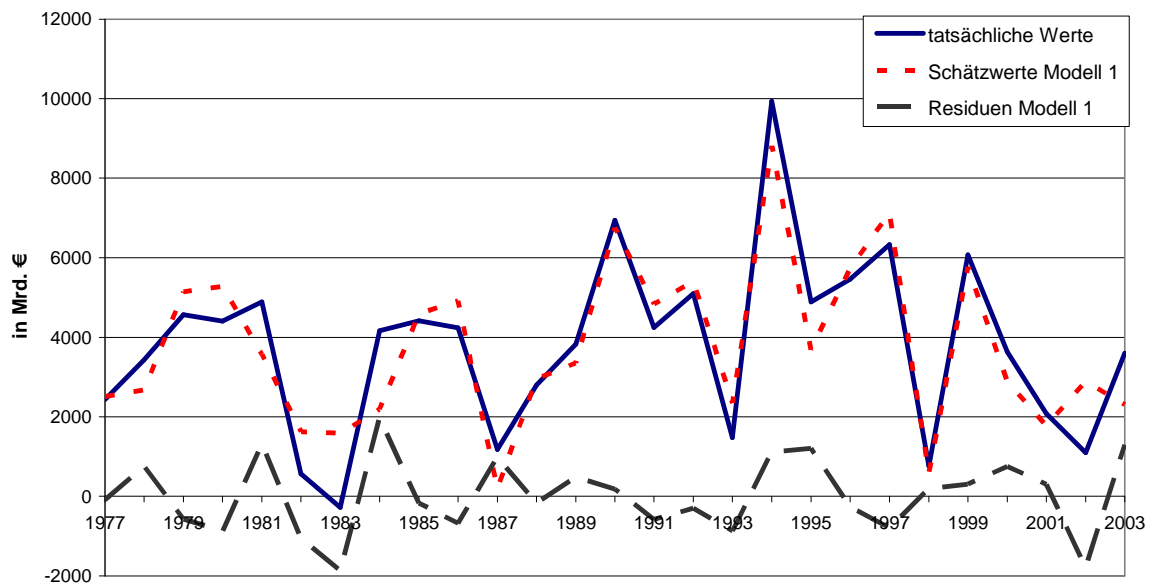
Möglicherweise hat sich der Einfluss der Regressoren nach 1990 generell verändert. Um eine solche denkbare Änderung in den Steigungsparametern zu erfassen, wird für den Regressor, bei dem man eine solche Änderung vermutet, eine zusätzliche Variable eingeführt, die diesen Regressor mit einer Dummyvariable verknüpft. Interessant war dies hier besonders für das BIP, für das nach 1991 nur noch gesamtdeutsche Werte vorliegen. Ein nach 1991 geänderter Effekt des BIP ließ sich jedoch nicht feststellen.

Möglicherweise gibt es aber andere Einflüsse, die sich nach 1990 geändert haben. Um dies zu modellieren, wurde eine „Trenddummy“ definiert, deren Werte bis 1990 Null sind und ab 1991 Jahr für Jahr um Eins zunehmen. Diese Variable sollte trendmäßige Einflüsse abbilden, die erst nach 1990 einsetzen. Jedoch war auch die „Trenddummy“ statistisch nicht signifikant ( $p = 0,075$ ).

Grafisch ist ebenfalls kein Trendbruch erkennbar (Abb. 6).

Die Residuen in Abb. 6 zeigen zugleich, dass Modell 1 die Entwicklung zeitweise nicht besonders gut erklären kann. In 1983/84 und 2002/2003 liegen die Residuen außerhalb der Marke eines Standardfehlers.

Abb. 6: Schätzwerte und Residuen im Modells 1 (Altersklassen)



Anmerkung: Residuum = tatsächlicher Wert - Schätzwert  
Quelle: Eigene Berechnungen.

### Eine parametrische Modellierung des Alterseinflusses

Die Modellierung mit den Altersklassen ist sehr flexibel und insbesondere auch für Simulationen geeignet, weil man die Werte einzelner Altersklassen unmittelbar gegeneinander verrechnen kann. Der Ansatz benötigt jedoch viele Freiheitsgrade, was Angesichts der relativ kurzen Zeitreihen einen prinzipiellen Nachteil darstellt. Eine Alternative ist eine Modellierung mit den Parametern der Altersverteilung: Mittelwert/Median, Standardabweichung, Schiefe und Kurtosis (Wölbung). Damit werden Freiheitsgrade gespart. Nachdem die Modellierung mit Altersklassen bereits ein sehr plausibles Modell mit guten statistischen Eigenschaften gebracht hat, setzt die folgende Darstellung unmittelbar darauf auf. Anstelle der Altersklassen wurden die Parameter der Altersverteilung als Variable verwendet.

Von den fünf parametrischen Altersvariablen Mittelwert, Median (für den Einfluss des Durchschnittsalters), Standardabweichung, Schiefe und Wölbung sind nur die Standardabweichung und die Wölbung signifikant. Das ändert sich auch nicht, wenn man einzelne Variable, bspw. die Schiefe, hinzunimmt (siehe Anhang A3).

Folgende Modellgleichung sei im Weiteren näher beschrieben:



## Modell 2:

$$D(\text{EINNAHMEN}) = 23966.14745 * D(\text{SVB\_AGE\_STDDEV}) + 86341.54591 * D(\text{SVB\_AGE\_KURT}) + 2021.028626 * D(\text{BS\_RV}) + 64722.56305 * W\_BIP + 687585.2052 * D(\text{QUAL\_UNI})$$

Erläuterung: siehe Tabelle 2

Tabelle 2 enthält die statistischen Kennwerte für Modell 2. Die Koeffizienten für den Beitragssatz, das Wirtschaftswachstum und die Qualifikationsstruktur haben wie schon beim obigen Modell 1 erwartungsgemäß ein positives Vorzeichen.

Die Koeffizienten der Standardabweichung und der Wölbung besagen, dass eine größere Streuung oder eine stärkere Steilheit der Altersverteilung die Einnahmeseite positiv beeinflusst.

Zu beachten sind die nicht unerheblichen Interpretationsprobleme mit diesem Modell. So kann sich die gleiche Streuung (Standardabweichung) aus ganz unterschiedlichen Verteilungen ergeben. Angenommen, zum Zeitpunkt 1 konzentrierte sich die gesamte Verteilung auf eine Altersklasse. Eine stärkere Streuung zum Zeitpunkt 2 folgt, wenn die Verteilung ausgeglichener wird. An welcher Stelle (Altersbereich) die Verteilung sich vorher konzentrierte – Jüngere oder Ältere -, ist bedeutungslos, wenn man sich nur auf die Zunahme der Streuung konzentriert.

Ähnliche Interpretationsprobleme wirft die Variable Wölbung auf. Eine stärkere Wölbung kann eine Folge der Zunahme des mittleren Altersintervalls zu Lasten der Ränder sein. Welche Stellen der Verteilung betroffen sind, bleibt offen.

Mit den üblichen Tests lassen sich keine Verletzungen der statistischen Annahmen (Homoskedastizität, keine Autokorrelation, Normalverteilung der Residuen, Linearität des Modells) feststellen. Eine gewisse Tendenz besteht zu einer Autokorrelation zweiter Ordnung.

Die Tests auf Strukturbruch ergaben keine signifikanten Hinweise. Variablen, die das Modell noch statistisch signifikant verbessern könnten, wurden nicht gefunden.

**Tabelle 2: Modell mit parametrischen Altersvariablen (Modell 2)  
– Schätzung mit 1. Differenzen**

Dependent Variable: D(EINNAHMEN)

Method: Least Squares

Sample (adjusted): 1977 2003

Included observations: 27 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SVB_AGE_STDDEV)	23966.15	8661.541	2.766961	0.0112
D(SVB_AGE_KURT)	86341.55	19805.13	4.359554	0.0003
D(BS_RV)	2021.029	475.3988	4.251228	0.0003
W_BIP	64722.56	9832.450	6.582547	0.0000
D(QUAL_UNI)	687585.2	194159.6	3.541341	0.0018
R-squared	0.713497	Mean dependent var		3787.464
Adjusted R-squared	0.661405	S.D. dependent var		2233.930
S.E. of regression	1299.899	Akaike info criterion		17.34354
Sum squared resid	37174221	Schwarz criterion		17.58351
Log likelihood	-229.1377	Durbin-Watson stat		1.994785

**White Heteroskedasticity Test:**

F-statistic	0.478775	Prob. F(10,16)	0.880260
Obs*R-squared	6.218528	Prob. Chi-Square(10)	0.796581

**Breusch-Godfrey Serial Correlation LM Test:**

F-statistic	2.668562	Prob. F(2,20)	0.093913
Obs*R-squared	5.582318	Prob. Chi-Square(2)	0.061350

**Ramsey RESET Test:**

F-statistic	1.893543	Prob. F(2,20)	0.176558
Log likelihood ratio	4.682086	Prob. Chi-Square(2)	0.096227

*Erläuterungen:*

*D(.): Veränderung der Variable gegenüber dem Vorjahr (1. Differenz)*

*SVB\_AGE\_Stddev: Standardabweichung der Altersverteilung der SVB*

*SVB\_AGE\_Kurt: Kurtosis (Wölbung) der Altersverteilung der SVB*

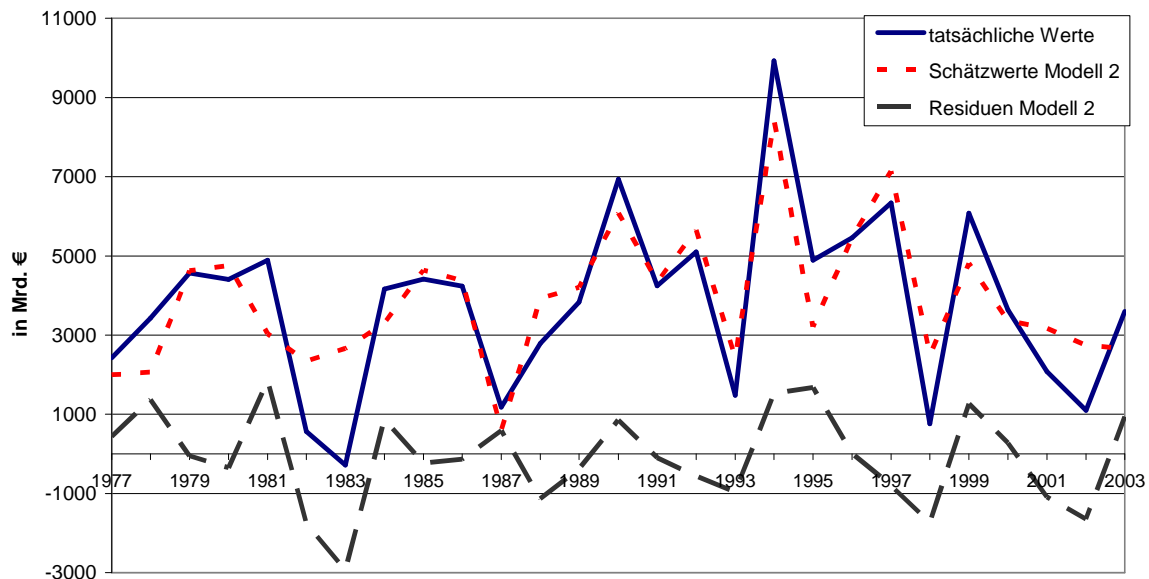
*BS\_RV: Beitragssatz zur GRV*

*W\_BIP: Wachstumsrate des BIP*

*Qual\_Uni : Anteil der SV-Beschäftigten mit (Fach-)Hochschulabschluss*

*Quelle: Eigene Berechnungen.*

Wie schon Modell 1 hat auch das Modell 2 mit den parametrischen Altersvariablen gewisse Probleme mit der Erklärung der Beitragsentwicklung in 1983/84 und am aktuellen Rand. Darüber hinaus liegen die Residuen in mehreren Jahren außerhalb der Marke eines Standardfehlers (Abb. 7).

**Abb. 7: Schätzwerte und Residuen im Modells 2 (Verteilungsparameter)**

Anmerkung: Residuum = tatsächlicher Wert - Schätzwert  
Quelle: Eigene Berechnungen.

#### Erläuterung:

durchgezogene Linie Schätzwert,  
gestrichelte Linien oberes und unteres Konfidenzintervall  
Quelle: Eigene Berechnungen

### Modellvergleich: Altersklassen vs. Parameter der Altersverteilung

Welcher der beiden Modelltypen ist nach den vorliegenden Ergebnissen vorzuziehen? Inhaltlich sind die Modelle ähnlich. Die Altersklassen dürften allerdings hinsichtlich ihres Einflusses eher einfacher zu interpretieren sein als die Verteilungsparameter.

Von den Kovariaten haben das BIP und der Beitragssatz in beiden Modellen vergleichbare Koeffizienten.

Der Koeffizient der Kovariaten Qualifikation unterscheidet sich dagegen deutlich zwischen Modell 1 und Modell 2. Dies dürfte mit der hohen Korrelation der Qualifikationsvariablen mit den Parametern der Altersverteilung zusammenhängen. Der Korrelationskoeffizient von Universitätsabschluss und Standardabweichung beträgt -0,95; der von Universitätsabschluss und Wölbung 0,7. Welcher inhaltlicher Zusammenhang sich dahinter verbirgt, ist unklar. Denkbar wäre das „Durchlaufen“ mehr oder weniger gut ausgebildeter Kohorten durch das Beschäftigungssystem. Genauso gut könnte aber auch lediglich ein gemeinsamer Trend vorliegen.

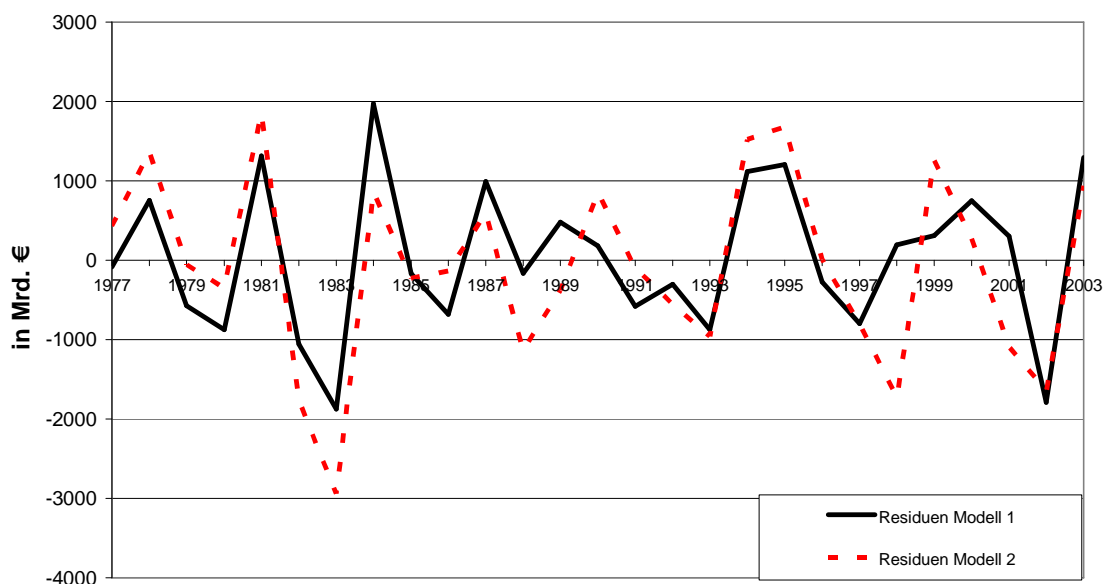
Im Folgenden soll untersucht werden, ob aus statistischer Sicht ein Modell vorzuziehen wäre.

Das korrigierte Bestimmtheitsmaß (adjusted R-squared in den Tabellen) beider Modelle ist vergleichbar gut: das korrigierte  $R^2$  von Modell 1 ist 0,68, das von Modell 2 0,66. In den Informationskriterien von Akaike und von Schwarz, die wie das korrigierte  $R^2$  die unterschiedliche Variablenzahl bei der Beurteilung der Güte der Schätzung berücksichtigen, weichen die Modelle kaum voneinander ab.

Vergleicht man die Entwicklung der geschätzten Differenzen über die Zeit, dann zeigt sich erneut eine weitgehende Übereinstimmung der Schätzungen beider Modelle mit den tatsächlichen Werten. Die Veränderungen in den Beitragseinnahmen in 2002 und 2003 werden jedoch von beiden Modellen deutlich unterschätzt.

Anhand der Residuen der Modelle gewinnt man den Eindruck, dass die flexiblere Modellierung mit Altersklassen (Modell 1) die Entwicklung genauer abbildet als Modell 2, weil die „Peaks“ schwächer ausfallen (Abb. 8).

Abb. 8: Vergleich der Residuen der Modelle 1 und 2

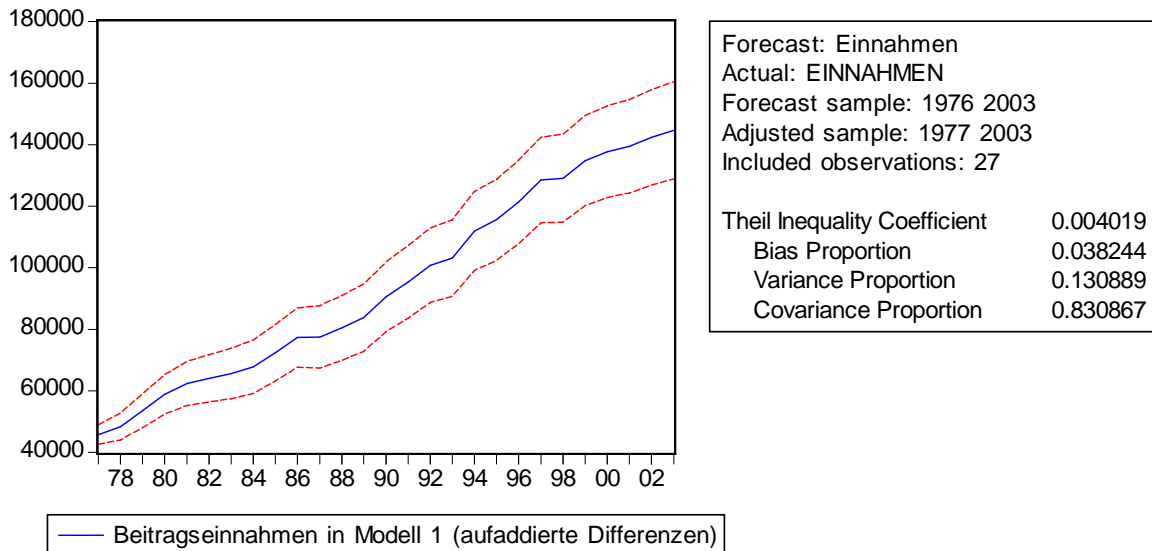


Anmerkung: Residuum = tatsächlicher Wert - Schätzwert  
Quelle: Eigene Berechnungen.

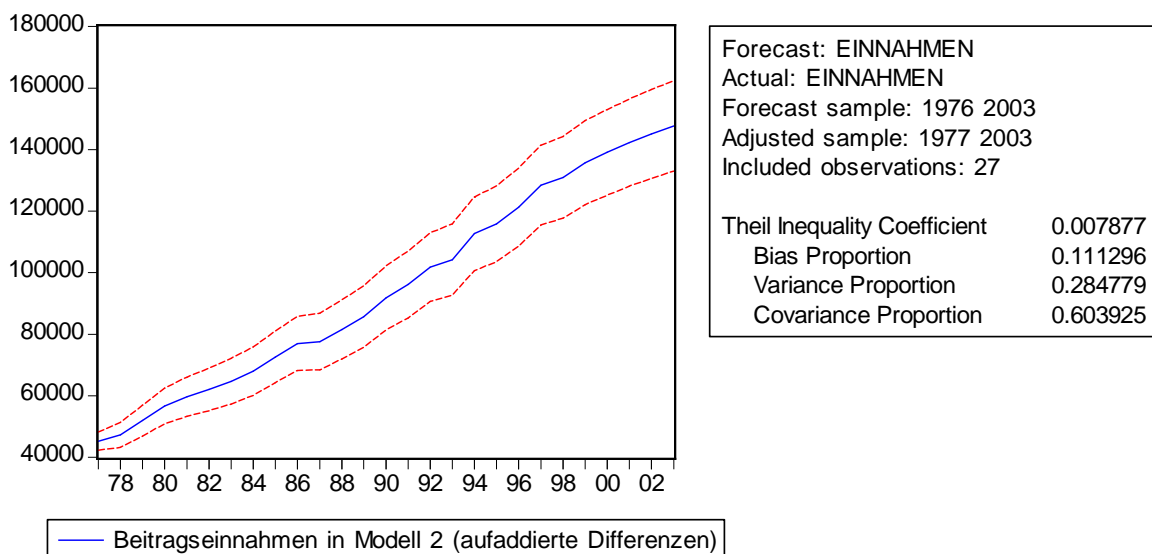
Mit den fortgeschriebenen geschätzten Differenzen lassen sich Niveaus errechnen, also die Einnahmen in einem Jahr. Grafisch sind keine relevanten Unterschiede zwischen den Modellen erkennbar (Abb. 9a und 9b).

**Abb. 9: Schätzungen für die jährlichen Einnahmen der GRV aus Modellen mit ersten Differenzen**

**a) Entwicklung des Niveaus der Einnahmen auf der Basis aufaddierter Differenzen: Modell 1 (Altersklassen)**



**b) Entwicklung des Niveaus der Einnahmen auf der Basis der aufaddierten Differenzen: Modell 2 (Parameter der Altersverteilung)**



*Erläuterung:*  
*durchgezogene Linie Schätzwert,*  
*gestrichelte Linien oberes und unteres Konfidenzintervall*  
*Quelle: Eigene Berechnungen*

Auch die Konfidenzintervalle sind vergleichbar breit. Bei dieser Fortschreibung addieren sich auch die Fehler für die einzelnen Jahre, was an der zunehmenden Breite der Konfidenzintervalle sichtbar wird.

Die Güte dieser Schätzung kann mit „Theil’s Inequality Coefficient“  $U$  quantifiziert werden. Theil’s  $U$  wird meist zur Beurteilung der Prognosefähigkeit von Modellen verwendet. Es misst, wie sehr zwei Zeitreihen voneinander abweichen, z.B. eine Schätzung oder Prognose von den tatsächlichen Werten. Theil’s  $U$  liegt immer zwischen Null (perfekte Prognose) und Eins (maximaler Fehler) (Pindyck/Rubinfeld 1998: 210 f.).<sup>13</sup>

Man kann Theil’s  $U$  in drei Fehlerkomponenten zerlegen: Die Komponente „Bias proportion“ bezeichnet den systematischen Fehler, der durch die Abweichung der durchschnittlichen Schätzung vom empirischen Durchschnitt entsteht. Dieser Fehler sollte bei beiden Modellen nahe Null sein, weil der Durchschnitt der geschätzten Differenzen mit dem empirischen Mittelwert der Differenzen übereinstimmt. Die Komponente „Variance Proportion“ zeigt die Fähigkeit eines Modells an, die Variabilität richtig abzubilden. In der „Covariance Proportion“ kommt der Anteil des Zufallsfehlers zum Ausdruck.

Theil’s  $U$  wurde hier für die Niveaus berechnet, die sich anhand der aufaddierten geschätzten jährlichen Veränderungen ergeben. Bezüglich Theil’s  $U$  selbst ist das parametrische Modell 2 kaum schlechter als das Modell 1 mit Altersklassen (siehe Abbildung 9a und 9b).

Die Komponentenzersetzung deutet jedoch auf eine spezifische Überlegenheit des flexiblen Altersklassenmodells hin. Beim Modell 1 (mit Altersklassen) sind 83 % des Fehlers unsystematischer Natur, also Zufallsfehler. Beim parametrischen Modell sind es nur 60 %. Ein Großteil des Fehlers des parametrischen Modells ist auf die Varianzkomponente zurückzuführen. Im Durchschnitt kann das parametrische Modell die Entwicklung durchaus abbilden, aber die Schätzungen für einzelne Jahre sind eher ungenau. Dies stimmt mit dem Eindruck überein, den bereits Abbildung 8 anhand der Residuen vermittelt.

$$U = \frac{\sqrt{\frac{1}{n} \sum (\hat{Y}_i - Y_i)^2}}{\sqrt{\frac{1}{n} \sum \hat{Y}_i^2 + \sqrt{\frac{1}{n} \sum Y_i^2}}}$$

<sup>13</sup> Theil’s  $U$  ist wie folgt definiert:  $U = \frac{\sqrt{\frac{1}{n} \sum (\hat{Y}_i - Y_i)^2}}{\sqrt{\frac{1}{n} \sum \hat{Y}_i^2 + \sqrt{\frac{1}{n} \sum Y_i^2}}}$ , mit den Schätzwerten  $\hat{Y}$  und den tatsächlichen Werten  $Y$ , jeweils für die Periode  $i$ .  $U$  setzt den durchschnittlichen Fehler in Bezug auf die Summe aus empirischer und geschätzter Varianz.

## 4.2 Modelle mit Bestandsgrößen: ein Kointegrationsansatz

Die bisher in ersten Differenzen geschätzten Modelle sind durchaus sinnvoll interpretierbar. Oft aber wünscht sich der Nutzer Aussagen über die Niveaus der Variablen. Ein einfaches Aufaddieren der ersten Differenzen ist nicht nur mühsam, sondern führt – wie gesehen – bereits nach einigen Perioden zu einer viel größeren Unsicherheit, weil sich die Schätzfehler des Differenzenmodells addieren.

Nun ist theoretisch zu erwarten, dass die Höhe der Beitragseinnahmen der GRV vom BIP, vom Beitragssatz und eventuell von der Altersstruktur abhängt. Im Folgenden wird deshalb ein Kointegrationsmodell geschätzt, das auf der Engle/Granger-Methodologie aufbauend einen Zusammenhang zwischen den Niveaus der Variablen untersucht. Ziel war es, erste Informationen über mögliche langfristige Zusammenhänge zu erhalten.<sup>14</sup>

Dieselbe Modellstruktur, die in ersten Differenzen geschätzt wurde, haben wir nun noch einmal für die Niveaus analysiert.

Inhaltlich sind die Modelle ähnlich. Das in Tabelle 3 angegebene Modell 3 für die Niveaus entspricht dem in Kapitel 4.1 analysierten Modell 1 für die Differenzen beim Altersklassenansatz. Das in Tabelle 4 dargestellte parametrische Modell 4 für Niveaus entspricht dem parametrischen Modell 2 für Differenzen. Weil die Modelle nicht frei von Autokorrelation sind, erfolgte die Schätzung wieder mit robusten Standardfehlern (Newey-West), außerdem enthält Modell 4 einen autoregressiven Term 1. Ordnung. Dies deutet – wie schon bei den Modellen mit Differenzen – auf das Fehlen irgendeiner wichtigen Einflussgröße hin.

Die Signifikanz der meisten Altersvariablen – sowohl von Modell 3 als auch von Modell 4 – ist gegenüber den Differenzenmodellen 1 und 2 etwas schlechter. Das extrem hohe  $R^2$  in den Tabellen darf nicht überinterpretiert werden. Dies ist für in Niveaus gemessene Variable nicht untypisch.

---

<sup>14</sup> Weitergehende Analysen (z.B. ein Fehlerkorrekturmodell, andere methodische Ansätze) waren aufgrund der Kürze der Bearbeitungszeit, die für die Vorstudie veranschlagt war, nicht möglich.

**Tabelle 3: Schätzung mit den Niveaus der Zeitreihen – Modell 3 mit Altersklassen**

Dependent Variable: EINNAHMEN

Method: Least Squares

Date: 06/26/07 Time: 09:01

Sample: 1976 2003

Included observations: 28

Newey-West HAC Standard Errors &amp; Covariance (lag truncation=3)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	184468.9	61908.83	2.979687	0.0094
SVB_20_24	-73269.66	93106.47	-0.786945	0.4436
SVB_25_29	-467651.6	160088.9	-2.921200	0.0105
SVB_30_34	-96814.46	73576.41	-1.315836	0.2080
SVB_35_39	-369196.5	75914.18	-4.863341	0.0002
SVB_40_44	-415670.3	137265.7	-3.028217	0.0085
SVB_45_49	-575415.8	121315.7	-4.743127	0.0003
SVB_50_54	-294503.3	164147.6	-1.794137	0.0930
SVB_55_59	88384.57	127867.4	0.691221	0.5000
SVB_60_65	335526.7	92804.54	3.615413	0.0025
BS_RV	2428.219	776.6834	3.126395	0.0069
BIP	21.17042	5.385620	3.930916	0.0013
QUAL_UNI	1636352.	241156.4	6.785437	0.0000
R-squared	0.999511	Mean dependent var		92554.97
Adjusted R-squared	0.999119	S.D. dependent var		33136.14
S.E. of regression	983.3718	Akaike info criterion		16.92427
Sum squared resid	14505300	Schwarz criterion		17.54279
Log likelihood	-223.9398	F-statistic		2553.514
Durbin-Watson stat	2.272742	Prob(F-statistic)		0.000000

*Erläuterungen:**SVB\_20\_24 ist der Anteil der 20- bis 24jährigen an allen SV-Beschäftigten; entsprechendes gilt für die übrigen Altersgruppen**BS\_RV: Beitragssatz zur GRV**BIP: Bruttoinlandsprodukt**Qual\_Uni : Anteil der SV-Beschäftigten mit (Fach-)Hochschulabschluss**Quelle: Eigene Berechnungen.*



**Tabelle 4: Schätzung mit den Niveaus der Zeitreihen – Modell 4 mit den Parametern der Altersverteilung**

Dependent Variable: EINNAHMEN

Method: Least Squares

Sample (adjusted): 1977 2003

Included observations: 27 after adjustments

Convergence achieved after 22 iterations

Newey-West HAC Standard Errors & Covariance (lag truncation=2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-185552.9	103974.9	-1.784592	0.0895
SVB_AGE_KURT	35637.36	17695.55	2.013916	0.0577
SVB_AGE_STDDEV	13800.66	9121.529	1.512976	0.1459
BS_RV	2658.483	547.9716	4.851498	0.0001
BIP	36.98030	6.870231	5.382687	0.0000
QUAL_UNI	843382.0	177252.6	4.758078	0.0001
AR(1)	0.593859	0.153419	3.870832	0.0010
R-squared	0.998462	Mean dependent var		94383.10
Adjusted R-squared	0.998001	S.D. dependent var		32296.41
S.E. of regression	1443.909	Akaike info criterion		17.60651
Sum squared resid	41697462	Schwarz criterion		17.94247
Log likelihood	-230.6879	F-statistic		2164.625
Durbin-Watson stat	1.829116	Prob(F-statistic)		0.000000

*Erläuterungen:*

*SVB\_AGE\_Stddev: Standardabweichung der Altersverteilung der SVB*

*SVB\_AGE\_Kurt: Kurtosis (Wölbung) der Altersverteilung der SVB*

*BS\_RV: Beitragssatz zur GRV*

*BIP: Bruttoinlandsprodukt*

*Qual\_Uni : Anteil der SV-Beschäftigten mit (Fach-)Hochschulabschluss*

*AR(1): Autoregressiver Term 1. Ordnung*

*Quelle: Eigene Berechnungen.*

Eine langfristige Beziehung zwischen der endogenen Variablen und den exogenen Variablen liegt vor, wenn für die Niveaus der Variablen ein Test auf Kointegration positiv ausfällt. Für diesen Test haben wir das Engle/Granger-Verfahren angewandt. Angenommen wird dazu die bereits früher angenommene Kausalstruktur, und dass die Regressoren selbst untereinander nicht kointegriert sind.

Die Kointegrationseigenschaft wird beim Engle/Granger-Verfahren anhand der Residuen festgestellt. Sind diese stationär, dann sind die Niveaus der Variablen kointegriert. Einen ersten Hinweis darauf gibt im Übrigen bereits die Durbin-Watson-Statistik (DW), die im Falle fehlender Kointegration ge-

gen Null strebt. Die DW-Werte beider Modelle sind weit von Null entfernt (siehe Tabellen 3 und 4).

Die Residuen aus Modell 3 bzw. Modell 4 wurden mit dem Dickey-Fuller-Test auf Stationarität getestet. Allerdings ist eine andere Testverteilung für die kritischen Werte erforderlich, da der Kointegrationstest von der Zahl der stochastischen Regressoren der Ausgangsgleichung abhängt (siehe Hassler 2004). Die kritischen Werte wurden anhand der bei McKinnon (1991: 267 ff.) angegebenen Parameter errechnet, wobei der Beitragssatz nicht als stochastische Variable aufgefasst wurde.

#### **Tabelle 5: Engle/Granger-Test auf Kointegration**

$H_0$ : die Residuen sind nicht stationär (hier: nicht kointegriert)

	Teststatistik	Testmodell
Modell 3 (Altersklassen)	-7,843285 *	k=11, Lag=3 n=24; kritischer Wert: -7,092
Modell 4 (Parameter)	-4,966011 *	k=5; Lag=0, n=26; kritischer Wert -4,55

\* Signifikant auf dem 10 %-Niveau; kritische Werte für den Kointegrationstest aus McKinnon 1991: 267ff.; für  $k = 11$  mittels linearer Regression extrapoliert<sup>15</sup>

$k$  = Zahl der stochastischen Variablen in der Schätzgleichung für die Niveaus

Lag-Länge des ADF-Tests

$n$  = Fallzahl (Zeitreihenwerte)

Quelle: Eigene Berechnungen.

Für beide Modelle kann die Nullhypothese, die Residuen sind nicht stationär, wohl abgelehnt werden. Dies spricht für die Existenz einer langfristigen Beziehung zwischen den beiden Seiten der Modellgleichung.

Dieses Modellergebnis bedarf allerdings zumindest im Falle des Altersklassenmodells noch einer gründlicheren Untersuchung, denn die Altersklassen könnten gemeinsam kointegriert sein. Insofern wäre die obige Modellannahme verletzt. Dies zu untersuchen war im Rahmen der vorliegenden Arbeit nicht möglich. (Es sei daran erinnert, dass 10 Altersklassen und nur 28 Werte vorliegen. Ein VAR-basierter Kointegrationstest kommt damit nicht in Frage.)

<sup>15</sup> Für mehr als 6 Regressoren standen keine Angaben zur Berechnung des kritischen Wertes zur Verfügung. Dieser wurde deshalb mit einem linearen Regressionsmodell auf der Basis der bei McKinnon angegebenen Daten extrapoliert ( $R^2=0,996$ ).

Ein weiteres methodisches Problem tritt auf, weil die parametrischen Altersvariablen Standardabweichung und Wölbung eher stationär sind – der Rest der Variablen nicht. Damit stehen rechts und links vom Gleichheitszeichen Variablen mit unterschiedlichen Differenzierungsgrad.

Die Verwendung von Niveauvariablen zur Analyse des Zusammenhangs von Beitragseinnahmen und Alterstruktur muss beim gegenwärtigen Stand der Analyse somit wohl noch kritisch hinterfragt werden.

## **5 Modelle mit der Altersstruktur der Bevölkerung bzw. des Erwerbspersonenpotenzials**

Die bisherigen Analysen beziehen sich auf die Altersstruktur der Beschäftigten in der Vergangenheit. Im Vergleich dazu werden für die Zukunft erheblich stärkere altersstrukturelle Verschiebungen erwartet. Es stellt sich die Frage, wie sich dies auf die Einnahmen der GRV auswirken könnte. Eine „Prognose“ der Altersstruktur der Beschäftigten liegt allerdings aktuell nicht vor. Außerdem wäre der Aufwand erheblich und die Resultate dürften sowieso nur sehr eingeschränkt zu interpretieren sein (siehe Fuchs 1992).

Es ist anzunehmen, dass sich die Altersstruktur der Beschäftigten nicht völlig abkoppeln kann von der der Bevölkerung oder der des Erwerbspersonenpotenzials. Tatsächlich korrelieren die Altersvariablen der Beschäftigten mit denen der Bevölkerung und des Erwerbspersonenpotenzials fast ausnahmslos sehr hoch. Lediglich bei den Ältesten (Altersgruppe 60-65 beziehungsweise 50-65) sowie beim Median und der Schiefe liegt der Korrelationskoeffizient unter 0,8 (siehe Korrelationsmatrix im Anhang).

Dies eröffnet eine interessante Perspektive. Für die Bevölkerung resp. das Erwerbspersonenpotenzial liegen langfristige Projektionen vor, bei denen insbesondere die Altersstruktur als gut gesichert angesehen werden kann, denn Szenarien mit erheblich unterschiedlichen Annahmen führen zu einer weitgehend ähnlichen alterstrukturellen Gliederung von Bevölkerung und Erwerbspersonenpotenzial (Fuchs/Dörfler 2005, Fuchs/Söhnlein 2005).

Im Folgenden wird deshalb der Frage nachgegangen, ob sich die gefundenen Modellbeziehungen auf die Bevölkerung oder das Erwerbspersonenpotenzial übertragen lassen. Dazu wurden die Modelle 1 und 2 noch einmal

mit den entsprechenden Altersvariablen der Bevölkerung bzw. des Erwerbspersonenpotenzials gerechnet.

## Bevölkerung

Zunächst seien die Schätzergebnisse für die Bevölkerung dargestellt; wieder für die 1. Differenzen der Altersklassen (Tabelle 6).<sup>16</sup>

**Tabelle 6: Schätzergebnisse für den Einfluss der Altersklassen der Bevölkerung im erwerbsfähigen Alter**

Dependent Variable: D(EINNAHMEN)

Method: Least Squares

Sample (adjusted): 1977 2003

Newey-West HAC Standard Errors & Covariance (lag truncation=2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(BEV_20_24)	-666136.7	332340.7	-2.004379	0.0648
D(BEV_25_29)	-708746.1	261240.9	-2.712998	0.0168
D(BEV_30_34)	-1360501.	368696.4	-3.690031	0.0024
D(BEV_35_39)	-969792.8	333857.5	-2.904810	0.0115
D(BEV_40_44)	-1165420.	378327.0	-3.080457	0.0081
D(BEV_45_49)	-1252664.	341212.7	-3.671212	0.0025
D(BEV_50_54)	-970597.5	429456.3	-2.260061	0.0403
D(BEV_55_59)	-509405.1	472989.1	-1.076991	0.2997
D(BEV_60_65)	197095.0	347237.1	0.567609	0.5793
D(BEV_AGE_STDDEV)	-61415.12	21546.51	-2.850352	0.0128
W_BIP	49830.76	11195.88	4.450814	0.0005
D(BS_RV)	1792.725	479.7742	3.736601	0.0022
D(QUAL_UNI)	833747.4	411531.0	2.025965	0.0623
R-squared	0.787199	Mean dependent var	3787.464	
Adjusted R-squared	0.604797	S.D. dependent var	2233.930	
S.E. of regression	1404.363	Akaike info criterion	17.63874	
Sum squared resid	27611302	Schwarz criterion	18.26266	
Log likelihood	-225.1230	Durbin-Watson stat	2.667915	

*Erläuterung:*

*D(..): Veränderung der Variable gegenüber dem Vorjahr (1. Differenz)*

*BEV\_20\_24 ist der Anteil der 20- bis 24jährigen in der Bevölkerung; entsprechendes gilt für die übrigen Altersgruppen*

*BS\_RV: Beitragssatz zur GRV*

*W\_BIP: Wachstumsrate des BIP*

*Qual\_Uni : Anteil der SV-Beschäftigten mit (Fach-)Hochschulabschluss*

*Quelle: Eigene Berechnungen.*

<sup>16</sup> Die Bevölkerungsdaten stammen vom Statistischen Bundesamt. Im Gegensatz zu den Beschäftigtenzahlen beziehen sie sich auf den 31.12. d.J.

Das Modell unterscheidet sich bezüglich der Kovariaten BIP, Beitragssatz und Qualifikation wenig von Modell 1, das mit den Beschäftigtenzahlen gerechnet wurde. Zusätzlich wurde die Altersvariable für die Standardabweichung in die Gleichung aufgenommen. Ohne sie wäre die Signifikanz der Altersvariablen erheblich schlechter.

Wichtiger ist aber, dass es Anzeichen für das Fehlen bedeutsamer Einflussgrößen gibt. Die autoregressiven Terme und der Test auf Fehlspezifikation (Reset) sind alle signifikant. Möglicherweise reicht die Bevölkerungsentwicklung nicht für die Erklärung der Beitragseinnahmen aus.

### **Erwerbspersonenpotenzial**

Das Erwerbspersonenpotenzial ist ein Maß für das Arbeitskräfteangebot. Es umfasst die Erwerbstätigen, die Erwerbslosen und die sogenannte Stille Reserve („verdeckte Arbeitslosigkeit“). Die aktuellste verfügbare Schätzung wurde vom IAB im Jahr 2005 veröffentlicht (Fuchs/Weber 2005). Das Erwerbspersonenpotenzial ist arbeitsmarktnäher als die Bevölkerung. Dies zeigt sich in den höheren bivariaten Korrelationen (siehe Anhang A4), die für die Altersvariablen der Beschäftigten, der Bevölkerung und des Erwerbspersonenpotenzials berechnet wurden. Insofern wäre denkbar, dass mit dem Erwerbspersonenpotenzial bessere Ergebnisse erzielt werden als mit der Bevölkerung.

Die Schätzung für das Erwerbspersonenpotenzial enthält Tabelle 7. Abhängige Variable ist wieder die jährliche Veränderung der Beitragseinnahmen. Als Regressoren wurden diesmal die Veränderungen in der Altersstruktur des Erwerbspersonenpotenzials eingesetzt. Die gemeinsame Signifikanz der Koeffizienten der Altersstrukturvariablen ist kaum besser oder schlechter als im Falle der Bevölkerung.

Im Modell mit dem Erwerbspersonenpotenzial wechselt gegenüber dem Modell mit den Beschäftigten das Vorzeichen der meisten Altersvariablen. Fast alle Koeffizienten der Altersklassen sind nun positiv, während beim entsprechenden Modell mit den Beschäftigten die Mehrzahl der Koeffizienten negativ war. Beim Erwerbspersonenpotenzial trägt also ein Anstieg des Anteils einer jeden einzelnen Altersklasse zunächst zu mehr Beitragseinnahmen bei. Die Nettowirkung ergibt sich wieder aus dem Gesamtzusammenhang, weil ein Mehr in einer Altersklasse mit einem sin-

kenden Anteil einer anderen einhergehen muss (außer der Anteil der Basisklasse der 15- bis 19-Jährigen sinkt).

**Tabelle 7: Schätzergebnisse für den Einfluss der Altersklassen des Erwerbersonenpotenzials**

Dependent Variable: D(EINNAHMEN)

Method: Least Squares

Sample (adjusted): 1977 2003

Newey-West HAC Standard Errors & Covariance (lag truncation=2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6067.170	2530.687	-2.397440	0.0322
D(EPP_20_24)	467600.6	258728.5	1.807302	0.0939
D(EPP_25_29)	39553.59	113209.3	0.349385	0.7324
D(EPP_30_34)	969370.4	288629.1	3.358533	0.0051
D(EPP_35_39)	697303.7	227222.7	3.068812	0.0090
D(EPP_40_44)	752734.0	212529.8	3.541781	0.0036
D(EPP_45_49)	1134160.	359109.5	3.158257	0.0076
D(EPP_50_54)	515017.2	236906.3	2.173927	0.0488
D(EPP_55_59)	64321.62	174594.4	0.368406	0.7185
D(EPP_60_64)	-271307.7	263566.0	-1.029373	0.3221
D(EPP_ANZ)	6.176515	2.113085	2.922985	0.0119
W_BIP	75683.98	40260.77	1.879845	0.0827
D(BS_RV)	1520.759	861.2360	1.765786	0.1009
D(QUAL_UNI)	1165703.	359566.5	3.241968	0.0064
R-squared	0.808806	Mean dependent var		3787.464
Adjusted R-squared	0.617612	S.D. dependent var		2233.930
S.E. of regression	1381.406	Akaike info criterion		17.60574
Sum squared resid	24807690	Schwarz criterion		18.27766
Log likelihood	-223.6775	F-statistic		4.230294
Durbin-Watson stat	2.425214	Prob(F-statistic)		0.007077

*Erläuterungen:*

*D(..): Veränderung der Variable gegenüber dem Vorjahr (1. Differenz)*

*EPP\_20\_24 ist der Anteil der 20- bis 24jährigen am Erwerbersonenpotenzial; entsprechendes gilt für die übrigen Altersgruppen*

*BS\_RV: Beitragssatz zur GRV*

*W\_BIP: Wachstumsrate des BIP*

*Qual\_Uni : Anteil der SV-Beschäftigten mit (Fach-)Hochschulabschluss*

*Quelle: Eigene Berechnungen.*

Wie beim Modell mit der Bevölkerung könnten wichtige Einflussgrößen fehlen. Eine denkbare Variable wäre die Arbeitslosigkeit. Wir haben dies sowohl für das Modell mit der Bevölkerung als auch mit dem Erwerbersonenpotenzial geprüft. In beiden Fällen trägt die Arbeitslosenquote nicht zur Verbesserung des Modells bei.

Die vorliegenden Ergebnisse sprechen nur eingeschränkt für die Möglichkeit, im Rahmen der hier modellierten Beitragseinnahmen die SV-Beschäftigten durch die Bevölkerung oder das Erwerbspersonenpotenzial zu substituieren. Insbesondere muss die Stabilität der Beziehungen noch intensiver geprüft werden. Denkbar wäre aber immer noch, die vorausgeschätzte Struktur (des Erwerbspersonenpotenzials oder der Bevölkerung) als Hilfsgröße im Rahmen einer längerfristigen Projektion zu verwenden – gewissermaßen als Proxy für die Altersstruktur der SV-Beschäftigten. Man würde in diesem Fall auf die für die SV-Beschäftigten ermittelten Gleichungen zurückgreifen.

## 6 Fazit

Mit der vorliegenden Studie wurde die Frage untersucht, ob sich altersstrukturelle Verschiebungen bei den sozialversicherungspflichtig Beschäftigten auf die Beitragseinnahmen der GRV auswirken könnten. Die Berechnungen wurden mit ersten Differenzen durchgeführt. Als Ergebnis stellte sich heraus, dass die jährliche Veränderung der Beitragseinnahmen vom Wirtschaftswachstum, dem Beitragssatz zur GRV und der Qualifikationsstruktur der Beschäftigten abhängt.

Die zwei verschiedenen Ansätze, mit denen der Einfluss der Altersstruktur modelliert wurde, belegen, dass sich bestimmte Eigenschaften der Altersstruktur der Beschäftigten auf die Einnahmen der GRV auswirken. Insbesondere wirkt ein höherer Anteil Älterer positiv auf die Einnahmenseite. Das Modell, das den Alterseinfluss mit Altersklassen schätzte, war dem alternativ berechneten Modell, das die Parameter der Altersverteilung als erklärende Variable zugrunde legte, tendenziell überlegen. Es kann vor allem die Variabilität der Einnahmenentwicklung genauer abbilden.

Weiterverfolgen sollte man die Modellierung mit den Niveaus der Variablen, die hier im Rahmen einer Kointegrationsmethodologie erfolgte. Möglicherweise bietet sich ein Fehlerkorrekturmodell an, mit dem Bestände und Veränderungen zusammen in einem Modell geschätzt werden. Dadurch ließen sich sowohl Aussagen über die langfristige Beziehung als auch die kurzfristige Dynamik treffen.

Sollte die Übertragung auf das Erwerbspersonenpotenzial bzw. die Bevölkerung mit weitergehenden Analysen gestützt werden, dann würde sich

der geschätzte Modelltypus sogar für „Prognosen“ eignen, bei denen man das prognostizierte Erwerbspotenzial oder die prognostizierte Bevölkerung vorgibt.

Es gibt aber eine Reihe offener Fragen, die unbedingt noch beantwortet werden müssen:

Die Beitragseinnahmen enthalten auch Beiträge von nicht-sozialversicherungspflichtig Beschäftigten. Insofern passen rechte und linke Seite des Regressionsmodells nicht zusammen. Vermutlich wäre eine Modellbildung mit den Beitragseinnahmen aus sozialversicherungspflichtiger Beschäftigung vorzuziehen. Damit könnte sich möglicherweise die serielle Korrelation in den Modellen verringern. Alternativ müsste eine Reihe weiterer Einflussfaktoren als exogene Variablen in die Modellbildung eingehen. Zu denken wäre dabei vor allem an den Einfluss der Arbeitslosigkeit.

In den hier geschätzten Modellen war der Beitragssatz zur GRV eine exogene Variable. Näher untersucht werden sollte aber vielleicht die wechselseitige Abhängigkeit von Einnahmen und Ausgaben der GRV.

Die Analysen basierten auf westdeutschen Zeitreihen. Es fanden sich keine Belege für einen Strukturbruch aufgrund der Wiedervereinigung, obwohl einzelne Zeitreihen (BIP, Lohnquote) nach 1990 nur für Gesamtdeutschland vorliegen. Zu prüfen wäre, ob sich die Modelle sehr ändern, wenn man bis 1990 westdeutsche Werte und ab 1991 gesamtdeutsche Werte zugrunde legt.

Ein methodisches Problem verursacht der Integrationsgrad der Zeitreihen, da nicht alle Zeitreihen vom selben Grad integriert sind. Vor allem die Modellbildung mit Altersklassen sollte in statistischer Hinsicht überdacht werden. Möglicherweise sind die Altersklassen als exogene Variablengruppe kointegriert. Ein alternativer Schätzansatz (z.B. im Rahmen eines VAR-Modells) wurde aufgrund der beschränkten Fallzahl nicht weiter verfolgt.



## Literatur

- Blos, Kerstin (2006): Die Bedeutung der Ausgaben und Einnahmen des Sozialversicherungssysteme für die Regionen in Deutschland. IAB-Forschungsbericht Nr. 8/2006.
- Börsch-Supan, Axel (2003): Gesamtwirtschaftliche Folgen des demographischen Wandels. IAW-Report, 31 (1), 1-26.
- BMAS (Bundesministerium für Arbeit und Soziales) (2006): Rentenversicherungsbericht 2006.
- Fuchs, Johann (1992): Zu- und Abgangsrechnung für die Erwerbstätigen nach Branchen. Ein Modell zur Fortschreibung sektoraler Beschäftigungsstrukturen nach Alter, Geschlecht und Stellung im Beruf. Nürnberg, Beiträge zur Arbeitsmarkt- und Berufsforschung (BeitrAB) 166.
- Fuchs, Johann; Dörfler, Katrin (2005): Projektion des Erwerbspersonenpotenzials bis 2050 \* Annahmen und Datengrundlage. IAB-Forschungsbericht Nr. 25/2005.
- Fuchs, Johann; Söhnlein, Doris (2005): Vorausschätzung der Erwerbsbevölkerung bis 2050. IAB-Forschungsbericht Nr. 16/2005.
- Fuchs, Johann, Brigitte Weber (2005): Neuschätzung der Stillen Reserve und des Erwerbspersonenpotenzials für Westdeutschland (inkl. Berlin-West). IAB-Forschungsbericht 15/2005;
- Hassler, Uwe (2004):. Leitfaden zum Testen und Schätzen von Kointegration. In: Werner Gaab, Ullrich Heilemann, Jürgen Wolters (Hrsg.): Arbeiten mit ökonomischen Modellen. Heidelberg, S. 85-115.
- MacKinnon, James (1991): Critical Values for Cointegration Tests. In: Engle, Robert F.; Granger, Clive W. J. (eds): Long-run Economic Relationships. Readings in Cointegration. Oxford, S. 267-276.
- OECD (2005): Alterung und Beschäftigungspolitik: Deutschland.
- Quantitative Micro Software (Hrsg.) (2004): EViews 5 User's Guide.
- Pindyck, Robert S.; Rubinfeld, Daniel L. (1998): Econometric Models and Economic Forecasts, New York u.a., 4th edition.
- Ulrich, Volker; Erbsland, Manfred (1997): Der langfristige Zusammenhang zwischen demographischen und budgetären Variablen. Finanzarchiv N.F., Bd. 54, S. 203-232.

## Anhang

### Anhang A1: Variablenbeschreibung

	Variable	Bezeichnung	Quelle	Bemerkung
GRV	Einnahmen der RV	EINNAHMEN	DRV	
	Beitragssatz in der RV	BS_RV	DRV	
Alter	Anteil der 14- bis 19-Jährigen	SVB_14_19	BA	in den alten Ländern sozialversicherungspflichtig Beschäftigte
	20- bis 24-Jährigen	SVB_20_24	BA	
	25- bis 29-Jährigen	SVB_25_29	BA	
	30- bis 34-Jährigen	SVB_30_34	BA	
	35- bis 39-Jährigen	SVB_35_39	BA	
	40- bis 44-Jährigen	SVB_40_44	BA	
	45- bis 49-Jährigen	SVB_45_49	BA	
	50- bis 54-Jährigen	SVB_50_54	BA	
	55- bis 59-Jährigen	SVB_55_59	BA	
	60- bis 64-Jährigen	SVB_60_65	BA	
	14- bis 29-Jährigen	SVB_AGE14_29	BA	
	30- bis 49-Jährigen	SVB_AGE30_49	BA	
	50- bis 65-Jährigen	SVB_AGE50_65	BA	
Alter	Standardabweichung	SVB_AGE_STDDEV	BA	
	Mittelwert	SVB_AGE_MEAN	BA	
	Median	SVB_AGE_MEDIAN	BA	
	Schiefe	SVB_AGE_SKEW	BA	
	Wölbung (Kurtosis)	SVB_AGE_KURT	BA	
Struktur	Teilzeitquote	TZ	BA	
	Teilzeitquote (unter 20 Wochenstd.)	TZ__20	BA	
	Teilzeitquote (ab 20 Wochenstd.)	TZ_M20	BA	
	Ausländeranteil	SVB_AUSL_SOZPFL	BA	
	Frauenanteil	SVB_FRAUEN_SOZPFL	BA	
	Anteil mit Uni-/FH-Abschluss	QUAL_UNI	BA	
	Anteil mit Berufsausbildung	QUAL_BER	BA	
Anteil ohne berufl. Qualifikation	QUAL_OHNE	BA		
Konjunktur	Anzahl SV-Beschäftigte	SVB_ANZ_SOZPFL	BA	
	Bruttoinlandprodukt	BIP	StBA	ab 1992 Gesamtdeutschland
	Lohnquote	LQ	StBA	ab 1991 Gesamtdeutschland
	Arbeitslosenquote	ALOQUOTE	BA	Basis: zivile Erwerbspersonen

#### Hinweise:

Jahreswerte oder Stichtagswert 30.6. von 1976 bis einschl. 2003 für die alten Bundesländer

BA: Statistiken der Bundesagentur für Arbeit

DRV: Statistiken der Deutschen Rentenversicherung Bund

StBA: Statistiken des Statistischen Bundesamtes

Die Altersstrukturvariablen für die Bevölkerung und das Erwerbspersonenpotenzial sind analog definiert  
Quellenangabe für das Erwerbspersonenpotenzial: Fuchs/Weber 2005.

## Tabelle A2: Tests auf Stationarität

Augmented Dickey-Fuller Unit Root Test

Nullhypothese  $H_0$ : Variable X ist nicht stationär

	Variable	Niveau	erste Differenzen
<b>Altersklassen</b> 5er Gruppen	14 bis 19	-3,58* (t, 4)	-1,511* (o, 4)
	20 bis 24	-3,511* (t, 1)	-1,687 n.s. (c, 1)
	25 bis 29	-4,43*** (c, 3)	-3,374*** (o, 3)
	30 bis 34	-3,166** (c, 2)	-0,989 n.s. (c, 3)
	35 bis 39	-4,338** (t, 1)	-1,426 n.s. (c, 0)
	40 bis 44	-2,328 n.s. (c, 1)	-2,076** (o, 1)
	45 bis 49	-3,812*** (c, 1)	-2,267** (o, 1)
	50 bis 54	-3,47** (c, 1)	-2,027** (o, 1)
	55 bis 59	-2,768* (c, 1)	-3,125*** (o, 1)
	60 bis 64	0,679 n.s. (c, 4)	-5,845*** (o, 0)
	14 bis 29	-1,883* (o, 1)	-2,464** (o, 0)
	30 bis 49	-3,036* (t, 1)	-1,43 n.s. (o, 1)
	50 bis 65	-1,474 n.s. (c, 6)	-3,414*** (o, 4)
<b>Statistiken</b> der Altersverteilung	Mittelwert	-0,358 n.s. (t, 4)	-4,59*** (o, 0)
	Median	-1,741 n.s. (t, 0)	-6,46*** (o, 0)
	Standardabweichung	-4,292** (t, 4)	-2,824* (c, 6)
	Schiefe	-0,8 n.s. (o, 0)	-4,368** (t, 6)
	Kurtosis (Wölbung)	-4,209*** (c, 1)	-1,943* (o, 6)
<b>sonstige</b> <b>Strukturmerkmale</b> der SVB	Anzahl SV-Beschäftigte	-1,277 n.s. (c, 2)	-4,236*** (o, 1)
	Ausländeranteil	-2,934* (c, 3)	-2,794*** (o, 0)
	Frauenanteil	-3,024 n.s. (t, 2)	-1,689* (o, 0)
	Anteil mit Uni-/FH-Abschluss	-1,388 n.s. (c, 1)	-2,412** (o, 0)
	Anteil mit Berufsausbildung	-2,908*** (o, 1)	-3,167** (c, 0)
	Anteil ohne berufl. Qualifikation	-1,122 n.s. (t, 1)	6,603*** (c, 0)
	TZ Quote	-1,622 n.s. (t, 0)	-2,091** (o, 0)
<b>Konjunktur</b>	Bruttoinlandsprodukt (BIP)	-2,034 n.s. (t, 1)	-2,664* (c, 0)
	BIP-Wachstumsrate	-1,603 (*) (o, 0)	
	Lohnquote	-2,285 n.s. (c, 1)	-3,856*** (o, 0)
	Arbeitslosenquote	-2,919* (c, 1)	-3,425*** (o, 1)
<b>RV</b>	Beitragssatz RV	-2,692 n.s. (t, 0)	-5,472*** (o, 0)
	Einnahmen RV	-1,785 n.s. (t, 0)	-4,955*** (c, 0)

\*Signifikanzniveau: 10% = \*; 5% = \*\*, 1% = \*\*\*, nicht signifikant = n.s.

in Klammer gesetzt: Signifikanzniveau knapp überschritten (..)

Modell (o = ohne Konstante, c = Konstante, t = Konstante und Trend) und Anzahl der Lags

Quelle: Eigene Berechnungen.

## Anhang A3: Alternative Modellspezifikationen

**Tabelle A30: Basismodell ohne Altersvariable**

Dependent Variable: D(EINNAHMEN)

Method: Least Squares

Sample (adjusted): 1977 2003

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(BS_RV)	2382.190	905.6901	2.630248	0.0147
W_BIP	69348.38	10061.85	6.892211	0.0000
D(SVB_ANZ_SOZPFL)	0.000249	0.001570	0.158449	0.8754
R-squared	0.100302	Mean dependent var		3787.464
Adjusted R-squared	0.025327	S.D. dependent var		2233.930
S.E. of regression	2205.459	Akaike info criterion		18.33970
Sum squared resid	1.17E+08	Schwarz criterion		18.48368
Log likelihood	-244.5859	Durbin-Watson stat		1.223915

Erläuterungen: siehe Anhang A1

### Modelle mit Altersklassen

**Tabelle A31: Autoregressive Terme 1. und 2. Ordnung AR(1) und AR(2) als zusätzliche Variable gegenüber Gleichung 1**

Dependent Variable: D(EINNAHMEN)

Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SVB_20_24)	-159199.8	113646.3	-1.400835	0.1888
D(SVB_25_29)	-445552.9	160373.4	-2.778222	0.0180
D(SVB_30_34)	-207196.4	99015.75	-2.092560	0.0604
D(SVB_35_39)	-275188.8	110522.1	-2.489898	0.0300
D(SVB_40_44)	-401105.5	151936.0	-2.639963	0.0230
D(SVB_45_49)	-576378.4	130959.8	-4.401187	0.0011
D(SVB_50_54)	-346802.5	145809.8	-2.378458	0.0366
D(SVB_55_59)	43608.15	147835.2	0.294978	0.7735
D(SVB_60_65)	549049.6	162548.9	3.377750	0.0062
D(BS_RV)	2418.613	513.1652	4.713128	0.0006
W_BIP	46413.32	9662.482	4.803458	0.0006
D(QUAL_UNI)	1198088.	276303.0	4.336139	0.0012
AR(1)	-0.522865	0.235131	-2.223718	0.0481
AR(2)	-0.698707	0.246119	-2.838906	0.0161
R-squared	0.906315	Mean dependent var		3856.024
Adjusted R-squared	0.795595	S.D. dependent var		2306.354
S.E. of regression	1042.729	Akaike info criterion		17.03609
Sum squared resid	11960119	Schwarz criterion		17.71866
Log likelihood	-198.9511	Durbin-Watson stat		2.175091

**Tabelle A32: Standardabweichung der Altersverteilung (SVB\_age\_stddev) als zusätzliche Variable gegenüber Gleichung 1**

Dependent Variable: D(EINNAHMEN)

Newey-West HAC Standard Errors & Covariance (lag truncation=2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SVB_20_24)	-201168.1	146185.0	-1.376120	0.1904
D(SVB_25_29)	-678598.4	279433.6	-2.428478	0.0292
D(SVB_30_34)	-622190.8	200998.3	-3.095503	0.0079
D(SVB_35_39)	-783369.3	294847.4	-2.656863	0.0188
D(SVB_40_44)	-753401.1	281811.2	-2.673425	0.0182
D(SVB_45_49)	-848313.3	219612.6	-3.862771	0.0017
D(SVB_50_54)	-579587.6	174505.7	-3.321311	0.0050
D(SVB_55_59)	594.0949	154441.2	0.003847	0.9970
D(SVB_60_65)	771976.4	164912.1	4.681139	0.0004
D(BS_RV)	2034.408	572.3910	3.554227	0.0032
W_BIP	40970.10	11483.57	3.567716	0.0031
D(QUAL_UNI)	983212.7	142601.5	6.894827	0.0000
D(SVB_AGE_STDDEV)	-45478.31	19611.82	-2.318924	0.0360
R-squared	0.842696	Mean dependent var		3787.464
Adjusted R-squared	0.707864	S.D. dependent var		2233.930
S.E. of regression	1207.430	Akaike info criterion		17.33656
Sum squared resid	20410428	Schwarz criterion		17.96048
Log likelihood	-221.0436	Durbin-Watson stat		2.707283

Erläuterungen: siehe Anhang A1

**Tabelle A33: Arbeitslosenquote als zusätzliche Variable gegenüber Gleichung 1**

Dependent Variable: D(EINNAHMEN)

Method: Least Squares

Newey-West HAC Standard Errors &amp; Covariance (lag truncation=2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SVB_20_24)	-396933.2	132712.0	-2.990936	0.0097
D(SVB_25_29)	-469219.7	197584.0	-2.374786	0.0324
D(SVB_30_34)	-212443.6	97458.01	-2.179847	0.0468
D(SVB_35_39)	-440658.7	120709.6	-3.650570	0.0026
D(SVB_40_44)	-468521.4	153018.0	-3.061872	0.0084
D(SVB_45_49)	-356440.0	145541.6	-2.449060	0.0281
D(SVB_50_54)	-253097.7	130464.2	-1.939979	0.0728
D(SVB_55_59)	78732.57	115528.7	0.681498	0.5067
D(SVB_60_65)	450322.7	97297.76	4.628295	0.0004
D(BS_RV)	2187.494	450.6600	4.853979	0.0003
W_BIP	44493.95	8409.794	5.290731	0.0001
D(QUAL_UNI)	1168080.	152194.1	7.674935	0.0000
D(ALOQUOTE)	-1372.570	633.7882	-2.165661	0.0481
R-squared	0.864699	Mean dependent var		3787.464
Adjusted R-squared	0.748726	S.D. dependent var		2233.930
S.E. of regression	1119.807	Akaike info criterion		17.18588
Sum squared resid	17555538	Schwarz criterion		17.80980
Log likelihood	-219.0094	Durbin-Watson stat		2.581028

Erläuterungen: siehe Anhang A1

**Modelle mit Parametern der Altersverteilung****Tabelle A34: Schiefe der Verteilung (age\_skew) als zusätzliche Variable gegenüber Gleichung 2**

Dependent Variable: D(EINNAHMEN)

Method: Least Squares

Date: 06/19/07 Time: 11:09

Sample (adjusted): 1977 2003

Included observations: 27 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SVB_AGE_STDDEV)	25073.93	8967.363	2.796132	0.0108
D(SVB_AGE_KURT)	86259.39	20090.15	4.293616	0.0003
D(SVB_AGE_SKEW)	12461.38	20184.90	0.617361	0.5436
D(BS_RV)	1994.144	484.1920	4.118499	0.0005
W_BIP	62673.68	10511.41	5.962445	0.0000
D(QUAL_UNI)	782622.1	249973.5	3.130821	0.0051
R-squared	0.718604	Mean dependent var		3787.464
Adjusted R-squared	0.651605	S.D. dependent var		2233.930
S.E. of regression	1318.577	Akaike info criterion		17.39962
Sum squared resid	36511562	Schwarz criterion		17.68759
Log likelihood	-228.8949	Durbin-Watson stat		1.940379

**Tabelle A35: Mittelwert (age\_mean) und Schiefe der Verteilung (age\_skew) als zusätzliche Variable gegenüber Gleichung 2**

Dependent Variable: D(EINNAHMEN)

Method: Least Squares

Date: 06/19/07 Time: 11:12

Sample (adjusted): 1977 2003

Included observations: 27 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SVB_AGE_STDDEV)	24922.73	9171.531	2.717401	0.0133
D(SVB_AGE_KURT)	87691.21	20933.99	4.188940	0.0005
D(SVB_AGE_SKEW)	10365.43	21485.34	0.482442	0.6347
D(SVB_AGE_MEAN)	-535.6738	1541.633	-0.347472	0.7319
D(BS_RV)	1997.117	494.7324	4.036763	0.0006
W_BIP	63005.48	10780.99	5.844126	0.0000
D(QUAL_UNI)	784853.9	255457.7	3.072344	0.0060
R-squared	0.720293	Mean dependent var		3787.464
Adjusted R-squared	0.636380	S.D. dependent var		2233.930
S.E. of regression	1347.080	Akaike info criterion		17.46768
Sum squared resid	36292470	Schwarz criterion		17.80364
Log likelihood	-228.8137	Durbin-Watson stat		1.991042

Erläuterungen: siehe Anhang A1

## Anhang A4: Korrelationsmatrix Beschäftigte, Bevölkerung, Erwerbspersonenpotenzial

### Altersstruktur von Beschäftigten und Bevölkerung

Bevölkerung	Sozialversicherungspflichtig Beschäftigte									
	14 - 19	20 - 24	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 65
14 - 19	0.950	0.525	-0.517	-0.817	-0.355	0.274	0.087	-0.549	-0.505	0.176
20 - 24	0.734	0.997	0.129	-0.865	-0.885	-0.489	0.236	0.224	-0.618	-0.367
25 - 29	-0.081	0.404	0.962	0.004	-0.576	-0.825	-0.410	0.630	0.188	-0.693
30 - 34	-0.877	-0.701	0.498	0.963	0.392	-0.072	-0.316	0.207	0.811	-0.233
35 - 39	-0.555	-0.851	-0.277	0.598	0.988	0.392	-0.154	-0.217	0.228	0.634
40 - 44	0.067	-0.466	-0.733	0.064	0.503	0.969	-0.164	-0.439	-0.011	0.518
45 - 49	0.230	0.289	-0.483	-0.324	-0.203	0.000	0.980	-0.295	-0.295	0.102
50 - 54	-0.205	0.357	0.512	-0.116	-0.253	-0.521	-0.229	0.942	-0.369	-0.239
55 - 59	-0.547	-0.504	0.566	0.724	0.130	-0.100	-0.403	-0.014	0.964	-0.509
60 - 65	-0.771	-0.660	-0.199	0.673	0.654	0.198	0.374	0.067	0.208	0.394

Bevölkerung	Sozialversicherungspflichtig Beschäftigte			Wölbung	Mittelwert	Median	Schiefe	Standardabweichung
	14 - 29	30 - 49	50 - 65					
14 - 29	0.967	-0.918	-0.519	-0.864	-0.973	-0.806	0.183	0.992
30 - 49	-0.979	0.994	0.218	0.973	0.919	0.877	-0.384	-0.908
50 - 65	-0.678	0.550	0.807	0.464	0.776	0.478	0.157	-0.833
Wölbung	-0.921	0.925	0.257	0.986	0.858	0.833	-0.337	-0.848
Mittelwert	-0.878	0.799	0.635	0.712	0.929	0.711	-0.095	-0.952
Median	-0.875	0.811	0.566	0.752	0.914	0.793	-0.342	-0.893
Schiefe	0.949	-0.955	-0.255	-0.873	-0.932	-0.931	0.586	0.873
Standardabw.	0.738	-0.620	-0.778	-0.599	-0.778	-0.480	-0.252	0.879

### Altersstruktur von Beschäftigten und Erwerbspersonenpotenzial

Erwerbspersonenpotential	Sozialversicherungspflichtig Beschäftigte									
	14 - 19	20 - 24	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 65
14 - 19	0.979	0.707	-0.378	-0.892	-0.528	0.024	0.204	-0.451	-0.563	0.048
20 - 24	0.771	0.995	0.143	-0.874	-0.881	-0.499	0.192	0.179	-0.602	-0.369
25 - 29	-0.080	0.345	0.970	0.054	-0.545	-0.787	-0.474	0.549	0.278	-0.712
30 - 34	-0.905	-0.816	0.332	0.996	0.542	0.079	-0.253	0.106	0.791	-0.069
35 - 39	-0.537	-0.854	-0.340	0.579	0.996	0.467	-0.153	-0.228	0.208	0.640
40 - 44	-0.017	-0.512	-0.743	0.134	0.500	0.995	-0.065	-0.434	0.064	0.494
45 - 49	-0.012	0.119	-0.462	-0.136	-0.004	-0.004	0.987	-0.175	-0.267	0.177
50 - 54	-0.466	0.058	0.551	0.171	-0.060	-0.363	-0.356	0.967	-0.121	-0.194
55 - 59	-0.741	-0.720	0.364	0.891	0.376	0.089	-0.259	-0.047	0.963	-0.320
60 - 65	-0.516	-0.634	-0.464	0.451	0.730	0.410	0.341	-0.102	-0.061	0.725

Erwerbspersonenpotential	Sozialversicherungspflichtig Beschäftigte			Wölbung	Mittelwert	Median	Schiefe	Standardabweichung
	14 - 29	30 - 49	50 - 65					
14 - 29	0.992	-0.957	-0.460	-0.918	-0.980	-0.851	0.295	0.980
30 - 49	-0.983	0.997	0.230	0.968	0.932	0.904	-0.439	-0.909
50 - 65	-0.716	0.575	0.880	0.522	0.804	0.466	0.144	-0.860
Wölbung	-0.928	0.957	0.139	0.992	0.856	0.879	-0.476	-0.817
Mittelwert	-0.966	0.907	0.567	0.848	0.985	0.806	-0.251	-0.986
Median	-0.364	0.338	0.233	0.318	0.420	0.404	-0.567	-0.284
Schiefe	0.853	-0.851	-0.260	-0.779	-0.852	-0.909	0.718	0.759
Standardabw.	0.954	-0.903	-0.524	-0.894	-0.939	-0.800	0.176	0.969

Quelle: Eigene Berechnungen.