

FDZ-Arbeitspapier
Nr.10

Timm Bönke
Frank Neher
Carsten Schröder



STATISTISCHE ÄMTER
DES BUNDES UND DER LÄNDER
FORSCHUNGSDATENZENTREN

Konstruktion und Bewertung
eines ökonomischen
Einkommens aus der
Faktisch Anonymisierten
Lohn- und Einkommen-
steuerstatistik

2006

FDZ-Arbeitspapier
Nr.10

Timm Bönke
Frank Neher
Carsten Schröder



STATISTISCHE ÄMTER
DES BUNDES UND DER LÄNDER
FORSCHUNGSDATENZENTREN

Konstruktion und Bewertung
eines ökonomischen
Einkommens aus der
Faktisch Anonymisierten
Lohn- und Einkommen-
steuerstatistik

2006

Herausgeber: Statistische Ämter des Bundes und der Länder
Herstellung: Statistisches Bundesamt

Fachliche Informationen

zu dieser Veröffentlichung:

Statistisches Bundesamt
Forschungsdatenzentrum
Tel.: 06 11 / 75 42 20
Fax: 06 11 / 72 40 00
forschungsdatenzentrum@destatis.de

Erscheinungsfolge: unregelmäßig
Erschienen im August 2006

Informationen zum Datenangebot:

Statistisches Bundesamt
Forschungsdatenzentrum
Tel.: 06 11 / 75 42 20
Fax: 06 11 / 72 40 00
forschungsdatenzentrum@destatis.de

Forschungsdatenzentrum der
Statistischen Landesämter
– Geschäftsstelle –
Tel.: 0211 / 9449 41 47
Fax: 0211 / 9449 40 77
forschungsdatenzentrum@lds.nrw.de

Diese Publikation wird kostenlos als **PDF-Datei** zum Download unter www.forschungsdatenzentrum.de angeboten.

© Statistisches Bundesamt, Wiesbaden 2006
(im Auftrag der Herausbergemeinschaft)

Für nichtgewerbliche Zwecke sind Vervielfältigung und unentgeltliche Verbreitung, auch auszugsweise, mit Quellenangabe gestattet. Die Verbreitung, auch auszugsweise, über elektronische Systeme/Datenträger bedarf der vorherigen Zustimmung. Alle übrigen Rechte bleiben vorbehalten.

Konstruktion und Bewertung eines ökonomischen Einkommens aus der Faktisch Anonymisierten Lohn- und Einkommensteuerstatistik

Timm Bönke[#], Frank Neher, Carsten Schröder

Zusammenfassung.

Atypische Engelkurven zeigen, dass die in der Faktisch Anonymisierten Lohn- und Einkommensteuerstatistik 1998 (FAST) verzeichneten steuerrechtlich definierten Einkommensaggregate ein verzerrtes Bild vom Lebensstandard der Zensenten geben. Wir zeigen, dass für einen überwiegenden Teil der Beobachtungen in FAST ausreichende Informationen zur Konstruktion von Einkommen im ökonomischen Sinne verfügbar sind. Eine exemplarische Analyse der Progressionswirkung der Einkommensteuer ergibt, dass Verteilungsanalysen, die direkt steuerliche Einkommensdefinitionen nutzen, zu systematischen Fehleinschätzungen führen können.

1 Einleitung

In den letzten Jahren hat sich in Deutschland die Datengrundlage für Analysen der Einkommensverteilung erheblich erweitert. So stehen mit dem Sozioökonomischen Panel (SOEP) und der Einkommens- und Verbrauchsstichprobe (EVS) umfangreiche Datensätze mit einer Vielzahl sozioökonomischer Merkmale privater Haushalte zur Verfügung. Untersuchungen zur Einkommensverteilung liegt die Vorstellung zugrunde, dass die Größe „Einkommen“ die Konsum- und Partizipationsmöglichkeiten determiniere und daher ein geeigneter Indikator des Lebensstandards sei. Dies setzt voraus, dass das verwendete Einkommensaggregat die konsumtiven Möglichkeiten der Haushalte tatsächlich widerspiegelt. Ein solches Einkommen wird als „ökonomisches Einkommen“ bezeichnet. Es liegt in der EVS (approximativ) in Form der „ausgabefähigen Einkommen und Einnahmen“ und des „verfügbaren Einkommens“ vor. Auch über die Variablen des SOEP lässt sich eine entsprechende Größe konstruieren.

Charakteristisch für SOEP und EVS ist ein umfangreiches Variablenspektrum, das es erlaubt, verschiedenste Fragestellungen im Verteilungskontext zu untersuchen. So können beispielsweise für verschiedene Haushaltstypen detaillierte Aussagen über die Einkommensverteilung gemacht werden. Ebenso ist es möglich, die Verteilungswirkungen insbesondere sozialpolitischer Variablen, wie z. B. des Kindergeldes, zu quantifizieren. Einschränkungen ergeben sich dadurch, dass Haushalte mit hohem Einkommen in SOEP und EVS unterrepräsentiert sind. Darüber hinaus fehlen steuerrechtlich relevante Variablen, woraus vor allem bei der Analyse der Verteilungswirkungen steuerpolitischer Maßnahmen Beschränkungen resultieren.

[#] Kontaktautor. Telefon: +49 (0)30 838 53331. E-Mail: tim.boenke@wiwi.fu-berlin.de
Free University Berlin, Department of Economics, Institute for Public Finance and Social Policy, Boltzmannstr. 20, 14195 Berlin, Germany

Mit FAST98, der faktisch anonymisierten Lohn- und Einkommensteuerstatistik des Jahres 1998, steht der Wissenschaft seit 2004 erstmals ein Mikrodatsatz zur Verfügung, in dem sowohl die Top-Einkommensperzentile für Deutschland repräsentativ vertreten als auch detaillierte Angaben zum Besteuerungsprozess der Mikroeinheiten enthalten sind. Dieser Querschnitt wird im 3-Jahresrhythmus erhoben.

Auf den ersten Blick erscheint FAST als ideale Grundlage für empirische Untersuchungen zu den redistributiven Wirkungen des steuerlichen Instrumentariums. Allerdings basieren die in FAST verzeichneten Einkommensaggregate, wie z. B. die „Summe der Einkünfte“, auf steuerrechtlichen Abgrenzungen und Definitionen. Steuerrechtlich nicht relevante Einkommen sind in diesem Aggregat nicht enthalten aber in FAST zum Teil ausgewiesen. Hierzu zählen Einkommen, die dem Progressionsvorbehalt unterliegen, oder steuerfreie Zuschläge für Sonntags-, Feiertags- und Nacharbeit. Andere Einkünfte sind unvollständig, im Sinne eines ökonomischen Einkommensbegriffs verzerrt oder unzureichend nachgewiesen. So sind Leibrenten nur mit ihrem Ertragsanteil ausgewiesen. Gewinneinkünfte oder Einkünfte aus Vermietung und Verpachtung können verzerrt sein, wenn die steuerrechtlichen Abschreibungsregelungen nicht den tatsächlichen Werteverbrauch widerspiegeln. Außerdem können Verluste periodenfremd veranlagt werden. Ein Problem im Verteilungskontext ergibt sich insbesondere dann, wenn sich die oben beschriebenen Verzerrungen nicht gleichmäßig auf alle Steuerpflichtigen auswirken.

Es stellt sich daher die Frage, ob steuerrechtlich definierte Einkommen, wie sie in FAST vorliegen, direkt in Verteilungsanalysen Verwendung finden sollten. Und falls nein, ob FAST hinreichende Informationen zur Konstruktion ökonomischer Einkommen enthält. Zur Beantwortung dieser Fragen konstruieren wir zunächst schrittweise ein möglichst umfassendes Bruttoeinkommensaggregat. Wir bestimmen für jeden der Korrekturschritte das gesamte Korrekturvolumen (in Euro), die Anzahl der von Korrekturen betroffenen Zensenten, und die Verteilungseffekte. Dieses Bruttoeinkommensaggregat liegt dann der Berechnung des „ökonomischen Einkommens“ zugrunde.

Die Beurteilung der Eignung alternativer Einkommensdefinitionen für Verteilungsanalysen erfolgt auf Basis von Engelkurvenverläufen, die sich für die in FAST verzeichneten Ausgabenkategorien „Steuerberatung“, „Spenden“, „Schulgeld“ und „Wohnen“ für verschiedene Einkommensaggregate ermitteln lassen. Nach empirischen Untersuchungen zum Konsumverhalten von Haushalten handelt es sich bei diesen Gütern um normale bzw. superiore Güter. Zeigen Engelkurven auf Grundlage eines steuerlichen Einkommenskonzepts dagegen einen atypischen Verlauf, deutet dies auf einen verzerrten Einkommensbegriff hin. So legt beispielsweise eine U-förmige Engelkurve nahe, dass der zugrunde gelegte Einkommensbegriff den tatsächlichen Lebensstandard eines Teils der dem unteren Einkommensbereich zugewiesenen Haushalte unterschätzt. Unseres Wissens existiert bisher keine Studie, in der solch ein „Gütemaß“ zur Bewertung alternativer Einkommensdefinitionen verwendet wurde.

Wenn sich atypische Engelkurven beobachten lassen, schließt sich die Frage an, wie bedeutsam die mit der Wahl des Einkommensaggregats verbundenen Verzerrungen für Verteilungsanalysen sind. Zur Beantwortung dieser Frage bestimmen wir exemplarisch die effektiven perzentilspezifischen Steuerbelastungen bei Verwendung alternativer Einkommensdefinitionen.

Die Arbeit ist wie folgt gegliedert: Im nächsten Abschnitt wird zunächst die Datengrundlage FAST vorgestellt. Abschnitt 3 beschreibt die Konstruktion eines ökonomischen Einkommens und die Effekte der einzelnen Korrekturschritte (Korrekturvolumen sowie Anzahl der betroffenen Zensenten). In Abschnitt 4 wird eine Bewertung der Einkommenskonzepte vorgenommen sowie die mit den Korrekturschritten einhergehenden Wanderungs- und Verteilungseffekte bestimmt. In Ab-

schnitt 5 werden die Progressionseigenschaften der deutschen Einkommensteuer auf Basis alternativer Einkommenskonzepte untersucht. Abschnitt 6 fasst die Ergebnisse zusammen.

2 Datengrundlage

Die faktisch anonymisierte Lohn- und Einkommensteuerstatistik ist eine geschichtete 10%ige Zufallsstichprobe der Lohn- und Einkommensteuerstatistik, die im 3-Jahres-Rhythmus als Sekundärstatistik aus dem Verwaltungsvollzug der Finanzämter erstellt wird. Sie enthält Informationen über die Einkommen und steuerlich relevanten sozio-ökonomischen Merkmale von ca. 3 Mio. Einkommensbezieher. FAST steht der Wissenschaft als Scientific-Use-File seit 2004 zur Verfügung.

2.1 Schichtung

Als Schichtungsmerkmale finden neben dem Gesamtbetrag der Einkünfte die regionale Herkunft, die Einkunftsarten, die Veranlagungsart sowie Anzahl und Alter der Kinder im Haushalt Berücksichtigung. Die Aufteilung des Stichprobenumfangs auf die einzelnen Schichten erfolgt nach dem „Prinzip der vergleichbaren Präzision für gegliederte Ergebnisse“. Hierdurch sollen Randgruppen, wie z. B. Bezieher hoher Einkommen, zuverlässig erfasst und der Stichprobenfehler minimiert werden.¹ Im Gegensatz beispielsweise zur nach oben zensierten EVS ist FAST unzensiert. Vielmehr enthält FAST *sämtliche* Einkommensbezieher ab einem Gesamtbetrag der Einkünfte von 102.258 Euro (200.000 DM).

Tabelle 1: Gesamtbetrag der Einkünfte, Hochrechnungsfaktoren

Bereich	Fälle	Gewichtet	Ø-Gewichtungsfaktor	GDE (in Mio. €)	Min GDE (in €)	Max GDE (in €)
Dezil: 1	432.119	2.787.651	6,45	-6.872	-349.765.949	4.507
2	482.710	2.787.040	5,77	19.860	4.508	9.825
3	477.623	2.787.277	5,84	35.471	9.826	15.541
4	110.910	2.787.332	25,13	50.528	15.542	20.557
5	92.475	2.787.357	30,14	63.405	20.558	24.906
6	86.329	2.787.323	32,29	75.569	24.907	29.449
7	85.530	2.787.054	32,59	89.942	29.450	35.433
8	82.704	2.787.292	33,70	110.056	35.434	44.024
9	137.019	2.787.418	20,34	140.506	44.025	58.624
10	805.441	2.787.099	3,46	312.542	58.625	162.793.412
Top 1,0%	278.729	278.729	1,00	113.422	146.203	162.793.412
0,5%	13.936	13.936	1,00	43.595	986.651	162.793.412
0,1%	2.787	2.787	1,00	24.699	3.599.330	162.793.412
Fehlend	74.477	800.069	10,74	—	—	—
Gesamt	2.867.337	28.672.912	10,00	891.006		
Anmerkung: · Einteilung nach dem Gesamtbetrag der Einkünfte (GDE); · Merkmal Gesamtbetrag der Einkünfte nicht besetzt; · Zusammen Veranlagte (Ehepaare) zählen als ein Fall. Quelle: Statistisches Bundesamt, FAST, eigene Berechnungen.						

¹ Zum Schichtungsverfahren vgl. Zwick (1998).

Tabelle 1 gibt einen Überblick über die Fallzahlen und Gewichtungsfaktoren für die nach dem „Gesamttrag der Einkünfte“ geordnete Stichprobe. Neben der gewichteten und ungewichteten Anzahl der Veranlagten sind der durchschnittliche Gewichtungsfaktor und die Summe des Gesamtbetrags der Einkünfte (GDE) für jeden Einkommensbereich ausgewiesen. Zusätzlich dazu findet sich in Spalte 6 (Min GDE) die Untergrenze und in Spalte 7 die Obergrenze (Max GDE) des jeweiligen Einkommensintervalls. Der durchschnittliche Hochrechnungsfaktor für die 2.867.337 Beobachtungen beträgt 10. Zählt man Ehepaare, die eine gemeinsame Veranlagung gewählt haben (Splittingtabelle) und damit nur einfach erfasst sind, doppelt, repräsentiert die Stichprobe ca. 40 Millionen veranlagte Personen. Wie die deutlich geringeren Gewichtungsfaktoren im ersten, zweiten, dritten und zehnten Dezil zeigen, sind die Randbereiche der Verteilung in der Stichprobe überrepräsentiert.

2.2 Anonymisierung

Um zu gewährleisten, dass die im Datensatz enthaltenen Einzelangaben nur mit hohem Ressourcenaufwand eine Identifikation einzelner Zensenten erlauben, wurde aus datenschutzrechtlichen Gründen eine Anonymisierung durchgeführt. Das Ausmaß der Anonymisierung ist aufgrund des überdurchschnittlichen Identifikationsrisikos für die Randbereiche der Einkommensverteilung am höchsten. Die Einkommensintervalle der insgesamt fünf Anonymisierungsbereiche können Tabelle 2 entnommen werden.

Tabelle 2: Einteilung der Anonymisierungsbereiche

Anonymisierungsbereich	Positives GDE /Einkommen in € (in DM)	Negatives GDE/Einkommen in € (in DM)
1	0 – 64.106 (0 – 125.381) (zweimal das Durchschnittliche GDE)	0 – -102.258 (0 – -200.000)
2	64.107 – 137.532 (125382 – 268.990) (99 % Perzentil)	—
3	137.533 – 970.202 (268.991 – 1.897.552) (99,95 % Perzentil)	-102.259 – -511.292 (-200.001 – -1.000.000)
4	970.203 – 7.354.714 (1.897.553 – 14.384.571) (bis zu den 1000 Reichsten)	—
5	> 7.354.714 (> 14.384.572)	< -511.292 (< -1.000.000)
Quelle: Vorgrimler und Zwick (2004), S. 2.		

Tabelle 3: Bedeutung der Anonymisierung

Anonym.-bereich	Fälle, ungewichtet	Anteil, ungewichtet	Fälle, gewichtet	Anteil, gewichtet	SDE (Mio. €)	SDE Anteil
1	2.094.791	73,06%	26.463.003	92,29%	625.355,04	69,96%
2	441.983	15,41%	1.868.912	6,52%	157.140,12	17,58%
3	311.150	10,85%	314.944	1,10%	71.634,25	8,01%
4	13.188	0,46%	13.188	0,05%	28.139,66	3,15%
5	6.219	0,22%	12.859	0,04%	10.929,90	1,22%
6	6	0,00%	6	0,00%	664,65	0,07%
Summe	2.867.337	100,00%	28.672.912	100,00%	893.863,61	100%
<p><i>Anmerkung:</i> · Im Datensatz nachgewiesenes Merkmal, die Bereiche 1, 3 und 5 umfassen sowohl negative und positive Einkünfte, die Bereiche 2, 4 und 6 nur positive (siehe Tabelle 2); · SDE entspricht hier der „Summe der Einkünfte“ korrigiert um Fälle, bei denen diese Variable unbesetzt ist, aber mindestens eine der Einkunftsarten ungleich Null ist.</p> <p><i>Quelle:</i> Statistische Bundesamt FAST, eigene Berechnungen.</p>						

Tabelle 3 zeigt, dass insgesamt 73,06% der ungewichteten Fälle dem Bereich niedrigster Anonymisierung (Anonymisierungsbereich 1) zuzuordnen sind. Dies entspricht einem gewichteten Anteil von 92,29% aller Steuerzahler. Der Anteil dieser Gruppe am Gesamtbetrag der „Summe der Einkünfte“ (SDE) nach §2 EStG beträgt 69,96%. Der den hohen Anonymisierungsgraden 4-6 zuzuordnende gewichtete Stichprobenanteil liegt bei 0,09%, sein Einkommensanteil bei 4,44%. Für diese fiskalisch bedeutsame Gruppe, über die ca. 10% des Einkommensteueraufkommens generiert wird, stehen aufgrund der vorgenommenen Anonymisierung kaum mehr Identifikationsmerkmale zur Verfügung. So fehlen hier Informationen z. B. über Einkommensart und Ort der Einkommenserzielung, steuerrelevante Ausgabenpositionen, die Inanspruchnahme von Sonderbesteuerungsregelungen oder über die Haushaltszusammensetzung.

2.3 Informationsgehalt

Eine Stärke von FAST im Vergleich zu anderen verfügbaren Statistiken besteht in der präzisen und detaillierten Abbildung der steuerlich relevanten Einkommen und Merkmale. Auf einen ersten Blick erscheint diese Statistik daher als ein idealer Kandidat zur Quantifizierung der Aufkommens- und Verteilungswirkungen der Einkommensbesteuerung.

Generell ist allerdings zu beachten, dass der untere Rand der Einkommensverteilung in FAST relativ zur Gesamtbevölkerung stark unterrepräsentiert ist, da viele Geringverdiener oder Bezieher von Sozialhilfe und Arbeitslosengeld nicht einkommensteuerpflichtig sind. Für Analysen, die sich insbesondere mit dem unteren Rand der Einkommensverteilung beschäftigen, sind jedoch spezielle Statistiken, wie z. B. das Niedrigeinkommenspanel vom Bundesministerium für Arbeit und Soziales oder die Sozialhilfestatistik des Statistischen Bundesamtes, verfügbar.

Einschränkungen in der Interpretierbarkeit der Ergebnisse können auch aus der steuerrechtlich motivierten Konstruktion der Einkommensaggregate resultieren. Vergleicht man etwa die in der Einkommensteuerstatistik nachgewiesenen Einkünfte mit den entsprechenden Aggregaten der Volkswirtschaftlichen Gesamtrechnung (VGR), so zeigt sich, dass erstgenannte signifikant niedri-

ger sind.² Beispielsweise summieren sich die in der Einkommensteuerstatistik ausgewiesenen Bruttolöhne auf nur etwa 90% ihres VGR-Pendants. Der Unterschied kann in diesem Fall auf die steuerlich nicht relevanten Einkünfte von geringfügig Beschäftigten sowie von Zivil- und Wehrdienstleistenden zurückgeführt werden. Bei Unternehmens- und Vermögenseinkommen ergeben sich Unterschiede zwischen FAST- und VGR-Aggregat aus unterschiedlichen Begriffsabgrenzungen und Abschreibungsvorschriften. Problematisch erscheint auch die Untererfassung laufender Übertragungen des Staates an die Haushalte, da diese in der Steuerstatistik gar nicht oder nur teilweise (z. B. Leibrenten) erfasst sind.

Weiter ist fraglich, ob es zulässig ist, Zensenten im Sinne von Haushalten zu interpretieren. Die Identifizierbarkeit zusammenlebender Haushaltsgemeinschaften ist aber für verschiedenste Fragestellungen notwendig. So greifen Verteilungsanalysen typischerweise auf „äquivalente Haushaltseinkommen“ zurück, die die Bedarfsunterschiede verschiedener Haushaltstypen berücksichtigen. Dies setzt zumindest die Identifizierbarkeit der im Haushalt lebenden Personenzahl und ihrer individuellen Merkmale voraus. Dies gelingt über FAST nur bedingt: Zwar kann die Anzahl der steuerlich relevanten Kinder sowie Ehepaare (über die Veranlagungsart nach Grund- oder Splittingtabelle) ermittelt werden, allerdings zeigt ein Vergleich der über EVS98 bzw. FAST98 berechneten Häufigkeiten verschiedener Haushaltstypen deutliche Unterschiede (vgl. Tabelle 4). Diese Differenzen haben mehrere Ursachen. Erstens können über FAST nur zusammenlebende Paare identifiziert werden, die auch gemeinsam veranlagt werden. Nichteheliche Lebensgemeinschaften erscheinen, da sie einzeln veranlagt werden, als zwei separate Haushalte. Ferner ist es mit FAST nicht möglich festzustellen, ob Kinder, die aufgrund eigener Einkünfte zur Einkommenssteuer veranlagt sind, bei ihren Eltern leben oder nicht. Weitere Fehlklassifikationen ergeben sich dadurch, dass getrennt lebende Ehepaare, die aus steuerlichen Gründen eine gemeinsame Steuererklärung einreichen, nicht als getrennt lebende Haushalte identifiziert werden können. Demgegenüber sind die EVS-Daten nach oben zensiert. Nach Merz (2001, S. 103) sind hiervon 37.345 Haushalte, überwiegend Mehrpersonenhaushalte, betroffen. Gravierender als die oben angeführten Einschränkungen, die bei Arbeiten auf Basis von FAST zu beachten sind und prinzipiell durch Kalibrierung und unter Verwendung zusätzlicher Statistiken weitgehend behoben werden könnten, sind die Informationsdefizite der verzeichneten Einkommen selbst.

² Hierzu auch der Vergleich von Daten der Lohn- und Einkommensteuerstatistik mit Aggregaten der Volkswirtschaftlichen Gesamtrechnung in Bork und Petersen (1997, S. 20). Nach dieser Studie sind in der Lohn- und Einkommensteuerstatistik 1992 nur 57,2% der gesamten Bruttoeinkommenssumme verzeichnet. Der niedrigste Erfassungsgrad liegt bei Transfereinkommen (3%) vor. Bei Vermögens- und Unternehmereinkommen liegt der Erfassungsgrad bei 45,2% bzw. 41,6%. Nach Bach et al. (2004) beträgt der Gesamterfassungsgrad in 1995 für Bruttolöhne und -gehälter in der Lohn- und Einkommensteuerstatistik 92,1%.

Tabelle 4: Haushaltstypen nach FAST und EVS 98

Haushaltstyp		Fälle	Gewichtet	Anteil
FAST				
Verheiratet (Splittingtabelle)	ohne Kinder	721.225	6.650.705	23,3%
	1 Kind	369.978	3.430.430	12,0%
	2 Kinder	334.043	3.483.135	12,1%
	3 Kinder	103.148	958.792	3,3%
	4 und mehr Kinder	32.781	281.909	1,0%
Unverheiratet (Grundtabelle)	ohne Kinder	875.432	11.592.988	40,3%
	1 Kind	244.429	1.487.089	5,2%
	2 Kinder	146.888	654.878	2,3%
	3 Kinder	31.950	133.556	0,5%
	4 und mehr Kinder	7.433	29.430	0,1%
Gesamt		2.867.337	28.672.912	100%
EVS98				
zusammenlebende Ehepaare	ohne Kinder	13.540	9.528.696	25,9%
	1 Kind	6.331	3.894.056	10,6%
	2 Kinder	8.837	4.020.673	10,9%
	3 Kinder	2.499	976.749	2,7%
	4 und mehr Kinder	714	274.002	0,7%
allein lebende Personen	ohne Kinder	10.891	12.989.851	35,3%
	1 Kind	1.618	1.176.056	3,2%
	2 und mehr Kinder	1.120	551.662	1,5%
zusammenlebende Paare	ohne Kinder	1.585	1.188.874	3,2%
	1 Kind	400	268.650	0,7%
	2 Kinder	218	119.050	0,3%
	3 und mehr Kinder	55	25.991	0,1%
Sonstige Haushalte		1.912	1.765.962	4,8%
Gesamt		49.720	36.780.272	100%
Anmerkung: · Veranlagte nach gemeinsamer oder auf Antrag getrennter Veranlagung; · Einzelveranlagung. Quelle: Statistisches Bundesamt FAST, eigene Berechnung; EVS, Drabinski und Schröder (2003), S. 10.				

3 Konstruktion: Von der „Summe der Einkünfte“ zum „ökonomischem Einkommen“

Prinzipiell liegt in FAST ein umfassendes Einkommensaggregat, die „Summe der Einkünfte“ nach §2 EStG, getrennt nach Geschlechtern vor. Tatsächlich entspricht die „Summe der Einkünfte“ aber nicht der Summe der Bruttoeinkünfte des Steuerpflichtigen, da letzte z. B. um Freibeträge und Werbungskosten vermindert werden können, verschiedene Einkünfte nicht oder nur zum Teil der Einkommensteuer unterliegen, und Einkünfte über steuerlich motivierte Abschreibungen und Verluste aus Beteiligungen vermindert werden können. Neben diesen steuerrechtlich induzierten Differenzen zwischen dem tatsächlichen Bruttoeinkommen und der „Summe der Einkünfte“ besteht ein grundsätzliches Erfassungsproblem. Liegen nur Einkünfte aus einer einzigen Einkunftsart vor, so haben die Finanzämter vielfach auf eine Berechnung der „Summe der Einkünfte“ verzichtet. Dies gilt insbesondere für Einkünfte aus nicht-selbständiger Tätigkeit.

Die Konstruktion eines ökonomischen Einkommens aus der „Summe der Einkünfte“ ist nachfolgend in drei Korrekturschritte zerlegt: a) Bereinigung vorhandener und Imputation fehlender Daten; b) Eliminierung der Effekte steuerrechtlicher Gestaltungsmöglichkeiten zwecks vollständiger Erfassung aller Einkünfte; c) Abzug tatsächlich gezahlter Steuern und Addition tatsächlich emp-

fangener Transferleistungen. Unsere Vorgehensweise ist weitgehend vergleichbar mit Merz (2001, Kapitel 7) und Merz et al. (2005).

Tabelle 5 stellt die Abfolge der Korrekturschritte dar, während in Tabelle 6 zum einen die absoluten und relativen Häufigkeiten der von Korrekturen betroffenen Fälle sowie das gesamte Korrekturvolumen in Mio. Euro per anno enthalten sind.³

Bereinigung und Imputation

Die „Summe der Einkünfte“ beinhaltet die um Freibeträge verminderten Gewinn- und die um Werbungskosten(pauschalen) verminderten Überschusseinkünfte. Liegt bei der Variablen „Summe der Einkünfte“ im Datensatz kein Eintrag vor (6,01% der Fälle), wird in einem ersten Korrekturschritt überprüft, ob dennoch Einkünfte aus mindestens einer der sieben Einkunftsarten erzielt werden. Dies trifft für 3,24% der Fälle, bzw. für 53,8% der Beobachtungen ohne Einträge zu,⁴ für die dann die „Summe der Einkünfte“ berechnet wird. Das Korrekturvolumen beträgt 4,279 Mrd. Euro. Für alle verbleibenden Fälle mit Missing Value wird SDE gleich Null gesetzt. Für den Anonymisierungsbereich 5 besteht zudem das Problem, dass die Einkunftsarten nicht mehr getrennt verzeichnet sind und nur in aggregierter Form als „Summe der Einkünfte“ vorliegen. Auch Informationen zu den in Anspruch genommenen steuerlichen Wahlmöglichkeiten fehlen. Unproblematisch ist dieses Informationsdefizit bei Zensenten, deren „Summe der Einkünfte“ positiv ist. Diese Fälle ändern ihre Position (zugewiesenes Perzentil) in der Einkommensverteilung bei der Durchführung der Korrekturschritte a)–c) allenfalls geringfügig, da sie für alle betrachteten Einkommensaggregate stets dem obersten Einkommensperzentil zugewiesen werden.⁵ Probleme ergeben sich aber bei Fällen mit negativen Einkommen. Unter der Arbeitshypothese, dass negative Einkommen vorwiegend steuerlich motiviert sind, besteht hier ein besonders hoher Korrekturbedarf, dem aufgrund des oben beschriebenen Informationsdefizits allerdings nicht entsprochen werden kann. Prinzipiell wäre es möglich, Schätzungen für die fehlenden Werte basierend auf vorliegenden Beobachtungen des Anonymisierungsbereichs 3 vorzunehmen. Da lediglich 2873 von ca. 28,6 Mio. Fällen von dieser Problematik betroffen sind, haben wir uns dafür entschieden, diese Fälle von der Analyse auszuschließen.

In einem weiteren Schritt werden die Angaben zur Höhe der Kapitaleinkünfte korrigiert.⁶ Schließt man von den in FAST angegebenen Kapitaleinkünften mittels einer unterstellten Durchschnittsverzinsung auf das zugrunde liegende Kapitalvermögen, zeigt sich, dass nur ein Teil des in der Bundesbankstatistik verzeichneten Geldvermögens in FAST ausgewiesen ist.⁷ Diese Differenz lässt sich unter anderem dadurch erklären, dass Steuerpflichtige, deren Kapitaleinkünfte die

³ Zur Begründung und zu alternativen Konzeptionen ökonomischer Einkommen vgl. auch Merz (2001, Kapitel 7) sowie Bach et al. (2006, S. 3–6).

⁴ Diese Missing Values resultieren offensichtlich aus Effizienzüberlegungen bei der Bearbeitung von Einkommensteuererklärungen bei den Finanzämtern: Erzielt ein Steuerpflichtiger nur Einkünfte aus einer einzigen Einkunftsart, ist die Berechnung der „Summe der Einkünfte“ zur Festlegung der Steuerschuld nicht notwendig.

⁵ Eine Ausnahme stellt der Sonderanonymisierungsbereich „Abgeordnete“ dar. Vergleiche hierzu Vorgrimler und Zwick (2004).

⁶ Merz et al. (2005, S.19) übernehmen die Einkünfte aus Kapitalvermögen unkorrigiert.

⁷ Die für den Sparerfreibetrag relevanten Vermögensaggregate nach EVS (Ammermüller et al. (2005)) wurden mit den durchschnittlichen Verzinsungen dieser Aggregate entsprechend der Zeitreihenstatistik der Bundesbank (2006) bewertet. Bankeinlagen, die 1998 etwa 77% des betrachteten Vermögens ausmachten, wurden demnach durchschnittlich mit 2,8% p.a. verzinst. Rentenwerte und Geldmarktpapiere mit einem Anteil von ca. 9% erfuhren durchschnittlich eine Verzinsung von 4,50% p.a., die verbleibenden 14% Aktien eine Verzinsung von 17,71%. Daraus resultiert eine durchschnittliche Vermögensverzinsung von 4,82%. Mit dieser Verzinsung kann über die in FAST verzeichneten Vermögenseinkommen der zugrunde liegende Vermögensbestand errechnet werden. Dieser Vermögensbestand ist nicht einmal halb so hoch wie das laut Bundesbankstatistik ausgewiesene Geldvermögen (Bundesministerium der Finanzen (2005)). Daraus kann geschlossen werden, dass die Vermögenseinkommen in FAST98 um mehr als 50% untererfasst sind.

Tabelle 5: Abfolge der Korrekturschritte

Gewinneinkünfte		Einkünfte und Veräußerungsgewinne aus Land- und Forstwirtschaft % Freibeträge für Land- und Forstwirtschaft und für Veräußerungsgewinne hieraus
	+	Einkünfte aus Gewerbebetrieb als Einzelunternehmer, aus Beteiligung und aus Veräußerungsgewinnen % Freibeträge für Veräußerungsgewinne
	+	Einkünfte und Veräußerungsgewinne aus selbständiger Tätigkeit % Freibeträge für selbständige Tätigkeit
Überschusseinkünfte	+	Bruttoarbeitslohn aus nichtselbständiger Arbeit % Werbungskosten % Versorgungsfreibetrag
	+	Einnahmen aus Kapitalvermögen % Werbungskosten % Sparerfreibetrag
	+	Einkünfte aus Vermietung und Verpachtung und aus Beteiligungen hieran
	+	Sonstige Einkünfte: Ertragsanteil Renten, Einnahmen aus Unterhaltsleistungen, Spekulationsgewinne, Einkünfte aus Abgeordnetentätigkeit % Werbungskosten(pauschbetrag)
	=	SUMME DER EINKÜNFTE (SDE)
	+	Bei Fällen, bei denen die „Summe der Einkünfte“ nicht besetzt ist, wird die „Summe der Einkünfte“ aus den Angaben zu den sieben Einkunftsarten nachberechnet
	+	Next Neighborhood Imputation für nicht angegebene Kapitaleinkünfte
	=	SUMME DER EINKÜNFTE 1 (SDE1)
	+	Alle Freibeträge für Land- und Forstwirtschaft sowie für selbständige Tätigkeit und für Gewerbebetrieb
	+	Werbungskosten für nicht-selbständige Arbeit sowie für Kapitalvermögen und für sonstige Einkünfte
	+	Versorgungs- und Sparerfreibetrag
	=	SUMME DER EINKÜNFTE 2 (SDE2)
	+	Steuerfreie Einkünfte: Lohn- und Einkommensersatzleistungen, Einkünfte aus DBA-Staaten
	+	Volle Rente abzüglich Ertragsanteil Rente
	=	SUMME DER EINKÜNFTE 3 (SDE3)
	+	Sonderabschreibungen laut Anlage St: Steuervergünstigungen nach Fördergebietsgesetz §§ 2,3,4,5 (bei §§ 2,3,4 zu 50%), Steuervergünstigungen nach § 7 EStG
	+	„Steuersparmodelle“: Verluste aus Beteiligungen bei Gewerbeeinkommen, selbständiger Tätigkeit und bei Vermietung und Verpachtung
	=	BRUTTOEINKOMMEN 1 (Brutto1)
	+	Verluste aus Gewinneinkünften und aus Vermietung und Verpachtung
	=	BRUTTOEINKOMMEN 2 (Brutto2)
Netto-transfer	%	Festgesetzte Einkommensteuer, festgesetzte Kirchensteuer, festgesetzter Solidaritätszuschlag
	%	Gezahlte Unterhaltsaufwendungen
	+	Empfangenes Kindergeld
	=	ÖKONOMISCHES EINKOMMEN (OeK)
Quelle: Eigene Darstellung.		

Tabelle 6: Absolute und relative Häufigkeit der korrigierten Fälle

Ausgangs- aggregat	Ergebnisaggregat	N (in Tsd.)	N in %	Korrekturvolumen in Mrd. €
SDE	SDE1	928	3,24	4,279
		24.279	84,68	20,676
		3	0,00	—
SDE1	SDE2	27.297	95,21	72,776
SDE2	SDE3	8.801	30,70	51,832
SDE3	Brutto1	1.640	5,72	19,380
	Brutto2	170	0,59	6,260
Brutto2	OekE	21.465	74,87	-182,893
Anmerkungen. · Anzahl der von der Korrekturmaßnahme betroffenen Fälle. N ist die Anzahl der Beobachtungen. Quelle: Eigene Berechnungen.				

geltenden Sparerfreibeträge (Ledige: 6.000 DM, Paare: 12.000 DM in 1998) nicht überstiegen, in der Regel keine Angabe zur Höhe ihrer Kapitaleinkünfte gemacht haben.⁸ Tatsächlich sind in FAST für nur ca. 10% der Zensenten positive Kapitaleinkünfte unterhalb des Sparerfreibetrags verzeichnet. Bei 84,68% der Zensenten liegen keine Informationen zur Höhe der Kapitaleinkünfte vor. Zur Imputation fehlender Beobachtungen gehen wir wie folgt vor: Für alle Fälle, die positive Kapitaleinkünfte unterhalb des Sparerfreibetrags angeben, schätzen wir ökonometrisch die Höhe dieser Einkünfte in Abhängigkeit verschiedener sozio-ökonomischer Variablen. Die Ergebnisse einer linearen Regression sind in Tabelle A1 im Anhang zusammengefasst. Auf Grundlage der ermittelten Koeffizienten wird unter Beachtung der individuellen sozio-ökonomischen Charakteristika jedes einzelnen Zensenten, der keine Angabe zur Höhe der Kapitaleinkünfte gemacht hat, ein Wert für die Höhe der Kapitaleinkünfte imputiert. Insgesamt werden 24,279 Mio. Fällen Kapitaleinkünfte in Höhe von insgesamt 20,676 Mrd. Euro zugewiesen. Das neue Einkommensaggregat wird nachfolgend als „Summe der Einkünfte 1“ (SDE1) bezeichnet.

Eliminierung der Effekte steuerlicher Gestaltungsmöglichkeiten

Im zweiten Korrekturschritt wird SDE1 um steuerrechtliche Effekte bereinigt, um so ein umfassenderes Bruttoeinkommensaggregat zu generieren. Zunächst werden alle Freibeträge und Werbungskosten den entsprechenden Einkunftsarten hinzugerechnet. Dieser Schritt lässt sich damit begründen, dass Werbungskostenpauschalen in der Regel keine entsprechenden Ausgaben gegenüberstehen und häufig eine konsumtive Komponente enthalten.⁹ Die Anrechnung von Freibeträgen ist unstrittig, da diesen keine tatsächlichen Ausgaben gegenüberstehen, die konsumtiven Möglichkeiten der Haushalte also nicht eingeschränkt werden. Korrigiert werden in diesem Schritt die Einkommen von 95,21% der verzeichneten Fälle bei einem Korrekturvolumen von insgesamt 72,776 Mrd. Euro. Das neue Einkommensaggregat „Summe der Einkünfte 2“ (SDE2) beläuft sich auf insgesamt 991,99 Mrd. Euro.

Darüber hinaus enthält das steuerliche Aggregat „Summe der Einkünfte“ nur solche Einkommen, die der Einkommensteuer unterliegen. Steuerfreie Einkünfte, die einem Progressionsvorbehalt unterliegen, sind jedoch in FAST nachgewiesen und können daher berücksichtigt werden. Hierzu zählen Lohn- und Einkommensersatzleistungen sowie alle Einkünfte aus Staaten, mit denen ein

⁸ Vgl. hierzu für 1995 auch Merz und Zwick (2004, S. 170).

⁹ Anstatt des Vollansatzes von Werbungskosten könnte auch ein teilweiser Ansatz erfolgen. Jedoch fehlen in FAST Informationen für eine geeignete empirische Abschätzung. Für eine solche müssten zusätzliche Statistiken, wie z B. die EVS, herangezogen werden.

Doppelbesteuerungsabkommen besteht. Weiterhin können Leibrenten identifiziert werden, da sie einer Ertragsanteilsbesteuerung unterliegen und somit anteilig in der „Summe der Einkünfte“ enthalten sind. Der bisher unberücksichtigte Ertragsanteil wird SDE2 ebenfalls hinzugerechnet. Insgesamt werden Korrekturen für 30,70% der Fälle mit einem Gesamtvolumen von 51,832 Mrd. Euro vorgenommen. Das neue Einkommensaggregat, die „Summe der Einkünfte 3“ (SDE3), beträgt insgesamt 1.043,83 Mrd. Euro.

Im letzten Schritt wird SDE3 um Sonderabschreibungen und „Steuersparmodelle“ korrigiert. Als Sonderabschreibungen gelten alle über die regulären steuerlichen Abschreibungen hinausgehenden Abschreibungsmodelle.¹⁰ Mit solchen Sonderabschreibungen möchte der Gesetzgeber besondere Investitionsanreize dadurch setzen, dass die angerechnete Abschreibung den ökonomischen Werteverzehr übersteigt. Als „Steuersparmodelle“ gelten Verluste aus Beteiligungen, da hier davon auszugehen ist, dass der überwiegende Teil derartiger Verluste nur steuerlicher Natur ist (vgl. Merz und Zwick (2001, S. 517) und Merz et al. (2005)). Das Korrekturvolumen von 19,380 Mrd. Euro verteilt sich auf 5,72% der Fälle. Das neue Einkommensaggregat „Bruttoeinkommen 1“ (Brutto1), das mit der „Summe der Einkünfte neu“ aus Merz et al. (2005, S. 25) vergleichbar ist, beläuft sich auf 1.063,21 Mrd. Euro.

Sind für einzelne Zensenten auch nach diesen Korrekturen sowohl Brutto1 als auch die darin enthaltenen bereinigten Gewinneinkünfte bzw. Einkünfte aus Vermietung und Verpachtung negativ, so bleiben Verluste bei diesen Einkunftsarten bei der Berechnung des „Bruttoeinkommens 2“ (Brutto2) unberücksichtigt.¹¹ Der Hintergrund dieser Korrektur ist, dass ein Einkommen im ökonomischen Sinne die Konsummöglichkeiten eines Haushalts abbilden soll. Null oder gar Negativkonsum ist in Deutschland aufgrund der Existenz der sozialen Sicherungssysteme ausgeschlossen. Daher ist ein negatives „Bruttoeinkommen 1“ nicht plausibel und eher Ausdruck steuerlicher Anreize. Von Korrekturen betroffen sind 0,59% der Fälle, das Korrekturvolumen beträgt 6,26 Mrd. Euro. Im neuen Einkommensaggregat „Bruttoeinkommen 2“ sind nunmehr Einkommen mit einem Gesamtvolumen von 1.069,47 Mrd. Euro verzeichnet. Zu einem vergleichbaren Wert von 1063,51 Mrd. Euro kommen trotz leicht abweichender Methodik insbesondere hinsichtlich des letzten Korrekturschritts auch Bach et al. (2006, S. 7).¹²

Korrektur um Transfersaldo

Nun wird das „Bruttoeinkommen 2“ um die festgesetzte Einkommen- und Kirchensteuer sowie um den festgesetzten Solidaritätszuschlag gemindert, da diese Beträge dem Haushalt nicht für konsumtive Zwecke zur Verfügung stehen. Gleiches gilt für gezahlte Unterhaltsaufwendungen. Dagegen erhöht empfangenes Kindergeld¹³ die Konsummöglichkeiten und ist somit dem „Bruttoeinkommen 2“ hinzu zu rechnen. Ergebnis ist das „ökonomische Einkommen“ (OekE), das mit 886,58 Mrd. Euro 182,90 Mrd. Euro niedriger als Brutto2 ist. Bei 74,87% der Zensenten wird Brutto2 effektiv gemindert. Nur bei knapp 1.000 Fällen übersteigt die Summe aus empfangenem Kindergeld und Unterhaltsleistungen die steuerliche Belastung. In den verbleibenden Fällen än-

¹⁰ Angaben über Sonderabschreibungen sind in FAST über die Anlage St zwar prinzipiell verfügbar, allerdings ist der Erfassungsgrad niedrig.

¹¹ Dies stellt einen wesentlichen Unterschied im Vergleich sowohl zur Vorgehensweise in Merz et al. (2005) als auch in Bach et al. (2006) dar. Während in Merz et al. (2005) Negativeinkünfte nicht korrigiert werden, bleiben in Bach et al. (2006) Einkünfte aus Vermietung und Verpachtung bis zu bestimmten ad hoc definierten Grenzen enthalten.

¹² Verluste aus Vermietung und Verpachtung berücksichtigen Bach et al. (2006) bis 5.000 Euro (bzw. 2.500 Euro bei Einkünften aus Beteiligungen). Zudem verzichten die Autoren auf eine Imputation von Kapitaleinkünften unterhalb des Freibetrags.

¹³ Unseren Berechnungen liegen die in FAST angegebenen Werte zum Kindergeld zugrunde, so dass Günstigerprüfungen zwischen Kindergeld und –freibetrag direkt berücksichtigt werden können, während Merz et al. (2005) im Hinblick auf ihre weitere Analyse eine Nachberechnung des Kindergeldes vornehmen.

dem Transfers und Steuern das Brutto2-Aggregat im Saldo nicht. Hierbei handelt es sich um Zensenten, bei denen das „zu versteuernde Einkommen“ unterhalb des einkommensteuerlichen Freibetrags liegt. Überwiegend betrifft dies bei ihren Eltern lebende Kinder, Geringverdiener und Rentner mit geringfügigen Nicht-Renteneinkommen.

Ein wesentlicher Unterschied in der Definition des „ökonomischen Einkommens“ in dieser Studie und der „Nettoeinkommensdefinition“ in Merz et al. (2005, S. 25) liegt in der Behandlung von Sozialversicherungsbeiträgen. Merz et al. (2005, S. 18) zielen bei ihrer Konstruktion des „Nettoeinkommens“ auf ein Einkommen „im Sinne eines verfügbaren Einkommens“ (ebenda) ab. Da Sozialversicherungsbeiträge zu einer direkten Minderung der Konsummöglichkeiten der Haushalte führen, sind sie in Merz et al. (2005, S. 22f.) nicht Bestandteil des „Nettoeinkommens“. Im Unterschied hierzu ist in unserer Definition des „ökonomischen Einkommens“ der Arbeitnehmeranteil zur Sozialversicherung enthalten. Dies lässt sich zum einen damit begründen, dass FAST keine Angaben zur Höhe der geleisteten Beiträge enthält.¹⁴ Zum anderen soll unser „ökonomisches Einkommen“ nicht das verfügbare Einkommen der Haushalte, sondern deren Lebensstandard abbilden. Da Sozialversicherungsbeiträgen Ansprüche auf Versicherungsleistungen gegenüberstehen, interpretieren wir sie als Lebensstandard erhöhenden Zukunfts- oder Zwangskonsum.

Während nach unserer OekE-Definition eine Korrektur um die Summe der festgesetzten Steuern auf Einkommen erfolgt, die neben der Einkommensteuer auch Solidaritätszuschlag und Kirchensteuer umfasst, greifen Bach et al. (2006) auf die festgesetzte Einkommensteuer zurück.

4 Bewertung und empirischer Befund

4.1 Ein Indikator zur Beurteilung alternativer Einkommenskonzepte

Die Konstruktion eines ökonomischen Einkommens ist – wie oben gezeigt – aufwendig und nicht werturteilsfrei. Es stellt sich daher die Frage, ob die vorgenommenen Korrekturen und die den Berechnungen implizit zugrunde liegenden Werturteile zweckkonform sind, also tatsächlich hin zu einem Einkommenskonzept führen, dass besser geeignet ist, den Lebensstandard der Haushalte zu beschreiben als das zugrunde liegende Einkommen im steuerrechtlichen Sinn. Als Indikator greifen wir auf das mikroökonomische Klassifikationsschema inferiorer, normaler und superiorer Güter zurück.

Der Zusammenhang zwischen Einkommen y und güterspezifischen Ausgaben $c(y)$, die so genannte Engelkurve, ist bei inferioren Gütern negativ. Bei normalen Gütern steigen die Ausgaben mit verfügbarem Einkommen unterproportional, bei superioren Gütern überproportional. Auch FAST enthält Informationen zum Ausgabeverhalten von Haushalten. Verzeichnet sind Ausgaben für „Steuerberatung“, „Spenden“, „Schulgeld“ und „Wohnen“¹⁵. Da nur die wenigsten Einkommensteuererklärungen positive Ausgaben für jede der vier Güterarten enthalten, verwenden wir die Summe über alle vier Ausgabenkategorien, um einen möglichst hohen Erfassungsgrad sicherzustellen und bezeichnen diese als „Summe der Ausgaben“ (SDA). Eine solche Aggregation

¹⁴ Prinzipiell könnten Beiträge auf Grundlage „durchschnittlicher“ Beitragssätze repliziert werden, wofür aber zusätzliche Annahmen insbesondere bei nicht-identifizierbaren Privatversicherten, nötig wären.

¹⁵ Informationen über die Ausgaben für „Wohnen“ sind approximativ über Angaben zur Wohnungsbauförderung in FAST enthalten.

erscheint unproblematisch, da verschiedene Studien gezeigt haben, dass es sich zumindest bei „Spenden“, „Schuldgeld“ und „Wohnen“ um normale bzw. superiore Güter handelt.¹⁶ Für verschiedene Einkommensaggregate ermitteln wir dann die perzentil-spezifischen Durchschnittsausgaben über alle vier Güterarten. Eine Grundvoraussetzung dafür, dass ein Einkommensaggregat einen geeigneten Proxy des Lebensstandards u darstellt, ist ein schwach positiver Engelkurvenverlauf über den gesamten Einkommensbereich. Darüber hinaus ist die durchschnittliche Steigung der Engelkurve dann maximal, wenn $y_i < y_j \leftrightarrow u_i < u_j \forall i, j$ gilt. Ist der Verlauf der SDA-Engelkurve aber beispielsweise U-förmig, weist dies darauf hin, dass der Lebensstandard eines signifikanten Anteils von Haushalten mit niedrigem Einkommen bei Verwendung des zugrunde liegenden Einkommensaggregats systematisch unterschätzt wird. Eine auf dem „ökonomischen Einkommen“ basierende Engel-Kurve sollte demgegenüber eine über alle Perzentile nicht-negative Steigung besitzen.

Eine grafische Illustration der Vorgehensweise für den Fall eines nicht-inferioren Gutes gibt Abbildung 1. Im oberen Diagramm sind die Haushalte nach dem Einkommensaggregat \tilde{y} geordnet. Auf der Ordinate sind die haushaltsspezifischen Ausgaben c_j abgetragen. Bei den durch schwarz ausgefüllte Punkte symbolisierten Haushalten liege eine Differenz zwischen \tilde{y} und dem „ökonomischen Einkommen“ y vor. Für diese gelte $y > \tilde{y}$. Eine mögliche Interpretation für \tilde{y} wäre, dass es sich hierbei um die „Summe der Einkünfte“ handelt, und diese Haushalte von steuerlichen Wahlmöglichkeiten Gebrauch gemacht haben. Sie wandern dann bei Verwendung des Aggregats y parallel zur Abszisse (mittleres Diagramm).

Der Zusammenhang zwischen Ausgaben und der Höhe des Einkommensaggregats y ist jetzt eindeutig positiv. Im unteren Diagramm ist die Differenz zwischen den beiden Funktionen $c_j(\tilde{y}) - c_j(y) \forall j$ abgetragen. Ist die Steigung dieser Differenzfunktion wie hier positiv, führt die Verwendung der „Summe der Einkünfte“ zu einer relativen Unterschätzung des Lebensstandards von Haushalten mit niedrigem Einkommen.

Empirischer Befund

Abbildung 2 zeigt die durchschnittliche Höhe der „Summe der Ausgaben“ für alle Perzentile bei Verwendung alternativer Ordnungskriterien. Die grau dargestellte Engelkurve ergibt sich bei Verwendung des Ordnungskriteriums „Summe der Einkünfte“, die schwarz dargestellte Engelkurve, wenn die Haushalte aufsteigend nach der Höhe ihres „ökonomischen Einkommens“ sortiert werden.

Die Steigung der Engelkurve auf Grundlage von SDE ist bis zum neunzehnten Perzentil negativ. Nach obigen Überlegungen ist SDE also ein verzerrter Proxy für den Lebensstandard der Haushalte. Insbesondere kommt es bei Verwendung dieses Einkommensaggregats zu einer Unterschätzung des Lebensstandards von Haushalten mit niedrigem Einkommen. Die Engelkurve auf Grund-

¹⁶ Empirische Schätzungen bezüglich der Ausgabenkategorie „Spenden“ ergeben durchweg positive, jedoch über die Studien stark variierende Einkommenselastizitäten. So ermitteln beispielsweise Manzoor und Straub (2005) für die USA eine Einkommenselastizität von ca. 0,40, Micklewright und Wright (2004) von 1,12, Ferris und West (2003) von 2,5. Nach einer Metadatenanalyse von Clotfelter (2002) beläuft sich die typische Einkommenselastizität für Spenden für den Zeitraum bis 1990 auf 0,40 bis 0,80. Ab 1990 steigt sie nach dieser Studie leicht an. Positive Einkommenselastizitäten werden auch für die Ausgabenkategorie „Wohneigentum“ beobachtet. Die Einkommenselastizität für Wohneigentum liegt nach einer Metaanalyse von Meen (1996) durchschnittlich bei rund 1,0 (vgl. auch European Central Bank (2003)). Das Ausgabeverhalten der Haushalte für „Bildung“ wird typischerweise für Länder mit (teil)privatisierten Bildungssystemen untersucht. So ermitteln Duncombe und Yinger (1998) und Acemoglu und Pischke (2001) für die USA Einkommenselastizitäten für „Bildung“ von 0,89 bzw. 0,14. Nicht ganz unproblematisch erscheint der Ansatz der Ausgabenkategorie „Steuerberatungskosten“, da diese bei Gewinneinkünften bereits als betrieblicher Aufwand geltend gemacht werden können und bei hohen Verlusten ein besonderer Beratungsbedarf besteht. Wir haben diese Ausgabenkategorie dennoch berücksichtigt, um einen ausreichenden Erfassungsgrad sicher zu stellen. Ab FAST01 stehen durch die teilweise steuerliche Abzugsfähigkeit für die private Altersvorsorge zusätzliche Ausgabenkategorien zur Verfügung.

lage des „ökonomischen Einkommens“ hat für die unteren Perzentile zwar ebenfalls eine negative Steigung, der quantitative Ausgabenrückgang mit steigender Einkommenshöhe ist jedoch

Abbildung 1:

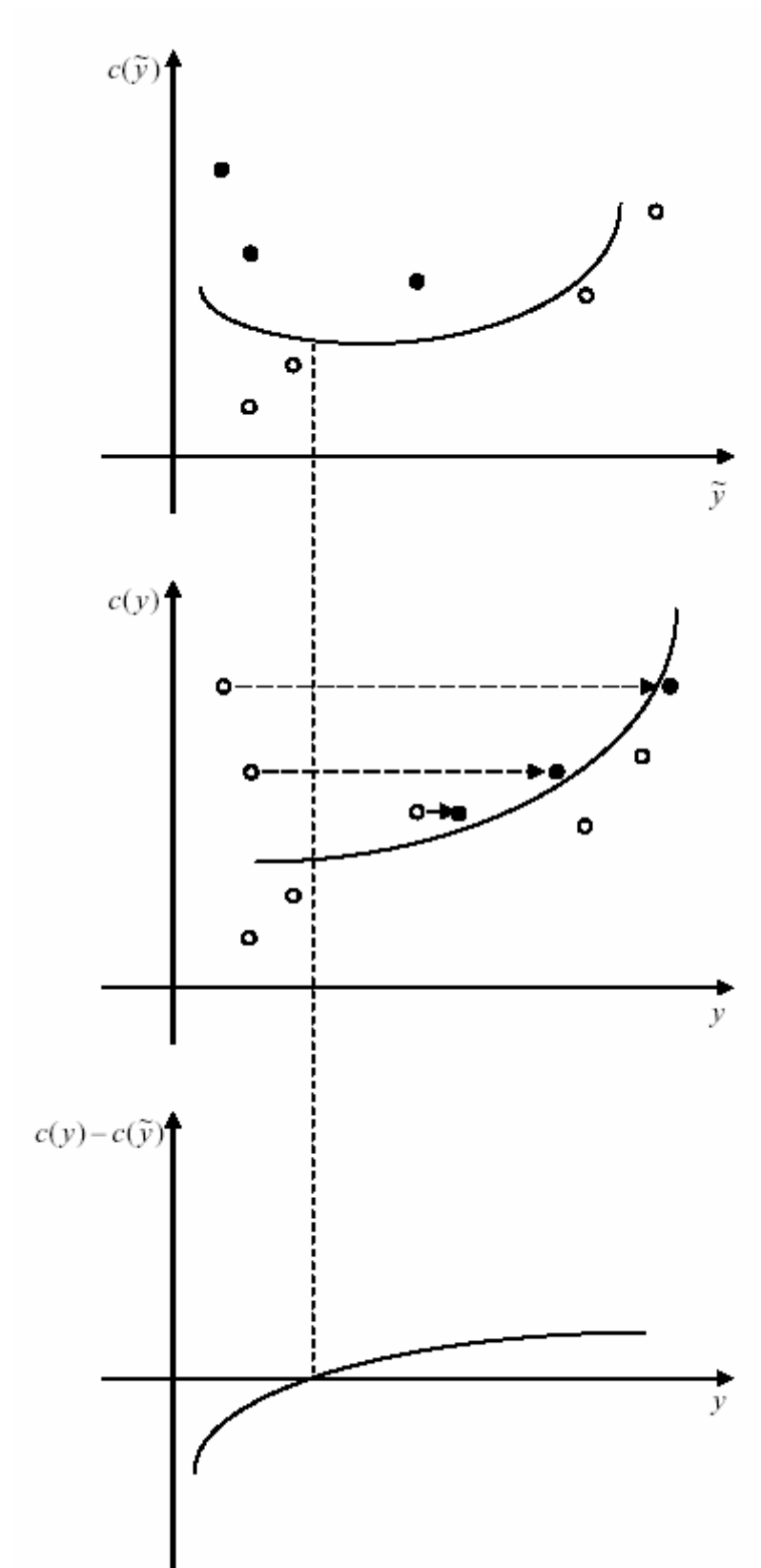
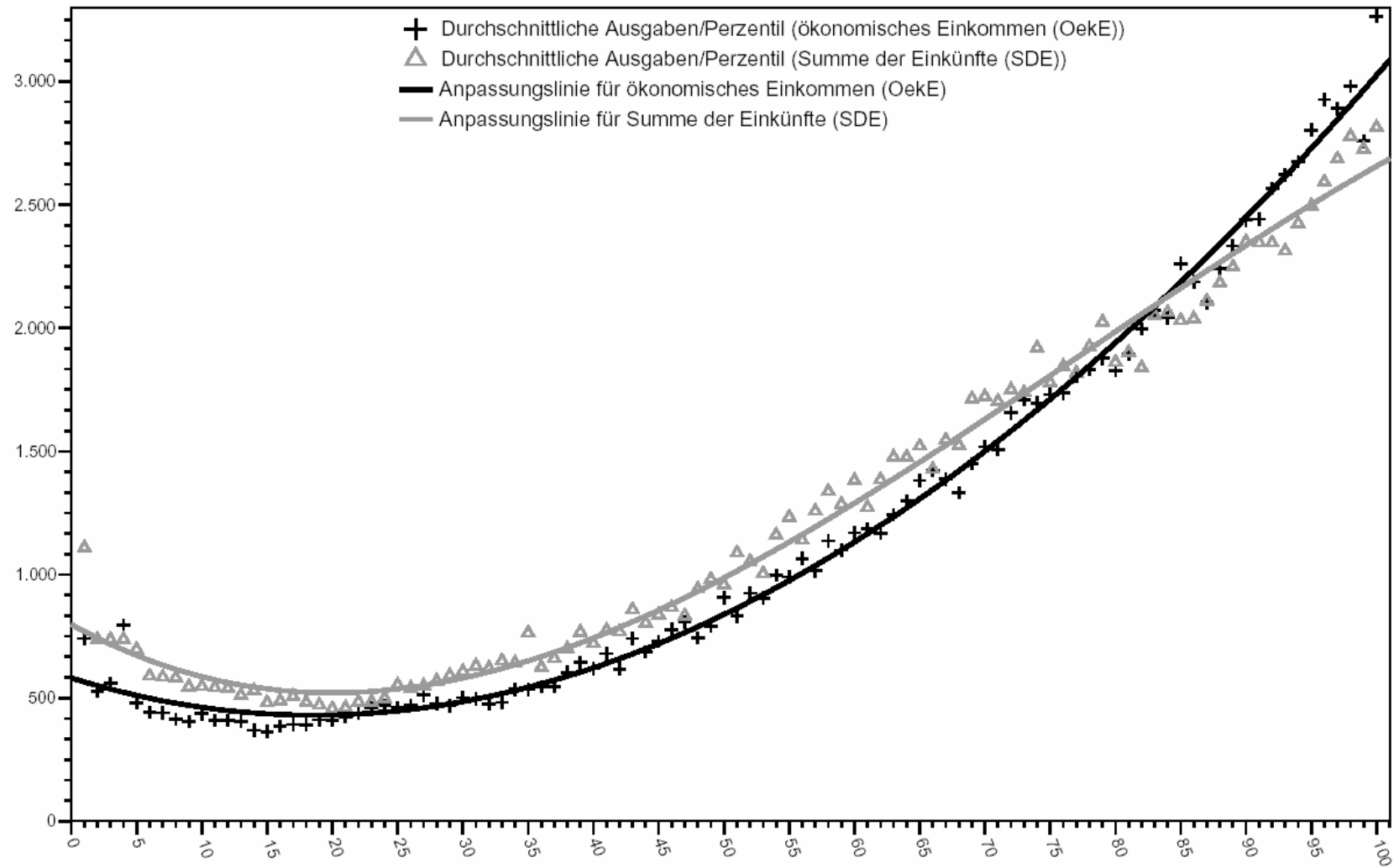


Abbildung 2:



geringer. Zudem schneiden sich beide Engelkurven: Bis zum 84. Perzentil weist die Engelkurve für OekE im Vergleich zu der für SDE ein niedrigeres Ausgabenniveau aus, danach ein höheres. Mit OekE lässt sich im Sinne der obigen Ausführungen somit eine Ordnung generieren, die den Lebensstandard der Haushalte exakter zu beschreiben in der Lage ist als bei Verwendung von SDE. Es bleibt aber zu konstatieren, dass die vorgenommenen Korrekturschritte – wie ein fallender Verlauf der Engelkurve auf Basis von OekE für die unteren Perzentile zeigt – für eine fehlerfreie Abbildung des materiellen Lebensstandards offensichtlich nicht ausreichen. Weitere Korrekturschritte, wie die Imputation unterstellter Einkommen aus selbst genutzten Vermögenswerten, die Berücksichtigung individueller Sozialversicherungssalden sowie die Vervollständigung der Statistik um einkommensschwache Haushalte, sind aber nur bei Verwendung von Informationen aus weiteren Statistiken möglich.¹⁷

4.2 Wanderungsbewegungen

Aufgrund der vorgenommenen Einkommenskorrekturen kann es zu Wanderungen der Zensenten über die Perzentile hinweg kommen. Abbildung 3 zeigt für die einzelnen Korrekturschritte die Wanderbewegungen exemplarisch für Zensenten des ersten Perzentils. Die Perzentileinteilung erfolgte anhand der „Summe der Einkünfte“ vor jeglicher Modifikation. Auf der Ordinate wird die Anzahl der Zensenten abgetragen, die nach Durchführung der Korrekturschritte eine entsprechende Anzahl von Perzentilen in der Verteilung aufsteigen. So wandern nach Durchführung aller Korrekturen z. B. rund 6000 Zensenten aus dem ersten Perzentil (nach SDE) um neunundneunzig Perzentilschritte ($dP = 99$) aufwärts in das oberste hundertste Perzentil (nach OekE). Bei diesen Zensenten handelt es sich überwiegend um Bezieher von Einkünften aus Kapital, Vermietung und Verpachtung oder aus Gewerbebetrieb, denen es gelingt, sich aufgrund steuerlicher Gestaltungsmöglichkeiten weitgehend der Besteuerung zu entziehen.

In Abbildung 4 sind die Wanderungsbewegungen für insgesamt 11 Perzentile vergleichend gegenübergestellt. Wie in den bisherigen Abbildungen ist auf der Abszisse die Wanderungsentfernung (dP) über die Perzentile dargestellt. Ein Wert von $dP = 1$ ($dP = -1$) entspricht einer Wanderung um ein Perzentil nach oben (unten). Für $dP = 0$ wandert der Zensent nicht, dargestellt durch eine weiße Markierung. Die Ordinate gibt die Zahl von Zensenten an, die eine bestimmte Anzahl von Perzentilen wandern. In der dritten Dimension sind die Einkommensperzentile nach SDE abgetragen. Die stärksten positiven Wanderungsbewegungen ergeben sich für die unteren SDE-Perzentile (hier dargestellt: erstes und zehntes Perzentil). Im 10., 20. und 30. Perzentil ist die Zahl der Fälle, die gar nicht wandern, minimal, danach nimmt – wie die steigende Fallzahl für $dP = 0$ zeigt – die Anzahl der nicht wandernden Fälle stetig zu.

4.3 Verteilungseffekte

Abbildung 5 zeigt die Dichtefunktionen der Einkommen von einzeln und getrennt Veranlagten bzw. für Zusammenveranlagte für die Einkommensaggregate SDE, Brutto2 und OekE. Auffällig ist, dass ein relativ hoher Anteil von ca. 6 bis 8% der einzeln und getrennt Veranlagten und ca. 2 bis 3% der Zusammenveranlagten eine Einkommensteuererklärung abgegeben haben, obgleich ihr zu versteuerndes Einkommen Null ist. Diese Steuerpflichtigen haben entweder tatsächlich keine

¹⁷ In diesem Zusammenhang schlägt Zwick (2004, S. 293) die Schaffung integrierter Mikrodatenfiles aus den Einzelstatistiken EVS, Einkommensteuerstatistik und Sozialhilfestatistik vor.

Abbildung 3:

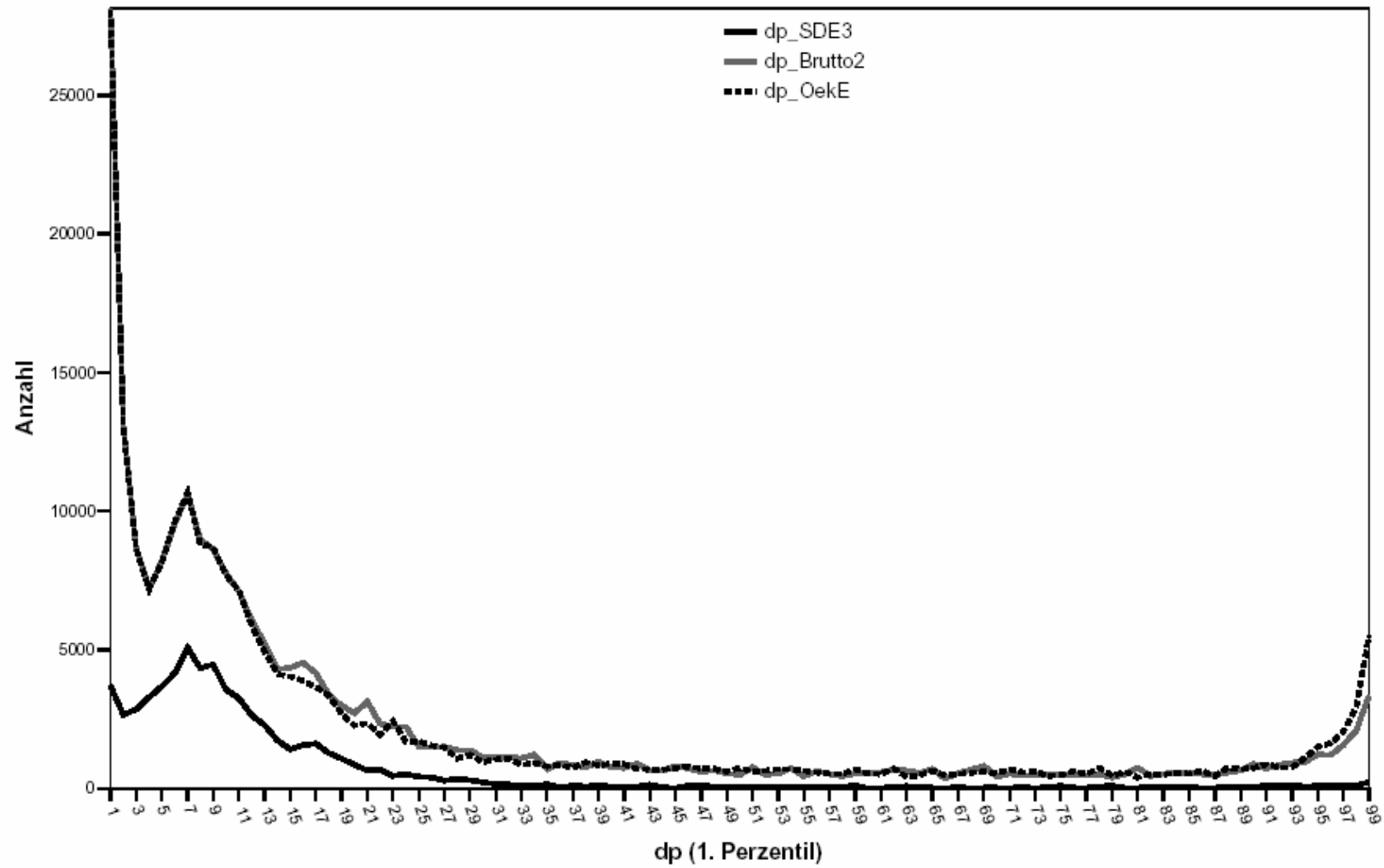


Abbildung 4:

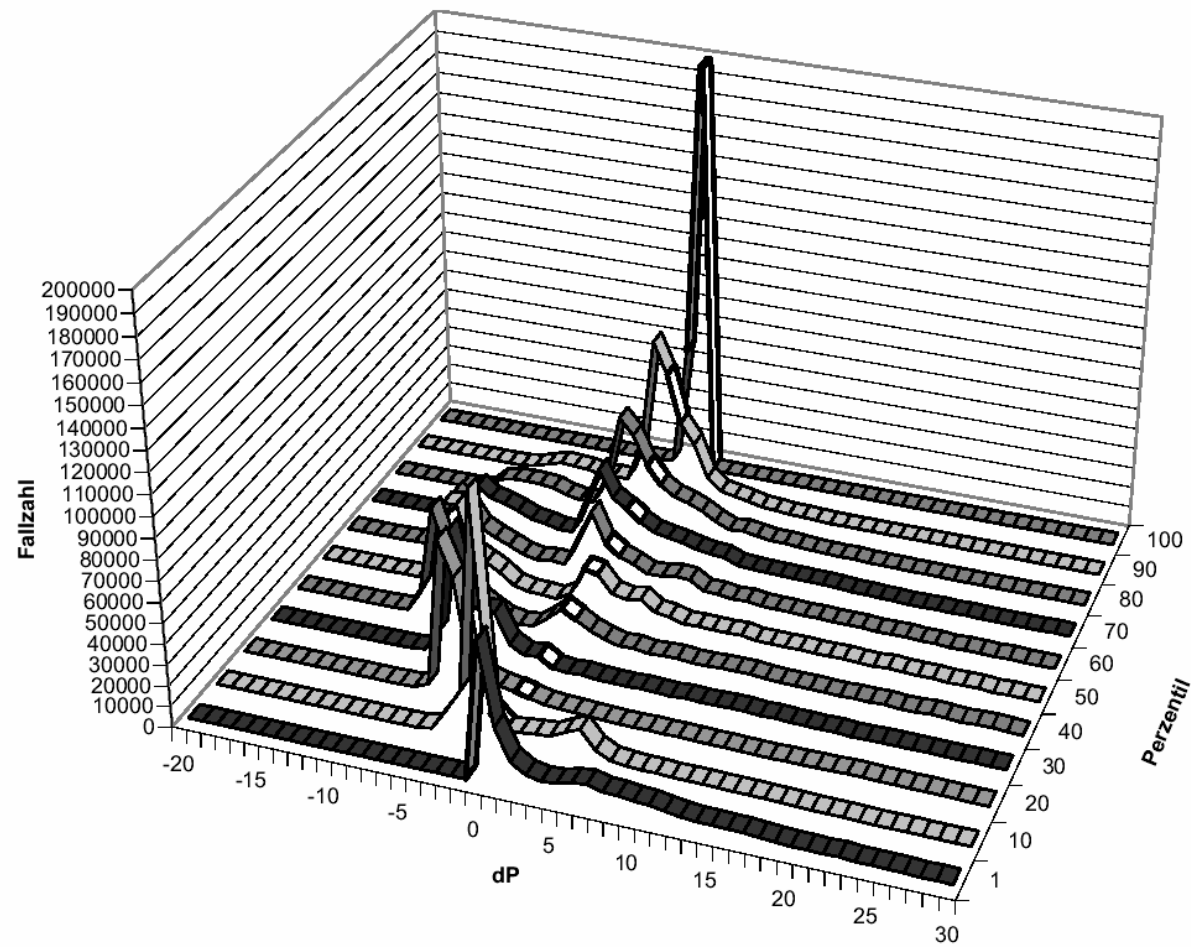


Abbildung 5a:

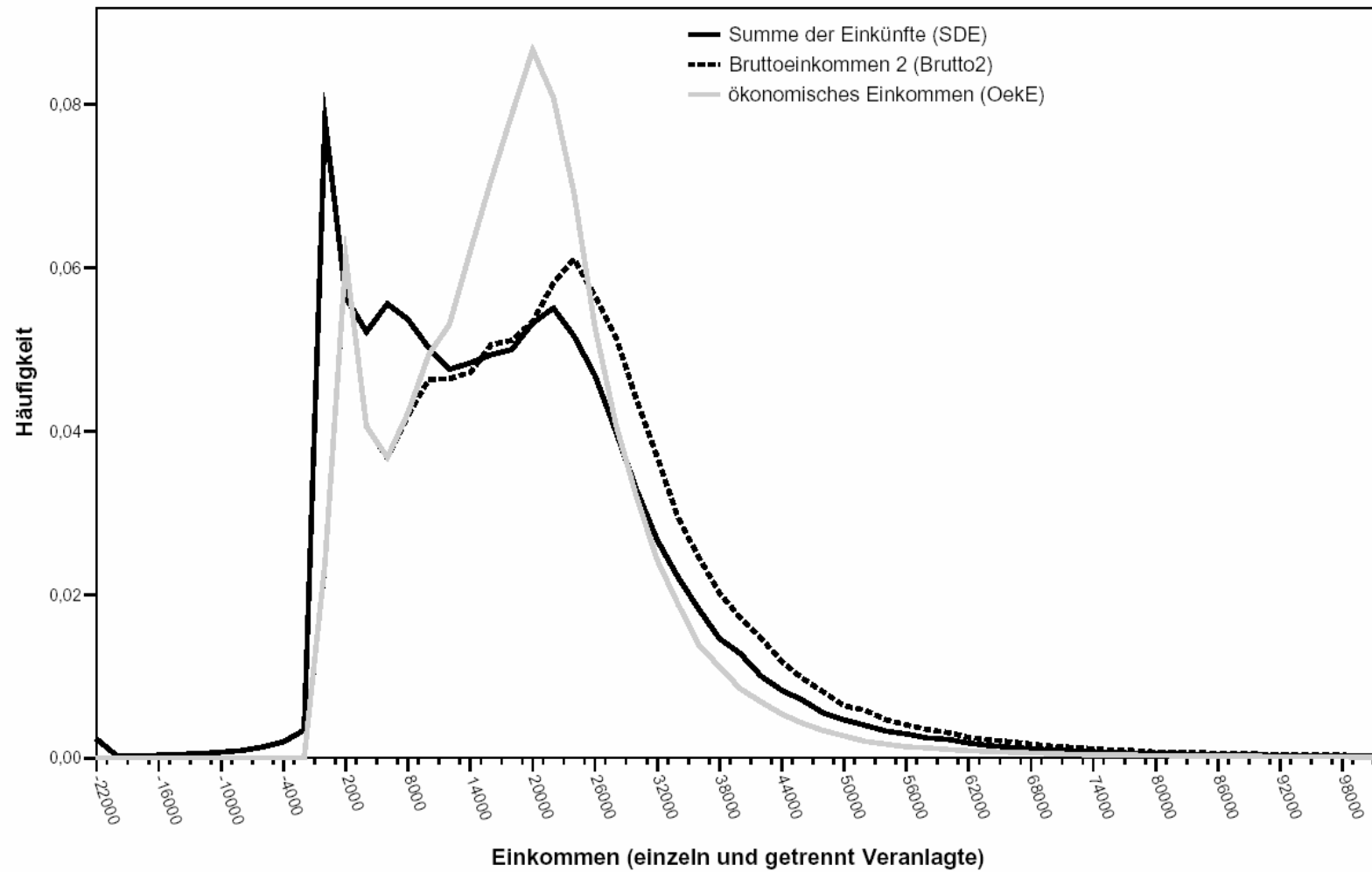
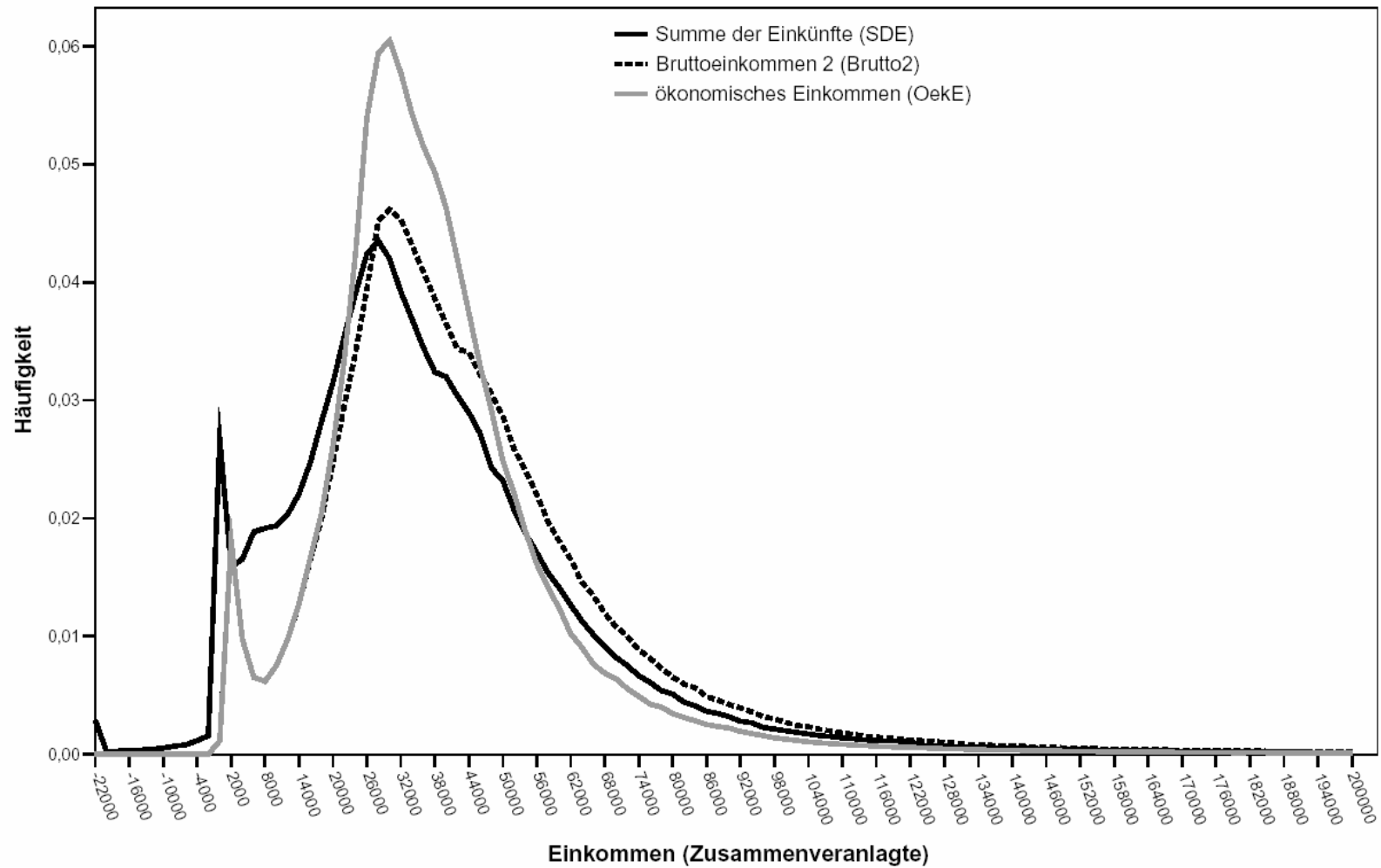


Abbildung 5b:



Einkünfte erzielt, oder sie konnten aufgrund steuerlicher Gestaltungsmöglichkeiten ihre steuerpflichtigen Einkünfte bis auf Null reduzieren.

Erwähnenswert sind auch die Unstetigkeiten der Dichtefunktionen für SDE und Brutto2, insbesondere bei einzeln und getrennt Veranlagten. Die Unstetigkeiten ergeben sich bei den Untergrenzen der einkommensteuerlichen Progressionsstufen.¹⁸ Dies ist ein Hinweis darauf, dass gerade für Steuerpflichtige, deren Einkommen nicht zu weit über dem Referenzwert für die nächste Progressionsstufe liegen, offensichtlich besondere Anreize bestehen, durch Ausnutzung steuerlicher Gestaltungsmöglichkeiten einer höheren steuerlichen Marginalbelastung auszuweichen. Dass dieses Muster bei Zusammenveranlagten weniger offensichtlich ist, lässt sich mit der Existenz von Freibeträgen (insbesondere Kinder- und Haushaltsfreibeträge) erklären, wodurch die Progressionsstufen verwischen.

Mit Ausnahme der Dichtefunktion für OekE unterscheiden sich die Dichtefunktionen für die beiden anderen Aggregate vom Verlauf her wenig. Bei Verwendung des „ökonomischen Einkommens“ ändert sich jedoch der für die beiden anderen Einkommensaggregate typische Verlauf. So sinken die Fallzahlen in der Nähe des Haushaltsfreibetrags stark ab. Dies lässt sich damit begründen, dass beim Übergang zu OekE viele Haushalte in höhere Einkommensbereiche wandern. Entsprechend dominiert die Dichtefunktion für OekE diejenigen für SDE und Brutto2. Mit Einsetzen der höchsten steuerlichen Progressionsstufe nähern sich die Dichtefunktionen wieder einander an.

Wie ändert sich die gemessene Einkommensungleichheit in Abhängigkeit von dem zugrunde liegenden Einkommensaggregat? Ein Index, der vielfach in Verteilungsanalysen Verwendung findet, ist der Gini-Koeffizient. Dieser auf das Intervall $[0, 1]$ normierte Koeffizient misst die doppelte Fläche zwischen Gleichverteilungsgerade und Lorenzkurve. Der Gini-Koeffizient steigt dabei mit zunehmender Ungleichheit in der Verteilung. Tabelle 7 fasst für alternative Einkommensdefinitionen (ZVE, Brutto2, OekE) die auf Basis von FAST ermittelten Gini-Koeffizienten zusammen (Zeile „eigene Berechnungen“) und stellt diese den Ergebnissen anderer Untersuchungen vergleichend gegenüber.

Da verschiedene Haushaltstypen verschiedene Einkommensbedürfnisse aufweisen, haben wir die Einkommen der Zensenten mit haushaltsspezifischen Äquivalenzskalen gewichtet. Eine Äquivalenzskala ist das Einkommensverhältnis zweier Haushalte, das beiden den gleichen Lebensstandard erlaubt. Da es keinen Konsens über die Wahl der „richtigen“ Äquivalenzskala gibt, wurden alternative Skalen-Spezifikationen verwendet, die in der Tabelle erläutert sind. Da in FAST keine „Haushalte“ sondern „Zensenten“ verzeichnet sind, liegt der Analyse die Konvention zugrunde, dass Zusammenveranlagte einen gemeinsamen Haushalt führen, während einzeln und getrennt Veranlagte als 1-Erwachsenen-Haushalte betrachtet werden. Jedes in einer Einkommensteuererklärung angegebene Kind wird dem Haushalt zugerechnet. Ein Vergleich der Gini-Koeffizienten auf Basis von SDE, Brutto2 und OekE (jeweils transformiert in äquivalente Einkommen) zeigt, dass die gemessene Ungleichheit in der Verteilung sinkt, wenn a) für steuerliche Gestaltungsmöglichkeiten (Vergleich SDE und Brutto2) und b) für den Nettotransfer (Einkommensteuer, Kirchensteuer, Solidaritätszuschlag, Unterhaltsaufwendungen und Kindergeld) korrigiert wird.

Beim Vergleich mit den Ergebnissen anderer Studien muss berücksichtigt werden, dass diese auf andere Datensätze zurückgreifen, sowie andere Einkommensdefinitionen und Äquivalenzskalen

¹⁸ In 1998 setzt die Besteuerung für einzeln bzw. getrennt Veranlagte ab 6322 Euro ein, die weiteren Tarifestufen folgen bei 29.983 Euro sowie bei 61.376 Euro. Für Zusammenveranlagte gelten die doppelten Beträge.

Tabelle 7: Gini-Koeffizienten

Quelle	Datengrundlage	Einkommensdefinition	Korrespondierendes Aggregat	Äquivalenzskala	Gini		
Bach et al. (2006)	Lohn- und Einkommensteuerstatistik 1998	Bruttoeinkommen mit Kapitalgewinn	Brutto2	keine	0,449		
		Nettoeinkommen mit Kapitalgewinn	OekE		0,411		
Merz et al. (2005)	Lohn- und Einkommensteuerstatistik 1998	Nettoeinkommen	OekE (abzgl. Sozialabgaben)	OECD, alt, angepasst ^b	0,418		
Bundesregierung (2005)	EVS 1998	Markteinkommen	Brutto2	OECD, neu	0,449		
		Verfügbares Einkommen	—		0,255		
Förster und Mira D’Ercole (2005) ^a	SOEP 2001	Verfügbares Einkommen	—	$[N^{0,5}]^c$	0,277		
Stauder und Hüning (2004)	EVS 1998	Nettoeinkommen	—	OECD, alt	0,28-0,29		
	Mikrozensus 1998	Nettoeinkommen	—	OECD, alt	0,27		
Eigene Berechnungen	FAST 1998	Summe der Einkünfte	SDE	OECD, alt OECD, neu	0,449 0,440		
		Bruttoeinkommen	Brutto2	OECD, alt OECD, neu keine	0,406 0,396 0,433		
				Ökonomisches Einkommen	OekE	OECD, alt OECD, neu OECD, alt, angepasst keine	0,356 0,344 0,353 0,389

Anmerkungen: Die Lohn- und Einkommensteuerstatistik 1998 ist eine nicht-anonymisierte 10%ige Stichprobe aller zur Einkommensteuer veranlagten Zensenten. Die alte OECD Skala weist dem ersten erwachsenen Haushaltsmitglied ein Gewicht von 1,0 zu. Jeder weiteren Person älter 14 Jahre wird ein Gewicht von 0,7 zugewiesen, jüngeren ein Gewicht von 0,5. Die neue OECD Skala weist jeder weiteren Personen älter als 14 Jahre ein Gewicht von 0,5 zu, jüngeren ein Gewicht von 0,3. ^a Die Autoren berücksichtigen nur Haushalte mit einem Einkommen von mindestens DM 5.000. ^b Da das Alter von Kindern in der Lohn- und Einkommensteuerstatistik nicht als Variable enthalten ist, weisen Merz et al. (2005) allen Kindern ein Gewicht von jeweils 0,5 zu.

^c N ist die Anzahl der Familienmitglieder.

verwenden. So werden in der Definition des „Nettoeinkommens“ in Merz et al. (2005) im Unterschied zu OeKE Pflichtbeiträge zur Sozialversicherung abgezogen. Unter Verwendung der alten OECD-Skala¹⁹ und auf Grundlage der nicht-anonymisierten Lohn- und Einkommensteuerstatistik finden Merz et al. (2005) leicht höhere Gini-Koeffizienten als unser Ergebnis für OeKE (OECD, alt, angepasst). Das Sozialversicherungssystem hat für die Gruppe der zur Einkommensteuer Veranlagten somit einen regressiven Charakter. Ursachen hierfür können Beitragsbemessungsgrenzen sowie die überproportionale Belastung der Arbeitseinkommen sein. Beides trägt zu einer relativen Beitragsentlastung „Besserverdienender“ bei. Zudem arbeiten Merz et al. (2005) mit nicht-anonymisierten Daten. Dies ermöglicht den Autoren exaktere Korrekturen als dies in FAST – insbesondere für Zensenten, die unter den Anonymisierungsbereich 5 fallen – möglich ist.

Ein Vergleich unserer Ergebnisse mit denen in Bundesregierung (2005) für die Aggregate „Markteinkommen“ und „verfügbares Einkommen“ unter Verwendung der neuen OECD-Skala zeigt, dass der Gini-Koeffizient von 0,3958 nach eigenen Berechnungen etwas niedriger ist als der in Bundesregierung (2005) mit 0,449. Eine mögliche Erklärung ist, dass den Ergebnissen in Bundesregierung (2005) die Einkommens- und Verbrauchsstichprobe zugrunde liegt, in der im Unterschied zu FAST auch solche Haushalte verzeichnet sind, die nicht zur Einkommensteuer veranlagt werden. Die Verteilung auf Grundlage des „Nettoeinkommens“ impliziert nach Bundesregierung (2005) eine deutlich weniger ungleiche Verteilung als unser Ergebnis für (OeKE). Dies ist wahrscheinlich darauf zurückzuführen, dass in Bundesregierung (2005) Sozialversicherungstransfers vollständig abgebildet werden.

Bach et al. (2006) verzichten – eventuell aufgrund der Identifizierungsproblematik von Haushaltstypen in der Lohn- und Einkommensteuerstatistik²⁰ – auf die Berechnung äquivalenter Einkommen. Ein direkter Vergleich mit unseren Ergebnissen (keine Skala) zeigt, dass die Gini-Koeffizienten in Bach et al. (2006) etwas höher sind als die von uns ermittelten. Dies lässt sich damit erklären, dass Bach et al. (2006) nicht-anonymisierte Daten verwenden und daher sowohl den oberen als auch unteren Rand der Verteilung exakter darstellen können. Die Veränderung des Ginis $G^{Brutto2} - G^{OeKE} \approx 0,04$ beim Übergang vom Bruttoeinkommensbegriff zum ökonomischen Einkommen in Bach et al. (2006) ist vergleichbar mit der Differenz nach unseren Berechnungen. Nicht direkt vergleichbar mit unseren Ergebnissen sind diejenigen in Förster und Mira D’Ercole (2005) sowie in Stauder und Hüning (2004), die mit dem Sozio-ökonomischen Panel bzw. Mikrozensus andere Datengrundlagen und andere Einkommensdefinitionen verwenden.

5 Tarifliche versus effektive Steuerbelastung

Der auf das zu versteuernde Einkommen (ZVE) anzuwendende Steuertarif ist in Deutschland progressiv, d.h., der Durchschnittssteuersatz steigt, wenn ZVE steigt. In der vorangegangenen Analyse haben wir gesehen, dass ZVE den tatsächlichen Mittelzufluss der Zensenten im Vergleich zu Brutto2 unterschätzt. Zudem steigt diese Unterschätzung mit der Höhe von ZVE. Es stellt sich daher die Frage, welcher effektiven Steuerbelastung die Haushalte tatsächlich unterliegen. Abbildung 6 zeigt für einzeln bzw. getrennt Veranlagte sowie für Zusammenveranlagte den Zusammenhang zwischen Durchschnittssteuersatz und Einkommenshöhe bei Verwendung alternativer Einkommensaggregate. Bei den angegebenen Durchschnittssteuersätzen handelt es sich um

¹⁹ Vgl. hierzu die Anmerkungen in Tabelle 8.

²⁰ Vgl. Kapitel 2.3.

Abbildung 6a:

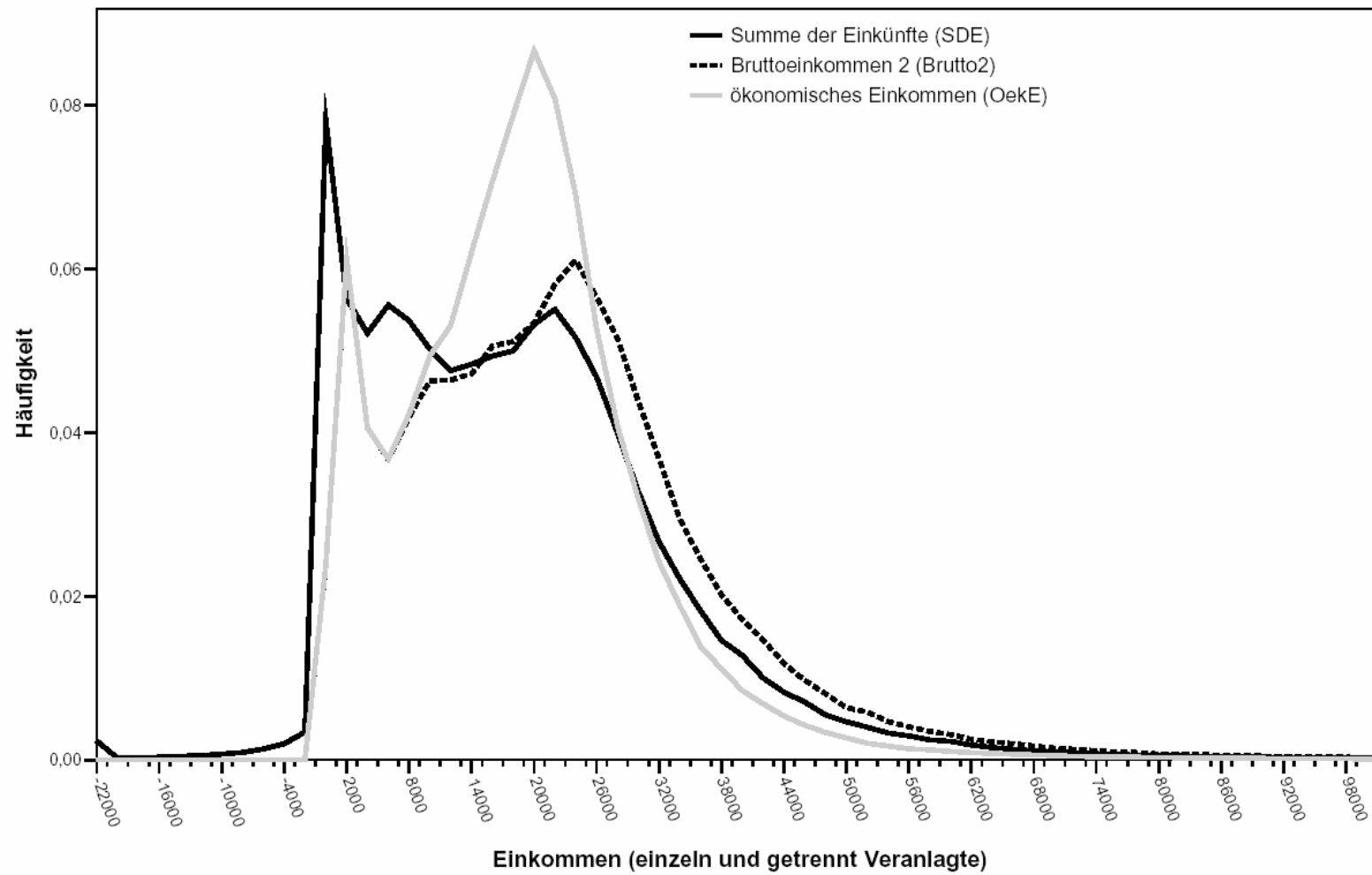


Abbildung 6b:

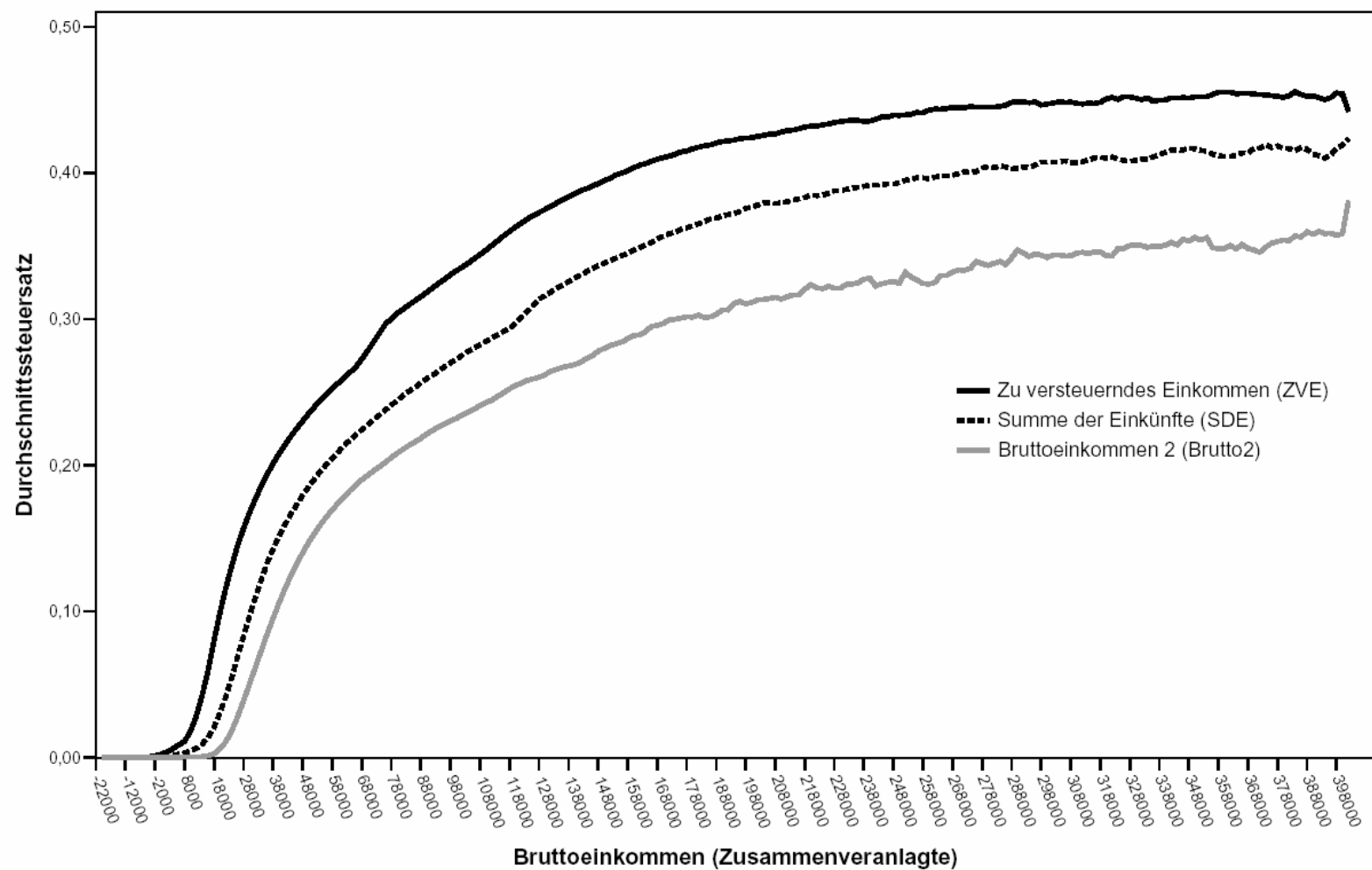


Abbildung 7a:

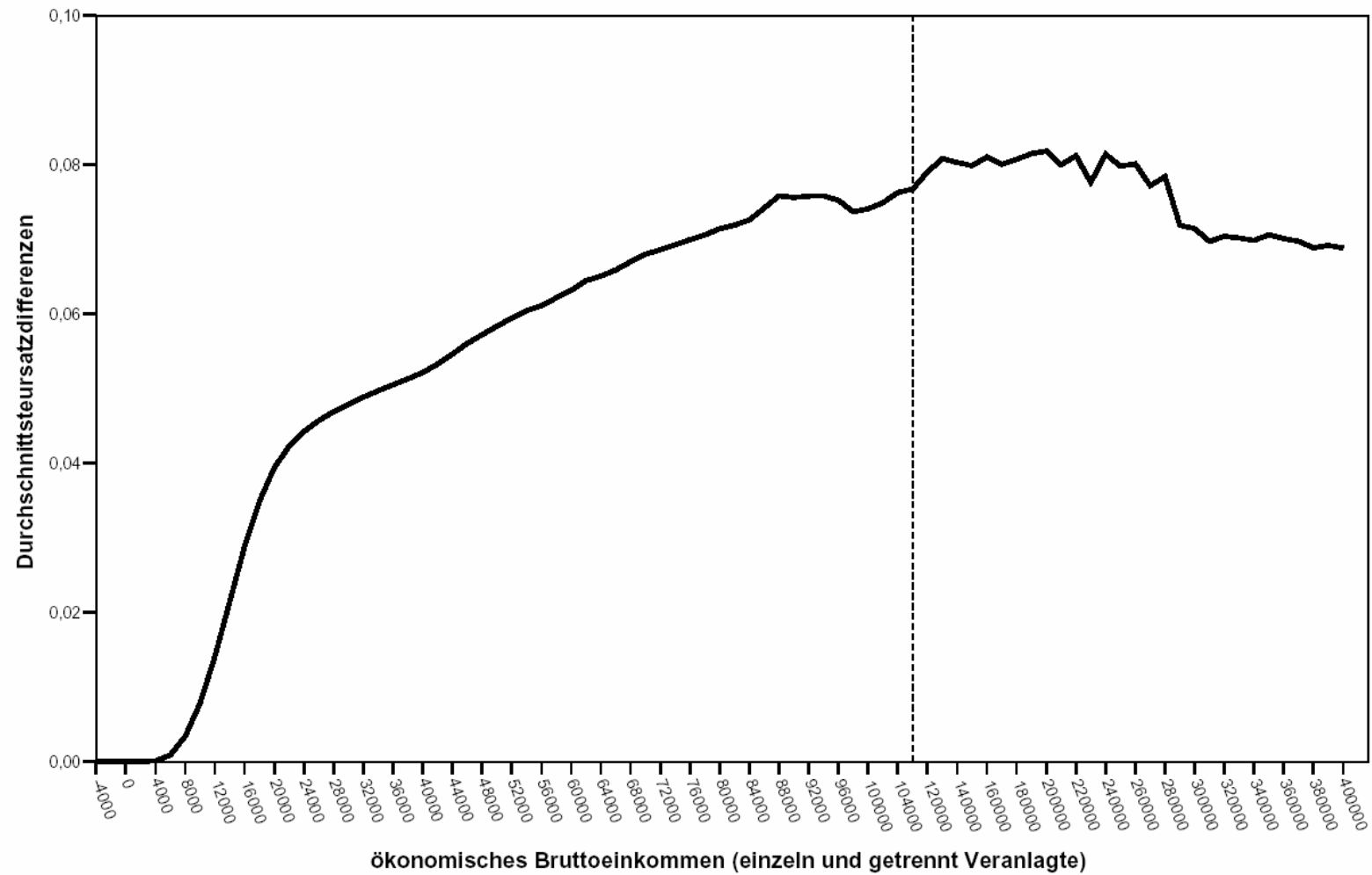
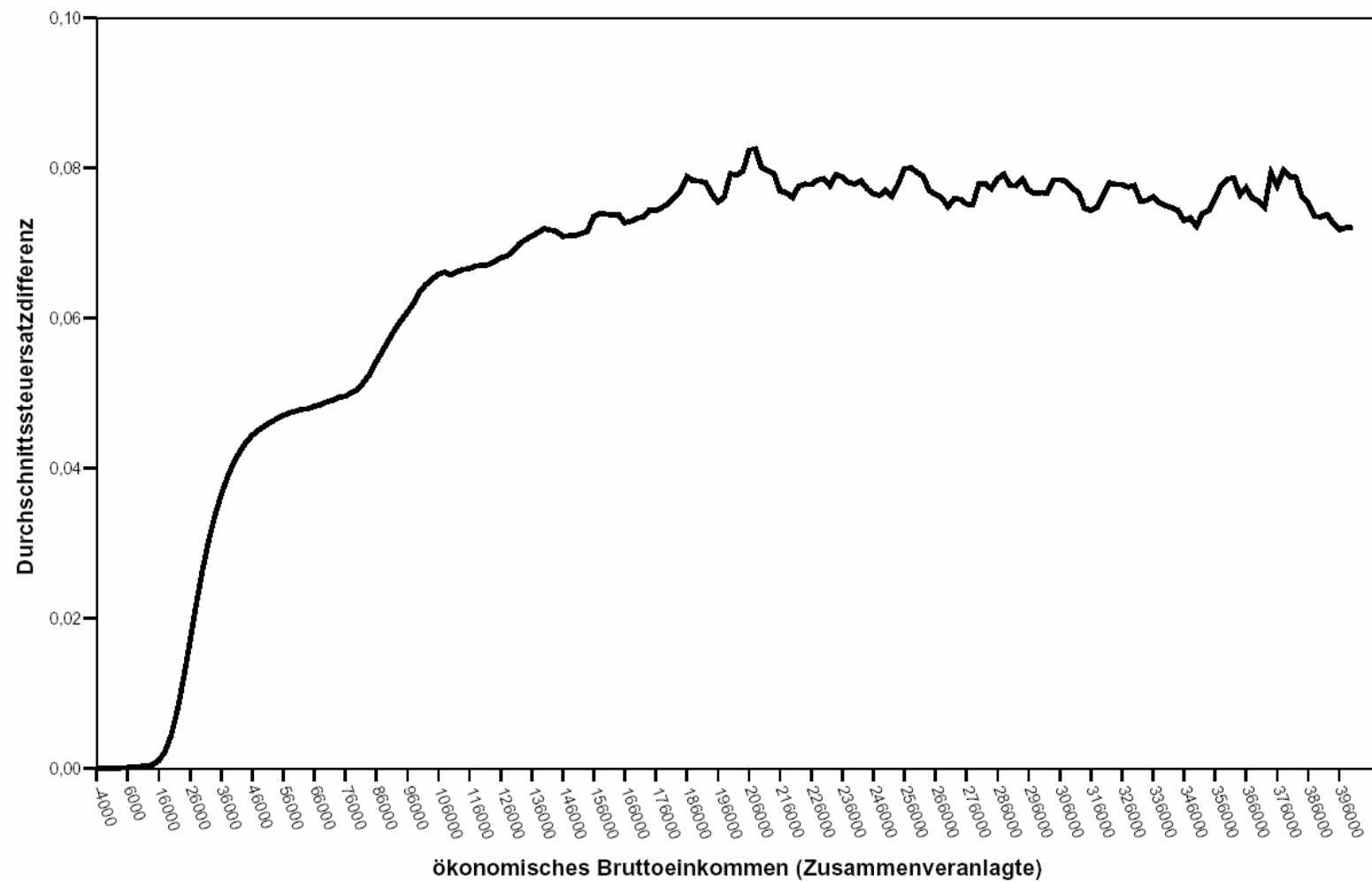


Abbildung 7b:



gleitende Durchschnitte fünfter Ordnung über die Einkommensperzentile. Es zeigt sich, dass der Durchschnittssteuersatz unabhängig vom zugrunde liegenden Einkommensaggregat mit dem Einkommensniveau ansteigt. Die Einkommensbesteuerung hat somit effektiv progressive Effekte, verringert also die Ungleichheit in der Einkommensverteilung.

Allerdings ist die effektive Steuerbelastung niedriger als es der Steuertarif selbst vermuten ließe. Abbildung 7 veranschaulicht dieses Ergebnis, wiederum differenziert nach einzeln bzw. getrennt Veranlagte sowie nach Zusammenveranlagten. Dargestellt ist die Differenz der durchschnittlichen Steuerbelastung unter Zugrundelegung von ZVE und Brutto2. Da Brutto2 einen weiter gefassten Einkommensbegriff als ZVE darstellt, ist die effektive, „gefühlte“ Steuerbelastung insgesamt niedriger als es der Steuertarif signalisiert. Die Stärke des Effekts ist aber einkommensabhängig: Bis zu einem Einkommen von ca. 150.000 Euro bei einzeln und getrennt Veranlagten bzw. von ca. 200.000 Euro bei Zusammenveranlagten steigt die Durchschnittssteuersatzdifferenz kontinuierlich bis auf ca. 8 Prozentpunkte. Danach nimmt die Differenz wieder leicht ab. Dies ist vor allem auf die oben angeführten Anonymisierungsprobleme, die höhere internationale Mobilität von Kapitaleinkommen (die vor allem in den oberen Perzentilen erzielt werden) mit den damit einhergehenden Möglichkeiten zur Steuerausweichung und auf die Konstanz des Marginalsteuersatzes ab Einsetzen der obersten Progressionsstufe zurückzuführen. Zu vergleichbaren Ergebnissen auf Basis der EVS kommen – für den dort verzeichneten Einkommensbereich – auch Lang et al. (1997).

Auch Bach et al. (2006) bestimmen die effektive Steuerprogression in Deutschland. Sie greifen bei ihrer Analyse auf die nicht-anonymisierte 10%ige Stichprobe der Lohn- und Einkommensteuerstatistik 1998 zurück. Ihr Fokus gilt insbesondere den Progressionswirkungen innerhalb des obersten Einkommensperzentils der Top-Einkommensbezieher. Tabelle 8 stellt die Ergebnisse von Bach et al. (2006) unseren Ergebnissen gegenüber. Um Vergleichbarkeit zu gewährleisten, haben wir die Steuerbelastungen wie in Bach et al. (2006) zunächst unter Verzicht auf eine Unterscheidung in einzeln bzw. getrennt sowie für Zusammenveranlagte bestimmt. Ein Vergleich der angegebenen Werte in den Spalte 1 und 2 (SDE bzw. AGI in Bach et al. (2006)) mit den eigenen Berechnungen (Spalten 4 (SDE) und 5 (Brutto2)) zeigt, dass die Ergebnisse insgesamt sowohl quantitativ als auch qualitativ sehr ähnlich sind.²¹ Stärkere Unterschiede lassen sich nur für das erste Dezil feststellen. Hier mussten aufgrund anonymisierungsbedingter Informationsmängel in FAST (vgl. Kapitel 3.1) viele Fälle eliminiert werden, die bei Bach et al. (2006), die nicht-anonymisierte Daten verwenden, enthalten sind. Die Spalten 7–9 und 11–13 geben die durchschnittlichen Steuerbelastungen innerhalb der Dezile für einzeln bzw. getrennt Veranlagte sowie für Zusammenveranlagte an. Insbesondere infolge des Ehegattensplittings ist die durchschnittliche Steuerbelastung für einzeln bzw. getrennt Veranlagte deutlich höher als für Zusammenveranlagte. Die entsprechenden Fallzahlen in Prozent der Gesamtfallzahl im jeweiligen Perzentil sind ebenfalls ausgewiesen (Spalten 6 und 10, *N in %*).

Zusammenfassend lässt sich festhalten, dass die deutsche Einkommensteuer eindeutig progressive Effekte hat. Allerdings ist die effektive Progression nicht so stark, wie es der Steuertarif zunächst vermuten ließe. Ursache hierfür ist die Diskrepanz zwischen der Höhe von ZVE und dem tatsächlichen Mittelzufluss der Zensenten (Brutto2), die ihren Ursprung vor allem in steuerrechtlichen Regelungen hat. Diese Ergebnisse sind für zukünftige Studien auf Basis von FAST von Bedeutung. Zwar sind in FAST alle Einkommensbezieher repräsentativ und ab einer bestimmten

²¹ Mögliche Ursachen für Differenzen sind neben den bereits diskutierten Unterschieden in Datengrundlage und Einkommensdefinitionen hier die verwendeten Ordnungskriterien. Im Gegensatz zu der hier vorgenommenen Sortierung nach Brutto2 ordnen Bach et al. (2006) die Einkommen aufsteigend nach „aggregated gross income without capital gains“.

Tabelle 8: Dezilspezifische Durchschnittssteuersätze

Bach et al. (2006)			Eigene Berechnungen										
Alle Veranlagten			Alle Veranlagten			Einzel- bzw. getrennt Veranlagte				Gemeinsam Veranlagte			
Durchschnittliche Steuerbelastung			Durchschnittliche Steuerbelastung			Durchschnittliche Steuerbelastung				Durchschnittliche Steuerbelastung			
Dezil	SDE	AGI	ZVE	SDE	Brutto2	N in %	ZVE	SDE	Brutto2	N in %	ZVE	SDE	Brutto2
1	-39,2	207,1	0,0	0,3	0,1	81,2	0,0	0,3	0,1	18,8	0,0	0,1	0,0
2	2,1	1,6	2,0	1,5	1,2	82,6	2,3	1,8	1,4	17,4	0,1	0,1	0,0
3	6,6	5,3	7,8	6,4	5,4	71,1	10,9	8,8	7,5	28,9	0,3	0,2	0,2
4	9,5	8,1	11,5	9,8	8,5	63,9	16,9	14,5	12,7	36,1	2,0	1,4	1,2
5	10,7	9,4	13,4	11,4	10,0	54,3	20,2	17,6	15,7	45,7	5,3	4,0	3,4
6	11,7	10,4	14,9	12,6	11,2	44,3	22,5	19,8	17,7	55,7	8,9	6,9	5,9
7	13,2	11,9	16,9	14,4	12,7	34,7	24,7	21,9	19,7	65,3	12,8	10,4	9,0
8	15,1	13,8	19,6	16,8	15,0	24,5	27,4	24,3	21,9	75,5	17,1	14,4	12,7
9	17,4	16,1	22,5	19,6	17,6	15,2	30,7	27,2	24,6	84,8	21,1	18,2	16,4
10	27,9	24,7	29,3	25,9	23,5	11,7	37,4	33,7	30,2	88,3	28,2	24,8	22,6
Top 1%	38,2	32,7	40,5	37,0	33,3	17,0	43,1	39,9	35,9	83,0	40,0	36,4	32,7

Anmerkung. „AGI“ (aggregated gross income) bei Bach et al. (2006) ist nahezu äquivalent mit Brutto2. N in % gibt die relative Fallzahl der jeweiligen Gruppe im betrachteten Dezil an.

Einkommenshöhe voll erfasst, allerdings nimmt der Informationsgehalt aufgrund der vorgenommenen Anonymisierung mit steigendem Einkommen ab. Unsere Ergebnisse für das oberste Einkommensperzentil zeigen jedoch, dass die Anonymisierung auf Perzentilebene keine wesentliche Hürde darstellt. Trotz des eingeschränkten Informationsgehaltes in FAST kommen wir mit Ausnahme des ersten Perzentils zu vergleichbaren Ergebnissen wie Bach et al. (2006) mit nicht-anonymisierten Daten.

6 Fazit

Nicht nur vor dem Hintergrund der derzeitigen Diskussion über die Neujustierung des Wohlfahrtsstaats mit den damit einhergehenden Eingriffen in die Systeme der sozialen Sicherung und der Besteuerung sind Informationen zur Einkommensverteilung und zur Wirkung staatlicher Umverteilungsmaßnahmen von zentralem Interesse. Die Faktisch Anonymisierte Lohn- und Einkommensteuerstatistik, die als erste Datenbasis repräsentative Informationen sowohl über die Top-Einkommensbezieher als auch über den Besteuerungsprozess der Zensenten enthält, erscheint als ideale Grundlage zur Generierung derartiger Informationen.

Allerdings sind in FAST einkommensteuerrechtlich definierte Einkommensaggregate verzeichnet, die die Konsummöglichkeiten der Zensenten – und damit deren Lebensstandard – in der Regel nicht adäquat beschreiben: Wie gezeigt, weisen Engelkurven, die auf einkommensteuerrechtlich definierten Einkommensaggregaten basieren, atypische Verläufe für klassischerweise als normal oder superior identifizierte Güter, wie z. B. „Spenden“ oder „Wohnen“, auf. Folglich sind auch Verteilungsanalysen, die diese Aggregate nutzen, fehlerbehaftet. Problematisch im Verteilungskontext ist insbesondere, dass sich die steuerrechtlich bedingten Verzerrungen nicht auf alle Steuerpflichtigen in gleicher Art und Weise auswirken.

Es erscheint daher sinnvoll, diese einkommensteuerrechtlichen Effekte zu eliminieren. Unsere Analyse zeigt, dass in FAST umfangreiche Informationen zur Konstruktion von (ökonomischen) Einkommen vorliegen, die es erlauben, ein exakteres Bild vom Lebensstandard der Zensenten zu entwerfen. Allerdings bleiben aufgrund der mit der Anonymisierung einhergehenden Informationsseinbußen insbesondere an den beiden Enden der Verteilung Verzerrungen bestehen. Zur Beseitigung derselben stehen jedoch prinzipiell Informationen aus zusätzlichen Statistiken zur Verfügung. Insbesondere erscheint es sinnvoll, eine Imputation der in dieser Studie unberücksichtigt gebliebenen Sozialtransfers vorzunehmen.

Dass Verteilungsanalysen auf Basis steuerrechtlich definierter Einkommen mit quantitativ bedeutsamen Fehleinschätzungen behaftet sein können, zeigen wir exemplarisch anhand der Progressionswirkungen der Einkommensteuer: Die effektive Progression ist wesentlich geringer als es der im Einkommensteuergesetz definierte Steuertarif zunächst vermuten ließe. Bei der Interpretation der Ergebnisse ist zu beachten, dass diese wegen der nicht eindeutigen Identifizierbarkeit zusammenlebender Haushaltsgemeinschaften fehlerbehaftet sind. Dies gilt auch aufgrund der mangelnden Repräsentativität der Stichprobe: Bevölkerungsgruppen, die der Einkommensbeteuerung nicht unterliegen, sind nicht erfasst. Dennoch stellt FAST aufgrund der detaillierten Erfassung von Hocheinkommensbeziehern ein sinnvolles Komplement zu bereits verfügbaren Statistiken (EVS, SOEP) dar.

Literaturverzeichnis

- ACEMOGLU, D. UND J.-S. PISCHKE (2001): Changes in the Wage Structure, Family Income, and Children's education, *European Economic Review*, 45, 890–904.
- AMMERMÜLLER, A., WEBER, A. M. UND P. WESTERHEIDE (2005): Die Entwicklung und Verteilung des Vermögens privater Haushalte unter besonderer Berücksichtigung des Produktivvermögens. Abschlussbericht zum Forschungsauftrag des BMGS, Zentrum für Europäische Wirtschaftsforschung, Mannheim.
- BACH, S. UND B. BARTHOLMAI (2000): Möglichkeiten zur Modellierung hoher Einkommen auf Grundlage der Einkommensteuerstatistik, *DIW Diskussionspapiere*, 212.
- BACH, S., CORNEO, C. UND V. STEINER (2006): Top Incomes and Top Taxes in Germany, *CESIFO Working Paper*, 1641.
- BACH, S., BUSLEI, H., RUDOLPH, H.J., SCHULZ, E. UND D. SVINDLAND (2004): Aufkommens- und Belastungswirkungen der Lohn- und Einkommensteuer 2003 bis 2005, *Materialien des DIW*, 38.
- BORK, C. UND H.G. PETERSEN (1997): Ein Vergleich möglicher Datensätze zur Eignung für steuerpolitische Simulationsrechnungen, *Finanzwissenschaftliche Diskussionsbeiträge der Universität Potsdam*, 20.
- BUNDESBANK (2006): Zeitreihendatenbank, http://www.bundesbank.de/statistik/statistik_zeitreihen.php.
- BUNDESMINISTERIUM DER FINANZEN (2005): Datensammlung zur Steuerpolitik, Berlin.
- BUNDESREGIERUNG (2005): Lebenslagen in Deutschland – Der zweite Armuts- und Reichtumsbericht der Bundesregierung, Bonn.
- BUNDESREGIERUNG (2001): Lebenslagen in Deutschland – Der erste Armuts- und Reichtumsbericht der Bundesregierung, Bonn.
- CLOTFELTER, C.T. (2002): The Economics of Giving, *Duke University Working Paper*, 97(19).
- DRABINSKI, T. UND C. SCHRÖDER (2003): Ein statisches Mikro-Simulationsmodell der sozialen Sicherung nach deutschem Recht, *Arbeitspapier des Lorenz-von-Stein-Instituts für Verwaltungswissenschaften*, 2. Auflage, Kiel.
- DUNCOMBE W. UND J. YINGER (1998): School Finance Reform: Aid Formulas and Equity Objectives. *National Tax Journal*, 51, 2239–62.
- EUROPEAN CENTRAL BANK (2003): Structural Factors in the EU Housing Markets, Frankfurt.
- FERRIS, J.S. UND E.G. WEST (2003): Private versus Public Charity: Reassessing Crowding out from the Supply Side. *Public Choice*, 116, 399–417.
- FÖRSTER, M. UND M. MIRA D'ERCOLE (2005): Income Distribution and Poverty in OECD Countries in the Second Half of the 1990s, *OECD Social Employment and Migration Working Papers*, 22.
- LANG, O., NÖHRBACH, K.-H. UND K. STAHL (1997): On Income Tax Avoidance: The Case of Germany, *Journal of Public Economics*, 66, 327–347.
- MANZOOR, S.H. UND J.D. STRAUB (2005): The Robustness of Kingma's Crowd-out Estimate: Evidence from Data on Contributions to Public Radio. *Public Choice*, 123, 463–476.
- MEEN, G. (1996): Ten Propositions in UK Housing Macroeconomics: An Overview of the 1980s and Early 1990s, *Urban Studies*, 33(3), 425–444.
- MERZ, J. (2001): Hohe Einkommen, ihre Struktur und Verteilung, Forschungsprojekt über Lebenslagen in Deutschland und Teil des Ersten Armuts- und Reichtumsbericht der Bundesregierung, Bundesministerium für Arbeit und Sozialordnung, Bonn.
- MERZ, J., HIRSCHEL, D. UND M. ZWICK (2005): Struktur und Verteilung hoher Einkommen – Mikroanalysen auf der Basis der Einkommensteuerstatistik, Beitrag zum zweiten Armuts- und Reichtumsbericht 2004 der Bundesregierung.
- MERZ, J. UND M. ZWICK (2004): Hohe Einkommen – Eine Verteilungsanalyse für Freie Berufe, Unternehmer und abhängig Beschäftigte. Eine Mikroanalyse auf Basis der Einkommensteuerstatistik, *Statistik und Wissenschaft*, 1, Statistisches Bundesamt, Wiesbaden, 167–193.
- MERZ, J. UND M. ZWICK (2001): über die Analyse hoher Einkommen mit der Einkommensteuerstatistik. Eine methodische Erläuterung zum Gutachten „Hohe Einkommen, ihre Struktur und Verteilung“ zum ersten Armuts- und Reichtumsbericht der Bundesregierung, *Wirtschaft und Statistik*, 7, 513–523.
- MICKLEWRIGHT J. UND A. WRIGHT, (2004): Private Donations for International Development. *CEPR Discussion Papers*, 4292.
- STAUDER, J. UND W. HÜNING (2004): Die Messung von Äquivalenzeinkommen und Armutsquoten auf der Basis des Mikrozensus, *Statistische Analysen und Studien Nordrhein Westfalen*, 13, 9–32.

VORGRIMLER, D. UND M. ZWICK (2004): Faktische Anonymisierung der Steuerstatistik (FAST) – Lohn- und Einkommensteuer 1998 –, Forschungsdatenzentrum des Statistischen Bundesamtes, Wiesbaden.

ZWICK, M. (2004): Integrierte Mikrodatenfiles, Statistik und Wissenschaft, 1, Statistisches Bundesamt, Wiesbaden, 287–294.

ZWICK, M. (1998): Einzeldatenmaterial und Stichproben innerhalb der Steuerstatistik, Statistisches Bundesamt, Wirtschaft und Statistik, 7.

Appendix

Variable	Koeffizient	SE	T	p
Konstante	441,43	1,93	228,57	0,00
Single, 1 Kind, Westdeutschland	-169,72	5,12	-33,18	0,00
Single, 2 Kinder, Westdeutschland	-188,14	7,43	-25,31	0,00
Single, 3 Kinder, Westdeutschland	-237,61	15,44	-15,39	0,00
Single, 4 und mehr Kinder, Westdeutschland	-374,57	33,85	-11,06	0,00
Paar, kinderlos, Westdeutschland	750,92	2,38	315,44	0,00
Paar, 1 Kind, Westdeutschland	299,53	3,01	99,46	0,00
Paar, 2 Kinder, Westdeutschland	244,36	2,99	81,71	0,00
Paar, 3 Kinder, Westdeutschland	202,68	4,80	42,25	0,00
Paar, 4 und mehr Kinder, Westdeutschland	105,75	8,67	12,20	0,00
Single, kinderlos, Ostdeutschland	-136,83	4,90	-27,91	0,00
Single, 1 Kind, Ostdeutschland	-271,33	8,16	-33,24	0,00
Single, 2 Kinder, Ostdeutschland	-283,02	11,64	-24,32	0,00
Single, 3 Kinder, Ostdeutschland	-318,55	27,61	-11,54	0,00
Single, 4 und mehr Kinder, Ostdeutschland	-390,76	67,79	-5,76	0,00
Paar, kinderlos, Ostdeutschland	335,51	4,52	74,30	0,00
Paar, 1 Kind, Ostdeutschland	30,68	5,45	5,63	0,00
Paar, 2 Kinder, Ostdeutschland	-51,90	5,57	-9,32	0,00
Paar, 3 Kinder, Ostdeutschland	-31,84	11,68	-2,73	0,01
Paar, 4 und mehr Kinder, Ostdeutschland	0,00	-1,20	0,00	1,00
Summe der Einkünfte (positiv)	0,01	0,00	206,45	0,00
Quadrierte Summe der Einkünfte (positiv)	0,00	0,00	-108,27	0,00
Summe der Einkünfte (negativ)	-0,01	0,00	-32,66	0,00
Quadrierte Summe der Einkünfte (negativ)	0,00	0,00	-17,87	0,00
Dummy Überwiegende Einkunftsart Gewinneinkünfte	374,97	3,40	110,31	0,00
Dummy Summe der Einkünfte negativ	-35,24	7,59	-4,65	0,00
Dummy überwiegende Einkunftsart Überschusseinkünfte	909,23	3,16	287,95	0,00
Interaktion (Summe der Einkünfte (negativ) x Gewinneinkünfte)	0,00	0,00	12,04	0,00
Interaktion (Summe der Einkünfte (negativ) x Überschusseinkünfte)	0,01	0,00	27,31	0,00
Interaktion (Summe der Einkünfte (positiv) x Gewinneinkünfte)	0,00	0,00	-70,49	0,00
Interaktion (Summe der Einkünfte (positiv) x Überschusseinkünfte)	0,00	0,00	-20,68	0,00

Variable	Koeffizient	SE	T	p
Land- und Forstwirtschaft, Mann	244,70	11,43	21,41	0,00
Land- und Forstwirtschaft, Frau	320,11	27,44	11,67	0,00
Bergbau und Verarbeitung, Mann	81,19	7,51	10,81	0,00
Bergbau und Verarbeitung, Frau	97,25	16,68	5,83	0,00
Energie, Wasser, Bau, Mann	0,00	1,11	0,00	0,94
Energie, Wasser, Bau, Frau	-173,75	31,13	-5,58	0,00
Handel, Mann	0,00	1,55	0,00	0,90
Handel, Frau	-44,53	7,76	-5,74	0,00
Gastgewerbe, Mann	-161,98	10,89	-14,88	0,00
Gastgewerbe, Frau	-227,40	14,84	-15,33	0,00
Verkehr- und Nachrichtenübermittlung, Mann	-185,87	12,92	-14,39	0,00
Verkehr- und Nachrichtenübermittlung, Frau	0,00	-0,83	0,00	1,00
Kredit und Versicherung, Mann	-108,46	9,35	-11,60	0,00
Kredit und Versicherung, Frau	-127,97	17,71	-7,23	0,00
Grund- und Wohnungswirtschaft, Mann	179,70	4,56	39,37	0,00
Grund- und Wohnungswirtschaft, Frau	131,21	6,74	19,46	0,00
Öffentliche Verwaltung, Mann	82,87	12,25	6,77	0,00
Öffentliche Verwaltung, Frau	48,11	17,72	2,71	0,01
Gesundheit, Veterinär und Gemeinwesen, Mann	120,56	7,88	15,29	0,00
Gesundheit, Veterinär und Gemeinwesen, Frau	21,32	10,19	2,09	0,04
Sonstige Dienstleistungen, Mann	-26,08	6,91	-3,78	0,00
Sonstige Dienstleistungen, Frau	-57,94	9,06	-6,40	0,00
Freie Berufe männlich (Dummy)	217,90	4,81	45,29	0,00
Freie Berufe weiblich (Dummy)	131,82	6,40	20,60	0,00
\overline{R}^2	0,12			
F-Statistik	7940,44			
<i>Anmerkung.</i> Schrittweise OLS-Schätzung mit Konstanter (Ausschluss einer Variablen falls $p>10\%$). Abhängige Variable sind die Kapitaleinkünfte unterhalb des Sparerfreibetrags. Berücksichtigt sind alle Zensenten, bei denen das Merkmal Kapitaleinkünfte besetzt war und den Sparerfreibetrag nicht überstieg. T ist die T-Statistik für einen 2-seitigen T-Test. SE ist der Standardfehler. Anzahl der Beobachtungen (gewichtet): 2.783.322.				
<i>Quelle:</i> Eigene Darstellung.				