



Kumulation von Querschnitten

Merz, Joachim; Stolze, Henning

Publication date:
2010

Document Version
Verlags-PDF (auch: Version of Record)

[Link to publication](#)

Citation for published version (APA):

Merz, J., & Stolze, H. (2010). *Kumulation von Querschnitten*. (FFB-Discussionpaper; Nr. 85). Forschungsinstitut Freie Berufe.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Kumulation von Querschnitten -

**Evaluierung alternativer Konzepte für die kumulierten
laufenden Wirtschaftsrechnungen 1999 bis 2003 im Vergleich
zur Einkommens- und Verbrauchsstichprobe 2003**

Joachim Merz und Henning Stolze

FFB-Diskussionspapier Nr. 85

Juli 2010



Fakultät II – Wirtschafts-, Verhaltens- und Rechtswissenschaften

Postanschrift:
Forschungsinstitut Freie Berufe
Postfach 2440
21314 Lüneburg

ffb@leuphana.de
<http://leuphana.de/ffb>
Tel.: +49 4131 677-2051
Fax: +49 4131 677-2059

**Kumulation von Querschnitten -
Evaluierung alternativer Konzepte für die kumulierten laufenden
Wirtschaftsrechnungen 1999 bis 2003 im Vergleich zur Einkommens- und
Verbrauchsstichprobe 2003**

Joachim Merz und Henning Stolze¹

FFB-Diskussionpapier Nr. 85

Juni 2010
ISSN 0942-2595

¹ Univ.-Prof. Dr. Joachim Merz, Leuphana Universität Lüneburg, Fakultät II: Wirtschafts,- Verhaltens- und Rechtswissenschaften, Forschungsinstitut Freie Berufe (FFB), 21332 Lüneburg, Professur ‚Statistik und Freie Berufe‘, Tel: 04131/677-2051, Fax: 04131/677-2059, e-mail: merz@uni.leuphana.de, www.leuphana.de/ffb

Dr. Henning Stolze, Forschungsinstitut Freie Berufe (FFB), Professur ‚Statistik und Freie Berufe‘, Wege&Gehege serverbased computing, e-mail: stolze@wegeundgehege.de

Besonders danken möchten wir Frau Brigitte Demant vom Statistischen Bundesamt für ihre aufwändigen Arbeiten bei der Erstellung der Auswahlfiles aller laufenden Wirtschaftsrechnungen und der EVS sowie der Mikrozensusgruppe des Statistischen Bundesamtes für die Mikrozensus-Sonderauswertungen.

**Kumulation von Querschnitten -
Evaluierung alternativer Konzepte für die kumulierten laufenden
Wirtschaftsrechnungen 1999 bis 2003 im Vergleich zur Einkommens- und
Verbrauchsstichprobe 2003****Joachim Merz und Henning Stolze**

FFB-Diskussionspapier Nr. 85, Juni 2010, ISSN 0942-2595

Zusammenfassung

Im Rahmen der Weiterentwicklung des Systems von Haushaltsstatistiken und unter Berücksichtigung der EU-Erfordernisse mit neuen EU-SILC Ansätzen wird die Kumulation von Querschnitts-Stichproben zu einem integrierten Informationssystem diskutiert und gefordert. Insbesondere sollen bei dem Umbau der Haushaltsbudgeterhebungen einerseits jährliche Ergebnisse produziert werden, sowie andererseits in mehrjährigem Abstand ausreichend hohe Stichprobenumfänge zur Verfügung stehen, um tiefgegliederte Analysen zu ermöglichen. Die vorliegende Studie ist dazu ein Beitrag mit einer generellen konzeptionellen Fundierung für die Kumulation von Querschnitten und einer Anwendung, die im Besonderen das neue Kumulationskonzept evaluiert mit großen offiziellen Stichproben/Umfragen: der Querschnitts-Kumulation von fünf jährlichen Laufenden Wirtschaftsrechnungen (LWR), die verglichen wird mit der fünfjährigen Einkommens- und Verbrauchsstichprobe (EVS) des deutschen Statistischen Bundesamtes. Dabei wird die Sensitivität des Kumulationskonzeptes mit seinen Alternativen evaluiert für ausgewählte Gruppen von Konsumausgaben privater Haushalte. Eine Empfehlung beschließt diese Studie.

JEL: C42, C81, D10, E20

***Keywords:** cumulation of cross sections, temporary cumulation, adjustment by information theory, consumption expenditures, Continuous Household Budget Surveys (Laufende Wirtschaftsrechnungen, LWR), Sample Survey of Income and Expenditures (Einkommens- und Verbrauchsstichprobe, EVS) of the German Federal Statistical Office*

Abstract

With the development of household budget systems and with regard to the requirements of the European Union with new EU-SILC approaches, the cumulation of cross-section surveys to an integrated information system is recently discussed and required. In particular the reconstruction of household budget surveys should deliver yearly results as well multi-annual sufficient large samples to allow in depth analyses. This study contributes by a general conceptual foundation of the cumulation of cross-sections and an application which in particular evaluates the new cumulation concept with actual large official samples: the cross sectional cumulation of five yearly Continuous Household Budget Surveys (Laufende Wirtschaftsrechnungen, LWR) which will be compared to the large quinquennial Sample Survey of Income and Expenditures (Einkommens- und Verbrauchsstichprobe, EVS) of the German Federal Statistical Office. Therewith the sensitivity of the cumulation concept with its alternatives is evaluated for private household consumption expenditures of selected expenditure groups. A recommendation concludes.

JEL: C42, C81, D10, E20

***Keywords:** cumulation of cross sections, temporary cumulation, adjustment by information theory, consumption expenditures, Continuous Household Budget Surveys (Laufende Wirtschaftsrechnungen, LWR), Sample Survey of Income and Expenditures (Einkommens- und Verbrauchsstichprobe, EVS) of the German Federal Statistical Office*

1 Einleitung

Im Rahmen der Weiterentwicklung des Systems von Haushaltsstatistiken und unter Berücksichtigung von EU-Erfordernissen wird die Kumulation von Stichproben zu einem integrierten Informationssystem diskutiert. Die bisherigen parallelen und nicht verbundenen Statistiken sollen geeignet zusammengeführt werden, um in einem integrierten System der Haushaltsstatistiken auch komplexe Fragestellungen bearbeitet zu können. Damit sollen flexible, zeitnahe und neue Datenanforderungen der interessierten Öffentlichkeit ermöglicht werden (Ehling 2002a).

Insbesondere sollen bei dem Umbau der Haushaltsbudgeterhebungen einerseits jährliche Ergebnisse produziert werden, sowie andererseits in mehrjährigem Abstand ausreichend hohe Stichprobenumfänge zur Verfügung stehen, um tiefgegliederte Analysen zu ermöglichen (Ehling 2002b, S. 22). Die Vor- und Nachteile einer präferierten Kumulation von Stichproben gegenüber alternierenden Stichproben haben beispielsweise Ehling (2002b, S. 24) oder Verma (2002, S. 51-52) in dem Tagungsband zu rotierenden Stichproben (Statistisches Bundesamt 2002) diskutiert.

In der vorliegenden Studie werden – aufbauend auf den Ausführungen des Autors zu den konzeptionellen Grundlagen einer Kumulation von Stichproben generell sowie einem Vorschlag zu ihrer Realisierung (Merz 2004) – alternative Vorgehensweisen für die Realisierung des Kumulationsvorschlags entwickelt und anhand aktueller amtlicher Stichproben evaluiert. Das Kumulationskonzept, zunächst allgemeiner diskutiert, wird hier konkret auf laufende Wirtschaftsrechnungen des Statistischen Bundesamtes mit einer Kumulation von jährlichen Haushaltsbudgetbefragungen angewendet. Mit den Stichproben der Wirtschaftsrechnungen 1999, 2000, 2001, 2002 und 2003 werden alternative Kumulationsvarianten über die einzelnen Jahresfiles simuliert und in ihren Auswirkungen auf ein entsprechend kumuliertes File analysiert und evaluiert. Da die Daten der laufenden Wirtschaftsrechnungen aus dem Jahr 2000 gegenüber den anderen Jahren ungewöhnlich hohe Varianzen aufweisen, haben wir allerdings das Jahr 2000 schließlich nicht in die Analyse einbeziehen können.

Mit den entwickelten alternativen Kumulationsvarianten analysieren wir die Wirkungen auf die Abschätzung des privaten Verbrauchs ausgewählter Gütergruppen und vergleichen die Ergebnisse der alternativen Kumulationsfiles mit einer anderen Großstichprobe, der Einkommens- und Verbrauchsstichprobe 2003. Damit wird die Sensitivität der Kumulationskonzepte hinsichtlich eines breiten Spektrums des Ausgabeverhaltens empirisch fundiert aufgezeigt und beurteilt. Eine Empfehlung schließt die Studie ab.

2 Ein Konzept für die Kumulation von jährlichen Haushaltsbudgetbefragungen

Auf der Basis allgemeiner theoretischer Ansätze wurde in Merz 2004 ein konkretes Kumulationskonzept für Haushaltsbudgetbefragungen entwickelt und der interessierten Öffentlichkeit zur Diskussion vorgestellt. Dieses Konzept – und seine Weiterentwicklung – wird zunächst in seinen grundlegenden Zügen zusammengefasst, wobei Weiterentwicklungen kursiv kenntlich gemacht sind. Die folgenden Abschnitte vertiefen dann die zentralen

Bausteine und Kumulationsalternativen, die dann schließlich in unsere Simulation und Evaluation eingehen.

Kumulationskonzept und Aufgaben:

- (1) **Inflationierung der ökonomischen Größen (Ausgaben, Einnahmen) aller Querschnitte auf $t=T$:** Mit geeigneten Preisindizes („economic multiplier“) werden alle ökonomischen (im Sinne von monetären) Größen auf das Jahr $t=T$ inflationiert und alle Ausgaben und Einnahmen (eventuell preisgruppenspezifisch) umgerechnet. Im Gegensatz zu den demographischen Gewichten, die abhängig sind von der soziodemographischen Struktur des jeweiligen Haushalts eines Querschnitts, ist ein „economic multiplier“ (Inflator/Deflator) unabhängig von den einzelnen Erhebungseinheiten (Haushalten).
- (2) **Demographische Strukturbestimmung und Rahmendaten:** Für den aktuellsten Querschnitt $t=T$, der in unserem Fall zugleich mit dem Jahr einer EVS als Vergleichsstichprobe übereinstimmen sollte, sowie für die zeitlich davor liegenden Querschnitte (hier laufende Wirtschaftsrechnungen) ist zu entscheiden, welche demographische Struktur sowohl auf der Haushalts- als auch auf der Personenebene für eine demographisch repräsentative Hochrechnung im Einzelnen heranzuziehen ist.² Die demographischen Rahmendaten der gewählten Anpassungsstruktur sollten aus den verschiedenen Mikrozensus als breitester Bevölkerungsstichprobe gewonnen werden.
- (3) **Kumulationsgewichtung:** Ziel einer Kumulationsgewichtung ist es, die Information aller älteren Stichproben einzubeziehen. Damit die unterschiedliche zeitliche Nähe/Aktualität und damit der unterschiedliche Informationsgehalt der Stichproben/Querschnitte bei der Kumulation berücksichtigt wird, werden wir verschiedene Alternativen zur Bestimmung geeigneter Abschreibungsraten, im Weiteren als Kumulationsgewichte (w_t , $t = 1, \dots, T$) bezeichnet, für die T Querschnitte vorschlagen. Dabei werden sowohl „gesetzte“ als auch durch die Daten selbst bestimmte Gewichte auf der Basis einer Clusteranalyse einbezogen.
- (4) **Neue Hochrechnungen für die kumulierte Stichprobe KUM zu $t=T$:** Mit den neuen Rahmendaten $r(t=T)$ sind die zuvor additiv zusammengeführten und bisherigen querschnittsbezogenen inflationierten Querschnitte t ($t=1, \dots, T$) – mit eventuell neu abgestimmten Hochrechnungsfaktoren – als neues File KUM theoriegeleitet, simultan und konsistent neu hochzurechnen.³ Mit der dafür gewählten Hochrechnung nach dem Prinzip des minimalen Informationsverlustes (realisiert mit dem Programmpaket ADJUST von Merz und Stolze 2004) gehen bereits vorhandene Hochrechnungsfaktoren in die informationstheoretisch fundierte Zielfunktion ein; die Zielfunktion verbindet die bisherigen Informationen der jeweiligen Gewichte minimierend mit den neu zu erstellenden Hochrechnungsfaktoren. Diese Vorgehensweise berücksichtigt damit eine bereits vorgenommene oder vorgegebene

² Beispielsweise gegliedert nach beruflicher Stellung des HH-Vorstands (HHV), Alter des HHV, Haushaltsstruktur: Haushaltsgröße, Anzahl der Erwerbstätigen, Anzahl der Kinder in Altersklassen etc. als auch die Personenstruktur (z.B. nach Altersklassen und Geschlecht).

³ Die additiv zusammengeführten Querschnitte erlauben eine Itembezogene Relativierung: Ist beispielsweise für Einkommensverteilungsanalysen eine Auswertung von relativierten Einkommen (z.B. relativiert auf das jeweilige Querschnittsmittel) vorgesehen, so kann dies mit den vorhandenen Informationen – beispielsweise entweder mit den originalen Querschnittsgewichten der jeweiligen Periode t oder mit den Gewichten der kumulierten Stichprobe KUM – erfolgen.

temporäre Repräsentativität über die entsprechenden Hochrechnungsfaktoren (Ausgleichsgewichte) und Informationen aus den einzelnen Querschnitten.

Prinzipiell alternative Hochrechnungen:

- a) Die Kumulationsgewichte w_i werden zuerst mit den *vorhandenen* Hochrechnungsfaktoren der einzelnen Querschnitte multipliziert. Hierbei kann als *vorhandenes* Gewicht auf die ursprünglichen Gewichte q_i der Querschnittsdaten, oder auf Hochrechnungsfaktoren p^* , die zuvor zur Anpassung der Querschnittsdaten an die Rahmen der jeweiligen Jahrgänge r_i gefunden wurden, zurückgegriffen werden. Danach wird die gesamte kumulierte Stichprobe KUM neu hochgerechnet und an die Rahmendaten zu T angepasst.
- b) Es werden erst die einzelnen Querschnitte aus der kumulierten Stichprobe KUM mit den Rahmendaten zu T neu hochgerechnet. Da jeder hochgerechnete Querschnitt die Population N_T ergibt, besteht die kumulierte Stichprobe dann aus TN_T Einheiten. Die Hochrechnungsfaktoren werden danach mit den Kumulationsgewichten w_i multipliziert. Die Kumulationsgewichtung sollte sich auf 1 summieren, damit die kumulierte Gesamtstichprobe durch die Kumulationsgewichte wieder den Umfang N ergibt.

Die zweite Hochrechnungsvariante (b) – Kumulationsgewichtung *nach* demographischer Hochrechnung – ist flexibler und erlaubt alternative Gewichtungen für den Nutzer/die Nutzerin auch später, ohne dass alle Querschnitte mit neuen Kumulationsgewichten jeweils neu hochgerechnet werden müssen.

- (5) **Modellgestützte Fortschreibung:** Wird eine modellgestützte Fortschreibung mit entsprechenden mikroökonomischen Schätzungen gewählt, dann ist diese Fortschreibung wie folgt mit der Hochrechnung verbunden:
 - Sind die inhaltlich interessanten Variablen (z.B. seltene Ereignisse) **unabhängig** von der demographischen Hochrechnung, kann die modellgestützte Fortschreibung auch nach der Hochrechnung erfolgen.
 - Sind diese inhaltlich interessanten Variablen aber **abhängig** von der demographischen Hochrechnung, dann ist die modellgestützte Fortschreibung direkt bei der Hochrechnung als ein weiteres Hochrechnungsmerkmal zu berücksichtigen.
- (6) **Sensitivitätsanalysen und Evaluation KUM vs. EVS:** Für eine Sensitivitätsanalyse und Evaluation der Kumulationsalternativen ist schließlich der Vergleich der Ergebnisse der alternativ gewonnenen kumulierten Stichprobe (KUM) mit einer Großstichprobe, hier mit der EVS, zum Zeitpunkt $t=T$ durchzuführen.

3 Konkretisierung des Kumulationskonzepts

Das obige zusammengefasste Kumulationskonzept basiert auf den vier zentralen Bausteinen

- Inflationierung der ökonomischen Größen (Ausgaben, Einnahmen)
- Alternative Kumulationsgewichtung
- Modellgestützte Fortschreibung

- Neue Hochrechnung der kumulierten Stichprobe(n)

auf die wir im Folgenden vertiefend eingehen wollen.

3.1 Inflationierung der ökonomischen Größen

Die Inflationierung ökonomischer Größen – hier die Ausgaben und Einnahmen privater Haushalte – berücksichtigt die Preisentwicklung durch geeignete Preisindizes. Ein solcher Preisindex passt unterschiedliche Preisniveaus durch eine multiplikative Verknüpfung einer monetären Größe mit einem Preisindex an. Der Preisindex („economic multiplier“) ist dabei – wenn nicht regional verschieden – für alle Haushalte gleich. Die Inflationierung ökonomischer Größen ist somit rechentechnisch unproblematisch.

Je nach Disaggregationsgrad wird ein allgemeiner Preisindex – wie z.B. der Preisindex für die private Lebenshaltung – oder es werden ausgabengruppenspezifische Preisindizes berücksichtigt.

3.2 Alternative Kumulationsgewichtung

Die von uns betrachtete temporäre Kumulation kombiniert alle T vorhandenen Querschnitte, hier die der laufenden jährlichen Wirtschaftsrechnungen. Diese jährlichen Querschnitte t ($t=1, \dots, T-1$) liegen gegenüber T somit um $T-t$ Jahre zurück. Damit haben wir es mit „veralterter“ Information im Vergleich zur aktuellen Situation in T zu tun.

Ziel einer Kumulationsgewichtung ist es, die Information aller Stichproben, also auch die der älteren Stichproben angemessen mit geeigneten Abschreibungsraten zu berücksichtigen. Die Gewichte, mit denen wir die einzelnen zu kumulierenden Stichproben (synonym: Files, Querschnitte, Erhebungen, hier: Jahrgänge der laufenden Wirtschaftsrechnungen) in der dann kumulierten Stichprobe berücksichtigen, werden im Weiteren als **Kumulationsgewichte** w_t bezeichnet. Sie sind nicht zu verwechseln mit den Gewichten aus einer Hochrechnung (Hochrechnungsfaktoren), die eine demographische Repräsentativität gewährleisten sollen.

Vier Ansätze für eine Bestimmung der Kumulationsgewichte sollen kurz charakterisiert werden:

- Ansätze aus der Informatik
- Informationstheoretischer Ansatz
- Alternative Distanzmessungen: Gewichtung generell mit Ähnlichkeitsmaßen
- Alternativ gesetzte temporäre Gewichtung.

Diese Ansätze werden im Folgekapitel verknüpft und gehen dann in die eigentlichen Simulationsalternativen ein.

3.2.1 Ansätze aus der Informatik: Informationswert von Datenbanken

Der Wert von Informationen in Datenbanken wird in der (Wirtschafts-)Informatik hinsichtlich deren Alterung und optimaler Aktualisierungsintervalle diskutiert.

Zum Beispiel kann aus einer Kundendatei ein bestimmter Nutzen gezogen werden, wenn der Datenbestand in Werbemaßnahmen Verwendung findet. Dieser Nutzen sinkt, wenn die Datei älter wird, und somit ein Teil der Adressen ungültig wird. Ein weiteres Beispiel sei aus der Informationsflussteuerung genommen: Die Zwischenspeicherung (das Cachen) von Netzwerkinformationen auf Servern. Die Übermittlung von Daten ist mit Kosten verbunden (Bandbreite, Rechenleistung, zugriffsbedingte Wartezeiten etc.); daher werden bestimmte Informationen zwischengespeichert. Sind diese jedoch falsch (da zu alt), verursacht diese Fehlinformation wiederum Kosten. Ab einem gewissen Zeitpunkt überschreitet das Risiko, durch veraltete Daten Kosten zu verursachen, die Chance, einen direkten Zugriff auf die benötigten Informationen zu bekommen und potenziell Kosten zu sparen.

Für die Bewertung des "Risikos" muss also ein Weg gefunden werden, die "Aktualität" der Daten zu ermitteln. Bei Adressdaten ist dies recht einfach: Man kann für ein paar Zeitpunkte die hinzugekommenen ungültigen Adressen erfassen und daraus eine Rate, die die Anzahl ungültig gewordener Adressen in einem bestimmten Zeitraum angibt, schätzen. Mit dieser Rate ließe sich die erwartete Anzahl ungültig gewordener Adressen zu jedem Zeitpunkt abschätzen beziehungsweise extrapolieren.

Die Idee einer Abschätzung des Risikos, veraltete Informationen zu verwenden, ist zentral für die temporäre Kumulation.

Übertragen auf unser Problem der temporären Kumulation wäre beispielsweise ein unterschiedliches Verbraucherverhalten aus den verschiedenen Jahren Grundlage einer Gewichtung. Man könnte daher aus Stichproben der Vergangenheit die Veränderungsrate des Verbraucherverhaltens über die Zeit schätzen, indem man beispielsweise über Distanzmaße (mehr dazu in 3.2.2 und 3.2.3) einen Ähnlichkeitsindex zu verschiedenen Zeitpunkten aufbaut oder ökonometrisch ein verändertes Verhalten (mehr dazu in 3.3) ermittelt. Hieraus könnte sich eine allgemeine, vermutlich exponentiell abnehmende, Abschreibungsrate in Abhängigkeit der Zeit $d(\Delta t)$ ergeben und diese für die verschiedenen Querschnitte der Kumulation Verwendung finden.

3.2.2 Informationstheoretischer Gewichtungsansatz

Wenn wir generell den Informationsaspekt weiter verfolgen, dann könnte der informationstheoretische Ansatz mit der Entropie als Maß einer Information weiterhelfen⁴.

Die **Entropie** oder der **Informationsgehalt** einer Menge von Objekten j ($j=1, \dots, n$) mit ihren anteiligem Vorkommen $\mathbf{p}=(p_1, \dots, p_n)'$, ($p_j > 0$), $\sum_j p_j = 1$, wird aus informationstheoretischer Sicht gemessen/charakterisiert durch

$$(1) \quad H(\mathbf{p}) = H(p_1, \dots, p_n) = \sum_j p_j \log(1/p_j).$$

Würden mit \mathbf{p} alle Variablen eines Querschnittes in seinen Ausprägungen gemessen, dann könnte die aggregierte Information dieses Querschnitts mit $H(\mathbf{p})$ eindimensional zusammengefasst werden.

Ein **Informationsverlust** (beziehungsweise Informationsgewinn) eines älteren Querschnitts - mit entsprechenden anteiligen Vorkommen $\mathbf{q}=(q_1, \dots, q_n)'$ - gegenüber der aktuellen Situation \mathbf{p} könnte gemessen/bewertet werden durch

4 Zum informationstheoretischen Hintergrund vgl. z.B. Golan, Judge und. Miller (1996).

$$(2) \quad I(\mathbf{p}:\mathbf{q}) = \sum_j p_j \log(1/q_j) - \sum_j p_j \log(1/p_j) \\ = \sum_j p_j \log(p_j/q_j),$$

wobei $\mathbf{p} = (p_1, \dots, p_n)'$, $\mathbf{q} = (q_1, \dots, q_n)'$ mit $(p_j, q_j > 0)$, $\sum_j p_j = \sum_j q_j = 1$, $(j=1, \dots, n)$.

Dieser Ansatz entspricht auch unserem Konzept der Hochrechnung/Kalibrierung von Stichproben nach dem Prinzip des minimalen Informationsverlusts.

Damit ergäbe sich für jede zeitlich zurückliegende Stichprobe ein Entropiewert H_t , beziehungsweise ein Abstandsmaß I_t zur jetzigen Situation.

Eine entsprechende informationstheoretisch begründete temporäre Kumulationsgewichtung ließe sich für die Querschnitte t ($t=1, \dots, T$) konstruieren mit

$$(3) \quad w_t = I(p_T : q_t) / \sum_{i=1}^T I(p_T : q_i).$$

Der Querschnitt also, der informationstheoretisch am stärksten von der aktuellen Situation abweicht, erhält somit das stärkste – oder eventuell auch invers das schwächste – Gewicht in der kumulierten Stichprobe.

Zur Verwendung der Entropie zur Charakterisierung einer Stichprobe sei das Folgende noch angemerkt: Der Entropiebegriff misst den Informationsgehalt einer Objektmenge. Wenn die Entropie=1 ist, sind die Informationen rein zufällig verteilt, bei kleinen Entropiewerten treten Redundanzen oder statistische Regelmäßigkeiten auf. Die Zahl $H(I)$ gibt intuitiv die durchschnittliche Information an, die in einem Symbol (z.B. einem Datensatz) der Quelle enthalten ist. Damit ist die Entropie ein Maß für die Struktur der Daten.

Es wäre daher kritisch zu fragen, ob ein solches Ordnungs- und Strukturmaß von der Intention her und inhaltlich angemessen zur Gewichtung geeignet ist, wenn auf eine weitere sozioökonomische Verhaltensfundierung verzichtet wird. Allerdings könnte man die funktionale Form der Entropie und daraus folgend den Informationsverlust als ein generelles Distanzmaß interpretieren. Dabei wären allerdings die ursprünglich verwendeten Anteile p und q als allgemein metrische Ausprägungen (nicht nur als Anteile) der Variablen der Stichprobe weiter zu entwickeln.

3.2.3 Datengenerierte alternative Distanzmessungen: Proximitätsmaße

Die eben angesprochene Gewichtung bewertet in ihrer Weise den ‚Abstand‘ zweier Datenmengen. Hierzu gibt es eine Reihe von alternativen Distanzmessungen, die als Proximitätsmaße die Distanz zwischen ganzen Datensätzen (Stichproben) ermitteln.

Unterteilt entsprechend des zugrunde liegenden Skalenniveaus kommen hierfür beispielsweise in Frage:

- **Proximitätsmaße bei Nominal-Skalen**
Tanimoto-Koeffizient, M-Koeffizient, Kulczynski-Koeffizient, RR-Koeffizient, Dice-Koeffizient, Chi²-Koeffizient, ...
- **Proximitätsmaße bei metrischen Skalen**
L₁- und L₂-Norm, Q-Korrelationskoeffizient, Mahalanobis-Distanz, Minkowski-Metrik (mit Sonderfall der quadratischen euklidischen Distanz), Generalized Least

Squares, Minimum Information Loss, Raking Ratio, Minimum Entropy, Hellinger-Distanz, Modified Chi-Square, ...

Jede darauf aufbauende Gewichtung ist eine durch die Daten selbst generierte Gewichtung – ähnlich der informationstheoretisch fundierten Gewichtung – und verwendet einen entsprechend gemessenen Abstand zwischen allen Ausprägungen zweier oder mehrerer Stichproben.

Dem temporären Gewichtungsaspekt einer Kumulation wird dabei über eine geänderte Ausprägung der Variablen als geändertes Verhalten Rechnung getragen. Der datengenerierte Proximitätsansatz liefert Distanzen zwischen jedem Querschnitt zu t zur aktuellen Situation zu T . Eine größere Distanz weist auf eine größere Änderung im (Konsumenten-)Verhalten hin. Wir argumentieren, dass dann die Situation zum Zeitpunkt t für die aktuelle Situation T (die sich ja stark geändert hat) weniger von Interesse ist; die Situation zu t sollte wegen ihres Aktualitätsverlusts deshalb weniger stark gewichtet werden. Da nicht die Distanz sondern die Ähnlichkeit schließlich interessiert wird die proximitätsbasierte Kumulationsgewichtung invers konstruiert: je ähnlicher eine Sub-Stichprobe ist desto stärker wird sie im gesamten Kumulationsfile dann gewichtet.

Unser Anliegen ist es letztlich zu überprüfen, welche Auswirkungen alternative Kumulationsgewichtungen auf die Ausgaben privater Haushalte in einem kumulierten File haben. Basis für eine Ähnlichkeits-, beziehungsweise Distanzbestimmung zwischen den einzelnen Querschnitten werden somit Ausgaben für bestimmte Warengruppen wie Nahrungsmittel und Alkoholfreie Getränke oder für Andere Waren und Dienstleistungen etc. sein.

Da es sich hierbei ausschließlich um metrische Größen handelt, kommen verschiedene metrische Distanz- (z.B. Minkowskimetrik, Kosinus-Distanz oder Tschebyscheff-Distanz) sowie Ähnlichkeitsmaße (z.B. Q-Korrelation) in Frage.

Konkret wollen wir die aus der Varianzanalyse her bekannte und universal verwendbare quadrierte euklidische Distanz als Spezialfall der Minkowski-Metrik anwenden.

Für unseren Fall berechnen wir (prinzipiell) insgesamt vier Distanzen als Distanz einer laufenden Wirtschaftsrechnung (LWR) 1999 bis 2002 ($t=1, \dots, 4$) im Vergleich zur letzten verfügbaren LWR 2003 ($t=5=T$).

Da eine Distanzmatrix nur zwischen den verschiedenen Querschnittsjahrgängen und nicht zwischen den einzelnen Merkmalsträgern benötigt wird, stellt sich die Frage der Behandlung von Gruppen (Clustern) hinsichtlich ihres Zentrums. Parallel zu Fusionierungsalgorithmen der Clusteranalyse können bekannte Verfahren wie das Single- oder Complete-Linkage, das Centroid- oder das Ward-Verfahren angewendet werden. Könnte man praktische Erwägungen wie die Gruppengröße und Umsetzbarkeit mit vorhandenen Statistikprogrammen außer Acht lassen, wäre das Ward-Verfahren die optimale Wahl; es ist robust und bestimmt zuverlässig Clusterzentren und Distanzen zu anderen Clustern, ohne Probleme wie beispielsweise Kettenbildungen zu verursachen.

Allerdings stoßen hierarchische Verfahren bei Clustergrößen von 20.000 und mehr Fällen wie in den Auszügen der LWR durchaus an die Grenzen der Leistungsfähigkeit von Desktop-Rechnern. Zudem wäre die Programmierung einer eigenen Fusionsroutine nötig, da ob der vorgegebenen Gruppenzugehörigkeit durch die Jahrgänge nicht auf die implementierten Fusionierungsroutinen der Clusteranalysefunktionen der gängigen Statistikprogramme zurückgegriffen werden könnte.

Aus Gründen der Transparenz und technischen Umsetzbarkeit, kann und soll stattdessen bei den Clustern eine Distanzmessung zwischen den Jahrgängen auf der Basis der jeweiligen arithmetischen Mittelwerte der oben genannten Ausgabenvariablen erfolgen.

Schließlich müssen die berechneten Distanzen in geeignete Alterungsgewichte überführt werden, die die Forderung $\sum_t w_t = 1$ erfüllen. Hierzu wird der jeweilige Anteil an der Gesamtdistanz als Informationsverlust verstanden. Die auf diese Weise entstandenen Kumulationsgewichte werden zusätzlich, wie die Kumulationsgewichte auf der Grundlage anderer Methoden auch, auf eine Summe von 1 normiert.

Es ergibt sich dann ein:

Datengeneriertes clusteranalytisches Kumulationsgewicht

$$w_t = \frac{1 - d_{t,T} / \sum_t d_{t,T}}{\sum_t \left(1 - d_{t,T} / \sum_t d_{t,T} \right)}$$

wobei es sich bei $d_{t,T}$ hier um die quadrierte euklidische Distanz zwischen dem Jahr t und dem Bezugsjahr 2003 (T) handelt.

Bestimmung der datengenerierten clusteranalytischen Kumulationsgewichte für unsere Simulationsrechnungen

Für die datengenerierte Kumulationsgewichtung wird eine Clusteranalyse vorgenommen, bei der für die Bestimmung der Zentren Mittelwerte verwendet werden (Cluster-Zentrum-Analyse). Zur Gewinnung der Kumulationsgewichte der Basis eines allgemeinen Distanzmaßes sind für unsere späteren Simulationsrechnungen dann folgende Schritte notwendig:

- Aggregation der ausgabenbezogenen Erhebungsmerkmale aus den LWR 1999-2003 nach gewünschten zentralen Ausgabegruppen (hier 12 aggregierte Ausgabegruppen)
- Bestimmung der arithmetischen Mittelwerte der o.g. 12 Ausgabegruppen für alle vier Jahrgänge als Grundlage der Distanzmatrix
- Spezifische Preisbereinigung der Mittelwerte durch den jeweiligen Preisindex der 12 Basisausgabegruppen und des Erhebungsjahres
- Clusteranalyse und Berechnung der Distanzen der Jahrgänge '99 bis '02 zu jeweils '03 (quadrierte euklidische Distanz)
- Überführung der Distanzmatrix in Kumulationsgewichte.

Die schließlich doch aufwändigen konkreten Berechnungen auf der Basis der laufenden Wirtschaftsrechnungen von 1999 bis 2003 ergaben mit T=5=2003 folgende Gewichte:

Datengenerierte clusteranalytische Kumulationsgewichte

$$w_t = \{0,156; 0,177; 0,194; 0,224; 0,250\}.$$

Der Stichprobe Z zeitlich nähere Stichproben haben somit ein höheres Gewicht erhalten, weil sie der Stichprobe zu T ähnlicher sind. Allerdings ist nicht immer ein solches Resultat zu erwarten, obgleich die Daten der aktuelleren Stichproben den Daten der Stichprobe zu T meist ähnlicher sein dürften.

Diese Gewichtung werden wir dann mit den im Folgenden alternativ gesetzten Gewichten als Variante (d) in unseren Simulationsrechnungen verwenden.

3.2.4 Alternativ gesetzte Gewichtung

Unter dem Aspekt einer „gesetzten“ alternativen Gewichtung, die sozusagen ohne Berücksichtigung der Datenstruktur von außen eine „Informations-Abschreibung“ vornimmt, ist natürlich eine Vielzahl an Kumulationsgewichten denkbar.

Diese denkbare Vielzahl eingrenzend werden für die Stichproben zu den Zeitpunkten $t=1, \dots, T$, wobei T die zeitlich aktuellste Stichprobe bezeichnet, drei Gewichtungsalternativen vorgeschlagen und in den Simulationsrechnungen verwendet:

- **Gleiche Kumulationsgewichtung:** eine gleiche Gewichtung aller Stichproben, die auch ältere Stichproben so wie die aktuellste gleichgewichtig berücksichtigt:

$$w_t = 1/T, \quad (t = 1, \dots, T),$$

mit hier fünf aufeinander folgenden Stichproben ($T=5$) wäre $w_t = 1/5 = 0,20$ für alle t .

- **Linear progressive Gewichtung,** die die älteste Stichprobe am Geringsten und die neueren Stichproben mit proportional zunehmenden Kumulationsgewichten einbezieht:

$$w_t = t / \sum_{i=1}^T i, \quad (t = 1, \dots, T),$$

mit $T=5$ wäre $w_t = \{0,067; 0,133; 0,200; 0,267; 0,333\}$.

- **Exponentiell wachsende Gewichtung,** die zwar wie bei der proportionalen Gewichtung der aktuellsten Stichprobe das größte Gewicht einräumt, das aber mit einem noch höheren Gewicht. Für ein exponentielles Wachstum zur Basis 2 und Basis 3 ergibt sich:

Wachstum zur Basis 2

$$w_t = 2^{t-1} / \sum_{i=0}^{T-1} 2^i \quad (t = 1, \dots, T)$$

mit $T=5$ wäre $w_t = \{0,032; 0,065; 0,129; 0,258; 0,516\}$.

Natürlich verstärkt sich ein exponentielles Wachstum bei größerer Basis. Wir werden im Weiteren als Alternative c eine exponentiell wachsende Kumulationsgewichtung zur Basis 2 verwenden, da eine höhere Basis die Informationen der

ersten beiden Jahrgänge nur unzureichend würdigen würde. Mit einer Gewichtung zur Basis 3 (und höher) würden die enthaltenen Informationen der ersten Jahre praktisch verloren gehen, da beispielsweise der letzte Jahrgang über 80-mal stärker ins Gewicht fallen würde als die Daten des ersten Jahrgangs.

3.2.5 Alternative Kumulationsgewichte im Überblick

Zusammengefasst werden wir für die folgenden Simulationsrechnungen sowohl die gesetzten Kumulationsgewichte als auch durch die Daten selbst bestimmte datengenerierte Kumulationsgewichte auf der Basis einer Clusteranalyse einsetzen.

Mit vier aufeinanderfolgenden Stichproben ($t=1, \dots, T=5$) sind dies:

a) Gleiche Kumulationsgewichtung:

$$w_t = 1/T, \quad (t = 1, \dots, T),$$

$$w_t = \{0,20; 0,20; 0,20; 0,20; 0,20\}.$$

b) Linear progressive Gewichtung

$$w_t = t / \sum_{i=1}^T i, \quad (t = 1, \dots, T),$$

$$w_t = \{0,067; 0,133; 0,200; 0,267; 0,333\}$$

c) Exponentiell wachsende Gewichtung (Basis 2)

$$w_t = t / \sum_{i=0}^{T-1} 2^i, \quad (t = 1, \dots, T)$$

$$w_t = \{0,032; 0,065; 0,129; 0,258; 0,516\}$$

d) Datengenerierte clusteranalytische Gewichtung (euklidische Distanz)

$$w_t = \{0,156; 0,177; 0,194; 0,224; 0,250\} \text{ (für die LWR 1999, 2000, '01, '02, '03).}$$

Anmerkung: hier könnte man eventuell schon die Gewichte der Folgeseite bringen, die sich zu Eins addieren.

Alternative Kumulationsgewichte im Überblick ohne LWR 2000

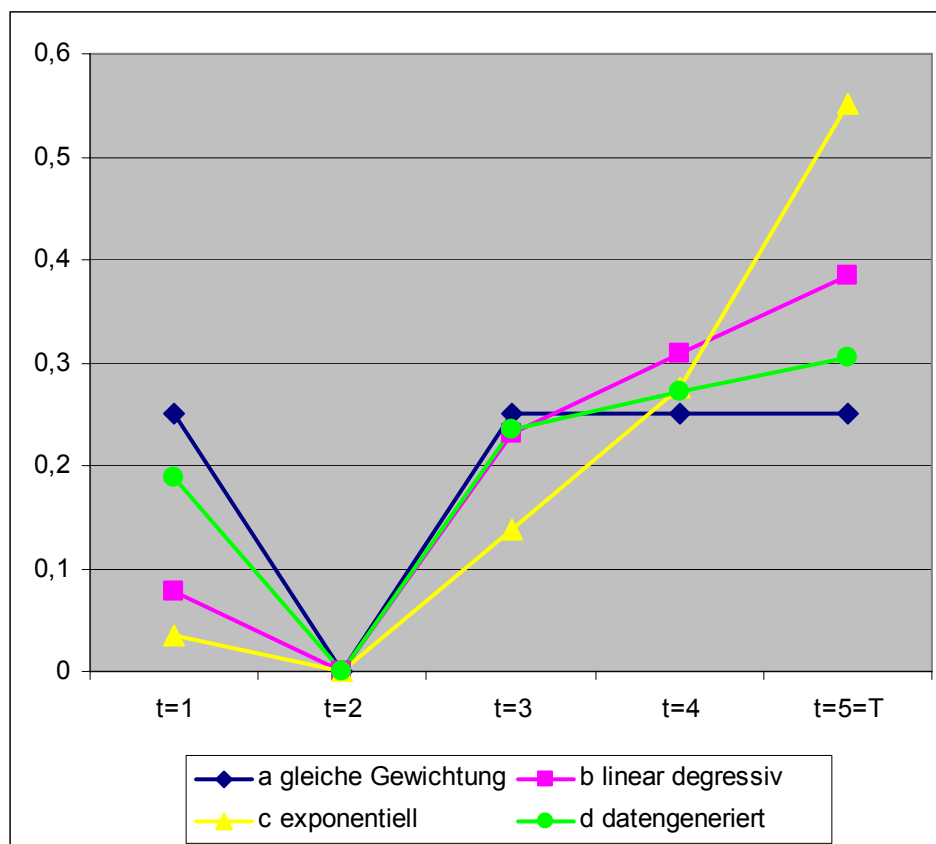
Bei einem Mittelwerts- und Streuungsvergleich der Ausgaben zwischen den laufenden Wirtschaftsrechnungen der Jahre 1999, 2000, 2001, 2002 und 2003 weichen die Informationen des Jahres 2000 auffällig von den Vergleichswerten der anderen Jahre ab. Eine genauere Inspektion der Einzeldaten aus den laufenden Wirtschaftsrechnungen 2000 (LWR 2000) ergibt beispielsweise, dass bei der Wahl einer 3-fachen Standardabweichung mehr als 35% aller Werte (und immer noch mehr als 15% bei einer 5-fachen Standardabweichung) ausserhalb dieses Streubereichs liegen.

Aufgrund der mangelhaften Datenqualität der LWR 2000 haben wir deshalb die laufenden Wirtschaftsrechnungen aus dem Jahr 2000 nicht weiter in unsere Analysen einbezogen. Damit wurden die Kumulationsgewichtungen, abweichend von den zuvor genannten theoretischen Überlegungen in der Weise abgeändert, dass die Daten aus der LWR 2000 mit einem Gewicht von Null bedacht wurden und somit in dem kumulierten File nicht zum Tragen kamen. Die übrigen Kumulationsgewichte wurden relativ zu ihrem bisherigen Gewicht neu bewertet, so dass sich die Kumulationsgewichte über die verschiedenen Jahrgänge hinweg wieder auf 100% addierten. Die folgende Tabelle 1 stellt die schließlich verwendeten Kumulationsgewichte in einer Übersicht dar. Diese Gewichte sind die Grundlage für alle folgenden Auswertungen.

Tabelle 1: Alternative Kumulationsgewichte im Überblick ohne LWR 2000

Neue Kumulationsgewichte (ohne LWR00, t=2)	a: gleiche Gewichtung	b: linear degressiv	c: exponentiell	d: datengeneriert
t=1 (1999)	25,0%	7,7%	3,4%	18,9%
2 (2000)	0,0%	0,0%	0,0%	0,0%
3 (2001)	25,0%	23,1%	13,8%	23,5%
4 (2002)	25,0%	30,8%	27,6%	27,2%
5 (2003)	25,0%	38,5%	55,2%	30,4%

Abbildung 1: Alternative Kumulationsgewichte im Überblick ohne LWR 2000



Wie mit der zur Tabelle 1 gehörenden Abbildung 1 deutlich wird, haben wir mit diesen vier alternativen Kumulationsgewichtungen ein breit gefächertes Gewichtungsspektrum erreicht,

das auch noch andere Gewichtungsschemata des Spektrums abdeckt und in ihren Wirkungen abzuschätzen erlaubt.

3.3 Modellgestützte Fortschreibung und Kalibrierung

Unter einem modellgestützten Ansatz sei ein durch die ökonomische Theorie begründeter Ansatz verstanden, auf dessen Basis mikroökonomische Schätzungen hier gemäß der Kumulierung der laufenden Wirtschaftsrechnungen für das Ausgabeverhalten vorgenommen werden. Aus der Vielzahl der mikroökonomisch fundierten Ansätze zur Erklärung des Ausgabeverhaltens⁵ wollen wir im Rahmen sogenannter kompletter Nachfragesysteme – stellvertretend für weitere Ansätze – auf einen Ansatz mit dem flexiblen AIDS-Nachfragesystem (Almost Ideal Demand System) etwas näher eingehen.

3.3.1 Modellgestützte Fortschreibung mit dem AIDS-Nachfragesystem und Kalibrierung

Eine Verbindung von modellgestützter Schätzung und Kalibrierung (Hochrechnung) haben Cassel, Granström, Lundquist und Selén 1997 bei ihrer Kumulation mit den schwedischen Haushaltsbefragungen HBS von 1985, 1988 und 1992⁶ vorgeschlagen. Bei der Kalibrierung unter Verwendung von modellgestützten Daten (CMD) verwenden sie das AIDS-Modell (Almost Ideal Demand System, Deaton 1990) für sieben Verbrauchsgütergruppen mit 6 Monatszeitpunkten unter Berücksichtigung von 10 Haushaltstypen.

Die Idee: Konsumausgabenanteile für bestimmte Konsumgüteruntergruppen werden aus einem Aggregat (z.B. Gesamtausgaben) mittels Regressionsanalyse geschätzt und dabei gleichzeitig kalibriert.

Die zentrale Gleichung eines generalisierten Regressionsschätzers ist

$$(5) \quad t_c(z) = t_z + (t_x^* - t_x)' \beta_{zx}$$

wobei $t_c(z)$ die geschätzten Konsumausgaben einer Subgruppe abhängig von den Gesamtausgaben z sind, $t_z = \sum_{i=1}^n (1/\pi_i) z_i = \sum_{i=1}^n d_i z_i$ die gewichtete Ausgabensumme (gewichtet mit dem Horvitz-Thompson Schätzer als Inverse der Auswahlwahrscheinlichkeit π) und β_{zx} ein Koeffizient zur Variablen x aus z mit $\beta_{zx} = (\sum d_i x_i z_i) / (\sum d_i x_i^2)$ ist.

Die Modellverknüpfung erfolgt über $t_x^* = R_m t_y$, den geschätzten Ausgaben aus einem Anteil R_m beispielsweise des Einkommens t_y .

⁵ Beispiele für Ausgabensysteme sind die vollständigen Nachfragesysteme mit flexibler funktionaler Form wie das Translog-Modell, das Almost Ideal Demand System (AIDS, QAIDS), das Rotterdam System etc., oder das Linear Expenditure System LES (Stone 1954) und seine Erweiterungen ELES Luch 1973 und FELES Merz 1983. Einen guten Überblick über Nachfragesysteme gibt z.B. Deaton 1990.

⁶ jeweils gleicher Stichprobenplan, gleicher Stichprobenumfang, Stichproben gezogen aus ‚Register of the Total Population‘, weitgehend Zufallsstichprobe

Ein bestimmter Modellansatz – hier das AIDS-Modell für die Ausgabenanteile R_m – geht über R_m als

$$(6) \quad R_m = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_j \log(x/P)$$

mit den geschätzten Parametern $\alpha_i, \gamma_{ij}, \beta_j$ (P ist das Preisniveau) dann direkt in den Gesamtansatz ein.

Die Resultate aus verschiedenen AIDS-Modellgestützten Varianten und ihrer Kalibrierung mit

- einfachen zufallsgestützten Stichprobenverfahren
- Kalibrierung unter Verwendung von Registerdaten (CRD)
- Kalibrierung unter Verwendung von modellgestützten Daten (CMD)
- Kalibrierung unter Verwendung von modellgestützten Daten und Registerdaten (CMRD)

lassen die Autoren zu folgendem Fazit kommen (Cassel et al. 1997, S. 19):

„Es kann offenbar zu Recht davon ausgegangen werden, dass die modellgestützten Kalibrierungsverfahren CMD und CMRD im Hinblick auf die Varianz und den systematischen Fehler zu guten Ergebnissen führen werden.“

Kritisch ist allerdings anzumerken, dass in vielen Ergebnissen aus ihrer Anwendung keine signifikanten Verbesserungen durch diesen modellgestützten Schätz- und Kalibrierungsansatz zu erkennen ist (vgl. auch die Diskussion dieser Ergebnisse in Selén 2002, S. 83 ff.).

Natürlich hängt eine Verbesserung der Ergebnisse von dem gewählten Modelltyp/-Nachfragesystem ab.

Da die Ergebnisse einer modellgestützten Fortschreibung modellabhängig sind, und die wissenschaftliche Diskussion über den zu verwendenden Ansatz in der Tat nicht abgeschlossen ist, ist es durchaus gerechtfertigt, wenn eine Institution wie das Statistische Bundesamt von einer solchen modellgestützten Fortschreibung Abstand nimmt.

Wir werden daher in unseren folgenden Simulationsrechnungen keine derart modellgestützten Fortschreibungen einbeziehen.

3.3.2 Modellgestützte Fortschreibung und Hochrechnung

Würde eine modellgestützte Fortschreibung allerdings gewählt, sei es aus einem Zeitreihenansatz, aus einer theoriegeleiteten Modellierung, dann wäre eine solche Fortschreibung generell wie folgt mit einer Hochrechnung als demographischer Anpassung verbunden:

Sind die inhaltlich interessanten Variablen (hier beispielsweise seltene Ereignisse) unabhängig von der demographischen Hochrechnung, kann die modellgestützte Fortschreibung aus den vorangegangenen Stichproben auch nach der Hochrechnung für die kumulierte Datenbasis erfolgen.

Sind diese inhaltlich interessanten Variablen aber abhängig von der demographischen Hochrechnung, dann ist die modellgestützte Fortschreibung direkt bei der Hochrechnung als ein weiteres Hochrechnungsmerkmal zu berücksichtigen.

3.3 Alternative Hochrechnungen

Eine neue Hochrechnung als demographische Gewichtung mit einer Anpassung an vorzugebende Rahmendaten ist im Allgemeinen immer dann nötig, wenn ein Datensatz nicht wirklich zufällig ist. Die Repräsentativität wird durch eine entsprechende individuelle Gewichtung erreicht, die mit der Berücksichtigung der individuellen Charakteristika jeder Mikroeinheit über einen für alle Mikroeinheiten identischen Hochrechnungsfaktor als reziproken Auswahlssatz weit hinausgeht.

Unsere neuen Hochrechnungen im Rahmen der alternativen Kumulationsansätze, die die demographische Repräsentativität sichern sollen, folgen einem informationstheoretischen Konzept nach dem Prinzip des minimalen Informationsverlustes (MIL-Prinzip). Ein informationstheoretischer Ansatz wurde bereits in Abschnitt 3.2.2 diskutiert, als die Information einer gesamten Stichprobe genutzt wurde, um eine Abschreibungsrate zu bestimmen. Wird die Informationstheorie auf das Hochrechnungsproblem angewendet, dann ergeben sich die neuen Hochrechnungsfaktoren nach dem MIL-Prinzip als Lösung des nichtlinearen Optimierungsproblems:

$$(10) \quad Z(\mathbf{p}, \mathbf{q}) = \min_{\mathbf{p}} \{ \sum_j p_j \log(p_j/q_j) \} \quad 0 < p_j, q_j < 1, \sum_j p_j = \sum_j q_j = 1,$$

unter der Nebenbedingung

$$(11) \quad \mathbf{S}\mathbf{p} = \mathbf{r}.$$

wobei \mathbf{p} der n (=Anzahl der Stichprobenhaushalte) Vektor der gesuchten Hochrechnungsfaktoren ist. Die zu optimierende Zielfunktion $Z(\mathbf{p}, \mathbf{q})$ minimiert die Abstände zu bereits vorhandenen Gewichten. Die Nebenbedingung sichert die Erfüllung extern vorgegebener Rahmenbedingungen \mathbf{r} aus m Hochrechnungscharakteristika über die mit \mathbf{p} gewichtete Aggregation der Stichprobeninformation \mathbf{S} (m, n -Matrix). Die Lagrangemultiplikatoren $\boldsymbol{\lambda}$ (m -Vektor) als Lösung bestimmen dann die neuen Hochrechnungsfaktoren mit

$$(12) \quad p_j = q_j \exp(\boldsymbol{\lambda}' \mathbf{s}^j - 1) \quad (j=1, \dots, n)$$

wobei \mathbf{s}^j die Charakteristika der j -ten Mikroeinheit (Haushalt) als Spaltenvektor beschreibt.

Jeder einzelne Hochrechnungsfaktor p_j (ein Faktor pro Haushalt j) gewichtet so simultan alle zu erfüllenden Hochrechnungscharakteristika. So werden beispielsweise für das Kumulationsfile KUMLWR 2003 über 24.000 individuelle Hochrechnungsfaktoren zu berechnen sein, die pro Haushalt 25 Hochrechnungscharakteristika simultan berücksichtigen.

Dieser informationstheoretische Ansatz sichert insbesondere die notwendige Positivität der neuen Gewichte (Beibehaltung der Mikroeinheiten) und erlaubt eine simultane Gewichtung auch hierarchisch geordneter Daten (hier Haushalts- und Personendaten). Eine ausführliche Diskussion der methodischen Grundlagen für die Hochrechnung nach dem Prinzip des minimalen Informationsverlustes und der effizienten globalen exponentiellen Approximation findet sich in Merz 1983a, 1985, 1994 und Merz und Stolze 2008. Informationen zu dem

verwendeten ADJUST Hochrechnungsprogramm finden sich in <http://ffb.uni-lueneburg.de/adjust> und Merz und Stolze 2004.

Dieser Hochrechnungsansatz wird – wie bei der Diskussion der Kumulationsgewichte angesprochen – eingesetzt für die neue Hochrechnung der einzelnen Stichproben $t=1,\dots,T$ und/oder für die neue Hochrechnung der kumulierten Stichprobe KUM mit alternativer Kumulationsgewichtung.

Es kommen hierbei drei grundlegende Varianten unter Berücksichtigung der in Abschnitt 3.2 diskutierten Kumulationsgewichte für die Hochrechnung in Frage (siehe auch Tabelle 2):

- I. Vor der eigentlichen Kumulationsgewichtung wird jeder Querschnitt für sich an die demographischen Rahmendaten r_t des Querschnittszeitpunkts nach dem MIL-Prinzip (Minimum Information Loss-Prinzip) hochgerechnet. Damit würden eventuell vorhandene Verzerrungen hinsichtlich der Charakteristika der neuen Hochrechnung aus t bereinigt.
- II. Die so gewonnenen Hochrechnungsfaktoren p_t^* werden dann multiplikativ mit den obigen alternativen Kumulationsgewichten verknüpft: $p_t^{*w_t}$ und als Ausgangsgewichte für eine abschließende Hochrechnung der gesamten kumulierten Stichprobe KUM mit den Rahmendaten des Jahres 2003 (T) r_T verwendet.
- III. Zur abschließenden Hochrechnung des kumulierten Files werden die mit den Kumulationsgewichten multiplizierten, ursprünglichen extern vorgegebenen Hochrechnungsfaktoren $q_t^{w_t}$ der laufenden Wirtschaftsrechnungen als Ausgangsgewichte verwendet.
- IV. Die Anpassung an die Rahmendaten von 2003 (T) r_T erfolgt in den einzelnen Querschnitten. So entsteht die Zielpopulation von N in jedem der Stichproben, so dass die kumulierte Stichprobe zunächst eine Population von TN_T Einheiten umfasst. Die Hochrechnungsfaktoren werden dann mit ihren jeweiligen Kumulationsgewichten w_t multipliziert. Die Kumulationsgewichte sollten sich auf 1 summieren, damit die kumulierte Stichprobe KUM schließlich die gewünschte Größe von N_T erreicht.

Die Alternative (III) ist – wie bereits erwähnt – flexibler und erlaubt später andere alternative Gewichtungen der einzelnen Querschnitte ohne eine neue demographische Hochrechnung insgesamt.

Tabelle 2: Alternative temporäre Kumulationsprozeduren

	Kumulationsgewichtung vor (abschließender) Anpassung an Rahmendaten		Kumulationsgewichtung nach Anpassung an Rahmendaten r_T
	Mit vorheriger Anpassung an r_t (Variante I)	Ohne vorherige Anpassung an r_t (Variante II)	
Hochrechnung (der einzelnen Files)	1. Anpassung der einzelnen Jahrgänge an die jeweiligen Rahmendaten q ↓ Anpassung an r_t $p^* \quad p^* = f(q; r_t)$		1. Jeweilige Anpassung der einzelnen Jahrgänge an die Rahmendaten von Jahrgang T q ↓ Anpassung an r_T $p^+ \quad p^+ = f(q; r_T)$
Kumulationsgewichtung	2. Berücksichtigung des Informationsverlustes durch Alterung (4 alternative Methoden) durch Kumulationsgewichtung der gefundenen Hochrechnungsfaktoren $p^* \cdot w_t$	1. Berücksichtigung des Informationsverlustes durch Alterung (4 alternative Methoden) durch Kumulationsgewichtung der vorhandenen Hochrechnungsfaktoren $q \cdot w_t$	2. Berücksichtigung des Informationsverlustes durch Alterung (4 alternative Methoden) durch Kumulationsgewichtung der gefundenen Hochrechnungsfaktoren im kumulierten File $p_{III} = f(p^+; w_t)$ $= p^+ \cdot w_t$
Hochrechnung (des kumulierten Files)	3. Anpassung des kumulierten Files an die Rahmendaten von Jahrgang T $p^* \cdot w_t$ ↓ Anpassung an r_T $p_I \quad p_I = f(p^* \cdot w_t; r_T)$	2. Anpassung des kumulierten Files an die Rahmendaten von Jahrgang T $q \cdot w_t$ ↓ Anpassung an r_T $p_{II} \quad p_{II} = f(q \cdot w_t; r_T)$	

q: ursprüngliche Gewichte der Querschnitte (hier laufende Wirtschaftsrechnungen LWR, Variable: HRD)

4 Zur Effizienz einer kumulierten Stichprobe

Generell ist aus Stichprobensicht eine Kumulation nicht überlappender Stichproben und damit unabhängiger Stichproben ohne wiederholte Befragung derselben Mikroeinheiten ideal, da nur sie effiziente Ergebnisse liefert. Zentrales Maß zur Signifikanzbestimmung ist die Varianz einer Größe. Wird nun eine vorangegangene Stichprobe mit einer laufenden Stichprobe kombiniert, wird die Varianz des Mittelwertes einer Variablen reduziert, je größer die Überlappungsrate P ist. Nach Cochran 1977 werden die Varianzen durch den Faktor

$$(4) \quad (1 - (1 - P)R^2) / (1 - (1 - P)^2 R^2),$$

reduziert, wobei R der Korrelationskoeffizient ist. Dies hat zur Folge, dass bei überlappender Kumulation eine kleinere Varianz ein verzerrtes höheres Signifikanzniveau anzeigt. Kordos (2002, S. 60) weist allerdings darauf hin, dass die maximale Varianzreduktion (bei optimalem P und optimalen Stichprobengewichten) durch den Faktor $(1 + ((1 - R^2)^{0.5})) / 2$ begrenzt wird.

Eine Varianzreduktion bei überlappender Kumulation gilt nicht nur für laufende Werte sondern auch für deren Veränderungen (Selén 2002, S. 75; Kish 1999, S. 136).

Da für unsere Analysen keinerlei Informationen über eine mögliche Überlappung vorliegen, können hierzu auch keine weiteren Abschätzungen vorgenommen werden; die kumulierte Stichprobe ist damit als eine in ihren Mikroeinheiten unabhängige Stichprobe anzusehen und entsprechend in ihrem Stichprobenfehler anzusehen.

Weitere Anmerkungen zur Genauigkeit einer kumulierten Stichprobe generell und hierzu sind in Merz 2004 zu finden.

5 Gewählte Kumulationsalternativen

Die Kumulationsalternativen vereinen einerseits die alternativen Verfahren der Kumulationsgewichte (Abschnitt 3.2) als auch die unterschiedlichen Varianten der Hochrechnung (Abschnitt 3.4).

Durch die **vier Verfahren der Kumulationsgewichtung** („Abschreibungsrate“) aus Abschnitt 3.2 mit

- a) Gleichgewichtung,
- b) Linear progressiver Gewichtung,
- c) Exponentieller Gewichtung und
- d) Datengenerierter clusteranalytischer Gewichtung

sowie mit den **drei verschiedenen Hochrechnungsvarianten**

- I) Kumulationsgewichtung *vor* (abschließender) Hochrechnung *mit* vorheriger Anpassung an laufende Rahmendaten r_t
- II) Kumulationsgewichtung *vor* (abschließender) Hochrechnung *ohne* vorherige Anpassung an laufende Rahmendaten r_t

III) Kumulationsgewichtung *nach* Hochrechnung
mit Anpassung nur an Rahmendaten r_T zum Zeitpunkt T

ergeben sich insgesamt 12 Simulationsalternativen mit **12 individuen-spezifischen Kumulationsfaktoren**

Kumulationsgewichtungsmethode
P_{Vorgehensweise Hochrechnung} \rightarrow

die über die Indizes den Kumulationsgewichts/Hochrechnungs-Kombinationen gekennzeichnet werden; P_{III}^d kennzeichnet beispielsweise eine datengenerierte Kumulationsgewichtung (Alterung, Abschreibung) nach der Anpassung an die gewünschten Rahmendaten aus T (Hochrechnungsvariante III).

Festzuhalten ist, dass diese 12 Kumulationsalternativen je nach demographischen Charakteristika der Mikroeinheiten (hier Haushalte) für jede Mikroeinheit in der kumulierten Stichprobe verschieden sind.

Diese 12 Verfahren der Kumulation sind Gegenstand der weiteren Simulationsrechnungen.

Gewählter Hochrechnungsrahmen: Haushalts- und Personeninformationen aus dem Mikrozensus 1999 bis 2003

Als geeignete Rahmendaten für die jeweilige Hochrechnung haben wir die Ergebnisse der Mikrozensus 1999, 2000, 2001, 2002 und 2003 dank der Sonderauswertungen des Statistischen Bundesamtes heranziehen können. Ziel dabei war es, für die zu untersuchenden Ausgaben privater Haushalte auch entsprechende Strukturdaten auf der Haushaltsebene einzubeziehen und damit eine diesbezügliche Repräsentativität zu erreichen. Die Strukturdaten umfassen so Informationen zu unterschiedlichen Haushaltstypen sowie für das Ausgabeverhalten wichtige weitere Merkmale wie die Erwerbstätigkeit als Kaufkraft-hintergrund, die soziale Stellung sowie die Altersstruktur nach Geschlecht:

- Soziale Stellung der Personen (Beamter, Angestellte, Arbeiter, Rentner)
- Privathaushalte nach Haushaltstyp (Einpersonenhaushalte, Haushalte mit 2, 3, 4 und mehr Personen)
- Privathaushalte nach der Zahl der Erwerbstätigen im Haushalt (keine Erwerbstätige, 1, 2ei Erwerbstätige)
- Bevölkerung nach Geschlecht und Alter (jeweils 7 Altersklassen).

Der Vektor \mathbf{r} der Rahmendaten besteht somit aus $m=25$ Hochrechnungscharakteristika, die die jeweilige Anzahl der Haushalte bzw. Personen für die Bundesrepublik Deutschland in den einzelnen Jahren angeben (detaillierte Angaben finden sich in Tabelle A1 des Anhangs).

Die Hochrechnungsprozedur liefert dann beispielsweise für das Kumulationsfile KUMLWR 2003 über 24.000 individuelle Hochrechnungsfaktoren, die pro Haushalt diese 25 Hochrechnungscharakteristika simultan berücksichtigen.

6 Ausgaben privater Haushalte im kumulierten File der laufenden Wirtschaftsrechnungen KUMLWR 2003 - Ergebnisse der Kumulationsalternativen

Für eine Beurteilung des Informationsgewinns einer Kumulation, hier der Kumulation aus einzelnen laufenden Wirtschaftsrechnungen (jeweils kleinere Stichproben), wählen wir den Vergleich zu den Ergebnissen einer inhaltlich passenden Großstichprobe, hier der Einkommens- und Verbrauchsstichprobe (EVS). Inhaltlich analysieren wir die individuellen Quartalsausgaben untergliedert in die Ausgaben der 12 Hauptausgabekategorien.

Das kumulierte Datenfile KUMLWR 2003 der fünf beziehungsweise vier laufenden Wirtschaftsrechnungen LWR 1999, (2000), 2001, 2002 und 2003 für $T=2003$ besteht schließlich aus 30.480 Haushaltsdatensätzen (davon 6.169, also 20,2% aus der LWR 2000, die durch passende Hochrechnungsfaktoren ausgeklammert werden). Die Vergleichsgroßstichprobe EVS 2003 umfasst 50.511 Haushalte mit den entsprechenden Ausgabeinformationen.

6.1 Vergleich der alternativen Kumulationsfaktoren in KUMLWR 2003

Beginnen wir mit einem Vergleich der für jeden Haushalt berechneten Kumulationsfaktoren aus den 12 Kumulationsalternativen (vier Kumulationsgewichtungen mal drei Hochrechnungsvarianten) für das Kumulationsfile KUMLWR 2003.

Mittelwertvergleich

Da alle Kumulationsfaktoren zu den gewünschten Rahmendaten gemäß r_T führen, weisen alle Verfahrensweisen einen gemeinsamen Mittelwert der Kumulationsfaktoren von 1.251 auf.

Varianzvergleich

Anhand von anderen statistischen Parametern lassen sich jedoch die verschiedenen Verteilungen der Faktoren beurteilen. Eine zentrale Rolle spielt hierbei die Varianz: bezüglich der Hochrechnungsfaktoren ist eine möglichst homogene Verteilung der Faktoren wünschenswert, damit das Gewicht, das einzelnen Informationen in einer Stichprobe zukommt, nicht über Gebühr streut. Die Werte der Tabelle beziehen sich auf die Hochrechnungsfaktoren ohne die LWR 2000, da diese (einheitlich Null) die Standardabweichung verzerren würde.

Grundsätzlich lässt sich über die Streuung (Standardabweichungen) der Kumulationsfaktoren (vgl. Tab. 3) sagen, dass diese bei den Verfahren I und III leicht unter denen der Faktoren liegen, die aus Verfahren II hervorgegangen sind.

Erwartungsgemäß wirkt sich die **Streuung der Kumulationsgewichte** auch auf die Streuung der gefundenen Kumulationsfaktoren aus. So sind die Standardabweichungen der Faktoren bei Gleichgewichtung (Variante a) bei allen drei Hochrechnungsvarianten am niedrigsten, gefolgt von den datengenerierten Kumulationsgewichten (Variante d) und den linear progressiven Kumulationsgewichten (Variante b). Die deutlich höchsten Varianzen weisen alle drei Hochrechnungsverfahren unter Verwendung der Variante c, der exponentiell progressiven Kumulationsgewichtung, auf.

Tabelle 3: Standardabweichungen der neuen Kumulationsfaktoren aus KUMLWR 2003

Standardabweichungen HRF			
Kumulationsgewicht	Hochrechnungsvariante		
	I	II	III
a (Gleichgewichtung)	2.011,1	2.075,6	2.012,2
b (linear progressiv)	2.317,0	2.401,7	2.306,2
c (exponentiell)	2.756,6	2.787,8	2.757,8
d (datengeneriert)	2.049,7	2.126,5	2.050,4

*Hochrechnungsfaktoren ohne LWR 2000

Minima, Maxima und Quartile

Die Kumulationsfaktoren geben schließlich an, wie viele Mikroeinheiten (hier Haushalte) der Grundgesamtheit durch die Mikroeinheit der Stichprobe repräsentiert werden. Ein wichtiger Beurteilungsaspekt sind in Ergänzung zu den Streuungsinformationen hierbei die Minima und Maxima der Hochrechnungsfaktoren. Generell geben ja die Hochrechnungsfaktoren an, wie viele weitere Mikroeinheiten durch einen Datensatz der Stichprobe repräsentiert werden. Nach Möglichkeit sollten beide Werte möglichst wenig vom Mittelwert abweichen, so dass zum einen die Informationen, die in der Stichprobe enthalten sind, entsprechend gewürdigt werden können (was nicht der Fall wäre, wenn die Hochrechnungsfaktoren zu niedrig sind), und zum anderen keine Mikroeinheiten ein übermäßig starkes Gewicht im Vergleich zu den anderen Einheiten bekommen.

Die Werte der Tabelle 4 beziehen sich wieder auf die Hochrechnungsfaktoren ohne LWR 2000, da diese die Ergebnisse verzerren würde.

Tabelle 4: Minima, Quartile und Maxima der neuen Hochrechnungsfaktoren aus KUMLWR 2003

Verteilungsparameter HRF				
Kumulationsgewicht	Parameter	Hochrechnungsvariante		
		I	II	III
a (Gleichgewichtung)	Min	1,49	4,22	1,45
	1. Quartal	486,42	495,55	488,65
	Median	976,75	971,19	970,45
	3.Quartal	1909,76	1887,77	1914,06
	Max	63730,49	95898,63	64329,70
b (linear progressiv)	Min	0,42	1,59	0,40
	1. Quartal	349,52	366,50	355,19
	Median	854,70	853,95	852,87
	3.Quartal	1862,04	1845,09	1844,59
	Max	59737,73	96207,64	59440,64

Verteilungsparameter HRF		Hochrechnungsvariante		
Kumulationsgewicht	Parameter	I	II	III
c (exponentiell)	Min	0,21	0,93	0,20
	1. Quartal	234,41	245,32	241,17
	Median	670,30	679,26	673,03
	3.Quartal	1759,18	1749,81	1769,01
	Max	55884,05	67272,68	56864,33
d (datengeneriert)	Min	1,12	3,56	1,09
	1. Quartal	468,24	482,40	474,48
	Median	956,94	969,42	960,97
	3.Quartal	1889,74	1900,41	1898,22
	Max	60059,02	93224,84	60469,92

*Hochrechnungsfaktoren ohne LWR 2000

Wie Tabelle 4 zu entnehmen ist, liegen bei allen Verfahren die Hochrechnungsfaktoren relativ eng zwischen dem ersten und dritten Quartil beieinander. Allerdings weisen die hohen Maximalwerte auf notwendige Anpassungen von stark unterrepräsentierten Gruppen in den Stichproben hin.

Konzentration der Kumulationsfaktorenverteilung

Mit Informationen zur Konzentration der Kumulationsfaktoren (Tabelle 5) mit den entsprechenden Gini-Koeffizienten zeigt sich erwartungsgemäß ein entsprechendes Bild zu den Streuungsergebnissen: Je größer die Streuung der Kumulationsgewichte, desto ungleicher (höherer Gini-Koeffizient) sind auch die endgültigen Hochrechnungsfaktoren verteilt.

Tabelle 5: Gini-Koeffizienten der neuen Kumulationsfaktoren aus KUMLWR 2003*

Gini HRF	Hochrechnungsvariante		
Kumulationsgewicht	I	II	III
a (Gleichgewichtung)	0,52	0,51	0,52
b (linear progressiv)	0,58	0,57	0,58
c (exponentiell)	0,65	0,64	0,64
d (datengeneriert)	0,53	0,52	0,52

*Hochrechnungsfaktoren ohne LWR 2000

Die exponentielle Gewichtung führt im Vergleich zu den anderen Kumulationsgewichtungen bei allen Hochrechnungsvarianten zu einer größeren Ungleichheit der Faktorverteilungen.

6.2 Vergleich der Ausgaben aus den Kumulationsalternativen in KUMLWR 2003

Im letzten Abschnitt haben wir die Verteilung der Kumulationsfaktoren selbst für das aus den laufenden Wirtschaftsrechnungen 1999 bis 2003 gewonnene Kumulationsfile KUMLWR 2003 verglichen. Diskutieren wir nun die Auswirkungen der Kumulationsalternativen auf die zentralen Ausgabekategorien der laufenden Wirtschaftsrechnungen in unserem Kumulationsfile KUMLWR 2003. Wir berücksichtigen neben der Summe der Ausgaben, die 12 Hauptausgabekategorien sowie die Ausgaben für zwei eher kleinere Warengruppen (Damenhosen (lang, aus Baumwolle) sowie PC-Komplettsysteme und Notebooks), um zudem auch die Auswirkungen auf eher seltene Ausgabekategorien einbeziehen zu können (vg. Tabelle 6).

Tabelle 6: Untersuchte Ausgabekategorien der Laufenden Wirtschaftsrechnungen in KUMLWR 2003

Summe W01 bis W12	Gesamtausgaben
W01	Nahrungsmittel und alkoholfreie Getränke
W02	Alkoholische Getränke und Tabakwaren
W03	Bekleidung und Schuhe
W04	Wohnung, Wasser, Strom, Gas u. a.
W05	Einrichtungsgegenstände, Apparate, Geräte u. a.
W06	Gesundheitspflege
W07	Verkehr
W08	Nachrichtenübermittlung
W09	Freizeit, Unterhaltung und Kultur
W10	Bildungswesen
W11	Beherbergungs- und Gaststättendienstleistungen
W12	Andere Waren und Dienstleistungen
w0312226	Damenhosen (lang, aus Baumwolle)
w0913011	PC-Komplettsysteme und Notebooks

Zur Beurteilung der zentralen Lage, der Streuung und der Ungleichheit/Konzentration der Ergebnisse verwenden wir in den folgenden drei Tabellen (7, 8 und 9) jeweils die arithmetischen Mittelwerte, die Standardabweichungen und die Gini-Koeffizienten der Gesamtausgaben, der einzelnen Ausgabengruppen W01 bis W12 sowie zwei beispielhafte Spezialfälle w0312226 (Damenhosen (lang, aus Baumwolle)) und w0913011 (PC-Komplettsysteme und Notebooks) für die insgesamt 12 verschiedenen Kumulationsalternativen. Eine Tabelle stellt dabei jeweils eine Hochrechnungsvariante mit ihren vier verschiedenen alternativen Kumulationsgewichten dar.

Die zum Teil hohen Gini-Koeffizienten signalisieren eine relativ hohe Konzentration beziehungsweise ungleiche Verteilung der Ausgaben, die darauf zurückzuführen ist, dass es sich um eher selten getätigte Ausgaben handelt (z.B. Ausgaben für Gesundheit, PC Komplettsysteme). In diesen Ausgabengruppen wurden für nur relativ wenige Haushalte

Ausgaben verzeichnet, während ein Großteil der Haushalte in diesen Gruppen überhaupt keine Ausgaben tätigte.

Insgesamt sind im Kumulationsfile die Abweichungen zwischen den verschiedenen Kumulationsalternativen innerhalb der Hochrechnungsvarianten nicht ausgeprägt. Lediglich in Warengruppe 6 (Gesundheitspflege) fallen in allen drei Hochrechnungsvarianten bei der exponentiellen Kumulationsgewichtung Abweichungen von über 5% über dem Mittelwert der Hochrechnungsvariante auf. Ebenso ist bei Warengruppe 12 (andere Waren und Dienstleistungen) innerhalb der Hochrechnungsvariante II eine stärkere Abweichung unter Verwendung der gleichverteilten Kumulationsgewichte zu verzeichnen. Im folgenden Ergebnisvergleich unseres Kumulationsfiles mit alternativen Informationen zum Jahr 2003 gehen wir detaillierter auch auf diese Ergebnisse ein.

Tabelle 7: Mittelwerte, Standardabweichungen und Gini-Koeffizienten verschiedener Ausgabentypen nach Hochrechnungsvariante I (KUMLWR 2003) Quartalswerte in €

		Hochrechnungsvariante I			
		a Gleich- gewichtung	b linear progressiv	c exponentiell progressiv	d datengeneri- ert
Warengruppe	Maß				
Summe W01 bis W12 Gesamtausgaben	Mittelwert	5.972,00	6.022,25	6.045,50	5.990,05
	Std.-Abw.	4.119,83	4.130,58	4.145,27	4.124,38
	Gini	0,325	0,323	0,323	0,324
W01 Nahrungsmittel und alkoholfreie Getränke	Mittelwert	686,41	697,02	700,03	690,07
	Std.-Abw.	396,19	401,43	399,66	397,85
	Gini	0,309	0,308	0,306	0,309
W02 Alkoholische Getränke und Tabakwaren	Mittelwert	134,65	132,82	131,86	133,98
	Std.-Abw.	179,97	170,53	170,47	176,91
	Gini	0,571	0,573	0,573	0,571
W03 Bekleidung und Schuhe	Mittelwert	337,57	333,21	330,22	335,94
	Std.-Abw.	327,59	326,06	321,80	327,01
	Gini	0,470	0,471	0,472	0,470
W04 Wohnung, Wasser, Strom, Gas u. a.	Mittelwert	1.792,28	1.814,30	1.823,80	1.800,29
	Std.-Abw.	880,84	879,43	866,02	879,80
	Gini	0,258	0,255	0,251	0,257
W05 Einrichtungsgegenst., Apparate, Geräte u. a.	Mittelwert	415,67	414,11	410,52	414,64
	Std.-Abw.	1.007,00	1.005,77	961,66	1.001,75
	Gini	0,656	0,657	0,656	0,656
W06 Gesundheitspflege	Mittelwert	265,40	276,54	281,49	269,43
	Std.-Abw.	683,62	724,59	755,11	699,72
	Gini	0,737	0,737	0,737	0,737

		Hochrechnungsvariante I			
Warengruppe	Maß	a Gleich- gewichtung	b linear progressiv	c exponentiell progressiv	d datengeneri ert
W07 Verkehr	Mittelwert	943,26	946,35	957,01	945,24
	Std.-Abw.	2.422,22	2.417,26	2.452,11	2.423,68
	Gini	0,656	0,653	0,654	0,655
W08 Nachrichtenübermittlung	Mittelwert	161,52	170,85	174,08	164,77
	Std.-Abw.	127,07	131,57	132,08	128,63
	Gini	0,370	0,368	0,366	0,370
W09 Freizeit, Unterhaltung und Kultur	Mittelwert	655,80	664,90	666,23	658,88
	Std.-Abw.	848,31	878,54	877,97	858,92
	Gini	0,506	0,510	0,510	0,507
W10 Bildungswesen	Mittelwert	301,66	305,01	297,04	302,65
	Std.-Abw.	760,59	816,42	765,04	780,94
	Gini	0,906	0,908	0,906	0,906
W11 Beherbergungs- und Gaststättendl.	Mittelwert	367,51	368,38	367,08	367,63
	Std.-Abw.	467,57	471,96	472,72	468,93
	Gini	0,585	0,589	0,591	0,587
W12 Andere Waren und Dienstleistungen	Mittelwert	250,29	243,81	243,84	248,14
	Std.-Abw.	433,86	405,93	403,31	424,77
	Gini	0,503	0,501	0,503	0,503
w0312226 Damenhosen (lang, aus Baumwolle)	Mittelwert	135,01	132,00	131,31	133,62
	Std.-Abw.	97,78	96,42	95,66	97,13
	Gini	0,937	0,914	0,886	0,928
w0913011 PC-Komplettsysteme und Notebooks	Mittelwert	2.377,93	2.277,60	2.262,16	2.337,07
	Std.-Abw.	1.630,60	1.606,43	1.575,59	1.617,93
	Gini	0,982	0,981	0,981	0,981

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

Tabelle 8: Mittelwerte, Standardabweichungen und Gini-Koeffizienten verschiedener Ausgabentypen nach Hochrechnungsvariante II (KUMLWR 2003) – Quartalswerte in €

		Hochrechnungsvariante II			
Warengruppe	Maß	a Gleich- gewichtung	b linear progressiv	c exponentiell progressiv	d datengeneri ert
Summe W01 bis W12	Mittelwert	5.981,60	6.036,07	6.056,53	6.000,61
Gesamtausgaben	Std.-Abw.	4.133,02	4.151,34	4.164,36	4.140,25
	Gini	0,325	0,324	0,323	0,324
W01	Mittelwert	688,42	698,84	701,37	691,98
Nahrungsmittel und alkoholfreie Getränke	Std.-Abw.	398,01	402,25	400,15	399,35
	Gini	0,310	0,308	0,306	0,309
W02	Mittelwert	135,08	133,09	132,03	134,36
Alkoholische Getränke und Tabakwaren	Std.-Abw.	178,46	170,08	170,28	175,69
	Gini	0,570	0,572	0,573	0,570
W03	Mittelwert	338,36	334,30	331,11	336,84
Bekleidung und Schuhe	Std.-Abw.	327,83	326,93	322,67	327,43
	Gini	0,470	0,471	0,472	0,470
W04	Mittelwert	1.796,45	1.819,88	1.828,62	1.804,85
Wohnung, Wasser, Strom, Gas u. a.	Std.-Abw.	881,23	883,24	870,13	881,23
	Gini	0,258	0,255	0,251	0,257
W05	Mittelwert	417,08	415,78	411,74	416,15
Einrichtungsgegenst., Apparate, Geräte u. a.	Std.-Abw.	1.024,72	1.028,66	979,46	1.021,20
	Gini	0,656	0,658	0,656	0,656
W06	Mittelwert	264,36	276,42	281,31	268,46
Gesundheitspflege	Std.-Abw.	669,35	714,14	746,03	685,58
	Gini	0,735	0,735	0,736	0,735
W07	Mittelwert	942,42	947,44	958,10	944,96
Verkehr	Std.-Abw.	2.425,31	2.429,86	2.466,34	2.430,00
	Gini	0,656	0,654	0,655	0,655
W08	Mittelwert	162,19	171,11	174,16	165,30
Nachrichtenübermittlung	Std.-Abw.	128,65	132,67	132,79	130,03
	Gini	0,372	0,369	0,367	0,371
W09	Mittelwert	657,75	666,28	667,10	660,62
Freizeit, Unterhaltung und Kultur	Std.-Abw.	856,57	883,17	880,40	866,12
	Gini	0,506	0,510	0,510	0,508

		Hochrechnungsvariante II			
Warengruppe	Maß	a Gleich- gewichtung	b linear progressiv	c exponentiell progressiv	d datengeneri ert
W10 Bildungswesen	Mittelwert	298,04	301,90	295,25	299,13
	Std.-Abw.	723,80	780,63	742,42	744,12
	Gini	0,905	0,907	0,905	0,906
W11 Beherbergungs- und Gaststättendl.	Mittelwert	367,42	368,81	367,38	367,64
	Std.-Abw.	469,63	475,01	475,17	471,26
	Gini	0,586	0,590	0,592	0,587
W12 Andere Waren und Dienstleistungen	Mittelwert	250,74	244,43	244,26	248,67
	Std.-Abw.	448,36	412,14	406,83	437,43
	Gini	0,503	0,501	0,503	0,502
w0312226 Damenhosen (lang, aus Baumwolle)	Mittelwert	135,44	132,36	131,57	134,04
	Std.-Abw.	97,59	96,37	95,60	96,98
	Gini	0,935	0,912	0,884	0,926
w0913011 PC-Komplettsysteme und Notebooks	Mittelwert	2.378,35	2.273,51	2.253,11	2.335,31
	Std.-Abw.	1.634,80	1.606,20	1.571,39	1.620,26
	Gini	0,982	0,981	0,981	0,981

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

Tabelle 9: Mittelwerte, Standardabweichungen und Gini-Koeffizienten verschiedener Ausgabentypen nach Hochrechnungsvariante III (KUMLWR 2003) – Quartalswerte in €

		Hochrechnungsvariante III			
Warengruppe	Maß	a Gleich- gewichtung	b linear progressiv	c exponentiell progressiv	d datengeneri ert
Summe W01 bis W12 Gesamtausgaben	Mittelwert	5.977,89	6.029,70	6.053,98	5.996,58
	Std.-Abw.	4.126,40	4.136,73	4.151,49	4.130,94
	Gini	0,325	0,324	0,323	0,324
W01 Nahrungsmittel und alkoholfreie Getränke	Mittelwert	687,68	698,60	701,66	691,46
	Std.-Abw.	397,92	403,58	402,02	399,75
	Gini	0,310	0,308	0,307	0,309
W02 Alkoholische Getränke und Tabakwaren	Mittelwert	134,81	132,95	131,99	134,13
	Std.-Abw.	181,34	171,06	170,78	178,04
	Gini	0,571	0,573	0,574	0,572
W03 Bekleidung und Schuhe	Mittelwert	338,24	333,87	330,87	336,63
	Std.-Abw.	327,98	326,45	322,15	327,43
	Gini	0,469	0,471	0,472	0,470

		Hochrechnungsvariante III			
Warengruppe	Maß	a Gleich- gewichtung	b linear progressiv	c exponentiell progressiv	d datengeneri- ert
W04 Wohnung, Wasser, Strom, Gas u. a.	Mittelwert	1.794,49	1.817,75	1.827,47	1.802,93
	Std.-Abw.	882,59	882,98	870,24	882,22
	Gini	0,259	0,256	0,252	0,258
W05 Einrichtungsgegenst., Apparate, Geräte u. a.	Mittelwert	416,71	414,42	410,81	415,47
	Std.-Abw.	1.010,76	1.004,77	959,44	1.003,88
	Gini	0,655	0,657	0,656	0,656
W06 Gesundheitspflege	Mittelwert	264,25	275,84	281,01	268,37
	Std.-Abw.	674,81	717,50	749,03	691,12
	Gini	0,736	0,737	0,737	0,736
W07 Verkehr	Mittelwert	943,55	947,07	958,19	945,69
	Std.-Abw.	2.424,60	2.420,19	2.455,25	2.426,26
	Gini	0,656	0,653	0,654	0,655
W08 Nachrichtenübermittlung	Mittelwert	161,68	171,10	174,38	164,96
	Std.-Abw.	127,63	132,20	132,76	129,22
	Gini	0,371	0,368	0,367	0,370
W09 Freizeit, Unterhaltung und Kultur	Mittelwert	656,13	665,19	666,52	659,21
	Std.-Abw.	848,95	878,31	876,79	859,31
	Gini	0,506	0,510	0,510	0,507
W10 Bildungswesen	Mittelwert	302,75	306,34	298,45	303,86
	Std.-Abw.	767,56	822,95	773,61	788,02
	Gini	0,906	0,908	0,906	0,906
W11 Beherbergungs- und Gaststättendl.	Mittelwert	367,50	368,36	367,16	367,62
	Std.-Abw.	467,54	472,22	472,96	468,98
	Gini	0,585	0,589	0,592	0,587
W12 Andere Waren und Dienstleistungen	Mittelwert	250,67	244,10	244,09	248,50
	Std.-Abw.	436,48	407,13	404,03	427,05
	Gini	0,503	0,501	0,503	0,503
w0312226 Damenhosen (lang, aus Baumwolle)	Mittelwert	135,10	132,17	131,49	133,74
	Std.-Abw.	97,90	96,58	95,81	97,26
	Gini	0,937	0,914	0,885	0,928
w0913011 PC-Komplettsysteme und Notebooks	Mittelwert	2.374,31	2.269,17	2.252,54	2.331,82
	Std.-Abw.	1.627,26	1.601,44	1.571,55	1.614,02
	Gini	0,982	0,981	0,981	0,981

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

7 Auswirkungen der Kumulationsalternativen auf die Ausgaben privater Haushalte - Vergleich KUMLWR 2003 mit EVS 2003 und LWR 2003

Für die Evaluation des Informationsgewinns einer Kumulation, hier der Kumulation aus einzelnen laufenden Wirtschaftsrechnungen (jeweils kleinere Stichproben), wählen wir den Vergleich zu den Ergebnissen einer inhaltlich passenden Großstichprobe, hier der Einkommens- und Verbrauchsstichprobe (EVS), sowie den Vergleich zu den laufenden Wirtschaftsrechnungen des Kumulationsjahres. Inhaltlich analysieren wir die individuellen Quartalsausgaben untergliedert in die Ausgaben der 12 Hauptausgabekategorien.

Das kumulierte Datenfile KUMLWR 2003 der fünf beziehungsweise vier laufenden Wirtschaftsrechnungen LWR 1999, (2000), 2001, 2002 und 2003 für T=2003 besteht schließlich aus 30.480 Haushaltsdatensätzen (davon 6.169, also 20,2% aus der LWR 2000, die durch passende Hochrechnungsfaktoren im Weiteren ausgeklammert werden). Die Vergleichsgroßstichprobe EVS 2003 umfasst 50.511 Haushalte mit den entsprechenden Ausgabeinformationen.

Welche Kumulationsalternative aus dem Kumulationsfile KUMLWR 2003 der laufenden Wirtschaftsrechnungen 1999, (ohne 2000, siehe oben), 2001, 2002, 2003 schneidet nun im Vergleich zu der Einkommens- und Verbrauchsstichprobe EVS 2003 am besten ab? Kann das Kumulationsfile dabei „besser“ sein als die gleichzeitige ursprüngliche laufende Wirtschaftsrechnung 2003? Diese Fragen wollen wir nun empirisch fundiert beantworten.

Das Kumulationsfile umfasst alle Querschnitte, wobei je nach Kumulationsgewicht (Abschreibungsrate) vergangene Querschnitte unterschiedlich gewichtet werden. Da die einzelnen Querschnitte hier voneinander unabhängig sind (prinzipiell und soweit bekannt), werden im Kumulationsfile alle Haushalte der T Querschnitte einbezogen. Prinzipiell und soweit bekannt sind im Vergleichsfile, hier die EVS, andere Haushalte einbezogen. Damit scheidet ein auf den jeweiligen Haushalt bezogener Vergleich der Ausgabenwerte aus dem Kumulationsfile und dem Vergleichsfile aus. Bekannte Maße zur Prognosegüte (vgl. z.B. Merz 1980) mit einem individuellen Vergleich der tatsächlichen und prognostizierten Werte können damit wegen der Unabhängigkeit der zu vergleichenden Stichproben zur Beurteilung der Anpassungsgüte nicht verwendet werden. Es verbleibt somit nur eine Beurteilung der über alle individuellen Haushaltswerte gebildeten Aggregate.

Ausgabekategorien

Wir können neben der Summe der Ausgaben die Aggregate der einzelnen zwölf Hauptausgabekategorien analysieren:

Gesamtausgaben	Summe W01 bis W12
Nahrungsmittel und alkoholfreie Getränke	W01
Alkoholische Getränke und Tabakwaren	W02
Bekleidung und Schuhe	W03
Wohnung, Wasser, Strom, Gas u. a.	W04
Einrichtungsgegenstände, Apparate, Geräte u. a.	W05
Gesundheitspflege	W06

Verkehr	W07
Nachrichtenübermittlung	W08
Freizeit, Unterhaltung und Kultur	W09
Bildungswesen	W10
Beherbergungs- und Gaststättendienstleistungen	W11
Andere Waren und Dienstleistungen	W12

Für die noch in den laufenden Wirtschaftrechnungen (und damit auch in KUMLWR) gegebenen eher seltenen Ausgabegruppen (Damenhosen sowie PC-Komplettsysteme) waren in dem uns zur Verfügung stehenden EVS 2003-File keine entsprechenden Ausgabegruppen gegeben; sie fallen also für eine Analyse der Prognosegüte bedauerlicherweise aus.

Kumulationsalternativen

Wir vergleichen die drei Hochrechnungsvarianten aus unterschiedlicher Hochrechnung (I, II, III)

- (I) Kumulationsgewichtung *vor* (abschließender) Hochrechnung
mit vorheriger Anpassung an laufende Rahmendaten r_t
- (II) Kumulationsgewichtung *vor* (abschließender) Hochrechnung
ohne vorheriger Anpassung an laufende Rahmendaten r_t
- (III) Kumulationsgewichtung *nach* Hochrechnung
mit Anpassung nur an Rahmendaten r_T zum Zeitpunkt T

mit den jeweils vier alternativen Kumulationsgewichtungen

- a Gleichgewichtung, b linear progressiv, c exponentiell, d datengeneriert.]

Die insgesamt zwölf Kumulationsalternativen ordnen jedem einzelnen Haushalt in jedem betrachteten Jahr ein jeweils individuelles Gewicht im aggregierten Kumulationsfile zu. Das Kumulationsfile KUMLWR 2003 enthält additiv alle fünf beziehungsweise vier laufenden Wirtschaftsrechnungen von 1999 bis 2003 mit 30.480 bzw. 24.311 (ohne LWR 2000) Haushalten.

Vergleich EVS 2003, LWR 2003 mit KUMLWR 2003

In Anlehnung an die Effizienz eines Schätzers als wünschenswerte Schätzeigenschaft mit Erwartungstreue und minimaler Varianz werden wir als Evaluationskriterium für die Anpassungsgüte von KUMLWR an die EVS das arithmetische Mittel sowie als Streuungsmaß die Standardabweichung der Ausgabenaggregate verwenden.

Die folgenden Tabellen enthalten somit als Gütemaße die arithmetischen Mittelwerte und die Standardabweichungen für alle Ausgabegruppen der EVS 2003, der LWR 2003 sowie die relativen Abweichungen zu allen Kumulationsalternativen aus KUMLWR 2003 (in Prozent zur EVS 2003) sowie Maximum, Minimum und die Spannweite (Range=Max-Min) dieser alternativenspezifischen Abweichungen (also nicht die Spannweite der individuellen Ausgabewerte).

Tabelle m1: Vergleich der Mittelwerte ausgewählter Konsumgüteraussgaben aus KUMLWR 2003 mit der EVS 2003 und der LWR 2003 – Hochrechnungsvariante I

Mittelwerte		EVS 2003* (Var.-Name)	LWR 2003	LWR-EVS	Kumulationsfile 2003 **						
					Hochrechnungsvariante I				Minimum	Maximum	Max-Min
					la	lb	lc	ld			
Summe W01 bis W12 (Gesamtausgaben)	Mittelwert Rel. Abw. zur EVS	6.392,70	5.858,18	-8,40%	5.972,00 -6,60%	6.022,25 -5,80%	6.045,50 -5,40%	5.990,05 -6,30%	5.972,00 -6,60%	6.045,50 -5,40%	73,51 1,15%
W01 Nahrungsmittel und alkoholfreie Getränke	Mittelwert Rel. Abw. zur EVS	778,08 (NG)	709,78	-8,80%	686,41 -11,80%	697,02 -10,42%	700,03 -10,00%	690,07 -11,30%	686,41 -11,80%	700,03 -10,00%	13,61 1,75%
W02 Alkoholische Getränke und Tabakwaren	Mittelwert Rel. Abw. zur EVS	119,47 (AlkGT)	143,45	20,10%	134,65 12,70%	132,82 11,17%	131,86 10,40%	133,98 12,10%	131,86 10,40%	134,65 12,70%	2,78 2,33%
W03 Bekleidung und Schuhe	Mittelwert Rel. Abw. zur EVS	326,25 (BekISch)	343,74	5,40%	337,57 3,50%	333,21 2,10%	330,22 1,20%	335,94 3,00%	330,22 1,20%	337,57 3,50%	7,34 2,25%
W04 Wohnung, Wasser, Strom, u.a.	Mittelwert Rel. Abw. zur EVS	2.043,75 (W04selbst)	1.844,24	-9,80%	1.792,28 -12,30%	1.814,30 -11,20%	1.823,80 -10,80%	1.800,29 -11,90%	1.792,28 -12,30%	1.823,80 -10,80%	31,51 1,54%
W05 Einrichtungsgegenst., Apparate, Geräte, u.a.	Mittelwert Rel. Abw. zur EVS	375,01 (W05selbst)	410,79	9,50%	415,67 10,80%	414,11 10,40%	410,52 9,50%	414,64 10,60%	410,52 9,50%	415,67 10,80%	5,16 1,38%
W06 Gesundheitspflege	Mittelwert Rel. Abw. zur EVS	251,55 (Gesundpflege)	343,72	36,60%	265,4 5,50%	276,54 9,90%	281,49 11,90%	269,43 7,10%	265,4 5,50%	281,49 11,90%	16,1 6,40%
W07 Verkehr	Mittelwert Rel. Abw. zur EVS	890,61 (Verkehr)	928,53	4,30%	943,26 5,90%	946,35 6,30%	957,01 7,50%	945,24 6,10%	943,26 5,90%	957,01 7,50%	13,75 1,54%
W08 Nachrichtenübermittlung	Mittelwert Rel. Abw. zur EVS	200,48 (Nachrichten)	173,19	-13,60%	161,52 -19,40%	170,85 -14,80%	174,08 -13,20%	164,77 -17,80%	161,52 -19,40%	174,08 -13,20%	12,56 6,26%
W09 Freizeit, Unterhaltung und Kultur	Mittelwert Rel. Abw. zur EVS	770,16 (Freizeit)	662,9	-13,90%	655,8 -14,80%	664,9 -13,70%	666,23 -13,50%	658,88 -14,40%	655,8 -14,80%	666,23 -13,50%	10,43 1,35%
W10 Bildungswesen	Mittelwert Rel. Abw. zur EVS	55,3 (Bildung)	262,83	375,30%	301,66 445,50%	305,01 451,60%	297,04 437,20%	302,65 447,30%	297,04 437,20%	305,01 451,60%	7,97 14,41%
W11 Begerbergungs- und Gaststättendienstleistungen	Mittelwert Rel. Abw. zur EVS	292,87 (Beherbg)	382,43	30,60%	367,51 25,50%	368,38 25,80%	367,08 25,30%	367,63 25,50%	367,08 25,30%	368,38 25,80%	1,3 0,44%
W12 Andere Waren und Dienstleistungen	Mittelwert Rel. Abw. zur EVS	289,17 (AndWarDI)	249,66	-13,70%	250,29 -13,40%	243,81 -15,70%	243,84 -15,70%	248,14 -14,20%	243,81 -15,70%	250,29 -13,40%	6,48 2,24%
W0312226 Damenhosen (lang, aus Baumwolle)	Mittelwert Rel. Abw. zur EVS	n/a***	132,25		135,01	132	131,31	133,62	131,31	135,01	3,7
W0913011 PC-Komplettsysteme und Notebooks	Mittelwert Rel. Abw. zur EVS	n/a***	2.404,34		2.377,93	2.277,60	2.262,16	2.337,07	2.262,16	2.377,93	115,77

* Daten gewichtet mit EVS-interner Variable "HRB" (führt zu N=35.899.946 Fällen)

** Kumuliertes File aus den LWR-Files 1999-2003 mit alternativen Vorgehensweisen und Kumulationsgewichten; Preisbereinigung auf 2003; Umgerechnet in Euro und Quartalswerte

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

Tabelle m2: Vergleich der Mittelwerte ausgewählter Konsumgüterausgaben aus KUMLWR 2003 mit der EVS 2003 und der LWR 2003 – Hochrechnungsvariante II

Mittelwerte		EVS 2003* (Var.-Name)	LWR 2003	LWR-EVS	Kumulationsfile 2003**				Minimum	Maximum	Max-Min
					Hochrechnungsvariante II						
					Ila	Ilb	Ilc	Ild			
Summe W01 bis W12 (Gesamtausgaben)	Mittelwert	6.392,70	5.858,18		5.981,60	6.036,07	6.056,53	6.000,61	5.981,60	6.056,53	74,93
	Rel. Abw. zur EVS			-8,40%	-6,40%	-5,60%	-5,30%	-6,10%	-6,40%	-5,30%	1,17%
W01 Nahrungsmittel und alkoholfreie Getränke	Mittelwert	778,08	709,78		688,42	698,84	701,37	691,98	688,42	701,37	12,95
	Rel. Abw. zur EVS	(NG)		-8,80%	-11,50%	-10,20%	-9,90%	-11,10%	-11,50%	-9,90%	1,66%
W02 Alkoholische Getränke und Tabakwaren	Mittelwert	119,47	143,45		135,08	133,09	132,03	134,36	132,03	135,08	3,05
	Rel. Abw. zur EVS	(AlkGT)		20,10%	13,10%	11,40%	10,50%	12,50%	10,50%	13,10%	2,55%
W03 Bekleidung und Schuhe	Mittelwert	326,25	343,74		338,36	334,3	331,11	336,84	331,11	338,36	7,25
	Rel. Abw. zur EVS	(BekISch)		5,40%	3,70%	2,50%	1,50%	3,20%	1,50%	3,70%	2,22%
W04 Wohnung, Wasser, Strom, u.a.	Mittelwert	2.043,75	1.844,24		1.796,45	1.819,88	1.828,62	1.804,85	1.796,45	1.828,62	32,16
	Rel. Abw. zur EVS	(W04selbst)		-9,80%	-12,10%	-11,00%	-10,50%	-11,70%	-12,10%	-10,50%	1,57%
W05 Einrichtungsgegenst., Apparate, Geräte, u.a.	Mittelwert	375,01	410,79		417,08	415,78	411,74	416,15	411,74	417,08	5,34
	Rel. Abw. zur EVS	(W05selbst)		9,50%	11,20%	10,90%	9,80%	11,00%	9,80%	11,20%	1,42%
W06 Gesundheitspflege	Mittelwert	251,55	343,72		264,36	276,42	281,31	268,46	264,36	281,31	16,95
	Rel. Abw. zur EVS	(Gesundpflege)		36,60%	5,10%	9,90%	11,80%	6,70%	5,10%	11,80%	6,74%
W07 Verkehr	Mittelwert	890,61	928,53		942,42	947,44	958,1	944,96	942,42	958,1	15,68
	Rel. Abw. zur EVS	(Verkehr)		4,30%	5,80%	6,40%	7,60%	6,10%	5,80%	7,60%	1,76%
W08 Nachrichtenübermittlung	Mittelwert	200,48	173,19		162,19	171,11	174,16	165,3	162,19	174,16	11,97
	Rel. Abw. zur EVS	(Nachrichten)		-13,60%	-19,10%	-14,60%	-13,10%	-17,50%	-19,10%	-13,10%	5,97%
W09 Freizeit, Unterhaltung und Kultur	Mittelwert	770,16	662,9		657,75	666,28	667,1	660,62	657,75	667,1	9,35
	Rel. Abw. zur EVS	(Freizeit)		-13,90%	-14,60%	-13,50%	-13,40%	-14,20%	-14,60%	-13,40%	1,21%
W10 Bildungswesen	Mittelwert	55,3	262,83		298,04	301,9	295,25	299,13	295,25	301,9	6,65
	Rel. Abw. zur EVS	(Bildung)		375,30%	439,00%	445,90%	433,90%	440,90%	433,90%	445,90%	12,02%
W11 Beverbergungs- und Gaststättendienstleistungen	Mittelwert	292,87	382,43		367,42	368,81	367,38	367,64	367,38	368,81	1,43
	Rel. Abw. zur EVS	(Beherbg)		30,60%	25,50%	25,90%	25,40%	25,50%	25,40%	25,90%	0,49%
W12 Andere Waren und Dienstleistungen	Mittelwert	289,17	249,66		250,74	244,43	244,26	248,67	244,26	250,74	6,49
	Rel. Abw. zur EVS	(AndWarDI)		-13,70%	-13,30%	-15,50%	-15,50%	-14,00%	-15,50%	-13,30%	2,24%
W0312226 Damenhosen (lang, aus Baumwolle)	Mittelwert	n/a***	132,25		135,44	132,36	131,57	134,04	131,57	135,44	3,87
	Rel. Abw. zur EVS										
W0913011 PC-Komplettssysteme und Notebooks	Mittelwert	n/a***	2.404,34		2.378,35	2.273,51	2.253,11	2.335,31	2.253,11	2.378,35	125,24
	Rel. Abw. zur EVS										

* Daten gewichtet mit EVS-interner Variable "HRB" (führt zu N=35.899.946 Fällen)

** Kumuliertes File aus den LWR-Files 1999-2003 mit alternativen Vorgehensweisen und Kumulationsgewichten; Preisbereinigung auf 2003; Umgerechnet in Euro und Quartalswerte

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

Tabelle m3: Vergleich der Mittelwerte ausgewählter Konsumgüterausgaben aus KUMLWR 2003 mit der EVS 2003 und der LWR 2003 – Hochrechnungsvariante III

Mittelwerte		EVS 2003* (Var.-Name)	LWR 2003	LWR-EVS	Kumulationsfile 2003**				Hochrechnungsvariante III		
					IIIa	IIIb	IIIc	IIId	Minimum	Maximum	Max-Min
Summe W01 bis W12 (Gesamtausgaben)	Mittelwert	6.392,70	5.858,18		5.977,89	6.029,70	6.053,98	5.996,58	5.977,89	6.053,98	76,09
	Rel. Abw. zur EVS			-8,40%	-6,50%	-5,70%	-5,30%	-6,20%	-6,50%	-5,30%	1,19%
W01 Nahrungsmittel und alkoholfreie Getränke	Mittelwert	778,08 (NG)	709,78		687,68	698,6	701,66	691,46	687,68	701,66	13,99
	Rel. Abw. zur EVS			-8,80%	-11,60%	-10,22%	-9,80%	-11,10%	-11,60%	-9,80%	1,80%
W02 Alkoholische Getränke und Tabakwaren	Mittelwert	119,47 (AlkGT)	143,45		134,81	132,95	131,99	134,13	131,99	134,81	2,82
	Rel. Abw. zur EVS			20,10%	12,80%	11,28%	10,50%	12,30%	10,50%	12,80%	2,36%
W03 Bekleidung und Schuhe	Mittelwert	326,25 (BeklSch)	343,74		338,24	333,87	330,87	336,63	330,87	338,24	7,37
	Rel. Abw. zur EVS			5,40%	3,70%	2,30%	1,40%	3,20%	1,40%	3,70%	2,26%
W04 Wohnung, Wasser, Strom, u.a.	Mittelwert	2.043,75 (W04selbst)	1.844,24		1.794,49	1.817,75	1.827,47	1.802,93	1.794,49	1.827,47	32,99
	Rel. Abw. zur EVS			-9,80%	-12,20%	-11,10%	-10,60%	-11,80%	-12,20%	-10,60%	1,61%
W05 Einrichtungsgegenst., Apparate, Geräte, u.a.	Mittelwert	375,01 (W05selbst)	410,79		416,71	414,42	410,81	415,47	410,81	416,71	5,9
	Rel. Abw. zur EVS			9,50%	11,10%	10,50%	9,50%	10,80%	9,50%	11,10%	1,57%
W06 Gesundheitspflege	Mittelwert	251,55 (Gesundpflege)	343,72		264,25	275,84	281,01	268,37	264,25	281,01	16,76
	Rel. Abw. zur EVS			36,60%	5,10%	9,70%	11,70%	6,70%	5,10%	11,70%	6,66%
W07 Verkehr	Mittelwert	890,61 (Verkehr)	928,53		943,55	947,07	958,19	945,69	943,55	958,19	14,64
	Rel. Abw. zur EVS			4,30%	5,90%	6,30%	7,60%	6,20%	5,90%	7,60%	1,64%
W08 Nachrichtenübermittlung	Mittelwert	200,48 (Nachrichten)	173,19		161,68	171,1	174,38	164,96	161,68	174,38	12,7
	Rel. Abw. zur EVS			-13,60%	-19,40%	-14,70%	-13,00%	-17,70%	-19,40%	-13,00%	6,34%
W09 Freizeit, Unterhaltung und Kultur	Mittelwert	770,16 (Freizeit)	662,9		656,13	665,19	666,52	659,21	656,13	666,52	10,39
	Rel. Abw. zur EVS			-13,90%	-14,80%	-13,60%	-13,50%	-14,40%	-14,80%	-13,50%	1,35%
W10 Bildungswesen	Mittelwert	55,3 (Bildung)	262,83		302,75	306,34	298,45	303,86	298,45	306,34	7,88
	Rel. Abw. zur EVS			375,30%	447,20%	454,00%	439,70%	449,50%	439,70%	454,00%	14,25%
W11 Bevergungs- und Gaststättendienstleistungen	Mittelwert	292,87 (Beherbg)	382,43		367,5	368,36	367,16	367,62	367,16	368,36	1,2
	Rel. Abw. zur EVS			30,60%	25,50%	25,80%	25,40%	25,50%	25,40%	25,80%	0,41%
W12 Andere Waren und Dienstleistungen	Mittelwert	289,17 (AndWarDI)	249,66		250,67	244,1	244,09	248,5	244,09	250,67	6,58
	Rel. Abw. zur EVS			-13,70%	-13,30%	-15,60%	-15,60%	-14,10%	-15,60%	-13,30%	2,28%
W0312226 Damenhosen (lang, aus Baumwolle)	Mittelwert	n/a***	132,25		135,1	132,17	131,49	133,74	131,49	135,1	3,62
	Rel. Abw. zur EVS										
W0913011 PC-Komplettsysteme und Notebooks	Mittelwert	n/a***	2.404,34		2.374,31	2.269,17	2.252,54	2.331,82	2.252,54	2.374,31	121,78
	Rel. Abw. zur EVS										

* Daten gewichtet mit EVS-interner Variable "HRB" (führt zu N=35.899.946 Fällen)

** Kumuliertes File aus den LWR-Files 1999-2003 mit alternativen Vorgehensweisen und Kumulationsgewichten; Preisbereinigung auf 2003; Umgerechnet in Euro und Quartalswerte

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

7.1 Kumulationsalternativen im Vergleich – Mittelwerte der Ausgaben privater Haushalte

Für die Evaluation der Kumulationsalternativen anhand eines Vergleiches der Mittelwerte der Ausgaben privater Haushalte sind die Ergebnisse für die Hochrechnungsvarianten I, II und III in den drei Evaluationstabellen m1, m2 und m3 aufgeführt.

Gesamtausgaben: Vergleich der Mittelwerte

Die EVS 2003 weist mittlere Gesamtausgaben von 6.392,70 EURO im Quartal aus. Insgesamt erreichen alle zwölf Kumulationsalternativen aus KUMLR 2003 diesen Wert recht gut. Zwar unterschätzen sie diesen Wert noch zwischen -5,3% und -6,6% je nach Kumulationsalternative, die Ergebnisse des kumulierten Files KUMLR sind den EVS-Werten aber alle näher im Vergleich zu den Werten der laufenden Wirtschaftsrechnung LWR 2003 (-8,4%).

Für die mittleren Gesamtausgaben schneiden damit im Vergleich zur EVS 2003 alle Kumulationsalternativen in KUMLR 2003 besser ab als die alleinige LWR 2003. Alle zwölf Kumulationsalternativen liegen mit ihren Mittelwerten der Gesamtausgaben sehr nahe beieinander: die alternativenspezifischen Aggregate weichen maximal 1,3 Prozentpunkte voneinander ab.

Welche Kumulationsgewichtung (Abschreibungsrate) empfiehlt sich?

Ausgabekategorien: Vergleich der Mittelwerte

Für die Gesamtausgaben liegen die Kumulationsalternativen nahe beieinander, Unterschiede in der Anpassungsgüte finden wir aber hinsichtlich der einzelnen Ausgabekategorien.

Auffallend sind die relativ geringen Bildungsausgaben in der EVS 2003. Mit durchschnittlich 55 EURO im Quartal ist dieser Wert gegenüber der LWR 2003 von 262 EURO beziehungsweise rund 300 EURO aus allen Alternativen aus KUMLR 2003 unverhältnismäßig klein; hier sind Abgrenzungsprobleme im vorhandenen EVS-File zu vermuten. Bei der weiteren Diskussion werden wir also die Bildungsausgaben außen vor lassen müssen.

Aus allen drei Hochrechnungsvarianten I, II und III wird mit Tabelle m4 ein identisches Muster der *besten*¹ *Kumulationsgewichtung* aus den Vergleichstabellen sichtbar. Die besten Ergebnisse werden erreicht mit der *exponentiellen Gewichtung (c)* und der *Gleichgewichtung (a)* jeweils für die folgenden Ausgabekategorien:

Beste Kumulationsgewichtung c (exponentiell) für

Nahrungsmittel und alkoholfreie Getränke	W01
Alkoholische Getränke und Tabakwaren	W02
Bekleidung und Schuhe	W03
Wohnung, Wasser, Strom, Gas u. a.	W04
Einrichtungsgegenstände, Apparate, Geräte u. a.	W05

¹ Im Sinne der kleinsten absoluten relativen Abweichung zum jeweiligen EVS-Wert.

Nachrichtenübermittlung	W08
Freizeit, Unterhaltung und Kultur	W09
Bildungswesen	W10
Beherbergungs- und Gaststättendienstleistungen	W11
<i>Beste Kumulationsgewichtung a (Gleichgewichtung) für</i>	
Gesundheitspflege	W06
Verkehr	W07
Andere Waren und Dienstleistungen	W12.

Alle anderen Kumulationsgewichte (b linear progressiv und d datengeneriert) ergeben in allen Hochrechnungsvarianten eine etwas größere relative Abweichung. Allerdings liegen die mittleren Ausgaben aller zwölf Kumulationsalternativen für alle Ausgabekategorien doch nahe beieinander: die Spannweiten der zwölf relativen Abweichungen für alle Ausgabekategorien liegen zwischen -0,41 Prozentpunkten (Beherbergung) und 6,74 Prozentpunkten (Gesundheitspflege).

Die Anpassungsgüte der Kumulationsalternativen ist also abhängig von den Ausgabegruppen. Eine gleiche Gewichtung der zeitlichen Abschreibungsraten (Kumulationsgewichte) – und damit eine stärkere Gewichtung auch der ältesten Information (LWR 1999) – führt für die Kategorien Gesundheitspflege, Verkehr und andere Waren und Dienstleistungen zu besseren Ergebnissen als für alle anderen neun Ausgabekategorien, für die eine starke zeitnahe exponentielle Gewichtung zu besseren Ergebnissen führt.

Möglicherweise unterliegt den drei Ausgabekategorien mit stärkerer Gewichtung der ältesten Information (Gesundheitspflege, Verkehr und andere Waren und Dienstleistungen) eine gewisse Gewohnheitsbildung, während bei den anderen Ausgabekategorien eine sich im Zeitablauf schnellere Verhaltensänderung aus der stärkeren zeitnahen Gewichtung abzeichnet. Hier sind natürlich noch weitere Untersuchungen notwendig.

Welche Hochrechnungsvariante empfiehlt sich?

Ausgabekategorien: Vergleich der Mittelwerte

Zentrales Ergebnis bisher: die besten Kumulationsgewichte sind c: exponentiell mit starker Gewichtung der aktuellsten Stichprobe und a: gleiche Gewichtung aller Stichproben je nach Ausgabekategorie.

Welche der drei Hochrechnungsvarianten ist nun diesbezüglich am besten? Fassen wir dazu die entsprechenden Teilergebnisse aus den Tabellen m1, m2 und m3 in Tabelle m4 zusammen. Die jeweilige Spalte zu den einzelnen Hochrechnungsvarianten vereint die eben diskutierten besten Werte der Kumulationsgewichte c beziehungsweise a.

Alle drei Hochrechnungsvarianten (I, II und III) liegen wieder im Ergebnis sehr nahe beieinander: die Spannweite zwischen bester und relative schlechtester Hochrechnungsvariante über alle Ausgabekategorien liegt nur zwischen 0,1% und 0,4%; für das Bildungswesen liegt sie allerdings bei 5,8%. Hier war ja die besondere Diskrepanz zwischen den EVS und LWR-Werten bereits angesprochen worden.

Wählen wir als zusätzliches Beurteilungskriterium den Erstellungsaufwand, so ist die Hochrechnungsvariante III, die eine Hochrechnung des gesamten Kumulationsfiles zum Zeitpunkt T=2003 zuerst vornimmt, und erst dann die alternativen Kumulationsgewichte

verwendet, die günstigste Variante. Sie erlaubt nämlich auch nachträgliche alternative Kumulationsgewichtungen der einzelnen Querschnitte, ohne dass mit jeder neuen Kumulationsgewichtung auch eine neue Hochrechnung erforderlich wäre (wie bei den Hochrechnungsvarianten I und II).

Eingedenk des erforderlichen Erstellungsaufwandes, der sehr nahen Ergebnisse zwischen den Verfahren I, II und III sowie der zudem besten Ergebnisse in einigen Ausgabekategorien empfiehlt sich damit das

Hochrechnungsverfahren III: Kumulationsgewichtung mit alternativen Kumulationsgewichten (Abschreibungsraten) erst nach der abschließenden Hochrechnung des gesamten Kumulationsfiles.

KUMLWR 2003 besser als LWR 2003 im Vergleich zur EVS 2003? Vergleich der Mittelwerte der einzelnen Ausgabekategorien

Gemessen wieder an der relativen Abweichung der jeweiligen Mittelwerte ist der Informationsgewinn der Kumulation mit KUMLWR 2003 gegenüber der alleinigen aktuellen laufenden Wirtschaftsrechnung LWR 2003 nicht nur insgesamt sondern auch für die einzelnen Ausgabekategorien durchgängig größer. Wenn auch für einzelne Ausgabekategorien der Unterschied nicht immer groß ist, so gibt es doch Gewinne bis zu über 30 Prozentpunkte (Gesundheitspflege), also eine bessere Anpassung der Ausgabesummen durch die Kumulation aus KUMLWR 2003 (II a,c) gegenüber LWR 2003.

Die Berücksichtigung vergangener Querschnitte verbessert also die Ergebnisse – sowohl für die Gesamtausgaben als auch für einzelne Ausgabekategorien – der alleinigen laufenden Wirtschaftsrechnungen und spricht deutlich für eine Kumulation der Stichprobe.

Tabelle m4: Vergleich der Mittelwerte ausgewählter Konsumgüteraussgaben aus KUMLWR 2003 mit der EVS 2003 und der LWR 2003 – Beste Ergebnisse der Hochrechnungsvarianten I, II, III

Mittelwerte		EVS 2003* (Var.-Name)	LWR 2003	LWR-EVS	KUMLWR Kumulations- gewichte c,a	2003 **			Minimum	Maximum	Max-Min
						I	II	III			
Summe W01 bis W12 (Gesamtausgaben)	Mittelwert Rel. Abw. zur EVS	6.392,70	5.858,18	-8,40%	c	6.045,50 -5,40%	6.056,53 -5,30%	6.053,98 -5,30%	6.045,50 -5,40%	6.056,53 -5,30%	11,03 0,17%
W01 Nahrungsmittel und alkoholfreie Getränke	Mittelwert Rel. Abw. zur EVS	778,08 (NG)	709,78	-8,80%	c	700,03 -10,00%	701,37 -9,90%	701,66 -9,80%	700,03 -10,00%	701,66 -9,80%	1,63 0,21%
W02 Alkoholische Getränke und Tabakwaren	Mittelwert Rel. Abw. zur EVS	119,47 (AlkGT)	143,45	20,10%	c	131,86 10,40%	132,03 10,50%	131,99 10,50%	131,86 10,40%	132,03 10,50%	0,17 0,14%
W03 Bekleidung und Schuhe	Mittelwert Rel. Abw. zur EVS	326,25 (BeklSch)	343,74	5,40%	c	330,22 1,20%	331,11 1,50%	330,87 1,40%	330,22 1,20%	331,11 1,50%	0,88 0,27%
W04 Wohnung, Wasser, Strom, u.a.	Mittelwert Rel. Abw. zur EVS	2.043,75 (W04selbst)	1.844,24	-9,80%	c	1.823,80 -10,80%	1.828,62 -10,50%	1.827,47 -10,60%	1.823,80 -10,80%	1.828,62 -10,50%	4,82 0,24%
W05 Einrichtungsgegenst., Apparate, Geräte, u.a.	Mittelwert Rel. Abw. zur EVS	375,01 (W05selbst)	410,79	9,50%	c	410,52 9,50%	411,74 9,80%	410,81 9,50%	410,52 9,50%	411,74 9,80%	1,22 0,33%
W06 Gesundheitspflege	Mittelwert Rel. Abw. zur EVS	251,55 (Gesundpflege)	343,72	36,60%	a	265,4 5,50%	264,36 5,10%	264,25 5,10%	264,25 5,10%	265,4 5,50%	1,14 0,45%
W07 Verkehr	Mittelwert Rel. Abw. zur EVS	890,61 (Verkehr)	928,53	4,30%	a	943,26 5,90%	942,42 5,80%	943,55 5,90%	942,42 5,80%	943,55 5,90%	1,13 0,13%
W08 Nachrichtenübermittlung	Mittelwert Rel. Abw. zur EVS	200,48 (Nachrichten)	173,19	-13,60%	c	174,08 -13,20%	174,16 -13,10%	174,38 -13,00%	174,08 -13,20%	174,38 -13,00%	0,3 0,15%
W09 Freizeit, Unterhaltung und Kultur	Mittelwert Rel. Abw. zur EVS	770,16 (Freizeit)	662,9	-13,90%	c	666,23 -13,50%	667,1 -13,40%	666,52 -13,50%	666,23 -13,50%	667,1 -13,40%	0,87 0,11%
W10 Bildungswesen	Mittelwert Rel. Abw. zur EVS	55,3 (Bildung)	262,83	375,30%	c	297,04 437,20%	295,25 433,90%	298,45 439,70%	295,25 433,90%	298,45 439,70%	3,2 5,79%
W11 Beverbergungs- und Gaststättendienstleistungen	Mittelwert Rel. Abw. zur EVS	292,87 (Beherbg)	382,43	30,60%	c	367,08 25,30%	367,38 25,40%	367,16 25,40%	367,08 25,30%	367,38 25,40%	0,3 0,10%
W12 Andere Waren und Dienstleistungen	Mittelwert Rel. Abw. zur EVS	289,17 (AndWarDI)	249,66	-13,70%	a	250,29 -13,40%	250,74 -13,30%	250,67 -13,30%	250,29 -13,40%	250,74 -13,30%	0,46 0,16%

* Daten gewichtet mit EVS-interner Variable "HRB" (führt zu N=35.899.946 Fällen)

** Kumuliertes File aus den LWR-Files 1999-2003 mit alternativen Vorgehensweisen und Kumulationsgewichten; Preisbereinigung auf 2003; Umgerechnet in Euro und Quartalswerte

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

7.2 Kumulationsalternativen im Vergleich – Streuungen der Ausgaben privater Haushalte

Während eine möglichst kleine Abweichung der Mittelwerte aus dem Kumulationsfile und dem Vergleichsfile im Sinne einer erwartungstreuen Schätzung als wünschenswert betrachtet werden kann, ist dies hinsichtlich der Streuung diskussionswürdig. Man könnte sogar argumentieren, dass eine größere Streuung im Kumulationsfile im Vergleich zu einer Großstichprobe (EVS) besser ist. Dahinter steht das Argument, dass die Großstichprobe EVS immer noch eine Stichprobe und nicht die Grundgesamtheit ist, so dass hinsichtlich der Grundgesamtheit die Heterogenität der Ausgaben besser abgedeckt wird; ein Argument gerade für dauerhafte Konsumgüter, die seltener gekauft werden und damit in einer Querschnittsperiode nur „ungenügend“ erfasst werden.

Bleiben wir – auch mangels Alternativen – aber bei der naheliegenden Annahme, dass die Großstichprobe näher der Grundgesamtheit als die kleineren Stichproben des Kumulationsfiles ist, und daher eine kleinere Varianzabweichung des Kumulationsfiles zur EVS als wünschenswerte Eigenschaft anzusehen ist.

Da wir der Erwartungstreue ein größeres Gewicht bei der Beurteilung des Kumulationsfiles geben – und damit den obigen Ergebnissen der Mittelwertvergleiche – sei die Diskussion der Streuungsergebnisse nicht vertieft geführt und nur zentrale Ergebnisse aus dem Streuungsvergleich zwischen KUMLWR 2003 und EVS 2003 sowie LWR 2003 in Anlehnung an den Mittelwertvergleich mit den Evaluationstabellen s1, s2, s3 und zusammengefasst mit Tabelle s4 im Folgenden betrachtet.

Gesamtausgaben: Vergleich der Streuungen

Die Standardabweichungen aller Kumulationsalternativen liegen zum einen nahe beieinander und weichen von der EVS-Streuung geringfügig zwischen -6% und - 5,4% ab. Dagegen überschätzt die LWR 2003 die Gesamtausgabenstreuung bemerkenswert mit 33%; die Kumulation mit KUMLWR ist hier also in allen Kumulationsalternativen deutlich besser.

Welche Kumulationsgewichtung (Abschreibungsrate) empfiehlt sich?

Ausgabekategorien: Vergleich der Streuungen

Wie schon bei der Analyse der Mittelwerte angesprochen, ist die Streuung der Bildungsausgaben in der uns vorliegenden EVS unverhältnismäßig klein, so dass wir auch hier diese Ausgabekategorie bei der Diskussion außen vor lassen.

Die hinsichtlich der geringsten Streuungsabweichung von KUMLWR 2003 zur EVS 2003 besten Kumulationsgewichtungen je Hochrechnungsvariante sind in ihren Häufigkeiten in Tabelle 9 zusammengefasst. Wie auch bezüglich der Mittelwerte – wenn auch nicht in der gleichen Eindeutigkeit – schält sich die exponentielle Gewichtung als die vergleichsweise beste Kumulationsgewichtung in allen drei Hochrechnungsvarianten heraus (vgl. dazu die Tabellen s1, s2 und s3).

Tabelle s1: Vergleich der Streuungen ausgewählter Konsumgüteraussgaben aus KUMLWR 2003 mit der EVS 2003 und der LWR 2003 – Hochrechnungsvariante I

Standardabweichung		EVS 2003* (Var. Name)	LWR 2003	LWR-EVS	Kumulationsfile 2003 **				Hochrechnungsvariante I		
					la	lb	lc	ld	Minimum	Maximum	Max-Min
Summe W01 bis W12 (Gesamtausgaben)	Standardabweichung Rel. Abw. zur EVS	4.382,21	5.865,27	33,80%	4.119,83 -6,00%	4.130,58 -5,70%	4.145,27 -5,40%	4.124,38 -5,90%	4.119,83 -6,00%	4.145,27 -5,40%	25,44 0,58%
W01 Nahrungsmittel und alkoholfreie Getränke	Standardabweichung Rel. Abw. zur EVS (NG)	450,02	436,95	-2,90%	396,19 -12,00%	401,43 -10,80%	399,66 -11,20%	397,85 -11,60%	396,19 -12,00%	401,43 -10,80%	5,24 1,16%
W02 Alkoholische Getränke und Tabakwaren	Standardabweichung Rel. Abw. zur EVS (AlkGT)	165,23	186,7	13,00%	179,97 8,90%	170,53 3,21%	170,47 3,20%	176,91 7,10%	170,47 3,20%	179,97 8,90%	9,5 5,75%
W03 Bekleidung und Schuhe	Standardabweichung Rel. Abw. zur EVS (BekISch)	339,62	425,26	25,20%	327,59 -3,50%	326,06 -4,00%	321,8 -5,20%	327,01 -3,70%	321,8 -5,20%	327,59 -3,50%	5,79 1,71%
W04 Wohnung, Wasser, Strom, u.a.	Standardabweichung Rel. Abw. zur EVS (W04selbst)	1.431,80	1.011,63	-29,30%	880,84 -38,50%	879,43 -38,60%	866,02 -39,50%	879,8 -38,60%	866,02 -39,50%	880,84 -38,50%	14,81 1,03%
W05 Einrichtungsgegenst., Apparate, Geräte, u.a.	Standardabweichung Rel. Abw. zur EVS (W05selbst)	892,86	1.603,00	79,50%	1.007,00 12,80%	1.005,77 12,60%	961,66 7,70%	1.001,75 12,20%	961,66 7,70%	1.007,00 12,80%	45,34 5,08%
W06 Gesundheitspflege	Standardabweichung Rel. Abw. zur EVS (Gesundpflege)	796,36	1.392,08	74,80%	683,62 -14,20%	724,59 -9,00%	755,11 -5,20%	699,72 -12,10%	683,62 -14,20%	755,11 -5,20%	71,49 8,98%
W07 Verkehr	Standardabweichung Rel. Abw. zur EVS (Verkehr)	2.411,58	4.558,18	89,00%	2.422,22 0,40%	2.417,26 0,20%	2.452,11 1,70%	2.423,68 0,50%	2.417,26 0,20%	2.452,11 1,70%	34,85 1,45%
W08 Nachrichtenübermittlung	Standardabweichung Rel. Abw. zur EVS (Nachrichten)	158,21	156,24	-1,20%	127,07 -19,70%	131,57 -16,80%	132,08 -16,50%	128,63 -18,70%	127,07 -19,70%	132,08 -16,50%	5,01 3,17%
W09 Freizeit, Unterhaltung und Kultur	Standardabweichung Rel. Abw. zur EVS (Freizeit)	866,61	1.286,38	48,40%	848,31 -2,10%	878,54 1,40%	877,97 1,30%	858,92 -0,90%	848,31 -2,10%	878,54 1,40%	30,23 3,49%
W10 Bildungswesen	Standardabweichung Rel. Abw. zur EVS (Bildung)	177,1	314,94	77,80%	760,59 329,50%	816,42 361,00%	765,04 332,00%	780,94 341,00%	760,59 329,50%	816,42 361,00%	55,83 31,52%
W11 Begerbergungs- und Gaststättendienstleistungen	Standardabweichung Rel. Abw. zur EVS (Beherbg)	404,9	692,91	71,10%	467,57 15,50%	471,96 16,60%	472,72 16,70%	468,93 15,80%	467,57 15,50%	472,72 16,70%	5,15 1,27%
W12 Andere Waren und Dienstleistungen	Standardabweichung Rel. Abw. zur EVS (AndWarDI)	417,58	642,48	53,90%	433,86 3,90%	405,93 -2,80%	403,31 -3,40%	424,77 1,70%	403,31 -3,40%	433,86 3,90%	30,55 7,32%
W0312226 Damenhosen (lang, aus Baumwolle)	Standardabweichung Rel. Abw. zur EVS	n/a***	105,41		97,78	96,42	95,66	97,13	95,66	97,78	2,12
W0913011 PC-Komplettssysteme und Notebooks	Standardabweichung Rel. Abw. zur EVS	n/a***	1.671,28		1.630,60	1.606,43	1.575,59	1.617,93	1.575,59	1.630,60	55,01

* Daten gewichtet mit EVS-interner Variable "HRB" (führt zu N=35.899.946 Fällen)

** Kumuliertes File aus den LWR-Files 1999-2003 mit alternativen Vorgehensweisen und Kumulationsgewichten; Preisbereinigung auf 2003; Umgerechnet in Euro und Quartalswerte

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

Tabelle s2: Vergleich der Streuungen ausgewählter Konsumgüteraussgaben aus KUMLWR 2003 mit der EVS 2003 und der LWR 2003 – Hochrechnungsvariante II

Standardabweichung		EVS 2003* (Var. Name)	LWR 2003	LWR-EVS	Kumulationsfile 2003 **				Hochrechnungsvariante II		
					Ila	Ilb	Ilc	Ild	Minimum	Maximum	Max-Min
Summe W01 bis W12 (Gesamtausgaben)	Standardabweichung Rel. Abw. zur EVS	4.382,21	5.865,27		4.133,02 -5,70%	4.151,34 -5,30%	4.164,36 -5,00%	4.140,25 -5,50%	4.133,02 -5,70%	4.164,36 -5,00%	31,34 0,72%
W01 Nahrungsmittel und alkoholfreie Getränke	Standardabweichung Rel. Abw. zur EVS (NG)	450,02	436,95	-2,90%	398,01 -11,60%	402,25 -10,60%	400,15 -11,10%	399,35 -11,30%	398,01 -11,60%	402,25 -10,60%	4,24 0,94%
W02 Alkoholische Getränke und Tabakwaren	Standardabweichung Rel. Abw. zur EVS (AlkGT)	165,23	186,7	13,00%	178,46 8,00%	170,08 2,90%	170,28 3,10%	175,69 6,30%	170,08 2,90%	178,46 8,00%	8,38 5,07%
W03 Bekleidung und Schuhe	Standardabweichung Rel. Abw. zur EVS (BeklSch)	339,62	425,26	25,20%	327,83 -3,50%	326,93 -3,70%	322,67 -5,00%	327,43 -3,60%	322,67 -5,00%	327,83 -3,50%	5,16 1,52%
W04 Wohnung, Wasser, Strom, u.a.	Standardabweichung Rel. Abw. zur EVS (W04selbst)	1.431,80	1.011,63	-29,30%	881,23 -38,50%	883,24 -38,30%	870,13 -39,20%	881,23 -38,50%	870,13 -39,20%	883,24 -38,30%	13,11 0,92%
W05 Einrichtungsgegenst., Apparate, Geräte, u.a.	Standardabweichung Rel. Abw. zur EVS (W05selbst)	892,86	1.603,00	79,50%	1.024,72 14,80%	1.028,66 15,20%	979,46 9,70%	1.021,20 14,40%	979,46 9,70%	1.028,66 15,20%	49,2 5,51%
W06 Gesundheitspflege	Standardabweichung Rel. Abw. zur EVS (Gesundpflege)	796,36	1.392,08	74,80%	669,35 -15,90%	714,14 -10,30%	746,03 -6,30%	685,58 -13,90%	669,35 -15,90%	746,03 -6,30%	76,67 9,63%
W07 Verkehr	Standardabweichung Rel. Abw. zur EVS (Verkehr)	2.411,58	4.558,18	89,00%	2.425,31 0,60%	2.429,86 0,80%	2.466,34 2,30%	2.430,00 0,80%	2.425,31 0,60%	2.466,34 2,30%	41,03 1,70%
W08 Nachrichtenübermittlung	Standardabweichung Rel. Abw. zur EVS (Nachrichten)	158,21	156,24	-1,20%	128,65 -18,70%	132,67 -16,10%	132,79 -16,10%	130,03 -17,80%	128,65 -18,70%	132,79 -16,10%	4,15 2,62%
W09 Freizeit, Unterhaltung und Kultur	Standardabweichung Rel. Abw. zur EVS (Freizeit)	866,61	1.286,38	48,40%	856,57 -1,20%	883,17 1,90%	880,4 1,60%	866,12 -0,10%	856,57 -1,20%	883,17 1,90%	26,6 3,07%
W10 Bildungswesen	Standardabweichung Rel. Abw. zur EVS (Bildung)	177,1	314,94	77,80%	723,8 308,70%	780,63 340,80%	742,42 319,20%	744,12 320,20%	723,8 308,70%	780,63 340,80%	56,83 32,09%
W11 Bevergungs- und Gaststättendienstleistungen	Standardabweichung Rel. Abw. zur EVS (Beherbg)	404,9	692,91	71,10%	469,63 16,00%	475,01 17,30%	475,17 17,40%	471,26 16,40%	469,63 16,00%	475,17 17,40%	5,54 1,37%
W12 Andere Waren und Dienstleistungen	Standardabweichung Rel. Abw. zur EVS (AndWarDI)	417,58	642,48	53,90%	448,36 7,40%	412,14 -1,30%	406,83 -2,60%	437,43 4,80%	406,83 -2,60%	448,36 7,40%	41,53 9,95%
W0312226 Damenhosen (lang, aus Baumwolle)	Standardabweichung Rel. Abw. zur EVS	n/a***	105,41		97,59	96,37	95,6	96,98	95,6	97,59	1,99
W0913011 PC-Komplettsysteme und Notebooks	Standardabweichung Rel. Abw. zur EVS	n/a***	1.671,28		1.634,80	1.606,20	1.571,39	1.620,26	1.571,39	1.634,80	63,41

* Daten gewichtet mit EVS-interner Variable "HRB" (führt zu N=35.899.946 Fällen)

** Kumuliertes File aus den LWR-Files 1999-2003 mit alternativen Vorgehensweisen und Kumulationsgewichten; Preisbereinigung auf 2003; Umgerechnet in Euro und Quartalswerte

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

Tabelle s3: Vergleich der Streuungen ausgewählter Konsumgüterausgaben aus KUMLWR 2003 mit der EVS 2003 und der LWR 2003 – Hochrechnungsvariante III

Standardabweichung		EVS 2003* (Var. Name)	LWR 2003	LWR-EVS	Kumulationsfile 2003 **						
					Hochrechnungsvariante III				Minimum	Maximum	Max-Min
					IIIa	IIIb	IIIc	IIId			
Summe W01 bis W12 (Gesamtausgaben)	Standardabweichung Rel. Abw. zur EVS	4.382,21	5.865,27		4.126,40 -5,80%	4.136,73 -5,60%	4.151,49 -5,30%	4.130,94 -5,70%	4.126,40 -5,80%	4.151,49 -5,30%	25,09 0,57%
W01 Nahrungsmittel und alkoholfreie Getränke	Standardabweichung Rel. Abw. zur EVS	450,02 (NG)	436,95	-2,90%	397,92 -11,60%	403,58 -10,30%	402,02 -10,70%	399,75 -11,20%	397,92 -11,60%	403,58 -10,30%	5,66 1,26%
W02 Alkoholische Getränke und Tabakwaren	Standardabweichung Rel. Abw. zur EVS	165,23 (AlkGT)	186,7	13,00%	181,34 9,70%	171,06 3,53%	170,78 3,40%	178,04 7,80%	170,78 3,40%	181,34 9,70%	10,56 6,39%
W03 Bekleidung und Schuhe	Standardabweichung Rel. Abw. zur EVS	339,62 (BekISch)	425,26	25,20%	327,98 -3,40%	326,45 -3,90%	322,15 -5,10%	327,43 -3,60%	322,15 -5,10%	327,98 -3,40%	5,83 1,72%
W04 Wohnung, Wasser, Strom, u.a.	Standardabweichung Rel. Abw. zur EVS	1.431,80 (W04selbst)	1.011,63	-29,30%	882,59 -38,40%	882,98 -38,30%	870,24 -39,20%	882,22 -38,40%	870,24 -39,20%	882,98 -38,30%	12,75 0,89%
W05 Einrichtungsgegenst., Apparate, Geräte, u.a.	Standardabweichung Rel. Abw. zur EVS	892,86 (W05selbst)	1.603,00	79,50%	1.010,76 13,20%	1.004,77 12,50%	959,44 7,50%	1.003,88 12,40%	959,44 7,50%	1.010,76 13,20%	51,32 5,75%
W06 Gesundheitspflege	Standardabweichung Rel. Abw. zur EVS	796,36 (Gesundpflege)	1.392,08	74,80%	674,81 -15,30%	717,5 -9,90%	749,03 -5,90%	691,12 -13,20%	674,81 -15,30%	749,03 -5,90%	74,22 9,32%
W07 Verkehr	Standardabweichung Rel. Abw. zur EVS	2.411,58 (Verkehr)	4.558,18	89,00%	2.424,60 0,50%	2.420,19 0,40%	2.455,25 1,80%	2.426,26 0,60%	2.420,19 0,40%	2.455,25 1,80%	35,06 1,45%
W08 Nachrichtenübermittlung	Standardabweichung Rel. Abw. zur EVS	158,21 (Nachrichten)	156,24	-1,20%	127,63 -19,30%	132,2 -16,40%	132,76 -16,10%	129,22 -18,30%	127,63 -19,30%	132,76 -16,10%	5,13 3,24%
W09 Freizeit, Unterhaltung und Kultur	Standardabweichung Rel. Abw. zur EVS	866,61 (Freizeit)	1.286,38	48,40%	848,95 -2,00%	878,31 1,40%	876,79 1,20%	859,31 -0,80%	848,95 -2,00%	878,31 1,40%	29,36 3,39%
W10 Bildungswesen	Standardabweichung Rel. Abw. zur EVS	177,1 (Bildung)	314,94	77,80%	767,56 333,40%	822,95 364,70%	773,61 336,80%	788,02 345,00%	767,56 333,40%	822,95 364,70%	55,39 31,27%
W11 Bevergungs- und Gaststättendienstleistungen	Standardabweichung Rel. Abw. zur EVS	404,9 (Beherbg)	692,91	71,10%	467,54 15,50%	472,22 16,60%	472,96 16,80%	468,98 15,80%	467,54 15,50%	472,96 16,80%	5,42 1,34%
W12 Andere Waren und Dienstleistungen	Standardabweichung Rel. Abw. zur EVS	417,58 (AndWarDI)	642,48	53,90%	436,48 4,50%	407,13 -2,50%	404,03 -3,20%	427,05 2,30%	404,03 -3,20%	436,48 4,50%	32,44 7,77%
W0312226 Damenhosen (lang, aus Baumwolle)	Standardabweichung Rel. Abw. zur EVS	n/a***	105,41		97,9	96,58	95,81	97,26	95,81	97,9	2,09
W0913011 PC-Komplettsysteme und Notebooks	Standardabweichung Rel. Abw. zur EVS	n/a***	1.671,28		1.627,26	1.601,44	1.571,55	1.614,02	1.571,55	1.627,26	55,7

* Daten gewichtet mit EVS-interner Variable "HRB" (führt zu N=35.899.946 Fällen)

** Kumuliertes File aus den LWR-Files 1999-2003 mit alternativen Vorgehensweisen und Kumulationsgewichten; Preisbereinigung auf 2003; Umgerechnet in Euro und Quartalswerte

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

Tabelle s4: Vergleich ausgewählter Konsumgüteraussgaben aus KUMLWR 2003 mit der EVS 2003 und LWR 2003 – Standardabweichungen: Hochrechnungsvarianten I, II, III

Standardabweichung		EVS 2003* (Var.-Name)	LWR 2003	LWR-EVS	KUMLWR Kumulations- gewichte c,a	2003 **			Minimum	Maximum	Max-Min
						I	II	III			
Summe W01 bis W12 (Gesamtausgaben)	Mittelwert	4.382,21	5.865,27			4.145,27	4.164,36	4.151,49	4.145,27	4.164,36	19,09
	Rel. Abw. zur EVS			33,80%	c	-5,40%	-5,00%	-5,30%	-5,40%	-5,00%	0,44%
W01 Nahrungsmittel und alkoholfreie Getränke	Mittelwert	450,02 (NG)	436,95			399,66	400,15	402,02	399,66	402,02	2,36
	Rel. Abw. zur EVS			-2,90%	c	-11,20%	-11,10%	-10,70%	-11,20%	-10,70%	0,52%
W02 Alkoholische Getränke und Tabakwaren	Mittelwert	165,23 (AlkGT)	186,7			170,47	170,28	170,78	170,28	170,78	0,5
	Rel. Abw. zur EVS			13,00%	c	3,20%	3,10%	3,40%	3,10%	3,40%	0,30%
W03 Bekleidung und Schuhe	Mittelwert	339,62 (BekISch)	425,26			321,8	322,67	322,15	321,8	322,67	0,87
	Rel. Abw. zur EVS			25,20%	c	-5,20%	-5,00%	-5,10%	-5,20%	-5,00%	0,26%
W04 Wohnung, Wasser, Strom, u.a.	Mittelwert	1.431,80 (W04selbst)	1.011,63			866,02	870,13	870,24	866,02	870,24	4,21
	Rel. Abw. zur EVS			-29,30%	c	-39,50%	-39,20%	-39,20%	-39,50%	-39,20%	0,29%
W05 Einrichtungsgegenst., Apparate, Geräte, u.a.	Mittelwert	892,86 (W05selbst)	1.603,00			961,66	979,46	959,44	959,44	979,46	20,02
	Rel. Abw. zur EVS			79,50%	c	7,70%	9,70%	7,50%	7,50%	9,70%	2,24%
W06 Gesundheitspflege	Mittelwert	796,36 (Gesundpflege)	1.392,08			683,62	669,35	674,81	669,35	683,62	14,27
	Rel. Abw. zur EVS			74,80%	a	-14,20%	-15,90%	-15,30%	-15,90%	-14,20%	1,74%
W07 Verkehr	Mittelwert	2.411,58 (Verkehr)	4.558,18			2.422,22	2.425,31	2.424,60	2.422,22	2.425,31	3,09
	Rel. Abw. zur EVS			89,00%	a	0,40%	0,60%	0,50%	0,40%	0,60%	0,16%
W08 Nachrichtenübermittlung	Mittelwert	158,21 (Nachrichten)	156,24			132,08	132,79	132,76	132,08	132,79	0,71
	Rel. Abw. zur EVS			-1,20%	c	-16,50%	-16,10%	-16,10%	-16,50%	-16,10%	0,45%
W09 Freizeit, Unterhaltung und Kultur	Mittelwert	866,61 (Freizeit)	1.286,38			877,97	880,4	876,79	876,79	880,4	3,62
	Rel. Abw. zur EVS			48,40%	c	1,30%	1,60%	1,20%	1,20%	1,60%	0,42%
W10 Bildungswesen	Mittelwert	177,1 (Bildung)	314,94			765,04	742,42	773,61	742,42	773,61	31,19
	Rel. Abw. zur EVS			77,80%	c	332,00%	319,20%	336,80%	319,20%	336,80%	17,61%
W11 Begerbergungs- und Gaststättendienstleistungen	Mittelwert	404,9 (Beherbg)	692,91			472,72	475,17	472,96	472,72	475,17	2,45
	Rel. Abw. zur EVS			71,10%	c	16,70%	17,40%	16,80%	16,70%	17,40%	0,61%
W12 Andere Waren und Dienstleistungen	Mittelwert	417,58 (AndWarDI)	642,48			433,86	448,36	436,48	433,86	448,36	14,5
	Rel. Abw. zur EVS			53,90%	a	3,90%	7,40%	4,50%	3,90%	7,40%	3,47%

* Daten gewichtet mit EVS-interner Variable "HRB" (führt zu N=35.899.946 Fällen)

** Kumuliertes File aus den LWR-Files 1999-2003 mit alternativen Vorgehensweisen und Kumulationsgewichten; Preisbereinigung auf 2003; Umgerechnet in Euro und Quartalswerte

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

Tabelle 9: Streuungsvergleich: Häufigkeiten der besten Kumulationsgewichtungen der Gesamtausgaben und aller elf Ausgabekategorien* von KUMLWR 2003 im Vergleich zur EVS 2003

Häufigkeiten der geringsten Streuungsabweichungen			
Kumulationsgewicht	Hochrechnungsvariante		
	I	II	III
a (Gleichgewichtung)	3	3	3
b (linear progressiv)	2	4	2
c (exponentiell)	5	4	5
d (datengeneriert)	2	1	2

*ohne Bildungsausgaben

Quelle: Kumulationsfile KUMLWR 2003 aus laufenden Wirtschaftsrechnungen 1999, 2001, 2002, 2003; Eigene Berechnungen

Allerdings sind die Streuungsunterschiede in KUMLWR 2003 zwischen allen Kumulationsgewichtungen doch gering (Tabellen s1, s2, s3 und s4). Bei den meisten Ausgabekategorien (9 von 12) sind zudem alle KUMLWR 2003-Streuungen gegenüber den nicht kumulierten LWR 2003-Streuungen den EVS-Streuungen wesentlich näher.

Halten wir also fest: Auch bezüglich der Streuungen in den einzelnen Ausgabekategorien gibt es einen Informationsgewinn durch die Kumulation der Stichproben.

Folgen wir den besten Kumulationsalternativen aus dem Mittelwertvergleich – exponentielle Gewichtung (c) und Gleichgewichtung (a) durchgängig für alle Hochrechnungsvarianten I, II und III (dort Tabelle m4) – so wird mit entsprechender Tabelle s4 wieder deutlich, dass zunächst auch hinsichtlich der Streuungen die Unterschiede zwischen allen Hochrechnungsvarianten gering sind (alle Spannweiten liegen zwischen 0,16 und 3,5 Prozentpunkten).

Die Streuungs-Anpassungsgüte der Kumulationsalternativen ist allerdings wieder abhängig von den einzelnen Ausgabekategorien. Es sind die Ausgaben für Gesundheitspflege und andere Waren und Dienstleistungen sowie hier auch Einrichtungsgegenstände, die im Vergleich zu allen anderen Warengruppen etwas weniger gut abschneiden. Allerdings liegen mit einer maximalen Abweichung von 7,4% die KUMLWR 2003-Werte immer noch sehr nahe bei den EVS-Werten. Weitere Einzelergebnisse finden sich dazu in Tabelle 10.

Die hinsichtlich der Streuung besten Kumulationsgewichte sind zusammengefasst

- c: exponentiell mit starker Gewichtung der aktuellen Stichprobe und
- a: gleiche Gewichtung aller Stichproben je nach Ausgabekategorie.

Welche Hochrechnungsvariante empfiehlt sich?

Ausgabekategorien: Vergleich der Streuungen

Wie bereits angesprochen, liegen alle drei Hochrechnungsvarianten (I, II und III) im Streuungsvergleich sehr nahe beieinander (Tabelle s4); ein dem Mittelwertvergleich ähnliches Ergebnis.

Wieder eingedenk des erforderlichen Erstellungsaufwandes, der sehr nahen Streuungsergebnisse zwischen den Hochrechnungsvarianten sowie der zudem besten Ergebnissen in mehreren Ausgabekategorien empfiehlt sich damit die

Hochrechnungsvariante III: Kumulationsgewichtung mit alternativen Abschreibungsraten nach der abschließenden Hochrechnung des gesamten Kumulationsfiles.

***KUMLWR 2003 besser als LWR 2003 im Vergleich zur EVS 2003?
Vergleich der Streuungen der einzelnen Ausgabekategorien***

Auffällig sind die relativ großen prozentualen Abweichungen zwischen der EVS 2003 und der alleinigen LWR 2003; die Abweichungen erreichen Unterschiede bis 89% (Verkehr, Tabelle s4). Für Verkehr beispielsweise liegen die Abweichungen aller Kumulationsalternativen aus KUMLWR 2003 unter 1%. Auch für andere Ausgabekategorien gilt, wenn auch nicht immer in so frappanter Weise, dass durch die Kumulation aus KUMLWR 2003 ein deutlicher Informationsgewinn gegeben ist.

Die Berücksichtigung vergangener Querschnitte verbessert also auch hinsichtlich der Streuung die Ergebnisse der alleinigen laufenden Wirtschaftsrechnung und spricht deutlich für eine Kumulation.

8 Zusammenfassende Beurteilung und Ausblick

Im Rahmen des Projektes „Amtliche Statistik und sozioökonomische Fragestellungen“ des Statistischen Bundesamtes und eingebunden in die EU-SILC-Ansätze haben wir für die Kumulation von laufenden Haushaltsbudgetbefragungen ein Konzept entwickelt und dieses Kumulationskonzept mit einer Kumulation von fünf (beziehungsweise vier) Querschnitten der laufenden Wirtschaftsrechnungen aus 1999, 2001, 2002, und 2003 praktisch umgesetzt und evaluiert. Ziel ist es, vertiefte Strukturdaten aus einer geeigneten Zusammenführung zu ermitteln und einen möglichen Informationsgewinn gegenüber anderen Großstichproben zu analysieren (Ehling 2002). Dafür vergleichen wir die die Ausgaben der kumulierten Stichprobe KUMLWR 2003 mit den Ausgaben der Einkommens- und Verbrauchsstichprobe 2003 (EVS 2003) als Großstichprobe sowie auch mit den Ausgaben der alleinigen laufenden Wirtschaftsrechnung LWR zum Zeitpunkt der Großstichprobe 2003.

Die theoretischen Grundlagen und Bausteine dazu wurden in Merz 2004 diskutiert und hier weiterentwickelt. Neben der Entwicklung geeigneter Kumulationsgewichte, die den Informationsgehalt vergangener Querschnitte unterschiedlich einbeziehen, wurde die zentrale Aufgabe einer strukturellen neuen demographischen Gewichtung mit einem Hochrechnungs-/Kalibrierungsansatz auf informationstheoretischer Basis (Minimum Information Loss (MIL)-Prinzip) gelöst. Der besondere Vorteil dieser Methodik liegt in der theoriegeleiteten repräsentativen Strukturübertragung und Strukturangepassung auch bezüglich hierarchischer Stichproben mit dem simultanen Ansatz. Neben der strukturellen Gewichtung nach demographischen Aspekten werden ‚economic multipliers‘ (‚inflatoren‘) für die Berücksichtigung von Preiseinflüssen sowie eine itembezogene Relativierung (Relativierung ökonomischer Größen auf periodenbezogene Mittelwerte) einbezogen und möglich.

Wir haben drei Varianten aus unterschiedlicher Hochrechnung (I, II, III)

- (I) Kumulationsgewichtung vor (abschließender) Hochrechnung mit vorheriger Anpassung an laufende Rahmendaten r_t

(II) Kumulationsgewichtung *vor* (abschließender) Hochrechnung
ohne vorheriger Anpassung an laufende Rahmendaten r_t

(III) Kumulationsgewichtung *nach* Hochrechnung
mit Anpassung nur an Rahmendaten r_T zum Zeitpunkt T

mit jeweils vier alternativen Kumulationsgewichtungen

a Gleichgewichtung, b linear progressiv, c exponentiell, d datengeneriert

untersucht. Die somit insgesamt zwölf Kumulationsalternativen ordnen jedem einzelnen Haushalt in jedem betrachteten Jahr ein jeweils individuelles Gewicht im aggregierten Kumulationsfile zu. Das Kumulationsfile KUMLWR 2003 enthält additiv alle fünf beziehungsweise vier laufenden Wirtschaftsrechnungen von 1999 bis 2003 mit 30.480 bzw. 24.311 Datensätzen.

Fazit: Vergleich KUMLWR 2003 mit der EVS 2003 und der LWR 2003

In Anlehnung an die Effizienz eines Schätzers (Erwartungstreue und minimale Varianz) als wünschenswerte Schätzeigenschaft haben wir als Beurteilungskriterium der Anpassungsgüte des kumulierten Files KUMLWR 2003 an die EVS 2003 und die alleinige LWR 2003 das arithmetische Mittel sowie die Standardabweichung für 12 ausgewählte Konsumausgabekategorien gewählt.

Sowohl für die Gesamtausgaben des privaten Verbrauchs als auch für die einzelnen Ausgabekategorien liegen alle zwölf Kumulationsalternativen sehr nahe beieinander und können die EVS 2003 besser erreichen als die alleinige LWR 2003. Das gilt sowohl für den Vergleich der jeweiligen Mittelwerte als auch durchgängig für den Vergleich der jeweiligen Streuungen. Über die Kumulation ist also ein so gemessener deutlicher Informationsgewinn festzuhalten.

Kumulationsgewichte: Mit allen drei Hochrechnungsvarianten (I, II und III) wird, sowohl nach dem Vergleich der Mittelwerte als auch der Streuungen, ein identisches Muster der besten Kumulationsgewichtung mit *c: exponentiell* und *a: gleiche Gewichtung* erreicht. Die Anpassungsgüte der Kumulationsalternativen hinsichtlich der Kumulationsgewichte ist somit abhängig von den Ausgabegruppen.

Eine gleiche Gewichtung der zeitlichen Abschreibungsraten (Kumulationsgewichte) – und damit eine stärkere Gewichtung auch der ältesten Information (LWR 1999) – führt für die Kategorien Gesundheitspflege, Verkehr und andere Waren und Dienstleistungen zu besseren Ergebnissen als für alle anderen neun Ausgabekategorien, für die eine starke zeitnahe exponentielle Gewichtung spricht.

Hochrechnungsvarianten: Alle drei Hochrechnungsvarianten mit ihren 4 alternativen Kumulationsgewichten liegen im Ergebnis nahe beieinander. Wird als zusätzliches Beurteilungskriterium der Erstellungsaufwand herangezogen, so ist die Hochrechnungsvariante III, die zuerst eine demographische Hochrechnung des gesamten Kumulationsfiles zum Zeitpunkt T=2003 vornimmt und dann erst jeden Querschnitt mit alternativen Kumulationsgewichten versieht, die günstigste Hochrechnungsvariante. Sie erlaubt auch eine nachträgliche andere Querschnittsgewichtung, ohne dass mit jeder neuen Kumulationsgewichtung auch eine neue Hochrechnung (Varianten I und II) nötig wird.

Diese Ergebnisse werden durch die Analyse der Mittelwert- als auch der Streuungsabweichungen getragen.

Fazit: Methodenvergleich alternativer Kumulationsfaktoren

Ein Kumulationsfaktor setzt sich aus dem Kumulationsgewicht für den einzelnen Querschnitt und der Hochrechnungsvariante zusammen und gibt schließlich an, wie viele Mikroeinheiten (hier Haushalte) in der Grundgesamtheit (hier Deutschland) durch einen Stichprobenhaushalt (hier KUMLR 2003) repräsentiert werden. Ergebnis: Bei allen zwölf Kumulationsalternativen liegen die Kumulationsfaktoren relativ eng zwischen dem ersten und dritten Quartil beieinander. Unterschiedlich hohe Maximalwerte weisen aber auf notwendige Anpassungen von stark unterrepräsentierten Gruppen in der Stichprobe hin.

Unsere Kumulationsgewichte haben gesetzte Verfahren (a Gleichgewichtung, b linear progressiv, c exponentiell) sowie ein datengeneriertes Verfahren einbezogen. Das datengenerierte Verfahren (ohne modellgestützte Ansätze) wäre eigentlich aufgrund seiner theoretischen Fundierung zu bevorzugen. Allerdings sind die dafür notwendigen clusteranalytischen Berechnungen relativ aufwendig. Als datengeneriertes Verfahren bestimmen die Daten selbst die zu wählende Abschreibungsgewichtung, also die Bewertung zeitlich zurückliegender Daten. Demgegenüber zeigen allerdings die gesetzten Alternativen transparent das Bewertungsschema des Betrachters an.

Fazit generell

Die Hochrechnungsvariante III, eine Hochrechnung/Anpassung des Kumulationsfiles an die Rahmendaten vor einer Kumulationsgewichtung, hat sich als beste Kumulationsalternative für die Abschätzung der Ausgaben privater Haushalte herausgestellt. Die Anpassungsgüte hinsichtlich der Kumulationsgewichte ist abhängig von den gewählten Ausgabegruppen; die exponentielle Gewichtung (c) mit hoher Bewertung der aktuellsten Informationen aber auch die Gleichgewichtung (a) vergangener Informationen schneiden am besten ab.

Mit dem nach unserer Analyse zu präferierenden Kumulationsalternative (Hochrechnungsvariante III) ist es schließlich mit Verwendung einer mächtigen und effizienten Hochrechnungsprozedur – wie die von uns verwendete nach dem MIL-Prinzip und der zugehörigen ADJUST-Software (<http://ffb.uni-lueneburg.de/adjust>) – relativ einfach, ein aggregiertes Kumulationsfile der aktuellen demographischen Struktur anzupassen. Nach einer solchen demographischen Hochrechnung des kumulierten Files können dann die individuellen Hochrechnungsfaktoren relativ leicht mit den gewählten alternativen Kumulationsgewichten zu individuellen abschließenden Hochrechnungsfaktoren modifiziert werden. Damit können dann die inhaltlich interessierenden Variablen (hier die Ausgaben privater Haushalte) abhängig von der gewählten Kumulationsgewichtung analysiert werden.

Mit unserem Kumulationskonzept hat sich ein deutlicher Informationsgewinn durch die Kumulation von Stichproben mit ihrer Berücksichtigung aufeinanderfolgender Stichproben ergeben. Die Ergebnisse haben aber auch einen Hinweis darauf gegeben, dass eine modellgestützte inhaltliche Fortschreibung auf der Basis zeitlich vorangegangener Informationen zu weiteren Informationsgewinnen führen wird. Das bleibt aber der inhaltlichen Analyse wie die des Kaufverhaltens für einzelne Ausgabegruppen des privaten Verbrauchs vorbehalten. Mit individuellen Paneldaten dürften sich schließlich weitere Möglichkeiten einer inhaltlich zielgenauen Mikroanalyse ergeben.

9 Anhang

Tabelle A1: Hochrechnungs-Rahmendaten aus den Mikrozensus 1999, 2000, 2001, 2002 und 2003

Hochrechnungs-Rahmendaten (r)	Mikrozensus 1999	Mikrozensus 2000	Mikrozensus 2001	Mikrozensus 2002	Mikrozensus 2003
Soziale Stellung					
Beamte	1692700	1652500	1614500	1619000	1606400
Angestellte	10014400	10020300	10646500	10795600	10800600
Arbeiter	7579300	7731400	7274100	7034100	6781800
Rentner	11908700	12180200	12309200	12419500	12557400
Privathaushalte* nach Haushaltstyp					
Einpersonenhaushalte	12956200	13204100	13504900	13658000	13848300
HH mit 2 Personen	12419200	12571300	12751400	12897900	13007900
HH mit 3 Personen	5616500	5565400	5470700	5454300	5428900
HH mit 4 und mehr Personen	6090600	6035500	5972900	5929600	5864700
Privathaushalte nach der Zahl der Erwerbstätigen im Haushalt					
keine Erwerbstätigen	13708400	13877200	14053700	14380400	14760300
ein Erwerbstätiger	12604100	12673900	12782300	12785300	12758600
zwei Erwerbstätige	9001800	9042400	9069100	9050200	8953600
Bevölkerung nach Geschlecht und Alter					
Männer					
0 bis unter 20 Jahre	8804900	8793200	8687100	8666200	8589100
20 bis unter 30 Jahre	4760800	4631800	4576100	4488200	4562700
30 bis unter 40 Jahre	6717000	6694300	6581600	6467100	6258600
40 bis unter 50 Jahre	5825500	5884700	6013000	6169200	6299700
50 bis unter 60 Jahre	5312500	5183700	5151700	5055800	5046600
60 bis unter 65 Jahre	2740900	2899600	2940800	2989700	2888100
65 Jahre und älter	5210100	5403000	5610300	5860700	6135400
Frauen					
0 bis unter 20 Jahre	8498800	8428900	8372800	8331900	8239400
20 bis unter 30 Jahre	4646500	4505600	4436600	4428400	4475300
30 bis unter 40 Jahre	6516200	6512100	6444600	6329200	6126200
40 bis unter 50 Jahre	5901000	5962800	6041800	6146100	6241000
50 bis unter 60 Jahre	5310600	5213900	5184100	5137800	5177700

Hochrechnungs-Rahmendaten (r)	Mikrozensus 1999	Mikrozensus 2000	Mikrozensus 2001	Mikrozensus 2002	Mikrozensus 2003
60 bis unter 65 Jahre	2887000	3019700	3068400	3111700	3038400
65 Jahre und älter	8072800	8232300	8356500	8507300	8655400
<i>Sonst. Nichterwerbstätige (West)</i>	<i>2506800</i>	<i>2354400</i>	<i>2355600</i>	<i>2525900</i>	<i>2761400</i>
<i>Sonst. Nichterwerbstätige (Ost)</i>	<i>745600</i>	<i>752400</i>	<i>812400</i>	<i>834200</i>	<i>877000</i>
<i>HH mit drei und mehr Erwerbstätigen</i>	<i>1768100</i>	<i>1782900</i>	<i>1794800</i>	<i>1724000</i>	<i>1677300</i>

* Privathaushalte am Ort der Hauptwohnung

Quelle: Statistisches Bundesamt, Sonderauswertungen des Mikrozensus für das Projekt „Kumulierung von Haushaltsbudgetdaten“

Tabelle A2: EVS 2003 - Generierung der Ausgabenkategorien für den Vergleich mit den laufenden Wirtschaftsrechnungen und dem kumulierten KUMLWR 2003 File

EVS 2003		
Abgegrenzte Größe	Ausgabenart	Code-Nrn.
Private Konsumausgaben =		lt. Code-Verzeichnis 01 bis 12
Nahrungsmittel, Getränke, Tabakwaren =		01, 02
	Nahrungsmittel, alkoholfreie Getränke	0110000H bis 0120000H
	+ Alkoholische Getränke, Tabakwaren u.ä.	0210000H bis 0230000H
Bekleidung, Schuhe =		03
	Σ Bekleidung, Schuhe	0311000H bis 0322000H
Wohnen, Wohnungsinstandhaltung =		041 bis 044
	Tatsächliche Mietzahlungen (Mieter, Unterm)	0411040H bis 0412900H
	+ Unterstellte Mietzahlungen (Deputat u.ä.)	0422040H
	+ Unterstellte Mietzahlungen (Eigentümer)	0421031H bis 0421033H, 0422050H
	+ Laufende Kosten für selbstgen. Eigentum	0445900H, 0445901H
	+ Instandhaltung und Reparatur	0431001H bis 0432900H
Energie =		045
	Σ Strom, Gas und andere Brennstoffe	0451010H bis 0456000H
Innenausstattung, Haushaltsgeräte u.ä. =		05
	Σ Möbel, Innenausstattung, Teppiche u.ä.	0511900H bis 0562000H
Gesundheitspflege =		06
	Σ Gesundheitspflege	0611010H bis 0630000H

Abgegrenzte Größe	Ausgabenart	Code-Nrn.
		It. Code-Verzeichnis
Verkehr =		07
	Σ Verkehr	0711100H bis 0733200H
Nachrichtenübermittlung =		08
	Σ Nachrichtenübermittlung	0810000H bis 0830900H
Freizeit, Unterhaltung, Kultur =		09
	Σ Freizeit, Unterhaltung, Kultur	0911100H bis 0962000H
Bildungswesen =		10
	Σ Bildungswesen	1010900H bis 1050900H
Beherbergungs- und Gaststättendienstleistungen =		11
	Σ Beherbergungs- und Gaststättendienstl.	1111000H bis 1120000H
Andere Waren und Dienstleistungen =		12
	Σ Andere Waren und Dienstleistungen	1211010H bis 1270900H
Andere Ausgaben (ohne Abzüge vom Einkommen)		
	Abzüge vom Einkommen	siehe ,Abgegrenzte Größen/ Einnahmen
Sonstige Steuern =		
	Erbschafts- und Schenkungssteuer	1512100H
	+ Sonstige Steuern, Fehlbelegungsabgabe	1512101H bis 1512900H
Versicherungsbeiträge =		
	Freiwillige Beiträge zu Sozialversicherungen	1522400P bis 1522800P
	+ Freiw. Beiträge und Beiträge zur privaten KV	1531000P, 1532100P
	+ Prämien für private Versicherungen	1532900H bis 1539906H
Sonstige geleistete Übertragungen =		
	Lohn-, Gehaltspfändungen (personenbez.)	1549002P
	+ Sonstige Übertragungen	1541000H bis 1549001H
	+ Sonstige Ausgaben	1591000H bis 1599000H

Abgegrenzte Größe	Ausgabenart	Code-Nrn.
		It. Code-Verzeichnis
Ausgaben für die Vermögensbildung	=	
	Ausgaben für die Bildung von Sachverm.	1561000H bis 1567000H
	+ Ausgaben für die Bildung von Geldverm.(P)	1581900P bis 1587900P
	+ Ausgaben für die Bildung von Geldverm.(H)	1579301H, 1579900H
	+ Ausgaben für die Bildung von Geldverm.(H)	1581100H bis 1588001H
Rückzahlung von Krediten	=	
	Tilgung und Verzinsung von Baudarlehen	1556000H
	+ Tilgung und Verzinsung von Konsumentenkr.	1559000H
	+ Überziehungszinsen (Dispokreditzinsen)	1554500H
Gesamtausgaben	=	
	Private Konsumausgaben	
	+ Abzüge vom Einkommen	
	+ Andere Ausgaben	
Ersparnis	=	
	Ausgaben für die Vermögensbildung	
	+ Rückzahlung von Krediten	
	- Einn. aus der Auflösung von Sachvermögen	0060100H bis 0060300H
	- Einn. aus der Auflösung von Geldvermögen	0071100H bis 0078001H
		0077100P
	- Einnahmen aus Kreditaufnahme	0081000H bis 0085900H
	- Zinsen für Baudarlehen, Hypotheken	1556500H
	- Zinsen für Konsumenten- und Dispokredite	1559500H, 1554500H

Tabelle A3: Dateiübersicht Kumulation (alle Files sind auf Anfrage erhältlich)

Dateiname	Pfad	Beschreibung
lwrXX_auszug_051205.csv	\Daten\Mai 2006	Auszug LWR, Jahrgang XX Original-Dateien im ASCII-Format
lwrXX_auszug_051205 – mit HRF.sav	\Daten\Mai 2006	Auszug LWR, Jahrgang XX Bearbeitete SPSS-Datenfiles Enthalten zusätzlich: <ul style="list-style-type: none"> - HRF an Rahmendaten XX - HRF an Rahmendaten 2002 - Summe Basisausgabentypen w01-w12 - Indikator für Jahrgang (t)
spezifikat_lwr.doc	\Daten\Mai 2006	Beschreibung der LWR-Datensätze Datei von DeStatis
Verbraucherpreise jan 2004 2170700041015.xls	Daten\Mai 2006	Verbraucherpreisindizes Datei von DeStatis
r-vector lwrXXXX.adr	\kumulation\Auswertung 2006	Adjust r-Vektor mit Rahmendaten zu Jahr XXXX
Build S-matrix for lwrXXXX auszug.SPS	\kumulation\Auswertung 2006	SPSS-Syntax zur Erstellung der S-Matrix- Files für Jahrgang XX XX aus: \Daten\Mai 2006\lwrXX_auszug_051205 – mit HRF.sav
S-Matrix XXXX q.ads	\kumulation\Auswertung 2006	Adjust S-Matrix-File mit ursprünglichen Gewichten der LWR als q für Jahrgang XXXX
S-Matrix psX.ads	\kumulation\Auswertung 2006	Adjust S-Matrix-File mit nach Methode X gealterten, zuvor hochgerechneten Ausgangsgewichten (p*; Vorehensweise I) als q für das aggregierte Datenfile der LWR
S-Matrix qwX.ads	\kumulation\Auswertung 2006	Adjust S-Matrix-File mit nach Methode X gealterten, ursprünglichen LWR- Ausgangsgewichten (qw; Hochrechnungsvariante II) als q für das aggregierte Datenfile der LWR
Adjust Logfile – [...].txt	\kumulation\Auswertung 2006	Adjust Logfile für alle 4 Hochrechnungsvorgänge, jeweils genau im Dateinamen spezifiziert: <ul style="list-style-type: none"> - ungealterte Gewichte an t - ungealterte Gewichte an T - gealterte Gewichte p* an T - gealterte Gewichte qw an T
Adjustment Statistics – [...].txt	\kumulation\Auswertung 2006	Adjust Statistiken zu den o.g. Hochrechnungsvorgängen, jeweils im Dateinamen genau spezifiziert

Dateiname	Pfad	Beschreibung
Adjustment Results – [...].txt	\\kumulation\Auswertung 2006	Adjust ASCII-Datei mit neuen Hochrechnungsfaktoren zu den o.g. Hochrechnungsvorgängen, jeweils im Dateinamen genau spezifiziert
Ausgabengruppen berechnen w01 bis w12 (für XX).SPS	\\kumulation\Auswertung 2006	SPSS Syntax zur Berechnung der Basisausgabengruppen w01 bis w12 für Jahrgang XX (da jeweils andere Zusammensetzung der Untergruppen für jeden Jahrgang separat)
Distanzmaße.doc	\\kumulation\Auswertung 2006	Word-Datei mit Dokumentation der Hochrechnungsvariante für Datengenerierte Alterung
Alterung.doc	\\kumulation\Auswertung 2006	Word-Datei mit Tabelle über Hochrechnungsvariante bezüglich Ablauf Alterung und Hochrechnung
Build S-matrix for Arbeitsfile.SPS	\\kumulation\Auswertung 2006	SPSS-Syntax zur Erstellung der S-Matrix-Files aus dem Arbeitsfile zur Berechnung von p_Ia ... p_Id sowie p_Ila ... p_Ild
Parameter der neuen Hochrechnungsfaktoren.doc	\\kumulation\Auswertung 2006	Text und Tabellen zum Vergleich der neuen Hochrechnungsfaktoren
Aggregate Arbeitsfile.SPS	\\kumulation\Arbeitsfile gesamt	SPSS-Syntax zur Aggregation der Jahrgänge in den Einzelfiles zu einem Arbeitsfile
Alterungsgewichte einfügen.SPS	\\kumulation\Arbeitsfile gesamt	SPSS-Syntax zum Einfügen der Alterungsgewichte entsprechend der Jahrgangskennung t nach den 4 Methoden a – d (→ w_a_gleich; w_b_degressiv; w_c_exponentiell; w_d_datengeneriert)
Alterungsstufe.SPS	\\kumulation\Arbeitsfile gesamt	SPSS-Syntax zur Alterung der jeweiligen Gewichte entsprechend den 3 Hochrechnungsvarianten und 4 Alterungsmethoden:
Arbeitsfile Auszug LVR 99 bis 02 aggregiert (Mittelwerte).sav	\\kumulation\Arbeitsfile gesamt	SPSS-Datenfile, alle Jahrgänge, Quartalsdaten über Mittelwertfunktion aggregiert.
Arbeitsfile Auszug LVR 99 bis 02.sav	\\kumulation\Arbeitsfile gesamt	SPSS-Datenfile, alle Jahrgänge, Quartalsdaten noch vorhanden
Euroumrechnung für t4.SPS	\\kumulation\Arbeitsfile gesamt	SPSS-Syntax zur Umrechnung von Euro in DM für die 12 Basisausgabengruppen in 2002
Mittelwerte w01-w12 aus allen Jahrgängen (Preisbereinigt).sav	\\kumulation\Arbeitsfile gesamt	SPSS-Datenfile mit Mittelwerten der 12 Basisausgabengruppen aus 4 Jahrgängen. Preisbereinigt entsprechend „V erbraucherpreise jan 2004 2170700041015.xl“; s.o.

Dateiname	Pfad	Beschreibung
Mittelwerte w01-w12 aus allen Jahrgängen und Preisbereinigung.xls	\kumulation\Arbeitsfile gesamt	Excel-Tabelle mit Berechnungen zur Ermittlung des datengenerierten Alterungsfaktors mit folgenden Tabellen: <ul style="list-style-type: none"> - Zusammenstellung aller alterungsfaktoren für alle 4 Jahrgänge und 4 Methoden - Übersicht Preisindizes der 12 Basisausgabengruppen - Mittelwerte der 12 Basisausgabengruppen aus 4 Jahrgängen, Originaldaten und Preisbereinigt
Min-Max-quart-Vergleich.xls	\kumulation\Arbeitsfile gesamt	Arbeitstabelle mit Minima, Maxima und Quartilen aus SPSS zur Übernahme in Word
Varianzvergleich.xls	\kumulation\Arbeitsfile gesamt	Arbeitstabelle mit Varianzen aus SPSS, berechneten F-Werten und berechneten Signifikanzen des Vrazianzgleichheitstests

Literatur

- Andreß, H.-J. und G. Lipsmeier (1995), *Was gehört zum notwendigen Lebensstandard und wer kann ihn sich leisten? Ein neues Konzept der Armutsmessung*, in: *Aus Politik und Zeitgeschehen*, B31-32, 35-49.
- Atkinson, A.B. (2003), *Multidimensional Deprivation: Contrasting Social Welfare and Counting Approaches*, in: *Journal of Economic Inequality*, 1, 51-65.
- Becker, Irene (2009), *Einkommens- und Vermögensverteilung*, in: Eicker-Wolf, K., Körzell, St., Nichoj, T. und A. Truger (Hrsg.), *In gemeinsamer Verantwortung. Die Sozial- und Wirtschaftspolitik der Großen Koalition 2005-2009*, Marburg: Metropolis-Verlag, 75-116.
- Bittman, M. (1999), *Social Participation and Family Welfare: The Money and Time Cost of Leisure*, SPRC Discussion Paper No. 95, Sydney.
- Bittman, M. und R.E. Goodin (1998), *An Equivalence Scale for Time*, SPRC Discussion Paper No. 85, Sydney.
- Bonke, J. und F. Gerstoft (2005), *Stress, Time Use and Gender*, in: *electronic Journal of Time Use Research*, Vol. 4, No. 1, 47-68.
- Bonke, J., Deding, M. und M. Lausten (2006), *Time and Money: Substitutes in Real Terms and Complements in Satisfaction*, The Levy Economics Institute, Economics Working Paper Archive wp_451, Annandale-on-Hudson, New York.
- Bonke, J., Deding, M. und M. Lausten (2009), *Time and Money - A Simultaneous Analysis of Men's and Women's Domain Satisfaction*, in: *Journal of Happiness Studies*, 10, 113-131.
- Bourguignon, F. und S.R. Chakravarty (1999), *A Family of Multidimensional Poverty Measures*, in: Slotte, D.-J. (Ed.), *Essays in Honor of Camilo Dagum*, Physica Verlag, Heidelberg, 331-344.
- Bourguignon, F. und S.R. Chakravarty (2003), *The Measurement of Multidimensional Poverty*, in: *Journal of Economic Inequality* 1, 1, 25-49.
- Bundesregierung (2005), *Lebenslagen in Deutschland – Der zweite Armuts- und Reichtumsbericht der Bundesregierung*, Berlin.
- Bundesregierung (2008), *Lebenslagen in Deutschland – Der dritte Armuts- und Reichtumsbericht der Bundesregierung*, Berlin.
- Cappellari, L. und S.P. Jenkins (2007), *Summarising Multiple Deprivation Indicators*, in: Jenkins, S.P. und J. Micklewright (2007), *Inequality and Poverty Re-examined*, Oxford University Press, Oxford, 166-184.
- Clark, A.E. und A.J. Oswald (1996), *Satisfaction and Comparison Income*, in: *Journal of Public Economics*, 61, 359-381.
- Clark, A.E., Frijters, P. und M. Shields (2008), *Relative Income, Happiness and Utility: An Explanation for the Easterlin Paradox and Other Puzzles*, in: *Journal of Economic Literature*, 46, 95-144.
- Deutsch, J. und J.G. Silber (2005), *Measuring Multidimensional Poverty: An Empirical Comparison of Various Approaches*, in: *Review of Income and Wealth*, Vol. 51, No. 1, 145-174.
- Diamond, P.A. (1994), *On Time – Lectures on Models of Equilibrium*, Cambridge University Press, Cambridge.
- Diener, E. (2009), *The science of well-being*, Springer, Dordrecht.

- Easterlin, R. A. (1974), *Does economic growth improve the human lot?*, in: Nations and Households in Economic Growth, eds. P. A. David und M. W. Reder, Academic Press, New York, 88-125.
- Fandel, G. (2005), *Produktion I: Produktions- und Kostentheorie*, 6th Edition, Berlin.
- Foster, J., Greer, J. und E. Thorbecke (1984), *A Class of Decomposable Poverty Measures*, in: *Econometrica*, Vol. 52, No. 3, 761-766.
- Freeman, R. (1978), *Job Satisfaction as an Economic Variable*, in: *American Economic Review*, Vol. 68, No. 2, 135-144.
- Frey, B.S. und A. Stutzer (2002), *What Can Economists Learn from Happiness Research?*, in: *Journal of Economic Literature*, Jg. 40, H. 2, 402-435.
- Frey, B.S. und A. Stutzer (2005), *Happiness Research: State and Prospects*, in: *Review of Social Economy*, Vol. LXII, No. 2, 207-228.
- Garhammer, M. (2002a), *Zeitarbeit - ein Muster für die Arbeits- und Betriebsorganisation der Zukunft?*, in: *Berl. J. Soziol.*, Heft 1 2002, Berlin, S. 109-126.
- Garhammer, M. (2002b), *Zeitwohlstand und Lebensqualität - ein interkultureller Vergleich*, in: Rinderspacher, J.P. (Hrsg.), *Zeitwohlstand. Ein Konzept für einen anderen Wohlstand der Nation*. Berlin, 165-205.
- Goodin, R.E., Rice, J.M., Parpo, A. und L. Erriksson (2008), *Discretionary time. A new measure of freedom*, Cambridge.
- Greene, W. (2008), *Econometric Analysis, 6th Edition*, Pearson Prentice Hall, Upper Saddle River, New Jersey, USA.
- Groh-Samberg, O. (2009), *Armut, soziale Ausgrenzung und Klassenstruktur – Zur Integration multidimensionaler und längsschnittlicher Perspektiven*, VS Verlag, Wiesbaden.
- Hagerty, M., Cummins, R., Ferries, A., Land, K., Michalos, A., Peterson, M., Sharpe, A., Sirgy, J. und J. Vogel (2002), *Quality of Life Indexes for National Policy: Review and Agenda for Research*, in: *Social Indicator Research*, 55, 1-96.
- Hamermesh, D.S. (2004), *Association Lecture Subjective Outcomes in Economics*, in: *Southern Economic Journal*, Vol. 71, No.1, 1-11.
- Hamermesh, D.S. und G.A. Pfann (2005), *The Economics of Time Use*, North-Holland.
- Hanglberger, D. (2010), *Arbeitszufriedenheit und flexible Arbeitszeiten*, FFB-Discussionpaper No. 86, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg.
- Harvey, A.S. und A.K. Mukhopadhyay (2007), *When Twenty-four Hours is Not Enough: Time Poverty of Working Parents*, in: *Social Indicators Research*, 82, 1, 57-77.
- Hauser, R. (1997), *Armutsberichterstattung*, in: Heinz-Herbert Noll (Hrsg.), *Sozialberichterstattung in Deutschland: Konzepte, Methoden und Ergebnisse für Lebensbereiche und Bevölkerungsgruppen*, Weinheim/München, 19-45.
- Hauser, R. (2008), *Das Maß der Armut: Armutsgrenzen im sozialstaatlichen Kontext: Der sozialstatistische Diskurs*, in: Huster, E.-U. et al. (Hrsg.), *Handbuch Armut und Soziale Ausgrenzung*, Wiesbaden.
- Hauser, Richard (2007), *Armut in Deutschland – versagt der Sozialstaat?*, in: *Wirtschaftsdienst*, 87 Jg., H. 3, 172-181.
- Hauser, Richard (2008), *Das Maß der Armut: Armutsgrenzen im sozialstaatlichen Kontext. Der sozialpolitische Diskurs*, in: Huster, E.-U., Boeckh, J. und H. Mogge-Grotjahn (Hrsg.), *Handbuch Armut und Soziale Ausgrenzung*, Wiesbaden, 94-117.

- Hoff, A. (2002), *The Translog Approximation of the Constant Elasticity of Substitution Production Function with more than two Input Variables*, Working Paper 14/2002, Danish Research Institute of Food Economics.
- Hoff, A. (2004), *The linear Approximation of the CES Function with n Input Variables*, in: *Marine Resource Economics*, Volume 19, 295-306.
- Hyytinen, A. und O. Ruuskanen (2007), *Time Use of the Self-Employed*, in: *KYKLOS*, Vol. 60, pp. 105-122.
- Jenkins, S. und L. Osberg (2005), *Nobody to Play With? The Implications of Leisure Co-Ordination*, in: Hamermesh, D.S. und G.A. Pfann (Eds.), *The Economics of Time Use*, Elsevier B.V., 113-145.
- Kahnemann, D. und A. B. Krueger (2006), *Developments in the measurement of subjective well-being*, in: *The Journal of Economic Perspectives*, 20, 1, 3-24.
- Kmenta, J. (1967), *On Estimation of the CES Production Function*, in: *International Economic Review*, Vol. 8, No. 2.
- Linder, St. (1970), *The Harried Leisure Class*, New York.
- Lugo, M.A. und E. Maasoumi (2008), *Multidimensional Poverty Measures from an Information Theory Perspective*, Paper Prepared for the 30th General Conference of The International Association for Research in Income and Wealth, Portoroz, Slovenia, August 24-30, 2008.
- Mayraz, G., Wagner, G. und J. Schupp (2009), *Life satisfaction and relative income: Perceptions and evidence*, IZA Discussionpaper No. 4390, Bonn.
- Merz, J. (2001), *Hohe Einkommen, ihre Struktur und Verteilung, Lebenslagen in Deutschland*, Der erste Armuts- und Reichtumsbericht der Bundesregierung, Bundesministerium für Arbeit und Sozialordnung, Bonn.
- Merz, J. (2002), *Time and Economic Well-Being – A Panelanalysis of Desired vs. Actual Working Hours*, in: *The Review of Income and Wealth*, 48/3, 317-346.
- Merz, J. (2007), *Polarisierung der Einkommen von Selbständigen? – Zur Dynamik der Einkommensverteilung von Freiberuflern und Unternehmern*, in: Merz, J. (Ed.), *Fortschritte der Mittelstandsforschung*, Merz, J., Schulte, R. und J. Wagner (Series Eds.), *Entrepreneurship, Professions, Small Business Economics*, CREPS-Schriftenreihe Vol. 3, Lit Verlag, Münster, 395-415.
- Merz, J. (2008), *Hohe und niedrige Einkommen – Neue Ergebnisse zu Freien und anderen Berufen mit den Mikrodaten der Einkommensteuerstatistik*, in: Merz, J. (Hrsg.), *Freie Berufe – Einkommen und Steuern (FBEST) – Beiträge aus Wissenschaft und Praxis*, Schriftenreihe Band 16, Nomos Verlagsgesellschaft, Baden-Baden, 69-92.
- Merz, J. und D. Burgert (2004a), *Arbeitszeitarrangements – Neue Ergebnisse aus der Zeitbudgeterhebung 2001/02 im Zeitvergleich zu 1991/92*, in: Statistisches Bundesamt (Ed.), *Alltag in Deutschland – Analysen zur Zeitverwendung*, FORUM Band 43, Wiesbaden, 304-336.
- Merz, J. und D. Burgert (2004b), *Wer arbeitet wann? – Arbeitszeitarrangements von Selbständigen und abhängig Beschäftigten: Eine mikroökonomische Analyse deutscher Zeitbudgetdaten*, in: Merz, J. and J. Wagner (Eds.), *Perspektiven der Mittelstandsforschung – Ökonomische Analysen zu Selbständigkeit, Freien Berufen und KMU*, Merz, J., Schulte, R. and J. Wagner (Series Eds.), *Entrepreneurship, Professions, Small Business Economics*, CREPS-Schriftenreihe Vol. 1, Lit-Verlag, Münster, 303-330.
- Merz, J. und L. Osberg (2009), *Keeping in Touch - A Benefit of Public Holidays Using Time Use Diary Data*, in: *electronic International Journal of Time Use Research* (in the referee process).

- Merz, J. und M. Zwick (2005), *Hohe Einkommen: Eine Verteilungsanalyse für Freie Berufe, Unternehmer und abhängig Beschäftigte mit Mikrodaten der Einkommensteuerstatistik*, in: Schmollers Jahrbuch – Journal of Applied Social Science Studies, Jg. 125/2, 269-298.
- Merz, J. und M. Zwick (2008), *Einkommens- und Verteilungsanalysen mit dem Taxpayer-Panel – Neue Möglichkeiten und erste Ergebnisse für Selbständige als Freiberufler und Unternehmer und abhängige Beschäftigte sowie für hohe Einkommen*, in: Rolf, Gabriele, Zwick, Markus und Gert G. Wagner (Hrsg.), Fortschritte der informationellen Infrastruktur in Deutschland, Festschrift für Johann Hahlen zum 65. Geburtstag und Hans-Jürgen Krupp zum 75. Geburtstag, Nomos Verlagsgesellschaft, Baden-Baden, 338-358.
- Merz, J. und P. Böhm (2008), *Tägliche Arbeitszeitmuster und Einkommen von Freien Berufen – Neue Ergebnisse auf der Basis von Zeittagebüchern*, in: Merz, J. (Hrsg.), Freie Berufe – Einkommen und Steuern (FBES) – Beiträge aus Wissenschaft und Praxis, Schriftenreihe Band 16, Nomos Verlagsgesellschaft, Baden-Baden, 93-118.
- Merz, J. und P. Paic (2005), *Zum Einkommen der Freien Berufe – Eine Ordered Probit-Analyse ihrer Determinanten auf Basis der FFB-Onlineumfrage*, in: Reinhard Schulte (Ed.), Ergebnisse der Mittelstandsforschung, Merz, J., Schulte, R. und J. Wagner (Series Eds.), Entrepreneurship, Professions, Small Business Economics, CREPS-Schriftenreihe Vol. 2, Lit-Verlag, Münster, 327-348.
- Merz, J. und T. Rathjen (2009), *Time and Income Poverty - An Interdependent Multidimensional Poverty Approach With German Time Use Diary Data*, FFB-Discussionpaper No. 79, Fakultät II, Wirtschafts- Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg 2009 (auch IZA Discussionpaper No. 4337, Bonn und ECINEQ Discussionpaper No. 126, Palma de Mallorca)
- Merz, J., P. Böhm, und D. Burgert (2009), *Timing and Fragmentation of Daily Working Hours Arrangements and Income Inequality – An Earnings Treatment Effects Approach with German Time Use Diary Data*, in: electronic International Journal of Time Use Research, 6/2, 200-239.
- Mückenberger, U. (2008), *Zeitpolitik in der Metropolregion Hamburg, Projektantrag Wissenschaftliche Begleitung des Projektverbunds ‚Zeitpolitik in der Metropolregion Hamburg‘*, Hamburg.
- Nolan, B. und V.T. Whelan (2007), *On the Multidimensionality of Poverty and Social Exclusion*, in: Jenkins, S.P. and J. Micklewright (2007), *Inequality and Poverty Re-examined*, Oxford University Press, Oxford, 146-165.
- Osberg, L. (2002), *Time, Money and Inequality in International Perspective*, Department of Economics, Manuscript, Dalhousie University, Halifax, Canada.
- Osberg, L. (Ed.) (2003), *The Economic Implications of Social Cohesion*, University of Toronto Press, Toronto.
- Osberg, L. und K. Xu (2002), *On Sen's Approach to Poverty Measures and Recent Developments*, Paper presented at the Sixth International Meeting of the International Association for Research of Income and Wealth, and China Economic Quarterly (published in Chinese) Vol. 1, No. 1, 151-170.
- Parker, S. (2004), *The Economics of Self-Employment and Entrepreneurship*, Cambridge University Press.
- Prahl, H. (2002), *Soziologie der Freizeit*, Paderborn.
- Rhein, T. (2009), *„Working poor“ in Deutschland und den USA - Arbeit und Armut im transatlantischen Vergleich*, IAB Kurzbericht, in: Aktuelle Analysen und Kommentare aus dem Institut für Arbeitsmarkt- und Berufsforschung 1/2009, Nürnberg.

- Rode, J.C. (2004), *Job Satisfaction and Life Satisfaction revisited: A Longitudinal Test of an Integrated Model*, in: *Human Relations*, 57 (9), 1205-1230.
- Rosa, Hartmut (2006), *Beschleunigung – Die Veränderung der Zeitstrukturen in der Moderne*, Suhrkamp, Frankfurt am Main.
- Sen, A.K. (1976), *Poverty: An Ordinal Approach*, in: *Econometrica*, Vol. 44, No. 2, pp. 219 – 231.
- Sen, A.K. (1985), *Commodities and Capabilities*, North-Holland, Amsterdam.
- Sen, A.K. (1999), *Development as Freedom*, Knopf Publishers, New York.
- Shorrocks, A.F. (1995), *Revisiting the Sen Poverty Index*, in: *Econometrica* 63 (5), S. 1225-1230.
- Statistisches Bundesamt (Ed.) (2006), *Compass 2001/02 Time Use Survey: Publications, Links and Prints incl. Data Files and Research Facilities*.
- Sullivan, O. (2007), *Cultural Voraciousness – A New Measure of the Pace of Leisure in a Context of 'Harriedness'*, in: *electronic International Journal of Time Use Research 2007*, Vol. 4, No. 1, 30-46.
- Thon, D. (1979), *On Measuring Poverty*, in: *Review of Income and Wealth* 25, S. 429-440.
- Thursby, J.G. und C.A.K Lovell (1978), *An Investigation of the Kmenta Approximation to the CES Function*, in: *International Economic Review*, Vol. 19, No. 2.
- Tsui, K. Y. (2002), *Multidimensional Poverty Indices*, in: *Social Choice and Welfare*, 19(1): 69-93.
- Vickery, C. (1977), *The Time-Poor: A New Look at Poverty*, in: *Journal of Human Resources*, 12, 1, 27-48.
- Wagner, G.G., Frick, J.R. und J. Schupp (2007), *The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements*, in: *Journal of Applied Social Science Studies*, 127, 1, 139-179.

Forschungsinstitut Freie Berufe (FFB), Universität Lüneburg

Publikationen

FFB-Jahresberichte

FFB-Bücher in der FFB-Schriftenreihe

FFB-Bücher

FFB-Gutachten

FFB-Artikel in wissenschaftlichen Zeitschriften und Büchern

FFB-Reprints

FFB-Dokumentationen, ISSN 1615-0376

FFB Sonstige Arbeitsberichte, ISSN 0175-7275

www.leuphana.de/ffb

FFB-Diskussionspapiere, ISSN 0942-2595

Merz, J. und Henning Stolze, D., 2010, Cumulation of Cross-Section Surveys - Evaluation of Alternative Concepts for the Cumulated Continuous Household Budget Surveys (LWR) 1999 until 2003 compared to the Sample Survey of Income and Expenditures (EVS) 2003, FFB-Diskussionpaper No. 84, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg.

FFB-DP Nr.84

Österbacka, E., Merz, J. and C.D. Zick., 2010, Human Capital Investments in Children: A Comparative Analysis of the Role of Parent-Child Shared Time in Selected Countries, FFB-Diskussionspapier Nr. 83, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg.

FFB-DP Nr. 83

Merz, J. und T. Rathjen, 2010, Sind Selbständige zeit- und einkommensarm? Eine Mikroanalyse der Dynamik interdependenter multidimensionaler Armut mit dem Sozio-ökonomischen Panel und den deutschen Zeitbudgeterhebungen, FFB-Diskussionspapier Nr. 82, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg.

FFB-DP Nr. 82

Rucha, R., 2010, Effekte einer freiwilligen Mitgliedschaft in Berufsverbänden auf das Einkommen - Eine Analyse der Einkommensdynamik bei Freiberuflern in Deutschland, FFB-Diskussionspapier Nr. 81, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg.

FFB-DP Nr. 81

Hanglberger, D., 2010, Arbeitszufriedenheit und flexible Arbeitszeiten – Empirische Analyse mit Daten des Sozio-oekonomischen Panels, FFB-Diskussionspapier Nr. 80, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg.

FFB-DP Nr. 80

Merz, J. and T. Rathjen, 2009, Time and Income Poverty – An Interdependent Multidimensional Poverty Approach with German Time Use Diary Data, FFB-Diskussionpaper No. 79, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg.

FFB-DP Nr. 79

Merz, J., 2009, Time Use and Time Budgets – Improvements, Future Challenges and Recommendations, FFB-Diskussionpaper, No. 78, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg.

FFB-DP Nr. 78

Merz, J., Hanglberger, D. and R. Rucha, 2009, The Timing of Daily Demand for Goods and Services – Multivariate Probit Estimates and Microsimulation Results for an Aged Population with German Time Use Diary Data, FFB-Diskussionpaper No. 77, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Leuphana Universität Lüneburg, Lüneburg.

FFB-DP Nr. 77

Merz, J., 2009, Zeitverwendungsforschung und Mediennutzung, FFB-Diskussionspapier Nr. 75, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg.

FFB-DP Nr. 76

- Böhm, P. und J. Merz, 2008, Reichtum in Niedersachsen und anderen Bundesländern – Ergebnisse aus der Steuer-Geschäftsstatistik 2003 für Selbständige (Freie Berufe und Unternehmer) und abhängig Beschäftigte, FFB-Diskussionspapier Nr. 75, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 75
- Merz, J. und M. Zwick, 2008, Einkommens- und Verteilungsanalysen mit dem Taxpayer-Panel – Neue Möglichkeiten und erste Ergebnisse für Selbständige als Freiberufler und Unternehmer und abhängig Beschäftigte sowie für hohe Einkommen, FFB-Diskussionspapier Nr. 74, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 74
- Böhm, P. und J. Merz, 2008, Zum Einkommensreichtum Älterer in Deutschland – Neue Reichtumskennzahlen und Ergebnisse aus der Lohn- und Einkommenssteuerstatistik (FAST 2001), FFB-Diskussionspapier Nr. 73, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 73
- Herrmann, H., 2007, Freie Berufe – Europäische Entwicklungen, FFB-Diskussionspapier Nr. 72, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 72
- Deneke, J.F.V., 2007, Freie Berufe – Herausforderungen an die Wissenschaft, FFB-Diskussionspapier Nr. 71, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 71
- Merz, J., Böhm, P., Hanglberger, D., Stolze, H. und R. Rucha, 2007, Wann werden Serviceleistungen nachgefragt? Eine Wirkungsanalyse anhand des neuen FFB-Mikrosimulationsmodells ServSim, FFB-Diskussionspapier Nr. 70, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 70
- Stolze H., 2007, Die Nachfrage nach Gesundheitsleistungen - Ergebnisse der FFB-Patientenumfrage, FFB-Diskussionspapier Nr. 69, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 69
- Burgert, D., 2006, Betriebliche Weiterbildung und Verbleib Älterer im Betrieb, FFB-Diskussionspapier Nr. 68, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 68
- Merz, J., 2006, Polarisierung der Einkommen von Selbständigen? – Zur Dynamik der Einkommensverteilung von Freiberuflern und Unternehmern, FFB-Diskussionspapier Nr. 67, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 67
- Paic, P., 2006, Informationelle Zugänge für die empirische Untersuchung freiberuflicher Existenzgründungen, FFB-Diskussionspapier Nr. 66, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 66
- Merz, J. und M. Zwick, 2006, Einkommensanalysen mit Steuerdaten – Mikroanalysen zu hohen Einkommen und Selbständigkeit und Mikrosimulation zu Politikalternativen der Einkommen-, Körperschaft- und Gewerbesteuer, FFB-Diskussionspapier Nr. 65, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 65
- Merz, J., Stolze, H. und M. Zwick, 2006, Wirkungen alternativer Steuerreformmodelle auf die Einkommensverteilung von Freien und anderen Berufen, FFB-Diskussionspapier Nr. 64, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 64
- Zwick, M., 2006, Gemeindefinanzreform und Freie Berufe, FFB-Diskussionspapier Nr. 63, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 63
- Burgert, D., 2006, Kündigungsschutz – Ein Hindernis für das Beschäftigungswachstum? FFB-Diskussionspapier Nr. 62, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 62
- Paic, P., 2006, Existenzgründung und Einkommen Freier Berufe – Ergebnisse aus dem Sozio-ökonomischen Panel und der FFB-Onlineumfrage Freie Berufe, FFB-Diskussionspapier Nr. 61, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 61

- Merz, J. und P. Böhm, 2006, Arbeitszeit und Einkommen – Neue Ergebnisse aus der deutschen Zeitbudgeterhebung, FFB-Diskussionspapier Nr. 60, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 60
- Merz, J., 2006, Hohe und niedrige Einkommen – Neue Ergebnisse zu Freien und anderen Berufen aus dem Armuts- und Reichtumsbericht der Bundesregierung, FFB-Diskussionspapier Nr. 59, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 59
- Merz, J., Vorgrimler, D. und M. Zwick, 2006, De facto Anonymised Microdata File on Income Tax Statistics 1998, FFB-Diskussionspapier Nr. 58, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 58
- Merz, J. and L. Osberg, 2006, Keeping in Touch – A Benefit of Public Holidays. FFB-Diskussionspapier Nr. 57, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 57
- Merz, J. and P. Paic, 2006, Start-up success of freelancers – New microeconomic evidence from the German Socio-Economic Panel. FFB-Diskussionspapier Nr. 56, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 56
- Merz, J. und P. Paic, 2006, Erfolgsfaktoren freiberuflicher Existenzgründung – Neue mikroökonomische Ergebnisse mit Daten des Sozio-ökonomischen Panels, FFB-Diskussionspapier Nr. 55, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 55
- Merz, J. and H. Stolze, 2006, Representative Time Use Data and Calibration of the American Time Use Studies 1965-1999, FFB-Diskussionspapier Nr. 54, Fakultät II, Wirtschafts-, Verhaltens- und Rechtswissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 54
- Paic, P. und H. Brand, 2005, Die Freien Berufe im Sozio-ökonomischen Panel - Systematische Berichtigung der kritischen Wechsel innerhalb der Selbständigengruppe, FFB-Diskussionspapier Nr. 53, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 53
- Merz, J. und P. Paic, 2005, Zum Einkommen der Freien Berufe – Eine Ordered Probit-Analyse ihrer Determinanten auf Basis der FFB-Onlineumfrage, FFB-Diskussionspapier Nr. 52, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 52
- Burgert, D., 2005, Schwellenwerte im deutschen Kündigungsschutzrecht – Ein Beschäftigungshindernis für kleine Unternehmen?, FFB-Diskussionspapier Nr. 51, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 51
- Merz, J., Böhm, P. und D. Burgert, 2005, Arbeitszeitarrangements und Einkommensverteilung – Ein Treatment Effects Ansatz der Einkommensschätzung für Unternehmer, Freiberufler und abhängig Beschäftigte, FFB-Diskussionspapier Nr. 50, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 50
- Burgert, D., 2005, The Impact of German Job Protection Legislation on Job Creation in Small Establishments – An Application of the Regression Discontinuity Design, FFB-Diskussionspapier Nr. 49, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 49
- Merz, J. und D. Burgert, 2005, Arbeitszeitarrangements – Neue Ergebnisse aus der nationalen Zeitbudgeterhebung 2001/02 im Zeitvergleich, FFB-Diskussionspapier Nr. 48, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 48
- Merz, J., Böhm, P. and D. Burgert, 2005, Timing, Fragmentation of Work and Income Inequality – An Earnings Treatment Effects Approach, FFB-Diskussionspapier Nr. 47, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 47
- Merz, J. und P. Paic, 2004, Existenzgründungen von Freiberuflern und Unternehmer – Eine Mikroanalyse mit dem Sozio-ökonomischen Panel FFB-Diskussionspapier Nr. 46, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 46
- Merz, J. und D. Burgert, 2004, Wer arbeitet wann? Arbeitszeitarrangements von Selbständigen und abhängig Beschäftigten – Eine mikroökonomische Analyse deutscher Zeitbudgetdaten, FFB-Diskussionspapier

- Nr. 45, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg.
FFB-DP Nr. 45
- Hirschel, D. und J. Merz, 2004, Was erklärt hohe Arbeitseinkommen der Selbständigen – Eine Mikroanalyse mit Daten des Sozio-ökonomischen Panels, FFB-Diskussionspapier Nr. 44, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 44
- Ackermann, D., Merz, J. und H. Stolze, 2004, Erfolg und Erfolgsfaktoren freiberuflich tätiger Ärzte – Ergebnisse der FFB-Ärztebefragung für Niedersachsen, FFB-Diskussionspapier Nr. 43, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 43
- Hirschel, D., 2003, Do high incomes reflect individual performance? The determinants of high incomes in Germany, FFB-Diskussionspapier Nr. 42, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 42
- Merz, J. and D. Burgert, 2003, Working Hour Arrangements and Working Hours – A Microeconomic Analysis Based on German Time Diary Data, FFB-Diskussionspapier Nr. 41, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 41
- Merz, J. und M. Zwick, 2002, Hohe Einkommen: Eine Verteilungsanalyse für Freie Berufe, Unternehmer und abhängig Beschäftigte, Eine Mikroanalyse auf der Basis der Einkommensteuerstatistik, FFB-Diskussionspapier Nr. 40, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 40
- Merz, J. and D. Hirschel, 2003, The distribution and re-distribution of income of self-employed as freelancers and entrepreneurs in Europe, FFB-Diskussionspapier Nr. 39, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 39
- Stolze, H., 2002, Datenbankgestützte Internetpräsenzen – Entwicklung und Realisation am Beispiel der Homepage des Forschungsinstituts Freie Berufe (FFB) der Universität Lüneburg <http://ffb.uni-lueneburg.de>, FFB-Diskussionspapier Nr. 38, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 38
- Merz, J., 2002, Zur Kumulation von Haushaltsstichproben, FFB-Diskussionspapier Nr. 37, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 37
- Merz, J., 2002, Reichtum in Deutschland: Hohe Einkommen, ihre Struktur und Verteilung – Eine Mikroanalyse mit der Einkommensteuerstatistik für Selbständige und abhängig Beschäftigte, FFB-Diskussionspapier Nr. 36, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 36
- Merz, J. und M. Zwick, 2002, Verteilungswirkungen der Steuerreform 2000/2005 im Vergleich zum ‚Karlsruher Entwurf‘ Auswirkungen auf die Einkommensverteilung bei Selbständigen (Freie Berufe, Unternehmer und abhängig Beschäftigte), FFB-Diskussionspapier Nr. 35, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 35
- Merz, J., Stolze, H. and M. Zwick, 2002, Professions, entrepreneurs, employees and the new German tax (cut) reform 2000 – A MICSIM microsimulation analysis of distributional impacts, FFB-Diskussionspapier Nr. 34, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 34
- Forschungsinstitut Freie Berufe, 2002, Freie Berufe im Wandel der Märkte - 10 Jahre Forschungsinstitut Freie Berufe (FFB) der Universität Lüneburg, Empfang am 4. November 1999 im Rathaus zu Lüneburg, FFB-Diskussionspapier Nr. 33, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 33
- Merz, J., 2002, Time Use Research and Time Use Data – Actual Topics and New Frontiers, FFB-Discussion Paper No. 32, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 32
- Merz, J., 2001, Freie Berufe im Wandel der Arbeitsmärkte, FFB-Diskussionspapier Nr. 31, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 31
- Merz, J., 2001, Was fehlt in der EVS? Eine Verteilungsanalyse hoher Einkommen mit der verknüpften Einkommensteuerstatistik für Selbständige und abhängig Beschäftigte, FFB-Diskussionspapier Nr. 30, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 30

- Merz, J., 2001, Informationsfeld Zeitverwendung – Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik, FFB-Diskussionspapier Nr. 29, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 29
- Schatz, C. und J. Merz, 2000, Die Rentenreform in der Diskussion – Ein Mikrosimulationsmodell für die Altersvorsorge in Deutschland (AVID-PTO), FFB-Diskussionspapier Nr. 28, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 28
- Merz, J., 2000, The Distribution of Income of Self-employed, Entrepreneurs and Professions as Revealed from Micro Income Tax Statistics in Germany, FFB-Discussion Paper No. 27, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 27
- Merz, J., Loest, O. und A. Simon, 1999, Existenzgründung – Wie werde ich selbständig, wie werde ich Freiberufler? Ein Leitfaden, FFB-Diskussionspapier Nr. 26, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 26
- Merz, J. und D. Kirsten, 1998, Extended Income Inequality and Poverty Dynamics of Labour Market and Valued Household Activities – A Ten Years Panelanalysis for Professions, Entrepreneurs and Employees in Germany, FFB-Discussion Paper No. 25, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 25
- Merz, J., Quiel, T. und K. Venkatarama, 1998, Wer bezahlt die Steuern? – Eine Untersuchung der Steuerbelastung und der Einkommenssituation für Freie und andere Berufe, FFB-Diskussionspapier Nr. 24, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 24
- Merz, J. and R. Lang, 1997, Preferred vs. Actual Working Hours – A Ten Paneleconometric Analysis for Professions, Entrepreneurs and Employees in Germany, FFB-Discussion Paper No. 23, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 23
- Merz, J., 1997, Privatisierung und Deregulierung und Freie und staatlich gebundene Freie Berufe – Einige ökonomische Aspekte, FFB-Diskussionspapier Nr. 22, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 22
- de Kam, C. A., de Haan, J., Giles, C., Manresa, A., Berenguer, E., Calonge, S., Merz, J. and K. Venkatarama, 1996, The Distribution of Effective Tax Burdens in Four EU Countries, FFB-Discussion Paper No. 21, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 21
- Deneke, J. F. V., 1996, Freie Berufe und Mittelstand – Festrede zur Verleihung der Ehrendoktorwürde, FFB-Diskussionspapier Nr. 20, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 20
- Merz, J., 1996, Die Freien Berufe – Laudatio zur Verleihung der Ehrendoktorwürde des Fachbereiches Wirtschafts- und Sozialwissenschaften der Universität Lüneburg an Prof. J. F. Volrad Deneke, FFB-Diskussionspapier Nr. 19, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 19
- de Kam, C. A., de Haan, J., Giles, C., Manresa, A., Berenguer, E., Calonge, S. and J. Merz, 1996, Who pays the taxes?, FFB-Discussion Paper No. 18, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 18
- Merz, J., 1996, Schattenwirtschaft und ihre Bedeutung für den Arbeitsmarkt, FFB-Diskussionspapier Nr. 17, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 17
- Merz, J. und D. Kirsten, 1995, Freie Berufe im Mikrozensus II – Einkommen und Einkommensverteilung anhand der ersten Ergebnisse für die neuen und alten Bundesländer 1991, FFB-Diskussionspapier Nr. 16, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 16
- Merz, J. und D. Kirsten, 1995, Freie Berufe im Mikrozensus I – Struktur und quantitative Bedeutung anhand der ersten Ergebnisse für die neuen und alten Bundesländer 1991, FFB-Diskussionspapier Nr. 15, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 15
- Merz, J., 1995, MICSIM – Concept, Developments and Applications of a PC-Microsimulation Model for Research and Teaching, FFB-Discussion Paper No. 14, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 14

- Rönnau, A., 1995, Freie Berufe in der DDR, der Bundesrepublik Deutschland und im wiedervereinten Deutschland: Auswertungen von Berufstätigenerhebung und Arbeitsstättenzählung, FFB-Diskussionspapier Nr. 13, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 13
- Burkhauser, R. V., Smeeding, T. M. and J. Merz, 1994, Relative Inequality and Poverty in Germany and the United States Using Alternative Equivalence Scales, FFB-Discussion Paper No. 12, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 12
- Widmaier, U., Niggemann, H. and J. Merz, 1994, What makes the Difference between Unsuccessful and Successful Firms in the German Mechanical Engineering Industry? A Microsimulation Approach Using Data from the NIFA-Panel, FFB-Discussion Paper No. 11, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 11
- Merz, J., 1994, Microdata Adjustment by the Minimum Information Loss Principle, FFB-Discussion Paper No. 10, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 10
- Merz, J., 1994, Microsimulation – A Survey of Methods and Applications for Analyzing Economic and Social Policy, FFB-Discussion Paper No. 9, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 9
- Merz, J., Garner, T., Smeeding, T. M., Faik, J. and D. Johnson, 1994, Two Scales, One Methodology – Expenditure Based Equivalence Scales for the United States and Germany, FFB-Discussion Paper No. 8, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 8
- Krickhahn, T., 1993, Lobbyismus und Mittelstand: Zur Identifikation der Interessenverbände des Mittelstands in der Bundesrepublik Deutschland, FFB-Diskussionspapier Nr. 7, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 7
- Merz, J., 1993, Market and Non-Market Labor Supply and Recent German Tax Reform Impacts – Behavioral Response in a Combined Dynamic and Static Microsimulation Model, FFB-Discussion Paper No. 6, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 6
- Merz, J., 1993, Microsimulation as an Instrument to Evaluate Economic and Social Programmes, FFB-Discussion Paper No. 5, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 5
- Merz, J., 1993, Statistik und Freie Berufe im Rahmen einer empirischen Wirtschafts- und Sozialforschung, Antrittsvorlesung im Fachbereich Wirtschafts- und Sozialwissenschaften der Universität Lüneburg, FFB-Diskussionspapier Nr. 4, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 4
- Merz, J. and J. Faik, 1992, Equivalence Scales Based on Revealed Preference Consumption Expenditure Microdata – The Case of West Germany, FFB-Discussion Paper No. 3, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 3
- Merz, J., 1992, Time Use Dynamics in Paid Work and Household Activities of Married Women – A Panel Analysis with Household Information and Regional Labour Demand, FFB-Discussion Paper No. 2, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg. FFB-DP Nr. 2
- Forschungsinstitut Freie Berufe, 1992, Festliche Einweihung des Forschungsinstituts Freie Berufe am 16. Dezember 1991 im Rathaus zu Lüneburg, FFB-Diskussionspapier Nr. 1, Fachbereich Wirtschafts- und Sozialwissenschaften, Universität Lüneburg, Lüneburg. FFB-DP Nr. 1

FFB-Jahresberichte und FFB-Reprints werden auf Anfrage kostenlos zugesandt.

FFB-Diskussionspapiere und FFB-Dokumentationen können über unsere FFB-Homepage kostenlos heruntergeladen werden.

FFB-Bücher können sowohl bei uns als auch beim NOMOS Verlag bestellt werden. Sonstige Arbeitsberichte und Sonstige Bücher sind über den jeweiligen Autor zu beziehen.

www.leuphana.de/ffb