

00

Statistische Grundlagen und Übersichten

338-0069

Methodenbericht

EVS I – Energieverbrauchsstatistik 2002 bis 2007

Stichprobenplan und Hochrechnung

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Statistik BFS

Neuchâtel, 2013

Die vom Bundesamt für Statistik (BFS)
herausgegebene Reihe «Statistik der Schweiz»
gliedert sich in folgende Fachbereiche:

0 Statistische Grundlagen und Übersichten

1 Bevölkerung

2 Raum und Umwelt

3 Arbeit und Erwerb

4 Volkswirtschaft

5 Preise

6 Industrie und Dienstleistungen

7 Land- und Forstwirtschaft

8 Energie

9 Bau- und Wohnungswesen

10 Tourismus

11 Mobilität und Verkehr

12 Geld, Banken, Versicherungen

13 Soziale Sicherheit

14 Gesundheit

15 Bildung und Wissenschaft

16 Kultur, Medien, Informationsgesellschaft, Sport

17 Politik

18 Öffentliche Verwaltung und Finanzen

19 Kriminalität und Strafrecht

20 Wirtschaftliche und soziale Situation der Bevölkerung

21 Nachhaltige Entwicklung und Disparitäten auf regionaler und internationaler Ebene

Methodenbericht

EVS I – Energieverbrauchsstatistik 2002 bis 2007

Stichprobenplan und Hochrechnung

Autorin Monika Ferster

Herausgeber Bundesamt für Statistik (BFS)

Office fédéral de la statistique (OFS)
Neuchâtel, 2013

Herausgeber:	Bundesamt für Statistik (BFS)
Auskunft:	Monika Ferster, Tel. 032 713 67 74, monika.ferster@bfs.admin.ch
Bearbeitung:	Dienst Statistische Methoden, BFS
Vertrieb:	Bundesamt für Statistik, CH-2010 Neuchâtel Tel. 032 713 60 60 / Fax 032 713 60 61 / E-Mail: order@bfs.admin.ch
Internet:	www.statistik.admin.ch
Bestellnummer:	338-0069
Preis:	gratis
Reihe:	Statistik der Schweiz
Fachbereich:	0 Statistische Grundlagen und Übersichten
Originaltext:	Deutsch
Titelgrafik:	BFS; Konzept: Netthoevel & Gaberthüel, Biel; Foto: © NorthShoreSurfPhotos – Fotolia.com
Grafik/Layout:	Sektion DIAM, Prepress/Print
Copyright:	BFS, Neuchâtel 2013 Abdruck – ausser für kommerzielle Nutzung – unter Angabe der Quelle gestattet
ISBN:	978-3-303-00497-5

Inhaltsverzeichnis

Vorwort 5

Zusammenfassung 5

1 Einleitung 7

2 Stichprobenplan und Ziehung 8

2.1 Stichprobenrahmen und Schichtung . . . 8

2.2 Stichprobenplan . . . 11

2.3 Ziehung . . . 14

3 Hochrechnung 16

3.1 Schätzer für das Total . . . 16

3.1.1 Robustifizierung . . . 17

3.1.2 Der Horvitz-Thompson-Schätzer . . . 19

3.1.3 Der Quotientenschätzer . . . 19

3.1.4 Ausreisserbehandlung . . . 19

3.2 Varianzschätzung . . . 21

4 Longitudinale Zerlegung 22

4.1 Ausgangslage . . . 23

4.2 Longitudinale Zerlegung der Stichprobenrahmen . . . 23

4.3 Longitudinale Zerlegung der Bruttostichproben . . . 25

4.4 Longitudinale Zerlegung der Nettostichproben . . . 27

4.5 Longitudinale Zerlegung der Schätzer . . . 29

5 Imputation 30

5.1 Imputation in der Phase der Datenerhebung . . . 30

6 Jährlich wiederkehrende Abläufe 33

7 Stichproben und Resultate 34

7.1 EVS 2002 . . . 35

7.2 EVS 2003 . . . 37

7.3 EVS 2004 . . . 39

7.4 EVS 2005 . . . 41

7.5 EVS 2006 . . . 43

7.6 EVS 2007 45

Anhang **47**

A Erwartete und effektive Stichprobengrösse **47**

B Imputation in der Phase der Hochrechnung **47**

B.1 Robuste Regression durch den Nullpunkt 47

B.2 Umsetzung am Beispiel EVS2007 48

C SAS-Programme und Daten **50**

C.1 Übersicht 50

C.2 Auszug SAS-Code zur Ziehung 50

C.3 Auszug SAS-Code zur Imputation in der Hochrechnungsphase 51

Literatur **52**

Methodenberichte des Dienstes Statistische Methoden des BFS **53**

Vorwort

Das Bundesamt für Energie erhebt jährlich den Energieverbrauch in der Industrie und im Dienstleistungssektor. In Zusammenarbeit mit dem Dienst Statistische Methoden (METH) des Bundesamtes für Statistik wurden Methoden entwickelt, um den Verbrauch von Elektrizität, Heizöl und Erdgas bei Betrieben in 19 Branchen gesamtschweizerisch zu schätzen.
Erstmals für den Energieverbrauch des Jahres 2002 wurden obige Methoden angewendet. Im Folgejahr wurde der Stichprobenplan noch leicht modifiziert.

Der Stichprobenplan und die Hochrechnung wurden vom ehemaligen METH-Mitarbeiter Paul-André Salamin entwickelt. Er war auch für die Anwendungen der ersten Jahre zuständig. Die Autorin dankt Paul-André Salamin für diese Entwicklungsarbeit und ihre Einführung in die Energieverbrauchsstatistik.

Ein weiterer Dank geht an Daniel Assoulin von METH für das aufmerksame und konstruktive Gegenlesen dieses Berichts.

Die Autorin bedankt sich ebenfalls bei Frau Jasmin Gülden und Herrn Felix Andrist vom Bundesamt für Energie für die gute Zusammenarbeit und die Erlaubnis, ihre Stichprobengrössen und Resultate in der gemachten Weise in diesen Bericht aufnehmen zu können.

Zusammenfassung

In diesem Methodenbericht sind die statistischen Methoden vom Stichprobenplan bis zur Hochrechnung beschrieben, welche im Bundesamt für Statistik für die Energieverbrauchsstatistik des Bundesamtes für Energie entwickelt wurden.

Das Bundesamt für Energie erhebt jährlich den Energieverbrauch in der Industrie und im Dienstleistungssektor. In Zusammenarbeit mit dem Dienst Statistische Methoden des Bundesamtes für Statistik wurde ein Stichprobenplan ausgearbeitet, der den unterschiedlichen Grössen und Organisationsformen der Betriebe - zum Beispiel ob zu einem Verein gehörend oder Arbeitsstätte eines Unternehmens - Rechnung trägt.

Die Hochrechnung im Bericht bezieht sich auf die Energieträger Elektrizität, Heizöl extra-leicht und Erdgas und beruht auf einem Quotientenschätzer mit der Hilfsvariable Vollzeitstellen. Für den Vergleich des Verbrauchs von zwei aufeinanderfolgenden Jahren wird eine longitudinale Zerlegung beschrieben, welche vor allem für die Plausibilisierung der erhobenen Daten herangezogen wird.

Der Bericht endet mit einigen hochgerechneten Resultaten der Jahre 2002 bis 2007 in den 19 für die Energieverbrauchsstatistik relevanten Branchen. Die Zuordnung der Betriebe zu den Branchen in den Jahren 2002 bis und mit 2007 basiert auf der NOGA2002, der Allgemeinen Systematik der Wirtschaftszweige, revidiert im Jahr 2002.

1 Einleitung

Das Bundesamt für Energie (BFE) erhebt jährlich den Energieverbrauch in der Industrie und im Dienstleistungssektor. Die Betriebe dieser beiden Sektoren werden nach dem Verbrauch aller verwendeten Energieträger befragt. Je nach Energieträger wird mit dieser Erhebung der Energieverbrauch unterschiedlich berechnet. Die Energieträger Elektrizität, Heizöl extra-leicht und Erdgas werden hochgerechnet. Das entsprechende Hochrechnungsmodell wurde von der Sektion Statistische Methoden (METH) des Bundesamtes für Statistik (BFS) ausgearbeitet. Die Energieträger Kohle, Industrieabfälle und Heizöl schwer werden in der Schweiz von wenigen grossen Arbeitsstätten beziehungsweise Unternehmungen verwendet, welche in energieintensiven Branchen zu finden sind. Da diese Arbeitsstätten alle in der Stichprobe der Erhebung enthalten sind und somit eine Vollerhebung vorliegt, kann dieser Energieverbrauch direkt aufsummiert werden. Bei den Energieträgern Fernwärme und Holz sind hingegen zu wenige Arbeitsstätten erfasst, um eine verlässliche Hochrechnung durchzuführen.

Der hier vorliegende Methodenbericht befasst sich mit der Hochrechnung der Energieträger Elektrizität, Heizöl extra-leicht und Erdgas und beschreibt die verwendeten statistischen Methoden. Detaillierte Angaben und Resultate sind in der Dokumentation "Energieverbrauchsstatistik in der Industrie und im Dienstleistungssektor" des BFE (siehe Gülden, J. et al. (2007)) zu finden.

Die Stichprobe auf Basis des Betriebs- und Unternehmensregisters (BUR) wird im BFS gezogen, die Datensammlung und -bereinigung wird von den beiden externen Firmen Polyquest AG und Helbling AG durchgeführt. Die Methodologie (Stichprobenplan und Hochrechnung) wurde im BFS von METH entwickelt. Die Hochrechnung wird sowohl von Polyquest wie auch von METH durchgeführt und es ist möglich, neue Methoden zu studieren und die Entwicklung des Energieverbrauchs über die Jahre hinweg zu verfolgen.

Das heutige Verfahren kam erstmals 2002 zur Anwendung. Die neueste Energieverbrauchsstatistik (EVS), die in diesem Methodenbericht berücksichtigt wird, ist die EVS2007. Sie umfasst den Stichprobenplan und die -ziehung gegen Ende 2007, die Erhebung der Daten über den Verbrauch im Jahr 2007 in der ersten Hälfte des Jahres 2008 sowie anschliessend die Bereinigung und Hochrechnung. Im vorliegenden Methodenbericht werden diese wiederkehrenden Abläufe vor allem aus methodologischer Sicht beschrieben. Der Schwerpunkt liegt weniger auf den Resultaten als vielmehr auf einem vertieften Überblick über die verwendete Methodik bei Ziehung und Hochrechnung.

Der vorliegende Methodenbericht umfasst die Betriebsdaten und Verbrauchsangaben der Jahre 2002 bis und mit 2007. Kapitel 2 und 3 beschreiben die Methoden vom Stichprobenplan bis zur Hochrechnung. In Kapitel 4 wird in einer sogenannten longitudinalen Zerlegung die Hochrechnung zweier aufeinanderfolgender Jahre detailliert miteinander verglichen. Kapitel 5 widmet sich den Imputationsverfahren. In Kapitel 6 wird der allgemeine, sich wiederholende Ablauf der Erhebungen beschrieben bevor abschliessend in Kapitel 7 die Stichprobenpläne und ausgewählte Resultate der sechs Jahre tabelliert werden.

Der Verbrauch der drei Energieträger Elektrizität, Heizöl und Erdgas wird pro Branche hochgerechnet. Die allgemeine Systematik der Wirtschaftszweige, in der Schweiz mit NOGA abgekürzt, im Europäischen Raum mit NACE, ist eine Nomenklatur der wirtschaftlichen Aktivität für Betriebe oder Unternehmen. Die hier verwendete Version des Jahres 2002, die NOGA2002, dient als Basis der Branchendefinition.

2 Stichprobenplan und Ziehung

Die Untersuchungseinheiten der Erhebung sind die Betriebe aus dem Betriebs- und Unternehmensregister. Beim Energieverbrauch handelt es sich um eine schief verteilte Grösse. Deswegen wird die Referenzpopulation in Schichten ähnlichen Verbrauchs unterteilt. Der Energieverbrauch korreliert gut mit der Betriebsgrösse, hier gemessen durch die Anzahl Vollzeitstellen. Unter Ausnützung dieser Registervariabeln wird mittels optimaler Allokation die Stichprobengrösse in den einzelnen Schichten ermittelt. Aus organisatorischen Gründen möchte man eine Stichprobe derart ziehen, dass eine maximale Überschneidung mit der Stichprobe des Vorjahres gewährleistet ist. Dies wird durch ein Poisson sampling mit festen Zufallszahlen (*permanent random numbers*) erreicht.

Bei der folgenden Beschreibung wird als anschauliches Beispiel vor allem auf die Zahlen der EVS2007 zurückgegriffen.

2.1 Stichprobenrahmen und Schichtung

Aus dem Betriebs- und Unternehmensregister BUR werden gegen Ende eines Jahres alle aktiven Betriebe (*STATUS = 1*) des privaten (*BETYP = L13*) und öffentlichen Sektors (*BETYP = L20, ... ,L27*) herausgefiltert. Die Betriebe dieses ersten Rahmens werden via NOGA-Code in Branchen und zum Teil in Unterbranchen eingeteilt. Mit dem Begriff Unterbranchen ist die gesamte Einteilung in Branchen und Unterbranchen gemeint. Die primäre Schichtung entspricht dieser Einteilung in Unterbranchen.

Die Zuordnung des NOGA2002-Codes zu den Unterbranchen (Tabelle 1) wurde nach energetischen Gesichtspunkten zusammengestellt. Zwei- bis vierstellige NOGA-Codes von Betrieben mit ähnlich grossem Energieverbrauch oder ähnlicher wirtschaftlicher Aktivität werden nach Möglichkeit derselben Unterbranche zugeteilt. Jedoch wird in gewissen NOGA-Codes der Energieverbrauch durch eine andere Erhebung als die hier beschriebene erfasst. Solche NOGA-Codes werden keiner Unterbranche zugeordnet. Oder NOGA-Codes gehören nicht zum Industrie- oder Dienstleistungssektor und sind deshalb von der EVS ausgeschlossen. In Tabelle 2 sind alle zweistelligen NOGA-Codes aufgeführt, die nicht zur EVS gehören.

Die Betriebe des Stichprobenrahmens U sind in zweistufige Schichten unterteilt. Die primäre Schichtung in Unterbranchen wird anhand der Vollzeitstellen weiter unterteilt in 4 Grössenklassen als sekundäre Schichtung (vgl. Tabelle 3). Dabei entspricht die Grössenklasse 3 der Aggregation der Grössenklassen 3 ($20 \leq VOLLZ \leq 49$) und 4 ($50 \leq VOLLZ \leq 99$).

Tabelle 4 zeigt den endgültigen Stichprobenrahmen U gemäss der zweitstufigen Stratifizierung nach Unterbranchen und Grössenklassen.

Unter den Betrieben des Stichprobenrahmens U werden noch drei Klassen von Betrieben definiert, die sich gegenseitig ausschliessen.

- Verbände: Hier liefert der Verband die Verbrauchszahlen seiner Mitglieder. Deshalb wird bei den Verbänden vollerhoben.

- AST-Gruppen: In diese Klasse werden diejenigen Unternehmen genommen, bei welchen der Hauptsitz des Unternehmens die Verbrauchszahlen aller seiner Arbeitsstätten resp. Betriebe liefert. Der erhobene Verbrauch wird gemäss Grösse auf die einzelnen Betriebe des Unternehmens aufgeteilt. Deshalb wird bei den AST-Gruppen ebenfalls vollerhoben.

Tabelle 1 Zuordnung des NOGA2002-Codes zu den 40 Unterbranchen

Branchengruppen	Branche	Unterbranche	NOGA 2002
Industrie			
Nahrungsmittel	1	-	15 16
Textil / Leder	2	1	17
		2	18 19
Papier / Druck	3	1	21.1
		2	21.2
		3	22
Chemie / Pharma	4	-	24
Zement / Beton	5	-	26.4 26.5
Andere NE-Mineralien	6	-	26.1 26.2 26.3 26.6 26.7 26.8
Metall / Eisen	7	-	27.1 27.2 27.3 27.5
NE-Metalle	8	-	27.4
Metall / Geräte	9	1	28.1 28.2 28.3 28.4 28.5 28.6 28.7
		2	30
		3	31.1 31.2 31.6
		4	31.3 31.4 31.5
		5	32
Maschinen	10	-	29
Andere Industrien	11	1	13 23 33 34 35 36 41
		2	14
		3	20
		4	25
		5	37
Bau	12	-	45
Dienstleistung			
Handel	13	1	50
		2	51
		3	52
Gastgewerbe	14	1	55.1
		2	55.2 55.3 55.4
		3	55.5
Kredit / Versicherungen	15	-	65 66 67
Verwaltung	16	-	75
Unterricht	17	1	80.1
		2	80.2
		3	80.3
		4	80.4
Gesundheits- / Sozialwesen	18	1	85.12 85.13 85.14 85.20 85.32
		2	85.11 85.31
Andere Dienstleistungen	19	1	60 61 62 63 64
		2	70 71 72 73 74
		3	90 91 92 93

Tabelle 2 In der EVS nicht enthaltene NOGA2002-Codes

NOGA-Code	Bedeutung
01	Landwirtschaft und Jagd
02	Forstwirtschaft
05	Fischerei und Fischzucht
10	Kohlenbergbau, Torfgewinnung
11	Gewinnung von Erdöl und Erdgas, Erbringung damit verbundener Dienstleistungen
12	Bergbau auf Uran- und Thoriumerze
40	Energieversorgung
95	Private Haushalte mit Hauspersonal
97	Erbringung von Dienstleistungen durch private Haushalte für den Eigenbedarf ohne ausgeprägte Schwerpunkte

Tabelle 3 Grössenklassendefinitionen gemäss Anzahl Vollzeitstellen

Grössenklasse (*clt*)	Vollzeitstellen (*VOLLZ*)	Bezeichnung
1	$1 \leq VOLLZ \leq 5$	kleine Betriebe
2	$6 \leq VOLLZ \leq 19$	kleine Betriebe
3	$20 \leq VOLLZ \leq 99$	mittlere Betriebe
5	$100 \leq VOLLZ$	grosse Betriebe

clt kommt von *classe de taille*.

Tabelle 4 EVS2007: Schichtgrössen des Rahmens

		clt				All
Branche	Unterbranche	1	2	3	5	
1	1	1515	525	255	122	2417
2	1	355	108	92	21	576
	2	740	72	36	9	857
3	1	10	3	5	8	26
	2	61	48	45	28	182
	3	2666	680	262	49	3657
4	1	417	194	213	102	926
5	1	5	20	10	7	42
6	1	841	259	156	18	1274
7	1	84	45	50	28	207
8	1	30	18	12	12	72
9	1	4999	1762	713	100	7574
	2	73	25	22	3	123
	3	412	181	119	45	757
	4	168	51	35	19	273
	5	366	135	89	39	629
10	1	2020	815	566	192	3593
11	1	5281	1118	629	224	7252
	2	174	123	66	2	365
	3	4694	1295	272	7	6268
	4	358	235	182	50	825
	5	178	79	42	2	301
12	1	25938	8093	2657	232	36920
13	1	11992	2607	505	35	15139
	2	15185	3450	1436	170	20241
	3	36235	5881	1205	84	43405
14	1	1959	1724	639	59	4381
	2	17375	3754	350	8	21487
	3	1198	250	49	8	1505
15	1	6869	2328	1047	301	10545
16	1	3989	2417	1083	197	7686
17	1	5216	1590	145	5	6956
	2	612	840	494	31	1977
	3	266	188	143	35	632
	4	2679	210	48	7	2944
18	1	15838	1519	254	22	17633
	2	1197	1082	941	245	3465
19	1	12225	2941	1303	294	16763
	2	63510	7551	1756	279	73096
	3	20988	2358	563	76	23985
All		268718	56574	18489	3175	346956

- Bundesbetriebe: Sie sind definiert durch den NOGA-Code 751 und den Betriebstyp L20. Die Daten werden welche vom Bundesamtes für Bauten und Logistik geliefert, weshalb im vorliegenden Bericht häufig die Abkürzung BBL-Betriebe verwendet wird. Aus Ressourcengründen wird hier eine Stichprobe der festen Grösse 50 gezogen.

Diese drei Klassen werden unter dem Oberbegriff "Spezialfälle" geführt. Neben den Spezialfällen gibt es noch den grossen Rest der Betriebe, die keine Spezialfälle sind. Unter letzteren wird eine Stichprobe von rund 12'000 Betrieben gezogen. Tabelle 5 zeigt die Rahmengrössen, aufgeschlüsselt nach den Spezialfällen, über die Jahre 2002 bis 2007 hinweg.

Tabelle 5 Anzahl Arbeitsstätten (*ast*) im Rahmen über die Jahre 2002 bis 2007, unterschieden nach Spezialfällen

Einteilung	ast02	ast03	ast04	ast05	ast06	ast07	Erhebungsart
Nicht-Spezialfälle	**354972**	**349238**	**346667**	**345875**	**320973**	**337046**	Stichprobe
Bundesbetriebe		217	216	159	125	126	Stichprobe
AST-Gruppen		8577	10847	10900	13334	9430	vollerhoben
Verbände	502	585	425	435	461	354	vollerhoben
Total	355474	358817	358155	357369	334893	346956	

Die Schichten der EVS bestehen also aus der Einteilung der Betriebe des Stichprobenrahmens U in Unterbranchen X Grössenklassen X (Nicht-)Spezialfälle.

2.2 Stichprobenplan

Alle Nicht-Spezialfall-Betriebe bilden die Population P, in welcher eine geschichtete Stichprobe gezogen wird. Die Population P wird in primäre Schichten g =Unterbranchen und sekundäre Schichten h =Grössenklassen zerlegt, also

$$P = \bigcup_g P_g = \bigcup_{g,h} P_{gh} \ .$$

Die Schichtgrössen werden mit $N_g = |P_g|$ und $N_{gh} = |P_{gh}|$ bezeichnet. In Tabelle 6 sind die Schichtgrössen der Nicht-Spezialfälle aufgelistet.

Bei jedem Element bzw. Betrieb $k \in P$ ist der Wert $x_k > 0$ der Anzahl Vollzeitstellen bekannt. Mittels dieser Variablen wird die Stichprobengrösse der einzelnen Schichten durch die im Folgenden beschriebene optimale Allokation berechnet.

Es soll die Stichprobengrösse n_g einer primären Schicht g so gewählt werden, dass bei der Schätzung des Totals $X_g = \sum_{k \in P_g} x_k$ aller Vollzeitstellen dieser Schicht eine vorgegebene Genauigkeit, ausgedrückt durch die Varianz V_g, erreicht wird. Bei optimaler Aufteilung nach Neyman in den Sekundarschichten P_{gh} ergibt sich

$$n_g = \frac{(\sum_h N_{gh} S_{gh})^2}{V_g + \sum_h N_{gh} S_{gh}^2} \ ,$$

$$\text{wobei } S_{gh}^2 = \frac{1}{N_{gh} - 1} \sum_{k \in P_{gh}} (x_k - \bar{X}_{gh})^2 \quad \text{mit } \bar{X}_{gh} = \frac{1}{N_{gh}} \sum_{k \in P_{gh}} x_k \ .$$

Wird die Varianz V_g durch den Variationskoeffizienten CV_g ausgedrückt mit $V_g = (CV_g X_g)^2$, so hat man

$$n_g = \frac{(\sum_h N_{gh} S_{gh})^2}{CV_g^2 X_g^2 + \sum_h N_{gh} S_{gh}^2} = n_g(CV_g). \tag{1}$$

Tabelle 6 EVS2007: Population P der Nicht-Spezialfälle

		clt				All
Branche	Unterbranche	1	2	3	5	
1	1	1512	516	237	104	2369
2	1	355	108	89	19	571
	2	740	72	36	9	857
3	**1**	**10**	**3**	**4**	**2**	**19**
	2	**61**	**48**	**44**	**23**	**176**
	3	2665	680	260	46	3651
4	1	414	190	207	95	906
5	**1**	**3**	**6**	**3**	**.**	**12**
6	1	839	258	153	14	1264
7	**1**	**84**	**44**	**47**	**18**	**193**
8	**1**	**30**	**18**	**12**	**7**	**67**
9	1	4999	1761	702	89	7551
	2	**73**	**24**	**22**	**3**	**122**
	3	405	178	113	35	731
	4	168	50	32	14	264
	5	363	133	83	30	609
10	1	2017	804	541	143	3505
11	1	5280	1115	620	204	7219
	2	171	121	66	1	359
	3	4694	1295	268	6	6263
	4	358	233	176	47	814
	5	178	79	42	2	301
12	1	25930	8088	2646	226	36890
13	1	11982	2579	467	27	15055
	2	15153	3429	1416	151	20149
	3	35611	5330	932	76	41949
14	1	1956	1721	630	59	4366
	2	17272	3688	345	8	21313
	3	1048	181	34	8	1271
15	1	6097	1896	783	163	8939
16	1	3847	2258	903	137	7145
17	1	5168	1568	142	5	6883
	2	603	812	449	30	1894
	3	221	150	118	30	519
	4	2677	206	48	7	2938
18	1	15812	1499	251	22	17584
	2	1176	1071	934	234	3415
19	1	9050	2082	839	150	12121
	2	63430	7510	1704	249	72893
	3	20953	2337	544	65	23899
All		263405	54141	16942	2558	337046

Die Gesamtstichprobengrösse n findet sich durch Summation über die Schichtstichprobengrössen:

$$n = \sum_g n_g = \sum_g n_g(CV_g) = \varphi(CV_1, ..., CV_G) \ .$$

Wenn für jede primäre Schicht dieselbe Genauigkeit angenommen wird, das heisst $CV_g = CV$ für alle $g \in G$, dann hat man $n = \varphi(CV)$ und kann umgekehrt die Genauigkeit CV in Abhängigkeit von der Stichprobengrösse n ausrechnen.

Konkret wird bei der EVS ein CV so gesucht, dass die Stichprobengrösse $n \approx 12000$ beträgt. Mit dem so gewählten CV lässt sich n_g gemäss Gleichung (1) berechnen. Mit optimaler Allokation wird nun n_g auf die sekundären Schichten aufgeteilt gemäss

$$n_{gh} = n_g \frac{N_{gh} S_{gh}}{\sum_i N_{gi} S_{gi}} \ . \tag{2}$$

Es kann sein, dass die auf diese Art zugeteilte Stichprobengrösse n_{gh} grösser ist als die Schichtgrösse N_{gh}. In diesem Fall wird diese Schicht gh vollerhoben und die überzähligen $n_{gh} - N_{gh}$ Betriebe wieder gemäss (2) auf die verbleibenden Schichten aufgeteilt. Diese Iteration wird so lange fortgesetzt, bis alle Stichprobengrössen kleiner oder gleich der Schichtgrössen sind.

Falls die primäre Schicht weniger als 200 Betriebe umfasst ($N_g < 200$), was bei sechs primären Schichten der Fall ist (hervorgehobene Zeilen in Tabelle 6), wird $n_g = N_g$ gesetzt und in allen Grössenklassen dieser primären Schicht vollerhoben.

Schichten der Grössenklasse 5 werden immer vollerhoben.

Nun ist die erste Phase der Allokation beendet. Die ermittelten Stichprobengrössen n_{gh}, die kleiner sind als die Schichtgrössen, werden nun noch gemäss folgenden Vorgaben angepasst.

1. Zu klein soll die Stichprobengrösse nicht sein. Wenn es möglich ist, wird ein $n_{gh} < 40$ auf $n_{gh} = 40$ gesetzt, andernfalls wird diese Schicht vollerhoben.

2. Viele Schichten der kleinen Betriebe sind sehr gross. Dann kann die Stichprobenrate sehr klein sein bzw. das Stichprobengewicht sehr gross. In diesem Fall wird die Stichprobengrösse so erhöht, dass das Gewicht höchstens 150 beträgt: $n_{gh} = N_{gh}/150$.

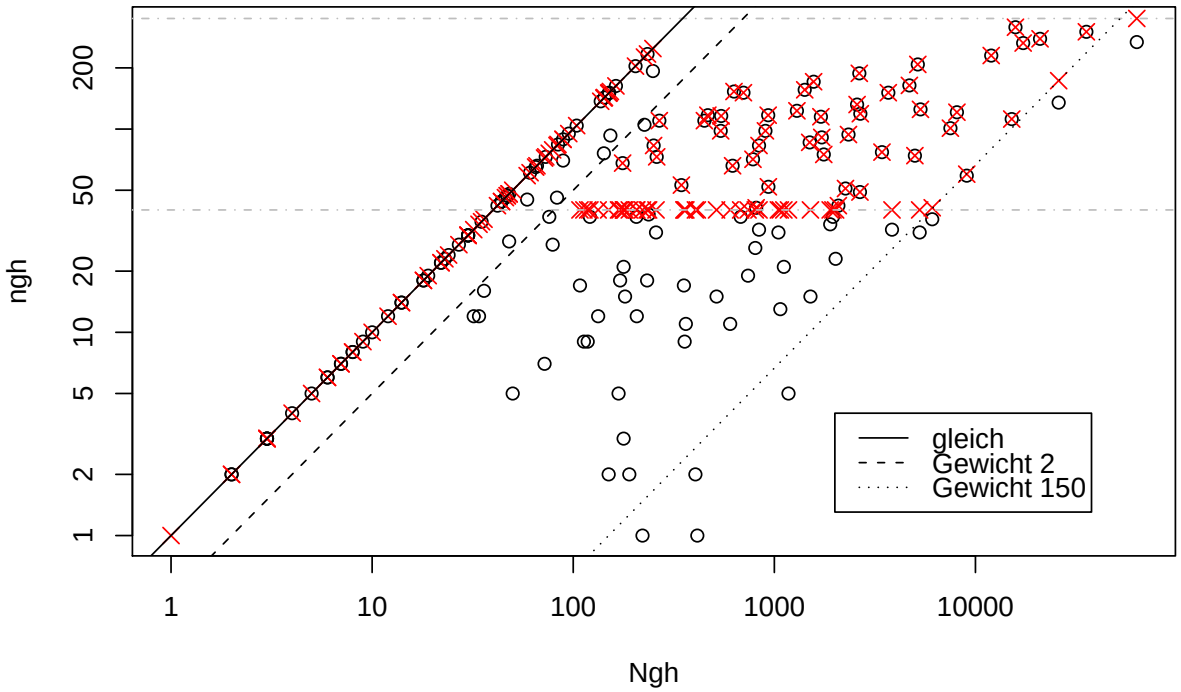
3. Damit nicht zuviele Betriebe Angaben liefern müssen, wird die maximale Stichprobengrösse auf 350 gesetzt: $n_{gh} = 350$. Dadurch kann in einigen wenigen Fällen das Stichprobengewicht 150 überschritten werden.

4. Schlussendlich wird in einer Schicht mit Stichprobenrate > 0.5 bzw. Stichprobengewicht < 2 vollerhoben.

In Abbildung 1 werden die Stichprobengrössen gegen die Schichtgrössen in logarithmischer Skala dargestellt. Dabei bedeuten die Kreise die durch das Verfahren der optimalen Allokation berechneten Stichprobengrössen und die Kreuze stehen für die angepassten Stichprobengrössen. Die Linien entsprechen den Grenzen gemäss den Anpassungskriterien.

Bei der EVS2007 erhält man mit Gleichung (1) und einem $CV = 0.02$ eine Stichprobengrösse von $n = 10092$ für diese Stichprobe der Nicht-Spezialfälle. Nach der Anpassung beläuft sie sich auf $n = 11664$.

Abbildung 1 Veranschaulichung der optimalen Allokation, $n = 11664$
$\circ$ entspricht Stichproben- vs Schichtgrösse vor der Anpassung,
$\times$ nach der Anpassung.
Die horizontalen Linien bedeuten die untere und obere Grenze der Stichprobengrösse.



2.3 Ziehung

Das Stichprobensystem ist so aufgebaut, dass eine maximale Überdeckung gegenüber dem Vorjahr gewährleistet ist. Tabelle 7 zeigt den Überdeckungsgrad der Jahre 2003 bis 2007 bei den Bruttostichproben. Die Bruttostichproben setzen sich zusammen aus den Nicht-Spezialfällen und den Spezialfällen. Vom Jahr 2002 zum Jahr 2003 verdoppelte sich die Bruttostichprobe, da neu im Jahr 2003 auch die AST-Gruppen- und Bundesbetriebe hinzukamen. Deswegen wird neben dem direkten Vergleich des Jahres 2003 mit dem Jahr 2002 auch das Jahr 2003, um AST-Gruppen und Bundesbetriebe verringert, mit dem Jahr 2002 verglichen. Dieser Vergleich ist in Tabelle 7 abgehoben in der letzten Zeile festgehalten.
Wenn die Betriebszählung (BZ) ins BUR integriert wurde, fiel der Unterschied zwischen der Stichprobe und derjenigen des Vorjahres deutlicher aus.

Tabelle 7 Überdeckungsgrad in der Bruttostichprobe

EVS	Brutto	davon im Vorjahr	Überdeckung	Bemerkung
2007	22098	20180	91%	
2006	25871	20437	79%	BZ 2005 ist ins BUR integriert
2005	23722	21581	91%	
2004	23738	18257	77%	andere AST-Gruppen-Definition im Vorjahr
2003	21472	9182	43%	nur Verbände als Spezialfälle im Vorjahr
2002	11723			
2003	12845	8463	66%	Bruttostichprobe 2003 ohne AST-Gruppen und Bundesbetriebe

Jeder Betrieb k des Rahmens U besitzt eine feste, über dem Intervall $(0, 1)$ uniform verteilte,

unabhängige Zufallszahl u_k oder erhält eine, falls es sich um einen neuen Betrieb handelt. Dieses Prinzip wird auch *permanent random numbers* genannt.

Da im Folgenden die Zweistufigkeit der Schichtung keine Rolle mehr spielt, wird der Einfachheit halber wieder allgemein von den Schichten $h \in H$ gesprochen. Die Stratifizierung H steht für die Einteilung des Rahmens U in Unterbranchen X Grössenklassen X (Nicht-)Spezialfälle.

In der Population P der Nicht-Spezialfälle wird pro Schicht h eine Bernoulli-Stichprobe gezogen, das heisst es werden alle Betriebe k mit der Zufallszahl $u_k < Stichprobenrate_h$ in die Bruttostichprobe genommen. Die Stichprobenrate ist der Quotient aus der durch die Allokation erwarteten Stichprobengrösse $\mathbf{E}(n_h)$ und der Schichtgrösse N_h. Die Anzahl Zufallszahlen kleiner als diese Stichprobenrate ergibt die effektive Stichprobengrösse n_h. Diese Stichprobengrösse n_h ist eine zufällige Grösse und weicht von der durch die optimale Allokation ermittelten festen Stichprobengrösse $\mathbf{E}(n_h)$ ab. Weitergerechnet wird nun mit den Gewichten N_h/n_h, wie wenn man pro Schicht eine einfache Zusfallsstichprobe der Grösse n_h hätte. Im Anhang A wird gezeigt, dass diese Gewichte N_h/n_h die auf die Schichtgrösse N_h kalibrierten Gewichte $N_h/\mathbf{E}(n_h)$ der ursprünglichen Bernoulli-Stichprobe sind.
Bei der EVS2007 beträgt die erwartete Stichprobengrösse $n = 11664$, die effektive Stichprobengrösse beläuft sich auf $n = 12264$. Die Differenzen zwischen effektiven und erwarteten Stichprobengrössen der Grössenklassen 1, 2 und 3, in welchen nicht vollerhoben wird, sind in Tabelle 8 durch ein Stamm-Blatt-Diagramm verdeutlicht.

Tabelle 8 EVS2007: effektive Stichprobengrösse - erwartete Stichprobengrösse für die 120 nicht vollerhobenen Schichten (ohne bbl-Betriebe)

```
Links von | stehen Zehner-Zahlen, rechts von | Einer-Zahlen
  -4 | 6      <--------- Bedeutung: einmal kommt die Differenz -46 vor
  -4 |
  -3 |
  -3 | 2
  -2 |
  -2 | 0
  -1 |
  -1 | 31
  -0 | 99777665
  -0 | 444332222222111111
   0 | 00000000000000000000000000000000000111112233444
   0 | 555566666788999
   1 | 0011111244
   1 | 5577899
   2 | 011    <--------- Bedeutung: vorkommende Differenzen 20, 21, 21
   2 | 6677
   3 | 3
   3 | 7
```

In den Schichten der Bundesbetriebe (*bbl*) wird gemäss vorgegebener Stichprobengrösse gezogen. Das heisst, dass in einer Schicht h die N_h Zufallszahlen der Grösse nach geordnet werden und dann die ersten n_h Betriebe in die Stichprobe aufgenommen werden. Auch hier werden kleine Schichten oder solche der Grössenklasse 5 vollerhoben. Somit fallen hier die erwarteten und effektiven Schichtgrössen zusammen. Über die Jahre können sie leicht modifiziert werden.

Die gesamte Bruttostichprobengrösse bestehend aus Spezialfällen und Nicht-Spezialfällen der Jahre 2002 bis 2007 ist in Tabelle 9 aufgelistet.

Tabelle 9 Bruttogrössen 2002 bis 2007, unterschieden nach Spezialfällen

Aufteilung	ast02	ast03	ast04	ast05	ast06	ast07
Stichprobe	11221	12260	12416	12337	12026	12264
BBL*		50	50	50	50	50
AST-Gruppen		8577	10847	10900	13334	9430
Verbände	502	585	425	435	461	354
Bruttostichprobe	11723	21472	23738	23722	25871	22098

*BBL ist die Abkürzung für das **B**undesamt für **B**auten und **L**ogistik und steht hier für den Datenlieferanten der Bundesbetriebe.

Die effektiven Stichprobengrössen pro Schicht bei den Nicht-Spezialfällen zeigt Tabelle 10.

Der SAS-Code zur Ziehung ist im Anhang B aufgelistet.

3 Hochrechnung

Bei den zu schätzenden Parametern handelt es sich um den Energieverbrauch in Terajoule (TJ) der drei Energieträger Elektrizität (Variable *bvm1*), Heizöl (Variable *cvm1*) und Erdgas (Variable *dvm1*) in den 19 Branchen. Mit der Hilfsgrösse *Vollzeitstellen*, die für jeden Betrieb des Rahmens bekannt ist, lassen sich für die gesuchten Totale Quotientenschätzer bilden. Dabei wird in den Schichten der kleineren Betriebe (Grössenklasse 1 und 2) ein robuster Quotientenschätzer verwendet, sonst der klassische Quotientenschätzer.

Die Zahlen in den nächsten beiden Abschnitten stammen von der EVS 2003 und beziehen sich auf den Stromverbrauch (Variable *bvm1*).

3.1 Schätzer für das Total

Die Population U besteht aus allen Betrieben des Stichprobenrahmens. Der Stichprobenplan ist nach Unterbranchen, Grössenklassen und Spezialfällen (Verbände, AST-Gruppen, Bundesbetriebe) respektive Nicht-Spezialfällen geschichtet: $U = \bigcup_{h \in H} U_h$ mit Stratifizierung H.

Die Bruttostichprobe S mit der Stichprobengrösse $|S| = n$ ist aufgeteilt in disjunkte Schichten $S_h = S \cap U_h$ der Grösse $|S_h| = n_h$. Der Teil der Bruttostichprobe, welcher Verbrauchsangaben liefert, die sogenannte Nettostichprobe $R \subset S$, umfasst $|R| = m$ Betriebe. In einer Schicht h besteht die Nettostichprobe $R_h = R \cap U_h$ mit $R_h \subset S_h$ aus $|R_h| = m_h$ Betrieben. Jeder Betrieb k des Rahmens hat ein Stichprobengewicht $1/\pi_k$, welches in einer Schicht konstant ist: $1/\pi_k = N_h/n_h$ für alle $k \in U_h$.

In der EVS wird davon ausgegangen, dass innerhalb einer Schicht h ein homogenes Antwortverhalten angenommen werden kann. Somit hat man pro Schicht h einen konstanten Antwortausfall-Korrektur-Faktor $\frac{n_h}{m_h}$ und erhält das damit korrigierte Anfangsgewicht

$$d_k = \frac{N_h}{n_h} \cdot \frac{n_h}{m_h} = \frac{N_h}{m_h} \ .$$

Falls in einer Schicht $m_h = 0$ ist, so fällt diese Schicht aus der Nettostichprobe heraus. Dies kommt bei den kleinen Spezialfall-Schichten manchmal vor. Auf diese Problematik wird im Methodenbericht EVS II ausführlich eingegangen.

Tabelle 10 EVS2007: Schichtgrössen der Nicht-Spezialfall-Bruttostichprobe

		clt				All
Branche	Unterbranche	1	2	3	5	
1	1	41	49	57	104	251
2	1	0	51	89	19	209
	2	44	72	36	9	161
3	1	10	3	4	2	19
	2	61	48	44	23	176
	3	70	49	72	46	237
4	1	33	46	50	95	224
5	1	3	6	3	.	12
6	1	51	42	153	14	260
7	1	84	44	47	18	193
8	1	30	18	12	7	67
9	1	75	94	170	89	428
	2	73	24	22	3	122
	3	41	41	48	35	165
	4	38	50	32	14	134
	5	40	52	83	30	205
10	1	40	46	118	143	347
11	1	51	44	92	204	391
	2	37	36	66	1	140
	3	158	77	111	6	352
	4	42	54	89	47	232
	5	40	79	42	2	163
12	1	168	130	221	226	745
13	1	221	100	154	27	502
	2	112	92	159	151	514
	3	304	152	115	76	647
14	1	57	91	167	59	374
	2	245	157	49	8	459
	3	38	39	34	8	119
15	1	40	44	98	163	345
16	1	48	56	104	137	345
17	1	223	162	142	5	532
	2	45	47	99	30	221
	3	38	39	38	30	145
	4	116	45	48	7	216
18	1	304	84	82	22	492
	2	45	38	51	234	368
19	1	71	49	101	150	371
	2	343	112	141	249	845
	3	271	88	112	65	536
All		3801	2550	3355	2558	12264

3.1.1 Robustifizierung

Bei Betrieben als Untersuchungseinheit streuen die Daten stark und haben einzelne Variablen schiefe Verteilungen. Es kann zum Teil massive Ausreisser geben. Deshalb macht es Sinn, robuste Schätzer zu verwenden.

Im Folgenden wird der Robustifizierungsfaktor u_k hergeleitet, mit welchem die Anfangsgewichte d_k bei den robusten Schätzern multipliziert werden (siehe Hulliger (1999)).

Der Energieverbrauch des Betriebes k wird mit y_k bezeichnet, seine Anzahl Vollzeitstellen mit

x_k. Pro Schicht h, in welcher ein robuster Schätzer zur Anwendung kommt (vgl. Tabelle 14), wird eine robuste Steigung $\hat{\beta}_h$ geschätzt durch

$$\hat{\beta}_h = \frac{median_{k \in R_h}(y_k)}{median_{k \in R_h}(x_k)} \ .$$

Damit erhält man die normalisierten Residuen r_k der Schicht h als

$$r_k = (y_k - \hat{\beta}_h x_k)/\sqrt{x_k} \ , \ k \in R_h \ .$$

Mit dem robusten Skalenfaktor

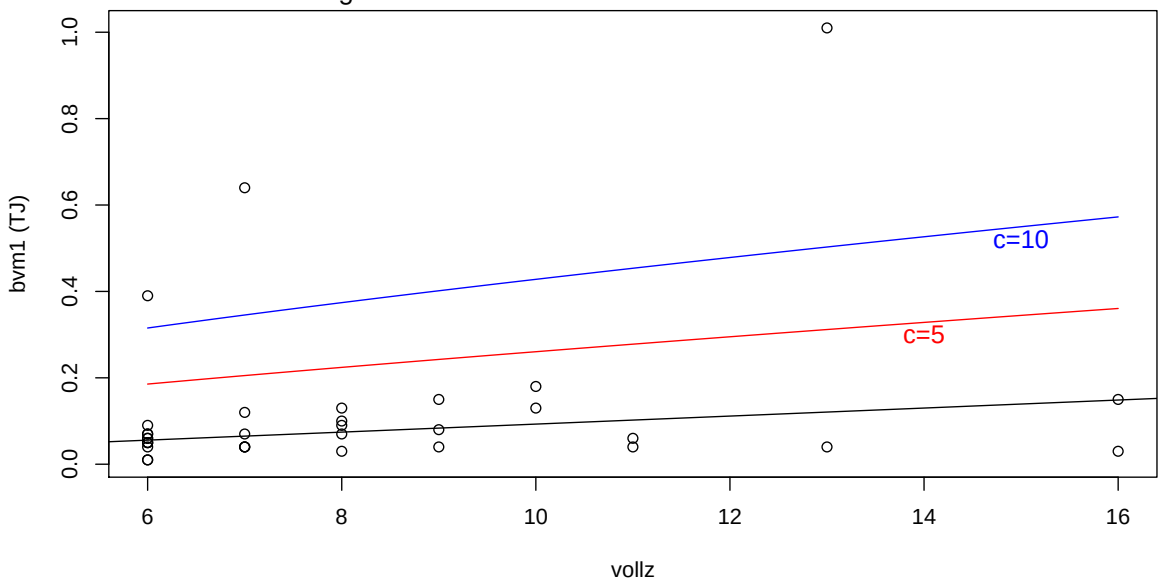
$$\hat{\sigma}_h = median_{k \in R_h}(|r_k|)$$

und der Robustifizierungskonstanten c lassen sich die Robustifizierungsfaktoren u_k für jede Schicht $h \in H$ aufschreiben als

$$u_k = \begin{cases} c\hat{\sigma}_h/|r_k| & \text{falls} \quad |r_k| > c\hat{\sigma}_h \\ 1 & \text{falls} \quad |r_k| \leq c\hat{\sigma}_h \end{cases} \quad \text{für } k \in R_h \ .$$

Je weiter entfernt $|r_k|$ von $c\hat{\sigma}_h$ liegt, desto kleiner wird u_k. Die minim gekrümmten Grenzkurven $f_c(x_k) = \hat{\beta}_h x_k \pm c\hat{\sigma}_h\sqrt{x_k}$ für $c = 5$ resp. $c = 10$ in Abbildung 2 folgen aus der Gleichung $|r_k| = |(y_k - \hat{\beta}_h x_k)/\sqrt{x_k}| = c\hat{\sigma}_h$. Die angepasste Gerade ist eine Gerade durch den Nullpunkt mit der robusten Steigung $\hat{\beta}_h$.

Abbildung 2 EVS 2003: Drei Ausreisser in Schicht 18.1.2 der Nicht-Spezialfälle mit c = 5 resp. c = 10; die Krümmung der Grenzkurven ist nicht offensichtlich.



In Abbildung 2 sieht man ferner, dass die Robustifizierungskonstanten $c = 5$ resp. $c = 10$ dieselben drei Beobachtungen als Ausreisser kennzeichnen. Je kleiner die Robustifizierungskonstante, desto mehr Beobachtungen werden als Ausreisser definiert und desto stärker werden diese extremen Beobachtungen heruntergewichtet.

3.1.2 Der Horvitz-Thompson-Schätzer

Der klassische Horvitz-Thompson-Schätzer für ein Total über die Schicht h ist die gewichtete Summe

$$\widehat{Y}_h = \sum_{k \in R_h} d_k y_k \ . \tag{3}$$

Mit den Gewichten nach Robustifizierung hat er die Form

$$\widehat{Y}_{rob,h} = \sum_{k \in R_h} d_k u_k y_k \ . \tag{4}$$

3.1.3 Der Quotientenschätzer

Mit dem wahren Total X_h der Vollzeitstellen in der Schicht h lautet der klassische Quotientenschätzer für ein Total über die Schicht h

$$\widehat{Y}_h^R = X_h \, \frac{\sum_{k \in R_h} d_k y_k}{\sum_{k \in R_h} d_k x_k} \ . \tag{5}$$

Mit $\widehat{X}_h = \sum_{k \in R_h} d_k x_k$ kann der Quotientenschätzer einer Schicht h auch wie folgt geschrieben werden:

$$\widehat{Y}_h^R = X_h \, \frac{\widehat{Y}_h}{\widehat{X}_h} = \frac{X_h}{\widehat{X}_h} \sum_{k \in R_h} d_k y_k = \sum_{k \in R_h} \underbrace{\frac{X_h}{\widehat{X}_h}}_{\text{Gewichte } g_k} d_k y_k = \sum_{k \in R_h} g_k d_k y_k \ .$$

Dabei werden die $g_k d_k$ als die Gewichte des Quotientenschätzers angeschaut. Mit diesen Gewichten hat auch der Quotientenschätzer die Form einer gewichteten Summe. Bei der EVS-Hochrechnung ist er mit diesen Gewichten $g_k d_k$ programmiert (vgl. Anhang C).

Der robuste Quotientenschätzer hat die Form

$$\widehat{Y}_{rob,h}^R = X_h \, \frac{\sum_{k \in R_h} d_k u_k y_k}{\sum_{k \in R_h} d_k u_k x_k} \ . \tag{6}$$

Unter Verwendung von $\widehat{X}_{rob,h} = \sum_{k \in R_h} d_k u_k x_k$ gilt für den robusten Quotientenschätzer in h:

$$\widehat{Y}_{rob,h}^R = X_h \, \frac{\widehat{Y}_{rob,h}}{\widehat{X}_{rob,h}} = \frac{X_h}{\widehat{X}_{rob,h}} \sum_{k \in R_h} d_k u_k y_k = \sum_{k \in R_h} \underbrace{\frac{X_h}{\widehat{X}_{rob,h}}}_{\text{Gewichte } g_{rob,k}} d_k u_k y_k = \sum_{k \in R_h} g_{rob,k} d_k u_k y_k \ ,$$

wobei die $g_{rob,k} u_k d_k$ als die Gewichte des robusten Quotientenschätzers angeschaut werden. Somit hat auch der robuste Quotientenschätzer die Form einer gewichteten Summe. Mit diesen Gewichten $g_{rob,k} u_k d_k$ ist er bei der EVS-Hochrechnung auch programmiert.

3.1.4 Ausreisserbehandlung

Tabelle 11 zeigt die ursprünglichen und die robustifizierten Gewichte des Horvitz-Thompson-Schätzers für die beiden Robustifizierungskonstanten $c = 5$ und $c = 10$. Bei nicht-extremen

Tabelle 11 EVS 2003: klassische und robustifizierte Gewichte des Horvitz-Thompson-Schätzers bei den drei Ausreisser in Schicht 18.1.2 der Nicht-Spezialfälle

					$c = 10$		$c = 5$	
ID	Schicht	vollz	bvm1	d_k	u_k	$d_k u_k$	u_k	$d_k u_k$
16	18.1.2	6	0.39	41.091	0.7594	31.204	0.3797	15.602
5	18.1.2	7	0.64	41.091	0.4823	19.818	0.2412	9.909
21	18.1.2	13	1.01	41.091	0.4241	17.427	0.2121	8.713

Beobachtungen hat der Robustifizierungsfaktor u_k den Wert 1. Der Effekt der Robustifizierung zeigt sich erst bei den drei extremen Beobachtungen der Schicht 18.1.2 der Nicht-Spezialfälle.

In Tabelle 12 wird deutlich, wie sich die Robustifizierung auf den Quotientenschätzer auswirkt und wie sich dadurch der gewichtete Beitrag $y_k gew$ einer extremen Beobachtung an das Stromverbrauchstotal dieser Schicht 18.1.2 der Nicht-Spezialfälle verändert.

Tabelle 12 EVS 2003: Die drei Ausreisser in Schicht 18.1.2 der Nicht-Spezialfälle mit den Gewichten der klassischen und robusten Quotientenschätzer und den zugehörigen gewichteten Summanden $y_k gew$

				$c = 10$			$c = 5$		
ID	bvm1	$g_k d_k$	$y_k gew$	$g_{rob,k} d_k$	$g_{rob,k} d_k u_k$	$y_k gew$	$g_{rob,k} d_k$	$g_{rob,k} d_k u_k$	$y_k gew$
16	0.39	42.84	16.71	44.85	34.06	13.35	46.00	17.47	6.85
5	0.64	42.84	27.42	44.85	21.63	13.80	46.00	11.10	7.08
21	1.01	42.84	43.27	44.85	19.02	19.17	46.00	9.76	9.83
	2.04		87.40			46.32			23.76

Als Beispiel werden in Tabelle 13 die verschiedenen Schätzer für das Elektrizitätsverbrauchstotal der Gruppe 18.1.2 angegeben. Um die Wirkung der Heruntergewichtung der drei Ausreisser zu sehen, wird ihr Beitrag am Gesamtverbrauch auch noch separat aufgeführt. Wie man sich mit Abbildung 2 gut vorstellen kann, umfasst der Nettobeitrag der drei Ausreisser fast die Hälfte des gesamten Nettobeitrages dieser Schicht.

Tabelle 13 EVS 2003: Geschätzte Totale in Schicht 18.1.2 der Nicht-Spezialfälle

Schicht	Arbeitsstätte				Vollzeit		Verbrauch an Elektrizität				
	pop	brut	net	d_k	pop	net	net	HT	ratio	$c = 10$	$c = 5$
18.1.2	1356	76	33	41.1	12039	281	4.06	166.8	180.4	142.5	120.0
3 Ausr.							2.04	83.8	87.4	46.3	23.8
in %							50	50	48	32	20

Achtung: Obige Resultate sind wegen Rundungen nicht exakt nachrechenbar.

In Schichten mit grossen Anfangsgewichten d_k wird ein robuster Quotientenschätzer verwendet.

Da man mit den Vollzeitstellen über eine Hilfsgrösse verfügt, die mit dem Energieverbrauch korreliert, nimmt man einen Quotientenschätzer für die Schätzung der einzelnen Totale in den Schichten. In Schichten mit grossen Anfangsgewichten d_k wird ein robuster Quotientenschätzer verwendet, wobei die Robustifizierungskonstante umso kleiner gewählt wird, je unterschiedlicher die Anfangsgewichte verteilt sind.
Grosse d_k kommen vor allem in den grossen Schichten der kleinen Betriebe (Grössenklasse 1 und 2) vor. Im Dienstleistungssektor (Branchen 13-19) gibt es viel mehr Betriebe als in der

Industrie, weshalb die Anfangsgewichte bei Dienstleistungsbetrieben sehr gross sein können. Deshalb sollen extreme Werte in diesem Sektor stärker heruntergewichtet werden, was für eine kleine Robustifizierungskonstante c spricht. Aus diesen Gründen wird bei der Wahl der Robustifizierungskonstanten c nach Sektoren und bei der Art des Schätzers nach Grössenklassen unterschieden, wie es in Tabelle 14 aufgelistet ist.

Tabelle 14 Wahl der Robustifizierungskonstanten c und Schätzer in Abhängigkeit der Branchen und Grössenklassen

Sektor	Grössenklasse	Anfangsgewicht	Schätzer
Industrie	1, 2	≥ 10	robuster Quotientenschätzer mit $c = 10$
	1, 2	< 10	Quotientenschätzer
	3, 5		Quotientenschätzer
Dienstleistungen	1, 2	≥ 10	robuster Quotientenschätzer mit $c = 5$
	1, 2	< 10	Quotientenschätzer
	3, 5		Quotientenschätzer

Bei der EVS werden die Verbrauchstotale für die Bereiche Branche und/oder Unterbranche berechnet. Da diese Bereiche Zusammensetzungen mehrerer ganzer Schichten h sind, können die einzelnen mit einem Quotientenschätzer geschätzten Schichttotale addiert werden. Dies wird auch als *separate-ratio*-Ansatz bezeichnet.

3.2 Varianzschätzung

Im Folgenden wird zuerst die Varianzschätzung des Horvitz-Thompson-Schätzers angegeben und dann, darauf aufbauend, diejenige des klassischen und robusten Quotientenschätzers hergeleitet.

Der Horvitz-Thompson-Schätzer eines Totals (3) in einer Schicht h hat die Form

$$\widehat{Y}_h = \sum_{k \in R_h} d_k y_k = \frac{N_h}{m_h} \sum_{k \in R_h} y_k \ .$$

Bei der robusten Form (4) des Horvitz-Thompson Schätzers wird y_k durch $u_k y_k$ ersetzt.

Hier werden die m_h Antwortenden als einfache Zufallsstichprobe der Schicht h betrachtet und so lässt sich die geschätzte Varianz des Horvitz-Thompson-Schätzers für eine Schicht h schreiben als (siehe Särndal (1992), (2.8.13))

$$\widehat{var}(\widehat{Y}_h) = N_h^2 (1 - \frac{m_h}{N_h}) \frac{1}{m_h} s^2_{y,R_h} \ ,$$

wobei s^2_{y,R_h} die Varianz der y in Schicht h schätzt:

$$s^2_{y,R_h} = \frac{1}{m_h - 1} \sum_{k \in R_h} (y_k - \bar{y}_{R_h})^2 \text{ mit } \bar{y}_{R_h} = \frac{1}{m_h} \sum_{k \in R_h} y_k \ .$$

Bei der Varianzschätzung des robusten Horvitz-Thompson Schätzers (4) wird wiederum y_k durch $u_k y_k$ ersetzt.

Um die Varianz des Quotientenschätzers herzuleiten, wird dieser linearisiert und die Varianz dieser Linearisierung geschätzt. Die Linearisierung lässt sich mit den Residuen ausdrücken.

Linearisiert lässt sich der Quotientenschätzer (5) schreiben als

$$\widehat{Y}_h^R \approx Y_h + \sum_{k \in R_h} d_k (y_k - \frac{Y_h}{X_h} x_k) = Y_h + \sum_{k \in R_h} d_k \epsilon_{hk} \ .$$

Die geschätzten Residuen in der Schicht h, welche für die Varianzschätzung benötigt werden, lauten:

$$\hat{\epsilon}_{hk} = y_k - \frac{\widehat{Y}_h}{\widehat{X}_h} x_k \ , \ k \in R_h \ .$$

Analog dem Horvitz-Thompson-Schätzer erhält man nun eine Schätzung der Varianz:

$$\widehat{var}(\widehat{Y}_h^R) = \widehat{var}(\sum_{k \in R_h} d_k \hat{\epsilon}_{hk}) = N_h^2 (1 - \frac{m_h}{N_h}) \frac{1}{m_h} s^2_{\hat{\epsilon}_h, R_h} \ ,$$

wobei $s^2_{\hat{\epsilon}_h, R_h}$ für die geschätzte Varianz der geschätzten Residuen in h steht:

$$s^2_{\hat{\epsilon}_h, R_h} = \frac{1}{m_h - 1} \sum_{k \in R_h} (\hat{\epsilon}_{hk} - \bar{\hat{\epsilon}}_{R_h})^2 \text{ mit } \bar{\hat{\epsilon}}_{R_h} = \frac{1}{m_h} \sum_{i \in R_h} \hat{\epsilon}_{hi} \ .$$

Mit den robust geschätzten und mit u_k gewichteten Residuen

$$\hat{\epsilon}_{rob,hk} = u_k (y_k - \frac{\widehat{Y}_{rob,h}}{\widehat{X}_{rob,h}} x_k) \ , \ k \in R_h$$

erhält man eine Schätzung der Varianz des robusten Quotientenschätzers (6) (Hulliger 1999):

$$\widehat{var}(\widehat{Y}^R_{rob,h}) = m_h (1 - \frac{m_h}{N_h}) \frac{N_h^2}{m_h^2} s^2_{\hat{\epsilon}_{rob,h}, R_h} = N_h^2 (1 - \frac{m_h}{N_h}) \frac{1}{m_h} s^2_{\hat{\epsilon}_{rob,h}, R_h} \ .$$

Der Varianzschätzer $s^2_{\hat{\epsilon}_{rob,h}, R_h}$ der robust geschätzten Residuen $\hat{\epsilon}_{rob,h}$ kann analog zu $s^2_{\hat{\epsilon}_h, R_h}$ aufgeschrieben werden.

4 Longitudinale Zerlegung

Im Rahmen der jährlichen Energieverbrauchsstatistik wird der Verbrauch an Elektrizität, Heizöl oder Erdgas pro Jahr für die Branchen des Dienstleistungssektors und der Industrie geschätzt. Dies ergibt sogenannte transversale Schätzungen. Häufig werden die Schätzungen des aktuellen Jahres mit denjenigen des Vorjahres verglichen und man möchte die beobachteten Veränderungen detailliert erklären. Neben dem Vergleich der Verbrauchsschätzungen betrachtet man hierzu auch die Entwicklung der Anzahl Betriebe des Rahmens, der Brutto- und Nettostichprobe. Diese Untersuchung kann in verschiedenen Phasen stattfinden: beim Datenbereinigungsprozess, bei der Schätzung erster Resultate für die Plausibilisierung oder um transversale Resultate besser interpretieren zu können.

Wir nehmen den Energieverbrauch zweier aufeinanderfolgender Jahre als Ausgangspunkt. Man möchte seine Veränderung in verschiedene Teile zerlegen, die strukturellen Veränderungen in der Population, Veränderungen der Bruttostichprobe oder den Antwortmustern entsprechen. Diese Zerlegung der Anzahl Betriebe einerseits und des geschätzten Energieverbrauchs andererseits in verschiedene Teile wird longitudinale Zerlegung genannt (siehe Salamin, P-A. (2006)).

4.1 Ausgangslage

Als Basis dient eine Folge von Stichprobenrahmen $(U_t, t \in T = \{2002, ..., 2007\})$, in welchen die jährlichen Stichproben S_t gezogen werden. Dabei bezeichnet S_t die Bruttostichprobe des Jahres t und R_t die Nettostichprobe der Betriebe, die Verbrauchsangaben liefern. Wir betrachten zwei aufeinanderfolgende Rahmen $(U_{t-1}, U_t) = (U_1, U_2)$ der Jahre 2003 und 2004. In Tabelle 15 sind einige dieser Ausgangszahlen aufgelistet.

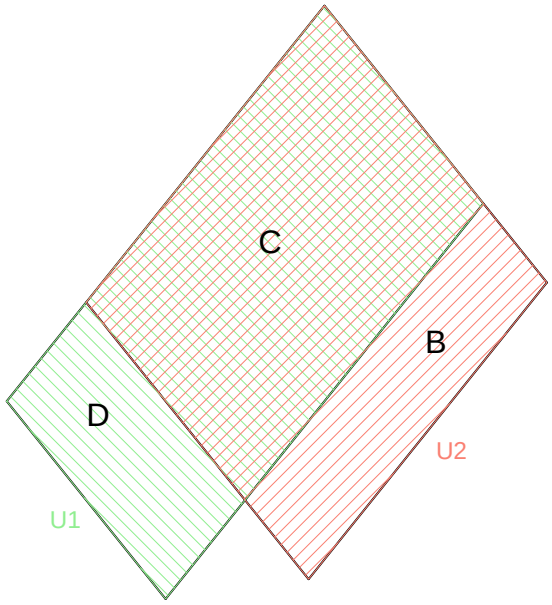
Tabelle 15 Ausgangsdaten für die longitudinale Zerlegung der Jahre 2003 und 2004: Anzahl Betriebe des Rahmens, der Brutto- sowie Nettostichprobe und geschätzter Stromverbrauch (*bvm*) in Terajoule.

Branche	pop03	pop04	brut03	brut04	net03	net04	bvm03	bvm04
1	2704	2615	296	317	213	206	5358	5543
2	1725	1658	377	355	196	186	1214	1182
...								
12	36823	36822	763	762	345	326	1498	1542
13	82591	81318	2974	3700	891	846	13956	14066
...								
19	116191	117457	6558	6614	5390	5382	17354	17890
total	358817	358155	21472	23738	12778	13329	112695	114915

4.2 Longitudinale Zerlegung der Stichprobenrahmen

Die Gesamtheit aller Betriebe zweier aufeinanderfolgender Rahmen $U_1 \cup U_2$ lässt sich in drei disjunkte Bereiche zerlegen, wie Abbildung 3 zeigt: Betriebe, die nur im Rahmen U_1 vorkommen (**D**$eaths$), Betriebe, die in beiden Rahmen vorkommen (**C**$ommon\ units$) und Betriebe, die nur im Rahmen U_2 vorkommen (**B**$irths$).

Abbildung 3 Longitudinale Zerlegung der Stichprobenrahmen $U_1 \cup U_2$



U_1	U_2	Mengenschreibweise	Bezeichnung
1	0	$U_1 \backslash U_2$	D $t-1$
1	1	$U_1 \cap U_2$	C $t-1, t$
0	1	$U_2 \backslash U_1$	B t

Im gemeinsamen Teil C interessiert auch, ob ein Betrieb die Branche wechselte oder während der beiden betrachteten Jahre derselben Branche zugeordnet blieb. Ein Stichprobenrahmen U_t des Jahres t kann disjunkt zerlegt werden in Branchen (oder andere Bereiche) $h \in H$, deren Definition in einem Jahr t fest bleibt:

$$U_t = \bigcup_{h \in H} U_{h,t} \ .$$

Mit der Definition

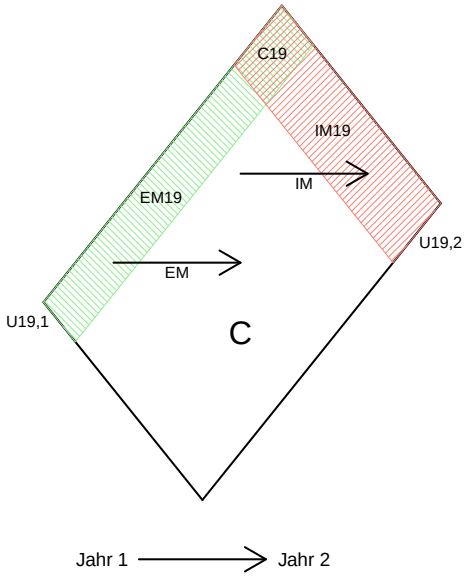
$$C_{ij} = U_{i,t-1} \cap U_{j,t}$$

für Betriebe, die im Jahr $t-1$ zur Branche i und im Jahr t zur Branche j gehören, lässt sich der gemeinsame Teil C bezüglich einer Branche h wie folgt zerlegen:

C$_h$ $t-1, t$	$=$	$C_{h,h}$	feste Einheiten bzgl. h (im Jahr $t-1$ in h, ebenso in t)
EM$_h$ $t-1$	$=$	$\cup_{j \in H, j \neq h} C_{h,j}$	Emigration aus h (im Jahr $t-1$ in h, nicht aber im Jahr t)
IM$_h$ t	$=$	$\cup_{i \in H, i \neq h} C_{i,h}$	Immigration in h (im Jahr $t-1$ nicht in h, im Jahr t schon).

In Abbildung 4 ist diese Branchenzerlegung des gemeinsamen Teils $C = U_1 \cap U_2$ schematisch dargestellt am Beispiel der Branche 19. Der Pfeil symbolisiert die zeitliche Veränderung: Pfeilanfang in derjenigen Menge, in welcher sich der Betrieb im Jahr $t-1$ befindet, Pfeilende in der Menge, welcher der Betrieb im Folgejahr t zugeordnet ist.

Abbildung 4 Longitudinale Zerlegung des gemeinsamen Teils $C = U_1 \cap U_2$ bezüglich der Branche 19



Somit lässt sich nun die Gesamtheit aller Betriebe zweier aufeinanderfolgender Rahmen in die drei disjunkten Bereiche D, C und B zerlegen. Der gemeinsame Bereich C wird noch weiter zerlegt, um das Verhalten bezüglich Branchenverbleib oder Branchenwechsel zu beschreiben.

Tabelle 16 Longitudinale Zerlegung der Stichprobenrahmen der Jahre 2003 und 2004: Anzahl Betriebe in den verschiedenen Teilmengen.

		weg	gemeinsam			neu	
Branche	pop03	D03	EM_h03	C_h0304	IM_h04	B04	pop04
1	2704	147	28	2529	18	68	2615
2	1725	138	18	1569	16	73	1658
...							
12	36823	1751	158	34914	91	1817	36822
13	82591	5127	356	77108	354	3856	81318
...							
19	116191	7087	365	108739	464	8254	117457
total	358817	18679	1418	338720	1418	18017	358155

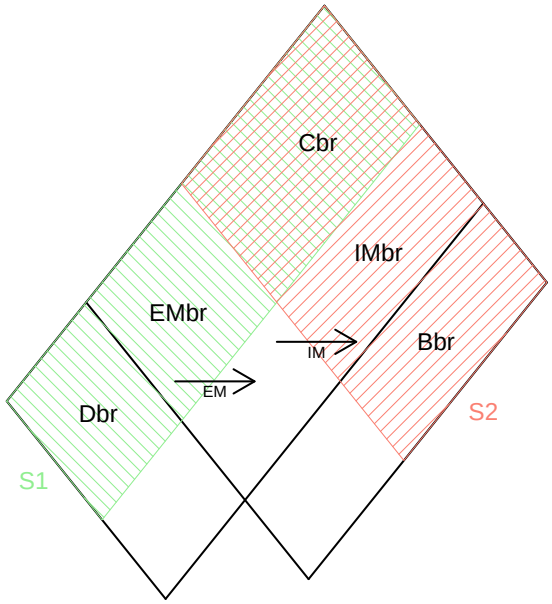
Die longitudinale Zerlegung für die Anzahl Betriebe der pop03-pop04-Kolonnen der Tabelle 15 ist in Tabelle 16 aufgeführt.
Offensichtlich ist die totale Anzahl EM_h03 aller Betriebe, die im Jahr 2003 die Branche wechseln (aus einer Branche emigrieren), gleich der totalen Anzahl IM_h04 aller Betriebe des Jahres 2004, die neu in eine Branche kommen (in die Branche immigrieren).

Diese longitudinale Zerlegung der Stichprobenrahmen kann benützt werden um Branchen zu entdecken, in welchen ungewöhnliche Ereignisse stattfinden, wie zum Beispiel unübliche Anteile Betriebe, die aufhören oder neugegründet werden (D03 und B04 in Branche 1). Auch unüblich viele Branchenwechsel können entdeckt und somit überprüft werden (EM_h03 und IM_h04 in Branche 12).

4.3 Longitudinale Zerlegung der Bruttostichproben

Nun wird in der Welt der beiden Stichprobenrahmen die Gesamtheit aller Betriebe zweier aufeinanderfolgender Bruttostichproben $S_1 \cup S_2$ betrachtet und in fünf disjunkte Bereiche zerlegt.

Abbildung 5 Longitudinale Zerlegung der Bruttostichproben $S_1 \cup S_2$



U_1	U_2	S_1	S_2	Mengenschreibweise	Bezeichnung
1	0	1	0	$S_1 \cap D$	$D_{br}\ t-1$
1	1	1	0	$(S_1 \backslash S_2) \cap C$	$EM_{br}\ t-1$
1	1	1	1	$S_1 \cap S_2$	$C_{br}\ t-1,t$
1	1	0	1	$(S_2 \backslash S_1) \cap C$	$IM_{br}\ t$
0	1	0	1	$S_2 \cap B$	$B_{br}\ t$

Dabei bedeuten

$D_{br}\ t-1$	Betrieb ist im Jahr $t-1$ in der Bruttostichprobe, im Folgejahr existiert er nicht mehr.
$EM_{br}\ t-1$	Betrieb ist nur im Jahr $t-1$ in der Bruttostichprobe, im Folgejahr nicht mehr
$IM_{br}\ t$	Betrieb ist erst im Jahr t in der Bruttostichprobe, im Jahr davor nicht.
$B_{br}\ t$	Betrieb ist im Jahr t in der Bruttostichprobe, im Jahr davor existiert er noch nicht.

Wieder lässt sich der gemeinsame Teil beider Bruttostichproben $S_1 \cap S_2 = C_{br}\ t-1,t$ noch weiter zerlegen, um das Verhalten bezüglich Branchenverbleib oder Branchenwechsel zu beschreiben (vgl. Abbildung 4)

$EM_h\ t-1$	Emigration aus h (im Jahr $t-1$ in h, nicht aber im Jahr t)
$C_h\ t-1,t$	in beiden Bruttostichproben in Branche h (im Jahr $t-1$ in h, ebenso in t)
$IM_h\ t$	Immigration in h (im Jahr $t-1$ nicht in h, im Jahr t schon).

Die longitudinale Zerlegung für die Anzahl Betriebe der brut03-brut04-Kolonnen der Tabelle 15 ist in Tabelle 17 aufgeführt.

Tabelle 17 Longitudinale Zerlegung der Bruttostichproben der Jahre 2003 und 2004:

		weg		gemeinsam			neu		
Branche	brut03	D_{br}03	EM_{br}03	EM_h03	C_h0304	IM_h04	IM_{br}04	B_{br}04	brut04
1	296	14	6	2	274	1	36	6	317
2	377	29	6	2	340	1	6	8	355
...									
12	763	30	25	3	705	3	38	16	762
13	2974	232	269	28	2445	10	1147	98	3700
...									
19	6558	882	195	15	5466	26	596	526	6614
total	21472	1595	1620	84	18173	84	4428	1053	23738

Wieder entspricht die totale Anzahl EM_h03 aller Betriebe der Bruttostichprobe des Jahres 2003, die im Jahr 2004 in eine andere Branche kamen, der totalen Anzahl IM_h04 aller Betriebe der Bruttostichprobe des Jahres 2004, die im Jahr 2003 in einer anderen Branche figurierten.

Lesebeispiel anhand der Branche 19:

- Von den 6558 Betrieben in Branche 19 der Bruttostichprobe des Jahres 2003 figurieren im folgenden Jahr 1077 Betriebe nicht mehr in dieser Branche 19, weil sie im Jahr 2004 nicht mehr existieren (882 Betriebe) oder weil sie nicht mehr in die Bruttostichprobe 2004 gezogen werden (195 Betriebe).
- Von den restlichen 5481 Betrieben erscheinen 5466 Betriebe auch im Jahr 2004 in Branche 19, 15 kommen in eine andere Branche zu liegen.
- Aus der Sicht des Jahres 2004 kommen 1122 Betriebe der gesamten 6614 Betriebe der Bruttostichprobe 2004 in Branche 19 neu in diese Branche hinein, einerseits weil sie

2003 noch nicht existieren (526 Betriebe) oder 2003 noch nicht in der Bruttostichprobe sind (596).

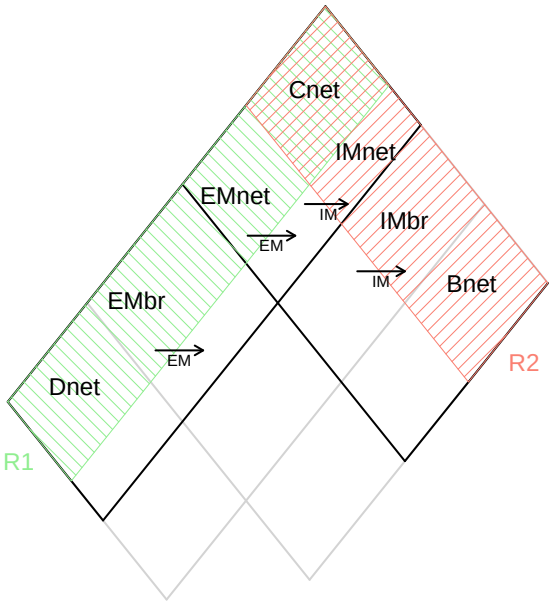
- Von den restlichen 5492 Betrieben gehören 5466 schon 2003 zur Branche 19 und 26 gehören 2003 zu einer andern Branche.

Eine Zerlegung wie in Tabelle 17 gibt Aufschluss über die Zusammensetzung der Bruttostichprobe.

4.4 Longitudinale Zerlegung der Nettostichproben

Nun wird in der Welt der beiden Bruttostichproben die Gesamtheit aller Betriebe zweier aufeinanderfolgender Nettostichproben $R_1 \cup R_2$ betrachtet und in sieben disjunkte Bereiche zerlegt.

Abbildung 6 Longitudinale Zerlegung der Nettostichproben $R_1 \cup R_2$



U_1	U_2	S_1	S_2	R_1	R_2	Mengenschreibweise	Bezeichnung
1	0	1	0	1	0	$R_1 \cap D$	$D_{net}\ t-1$
1	1	1	0	1	0	$R_1 \cap (S_1 \backslash S_2) \cap C$	$EM_{br}\ t-1$
1	1	1	1	1	0	$(R_1 \backslash R_2) \cap (S_1 \cap S_2) \cap C$	$EM_{net}\ t-1$
1	1	1	1	1	1	$R_1 \cap R_2$	$C_{net}\ t-1, t$
1	1	1	1	0	1	$(R_2 \backslash R_1) \cap (S_1 \cap S_2) \cap C$	$IM_{net}\ t$
1	1	0	1	0	1	$R_2 \cap (S_2 \backslash S_1) \cap C$	$IM_{br}\ t$
0	1	0	1	0	1	$R_2 \cap B$	$B_{net}\ t$

Dabei bedeuten

$D_{net}\ t-1$	Betrieb ist im Jahr $t-1$ in der Nettostichprobe, im Folgejahr existiert er nicht mehr.
$EM_{br}\ t-1$	Betrieb ist im Jahr $t-1$ in der Nettostichprobe, im Folgejahr nur im Rahmen.
$EM_{net}\ t-1$	Betrieb ist im Jahr $t-1$ in der Nettostichprobe, im Folgejahr nur in der Bruttostp.
$IM_{net}\ t$	Betrieb ist im Jahr t in der Nettostichprobe, im Jahr davor nur in der Bruttostp.
$IM_{br}\ t$	Betrieb ist im Jahr t in der Nettostichprobe, im Jahr davor nur im Rahmen.
$B_{net}\ t$	Betrieb ist im Jahr t in der Nettostichprobe, im Jahr davor existiert er noch nicht.

Wieder lässt sich der gemeinsame Teil beider Nettostichproben $R_1 \cap R_2 = C_{net}\ t-1,t$ noch weiter zerlegen, um das Verhalten bezüglich Branchenverbleib oder Branchenwechsel zu beschreiben (vgl. Abbildung 4)

$EM_h\ t-1$	Emigration aus h (im Jahr $t-1$ in h, nicht aber im Jahr t)
$C_h\ t-1,t$	in beiden Nettostichproben in Branche h (im Jahr $t-1$ in h, ebenso in t)
$IM_h\ t$	Immigration in h (im Jahr $t-1$ nicht in h, im Jahr t schon).

Die longitudinale Zerlegung für die Anzahl Betriebe der net03-net04-Kolonnen der Tabelle 15 ist in Tabelle 18 aufgeführt.

Tabelle 18 Longitudinale Zerlegung der Nettostichproben der Jahre 2003 und 2004

		weg			gemeinsam			neu			
Branche	net03	$D_{net}03$	$EM_{br}03$	$EM_{net}03$	EM_h03	C_h0304	IM_h04	$IM_{net}04$	$IM_{br}04$	$B_{net}04$	net04
1	**213**	3	0	30	1	179	0	18	5	4	206
2	196	8	5	28	1	154	1	25	3	3	**186**
...											
12	345	5	5	82	2	251	2	58	11	4	326
13	891	14	44	200	20	613	6	167	38	22	846
...											
19	5390	798	27	183	4	4378	21	159	366	458	5382
total	12778	955	344	1390	51	10038	51	1159	1425	656	13329

Wieder müssen die beiden gesamthaften Branchen-Emigrations- bzw. Branchen-Immigrationszahlen gleich sein (hier 51 Betriebe).

Lesebeispiel:

- Von den 5390 Betrieben der Nettostichprobe des Jahres 2003 der Branche 19 fielen 908 Betriebe aus dieser Branche heraus, weil sie ein Jahr später nicht mehr existierten (798 Betriebe), nicht mehr in die Bruttostichprobe gezogen wurden (27) oder keine Verbrauchsangaben lieferten (183).

- Von den 4382 $(= 4378 + 4)$ Betrieben der Branche 19, die beide Jahre die Verbrauchszahlen lieferten, figurierten 4 Betriebe im Jahr 2004 in einer andern Branche.

- Die 213 Betriebe der Nettostichprobe des Jahres 2003 in Branche 1 werden auf fünf disjunkte Teilbereiche aufgeteilt (Spalte 3 bis 7).

- Die 186 Betriebe der Nettostichprobe des Jahres 2004 in Branche 2 werden auf fünf disjunkte Teilbereiche aufgeteilt (Spalte 7 bis 11).

4.5 Longitudinale Zerlegung der Schätzer

Nimmt man die Betriebe der longitudinalen Zerlegung der Nettostichproben (Tabelle 18) und ihre Gewichte, so kann man damit in allen disjunkten Bereichen den Energieverbrauch schätzen. Zum Beispiel berechnet sich der Energieverbrauch aller Betriebe, die im Vorjahr Verbrauchszahlen lieferten, aber im aktuellen Jahr verschwunden sind, durch

$$\widehat{Y}_D = \sum_{k \in D_{net}\ t-1} w_k y_k$$

mit den robustifizierten Gewichten w_k. Schränkt man sich auf die Betriebe einer Branche h ein, so hat man die entsprechenden Schätzungen für eine Branche.

Tabelle 19 zeigt die longitudinale Zerlegung der Schätzer des Stromverbrauchs. Die letzten vier Kolonnen geben relative Veränderungen vom Jahr 2003 zum Jahr 2004 in Prozent vom Gesamtverbrauch des Jahres 2003 an. Da einzelne disjunkte Bereiche nur über wenig Beobachtungen verfügen können (vgl Tabelle 18), sind diese Schätzungen zum Teil mit grosser Ungenauigkeit behaftet und eignen sich nicht einfach so für eine Publikation. Sie dienen intern vor allem dazu, Inkonsistenzen oder Ungereimtheiten auf die Spur zu kommen.

Im Folgenden werden für die disjunkten Bereiche dieselben Bezeichnungen wie in den vorangegangenen Unterkapiteln verwendet, jedoch sind nicht mehr die Anzahl Betriebe eines disjunkten Bereichs gemeint, sondern die Schätzung des Energieverbrauchs eines bestimmten Energieträgers.

So wird die prozentuale Veränderung des Gesamtverbrauchs $ev(\%)$ eines Energieträgers E (Elektrizität, Heizöl oder Erdgas) definiert als

$$ev(\%) = 100 \cdot \frac{E04 - E03}{E03} = 100 \cdot \frac{\Delta E}{E03} \ .$$

Der Ausdruck ΔE lässt sich gemäss der Nettostichprobe in Teilverbräuche zerlegen.

$$\begin{aligned}
\Delta E &= (C_h04 + IM_h04 + IM_{net}04 + IM_{br}04 + B_{net}04) \\
&\quad -(C_h03 + EM_h03 + EM_{net}03 + EM_{br}03 + D_{net}03) \\
&= (C_h04 + IM_{net}04 + IM_{br}04 - C_h03 - EM_{net}03 - EM_{br}03) \\
&\quad +(IM_h04 - EM_h03) + (B_{net}04 - D_{net}03) \\
&= \Delta C + \Delta ST + \Delta DEM \ .
\end{aligned}$$

Die relative Veränderung des Energieverbrauchs evC der Betriebe, die die Branche nicht wechselten aber vielleicht das andere Jahr nicht mehr oder noch nicht in der Brutto- oder Nettostichprobe waren lässt sich schreiben als:

$$evC = 100 \cdot \frac{\Delta C}{E03} = 100 \cdot \frac{C_h04 + IM_{net}04 + IM_{br}04 - C_h03 - EM_{net}03 - EM_{br}03}{E03} \ .$$

Die relative Veränderung des Energieverbrauchs $evST$ der Betriebe, die jeweils das andere Jahr die Branche gewechselt hatten, wird angegeben durch:

$$evST = 100 \cdot \frac{\Delta ST}{E03} = 100 \cdot \frac{IM_h04 - EM_h03}{E03} \ .$$

Der letzte Summand $evDEM$ beschreibt die relative Veränderung des Energieverbrauchs der neu entstandenen zu den verschwundenen Betrieben:

$$evDEM = 100 \cdot \frac{\Delta DEM}{E03} = 100 \cdot \frac{B04 - D03}{E03} \ .$$

Tabelle 19 Longitudinale Zerlegung der Schätzer der Verbrauchstotale bei der Elektrizität (*bvm*) in den Jahren 2003 und 2004

Branche	bvm03	$D_{net}03$	$EM_{br}03$	$EM_{net}03$	EM_h03	C_h03	...
1	**5358.2**	18.4	0	560	1.8	4778.1	
2	1213.6	92.7	4.6	78.7	0.4	1037.1	
...							
12	1498.5	17.9	13.2	317.0	8.2	1142.2	
13	13956.2	272.1	191.6	3712.0	497.0	9283.6	
...							
19	17353.8	505.5	172.2	2835.8	120.6	13719.7	
total	112694.7	2020.7	1981.5	15989.8	910.8	91792.1	

...	Branche	C_h04	IM_h04	$IM_{net}04$	$IM_{br}04$	$B_{net}04$	bvm04	ev(%)	evC	evST	evDEM
	1	5062.1	0	346.9	41.6	92.8	**5543.4**	3.5	2.1	0.0	1.4
	2	1015	1.4	129.3	22.3	14.4	1182.4	-2.6	3.8	0.1	-6.5
	...										
	12	1281.3	32.5	175.6	41.3	10.9	1541.5	2.9	1.7	1.6	-0.5
	13	8780.6	34.5	4426.2	609.8	214.5	14065.6	0.8	4.5	-3.3	-0.4
	...										
	19	13625.9	314.6	2690.7	910.5	348	17889.7	3.1	2.9	1.1	-0.9
	total	92681.9	855.9	16891.9	3163.8	1321.9	114915.4	2.0	2.6	0.0	-0.6

Die Interpretation der Zahlen der Tabelle 19 liefert interessante Einsichten zur Veränderung des Stromverbrauchs in zwei aufeinanderfolgenden Jahren. So war in Branche 2 der 2004-Stromverbrauch der neu entstandenen Betriebe um 6.5 Prozent (Kolonne evDEM) kleiner als der 2003-Stromverbrauch der verschwundenen Betriebe, gemessen am gesamten Stromverbrauch des Jahres 2003 in Branche 2.
Im Rahmen des *Macro Editing* kann Tabelle 19 auch verwendet werden, um Widersprüchlichkeiten in den Daten zu finden.

5 Imputation

In diesem Kapitel wird das Vorgehen und die Anwendung von Imputationen beschrieben. Bei Mehrbetriebsunternehmen muss in der Phase der Datenerhebung manchmal ein Verbrauchswert, der von einem Betrieb für das ganze Mehrbetriebsunternehmen geliefert wird, gemäss Grösse auf die einzelnen Betriebe aufgeteilt werden.

In der Phase der Hochrechnung werden fehlende Werte durch nonresponse-Faktoren respektive Hochrechnungsgewichte korrigiert, so wie es im Kapitel 3, Hochrechnung, beschrieben ist. In dieser Hochrechnungs-Phase könnten die fehlenden Werte auch imputiert werden. Diese Imputation ist im Anhang ausgeführt, da sie hier nicht zur Anwendung kommt.

5.1 Imputation in der Phase der Datenerhebung

Anhand der Bruttostichprobe wird von den Firmen Polyquest und Helbling der Versand des Fragebogens zum Energieverbrauch bei den Betrieben vorbereitet und durchgeführt. Mit den retournierten, ausgefüllten Fragebögen wird die Nettostichprobe mit den Verbrauchswerten erstellt. Es kann sein, dass bei Unternehmen (AST-Gruppen) nur ein Betrieb Verbrauchszahlen für das gesamte Unternehmen liefert. Dann werden die Verbrauchszahlen von Polyquest/Helbling gemäss Vollzeitstellen auf die einzelnen Betriebe des Unternehmens aufgeteilt.

Diese Situation tritt auch bei Verbänden auf. Dieser Vorgang wird *Imputation* genannt. Die Variable agg der Nettostichprobe hält fest, ob und welche dieser Situationen vorliegt. Die möglichen Werte der Variablen agg sind in Tabelle 20 aufgelistet, ihre Häufigkeiten bei der Nettostichprobe der EVS2007 bezüglich des Elektrizitätskonsums in Tabelle 21. In der Variable $burnr_agg$ ist jeweils die burnr des Spender-Betriebs festgehalten.

Tabelle 20 Die kategorielle Variable agg

agg	Bedeutung
20	Verbrauchsangabe gilt für den Betrieb selber und andere Betriebe
10	Betrieb erhält Verbrauchsanteil von einem Betrieb mit $agg = 20$
40	Verbrauchsangabe gilt nur für andere Betriebe*
30	Betrieb erhält Verbrauchsanteil von einem Betrieb mit $agg = 40$
0 oder .	sonst

*$agg = 40$ sollte also in der Nettostichprobe nicht vorkommen.

Tabelle 21 EVS2007: Häufigkeiten der Variablen agg bei der Elektrizität ($clt = 5$ entspricht der grössten Grössenklasse, $spf = 1$ den Spezialfällen AST-Gruppen, Verbänden und BBL)

agg	ganze Nettostichprobe	nur $clt = 5$ oder $spf = 1$
20	156	129
10	5560	5526
40	0	0
30	6	5
0 oder .	6218	1823
all	11940	7483

Es kann auch vorkommen, dass in einem Gebäude mehrere Betriebe verschiedener Unternehmen ihren Sitz haben und für das ganze Gebäude nur eine Verbrauchszahl geliefert wird. Auch hier versucht Polyquest/Helbling den gelieferten Wert möglichst vernünftig auf die einzelnen Betriebe des Gebäudes aufzuteilen. Sie müssen dabei berücksichtigen, dass einzelne Betriebe dieses Gebäudes eventuell selber einen Wert liefern. Auch ist es nicht immer gegeben, alle Betriebe eines Gebäudes zu kennen, da nicht alle in der Stichprobe sein müssen.

Die Beurteilung, ob eine Verbrauchsangabe für sich alleine steht oder für mehrere Betriebe gilt, ist nicht immer einfach. Manchmal wird dies erst beim Vergleich der hochgerechneten Verbrauchszahlen über zwei Jahre erkannt. Dann folgt eine Bereinigungsrunde und die Hochrechnung wird wiederholt. Es kann sein, dass es mehrere solche Runden gibt, bis die Daten und infolgedessen auch die Hochrechnung stabil sind.

Die bereinigte Nettostichprobe, welche dem BFS für die Hochrechnung geliefert wird, enthält also schon imputierte Werte, die anhand der Variablen agg feststellbar sind.

Tabelle 22 zeigt die Antwortquoten bei den Nicht-Spezialfällen und bei den Spezialfällen pro Branche. In Tabelle 23 wird die Antwortquote bei den Nicht-Spezialfällen noch nach den Grössenklassen aufgeschlüsselt. Im Allgemeinen ist die Rücklaufquote umso höher, je grösser der Betrieb ist.

In den Tabellen 22 und 23 zu den fehlenden Werten sind die 22098 Betriebe der Bruttostichprobe der EVS2007 schon um die 48 Betriebe, die Konkurs gingen, reduziert.

Tabelle 22 EVS 2007 Elektrizität: Antwortquoten pro Spezialfälle (SPF)

	Nicht-SPF		nur SPF	
Branche	brut	resp	brut	resp
1	251	0.73	48	0.67
2	369	0.49	5	0.60
3	427	0.51	19	0.79
4	224	0.67	20	0.45
5	12	0.42	30	0.93
6	260	0.55	10	0.90
7	193	0.54	14	0.71
8	67	0.48	5	1.00
9	1052	0.58	79	0.56
10	347	0.67	88	0.53
11	1274	0.54	55	0.44
12	745	0.44	30	0.17
13	1651	0.44	1632	0.08
14	942	0.38	423	0.01
15	344	0.61	1606	0.34
16	344	0.52	465	0.54
17	1110	0.51	275	0.15
18	859	0.54	99	0.23
19	1745	0.46	4931	0.92
all	12216	0.51	9834	0.58

Tabelle 23 EVS 2007 Elektrizität: Antwortquoten pro Nicht-Spezialfälle

	Grössenklasse 1			Grössenklasse 2			Grössenklasse 3			Grössenklasse 5		
Branche	brut	net	resp	brut	net	resp	brut	net	resp	brut	net	resp
1	41	23	0.56	49	27	0.55	57	44	0.77	104	88	0.85
2	93	33	0.35	123	59	0.48	125	70	0.56	28	20	0.71
3	139	44	0.32	97	43	0.44	120	76	0.63	71	54	0.76
4	33	11	0.33	46	25	0.54	50	34	0.68	95	79	0.83
5	3	1	0.33	6	1	0.17	3	3	1.00	0	0	0.00
6	51	21	0.41	42	23	0.55	153	87	0.57	14	11	0.79
7	84	32	0.38	44	24	0.55	47	32	0.68	18	16	0.89
8	30	9	0.30	18	7	0.39	12	9	0.75	7	7	1.00
9	266	99	0.37	260	145	0.56	355	226	0.64	171	140	0.82
10	40	14	0.35	46	19	0.41	118	87	0.74	143	112	0.78
11	326	106	0.33	289	147	0.51	400	246	0.62	259	194	0.75
12	168	42	0.25	130	47	0.36	221	119	0.54	226	119	0.53
13	626	207	0.33	343	146	0.43	428	218	0.51	254	162	0.64
14	337	89	0.26	283	102	0.36	249	118	0.47	73	50	0.68
15	40	20	0.50	44	32	0.73	98	59	0.60	162	100	0.62
16	48	32	0.67	56	34	0.61	103	47	0.46	137	67	0.49
17	420	187	0.45	292	158	0.54	326	189	0.58	72	32	0.44
18	348	124	0.36	122	69	0.57	133	83	0.62	256	191	0.75
19	681	211	0.31	249	106	0.43	353	191	0.54	462	289	0.63
all	3774	1305	0.35	2539	1214	0.48	3351	1938	0.58	2552	1731	0.68

6 Jährlich wiederkehrende Abläufe

Jeweils gegen Ende Jahr steht die Ziehung der EVS-Stichprobe an. Dafür wird mit einem Auszug aus dem aktuellen Betriebs- und Unternehmensregister BUR der Rahmen erstellt und die Schichten definiert: Branchen, Unterbranchen, Grössenklassen, Unternehmen (*astgr*), Verbände (*assoc*) und die Bundesbetriebe (*bbl*). Die letzten drei Unterteilungen bilden die sogenannten Spezialfälle. Vor der eigentlichen Ziehung besteht der grösste Teil der Arbeit darin, die Spezialfälle zu erkennen. Dies geschieht aufgrund der letztjährigen Spezialfälle, der Erfahrung damit bei der letztjährigen Datenerhebung und durch Vergleich mit Profiling-Betrieben.

Sind die Schichten definiert, wird die Bruttostichprobe gezogen. Polyquest wird der Stichprobenplan und die Bruttostichprobe, versehen mit den Adressen, geliefert. Dabei wird die Bruttostichprobe insofern *aufgeblasen*, als jeder in der Bruttostichprobe vorkommende Unternehmensidentifikator (*entid*) mit allen zugehörigen Betrieben resp. Arbeitsstätten (*burnr*) in der Stichprobe erscheint, wobei natürlich erkennbar ist, ob er in die Stichprobe gezogen wurde (*brut=1*) oder nicht (*brut=0*). Für die datenerhebende Stelle ist es im Hinblick auf die Datenaufbereitung wichtig, die Struktur der Unternehmen zu kennen, da sie zum Teil Energieverbrauchsangaben eines Betriebes erhalten, der für mehrere Betriebe eines Unternehmens geantwortet hat. Ihre Arbeit besteht dann darin, diese Angaben gemäss der Betriebsgrösse (Anzahl Vollzeitstellen) auf die einzelnen Betriebe aufzuteilen. Dies ist die im vorhergehenden Kapitel erwähnte Imputation in der Phase der Datenerhebung.

Während der Datenerhebung erfolgen zum Teil Plausibilisierungsanfragen an METH, weil ein Betrieb seine Adresse änderte, nicht mehr existiert, die Aktivität änderte oder andere Widersprüchlichkeiten aufgetreten sind.

Gegen Mitte Jahr ist die Datenerhebung abgeschlossen und wird mit der bereinigten Nettostichprobe sowohl bei METH wie bei Polyquest die Hochrechnung durchgeführt. Es wird der Energieverbrauch pro Branche und auch feiner pro Schicht geschätzt. Ersteres ist für die Veröffentlichung bestimmt, letzteres gibt eine Übersicht über die Schichtgrössen, den Hochrechnungsfaktor und die Genauigkeit und dient der Datenbereinigung.
METH erstellt auch eine longitudinale Zerlegung, das heisst einen Vergleich zweier Jahre, in welcher Abgänge, Neuzugänge und Branchenwechsel pro Branche erkennbar sind. Diese dient vor allem der Datenbereinigung respektive dem sogenannten *macro editing*.

7 Stichproben und Resultate

In der jeweils ersten Tabelle pro Jahr wird die Schicht- und Stichprobengrösse pro primärer und sekundärer Schicht angegeben. Es wird hier nicht nach Spezialfällen und Nicht-Spezialfällen unterschieden.

Die aufgeführten Hochrechnungsresultate beziehen sich jeweils auf die jährliche Rechnung circa Mitte Jahr, nachdem die Daten erhoben, aufbereitet und bereinigt wurden. Spätere Revisionen und Korrekturen sind nicht berücksichtigt, weswegen diese Werte nicht mit denjenigen in der Dokumentation "Energieverbrauchsstatistik in der Industrie und im Dienstleistungssektor" des BFE (siehe Gülden, J. et al. (2007)) übereinstimmen müssen.

Es werden einerseits die hochgerechneten Verbrauchszahlen pro Branche tabelliert und andererseits auch als Beispiel für die unterschiedlichen Schätzer die Werte der Unterbranche 18.1 für jede ihrer Schichten detailiert aufgeführt. Insbesondere werden der klassische Horvitz-Thompson-Schätzer (Endung *_hr*) und der klassische (Endung *_r*) und robuste Quotientenschätzer (Endung *_rr*) samt ihren Standardabweichungen angegeben. Es zeigen sich zum Teil massive Unterschiede. Der endgültig verwendete Schätzer pro Schicht richtet sich nach Tabelle 14.

In den 18.1-Beispiel-Tabellen 25, 28, 31, 34, 37 und 40 werden jeweils alle Schichten der Nicht-Spezialfälle, deren Zwischentotal und das 18.1-Gesamttotal samt Standardabweichungen aufgeführt. Nur bei Tabelle 25 zum Jahr 2002 fallen das 18.1-Zwischentotal und das 18.1-Gesamttotal zusammen, da in der Unterbranche 18.1 im Jahr 2002 keine Spezialfälle vorhanden sind.

In den Hochrechnungstabellen über alle Branchen sind die maximalen Schlussgewichte pro Branche aufgeführt. Generell sind sie in der Grössenklasse 1 bei weitem am grössten, bei den Spezialfällen sind sie kleiner. Die ursprünglich auf 150 beschränkten Anfangsgewichte (vgl Abbildung 1) können durch die Antwortausfall-Korrektur massiv grösser werden.

Der totale hochgerechnete Elektrizitätsverbrauch bei den Spezialfällen ab 2003 beträgt jeweils knapp ein Drittel des Gesamtelektrizitätsverbrauchs (vgl. Tabellen 29, 32, 35, 38 und 41).

7.1 EVS 2002

Nur Verbände ($assoc$) werden als Spezialfälle geführt. Sie werden aufgrund einer Liste von Polyquest identifiziert und vollerhoben.

Tabelle 24 EVS 2002: Anzahl Betriebe im Rahmen und in der Bruttostichprobe

(Unter-)		clt 1		clt 2		clt 3		clt 5	Total	
Branche		pop	brut	pop	brut	pop	brut	pop	pop	brut
1	1	1877	40	532	35	300	54	119	2828	248
2	1	397	34	138	37	130	130	29	694	230
	2	853	37	114	38	61	61	15	1043	151
3	1	11	11	4	4	8	8	9	32	32
	2	56	56	60	60	57	57	35	208	208
	3	3328	49	850	40	322	50	59	4559	198
4	1	352	69	173	81	199	133	97	821	380
5	1	8	8	18	18	15	15	5	46	46
6	1	890	49	303	44	157	157	23	1373	273
7	1	63	63	53	53	48	48	30	194	194
8	1	25	25	21	21	16	16	13	75	75
9	1	5295	69	1660	50	725	90	86	7766	295
	2	65	65	33	33	23	23	5	126	126
	3	435	35	171	36	128	40	54	788	165
	4	152	35	54	54	41	41	17	264	147
	5	314	41	118	42	90	36	40	562	159
10	1	2083	42	820	44	589	51	209	3701	346
11	1	5894	53	1086	41	608	59	205	7793	358
	2	179	40	128	43	53	53	4	364	140
	3	5025	116	1359	39	247	247	8	6639	410
	4	387	39	231	67	175	56	48	841	210
	5	186	32	75	37	35	35	2	298	106
12	1	25773	178	8373	80	2567	88	188	36901	534
13	1	12429	136	2355	77	445	445	29	15258	687
	2	16883	131	3644	76	1431	107	172	22130	486
	3	41076	242	5821	78	1106	76	114	48117	510
14	1	2513	57	2176	63	714	90	56	5459	266
	2	16444	110	4020	74	355	355	9	20828	548
	3	978	46	215	47	55	55	6	1254	154
15	1	5511	62	2329	45	1046	76	303	9189	486
16	1	4552	64	2517	52	1019	89	151	8239	356
17	1	5819	126	1806	85	188	188	3	7816	402
	2	633	50	895	47	466	65	31	2025	193
	3	233	47	173	42	91	37	36	533	162
	4	2976	159	189	71	43	43	3	3211	276
18	1	17033	150	1219	65	258	84	33	18543	332
	2	1182	43	1072	40	805	46	230	3289	359
19	1	12708	86	3039	34	1284	43	263	17294	426
	2	60759	268	6517	88	1290	125	188	68754	669
	3	22794	167	2307	65	465	95	53	25619	380
Total		278171	3130	56668	2046	17655	3567	2980	355474	11723

Tabelle 25 EVS 2002: Beispiel - Extrapolation in Schicht 18.1

strate	ast_pop	ast_brut	ast_net	Elek_net	Elek_hr	Elek_r	Elek_rr	cv_hr%	cv_r%	cv_rr%
18.1.1	17033	150	50	5.37	1829.22	1700.76	393.32	73.22	78.36	10.54
18.1.2	1219	65	39	9.59	299.74	273.88	144.91	30.16	31.65	15.76
18.1.3	258	84	48	31.75	170.67	161.69	161.69	20.67	21.61	21.61
18.1.5	33	33	22	51.94	77.91	93.44	93.44	15.30	11.53	11.53
18.1.total	18543	332	159	98.65	2377.54	2229.77	793.36	56.48	59.92	7.54

In 18.1. gibt es keine Spezialfälle, das heisst keine Verbände.

In Tabelle 25 fällt der grosse Unterschied bei den drei Schätzern in den Schichten 18.1.1 und 18.1.2 der kleinen Betriebe auf. Bei der EVS2002 waren nur die Verbände als Spezialfälle definiert. Mehrbetriebsunternehmen wurden noch nicht als Astgruppen erkannt. Es kann also der Fall vorkommen, dass bei einem Mehrbetriebsunternehmen ein kleiner Betrieb die allenfalls sehr grosse Verbrauchszahl für das ganze Unternehmen liefert. Somit ist dieser Betrieb zum Beispiel in Schicht 18.1.1 dann ein gewaltiger Ausreisser, der den hohen Wert und die grosse Ungenauigkeit beim Horvitz-Thompson-Schätzer verursacht. Erst beim robusten Quotientenschätzer wird die Wirkung dieses übergrossen Verbrauchswertes heruntergewichtet.

In den Folgejahren wird der Unterschied bei den drei Schätzern kleiner, da durch die Astgruppen-Definition einige dieser Ausreisser erkannt und grosse Verbrauchswerte auf die einzelnen Betriebe des Unternehmens aufgeteilt werden können.

Tabelle 26 EVS 2002: Hochrechnung des Verbrauchs an Elektrizität (Elek)

							davon Spezialfälle				
Branche	pop	brut	net	maxgew	Elek	cv_%	pop	net	Elek	cv_%	maxgew
1	2828	248	180	66.0	5440.2	4.9	19	14	383.8	1.7	1.1
2	1737	381	198	67.3	1370.7	7.0	1	.	.	.	.
3	4799	438	235	125.6	8556.5	13.1	47	35	2998.1	31.9	2.5
4	821	380	253	21.6	10774.5	11.0	217	157	5769.3	12.6	1.9
5	46	46	35	3.1	1494.2	2.9	19	16	568.6	6.3	1.2
6	1373	273	160	28.2	1289.2	9.6	9	5	46.9	34.9	2.1
7	194	194	101	2.1	3740.3	3.4	3	3	2321.4	0.0	1.0
8	75	75	41	4.4	3539.7	19.3	1	1	4.4	0.0	1.0
9	9506	892	504	105.4	6189.7	4.9	25	20	326.4	3.9	1.5
10	3701	346	229	69.7	3204.0	4.5	30	22	421.0	8.6	2.0
11	15935	1224	678	130.3	8496.4	3.9	28	20	1064.2	1.6	4.3
12	36901	534	228	340.6	1535.7	11.9	3	1	0.1	0.0	1.0
13	85505	1683	865	286.6	13393.0	4.7	77	41	188.0	18.8	1.9
14	27541	968	376	404.5	8256.1	4.7	2	.	.	.	.
15	9189	486	348	142.6	4512.0	5.5	2	.	.	.	.
16	8239	356	219	110.1	2190.9	6.6	.	.	.	.	.
17	13585	1033	532	77.1	3818.0	6.6	.	.	.	.	.
18	21832	691	442	331.4	5034.4	4.2	.	.	.	.	.
19	111667	1475	786	585.2	17146.1	5.5	19	10	129.4	24.0	3.9
Total	355474	11723	6410		109981.6		502	345	14221.6		

7.2 EVS 2003

Neben den Verbänden ($assoc$) werden auch Mehrbetriebsunternehmen ($astgr$) und Grossverbraucher ($AstGrBund$) vollerhoben. Diese drei Bereiche bilden die Spezialfälle. Zu den Grossverbrauchern zählen VBS, Bund, Swisscom, Post und SBB. 41 Betriebe erfüllen die Kriterien "Verband" und "AST-Gruppen" gleichzeitig.

Tabelle 27 EVS 2003: Anzahl Betriebe im ursprünglichen Rahmen und in der Bruttostichprobe

(Unter-) Branche		clt 1 pop	clt 1 brut	clt 2 pop	clt 2 brut	clt 3 pop	clt 3 brut	clt 5 pop	Total pop	Total brut
1	1	1820	49	481	49	286	81	117	2704	296
2	1	405	43	122	48	115	115	24	666	230
	2	893	40	97	38	56	56	13	1059	147
3	1	9	9	4	4	6	6	10	29	29
	2	65	65	52	52	52	52	37	206	206
	3	3104	64	775	50	319	92	58	4256	264
4	1	446	58	193	82	211	144	105	955	389
5	1	5	5	19	19	17	17	6	47	47
6	1	1001	58	265	46	148	148	22	1436	274
7	1	88	88	48	48	48	48	33	217	217
8	1	26	26	16	16	16	16	12	70	70
9	1	5442	70	1781	89	777	191	94	8094	444
	2	68	68	27	27	28	28	4	127	127
	3	412	41	187	54	108	47	56	763	198
	4	170	36	65	65	36	36	21	292	158
	5	377	43	112	42	94	41	45	628	171
10	1	1991	40	747	51	593	119	219	3550	429
11	1	5935	41	1087	49	657	99	230	7909	419
	2	190	41	134	52	60	60	2	386	155
	3	4862	148	1363	86	259	105	8	6492	347
	4	371	37	237	61	190	97	51	849	246
	5	210	32	89	32	38	38	5	342	107
12	1	25863	152	8163	146	2573	241	224	36823	763
13	1	12318	256	2489	145	465	183	37	15309	621
	2	14912	152	3375	116	1437	184	180	19904	632
	3	40103	628	5929	597	1245	395	101	47378	1721
14	1	2404	45	2056	93	726	173	63	5249	374
	2	17011	289	4214	217	400	74	9	21634	589
	3	1082	46	246	40	56	56	8	1392	150
15	1	6250	498	2335	249	1085	241	333	10003	1321
16	1	4614	398	2505	376	1088	442	154	8361	1370
17	1	5834	300	1809	220	175	84	2	7820	606
	2	621	81	901	93	464	150	37	2023	361
	3	248	104	183	96	121	65	35	587	300
	4	2980	256	213	73	56	56	3	3252	388
18	1	16840	329	1367	87	247	76	24	18478	516
	2	1200	66	1039	46	855	45	242	3336	399
19	1	12848	3155	3227	964	1445	575	280	17800	4974
	2	64089	343	7395	159	1699	199	267	73450	968
	3	22047	317	2288	109	545	129	61	24941	616
Total		279154	8517	57635	4886	18796	5004	3232	358817	21639*

* Im revidierten Rahmen sind es 167 Betriebe weniger.

Im Oktober 2005 wurde der Stichprobenrahmen der EVS2003 gemäss der zukünftigen Struktur revidiert. Die Bundesgrossbetriebe ($AstGrBund$ wurden in die AST-Gruppen ($astgr$) integriert mit Ausnahme der Bundesbetriebe. Diese wurden zu bbl-Betrieben. Im Rahmen gibt es 217 solche bbl-Betriebe, die ursprünglich vollerhoben wurden. Bei der Revidierung wurde aus diesen 217 bbl-Betrieben eine Stichprobe der festen Grösse 50 gezogen. Deswegen ist die ursprüng-

liche Bruttostichprobe um 167 Betriebe grösser als die revidierte, mit welcher im Folgenden (Tabellen 28 und 29) weitergerechnet wird.

Tabelle 28 EVS 2003: Beispiel - Extrapolation in Schicht 18.1 der Nicht-Spezialfälle

strate	ast_pop	ast_brut	ast_net	Elek_net	Elek_hr	Elek_r	Elek_rr	cv_hr%	cv_r%	cv_rr%
18.1.1	16786	275	106	5.81	919.88	809.72	435.10	40.77	46.41	8.75
18.1.2	1356	76	33	4.06	166.82	180.35	120.02	27.68	25.03	11.76
18.1.3	241	70	34	26.63	188.76	154.46	154.46	21.72	26.11	26.11
18.1.5	24	24	19	52.81	66.71	63.76	63.76	13.48	10.41	10.41
Zwischentotal	18407	445	192	89.31	1342.16	1208.28	773.34	28.33	31.51	7.45
18.1.total	18478	516	201	91.48	1352.73	1215.57	780.63	28.10	31.32	7.38

Vom Zwischentotal zum 18.1.total kommt noch der Energieverbrauch der Spezialfälle hinzu.

Tabelle 29 EVS 2003: Hochrechnung des Verbrauchs an Elektrizität (Elek)

							davon Spezialfälle				
Branche	pop	brut	net	maxgew	Elek	cv_%	pop	net	Elek	cv_%	maxgew
1	2704	296	213	58.1	5358.2	4.0	32	19	503.3	5.2	5.0
2	1725	377	196	49.5	1213.6	5.3	12	7	63.9	2.1	1.7
3	4491	499	276	115.4	8505.1	10.9	75	50	5992.7	15.2	5.6
4	955	389	275	27.4	10798.5	9.0	241	181	9116	10.4	1.4
5	47	47	32	8	1592.9	2.5	26	22	1258.2	0.6	1.9
6	1436	274	148	37.3	1292.3	5.9	23	14	432.1	5.5	1.8
7	217	217	117	2.1	4207.1	12.1	11	10	3269.5	15.4	1.2
8	70	70	30	3.6	3542.6	0.4	10	8	3353.5	0.0	1.0
9	9904	1098	636	122.6	6587.6	4.8	67	38	832.3	4.9	2.8
10	3550	429	295	94.8	3085.6	3.6	70	47	969.5	6.1	4.8
11	15978	1274	721	255.8	8913.7	6.3	95	48	1642.3	4.4	3.5
12	36823	763	345	492.7	1498.5	9.5	37	9	53.4	21.7	20.0
13	82591	2974	891	328	13956.2	7.0	1310	84	959.6	7.4	138.4
14	28275	1113	368	262.8	8668.9	4.2	183	5	229.3	2.2	75.7
15	10003	1321	1024	251.9	4359.2	4.9	871	737	669.5	0.7	1.4
16	8361	1203	563	103.6	2359.6	4.9	946*	341	904.3	6.1	57.0
17	13682	1655	755	61.8	3680.5	4.2	383	132	174.6	18.6	31.0
18	21814	915	509	143.3	5381.9	5.4	117	32	68.7	14.2	6.1
19	116191	6558	5392	742.6	17097.2	5.1	4870	4617	3303.4	5.3	23.2
	358817	21472	12786		112099.2		9379	6401	33796.1		

* hier wären wegen der bbl-Stichprobe 779 Betriebe in der Bruttostichprobe.

7.3 EVS 2004

Die Grossverbraucher ($AstGrBund$) werden abgeschafft. Post, Swisscom, SBB und VBS kommen zu den Mehrbetriebsunternehmen ($astgr$). Die Betriebe mit Noga-Code 751 und Betyp $L20$ werden als Bundesbetriebe (bbl) definiert. Innerhalb dieser Bundesbetriebe wird eine Stichprobe vorgegebener Grösse gezogen (vgl Kapitel 2.1 und 2.3).

Tabelle 30 EVS 2004: Anzahl Betriebe im Rahmen und in der Bruttostichprobe

(Unter-) Branche		clt 1		clt 2		clt 3		clt 5	Total	
		pop	brut	pop	brut	pop	brut	pop	pop	brut
1	1	1753	46	467	71	278	83	117	2615	317
2	1	400	42	119	49	107	107	25	651	223
	2	854	36	93	36	50	50	10	1007	132
3	1	9	9	4	4	6	6	9	28	28
	2	59	59	52	52	48	48	34	193	193
	3	3074	54	738	47	301	77	56	4169	234
4	1	442	36	189	45	214	52	110	955	243
5	1	4	4	17	17	15	15	6	42	42
6	1	987	56	262	47	146	146	19	1414	268
7	1	87	87	45	45	45	45	33	210	210
8	1	26	26	14	14	14	14	13	67	67
9	1	5447	70	1738	86	757	187	88	8030	431
	2	76	76	31	31	30	30	6	143	143
	3	413	41	185	49	110	49	55	763	194
	4	176	35	59	59	32	32	22	289	148
	5	391	43	114	45	89	43	47	641	178
10	1	2002	46	754	56	582	129	216	3554	447
11	1	5871	49	1065	50	629	89	228	7793	416
	2	189	43	134	44	60	60	3	386	150
	3	4893	162	1327	90	256	112	6	6482	370
	4	371	36	230	56	189	89	51	841	232
	5	205	33	87	33	40	40	4	336	110
12	1	26171	150	7946	149	2485	243	220	36822	762
13	1	12306	232	2446	115	458	143	36	15246	526
	2	15024	156	3319	124	1395	208	170	19908	658
	3	39090	874	5786	946	1196	604	92	46164	2516
14	1	2366	45	2020	87	722	167	64	5172	363
	2	16824	351	4148	305	382	74	9	21363	739
	3	1094	40	233	40	51	51	6	1384	137
15	1	6313	1125	2231	648	1037	403	310	9891	2486
16	1	4604	260	2506	303	1090	371	154	8354	1088
17	1	5858	304	1809	271	173	85	2	7842	662
	2	616	116	899	140	471	232	31	2017	519
	3	248	182	178	140	123	123	36	585	481
	4	2998	280	194	75	53	53	3	3248	411
18	1	17100	352	1363	89	239	76	27	18729	544
	2	1217	72	1052	68	848	69	247	3364	456
19	1	12653	3105	3082	971	1357	605	305	17397	4986
	2	65756	423	7239	175	1600	207	255	74850	1060
	3	22333	288	2289	105	532	119	56	25210	568
Total		280300	9444	56464	5777	18210	5336	3181	358155	23738

Tabelle 31 EVS 2004: Beispiel - Extrapolation in Schicht 18.1 der Nicht-Spezialfälle

strate	ast_pop	ast_brut	ast_net	Elek_net	Elek_hr	Elek_r	Elek_rr	cv_hr%	cv_r%	cv_rr%
18.1.1	17077	329	94	7.24	1315.00	1191.37	460.03	54.57	60.33	9.20
18.1.2	1350	76	29	5.13	238.65	248.53	121.38	39.04	36.61	14.56
18.1.3	232	69	38	33.25	202.97	173.28	173.28	17.77	20.36	20.36
18.1.5	27	27	18	45.71	68.56	74.20	74.20	17.34	12.13	12.13
Zwischentotal	18686	501	179	91.3182	1825.18	1687.38	828.895	39.70	42.99	7.06
18.1.total	18729	544	183	93.4032	1840.32	1695.43	836.945	39.37	42.79	7.00

Vom Zwischentotal zum 18.1.total kommt noch der Energieverbrauch der Spezialfälle hinzu.

Tabelle 32 EVS 2004: Hochrechnung des Verbrauchs an Elektrizität (Elek)

							davon Spezialfälle				
Branche	pop	brut	net	maxgew	Elek	cv_%	pop	net	Elek	cv_%	maxgew
1	2615	317	206	56.7	5543.4	4.6	70	29	1131.3	8.0	11.3
2	1658	355	186	80.1	1182.4	5.2	.	.	.	.	.
3	4390	455	241	139.5	8495.9	11.0	25	16	5670.5	16.0	1.1
4	955	243	164	32.8	10697.3	8.6	13	8	1231.9	3.1	1.1
5	42	42	36	6	1655.2	0.0	32	32	1603.3	0.0	1.0
6	1414	268	145	40.8	1248.7	10.3	11	8	210.2	40.9	2.5
7	210	210	104	2.4	4212.6	12.9	11	9	3283.4	16.4	1.1
8	67	67	31	8.1	3917.7	40.1	5	4	3715.2	42.3	1.1
9	9866	1094	588	160.6	6160.7	5.2	66	28	724.1	14.1	2.6
10	3554	447	289	87.8	3304.2	5.8	77	44	771.2	6.7	6.0
11	15838	1278	679	217	8951.2	3.5	28	16	819.3	0.5	2.2
12	36822	762	326	583.1	1541.5	10.8	24	3	19.6	10.4	2.4
13	81318	3700	847	340.2	14072.3	4.4	1990	81	3653.1	7.2	139.5
14	27919	1239	343	1189.5	8621.2	4.5	326	5	107.9	10.7	86.6
15	9891	2486	1897	189.9	4653.5	6.1	2121	1676	1800.8	0.8	1.4
16	8354	1088	589	98.4	2405.2	5.4	850*	362	817.5	4.3	57.0
17	13692	2073	786	76.4	3685.5	2.7	790	193	989.8	3.9	58.4
18	22093	1000	490	173.6	5645.4	5.5	132	31	212.2	3.8	34.5
19	117457	6614	5384	2015.3	17804.6	4.7	4917	4660	3809.4	3.3	31.8
	358155	23738	13331		113798.5		11488	7205	30570.7		

* hier wären wegen der bbl-Stichprobe 684 Betriebe in der Bruttostichprobe.

7.4 EVS 2005

Verfahren gleich wie im Vorjahr.

Tabelle 33 EVS 2005: Anzahl Betriebe im Rahmen und in der Bruttostichprobe

(Unter-) Branche		clt 1		clt 2		clt 3		clt 5	Total	
		pop	brut	pop	brut	pop	brut	pop	pop	brut
1	1	1703	44	465	73	268	88	113	2549	318
2	1	402	46	112	46	104	104	23	641	219
	2	815	40	93	37	46	46	9	963	132
3	1	10	10	4	4	5	5	9	28	28
	2	60	60	52	52	46	46	33	191	191
	3	2984	52	707	51	285	75	56	4032	234
4	1	447	35	189	47	212	55	112	960	249
5	1	3	3	15	15	14	14	6	38	38
6	1	960	59	255	52	142	142	19	1376	272
7	1	79	79	46	46	45	45	36	206	206
8	1	26	26	15	15	16	16	13	70	70
9	1	5438	70	1711	84	745	184	89	7983	427
	2	78	78	28	28	25	25	3	134	134
	3	424	49	187	50	105	51	50	766	200
	4	174	34	58	58	28	28	23	283	143
	5	390	44	114	48	87	87	46	637	225
10	1	2065	47	751	61	579	103	207	3602	418
11	1	5848	49	1054	47	619	90	223	7744	409
	2	173	44	131	45	57	57	4	365	150
	3	4925	172	1321	92	250	115	6	6502	385
	4	370	35	224	54	184	87	51	829	227
	5	189	38	76	76	38	38	3	306	155
12	1	26596	153	7799	146	2423	249	214	37032	762
13	1	12377	252	2410	131	456	162	35	15278	580
	2	15613	163	3233	115	1362	202	167	20375	647
	3	38091	893	5648	942	1164	595	91	44994	2521
14	1	1961	49	1872	87	689	174	62	4584	372
	2	16793	361	4142	321	380	78	10	21325	770
	3	1065	73	230	48	48	48	6	1349	175
15	1	6280	956	2143	618	1001	387	302	9726	2263
16	1	4307	190	2381	234	1023	311	146	7857	881
17	1	5668	377	1790	301	169	84	2	7629	764
	2	611	132	864	158	467	235	29	1971	554
	3	278	210	184	184	122	122	36	620	552
	4	2993	242	193	70	49	49	4	3239	365
18	1	17191	363	1349	93	228	78	28	18796	562
	2	1188	76	1038	70	846	69	242	3314	457
19	1	12744	3202	2937	923	1305	574	291	17277	4990
	2	67546	438	7129	184	1540	227	255	76470	1104
	3	22510	301	2243	103	521	115	54	25328	573
Total		281375	9545	55193	5809	17693	5260	3108	357369	23722

Tabelle 34 EVS 2005: Beispiel - Extrapolation in Schicht 18.1 der Nicht-Spezialfälle

strate	ast_pop	ast_brut	ast_net	Elek_net	Elek_hr	Elek_r	Elek_rr	cv_hr%	cv_r%	cv_rr%
18.1.1	17157	329	106	4.26	689.18	609.57	480.49	15.12	17.10	9.65
18.1.2	1332	76	35	5.98	227.50	230.37	121.53	39.52	38.96	12.71
18.1.3	221	71	43	42.48	218.32	179.65	179.65	16.27	19.20	19.20
18.1.5	28	28	19	63.12	93.02	85.55	85.55	13.49	10.65	10.65
Zwischentotal	18738	504	203	115.84	1228.02	1105.13	867.20	11.62	12.86	6.98
18.1.total	18796	562	207	121.96	1262.90	1140.18	902.25	11.31	12.52	6.87

Vom Zwischentotal zum 18.1.total kommt noch der Energieverbrauch der Spezialfälle hinzu.

Tabelle 35 EVS 2005: Hochrechnung des Verbrauchs an Elektrizität (Elek)

							davon Spezialfälle				
Branche	pop	brut	net	maxgew	Elek	cv_%	pop	net	Elek	cv_%	maxgew
1	2549	318	218	62.4	5364.3	4.4	75	35	1157.5	5.4	5.3
2	1604	351	185	65	1100.5	5.4	3	3	31.2	0.0	1.0
3	4251	453	264	133.2	7735.6	2.5	32	20	5176.7	0.0	2.7
4	960	249	172	26.9	11327.4	8.7	28	17	2399.5	27.6	2.8
5	38	38	28	1.2	1751.8	2.3	29	25	1697.7	2.4	1.2
6	1376	272	143	40.7	1267.4	9.7	25	7	249.4	38.3	4.1
7	206	206	109	2.5	4180.9	11.8	13	11	3172.4	15.4	1.1
8	70	70	37	4.9	3352.8	0.4	6	5	3130.3	0.0	1.0
9	9803	1129	649	148.8	6448.6	4.9	80	43	848.2	6.7	4.7
10	3602	418	280	74.9	3209.5	7.5	82	45	831.9	6.4	2.6
11	15746	1326	749	256.9	9851.8	7.9	45	14	887.6	2.2	3.4
12	37032	762	363	395.3	1456.2	9.8	24	2	1.5	.	3.1
13	80647	3748	896	313.5	14668	3.8	2005	82	3786.9	6.7	84.2
14	27258	1317	377	251.1	8543.6	4.7	406	9	197.5	40.3	107.0
15	9726	2263	1681	297.9	4722.4	7.2	1917	1467	1838	1.6	2.7
16	7857	881	417	88.2	2795.1	4.5	600*	188	1351.6	2.4	10.4
17	13459	2235	722	48.9	3745	3.3	965	73	1124.2	7.8	42.3
18	22110	1019	515	148.7	5730.9	5.6	161	20	390.9	6.4	38.0
19	119075	6667	5481	1371	19091.6	8.6	4998	4705	5738.1	25.3	55.0
Total	357369	23722	13286		116343.4		11494	6771	34011.1		

* hier wären wegen der bbl-Stichprobe 491 Betriebe in der Bruttostichprobe.

7.5 EVS 2006

Im Jahr 2005 fand eine Betriebszählung statt, die zum Zeitpunkt der Ziehung fast vollständig im BUR integriert war.

Tabelle 36 EVS 2006: Anzahl Betriebe im Rahmen und in der Bruttostichprobe

(Unter-)		clt 1		clt 2		clt 3		clt 5	Total	
Branche		pop	brut	pop	brut	pop	brut	pop	pop	brut
1	1	1496	72	528	122	260	83	124	2408	401
2	1	345	46	110	51	94	94	21	570	212
	2	712	42	79	79	38	38	9	838	168
3	1	10	10	3	3	8	8	7	28	28
	2	69	69	46	46	44	44	29	188	188
	3	2651	72	681	52	280	77	47	3659	248
4	1	402	35	194	51	218	58	101	915	245
5	1	6	6	20	20	10	10	7	43	43
6	1	842	62	256	54	157	157	18	1273	291
7	1	79	79	43	43	48	48	29	199	199
8	1	28	28	18	18	12	12	13	71	71
9	1	4813	64	1766	91	726	174	98	7403	427
	2	74	74	28	28	21	21	3	126	126
	3	404	49	178	43	120	55	43	745	190
	4	171	39	50	50	35	35	20	276	144
	5	375	43	135	55	91	91	40	641	229
10	1	1983	49	819	59	572	146	196	3570	450
11	1	5095	55	1121	51	635	124	222	7073	452
	2	184	44	129	45	64	64	1	378	154
	3	4453	155	1303	78	278	120	7	6041	360
	4	347	43	246	59	185	94	50	828	246
	5	166	40	79	79	41	41	3	289	163
12	1	24111	163	8111	144	2709	244	238	35169	789
13	1	11708	223	2624	149	509	189	34	14875	595
	2	14368	157	3449	128	1446	210	167	19430	662
	3	35490	911	5848	969	1199	606	88	42625	2574
14	1	2021	59	1744	102	636	180	60	4461	401
	2	16946	428	3692	319	341	79	9	20988	835
	3	1137	487	246	196	48	48	6	1437	737
15	1	6364	982	2321	632	1036	400	295	10016	2309
16	1	4147	467	2481	452	1118	412	184	7930	1515
17	1	5282	521	1614	225	149	74	5	7050	825
	2	632	196	845	227	508	314	33	2018	770
	3	268	214	197	197	148	148	34	647	593
	4	2463	144	205	58	53	53	6	2727	261
18	1	15267	380	1469	122	260	88	25	17021	615
	2	1160	82	1053	73	929	102	239	3381	496
19	1	11791	3171	3010	966	1367	603	297	16465	5037
	2	58649	455	7545	208	1793	247	277	68264	1187
	3	19849	316	2344	120	561	126	73	22827	635
Total		256358	10532	56630	6464	18747	5717	3158	334893	25871

Tabelle 37 EVS 2006: Beispiel - Extrapolation in Schicht 18.1 der Nicht-Spezialfälle

strate	ast_pop	ast_brut	ast_net	Elek_net	Elek_hr	Elek_r	Elek_rr	cv_hr%	cv_r%	cv_rr%
18.1.1	15190	303	106	4.77	684.13	648.42	455.50	25.19	26.33	7.90
18.1.2	1432	85	43	13.75	457.93	502.45	208.64	45.60	41.42	11.41
18.1.3	245	73	47	36.95	192.62	164.97	164.97	16.62	18.94	18.94
18.1.5	24	24	15	56.85	90.96	77.73	77.73	14.99	14.29	14.29
Zwischentotal	16891	485	211	112.32	1425.64	1393.57	906.85	19.15	19.46	6.00
18.1.total	17021	615	214	113.17	1457.14	1440.20	953.47	18.73	18.83	5.71

Vom Zwischentotal zum 18.1.total kommt noch der Energieverbrauch der Spezialfälle hinzu.

Tabelle 38 EVS 2006: Hochrechnung des Verbrauchs an Elektrizität (Elek)

							davon Spezialfälle				
Branche	pop	brut	net	maxgew	Elek	cv_%	pop	net	Elek	cv_%	maxgew
1	2408	401	204	67.4	5761.3	4.2	163	40	1514.5	7.2	27.7
2	1408	380	170	88.6	1051.7	5.9	10	5	129	5.3	2.0
3	3875	464	239	101.4	7598.2	8.6	29	19	5323.4	12.0	1.6
4	915	245	160	24.4	11388.2	13.4	45	27	4209.5	27.2	2.7
5	43	43	30	1.5	1825	2.4	34	27	1770.6	2.4	1.5
6	1273	291	138	33.2	1541.4	6.6	39	9	424	4.0	10.8
7	199	199	98	3	4930.7	14.7	15	11	3931.1	18.4	1.3
8	71	71	37	2.3	1135.1	4.7	6	4	681.3	7.1	1.2
9	9191	1116	593	148.2	7070.3	4.7	87	41	794.4	7.4	4.4
10	3570	450	281	89.3	3457.9	6.1	105	52	1044.2	10.1	5.2
11	14609	1375	721	182.3	10150.6	6.7	102	36	1519.8	1.6	11.7
12	35169	789	328	415.9	1452.6	10.0	48	11	16	2.8	19.6
13	76930	3831	913	320.9	15113.3	3.6	2168	121	3995.4	7.2	101.2
14	26886	1973	377	237.3	8235.1	4.7	1048	9	148	22.5	7.3
15	10016	2309	1662	243.1	4402.6	3.0	1999	1460	2168.2	0.5	3.3
16	7930	1515	454	91.9	3101.7	3.9	1277*	264	1784.6	3.2	32.0
17	12442	2449	779	54.4	3809	3.7	1319	166	1361.7	8.4	18.9
18	20402	1111	531	140.2	5695.4	3.1	270	34	880	10.5	58.3
19	107556	6859	5432	757.4	17190.7	7.3	5156	4654	5333.7	17.7	15.6
Total	334893	25871	13147		114910.8		12643	6990	37029.4		

* hier wären wegen der bbl -Stichprobe 1202 Betriebe in der Bruttostichprobe.

7.6 EVS 2007

Das Verfahren, um die *AST-Gruppen* festzustellen, wird einfacher. Einerseits arbeitet man mit der AST-Gruppen-Liste des vergangenen Jahres, andererseits hat man durch Profiling-Angaben einen besseres Kriterium, was eine AST-Gruppe sein kann.

Tabelle 39 EVS 2007: Anzahl Betriebe im Rahmen und in der Bruttostichprobe

(Unter-) Branche		clt 1		clt 2		clt 3		clt 5	Total	
		pop	brut	pop	brut	pop	brut	pop	pop	brut
1	1	1515	44	525	58	255	75	122	2417	299
2	1	355	50	108	51	92	92	21	576	214
	2	740	44	72	72	36	36	9	857	161
3	1	10	10	3	3	5	5	8	26	26
	2	61	61	48	48	45	45	28	182	182
	3	2666	71	680	49	262	74	49	3657	243
4	1	417	36	194	50	213	56	102	926	244
5	1	5	5	20	20	10	10	7	42	42
6	1	841	53	259	43	156	156	18	1274	270
7	1	84	84	45	45	50	50	28	207	207
8	1	30	30	18	18	12	12	12	72	72
9	1	4999	75	1762	95	713	181	100	7574	451
	2	73	73	25	25	22	22	3	123	123
	3	412	48	181	44	119	54	45	757	191
	4	168	38	51	51	35	35	19	273	143
	5	366	43	135	54	89	89	39	629	225
10	1	2020	43	815	57	566	143	192	3593	435
11	1	5281	52	1118	47	629	101	224	7252	424
	2	174	40	123	38	66	66	2	365	146
	3	4694	158	1295	77	272	115	7	6268	357
	4	358	42	235	56	182	95	50	825	243
	5	178	40	79	79	42	42	2	301	163
12	1	25938	176	8093	135	2657	232	232	36920	775
13	1	11992	231	2607	128	505	192	35	15139	586
	2	15185	144	3450	113	1436	179	170	20241	606
	3	36235	928	5881	703	1205	388	84	43405	2103
14	1	1959	60	1724	94	639	176	59	4381	389
	2	17375	348	3754	223	350	54	8	21487	633
	3	1198	188	250	108	49	49	8	1505	353
15	1	6869	812	2328	476	1047	362	301	10545	1951
16	1	3989	169	2417	190	1083	256	195	7686	810
17	1	5216	271	1590	184	145	145	5	6956	605
	2	612	54	840	75	494	144	31	1977	304
	3	266	83	188	77	143	63	35	632	258
	4	2679	118	210	49	48	48	7	2944	222
18	1	15838	330	1519	104	254	85	22	17633	541
	2	1197	66	1082	49	941	58	245	3465	418
19	1	12225	3246	2941	908	1303	565	294	16763	5013
	2	63510	423	7551	153	1756	193	279	73096	1048
	3	20988	306	2358	109	563	131	76	23985	622
Total		268718	9093	56574	4958	18489	4874	3173	346956	22098

Tabelle 40 EVS 2007: Beispiel - Extrapolation in Schicht 18.1 der Nicht-Spezialfälle

strate	ast_pop	ast_brut	ast_net	Elek_net	Elek_hr	Elek_r	Elek_rr	cv_hr%	cv_r%	cv_rr%
18.1.1	15812	304	98	3.34	539.52	495.11	454.37	9.87	10.46	7.67
18.1.2	1499	84	42	12.85	458.77	490.91	180.22	51.78	47.91	10.92
18.1.3	251	82	40	34.71	217.83	171.72	171.72	18.06	22.13	22.13
18.1.5	22	22	11	50.43	100.85	82.75	82.75	21.27	19.57	19.57
Zwischentotal	17584	492	191	101.34	1316.97	1240.49	889.07	18.80	19.70	6.47
18.1.total	17633	541	193	101.87	1327.82	1248.21	896.78	18.64	19.57	6.41

Vom Zwischentotal zum 18.1.total kommt noch der Energieverbrauch der Spezialfälle hinzu.

Tabelle 41 EVS 2007: Hochrechnung des Verbrauchs an Elektrizität (Elek)

							davon Spezialfälle				
Branche	pop	brut	net	maxgew	Elek	cv_%	pop	net	Elek	cv_%	maxgew
1	2417	299	214	59.3	5815	3.8	48	32	1010.9	9.4	2.7
2	1433	375	185	79.4	999.7	5.1	5	3	110.3	0.0	3.5
3	3865	451	232	111.3	7813.8	10.8	19	15	5503.2	15.2	4.9
4	926	244	158	35.5	11168.6	10.1	20	9	621.9	0.0	3.4
5	42	42	33	4.9	1800.9	2.3	30	28	1737.6	2.4	1.1
6	1274	270	151	46.4	1504.9	5.8	10	9	413.7	1.7	1.2
7	207	207	114	2.3	5049.7	15.1	14	10	3954.3	19.2	1.8
8	72	72	37	2.5	1032	1.2	5	5	618	0.0	1.0
9	9356	1133	654	142.2	7082.2	3.5	79	44	944.8	3.9	4.8
10	3593	435	279	109	3340.8	4.9	88	47	819.9	6.5	6.5
11	15011	1333	717	185.4	10797	7.1	55	24	951.1	2.0	5.4
12	36920	775	332	510	1646.6	11.7	30	5	13.2	39.9	4.8
13	78785	3295	869	404.8	14477.7	4.5	1632	136	1585.2	16.1	53.1
14	27373	1375	364	281	8343.6	5.0	423	5	20.9	0.0	1.9
15	10545	1951	760	236.3	4131.1	3.2	1606	549	1653.7	0.8	3.7
16	7686	810	432	108.1	2954.1	9.4	541*	252	1015	7.0	16.9
17	12509	1389	606	67.7	3869.9	5.6	275	40	286.4	5.8	24.1
18	21098	959	490	150.5	5807.8	3.2	99	23	338.6	1.7	63.0
19	113844	6683	5313	761.9	17152.5	5.4	4931	4541	3891.6	7.9	26.0
Total	346956	22098	11940		114787.9		9910	5751	25490.3		

* hier wären wegen der bbl-Stichprobe 465 Betriebe in der Bruttostichprobe.

Anhang

A Erwartete und effektive Stichprobengrösse

Unter einem Bernoulli Stichprobenplan haben alle Einheiten $k \in U$ dieselbe Einschlusswahrscheinlichkeit $\pi_k = \pi = \mathbf{E}(n)/N$, wobei $\mathbf{E}(n)$ die erwartete Stichprobengrösse und N die Gesamtgrösse ist. Damit ergibt sich, dass im Allgemeinen

$$\widehat{N} = \sum_{k \in S} \frac{1}{\pi_k} = \frac{1}{\pi} \sum_{k \in S} 1 = \frac{1}{\pi} \cdot n = \frac{N}{\mathbf{E}(n)} \cdot n \neq N \ . \tag{7}$$

Um die Gleichheit zu erhalten, wird auf N kalibriert, das heisst es wird mit dem Faktor $N/\widehat{N}$ multipliziert:

$$\widehat{N}_{cal} = \frac{N}{\widehat{N}} \sum_{S} \frac{1}{\pi_k} \overset{(7)}{=} \frac{N}{\widehat{N}} \cdot \widehat{N} = N \ . \tag{8}$$

Gleichung (8) lässt sich auch etwas umformen, sodass man die Gewichte d_k darin erkennt:

$$\widehat{N}_{cal} = \sum_{S} \frac{N}{\widehat{N}} \cdot \frac{1}{\pi_k} \overset{(7)}{=} \sum_{S} N \cdot \frac{\pi}{n} \cdot \frac{1}{\pi_k} \overset{\pi_k = \pi}{=} \sum_{S} \frac{N}{n} = \sum_{S} d_k \tag{9}$$

Analoges lässt sich für jede Schicht h herleiten, indem in obigen Gleichungen überall der Index h mitgeführt wird.

Schlussfolgerung: Die mit der optimalen Allokation berechneten festen Stichprobengrössen $\mathbf{E}(n_h)$ in den einzelnen Schichten h werden als Bernoulli-Stichprobe mit festen Zufallszahlen (*permanent random numbers*) gezogen. Dadurch erhält man die Stichprobengrössen $n_h \neq \mathbf{E}(n_h)$. Fasst man diese Bernoulli-Stichprobe als geschichtete Zufallsstichprobe mit konstanten Einschlusswahrscheinlichkeiten auf, so kann man $\frac{N_h}{n_h}$ als Anfangsgewichte d_k anschauen. Eigentlich sind sie die kalibrierten ursprünglichen Bernoulli-Anfangsgewichte $N_h/\mathbf{E}(n_h)$.

B Imputation in der Phase der Hochrechnung

Ein Betrieb hat einen fehlenden Wert, wenn der Betrieb in der Bruttostichprobe ist, aber keine Verbrauchsangaben liefert oder ihm in der Phase der Datenerhebung (vgl. Kapitel 5, Imputation) keine zugeordnet werden konnten.

B.1 Robuste Regression durch den Nullpunkt

Konkret stellen sich zwei grundsätzliche Fragen:

1. Mit welcher Methode soll imputiert werden? Nearest neighbor, Modellierung, usw
2. Wo soll imputiert werden? Nur in Grössenklasse 5 oder in allen vollerhobenen Schichten?

Mit den Daten der Wertschöpfung 2002, 2003 und 2004 wurden von Daniel Assoulin verschiedene Imputationsmethoden getestet (siehe Assoulin, D. (2007)). Es zeigte sich, dass ein Quotientenschätzer die besten Resultate lieferte, dh ein Regressionsmodell durch den Nullpunkt.

Für die Imputation bei der EVS wurden auch Regressionsmodelle mit den zwei Parametern y-Achsenabschnitt und Steigung getestet. Der y-Achsenabschnitt war meist nicht signifikant und der Genauigkeitsgewinn auch nicht besonders gross. So ist man beim Quotientenschätzer geblieben, hauptsächlich darum, weil dann die Varianzberechnung einfacher ist und es auch Literatur dazu gibt. Zudem sind z.T. nur wenig Daten vorhanden, sodass es sinnvoller ist, damit nur einen Parameter zu schätzen.

Konkret wird ein fehlender Energieverbrauchswert $y(t, k, C)$ in Branche C des Betriebes k im Jahr t imputiert, wenn sein Energieverbrauchswert im Jahr $t-1$ existiert (kann auch 0 sein, für 0 Verbrauch). Dazu wird mit allen Betrieben, die im Jahr $t-1$ und t in Branche C sind und die nicht schon als Spender oder Empfänger bei der grundlegenden Imputation bei der Datenerhebung dienen ($agg \in \{0, .\}$), ein robustes Regressionsmodell (vgl. SAS-Code im Anhang) durch den Nullpunkt gerechnet: Auf der x-Achse liegen die Energieverbrauchswerte des Jahres $t-1$, auf der y-Achse diejenigen des Jahres t. Damit erhält man die Steigung $\beta_{t-1,t}$ für die Branche C und somit kann $\hat{y}(t, k, C) = \beta_{t-1,t} \cdot y(t-1, k, C)$ imputiert werden. Falls keine robuste Regression gerechnet werden kann, weil nur zwei Werte vorhanden sind, wird gewöhnliche lineare Regression durch den Nullpunkt gerechnet. Falls nur ein Wert vorhanden ist, sollte nicht imputiert werden. Dieser Fall kommt bisher nicht vor.

Für die Beantwortung der Frage, wo imputiert werden soll, wurden drei Situationen in Betracht gezogen:

- Variante 1a: alle Betriebe in Schichten, die vollerhoben werden plus die BBL-Spezialfälle ($clt = 5$ oder AST-Gruppen oder Verbände oder Bundesbetriebe *BBL*)
- Variante 1b: nur die Betriebe der grössten Grössenklasse ($clt = 5$)
- Variante 2: die Betriebe der zwei grössten Grössenklassen ($clt = 5$ oder $clt = 3$)

Schlussendlich wird gemäss Variante 1a verfahren. Dabei wird ein Betrieb als zu 1a zugehörig angeschaut, falls er im betrachteten Zeitraum wenigstens in einem Jahr in einer dieser Schichten ist. Damit lässt sich auch die Entwicklung eines einstmal grossen Betriebes, der in Folgejahren stark abbaute, verfolgen.

Es kann auch sein, dass ein fehlender Wert nicht imputiert werden soll, weil der Betrieb Konkurs ging. Dies festzustellen ist nicht einfach. Es gibt zwar eine Variable $edit$, doch ist sie sehr unvollständig. Die für METH relevante Stufe 2 der kategoriellen Variablen $edit$ bedeutet "*Firma erloschen, aufgehoben, Konkurs*".

B.2 Umsetzung am Beispiel EVS2007

Die Imputation wird jeweils auf zwei Jahre angewendet: Einerseits werden fehlende Werte des aktuellen Jahres mittels Vorjahreswerten imputiert und andererseits fehlende Werte des Vorjahres mittels Verbrauchswerten des aktuellen Jahres. Beides macht Sinn, da jeweils mit dem aktuellen Jahr die Vorjahreswerte korrigiert oder Vorjahreswerte zur Plausibilisierung der aktuellen Werte herbeigezogen werden. Auch bezieht sich die longitudinale Zerlegung jeweils auf diese beiden Jahre.

Gemäss Tabellen 22 und 23 wären bei den Spezialfällen und den Betrieben der Grössenklasse 5 (Variante 1a) 4904 fehlende Werte zu imputieren. Da es doch einige Branchenwechsel (Betrieb im Vorjahr und im aktuellen Jahr nicht in derselben Branche) gibt oder einige Betriebe

im Vorjahr keinen Verbrauchswert angaben, werden schlussendlich noch 1266 fehlende Werte mit den Verbrauchsangaben des Vorjahres imputiert.

Tabelle 42 EVS2007: Extrapolation ohne und mit imputierten Werten

			ohne Imputationen			mit Imputationen			
Branche	ast_pop	ast_brut	ast_net	Elek_rr	cv*_Elek_rr	ast_net	Elek_rr	cv_Elek_rr	Anz.Imp.
1	2417	299	214	5815	3.80	221	5777	3.49	7
2	1433	375	185	1000	5.10	189	1009	5.03	4
3	3865	451	232	7814	10.80	241	7508	1.38	9
4	926	244	158	11169	10.14	166	10862	9.04	8
5	42	42	33	1801	2.32	33	1801	2.32	0
6	1274	270	151	1505	5.75	152	1506	5.30	1
7	207	207	114	5050	15.09	117	5125	14.94	3
8	72	72	37	1032	1.16	37	1032	1.16	0
9	9356	1133	654	7082	3.54	664	7141	3.46	10
10	3593	435	279	3341	4.93	301	3629	8.60	22
11	15011	1333	717	10797	7.10	753	10747	7.09	36
12	36920	775	332	1647	11.67	361	1650	11.63	29
13	78785	3295	869	14478	4.50	958	15043	4.07	89
14	27373	1375	364	8344	4.96	372	8396	4.92	8
15	10545	1951	760	4131	3.16	1494	4188	3.75	734
16	7686	810	432	2954	9.42	462	2977	8.33	30
17	12509	1389	606	3870	5.64	723	4050	5.27	117
18	21098	959	490	5808	3.21	534	5825	3.20	44
19	113844	6683	5313	17152	5.37	5428	17093	4.52	115
total	346956	22098	11940	114788	mean 6.20	13206	115361	mean 5.60	1266

**mit cv ist hier der Variationskoeffizient (Standardabweichung/Schätzwert) in Prozent gemeint.*

Die Extrapolation wird mit dem Schätzer gemäss Tabelle 14 durchgeführt. Die Endung _rr steht für den robusten Quotientenschätzer bzw. bedeutet, dass in dieser Branche je nach Schicht ein klassischer oder robuster Quotientenschätzer verwendet wird.

Die Extrapolation mit imputierten Werten (vgl. Tabelle 42) liefert einen um 0.5 Prozent höheren Verbrauchswert bei der Elektrizität als die Extrapolation ohne imputierte Werte (vgl. Tabelle 42). Die Genauigkeit der Energieverbrauchsschätzung pro Branche erhöht sich, wenn imputierte Werte verwendet werden. Weil imputierte Werte bei der jetzigen Berechnung jedoch wie feste Werte angeschaut werden, wird die Varianz unterschätzt.

Es ist schwierig zu beurteilen, ob in der Phase der Hochrechnung fehlende Werte durch das beschriebene Regressionsmodell imputiert oder durch entsprechende Gewichtungsfaktoren korrigiert werden sollen. In der Erhebungsphase wird schon imputiert. Fehlende Werte in der Hochrechnungsphase fliessen in die Gewichtung ein. Die Entstehung und Wirkung der Hochrechnungsgewichte auf die Schätzung kann explizit verfolgt und beurteilt werden. Besonders in den kleineren Schichten mit grossem Energieverbrauch, wo fehlende Werte die Hochrechnungsfaktoren stark beeinflussen, könnte das Antwortverhalten verbessert und somit die Hochrechnungsfaktoren reduziert werden. Im Moment wird diese Schiene verfolgt und - weil überzeugende Argumente gegen das Bisherige fehlen - die Hochrechnung so wie in Kapitel 3 beschrieben, durchgeführt.

C SAS-Programme und Daten

C.1 Übersicht

Tabelle 43 listet alle SAS-Programme und Daten auf, welche bei der Ziehung und Hochrechnung verwendet wurden. Die Auflistung dient rein internen Zwecken, die Programme und Daten können aus Datenschutzgründen nicht öffentlich gemacht werden.

Tabelle 43 Verwendete Programme und Daten bei der EVS2007

Programme/Daten	**Bedeutung**
1_ExtractInitialFrame.sas	Bearbeitung des BUR-Auszugs
2_InputStratification.sas	ordnet Branchen zu und führt Grössenklassen ein
3_CompareFrames.sas	vergleicht aktuellen mit Vorjahres-Rahmen
4_IdentifySpecialCases.sas	identifiziert die Spezialfälle Verbände, Unternehmen und BBL
5_Allocation.sas	bereitet csv-datasets für die Allokation in R vor
allocation07.R	Allokation, liefert Graphik und csv-file mit Schichtgrössen
6_MakeSamplingPlan.sas	erstellt Stichprobenplan
7_SelectSample.sas	zieht Stichprobe
8_SampleDelivery.sas	bereitet Stichprobe und Adressfiles für Lieferung an Polyquest auf
evs2007_extrapolation.sas	liefert Resultate für die Schichten und pro Branche; enthält die macros
- *weight*	Gewichte für HR- und Quotientenschätzer, standard und robust
- *fweight*	Endgewichte und Residuen
- *extrap*	Hochrechnungsresultate pro Schicht
- *extrap_gr*	Hochrechnungsresultate aggregiert auf die Branchen
evs2006evs2007_extrapolation _AnalyseLongitudinale_neu.sas	longitudinale Zerlegung von zwei Jahren
week_loc.sas7bdat, week_ent.sas7bdat	BUR-Auszug ca. Ende Herbst, wird nicht gespeichert
Verbandsliste, Unternehmensliste	von Polyquest, dient der Spezialfallidentifizierung
primarystratificationnoga2002.sas7bdat	Zuordnung NOGA2002 zu (Unter-)Branchen
evs2007_frame20071221.sas7bdat	Schlussrahmen, mit Spezialfällen, Zufallszahlen usw
evs2007_plan.csv*	Stichprobenplan
evs2007_brutadr_20071221.sas7bdat	Adressen in der Bruttostichprobe, für Versand von PQ
evs2007_brut_20071221.sas7bdat	'aufgeblasene' Bruttostichprobe (jeweils alle burnr einer entid)
net_bvm1_corr.sas7bdat	Nettostichprobe bei der Elektrizität (von Polyquest)
net_cvm1_corr.sas7bdat	Nettostichprobe bei Heizöl extra-leicht (von Polyquest)
net_dvm1_corr.sas7bdat	Nettostichprobe beim Erdgas (von Polyquest)
evs2007_extrap_bvm1_20080707.csv	geschätzter Elektrizitätsverbrauch pro Schicht
evs2007_extrap_gr_bvm1_20080707.csv	geschätzter Elektrizitätsverbrauch auf Branchen aggregiert
evs2007_extrap_cvm1_20080707.csv	geschätzter Heizölverbrauch pro Schicht
evs2007_extrap_gr_cvm1_20080707.csv	geschätzter Heizölverbrauch auf Branchen aggregiert
evs2007_extrap_dvm1_20080707.csv	geschätzter Erdgasverbrauch pro Schicht
evs2007_extrap_gr_dvm1_20080707.csv	geschätzter Erdgasverbrauch auf Branchen aggregiert
evs20062007_evolution_data20080707.csv	longitudinale Zerlegung

*Die *csv-Files existieren jeweils auch als sas-Files.*

C.2 Auszug SAS-Code zur Ziehung

Untenstehend ist der Code der Ziehung über den Verbrauch des Jahres 2007 aufgelistet.

```
* Zufallszahlen u für Betriebe, die neu im Rahmen des Jahres 2007 sind
data frame; set frame;
```

```
    if frame06 eq 0 then u=ranuni(2007);
run;

* in der Stichprobe 2007, falls ZZ <= taux (=pop/ebrut pro Schicht)
data frame; set frame;
    if (bbl07 eq 0) and (u le taux07) then brut07=1;
    if (bbl07 eq 0) and (u gt taux07) then brut07=0;
run;
```

Bei der EVS2007 hat man 126 Bundesbetriebe ($bbl07 = 1$) im Rahmen. Es wird eine BBL-Stichprobe von 50 Betrieben gezogen. Wegen der Art der Ziehung (feste Stichprobengrösse) ist die erwartete Stichprobengrösse hier gleich der gezogenen.

```
********** Wahl der festen BBL-Stichprobengrössen ********************;
data plan; set plan;
  if bbl
  07 eq 1 and clt07 eq 1 then ast_ebrut07=  5;
  if bbl07 eq 1 and clt07 eq 2 then ast_ebrut07=  5;
  if bbl07 eq 1 and clt07 eq 3 then ast_ebrut07= 15;
  if bbl07 eq 1 and clt07 eq 5 then ast_ebrut07= ast_pop;   * vollerhoben;
run;
proc sort data=plan out=plan;
  by group07 subgroup07 assoc07 astgr07 bbl07 clt07;
run;

********** Ziehung der BBL-Stichprobe ********************************;
* BBL-Stichprobenrahmen für die EVS2007;
data frame_bbl
; set frame;
  keep burnr bbl07 group07 subgroup07 clt07 ast_ebrut07 u;
  if bbl07 eq 1 then output;
run;
* Ziehung der BBL-Stichprobe;
proc sort data=frame_bbl out=frame_bbl;
  by group07 subgroup07 clt07 u;
run;
data frame_bbl; set frame_bbl;
  by group07 subgroup07 clt07; * pro group x subgroup x clt nach Rängen von u ordnen;
  retain rank;
  if first.group07 or first.subgroup07 or first.clt07 then rank=1;
  else rank=rank+1;
  if rank <= ast_ebrut07 then brut07=1; else brut07=0;
run;
```

C.3 Auszug SAS-Code zur Imputation in der Hochrechnungsphase

SAS-Prozedur für die robuste Regression durch den Nullpunkt

```
* egt für Energieträger, t1, t2 für Jahre;
proc robustreg data=check_&egt method=mm(chif=yohai initest=s) seed=20061023;
  by t1 t2 group;
  model &egt._t2 = &egt._t1 / noint;
  ods output ParameterEstimates=est_beta;
run;
```

Literatur

[Assoulin, D. 2007] ASSOULIN, D.: Wertschöpfungsstatistik. Einsetzungsversuche für fehlende Antworten grosser Unternehmen / Bundesamt für Statistik. Neuchâtel, 2007. – BFS Methodenbericht. – URL http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen__quellen/methodenberichte.html?publicationID=3007

[Gülden, J. et al. 2007] GÜLDEN, J. ET AL.: Energieverbrauch in der Industrie und im Dienstleistungssektor / Bundesamt für Energie. Bern, 2007. – BFE Energiestatistiken/Teilstatistiken. – URL http://www.bfe.admin.ch/themen/00526/00541/00543/index.html?lang=de&dossier_id=00775

[Hulliger 1999] HULLIGER, B.: Simple and Robust Estimators for Sampling. In: *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 1999, S. 54–63

[Salamin, P-A. 2006] SALAMIN, P-A.: Longitudinal Decomposition of Sampling Frames and Estimations as a Tool for Improving Data Quality. In: *Proceedings of Q2006*, European Conference on Quality in Survey Statistics, 2006

[Särndal 1992] SÄRNDAL, Swensson B. Wretman J.: *Model Assisted Survey Sampling*. Springer-Verlag New York, 1992. – ISBN 0387406204

Methodenberichte des Dienstes Statistische Methoden des BFS
Rapports de méthodes du Service de méthodes statistiques de l'OFS
Methodology reports published by the SFSO's Statistical Methods Unit

Ferster, M. (2013). EVS I - Energieverbrauchsstatistik 2002 bis 2007: Stichprobenplan und Hochrechnung. Bestellnummer: 338-0069

Assoulin, D. (2013). Zusatzerhebung für die landwirtschaftliche Betriebszählung 2010: Stichprobenplan und Hochrechnung. Bestellnummer: 338-0068

Potterat, J. (2012). Use of conversion keys for NOGA2002-2008. Order number: 338-0067-05

Andrade, B., Salamin P.-A. (2012). Enquête sur la situation sociale et économique des étudiant-e-s des hautes écoles suisses 2009. Cadre de sondage, plan d'échantillonnage et méthodes d'estimation. Numéro de commande: 338-0066

Potterat, J., Panchard, C., Kilchmann, D. (2012). Umweltschutzausgaben der Unternehmen 2009 (UWSA2009). Stichprobenplan, Einsetzungen, Gewichtung und Schätzverfahren. Bestellnummer: 338-0065

Potterat, J. (2012). Benutzung der Umsteigeschlüssel NOGA 2002-2008. Bestellnummer: 338-0064

Potterat, J. (2011). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2009 (KNBB09). Stichprobenplan, Gewichtung und Schätzverfahren. Bestellnummer: 338-0063

Kilchmann, D., Potterat, J., Genoud, S. (2011). Gütertransporterhebung 2008. Stichprobenplan, Datenaufbereitung, Gewichtung und Schätzverfahren. Bestellnummer: 338-0062

Graf, E. (2010). Enquête suisse sur la santé 2007. Plan d'échantillonnage, pondérations et analyses pondérées des données. Numéro de commande: 338-0061

Eichenberger P., Hulliger B., Potterat J. (2010). Describing the Anticipated Accuracy of the Swiss Population Survey. Order number: 338-0060

Graf, E. (2010). Étude empirique de l'attrition du Panel Suisse de Ménages : vers une caractérisation du profil du non-répondant. Numéro de commande: 338-0059

Graf, E. (2009). Weightings of the Swiss Household Panel: SHP_I wave 9, SHP_II wave 4, SHP_I et SHP_II combined. Order number: 338-0058

Graf, E. (2009). Pondérations du Panel Suisse de Ménages: PSM_I vague 9, PSM_II vague 4, PSM_I et PSM_II combinés. Numéro de commande: 338-0057-05

Qualité, L., Tillé, Y. (2009). Estimation de la précision d'évolutions dans l'enquête sur la valeur ajoutée. Numéro de commande: 338-0056

Renaud, A., Panchard, C. et Potterat, J. (2008). Statistique de l'emploi. Révision 2007 : méthodes d'estimation. Numéro de commande: 338-0055

Graf, E. (2008). Pondérations du PSM. PSM_I vague 8, PSM_II vague 3, PSM_I et PSM_II combinés. Numéro de commande: 338-0054

Andrade, B., Graf, M. (2008). Enquête suisse sur la structure des salaires 2006. Aspects méthodologiques du modèle des salaires SSalarium". Numéro de commande: 338-0053

Renaud, A. (2008). Statistique de l'emploi. Révision 2007 : cadre de sondage et échantillonage. Numéro de commande: 338-0052

Graf, E. (2008). Pondérations du SILC pilote. SILC_I vague 2, SILC_II vague 1, SILC_I et SILC_II combinés. Numéro de commande: 338-0051

Kilchmann, D. (2008). Statistik der sozialmedizinischen Institutionen 1999-2004 und Krankenhausstatistik 1999-2002. Einsetzungen für fehlende Daten. Bestellnummer: 338-0050

Renaud, A. (2008). Technologies de l'information et de la communication. Estimations sur la base de la statistique de la valeur ajoutée. Numéro de commande: 338-0049

Assoulin, D. (2007). Wertschöpfungsstatistik. Einsetzungsversuche für fehlende Antworten grosser Unternehmen. Bestellnummer: 338-0048

Kilchmann, D. (2007). Beherbergungsstatistik Campingplätze. Stichprobenrahmen und Schätzverfahren 2005/06. Bestellnummer: 338-0047

Gabler, S., Häder, S. (2007). Haushalts- und Personenerhebungen. Machbarkeit von Random Digit Dialing in der Schweiz. Bestellnummer: 338-0046

Ferrez, J., Graf, M. (2007). Enquête suisse sur la structure des salaires. Programmes R pour l'intervalle de confiance de la médiane. Numéro de commande: 338-0045

Renaud, A. (2007). Harmonisation de la scolarité obligatoire en Suisse (HarmoS). Design général de l'enquête et échantillon des écoles. Numéro de commande: 338-0044

Potterat, J. (2007). Betriebszählung 2005. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer: 338-0043

Hulliger, B. (2006). Umweltschutzausgaben der Unternehmen 2003, Stichprobenplan, Datenaufbereitung und Schätzverfahren. Bestellnummer: 338-0042

Renfer, J.-P. (2006). Enquête sur les chiffres d'affaires du commerce de détail. Plan d'échantillonnage et méthodes d'estimation. Numéro de commande: 338-0041

Salamin, P.-A. (2006). Statistique de l'aide sociale dans le domaine de l'asile. Plan de sondage et extrapolations pour l'enquête pilote 2005. Numéro de commande: 338-0040

Renaud, A. (2006). Statistique suisse des bénéficiaires de l'aide sociale. Pondération des communes 2004. Numéro de commande: 338-0039

Graf, M. (2006). Swiss Earnings Structure Survey 2002-2004. Compositional data in a stratified two-stage sample: Analysis and precision assessment of wage components. Order number: 338-0038

Potterat, J. (2006). Pensionskassenstatistik 2004. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer: 338-0037

Potterat, J. (2006). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2004. Stichprobenplan, Gewichtung und Schätzverfahren. Bestellnummer: 338-0036

Kilchmann, D. (2006). Vierteljährliche Wohnbaustatistik. Stichprobenplan, statistische Datenaufarbeitung und Schätzverfahren 2005. Bestellnummer: 338-0035

Kilchmann, D. (2006). Erhebung über Forschung und Entwicklung in der schweizerischen Privatwirtschaft 2004. Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Bestellnummer: 338-0034

Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 2. Bestellnummer: 338-0033

Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 1. Bestellnummer: 338-0032

Graf, M., Matei, A. (2005). Enquête suisse sur la structure des salaires 2002. La précision du salaire brut standardisé médian. Numéro de commande: 338-0031

Graf, E., Renfer, J.-P. (2005). Enquête suisse sur la santé 2002. Plan d'échantillonnage, pondération et estimation de la précision. Numéro de commande: 338-0030

Potterat, J. (2005). Mietpreis-Strukturerhebung 2003. Gewichtung und Schätzverfahren. Bestellnummer: 338-0029

Potterat, J. (2005). Landwirtschaftliche Betriebszählung 2003. Schätzverfahren für die Zusatzerhebung. Bestellnummer: 338-0028

Renaud, A. (2004). Coverage estimation for the Swiss population census 2000. Estimation methodology and results. Order number: 338-0027

Kilchmann, D. (2004). Revision des Schweizerischen Lohnindex. Schätzmethoden der Lohnindices und deren Varianzschätzer. Bestellnummer: 338-0026

Graf, M. (2004). Enquête suisse sur la structure des salaires 2002. Plan d'échantillonnage et extrapolation pour le secteur privé. Numéro de commande: 338-0025

Renaud, A. (2004). Analyse de données d'enquêtes. Quelques méthodes et illustration avec des données de l'OFS. Numéro de commande 338-0024

Renaud, A., Potterat, J. (2004). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sous-couverture (P-sample) et qualité du cadre de sondage des bâtiments. Numéro de commande: 338-0023

Graf, M. (2004). Fusion de données. Etude de faisabilité. Numéro de commande: 338-0022

Potterat, J. (2003). Mietpreis-Strukturerhebung 2003. Entwicklung des Stichprobenplans und Ziehung der Stichprobe. Bestellnummer: 338-0021

Potterat, J. (2003). Landwirtschaftliche Betriebszählung 2003. Stichprobenplan der Zusatzerhebung. Bestellnummer: 338-0020.

Renaud, A. (2003). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sur-couverture (E-sample). Numéro de commande: 338-0019

Hulliger, B. (2003). Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Erhebung über F+E in der schweizerischen Privatwirtschaft 2000. Bestellnummer: 338-0018

Renfer, J.-P. (2003). Enquête 2000 sur la recherche et le développement dans l'économie privée en Suisse. Plan d'échantillonnage. Numéro de commande: 338-0017

Potterat, J. (2003). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe. Schätzverfahren. Bestellnummer: 338-0016

Graf, M., Matei, A. (2003). Stratégie de choix des modèles de désaisonnalisation. Application aux séries de l'emploi total. Numéro de commande: 338-0015

Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Methoden für die Datenbereinigung. Bestellnummer: 338-0014

Renaud, A. (2002). Programme international pour le suivi des acquis des é lèves (PISA). Plans d'échantillonnage pour PISA 2000 en Suisse. Numéro de commande: 338-0013

Renfer, J.-P. (2002). Enquête 2001 sur les coûts et l'utilité de la formation des apprentis du point de vue des établissements. Plan d' échantillonnage. Numéro de commande: 338-0012

Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Stichprobenplan und Schätzverfahren für die provisorischen Ergebnisse. Bestellnummer: 338-0011

Graf, M. (2002). Enquête suisse sur la structure des salaires 2000. Plan d'échantillonnage, pondération et méthode d'estimation pour le secteur privé. Numéro de commande: 338-0010

Renaud, A., Eichenberger P. (2002). Estimation de la couverture du recensement de la population de l'an 2000. Procédure d'enquête et plan d'échantillonnage de l'enquête de couverture. Numéro de commande: 338-0009

Kilchmann, D., Hulliger, B. (2002). Stichprobenplan für die Obstbaumzählung 2001. Bestellnummer: 338-0008

Graf, M. (2002). Passage du concept établissement au concept entreprise. Numéro de commande: 338-0007

Salamin, P.A. (2001). La technique de la double enquête pour la statistique du transport routier de marchandise. Numéro de commande: 338-0006

Peters, R., Renfer, J.-P. et Hulliger, B. (2001). Statistique de la valeur ajoutée 1997-1998. Procédure d'extrapolation des données. Numéro de commande: 338-0005

Potterat, J., Hulliger, B. (2001). Schätzung der Sägereiproduktion mit der Sägerei-Erhebung PAUL. Bestellnummer: 338-0004

Graf, M. (2001). Désaisonnalisation. Aspects méthodologiques et application à la statistique de l'emploi. Numéro de commande: 338-0003

Hüsler, J., Müller, S. (2001). Schlussbericht Betriebszählung 1995 (BZ 95), Mehrfach imputierte Umsatzzahlen. Bestellnummer: 338-0002

Renaud, A. (2001). Statistique suisse des bénéficiaires de l'aide sociale. Plan d'échantillonnage des communes. Numéro de commande: 338-0001

Hulliger, B., Eichenberger, P. (2000). Stichprobenregister für Haushalterhebungen: Umstellung auf Telefonnummern ohne Namen und Adressen, Abläufe für Erstellung und Stichprobenziehung. Bestellnummer: 338-0000

de Rossi, F.-X. (1998). Méthodes statistiques pour le compte routier suisse.

Hulliger, B., Kassab, M. (1998). Evaluation of Estimation Methods for the Survey on Environment Protection Expenditures of Swiss Communes.

Salamin, P.A. (1998). Etablissement d'une clef de passage pondérée entre l'ancienne (NGAE 85) et la nouvelle nomenclature (NOGA 95) générale des activités économiques.

Peters, R. (1998). Extrapolation des données de l'enquête de structure sur les loyers.

Bender, A., Hulliger, B. (1997). Enquête suisse sur la population active: rapport de pondération pour 1996.

Salamin, P.A. (1997). Evaluation de la Statistique de l'emploi.

Peters, R. (1997). Etablissement du plan d'échantillonnage pour l'enquête 1996 sur la recherche et le développement dans l'économie privée en Suisse.

Peters, R. (1997). Enquête 1996 sur la structure des salaires en Suisse: établissement du plan d'échantillonnage.

Peters, R. (1996). Pondération des données de l'enquête sur la famille en Suisse.

Comment, T., Hulliger, B., Ries, A. (1996). Gewichtungsverfahren für die Schweizerische Arbeitskräfteerhebung (1991-1995).

Hulliger, B. (1996). Haushalterhebung Familie 1994: Stichprobenplan, Stichprobenziehung und Reservestichproben.

Peters, R., Hulliger, B. (1996). Schätzverfahren für die Lohnstruktur-Erhebung 1994 / Procédure d'estimation pour l'enquête de 1994 sur la structure des salaires.

Peters, R. (1996). Schéma de pondération des indices PAUL.

Hulliger, B., Peters, R. (1996). Enquête sur le comportement de la population suisse en matière de transport en 1994: plan d'échantillonnage et pondération.

Hulliger, B. (1996). Gütertransportstatistik 1993: Schätzverfahren mit Kompensation der Antwortausfälle.

Salamin, P.A. (1995). Estimation des flux pour le module II des comptes globaux du marché de travail.

Peters, R. (1995). Enquête de structure sur les loyers: établissement d'un plan d'échantillonnage stratifié.

Hulliger, B. (1995). Konjunkturelle Mietpreiserhebung: Stichprobenplan und Schätzverfahren.

Schwendener, P. (1995). Verbrauchserhebung 1990 - Vertrauensintervalle.

Peters, R., Hulliger, B. (1994). La technique de pondération des données: application à l'enquête suisse sur la santé.

Hulliger, B., Peters, R. (1994). Enquête sur la structure des salaires en Suisse: stratégie d'échantillonnage pour le secteur privé.

Publikationsprogramm BFS

Das Bundesamt für Statistik (BFS) hat – als zentrale Statistikstelle des Bundes – die Aufgabe, statistische Informationen breiten Benutzerkreisen zur Verfügung zu stellen.

Die Verbreitung der statistischen Information geschieht gegliedert nach Fachbereichen (vgl. Umschlagseite 2) und mit verschiedenen Mitteln:

Diffusionsmittel	*Kontakt*
Individuelle Auskünfte	032 713 60 11 info@bfs.admin.ch
Das BFS im Internet	www.statistik.admin.ch
Medienmitteilungen zur raschen Information der Öffentlichkeit über die neusten Ergebnisse	www.news-stat.admin.ch
Publikationen zur vertieften Information	032 713 60 60 order@bfs.admin.ch
Online Datenrecherche (Datenbanken)	www.statdb.bfs.admin.ch

Nähere Angaben zu den verschiedenen Diffusionsmitteln im Internet unter der Adresse www.statistik.admin.ch→Dienstleistungen→Publikationen Statistik Schweiz

Methodenberichte des Dienstes Statistische Methoden

Die Methodenberichte beschreiben die mathematischen und statistischen Methoden, die den Resultaten und Analysen der öffentlichen Statistik zugrunde liegen. Sie enthalten ausserdem die Evaluation und Entwicklung von neuen Methoden im Hinblick auf eine zukünftige Anwendung. Diese Publikationen sollen einerseits die verwendeten Methoden dokumentieren, um Transparenz und Wissenschaftlichkeit sicherzustellen, und sie sollen andererseits die Zusammenarbeit mit den Hochschulen und der Wissenschaft fördern.

Zur Illustration der beschriebenen mathematischen Konzepte werden im Bericht numerische Resultate aufgeführt. Diese sind allerdings nicht als offizielle Resultate der betreffenden Erhebungen zu verstehen. Ebenfalls können die tatsächlich angewendeten Methoden leicht von den hier beschriebenen abweichen.

Die Methodenberichte sind auf der Internetseite des BFS in elektronischer Form verfügbar.

In diesem Methodenbericht sind die statistischen Methoden vom Stichprobenplan bis zur Hochrechnung beschrieben, welche im Bundesamt für Statistik für die jährliche Energieverbrauchsstatistik des Bundesamtes für Energie entwickelt wurden.

Der Stichprobenplan berücksichtigt die unterschiedlichen Grössen und Organisationsformen der Betriebe.

Die Hochrechnung im Bericht bezieht sich auf die Energieträger Elektrizität, Heizöl extra-leicht und Erdgas und beruht auf einem Quotientenschätzer mit der Hilfsvariable Vollzeitstellen.

Für den Vergleich des Verbrauchs von zwei aufeinanderfolgenden Jahren wird eine longitudinale Zerlegung beschrieben, welche vor allem für die Plausibilisierung der erhobenen Daten herangezogen wird.

Der Bericht endet mit einigen hochgerechneten Resultaten der Jahre 2002 bis 2007 in den 19 für die Energieverbrauchsstatistik relevanten Branchen, die auf der NOGA2002 basieren.

Bestellnummer
338-0069

Bestellungen
Tel.: 032 713 60 60
Fax: 032 713 60 61
E-Mail: order@bfs.admin.ch

Preis
gratis

ISBN 978-3-303-00497-5