

Ralph Wenzl, IHA-GfK Schweiz

Fundierte Stichproben der IHA-GfK auf Grundlage der Schweizer Volkszählung 2000 und einer jährlichen Fortschreibung in der GfK

Schweizer Tage der öffentlichen Statistik
8.-10. September 2008
Davos, Schweiz

Agenda

2

1. Grundgesamtheit
2. Datengrundlage Fortschreibung (VZ 2000, ESPOP)
3. Fortschreibung VZ 2000
 - 3.1 Methode
 - 3.2 Qualität
 - 3.3 Genauigkeit
 - 3.4 Notwendige Gewichtung
 - 3.5 Einfluss
4. Fazit / Ausblick

1. Grundgesamtheit

3

Warum brauchen wir Grundgesamtheiten für Stichprobenziehungen?

- Ausweis von Potentialen
- Schichtung von Stichproben nach aktuellen Strukturmerkmalen
- Stichprobendesign und –Steuerung
- Inhaltliche Klärung der Definition der Grundgesamtheit je Erhebung

2. Datengrundlage Fortschreibung (VZ 2000, ESPOP)

	ESPOP	VZ 2000
Grundgesamtheit	Ständige Wohnbevölkerung ESPOP Ansatz: Die Bevölkerung am 31. Dezember des Jahres entspricht der Bevölkerung am 1. Januar des Jahres zuzüglich der Zugänge (Geburten und Zuwanderung) und abzüglich der Abgänge (Todesfälle und Abwanderung) des Jahres.	Personen, Haushalte und Erwerbstätige der Wohnbevölkerung (sowohl wirtschaftlicher als auch zivilrechtlicher Wohnsitz), verknüpft mit georeferenzierten Wohnungen und Gebäuden, sowie Arbeits- und Schulorten
Erfasste Merkmale	Bevölkerungsstand und Bevölkerungsbewegung nach: <ul style="list-style-type: none"> - Geschlecht - Alter - Zivilstand - Staatsangehörigkeit - Wohngemeinde (zivilrechtlicher Wohnsitz) 	- Personen: <ul style="list-style-type: none"> - Geschlecht - Geburtsdatum - Zivilstand - Staatsangehörigkeit - Wohnort - Geburtsort, Stellung im Haushalt, - Anzahl Kinder, Religion, Sprache etc.

2. Datengrundlage Fortschreibung (VZ 2000, ESPOP)

5

- Rechtliche Grundlage und Datenschutz
 - Vertrag zwischen der IHA und dem BfS
- VZ-Daten als Personen-Individualdaten
 - Nutzung nur ausgewählter Datenfelder, ohne persönliche Identifizierung
 - Beschränkung auf offensichtliche und wenig sensible Datenfelder
- ESPOP-Daten
 - Jährliche Fortschreibung
 - Aggregierte Tabellen

2. Datengrundlage Fortschreibung (VZ 2000, ESPOP)

6

Vorteile der Nutzung von VZ-Individualdaten für die Stichprobenbildung

- Freiheit der Definition der Grundgesamtheit:
 - Mediengrundgesamtheit enthält eine sehr spezifische Bevölkerungsdefinition
 - Mediengrundgesamtheit nutzt die Merkmale zur Sprachbeherrschung sehr differenziert
- Freiheit der Bildung von Stichprobenstrukturen:
 - beliebige räumliche Gliederungen ausgehend von der Gemeindeebene
 - beliebige Altersklassifizierung
- Verfügbarkeit von Strukturmerkmalen:
 - begrenzte Auswahl, aber trotzdem vielseitig

2. Datengrundlage Fortschreibung (VZ 2000, ESPOP)

Vorteile der Nutzung von VZ-Individualdaten für die Stichprobenbildung

Beispiel :

Region	Summe	Alter und Geschlecht						Erwerbstätigkeit		
		Männlich, 15-34 Jahre	Männlich, 35-54 Jahre	Männlich, 55-74 Jahre	Weiblich, 15-34 Jahre	Weiblich, 35-54 Jahre	Weiblich, 55-74 Jahre	Vollzeit (90 bis 100%)	Teilzeit(bis 89%)	Keine Beschäftigung
Zürich Agglo (ohne Stadt)	511,852	83,985	106,562	64,187	83,572	107,068	66,478	278,782	96,420	136,650
Zürich Stadt	250,796	50,403	46,714	27,156	47,231	45,741	33,551	128,531	49,516	72,749
	762,648	134,388	153,276	91,343	130,803	152,809	100,029	407,313	145,936	209,399

3.1 Methode

8

Methode*

- Randvorgaben ESPOP => Fortschreibung nur nach Personenvorgaben
- Methode Gewichtungsfaktoren je Individualdatensatz
- Sonderaspekt wirtschaftliche und zivilrechtliche Bevölkerung

- Entwicklung und Konzeption 2003
Stefan Klug, IHA-GfK Schweiz, Statistik
Klaus Bergmann, GfK Deutschland, Methods and Product Development

* Nähere Details können nicht dargestellt werden

3.2 Qualität

9

3.2 Qualität der Fortschreibung

- Problematik der HH und Familien-Fortschreibung
- Nullzellen
- Nur Fortschreibung Struktureffekte, keine Verhaltenseffekte möglich:
z.B. Entwicklung zu kleinen HH, Veränderungen im Bildungsverhalten
- Faktoren nahe 1,0 – sehr hohe RIM-Efficiency, einzelne Ausreisser

3. Fortschreibung VZ 2000

10

- **3.3 Genauigkeit der Gewichtung**
 - Vergleich der Soll-Vorgaben mit den gewichteten Ist-Vorgaben: Mit welcher Genauigkeit werden die Ist-Vorgaben erreicht?
- **3.4 Notwendige Gewichtung**
 - Gesamt: Wie gross sind die Gewichtungsfaktoren?
 - Wie ist die Verteilung der Gewichtungsfaktoren nach Gewichtungsmerkmalen?
- **3.5 Einfluss der Gewichtung**
 - Bevölkerungskonzepte: Total
 - Soziodemographische Variablen

3.3 Genauigkeit der Gewichtung (2000-2003)

11

- **Vergleich der Soll-Vorgaben mit den gewichteten Ist-Vorgaben 2003**

Das SOLL-Total von 7'364'300 wird bis auf **58 Personen ($7.8 * 10^4 \%$)** genau erreicht.

Gewichtetes IST-Total = 7'364'242.

- Maximale absolute Abweichungen nach Gewichtung / Beispiele:

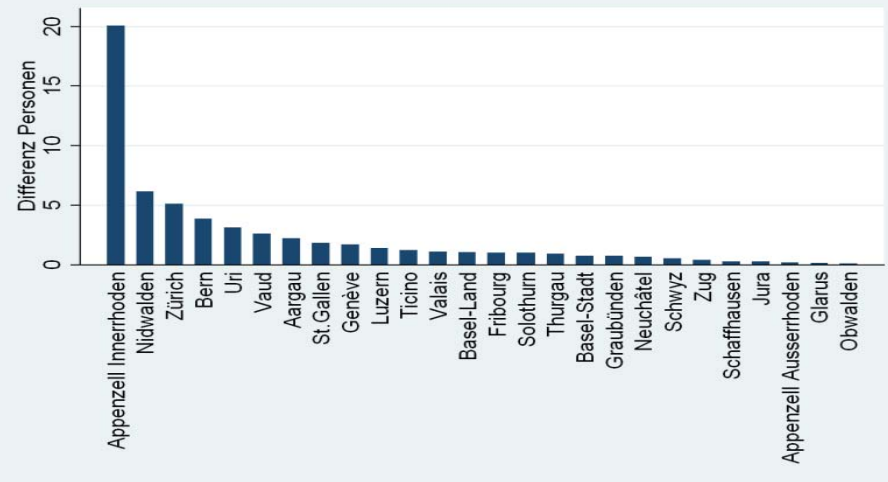
- Alter 8 Personen
- Kanton 20 Personen
- Gemeinde 13 Personen

3.3 Genauigkeit der Gewichtung (2000-2003)

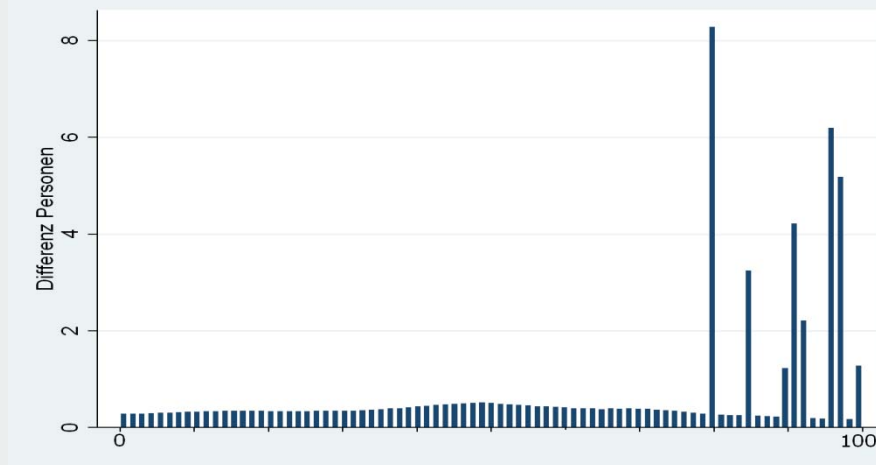
Säulendiagramm : Differenz Personen nach Gemeindenummer



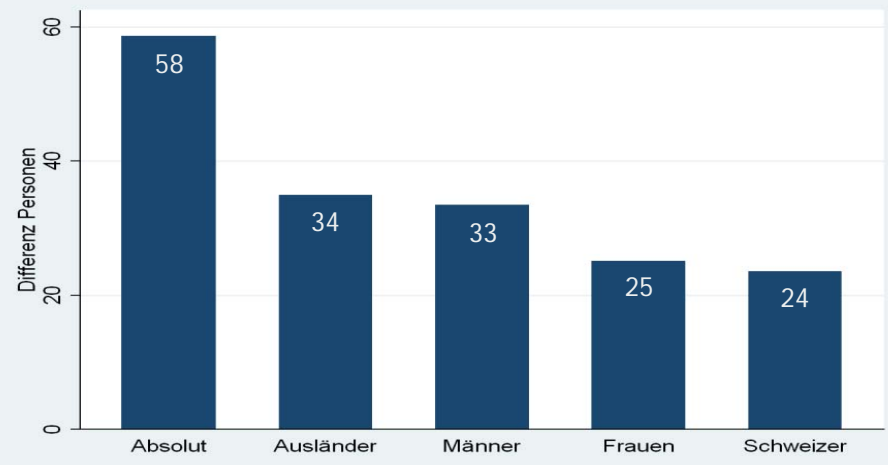
Säulendiagramm : Differenz Personen nach Kanton



Säulendiagramm : Differenz Personen nach Alter



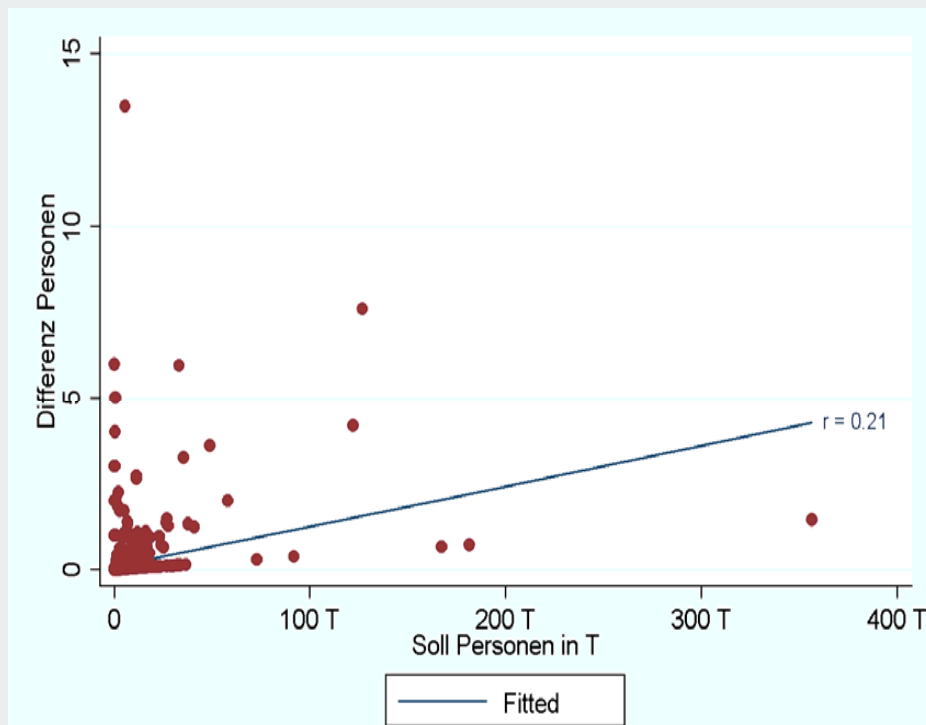
Säulendiagramm : Differenz Personen nach Geschlecht / Nationalität



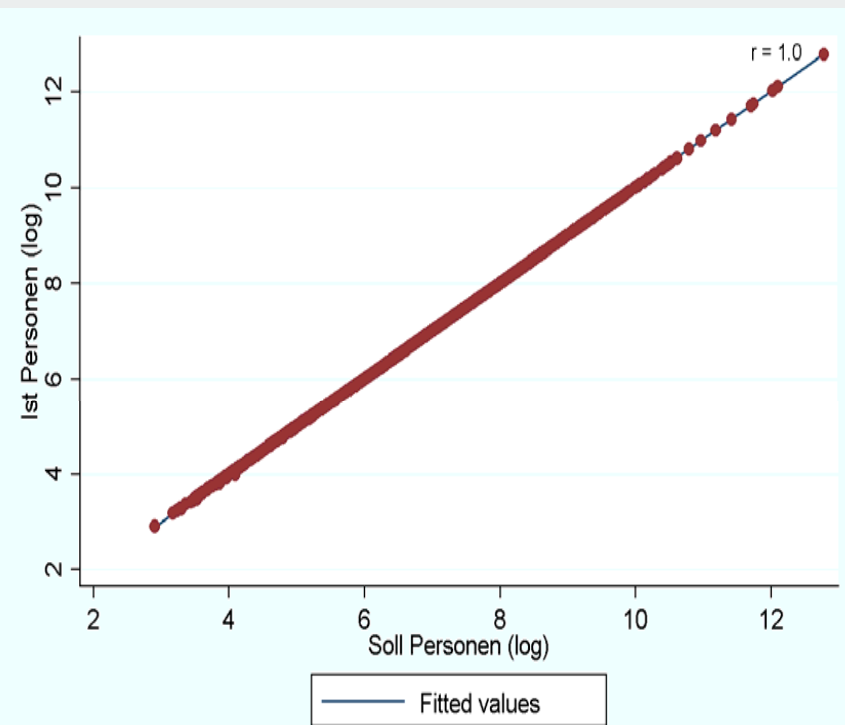
3.3 Genauigkeit der Gewichtung (2000-2003)

13

Scatterplot : Differenz Ist Personen / Soll Personen
Fortschreibung 2003, nach Gemeinden



Scatterplot : Ist Personen / Soll Personen
Fortschreibung 2003, nach Gemeinden



3.4 Notwendige Gewichtung

- **Wie gross sind die Gewichtungsfaktoren ?**
- Die Gewichtungen sind insgesamt höher für Ausländer als für Schweizer und für Männer höher als für Frauen.

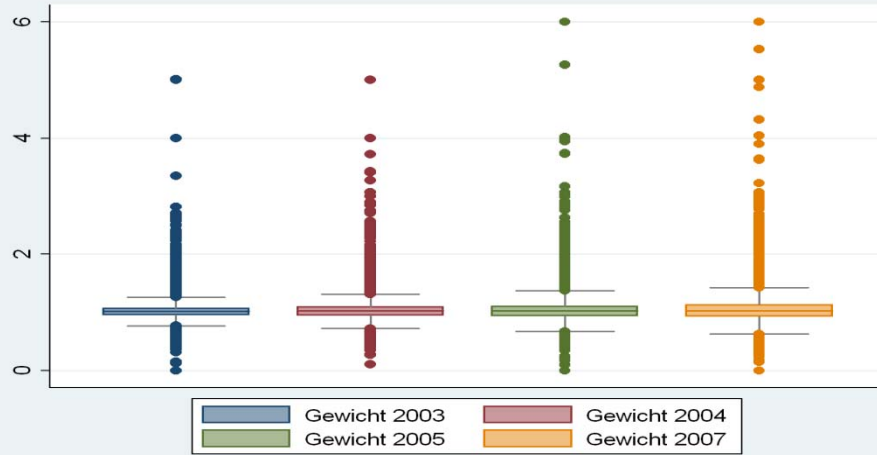
	Gewicht 2003	Gewicht 2004	Gewicht 2005	Gewicht 2007
Konstante	0.889**	0.855**	0.828**	0.811**
Alter (sqrt)	0.012**	0.017**	0.022**	0.026**
Geschlecht	-0.002**	-0.005**	-0.007**	-0.009**
Nationalität	0.056**	0.074**	0.085**	0.088**
Alpen und Voralpen	-0.011**	-0.015**	-0.021**	-0.027**
Westmittelland	-0.015**	-0.021**	-0.028**	-0.036**
Ostmittelland	-0.004**	-0.007**	-0.009**	-0.011**
Svizzera Italiana	-0.004**	-0.007**	-0.010**	-0.013**
Adj. R²	0.08	0.11	0.12	0.13

- **Gewichtungsfaktor**

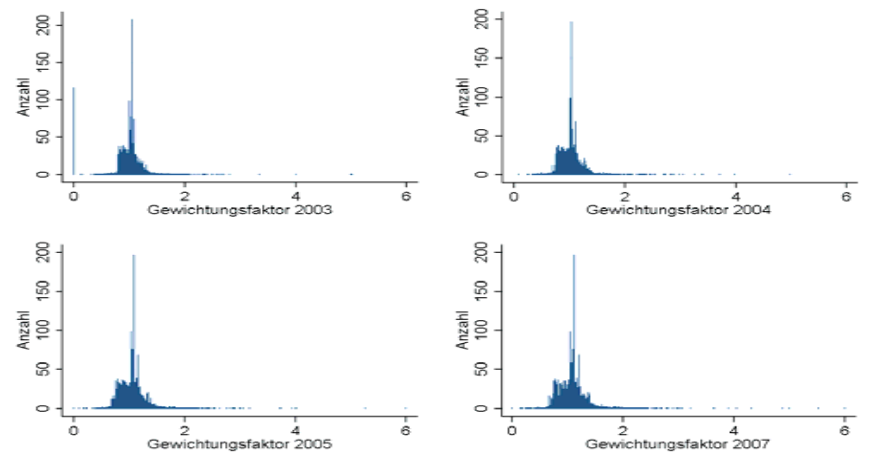
	Gewicht 2003	Gewicht 2004	Gewicht 2005	Gewicht 2007
Mittelwert	1.0224	1.0301	1.0362	1.0431
Standardabweichung	0.1195	0.1339	0.1560	0.1760
Cohens d (pooled)	-	0.0568	0.0973	0.1352

3.4 Notwendige Gewichtung

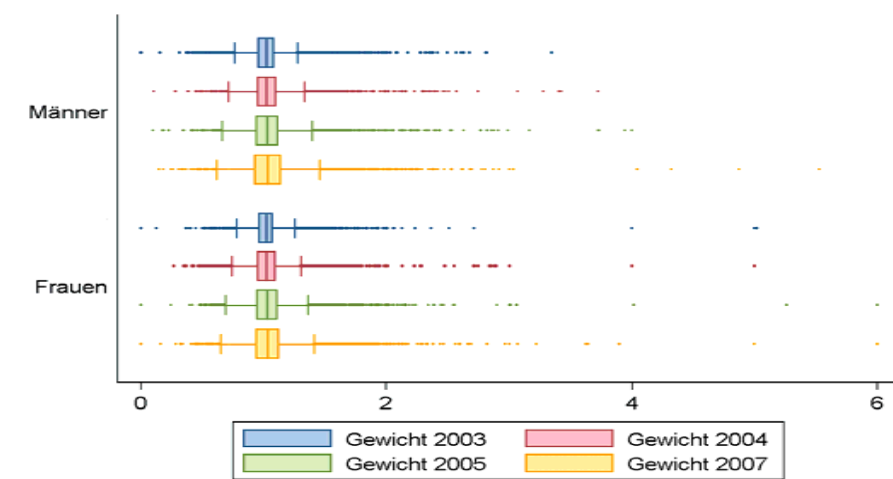
Boxplot : Jahre (n=100.000)



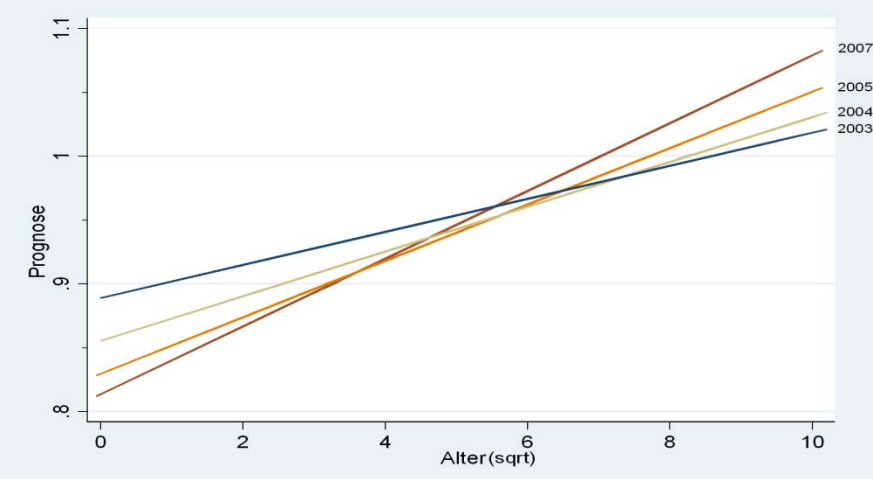
Verteilung : Jahre (n=100.000)



Boxplot : Jahre nach Geschlecht (n=100.000)



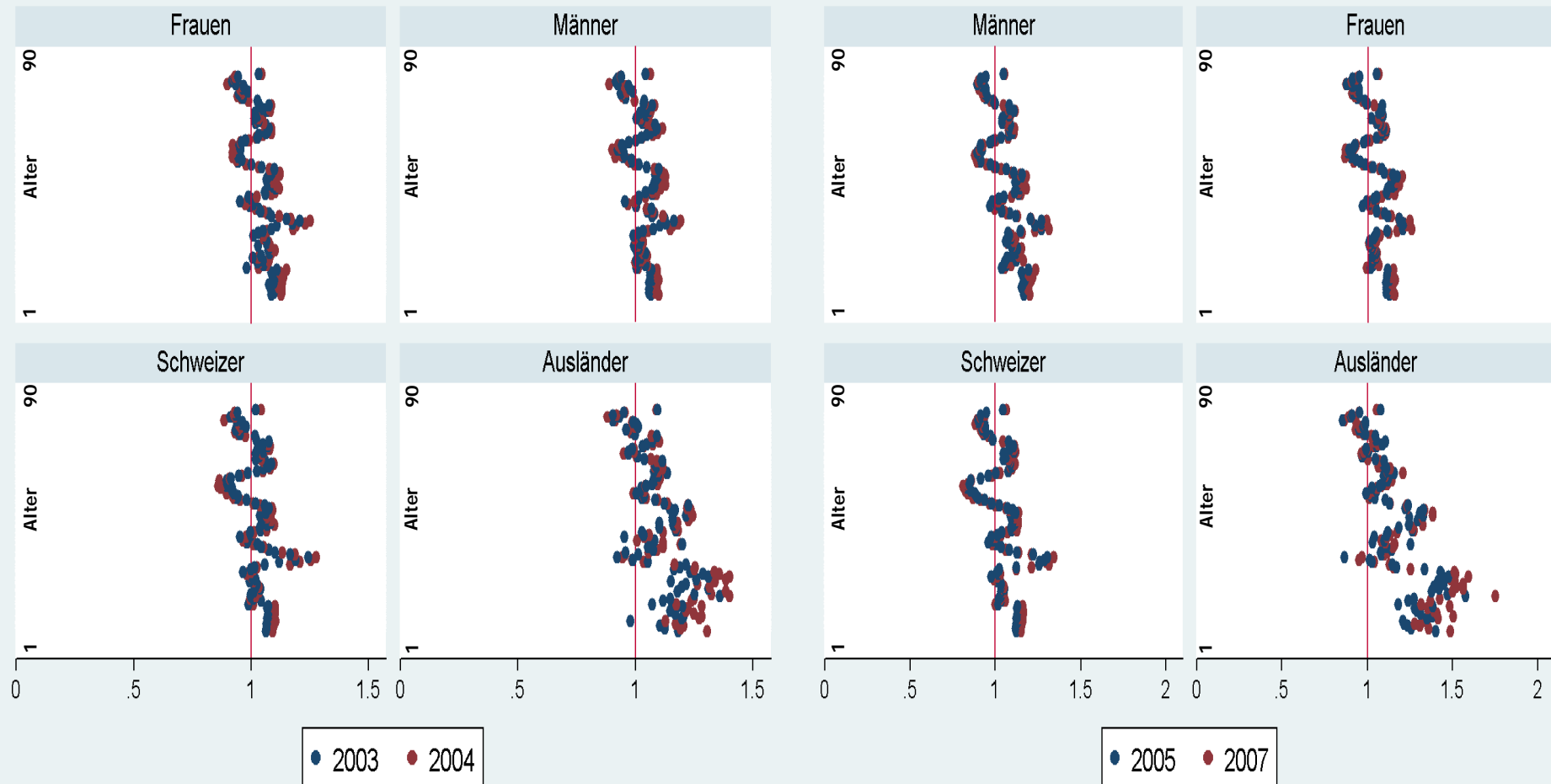
Conditional Effect Plot : Alter (Berechnung aufgrund Regression Slide 14)



3.4 Notwendige Gewichtung

16

Mittelwert der Gewichte*Geschlecht / Nationalität, sortiert nach Alter (n=100.000)

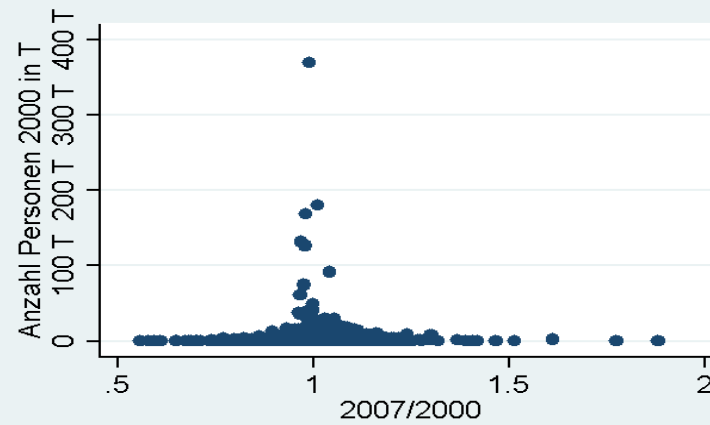
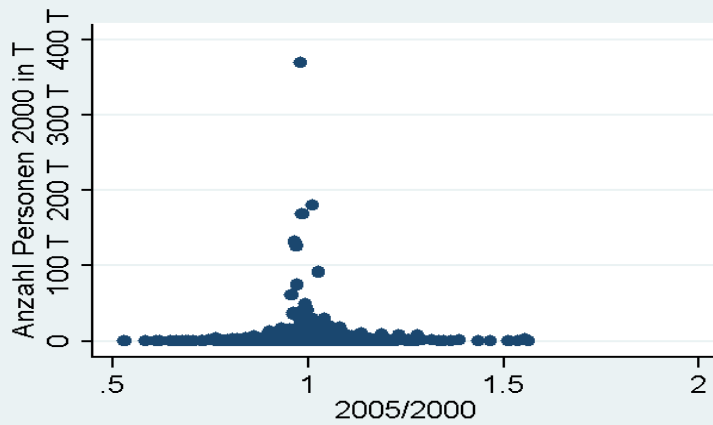
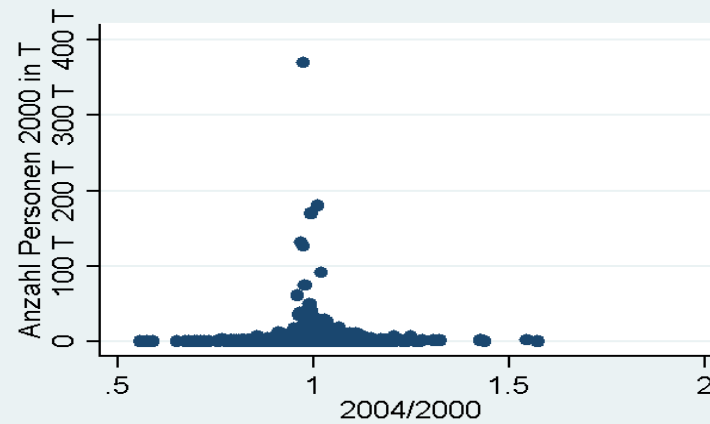
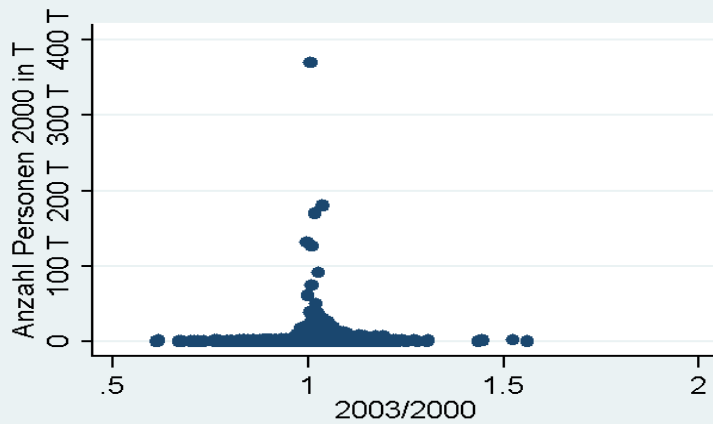


Nach Ebene

3.4 Notwendige Gewichtung

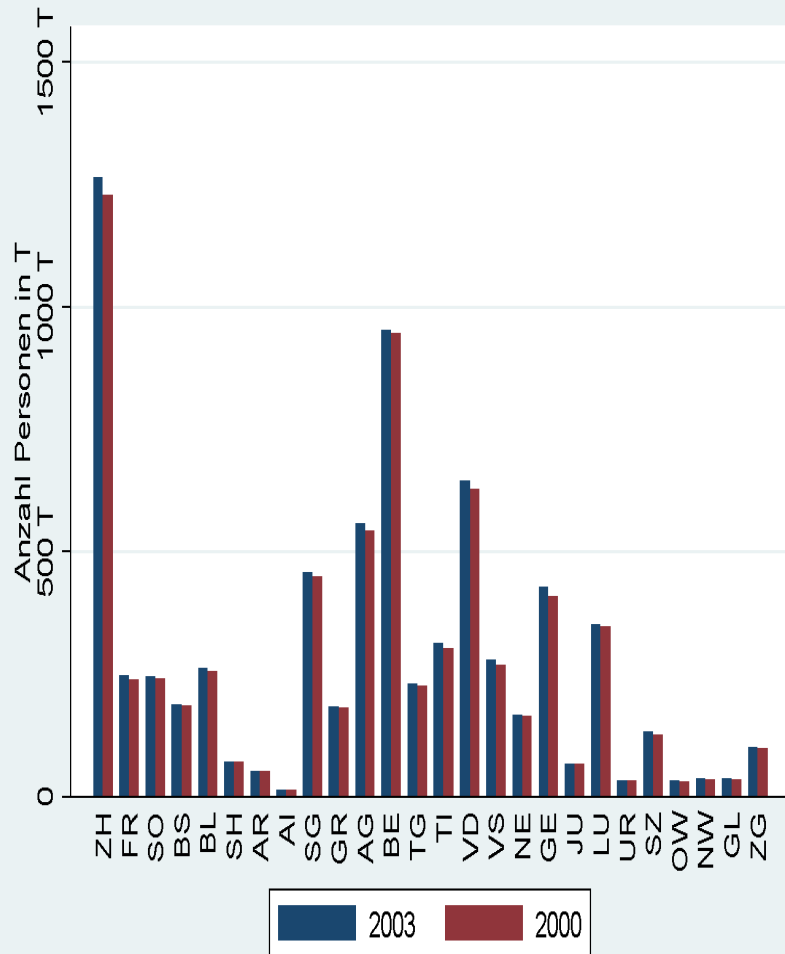
17

Scatterplot : Zusammenhang Gemeindegrosse 2000 / Veränderung Gewichtungsfaktor auf Gemeindeebene

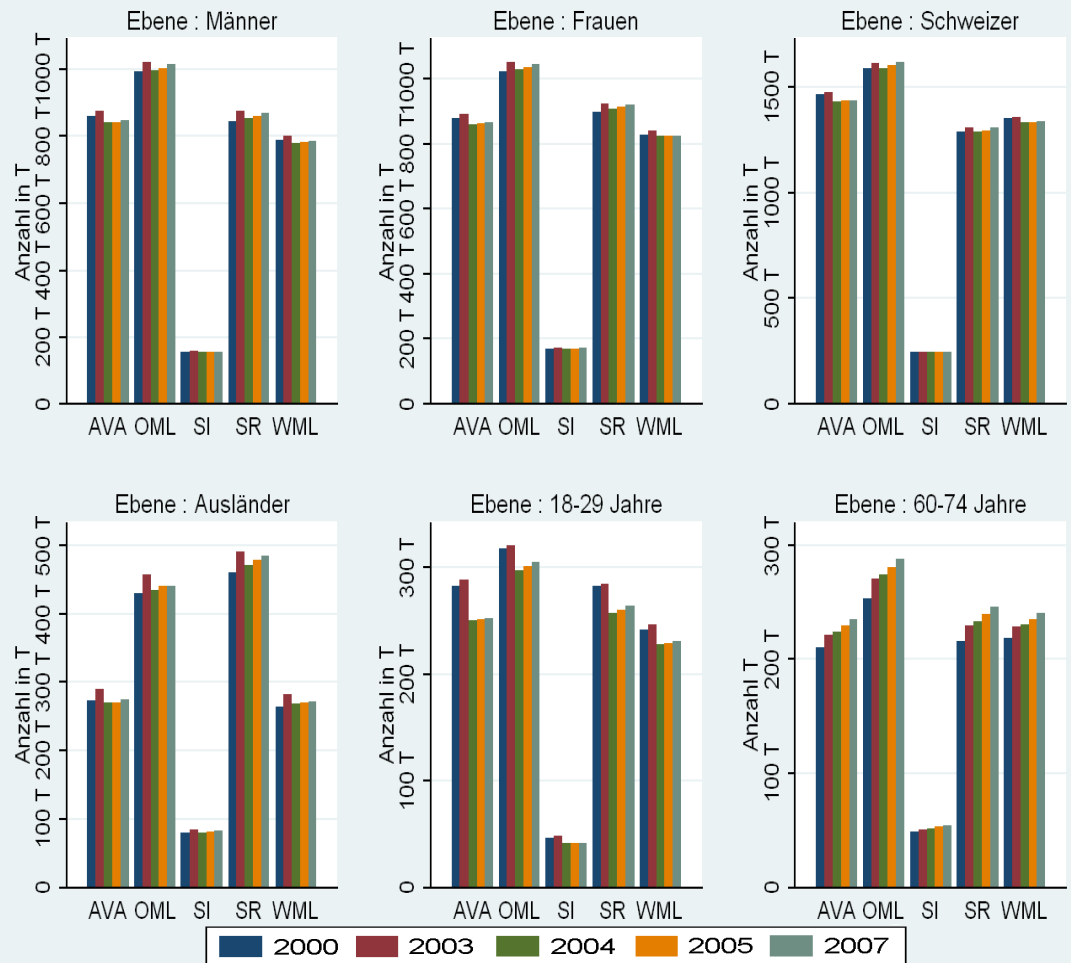


3.5 Einfluss der Gewichtung

Strukturveränderung nach Kanton 2000/2003



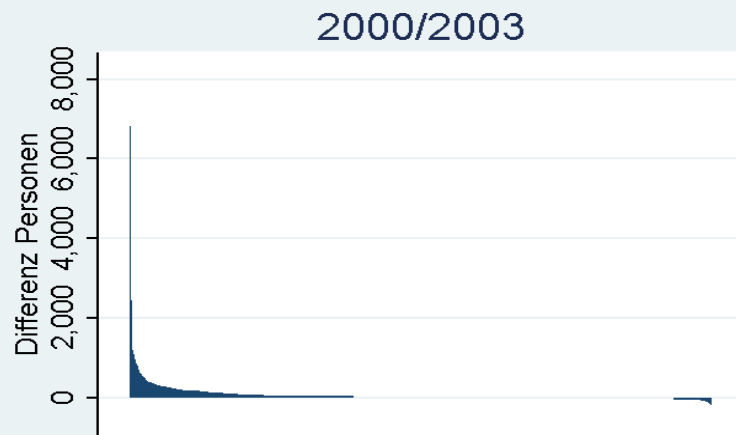
Strukturveränderung nach WEMF* Ebenen



3.5 Einfluss der Gewichtung

Veränderung nach Gemeindenummer zu 2000, absteigend sortiert

19

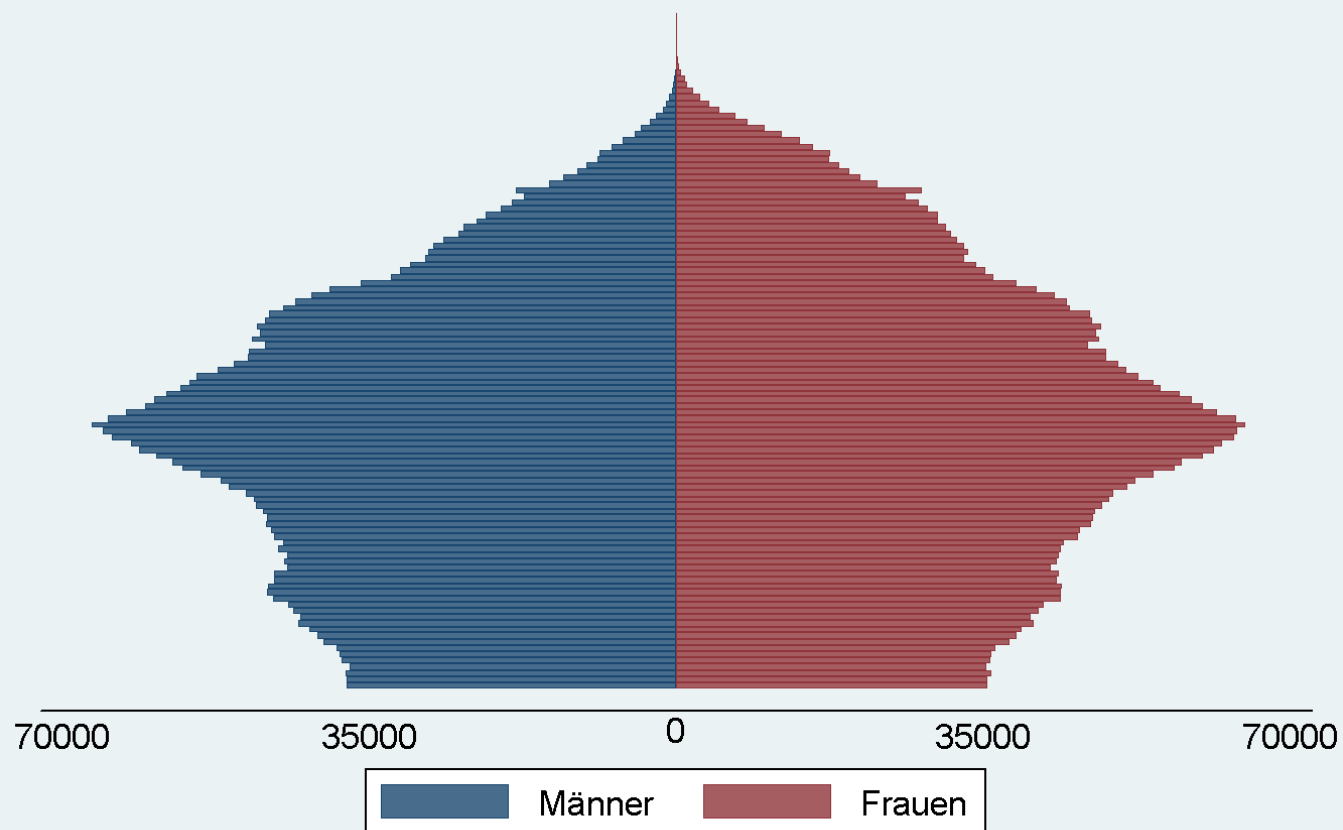


3.5 Einfluss der Gewichtung

Population nach Definition für Fortschreibung

20

Population in 2007



3.5 Einfluss der Gewichtung

2000

Region	Summe	Alter und Geschlecht					
		Männlich, 15-34 Jahre	Männlich, 35-54 Jahre	Männlich, 55-74 Jahre	Weiblich, 15-34 Jahre	Weiblich, 35-54 Jahre	Weiblich, 55-74 Jahre
Zürich Agglo (ohne Stadt)	511'852	83'985	106'562	64'187	83'572	107'068	66'478
Zürich Stadt	250'796	50'403	46'714	27'156	47'231	45'741	33'551
	762'648	134'388	153'276	91'343	130'803	152'809	100'029

2003

Region	Summe	Alter und Geschlecht					
		Männlich, 15-34 Jahre	Männlich, 35-54 Jahre	Männlich, 55-74 Jahre	Weiblich, 15-34 Jahre	Weiblich, 35-54 Jahre	Weiblich, 55-74 Jahre
Zürich Agglo (ohne Stadt)	533'369	82'521	113'286	69'784	83'409	113'343	71'025
Zürich Stadt	251'643	47'462	48'549	28'597	45'238	47'184	34'612
	785'012	129'984	161'835	98'381	128'647	160'527	105'637

2004

Region	Summe	Alter und Geschlecht					
		Männlich, 15-34 Jahre	Männlich, 35-54 Jahre	Männlich, 55-74 Jahre	Weiblich, 15-34 Jahre	Weiblich, 35-54 Jahre	Weiblich, 55-74 Jahre
Zürich Agglo (ohne Stadt)	535'534	81'270	114'246	71'005	82'335	114'542	72'137
Zürich Stadt	248'144	45'469	48'209	28'954	43'595	47'098	34'818
	783'678	126'739	162'456	99'959	125'930	161'640	106'955

3.5 Einfluss der Gewichtung

2005

Region	Summe	Alter und Geschlecht					
		Männlich, 15-34 Jahre	Männlich, 35-54 Jahre	Männlich, 55-74 Jahre	Weiblich, 15-34 Jahre	Weiblich, 35-54 Jahre	Weiblich, 55-74 Jahre
Zürich Agglo (ohne Stadt)	541'347	81'590	115'609	72'264	82'821	115'847	73'216
Zürich Stadt	249'686	45'676	48'781	29'471	43'474	47'260	35'024
	791'033	127'266	164'390	101'735	126'295	163'107	108'240

2007

Region	Summe	Alter und Geschlecht					
		Männlich, 15-34 Jahre	Männlich, 35-54 Jahre	Männlich, 55-74 Jahre	Weiblich, 15-34 Jahre	Weiblich, 35-54 Jahre	Weiblich, 55-74 Jahre
Zürich Agglo (ohne Stadt)	547'116	82'135	117'141	73'318	83'185	117'245	74'093
Zürich Stadt	251'689	45'896	49'373	29'964	43'395	47'666	35'395
	798'805	128'031	166'514	103'282	126'580	164'910	109'488

Veränderung 2000-2007

Region	Summe	Alter und Geschlecht					
		Männlich, 15-34 Jahre	Männlich, 35-54 Jahre	Männlich, 55-74 Jahre	Weiblich, 15-34 Jahre	Weiblich, 35-54 Jahre	Weiblich, 55-74 Jahre
Zürich Agglo (ohne Stadt)	35'264	-1'850	10'579	9'131	-387	10'177	7'615
Zürich Stadt	893	-4'507	2'659	2'808	-3'836	1'925	1'844
	36'157	-6'357	13'238	11'939	-4'223	12'101	9'459

4. Fazit / Ausblick

23

- **Fazit**
 - Qualität – Hoch : Mittelwert/Verteilung der Gewichte
 - Genauigkeit – Hoch : Bis auf 0.0000078 % genau (2003)
- **Aus Marktforschungssicht**
 - Ermöglicht flexible und studienangepasste Stichprobenziehung
 - Räumliche, zeitliche und sachliche Komponente wird berücksichtigt
- **Ausblick**
 - Evaluation anhand neuer Volkszählung

Anhang : Abkürzungen (für Nichtschweizer😊)

24

ABKÜRZUNGEN DER KANTONSNAMEN DER SCHWEIZ

Name	Abkürzung	Name	Abkürzung
Aargau	AG	Nidwalden	NW
Appenzell Innerrhoden	AI	Obwalden	OW
Appenzell Ausserrhoden	AR	Sankt Gallen	SG
Bern	BE	Schaffhausen	SH
Basel-Land	BL	Solothurn	SO
Basel-Stadt	BS	Schwyz	SZ
Fribourg	FR	Thurgau	TG
Genève	GE	Ticino	TI
Glarus	GL	Uri	UR
Graubünden	GR	Valais	VS
Jura	JU	Vaud	VD
Luzern	LU	Zug	ZG
Neuchâtel	NE	Zürich	ZH

Quelle: <http://ead.nb.admin.ch/web/cats/schweiz.pdf>

Anhang : 3.5 Einfluss der Gewichtung

25

Population

