

2010–  
2018

00

Statistische  
Grundlagen und  
Übersichten

Neuchâtel 2019

# Strukturerhebung der eidgenössischen Volkszählung

Stichprobenplan und -ziehung, Gewichtung, Schätzverfahren  
und Pooling 2010–2018

## Methodenberichte der Sektion 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 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.

# Strukturerhebung der eidgenössischen Volkszählung

Stichprobenplan und -ziehung, Gewichtung, Schätzverfahren  
und Pooling 2010–2018

**Redaktion** Jann Potterat, BFS; Lionel Qualité, BFS;  
Daniel Assoulin, BFS  
**Herausgeber** Bundesamt für Statistik (BFS)

Neuchâtel 2019

**Herausgeber:** Bundesamt für Statistik (BFS)

**Auskunft:** Lionel Qualité, Tel. 058 467 24 54,  
lionel.qualite@bfs.admin.ch;  
Jann Potterat, Tel. 058 463 65 90  
jann.potterat@bfs.admin.ch

**Redaktion:** Jann Potterat, BFS; Lionel Qualité, BFS;  
Daniel Assoulin, BFS

**Inhalt:** Sektion Statistische Methoden, BFS

**Reihe:** Statistik der Schweiz

**Themenbereich:** 00 Statistische Grundlagen und Übersichten

**Originaltext:** Deutsch

**Layout:** Sektion DIAM, Prepress/Print  
Dieses Dokument wurde automatisch aus einer Datenbank  
heraus erzeugt. Es entspricht daher nicht dem üblichen  
typografischen Standard der BFS-Publikationen.

**Online:** [www.statistik.ch](http://www.statistik.ch)

**Print:** [www.statistik.ch](http://www.statistik.ch)  
Bundesamt für Statistik, CH-2010 Neuchâtel,  
order@bfs.admin.ch, Tel. 058 463 60 60  
Druck in der Schweiz

**Copyright:** BFS, Neuchâtel 2019  
Wiedergabe unter Angabe der Quelle  
für nichtkommerzielle Nutzung gestattet

**BFS-Nummer:** 338-0080

**ISBN:** 978-3-303-00631-3

# Inhaltsverzeichnis

<b>Vorwort und Zusammenfassung</b>	<b>5</b>
<b>1 Einleitung</b>	<b>7</b>
<b>2 Stichprobenrahmen</b>	<b>7</b>
<b>3 Stichprobenplan, Ziehung und Antwortverhalten</b>	<b>9</b>
3.1 Stichprobenplan . . . . .	9
3.2 Ziehung und Koordinierung . . . . .	10
3.3 Analyse Antwortverhalten . . . . .	12
<b>4 Gewichtung der Personen</b>	<b>12</b>
4.1 Initialgewichtung . . . . .	12
4.2 Kalibrierung . . . . .	14
<b>5 Gewichtung der Haushalte und der Zielpersonen der Haushalte</b>	<b>16</b>
5.1 Initialgewichtung . . . . .	16
5.2 Kalibrierung . . . . .	19
<b>6 Schätzungen</b>	<b>23</b>
6.1 Punktschätzungen . . . . .	23
6.1.1 Die Schätzung von Totalen, Anzahlen, Quotienten . . . . .	23
6.2 Varianzschätzung und Konfidenzintervall . . . . .	23
6.2.1 Einleitung . . . . .	23
6.2.2 Schätzung mit SURVEYFREQ (für Personen und Haushalte) . . . . .	24
6.2.3 Schätzung mit %CalVarSE . . . . .	25
<b>7 Pooling</b>	<b>27</b>
7.1 Einführung . . . . .	27
7.2 Ziele des Poolings . . . . .	27
7.3 Schätzer des Poolings . . . . .	29
7.3.1 Eigenschaften . . . . .	29
7.3.2 Genauigkeit . . . . .	30
7.4 Pooling über 5 Jahre, Entwicklungsmasse . . . . .	30
7.4.1 Pooling über 5 Jahre . . . . .	30
7.4.2 Vergleich von sukzessiven Pooling-Resultaten . . . . .	31
7.5 Empfehlungen . . . . .	31
<b>8 Schlussfolgerungen</b>	<b>32</b>
<b>Anhang</b>	<b>33</b>
<b>A Approximation der Varianz für das Pooling</b>	<b>33</b>
A.1 Aspekte über die Kovarianzen . . . . .	33
A.2 Vereinfachte Schätzung der Genauigkeit . . . . .	34
<b>B Fragebogen der Strukturerhebung 2015</b>	<b>35</b>
<b>Literatur</b>	<b>42</b>
<b>Methodenberichte der Sektion Statistische Methoden des BFS</b>	<b>43</b>



## Vorwort

Die Sektion Statistische Methoden METH des Bundesamts für Statistik wurde mit der Erarbeitung der statistischen Methoden für die Strukturhebung der eidgenössischen Volkszählung beauftragt.

Erstmals mussten die Methoden für das Erhebungsjahr 2010 entwickelt werden. Die Methoden für die jährlich stattfindende Erhebung wurden bis zum Jahr 2015 teilweise leicht angepasst, beziehungsweise optimiert.

Im Lauf der Zeit waren dabei mehrere Personen von METH involviert, insbesondere D. Assoulin, D. Kilchmann, D. Nedyalkova, Ch. Panchard, J. Potterat und L. Qualité.

Die methodischen Arbeiten rund um die Strukturhebung sind umfangreich und vielseitig. Dieser Bericht konzentriert sich auf die Aspekte Stichprobenplan, Gewichtung, Varianzschätzung und Pooling.

Die Autoren danken der Sektion Bevölkerung POP, welche zuständig für die Durchführung der Erhebung ist, insbesondere Ch. Freymond, J.-P. Kauthen, J. Voronkova und O. Wirz, für die sehr gute und angenehme Zusammenarbeit. Ein weiterer Dank geht an das Leitungsteam von METH (J.-P. Renfer, D. Kilchmann, D. Assoulin) für die methodische Unterstützung und die aufmerksame Durchsicht des Berichts.

## Zusammenfassung

Die Strukturhebung ist ein Element der seit 2010 jährlich stattfindenden eidgenössischen Volkszählung und ergänzt die Informationen aus den Registern mit zusätzlichen Statistiken zur Bevölkerungsstruktur.

Dieser Methodenbericht beschreibt die verwendeten statistischen Methoden rund um den Stichprobenplan, die Hochrechnung sowie das Pooling (Kombinieren mehrerer Erhebungsjahre).

Die Stichprobe der Strukturhebung wird im Stichprobenrahmen für Personen und Haushaltserhebungen SRPH gezogen. Die erwünschte nationale Netto-Stichprobe (verwendbare Antworten) wird auf 200'000 Personen festgelegt. Verschiedene Gebiete wie Kantone oder grössere Städte lassen die Stichprobe in ihrem Gebiet aufstocken. Zusammen mit der Antizipation von erwarteten Antwortausfällen führte dies zu einer Bruttostichprobengrösse von ca. 330'000 Personen. Eine ebenfalls vorgesehene Reservestichprobe musste nie aktiviert werden, da die Antwortrate ungefähr den Erwartungen von 85% entsprach.

Für die antwortenden Einheiten werden in mehreren Schritten unter Berücksichtigung der Einschliesswahrscheinlichkeiten, von geschätzten Antwortwahrscheinlichkeiten in einzelnen Unterkategorien und von im Bevölkerungsregister (STATPOP) bekannten Totalen Hochrechnungsgewichte berechnet. Es wurden Gewichte für Personen (Personengewichte) und Gewichte für Auswertungen auf Ebene Haushalte (Haushaltsgewichte) entwickelt.

Zu jedem geschätzten Merkmal werden im BFS mit der Software SAS Varianzen und Konfidenzintervalle berechnet. Die Formeln dazu sind in diesem Bericht ebenso beschrieben wie die Methoden zur Schätzung von Resultaten, basierend auf gepoolten Daten inklusive Varianzen. Die Schätzwerte von mehreren Jahren können zusammen kombiniert (gepoolt) werden. Dies hat den Vorteil von genaueren Resultaten oder erlaubt die Analyse für kleinere Untersuchungsgebiete, da die Stichprobengrösse durch das Pooling vergrössert wird.



# 1 Einleitung

Die Strukturhebung (SE) ist ein Element der seit 2010 jährlich stattfindenden eidgenössischen Volkszählung und ergänzt die Informationen aus den Registern mit zusätzlichen Statistiken zur Bevölkerungsstruktur. Sie ist aktuell die grösste Befragung bei Personen und Haushalten des Bundesamts für Statistik BFS und erlaubt es, jährliche Resultate mit einer guten Genauigkeit für Untersuchungsbereiche in der Grössenordnung ab 15'000 Personen (siehe [BFS, 2011, 2008](#)) zu publizieren. Der Fragebogen, die Stichprobengrössen und die statistischen Methoden der Erhebung sind von Jahr zu Jahr sehr ähnlich. Die Koordinierung der Stichproben (siehe [Qualité, 2009](#)) gewährleistet, dass die Überlappung von zwei aufeinanderfolgenden Erhebungen minimal ist, d.h. die gleichen Haushalte bzw. Personen werden nur in sehr seltenen Fällen mehrmals innerhalb der Zeitspanne des Poolings in die Stichprobe der SE gezogen.

Die in diesem Bericht vorgestellten statistischen Methoden der Strukturhebung haben sich seit 2015 nicht mehr wesentlich geändert. Da einige methodischen Aspekte in den Jahren 2014 und 2015 entwickelt worden sind, beziehen sich numerische Beispiele und Abbildungen zur Illustration der Methoden auf diese beiden Jahre. Infolge methodischer Anpassungen können die in den Vorgängererhebungen 2010-2013 verwendeten Methoden teilweise von den in diesem Bericht beschriebenen abweichen.

Der Stichprobenrahmen bzw. die Basis zur Ziehung der Stichprobe der Erhebung wird in Abschnitt [2](#) beschrieben. Der Stichprobenplan, d.h. der Beschrieb, wie viele Einheiten je Kanton bzw. Gebiet gezogen werden sollen, ist in Kapitel [3](#) dargestellt. Ebenso werden die Stichprobenziehung und die Analyse des Antwortverhaltens beschrieben.

Es werden unterschiedliche Hochrechnungsgewichte berechnet. Diese werden in den Kapiteln [4](#) und [5](#) beschrieben. Im Abschnitt [6](#) werden die Schätzverfahren und die verwendeten SAS-Prozeduren beschrieben.

Nach einigen Jahren können die Resultate von mehreren Strukturhebungen kombiniert und gemeinsam analysiert werden. Dies geschieht mit der Methode eines Poolings, welche im Kapitel [7](#) beschrieben wird.

## 2 Stichprobenrahmen

Die Stichprobe der Strukturhebung wird im Stichprobenrahmen für Personen- und Haushaltshebungen (SRPH) gezogen. Dieser Rahmen wird für alle Personen- bzw. Haushalts-Ziehungen viermal im Jahr (Mitte Februar, Mai, August und November) aufgrund neuer Daten der kantonalen und kommunalen Einwohnerregister für Ziehungen bereitgestellt und aktualisiert. Die Strukturhebung wird dabei jährlich aus dem November-Rahmen gezogen, welcher sich auf die genannten Register mit Stand 30. September des Referenzjahres bezieht. Dabei werden diese noch mit Informationen aus Bundesregistern wie dem Zivilstandsregister, dem Asyl-Register, dem Register der Ausländer und dem Register der Diplomaten und internationalen Funktionäre angereichert.

Folgende Kriterien sind massgebend für die Bestimmung der Grundgesamtheit der Erhebung.

- Kategorie der Bevölkerung: keine Personen unter Vormundschaft und seit 2011 keine Diplomaten und internationale Funktionäre, sowie deren Familien.

- Aufenthaltsstatus: nur ständige Wohnbevölkerung, bzw. Aufenthaltsstatus mindestens 12 Monate für Personen im Asylbereich.
- Geburtsdatum: Die Zielpersonen müssen am Stichtag 31. Dezember des Referenzjahres 15-jährig oder älter sein. Zum Beispiel für die SE2015: die Zielpersonen müssen am oder vor dem 31.12.2000 geboren sein.
- Haushaltsart: Privat- Kollektiv- oder Sammelhaushalt; nur Privathaushalte sind für die Erhebung relevant.

Zudem werden folgende beiden Registervariablen für den Stichprobenplan (siehe 3) benutzt:

- Ort des Hauptwohnsitzes.
- Haushaltszusammensetzung /-bildung (seit Ende 2012 obligatorisch in den Einwohnerregistern enthalten).

Die Grundgesamtheit der Strukturhebung ist also definiert durch die Personen, die in Privathaushalten mit einer Schweizer Adresse leben, am 31.12. des Referenzjahres 15 oder älter sind, mit einer (ständigen) permanenten Aufenthaltsbewilligung, oder einer von mindestens 12 Monaten für Personen im Asylbereich. Ausgeschlossen werden Diplomaten und internationale Funktionäre (inkl. deren Familienangehörige) sowie alle Personen unter Vormundschaft. Die Variablen im SRPH, die diese Selektion vornehmen, sind in Tabelle 1 dargestellt.

**Tabelle 1** Bestimmung der Grundgesamtheit/Variablen in Datei

Variable	Beschrieb
TYPEOFHOUSEHOLD=1 (oder indic_typeMenage=1) indic_Tutelle=0	Nur Privathaushalte (1 sind die Privathaushalte) Keine Personen unter Vormundschaft (gemäss Kontaktadresse)
indic_AdresseEtranger=0 indic_ordipro=0	Nur Personen mit einer Schweizer Adresse Keine Diplomaten und internationale Funktionäre (inkl. deren Familienangehörige)
indic_organ=0	Ohne Bezeichnung «Organisation» in den Variablen der Kontaktadresse
deliveryKanton ne " "	Kanton vorhanden (immer der Fall, früher gab es noch Missings)
indic_Pop_SE20XX=1 beinhaltet: populationType = '1'	Nur die ständige Wohnbevölkerung (am Hauptwohnsitz). D.h. die nicht ständige Wohnbevölkerung (2) und die Population am letzten Nebenwohnsitz (3) und die übrigen Nebenwohnsitze (4) werden ausgeschlossen.
dateofBirth <= 31.12.yyyy	Personen, die am 31.12. des entsprechenden Jahres mindestens 15-jährig sind (yyyy = Jahr der SE minus 15).
indic_Protection=0	Die Personen, die noch vor Einführung des SRPH (im Nov. 2010) in Erhebungen gezogen wurden, erhalten einen Sperrflag (indic_Protection=1) und werden für eine bestimmte Dauer für zukünftige Ziehungen gesperrt.

**Bemerkung:** Im ersten Jahr SE2010 wurde noch eine Stichprobe von Diplomaten und internationalen Funktionären gezogen ( $n \approx 1500$ ). Da für diese Subpopulation kein Analyseauftrag bestand und die Antwortquote für diese Kategorie sehr schlecht war (16%), wurde diese Gruppe in den nächsten Jahren nicht mehr in den Stichprobenrahmen integriert.

### 3 Stichprobenplan, Ziehung und Antwortverhalten

#### 3.1 Stichprobenplan

Generell wurde eine Einschusswahrscheinlichkeit proportional zur Bevölkerungsgrösse in den Kantonen so festgelegt, dass die erwartete Netto-Stichprobe für die Erhebung rund 200'000 Personen enthält. Die Bevölkerungsgrösse wurde aus der Anzahl Personen gemäss STATPOP aus dem Vorjahr entnommen. Diese Stichprobe konnte in einzelnen Gebieten (Kantonen oder Gemeinden) auf Wunsch und Kosten dieser Gebiete erhöht werden (verdoppelt oder bis zu vervierfacht). Zudem wurde eine Antwortquote von 85% angenommen und in jedem Jahr bis 2015 noch eine Reservestichprobe von 10% gezogen, welche im Falle einer schlechten Antwortquote hätte aktiviert werden können. Da dies aber nie der Fall war, wurde ab 2016 keine Reserve mehr gezogen. Die so resultierenden Bruttostichprobengrössen und andere Eckwerte für die Jahre 2010 bis 2018 sind aus der Tabelle 2 ersichtlich. Es ist zu sehen, dass die effektive Antwortquote bis 2017 über 85% liegt, daher ist auch die effektive Anzahl Antwortende generell leicht höher als die geplante Netto-Stichprobengrösse (inkl. Aufstockung).

**Tabelle 2** Rahmen, Brutto- und Nettostichprobengrössen 2010-2018

Jahr	STATPOP <sup>1</sup> Vorjahr	Basis Stichprobe	Aufstockung Kantone	Stichprobe netto (plan)	Zuschlag Antwortausfälle	Stichprobe brutto (o.R.)	Zuschlag Reserve
2010 <sup>2</sup>	5'767'737	200'000	119'680	319'680	41'820	361'500	40'000
2011	6'548'570	200'000	76'218	276'218	50'782	327'000	33'000
2012	6'608'399	200'000	79'141	279'141	50'859	330'000	33'000
2013	6'671'959	200'000	76'105	276'105	50'895	327'000	32'700
2014	6'756'402	200'000	76'289	276'289	52'711	329'000	32'900
2015	6'839'898	200'000	63'903	263'903	50'297	314'200	31'420
2016	6'917'839	200'000	71'880	271'880	51'820	323'700	-
2017	6'990'692	200'000	62'284	262'284	54'716	317'000	-
2018	7'040'745	200'000	55'795	255'795	56'205	312'000	-

  

Jahr	Stichpr.brutto (inkl. Res.) Plan erwartet	Stichprobe gezogen	Aktivierte effekt. Stichpr.	Antwortende netto effektiv	Ziel-Popul Erwachsene	Kalibrierung inkl. Kinder	effekt. Antw.- quote in %
2010	401'500	406'351	365'883	317'221	6'519'253	7'733'493	86.70
2011	360'000	359'122	326'251	282'749	6'587'556	7'770'159	86.67
2012	363'000	362'193	329'024	286'015	6'662'333	7'850'782	86.93
2013	359'700	358'754	325'915	281'990	6'744'794	7'944'566	86.52
2014	361'900	366'528	333'054	287'842	6'829'610	8'041'310	86.43
2015	345'620	345'731	314'109	269'509	6'907'818	8'131'033	85.80
2016	323'700	322'431	322'431	274'213	6'981'381	8'224'553	85.05
2017	317'000	314'464	314'464	268'495	7'036'199	8'293'618	85.38
2018	312'000	314'155	314'155	266'537	7'084'068	8'356'225	84.84

<sup>1</sup> im 2010: Basis VZ2000, <sup>2</sup> inkl. Zusatzstichprobe für Diplomaten und internationale Funktionäre (Ordipro)

Die Tabelle 3 zeigt die Kantone (oder Teile davon), welche für gewisse Jahre eine Erhöhung, meistens eine Verdoppelung, ihrer Stichprobe beantragt haben.

**Tabelle 3** Erhöhung der Stichprobe in speziellen Kantonen (Gebieten) 2010-2018

Jahr	Gebiete mit Erhöhung (Verdoppelung)
2010	AG, BE-Jura, BE-Stadt (x4), GE, JU, LU, NE, TG, TI, VD, ZG, ZH-Stadt (x4)
2011	AG, BE-Jura, GE, JU, LU, NE, TG, TI, VD
2012	AG, BE-Jura, GE, JU, LU, NE, TG, TI, VD, ZG
2013/2014	AG, BE-Jura, BE-Stadt, GE, JU, LU, NE, TI, VD, ZG
2015	BE-Stadt, GE, LU, NE, TG, TI, VD, ZG
2016	BE-Stadt, GE (x1.98), LU (x1.83), NE, TG, TI, VD, ZG, ZH-Stadt
2017	BE-Stadt, GE (x1.97), LU (x1.83), NE, TG, TI, VD, ZG
2018	BE-Stadt, GE (x1.97), LU (x1.83), NE, TI, VD, ZG

## 3.2 Ziehung und Koordinierung

Damit die Belastung der Personen, welche an den Erhebungen teilnehmen müssen, möglichst gut unter der ganzen Bevölkerung verteilt wird, werden im BFS alle Stichproben im Stichprobenverwaltungssystem (negativ) koordiniert. Die Methode wird in [Qualité \(2009\)](#) beschrieben und beruht auf einer Poisson-Ziehung, wobei die (Nicht-)Ziehung einer Einheit in die Stichprobe unabhängig ist von der (Nicht-)Ziehung einer anderen Einheit der Grundgesamtheit.

Die Koordinierungsmethode entwickelte sich im Verlaufe der ersten Strukturhebungen. Für die SE2010 und SE2011 erfolgte die Koordination ausschliesslich auf der Ebene der Einzelpersonen. Bis auf wenige Ausnahmen konnte für die SE2010 und SE2011 nicht die gleiche Person ausgewählt werden. Ausnahmen sind: Personen, die ihre AHV-Sozialversicherungsnummer geändert haben, Personen, deren AHV-Nummer in den Erhebungsrahmen verschwunden und wiederaufgetaucht sind (z.B. Personen, die einmal ins Ausland gegangen sind), und aus technischen Gründen Personen, die in sehr grossen Haushalten leben.

Eine zusätzliche Einschränkung wurde der SE2012-Ziehung auferlegt: Zusätzlich zur Ziehung einer mit früheren Erhebungen negativ koordinierten Stichprobe von Personen wurde sichergestellt, dass die für die SE2012 ausgewählten Haushalte (Haushalte mit gezogenen Personen) keine bereits für die SE2011 ausgewählten Personen enthielten.

Ein Problem war aber, dass bis Ende 2012 die Haushaltsbildung mittels EGID und EWID in den Einwohnerregistern noch nicht obligatorisch war. Daher konnte diesem Wunsch nicht in allen Kantonen zu 100% Rechnung getragen werden. Die Personen, bei denen die Haushaltsbildung nicht klar war, wurden als Einpersonenhaushalte definiert und deren Einschlusswahrscheinlichkeiten bei der Ziehung leicht erhöht. Bei der Erhebung wurden Filterfragen eingebaut, mit welchen erreicht werden sollte, dass doppelt befragte Haushalte nur einmal antworten. Falls in einem Haushalt mehrere Personen gezogen wurden, wurde aufgrund der Geburtstage der Personen, zufällig jene ausgewählt, nämlich jene, die als Erste im Jahr Geburtstag hat.

Das Problem der Mehrfachbefragung innerhalb von Haushalten ist seit der SE2013 gelöst, da seither die Koordination vollständig auf Haushaltsebene und nicht mehr auf Einzelpersonenebene geschieht. Die Haushaltsstichproben sind nach Möglichkeit für Haushalte, die ihre Zusammensetzung nicht ändern, disjunkt. Andererseits wird akzeptiert, dass Personen, die den Haushalt wechseln, wieder gezogen werden können, was einigen hundert Fällen pro Jahr entspricht. Das beobachtete Ergebnis der personenbezogenen koordinierten Ziehungen ist, dass sich die Haushaltsstichproben zwischen den SE2010 und SE2011 (etwa 26'000 betroffene Haushalte) und zwischen den SE2010 und SE2012 (etwa 28'000 Haushalte) ziemlich stark überschneiden. Ihre Überlappung ist Null zwischen den SE2011 und SE2012 und Null oder minimal zwischen den folgenden Strukturhebungen. Es sei darauf hingewiesen, dass sich die Zusammensetzung der Haushalte im Laufe der Zeit ändert, so dass das Konzept der Überlappung von Haushaltsstichproben nicht trivial ist.

Da seit der SE2013 Haushalte gezogen und koordiniert werden, die Grundgesamtheit und der Plan jedoch weiterhin auf Niveau Personen definiert sind, gibt es mehrere Schritte, welche notwendig sind, um die Stichprobenziehung vorzunehmen. Dies geschieht wie folgt:

1. Aufteilung der Population nach den Schichten. In der SAS-Datei werden diese Schichten als Zonen bezeichnet. Im 2014 gab es 28 Zonen mit den Kantonen, wobei Bern in 3 Gebiete aufgeteilt wird. Die Anzahl Personen je Schicht  $h$  wird mit  $N_h$  bezeichnet (im SAS: Nh\_Pers\_popul). Die Anzahl Personen je Schicht, die gemäss Plan in die Brutto-Stichprobe zu ziehen sind, wird mit  $n_h$  bezeichnet.

2. Berechnung der Einschusswahrscheinlichkeiten der Personen  $P$ :  $\pi_P = n_h/N_h$ .  
(SAS:<sup>1</sup> tx\_erheb\_40\_P= brutto\_inkl\_res / Nh\_Pers\_popul)
3. Berechnung der Einschusswahrscheinlichkeiten der Haushalte durch Summierung aller Personen-Einschlusswahrscheinlichkeiten innerhalb desselben Haushalts  $H$ :  
 $\pi_H^0 = \sum_{P \in H} \pi_P$ . (SAS: tv\_erheb\_40\_H = sum\_P {tx\_erheb\_40\_P})
4. Diese Summierung könnte bei grossen Haushalten grösser als 1 sein, daher wird Summe begrenzt auf 1:  $\pi_H = \min(1, \pi_H^0)$ . (SAS: tx\_erheb\_40\_H = min{tv\_erheb\_40\_H,1})
5. Bestimmung der ziehbaren Haushalte: jene mit  $\pi_H > 0$ .  
(SAS: if tx\_erheb\_40\_H >0 then popul\_erheb\_40=1)
6. Durchführung der Haushaltsziehung im SRPH mittels einer Poissonziehung, d.h. die effektive Stichprobengrösse ist zufällig.
7. Personenziehung: 1 Person pro gezogenen Haushalt mit proc surveyselect im SAS. Es wird der Haushaltsidentifikator  $hn$  als Schichtvariable angegeben und als *size* Variable wird die Einschusswahrscheinlichkeit der Personen angegeben. Letztere hat aber keinen Einfluss auf die pps<sup>2</sup>-Ziehung, da alle Personen im Haushalt dieselbe Wahrscheinlichkeit besitzen. Es wird also pro Haushalt eine Person zufällig unter allen Personen ausgewählt. Die Stichprobengrösse  $\_nsize_=1$  wird in der Datei *tailles\_strates* pro Haushalt angegeben.

```
proc surveyselect data=pers_pr_tirage out=output.ech_pers_40 seed=xxxx
method=pps sampsize=tailles_strates noprint;
strata hn;
size tx_erheb_40_P;
run;
```

8. Zufällige Aufteilung der gezogenen Personen in die effektiv aktivierte Brutto-Stichprobe, welche zwingend versandt wird, und die Reservestichprobe. Letztere hätte bei Bedarf auch aktiviert werden können, wenn die Antwortquoten zu tief gewesen wären, doch dies war nie der Fall<sup>3</sup>. Die Zuordnung, ob eine Person zur aktivierten Stichprobe oder zur Reserve gehört, erfolgte innerhalb der Schichten. In jeder Schicht ist ein Reserve-Anteil  $R_h = Z_h/n_h$  definiert, mit  $Z_h$  der erwünschten Reservegrösse und  $n_h$  der geplanten Brutto-Stichprobengrösse inkl. Reserve. Für jede Person  $k$  in der Schicht  $h$  wurde eine Zufallszahl  $u_{k,h}$  definiert. Falls  $u_{k,h} < R_h$  gehört die Person zur Reserve, sonst zur aktivierten Stichprobe. Die resultierenden Stichprobengrössen sind somit zufällig.  
Dieses gewählte Vorgehen erklärt, wieso in der Tabelle 2 die Differenz zwischen den Spalten «Stichpr.brutto (inkl. Res.) gezogen» und «Aktivierte effekt. Stichpr.», also die effektive Reserve, nicht exakt der Spalte «Zuschlag Reserve» entspricht.

<sup>1</sup>Die SE2014 ist die vierzigste Erhebung, welche im SRPH gezogen wurde. Daher steht in den SAS-Variablen die Zahl 40.

<sup>2</sup>für «probability proportional to size».

<sup>3</sup>Daher wurde ab der SE2016 keine Reservestichprobe mehr gezogen.

### 3.3 Analyse Antwortverhalten

In der letzten Spalte der Tabelle 2 sind die Antwortquoten der Erhebungen 2010 bis 2018 ersichtlich. Obwohl die Erhebung obligatorisch ist, betragen die Antwortausfälle rund 14%. Mögliche Erklärungen:

1. Die Zielpopulation im Ziehungsrahmen von Ende September ist nicht genau jene im Moment der Erhebung (ab Januar). Einige Personen sind der Zielpopulation entfallen: Todesfälle, Auswanderung aus der Schweiz, Eintritt in einen Kollektivhaushalt etc. Zudem können Personen bereits im Ziehungsrahmen fälschlicherweise der Zielpopulation zugeordnet worden sein.
2. Für 2010: Mehrfache Antworten innerhalb der Haushalte wurden gelöscht, falls Haushalte, trotz Anweisung, mehrere Fragebogen ausgefüllt haben.
3. Antwortverweigerung: Dies ist der häufigste Grund.
4. Obwohl ein umfangreicher **Statistischer DatenAufarbeitungsProzess** (SDAP) definiert und Einsetzungen vorgenommen wurden, gab es einzelne Fragebogen, deren Qualität nicht ausreichend war (zu viele fehlende Werte oder Inkohärenzen) und deshalb eliminiert wurden.

## 4 Gewichtung der Personen

Die Personengewichtung geschieht in mehreren Schritten. Diese werden in den folgenden Abschnitten erläutert.

### 4.1 Initialgewichtung

Die Initialgewichtung basiert auf den Einschlusswahrscheinlichkeiten der Zielpersonen gemäss Stichprobenplan. Die Einschlusswahrscheinlichkeit wird angepasst, da nur die Stichprobe ohne Reserve aktiviert wurde. Die Tabelle 4 zeigt die Schritte zur Berechnung der Initialgewichtung.

**Tabelle 4** Berechnung der Initialgewichtung Personen

Notation	SAS	Bedeutung
In Abschnitt 3.2 wird die Ziehungswahrscheinlichkeit (inkl. Reserve) wie folgt bezeichnet:		
$\pi_P = n_h/N_h$ $N_h$ $n_h$	tx_erheb_40_P Nh_Pers_popul brutto_inkl_res	Ziehungswahrscheinlichkeit inkl Reserve, mit Personen je Schicht $h$ Erwartete Anzahl Personen Brutto-Stichprobe je Schicht gemäss Plan.
Die angepasste Ziehungswahrscheinlichkeit wird wie folgt bezeichnet		
$\tilde{\pi}_P = \pi_P(1 - \tilde{n}_h/n_h)$ $\tilde{n}_h$ $w_{k,P}^{des} = \tilde{\pi}_P^{-1}$	tb_erheb_40_P Reserve w_design	Angepasste Ziehungswahrscheinlichkeit ohne Reserve, mit Erwartete Anzahl Personen Reservestichprobe je Schicht gemäss Plan. Design-Personengewicht (Initialgewicht Bruttostichprobe vor allen Korrekturen)
Es folgt eine Anpassung für totale Antwortausfälle mittels logistischer Regression (Details s. unten)		
$r_{k,P}$ $\tilde{\pi}_{k,P} = \tilde{\pi}_P r_{k,P}$ $w_{k,P}^{ini} = \tilde{\pi}_{k,P}^{-1}$	p_rep_zp pikpk_p w_ini_zp_rep	Vorhergesagte Antwortwahrscheinlichkeit der Person $k$ Für Antwortausfälle korrigierte Einschlusswahrscheinlichkeit Für Antwortausfälle korrigiertes Initialgewicht (Input %ca1mar2)

Mit einer logistischen Regression wird die Wahrscheinlichkeit modelliert, dass eine Einheit antwortet ( $p\_rep\_zp=1$ ), unter Verwendung der Regressionsvariablen Geschlecht, Zone, Zivilstand,

Haushaltsgrösse und -zusammensetzung, Nationalität, Altersklassen, Aufenthaltsbewilligung. Die Codes bzw. Level aller Variablen sind im Abschnitt 4.2 beschrieben. Der entsprechende SAS-Code ist der folgende.

Logistische Regression:

```
proc logistic data=brut_pour_nr outmodel=out_zp.se20yy_model_nr_zp;
class sex cl_maritalstatus compresofphperm zonere nationalit_grob
altklass_grob permis cl_men;
model in_net (event='1')=
sex cl_maritalstatus compresofphperm zonere nationalit_grob
altklass_grob permis cl_men /rsquare selection=stepwise lackfit;
quit;
```

Anwenden und Prädiktion:

```
proc logistic inmodel=out_zp.se20yy_model_nr_zp(type=logismod);
score clm data = brut_pour_nr
out = out_zp.se20yy_predict_pk_zp (rename=(P_1=p_rep_zp)) ;
run;
```

Die Abbildungen 1 und 2 zeigen die Outputs der logistischen Regressionen

**Abbildung 1** Logistische Regression: Output 2014 und 2015

Analyse Stepwise Selection						Verteilung geschätzte Antwortwahrscheinlichkeiten Personen (Bruttostichprobe)		Analyse der Effekte																																																																																																																																																
<b>SE 2015</b>						<b>Quantiles (Definition 5)</b>		<b>Type 3 Analysis of Effects</b>																																																																																																																																																
<table border="1"> <thead> <tr> <th colspan="6">Summary of Stepwise Selection</th> </tr> <tr> <th rowspan="2">Step</th> <th colspan="2">Effect</th> <th colspan="2">Number</th> <th>Score</th> <th rowspan="2">Pr &gt; ChiSq</th> </tr> <tr> <th>Entered</th> <th>DF</th> <th>In</th> <th>Out</th> <th>Chi-Square</th> </tr> </thead> <tbody> <tr><td>1</td><td>nationalit_grob</td><td>3</td><td>1</td><td></td><td>4830.2383</td><td>&lt;.0001</td></tr> <tr><td>2</td><td>cl_maritalstatus</td><td>2</td><td>2</td><td></td><td>4233.1232</td><td>&lt;.0001</td></tr> <tr><td>3</td><td>cl_men</td><td>5</td><td>3</td><td></td><td>1066.4083</td><td>&lt;.0001</td></tr> <tr><td>4</td><td>compResOfPHHPerm</td><td>3</td><td>4</td><td></td><td>731.6098</td><td>&lt;.0001</td></tr> <tr><td>5</td><td>altklass_grob</td><td>3</td><td>5</td><td></td><td>526.3084</td><td>&lt;.0001</td></tr> <tr><td>6</td><td>sex</td><td>1</td><td>6</td><td></td><td>338.9155</td><td>&lt;.0001</td></tr> <tr><td>7</td><td>zoneRE</td><td>30</td><td>7</td><td></td><td>410.3767</td><td>&lt;.0001</td></tr> <tr><td>8</td><td>Permis</td><td>2</td><td>8</td><td></td><td>121.9887</td><td>&lt;.0001</td></tr> <tr><td colspan="6">removed: keine</td></tr> </tbody> </table>						Summary of Stepwise Selection						Step	Effect		Number		Score	Pr > ChiSq	Entered	DF	In	Out	Chi-Square	1	nationalit_grob	3	1		4830.2383	<.0001	2	cl_maritalstatus	2	2		4233.1232	<.0001	3	cl_men	5	3		1066.4083	<.0001	4	compResOfPHHPerm	3	4		731.6098	<.0001	5	altklass_grob	3	5		526.3084	<.0001	6	sex	1	6		338.9155	<.0001	7	zoneRE	30	7		410.3767	<.0001	8	Permis	2	8		121.9887	<.0001	removed: keine						<table border="1"> <thead> <tr> <th>Level</th> <th>Quantile</th> </tr> </thead> <tbody> <tr><td>100% Max</td><td>95.17%</td></tr> <tr><td>99%</td><td>94.36%</td></tr> <tr><td>95%</td><td>93.85%</td></tr> <tr><td>90%</td><td>93.09%</td></tr> <tr><td>75% Q3</td><td>91.09%</td></tr> <tr><td>50% Median</td><td>87.40%</td></tr> <tr><td>25% Q1</td><td>82.33%</td></tr> <tr><td>10%</td><td>76.84%</td></tr> <tr><td>5%</td><td>71.96%</td></tr> <tr><td>1%</td><td>62.26%</td></tr> <tr><td>0% Min</td><td>0.02%</td></tr> </tbody> </table>	Level	Quantile	100% Max	95.17%	99%	94.36%	95%	93.85%	90%	93.09%	75% Q3	91.09%	50% Median	87.40%	25% Q1	82.33%	10%	76.84%	5%	71.96%	1%	62.26%	0% Min	0.02%	<table border="1"> <thead> <tr> <th rowspan="2">Effect</th> <th rowspan="2">DF</th> <th colspan="2">Wald</th> </tr> <tr> <th>Chi-Square</th> <th>Pr &gt; ChiSq</th> </tr> </thead> <tbody> <tr><td>sex</td><td>1</td><td>350.38</td><td>&lt;.0001</td></tr> <tr><td>cl_maritalstatus</td><td>2</td><td>927.55</td><td>&lt;.0001</td></tr> <tr><td>compResOfPHHPerm</td><td>3</td><td>802.15</td><td>&lt;.0001</td></tr> <tr><td>zoneRE</td><td>30</td><td>411.32</td><td>&lt;.0001</td></tr> <tr><td>nationalit_grob</td><td>3</td><td>558.11</td><td>&lt;.0001</td></tr> <tr><td>altklass_grob</td><td>3</td><td>445.40</td><td>&lt;.0001</td></tr> <tr><td>Permis</td><td>2</td><td>109.69</td><td>&lt;.0001</td></tr> <tr><td>cl_men</td><td>5</td><td>716.53</td><td>&lt;.0001</td></tr> </tbody> </table>				Effect	DF	Wald		Chi-Square	Pr > ChiSq	sex	1	350.38	<.0001	cl_maritalstatus	2	927.55	<.0001	compResOfPHHPerm	3	802.15	<.0001	zoneRE	30	411.32	<.0001	nationalit_grob	3	558.11	<.0001	altklass_grob	3	445.40	<.0001	Permis	2	109.69	<.0001	cl_men	5	716.53	<.0001
Summary of Stepwise Selection																																																																																																																																																								
Step	Effect		Number		Score	Pr > ChiSq																																																																																																																																																		
	Entered	DF	In	Out	Chi-Square																																																																																																																																																			
1	nationalit_grob	3	1		4830.2383	<.0001																																																																																																																																																		
2	cl_maritalstatus	2	2		4233.1232	<.0001																																																																																																																																																		
3	cl_men	5	3		1066.4083	<.0001																																																																																																																																																		
4	compResOfPHHPerm	3	4		731.6098	<.0001																																																																																																																																																		
5	altklass_grob	3	5		526.3084	<.0001																																																																																																																																																		
6	sex	1	6		338.9155	<.0001																																																																																																																																																		
7	zoneRE	30	7		410.3767	<.0001																																																																																																																																																		
8	Permis	2	8		121.9887	<.0001																																																																																																																																																		
removed: keine																																																																																																																																																								
Level	Quantile																																																																																																																																																							
100% Max	95.17%																																																																																																																																																							
99%	94.36%																																																																																																																																																							
95%	93.85%																																																																																																																																																							
90%	93.09%																																																																																																																																																							
75% Q3	91.09%																																																																																																																																																							
50% Median	87.40%																																																																																																																																																							
25% Q1	82.33%																																																																																																																																																							
10%	76.84%																																																																																																																																																							
5%	71.96%																																																																																																																																																							
1%	62.26%																																																																																																																																																							
0% Min	0.02%																																																																																																																																																							
Effect	DF	Wald																																																																																																																																																						
		Chi-Square	Pr > ChiSq																																																																																																																																																					
sex	1	350.38	<.0001																																																																																																																																																					
cl_maritalstatus	2	927.55	<.0001																																																																																																																																																					
compResOfPHHPerm	3	802.15	<.0001																																																																																																																																																					
zoneRE	30	411.32	<.0001																																																																																																																																																					
nationalit_grob	3	558.11	<.0001																																																																																																																																																					
altklass_grob	3	445.40	<.0001																																																																																																																																																					
Permis	2	109.69	<.0001																																																																																																																																																					
cl_men	5	716.53	<.0001																																																																																																																																																					
<b>SE 2014</b>						<b>Quantiles (Definition 5)</b>		<b>Type 3 Analysis of Effects</b>																																																																																																																																																
<table border="1"> <thead> <tr> <th colspan="6">Summary of Stepwise Selection</th> </tr> <tr> <th rowspan="2">Step</th> <th colspan="2">Effect</th> <th colspan="2">Number</th> <th>Score</th> <th rowspan="2">Pr &gt; ChiSq</th> </tr> <tr> <th>Entered</th> <th>DF</th> <th>In</th> <th>Out</th> <th>Chi-Square</th> </tr> </thead> <tbody> <tr><td>1</td><td>nationalit_grob</td><td>3</td><td>1</td><td></td><td>5404.3511</td><td>&lt;.0001</td></tr> <tr><td>2</td><td>cl_maritalstatus</td><td>2</td><td>2</td><td></td><td>4139.0137</td><td>&lt;.0001</td></tr> <tr><td>3</td><td>cl_men</td><td>5</td><td>3</td><td></td><td>790.5315</td><td>&lt;.0001</td></tr> <tr><td>4</td><td>sex</td><td>1</td><td>4</td><td></td><td>412.8999</td><td>&lt;.0001</td></tr> <tr><td>5</td><td>zoneRE</td><td>31</td><td>5</td><td></td><td>450.87</td><td>&lt;.0001</td></tr> <tr><td>6</td><td>altklass_grob</td><td>3</td><td>6</td><td></td><td>324.548</td><td>&lt;.0001</td></tr> <tr><td>7</td><td>compResOfPHHPerm</td><td>3</td><td>7</td><td></td><td>419.4589</td><td>&lt;.0001</td></tr> <tr><td>8</td><td>Permis</td><td>2</td><td>8</td><td></td><td>164.5966</td><td>&lt;.0001</td></tr> <tr><td colspan="6">removed: keine</td></tr> </tbody> </table>						Summary of Stepwise Selection						Step	Effect		Number		Score	Pr > ChiSq	Entered	DF	In	Out	Chi-Square	1	nationalit_grob	3	1		5404.3511	<.0001	2	cl_maritalstatus	2	2		4139.0137	<.0001	3	cl_men	5	3		790.5315	<.0001	4	sex	1	4		412.8999	<.0001	5	zoneRE	31	5		450.87	<.0001	6	altklass_grob	3	6		324.548	<.0001	7	compResOfPHHPerm	3	7		419.4589	<.0001	8	Permis	2	8		164.5966	<.0001	removed: keine						<table border="1"> <thead> <tr> <th>Level</th> <th>Quantile</th> </tr> </thead> <tbody> <tr><td>100% Max</td><td>96.28%</td></tr> <tr><td>99%</td><td>94.30%</td></tr> <tr><td>95%</td><td>93.72%</td></tr> <tr><td>90%</td><td>93.14%</td></tr> <tr><td>75% Q3</td><td>91.62%</td></tr> <tr><td>50% Median</td><td>87.88%</td></tr> <tr><td>25% Q1</td><td>83.41%</td></tr> <tr><td>10%</td><td>77.93%</td></tr> <tr><td>5%</td><td>73.24%</td></tr> <tr><td>1%</td><td>63.40%</td></tr> <tr><td>0% Min</td><td>0.05%</td></tr> </tbody> </table>	Level	Quantile	100% Max	96.28%	99%	94.30%	95%	93.72%	90%	93.14%	75% Q3	91.62%	50% Median	87.88%	25% Q1	83.41%	10%	77.93%	5%	73.24%	1%	63.40%	0% Min	0.05%	<table border="1"> <thead> <tr> <th rowspan="2">Effect</th> <th rowspan="2">DF</th> <th colspan="2">Wald</th> </tr> <tr> <th>Chi-Square</th> <th>Pr &gt; ChiSq</th> </tr> </thead> <tbody> <tr><td>sex</td><td>1</td><td>402.90</td><td>&lt;.0001</td></tr> <tr><td>cl_maritalstatus</td><td>2</td><td>1189.26</td><td>&lt;.0001</td></tr> <tr><td>compResOfPHHPerm</td><td>3</td><td>403.22</td><td>&lt;.0001</td></tr> <tr><td>zoneRE</td><td>31</td><td>433.26</td><td>&lt;.0001</td></tr> <tr><td>nationalit_grob</td><td>3</td><td>672.42</td><td>&lt;.0001</td></tr> <tr><td>altklass_grob</td><td>3</td><td>443.83</td><td>&lt;.0001</td></tr> <tr><td>Permis</td><td>2</td><td>158.74</td><td>&lt;.0001</td></tr> <tr><td>cl_men</td><td>5</td><td>545.35</td><td>&lt;.0001</td></tr> </tbody> </table>				Effect	DF	Wald		Chi-Square	Pr > ChiSq	sex	1	402.90	<.0001	cl_maritalstatus	2	1189.26	<.0001	compResOfPHHPerm	3	403.22	<.0001	zoneRE	31	433.26	<.0001	nationalit_grob	3	672.42	<.0001	altklass_grob	3	443.83	<.0001	Permis	2	158.74	<.0001	cl_men	5	545.35	<.0001
Summary of Stepwise Selection																																																																																																																																																								
Step	Effect		Number		Score	Pr > ChiSq																																																																																																																																																		
	Entered	DF	In	Out	Chi-Square																																																																																																																																																			
1	nationalit_grob	3	1		5404.3511	<.0001																																																																																																																																																		
2	cl_maritalstatus	2	2		4139.0137	<.0001																																																																																																																																																		
3	cl_men	5	3		790.5315	<.0001																																																																																																																																																		
4	sex	1	4		412.8999	<.0001																																																																																																																																																		
5	zoneRE	31	5		450.87	<.0001																																																																																																																																																		
6	altklass_grob	3	6		324.548	<.0001																																																																																																																																																		
7	compResOfPHHPerm	3	7		419.4589	<.0001																																																																																																																																																		
8	Permis	2	8		164.5966	<.0001																																																																																																																																																		
removed: keine																																																																																																																																																								
Level	Quantile																																																																																																																																																							
100% Max	96.28%																																																																																																																																																							
99%	94.30%																																																																																																																																																							
95%	93.72%																																																																																																																																																							
90%	93.14%																																																																																																																																																							
75% Q3	91.62%																																																																																																																																																							
50% Median	87.88%																																																																																																																																																							
25% Q1	83.41%																																																																																																																																																							
10%	77.93%																																																																																																																																																							
5%	73.24%																																																																																																																																																							
1%	63.40%																																																																																																																																																							
0% Min	0.05%																																																																																																																																																							
Effect	DF	Wald																																																																																																																																																						
		Chi-Square	Pr > ChiSq																																																																																																																																																					
sex	1	402.90	<.0001																																																																																																																																																					
cl_maritalstatus	2	1189.26	<.0001																																																																																																																																																					
compResOfPHHPerm	3	403.22	<.0001																																																																																																																																																					
zoneRE	31	433.26	<.0001																																																																																																																																																					
nationalit_grob	3	672.42	<.0001																																																																																																																																																					
altklass_grob	3	443.83	<.0001																																																																																																																																																					
Permis	2	158.74	<.0001																																																																																																																																																					
cl_men	5	545.35	<.0001																																																																																																																																																					

\* Diese Personen haben nicht geantwortet

Alle eingegebenen Variablen werden bei der Stufenweise-Selektion als signifikante Variablen in das Modell aufgenommen. Aus der Abbildung 2 kann zusammenfassend gesagt werden, dass in beiden analysierten Jahren folgende Personengruppen eher bzw. leicht besser antworten als andere: Zweipersonenhaushalte, Schweizer, ältere Personen, Verheiratete, Frauen. Ihr Gewicht wird bei der Anpassung für die Antwortausfälle weniger nach oben angepasst.

**Abbildung 2** Logistische Regression: Mittelwerte der geschätzten Antwortwahrscheinlichkeiten 2014 und 2015 je Ausprägung der Variablen im Modell

	SE2014	SE2015		SE2014	SE2015
<b>cl_men</b>			<b>zoneRE</b>		
1	84.1%	83.3%	AG00	86.4%	85.6%
2	89.1%	88.8%	AI00	89.4%	86.0%
3	85.7%	85.0%	AR00	87.3%	87.1%
4	87.3%	86.8%	BE00	87.6%	87.3%
5	84.0%	83.1%	BE01	84.4%	83.8%
6	77.9%	76.1%	BE02	88.3%	
<b>Permis</b>			BL00	86.9%	86.7%
B_	75.8%	75.5%	BS00	82.2%	81.8%
CH	88.7%	88.1%	FR00	89.5%	88.9%
C_	81.2%	80.8%	GE00	84.4%	82.4%
<b>altklass_grob</b>			GE01	79.3%	78.4%
15	81.2%	80.3%	GL00	84.9%	83.5%
25	83.8%	82.9%	GR00	87.5%	87.3%
45	88.8%	88.2%	JU00	87.7%	88.3%
65	90.1%	89.9%	LU00	87.2%	86.5%
<b>nationalit_grob</b>			NE00	87.4%	86.7%
1000	88.7%	88.1%	NW00	90.0%	89.1%
2000	81.7%	81.1%	OW00	89.6%	88.1%
3000	72.7%	72.8%	SG00	86.4%	85.4%
4000	75.8%	74.6%	SH00	85.5%	85.8%
<b>compResOfPHHPerm</b>			SO00	86.2%	86.2%
1	87.4%	87.5%	SZ00	88.0%	87.1%
2	84.4%	82.6%	TG00	86.6%	85.9%
3	88.3%	87.7%	TI00	87.6%	86.9%
4	82.5%	80.9%	UR00	89.0%	88.2%
<b>cl_maritalstatus</b>			VD00	87.3%	86.9%
1	82.0%	81.0%	VDLA	83.5%	82.5%
2	89.5%	89.2%	VS00	86.9%	86.5%
3	85.7%	85.0%	ZG00	88.6%	87.4%
<b>sex</b>			ZH00	86.5%	86.5%
1	85.0%	84.4%	ZHWI	86.1%	85.8%
2	87.8%	87.1%	ZHZH	83.5%	83.0%

## 4.2 Kalibrierung

Die berechneten Initialgewichte  $w_{k,P}^{ini}$  wurden zwecks Übereinstimmung mit den Totalen der STATPOP und zur Reduktion der Varianz und eines möglichen Bias kalibriert. Auf Ebene Schweiz wurde auf die Totale (STATPOP per Ende Jahr) folgender Variablen kalibriert (x=Kreuzung):

**CHVar1** Geschlecht x Altersklassen fein

**CHVar2** Nationalitätengruppen fein

**CHVar3** Geschlecht x Altersklassen grob x Nationalitätengruppen grob

**CHVar4** Haushaltsgrösse

**CHVar5** Geschlecht x Altersklassen grob x Zivilstand<sup>4</sup>

**CHVar6** Geschlecht x Altersklassen grob x Ausländerkategorie

**CHVar7** ZoneRE (=erweiterte Zonen: Kantone, jedoch in BE: Berner Jura, Stadt Bern, Rest Bern. In ZH: Stadt Zürich, Winterthur, Rest ZH. In GE: Stadt Genf, Rest GE. In VD: Lausanne, Rest VD)

**CHVar8** Haushaltszusammensetzung (compResOfPHHPerm, 4 Gruppen) seit 2013

Sofern die Netto-Stichprobengrösse es erlaubt, d.h. wenn mindestens 100 Antworten pro Kreuzung vorliegen, wurden zusätzlich auf Niveau der Kantone und für die Grossstädte Bern, Genf, Lausanne, Winterthur und Zürich die folgenden Variablen kalibriert:

<sup>4</sup>Aufgrund geringer Anzahl Beobachtungen werden bei den 15-24-Jährigen die Zivilstandsklassen 2 und 3-7 zusammengefasst.

- KtVar1** Geschlecht
- KtVar2** Altersklassen grob
- KtVar3** Nationalität
- KtVar4** Ausländerkategorie
- KtVar5** Nationalitätengruppen grob
- KtVar6** Geschlecht x Altersklassen grob
- KtVar7** Geschlecht x Nationalität
- KtVar8** Geschlecht x Ausländerkategorie
- KtVar9** Geschlecht x Nationalitätengruppen grob
- KtVar10** Altersklassen grob x Nationalität

Dabei sind die Variablen wie folgt definiert:

- Geschlecht = 1 (Mann) / 2 (Frau)
- Altersklassen fein = 15-19 / 20-24 / 25-29 / ... / 80-84 / 85+
- Altersklassen grob = 15-24 / 25-44 / 45-64 / 65+
- Nationalitätengruppen fein = Schweiz / Deutschland / Frankreich / Italien / Österreich / Spanien / Portugal / Benelux / nordische Staaten / restliche EU und EFTA / Türkei / Kosovo / Serbien / Rest Europas / Rest der Welt
- Nationalitätengruppen grob = Schweiz / EU und EFTA / restliches Europa / Rest der Welt
- Nationalität = Schweiz / Ausland
- Ausländerkategorie = Schweiz / C / B und Rest
- Zivilstand = 1 (Ledig) / 2 (Verheiratet) / 3-7 (Rest)
- Haushaltsgrosse = 1 / 2 / 3 / 4 / 5 / 6 und mehr
- Haushaltszusammensetzung (1 bis 4, compResOfPHHPerm):

compResOf PHHPerm	compBasic OfPHHPerm	Beschrieb
1	101	Mann (erwachsen, alleinstehend, ohne Kinder)
	102	Frau (erwachsen, alleinstehend, ohne Kinder)
	110	Zwei erwachsene Personen untersch. Geschlechts, ohne Kinder
	201	Eine minderjährige Person
	202	Zwei oder mehrere minderjährige Personen
2	111	Zwei erwachsene Personen gleichen Geschlechts, ohne Kinder
	120	Drei oder mehr erwachsene Personen, ohne Kinder
3	211	Frau (erwachsen, alleinstehend, mit mind. 1 Kind)
	220	Zwei erwachsene Personen untersch. Geschlechts mit mind. 1 Kind
4	210	Mann (erwachsen, alleinstehend, mit mind. 1 Kind)
	221	Zwei erwachsene Personen gleichen Geschlechts mit mind. 1 Kind
	230	Drei oder mehr erwachsene Personen mit mind. 1 Kind

Die genaue Bedeutung der Codierung ist im Beschrieb (Synopsis) der Strukturhebung nachzulesen (siehe [BFS, 2016](#)).

Die Kalibrierung der Personengewichte mit dem SAS-Makro %calmar2 (siehe INSEE, 2005) wird wie folgt aufgerufen. Dabei heisst m=1, dass die «lineare Methode» verwendet wird.

```
%calmar2(DATAMEN =out_zp.se20yy_ech_pour_cal,
MARMEN =out_zp.se20yy_marges,
DATAPOI =out_zp.se20yy_Poids_sortie,
IDENT =vn,
Poids =w_ini_zp_rep,
Poidsfin=w_zp_cal,
Obseli =oui, colin = oui, misajour = non,
Editpoi =non, M=1, maxiter= 25);
```

Das kalibrierte Gewicht, welches für die Hochrechnung im Personendatensatz verwendet wird heisst w\_zp\_cal und ist in der Outputdatei se20yy\_poids\_sortie gespeichert. In der Tabelle 5 werden die Verteilungen der einzelnen Gewichte verglichen. Anhand des Quotienten der Gewichte nach und vor der Kalibrierung (w\_zp\_cal / w\_ini\_zp\_rep) ist der Einfluss dieser Kalibrierung ersichtlich.

**Tabelle 5** Verteilung der Personengewichte 2014 und 2015

Variable	Anzahl	Min.	p1	p5	p10	p25	Mittelwert	Median	p75	p90	p95	p99	Max.	Summe	Max./Min.
<b>2014</b>															
w_design	333'054	13.40	13.94	13.94	13.97	14.15	20.43	14.21	28.30	28.40	28.40	28.45	28.48	6'803'464	2.13
w_ini_zp_rep	287'842	14.02	14.91	15.12	15.23	15.87	23.64	17.96	31.67	33.82	35.31	40.40	62.85	6'805'478	4.48
w_zp_cal	287'842	10.94	13.84	14.63	14.99	15.74	23.73	18.61	31.51	34.53	36.64	42.78	79.87	6'829'610	7.30
w_zp_cal / w_ini_zp_rep	287'842	0.45	0.86	0.91	0.94	0.97	1.00	1.00	1.03	1.07	1.10	1.19	1.68		3.73
<b>2015</b>															
w_design	314'109	14.28	14.28	14.42	14.46	14.50	21.94	28.69	28.95	28.99	29.04	29.04	29.11	6'891'735	2.04
w_ini_zp_rep	269'509	15.04	15.34	15.53	15.79	16.64	25.58	30.73	32.91	35.27	36.79	42.19	68.10	6'893'001	4.53
w_zp_cal	269'509	13.00	14.79	15.21	15.64	16.66	25.63	29.43	32.81	35.80	38.06	44.42	81.85	6'907'818	6.29
w_zp_cal / w_ini_zp_rep	269'509	0.58	0.87	0.92	0.95	0.97	1.00	1.00	1.03	1.07	1.10	1.16	1.52		2.61

## 5 Gewichtung der Haushalte und der Zielpersonen der Haushalte

### 5.1 Initialgewichtung

Als Startpunkt der Gewichtung der Haushalte bzw. der Zielpersonen der Haushalte wird ab 2013 die mittels logistischer Regression für Antwortausfälle korrigierte Einschlusswahrscheinlichkeit der Personen  $\tilde{\pi}_{k,P}$  verwendet, welche im Abschnitt Personengewichtung 4.1 beschrieben wurde.

Wie in den Vorjahren gibt es auch für die neueren SE (ab 2013) einige Zielpersonen, die in der SE ihre Haushaltsmitglieder nicht so angegeben haben, wie es eigentlich gemäss STATPOP zu erwarten gewesen wäre. In der Regel werden in diesen Fällen in der SE weniger Personen angegeben, die unter der Annahme, dass in STATPOP alles richtig erfasst ist, Teilantwortausfälle wären.

Da eine Einsetzung zusätzlicher Haushaltsmitglieder relativ kompliziert wäre, wurde entschieden, die betroffenen Haushalte zu entfernen. Dafür müssen diese (Teil-)Antwortausfälle, welche als eine Ziehung einer Unterstichprobe von der eigentlichen Nettostichprobe betrachtet werden können, vor der Kalibrierung modelliert werden. Für die Modellierung gehen wir davon aus, dass die Angabe des in der SE deklarierten Haushalts identisch zum STATPOP-Haushalt zufällig und Bernoulli verteilt ist. Für die Schätzung der entsprechenden bedingten Wahrscheinlichkeiten wird wiederum eine logistische Regression berechnet. Für die Modellkonstruktion wird ein Teil der erklärenden Variablen wie für die Schätzung der Antwortwahrscheinlichkeit auf Niveau

Personen verwendet. Die erklärte Variable ist also die Wahrscheinlichkeit, dass eine Zielperson, gegeben sie hat bei der SE geantwortet, einen SE-Haushalt deklariert, der identisch zum STATPOP-Haushalt ist. Diese Anpassung kann somit nur für die Nettostichprobe der Zielpersonen gemacht werden.

Für SE2013 und 2014:

```
proc logistic data=predict_sk_zp outmodel=out_hh.se20yy_model_sel_hh_zp;
class compresofphhperm zonere permis;
model indic_hh_sel(event='1')= compresofphhperm zonere permis
/rsquare selection=stepwise lackfit;quit;
```

Ab der Strukturhebung 2015 wurde die logistische Regression erweitert. Es wurde beobachtet, dass in fünf mittleren bis grösseren Gemeinden und zusätzlich bei einigen Gemeinden im Kanton Genf, die obige Modellierung nur bedingt zufriedenstellende Schätzungen ergab. Mit den resultierenden Haushaltsgewichten wurde eine Unterschätzung der Anzahl Haushalte im Vergleich zur STATPOP festgestellt. Daher wurde eine weitere erklärende Variable (*fehlhh*) eingeführt, welche für diese Gemeinden gekreuzt mit bestimmten Haushaltszusammensetzungen (*compbasicofphhperm*) eine separate Schätzung der Antwortwahrscheinlichkeiten zulässt. Diese Variable, *fehlhh*, ist in Tabelle 6 definiert.

**Tabelle 6** Definition Variablen *fehlhh* nach Gemeinde und HH-Zusammensetzung.

Gemeinde	HH-Zusammens. Typ «speziell»	HH-Zusammens. Typ «normal»
Bern	351	3510001
Biel	371	3710001
Monthey	6153	61530001
Genf	6621	66210001
Delemont	6711	67110001
Gemeinden Kanton Genf (ohne Stadt)	9999	000
übrige Gemeinden	000	000
Typ «speziell»: falls <i>compbasicofphhperm</i> in ('111','120','210','221','230')		

Die logistische Regression wurde somit wie folgt angepasst.

Für SE2015:

```
proc logistic data=predict_sk_zp outmodel=out_hh.se20yy_model_sel_hh_zp nocov;
class compresofphhperm zonere permis fehlhh;
model indic_hh_sel(event='1')=compresofphhperm zonere permis fehlhh
/rsquare selection=stepwise lackfit;
quit;
Outputdatei: out_hh.se20yy_model_sel_hh_zp;
```

Anwenden und Prädiktion:

```
proc logistic inmodel=out_hh.se20yy_model_sel_hh_zp(type=logismod);
score clm data = predict_sk_zp
out =out_hh.se20yy_predict_sel_hh_zp(rename=(P_1=p_sel_hh_zp)) ;
run;
```

Outputdatei: *out\_hh.se20yy\_predict\_sel\_hh\_zp*

Die Abbildungen 3 und 4 zeigen die Outputs der logistischen Regressionen

**Abbildung 3** Logistische Regression: Output 2014 und 2015

Analyse Stepwise Selection  
SE 2015

Summary of Stepwise Selection					
Step	Effect	DF	Number	Score	
			In	Chi-Square	Pr > ChiSq
1	compResOfPHHPerm	3	1	31126.4676	<.0001
2	zoneRE	30	2	4299.2373	<.0001
3	Permis	2	3	1741.6882	<.0001
4	fehlhh	9	4	556.2569	<.0001

removed: keine

Verteilung geschätzte Wahrscheinlichkeiten Nicht-Ausschluss HH

Quantiles (Definition 5)	
Level	Quantile
100% Max	96.66%
99%	96.56%
95%	96.36%
90%	95.78%
75% Q3	94.94%
50% Median	93.42%
25% Q1	76.22%
10%	68.80%
5%	54.99%
1%	37.92%
0% Min	19.85%

Analyse der Effekte

Type 3 Analysis of Effects			
Effect	DF	Wald	
		Chi-Square	Pr > ChiSq
compResOfPHHPerm	3	19636.52	<.0001
Permis	2	1652.59	<.0001
fehlhh	11	1220.30	<.0001
zoneRE	28	747.07	<.0001

SE 2014

Summary of Stepwise Selection					
Step	Effect	DF	Number	Score	
			In	Chi-Square	Pr > ChiSq
1	compResOfPHHPerm	3	1	40219.9743	<.0001
2	zoneRE	31	2	3922.1236	<.0001
3	Permis	2	3	1426.5774	<.0001

removed: keine

Quantiles (Definition 5)	
Level	Quantile
100% Max	96.89%
99%	96.50%
95%	95.98%
90%	95.66%
75% Q3	94.70%
50% Median	93.67%
25% Q1	70.54%
10%	65.53%
5%	52.63%
1%	42.26%
0% Min	24.95%

Type 3 Analysis of Effects			
Effect	DF	Wald	
		Chi-Square	Pr > ChiSq
compResOfPHHPerm	3	32159.60	<.0001
Permis	2	1393.69	<.0001
zoneRE	31	3511.19	<.0001

**Abbildung 4** Logistische Regression: Mittelwerte der geschätzten Wahrscheinlichkeiten 2014 und 2015 je Ausprägung der Variablen im Modell

	SE2014	SE2015
<b>Permis</b>		
B_	76.9%	77.3%
CH	85.5%	87.2%
C_	82.9%	84.1%
<b>compBasicOfPHHPerm</b>		
1	93.1%	93.3%
2	64.4%	69.4%
3	94.7%	95.0%
4	64.6%	66.1%
<b>fehlhh</b>		
000		87.7%
351		35.0%
371		38.6%
6153		50.5%
6621		45.0%
6711		42.2%
9999		50.3%
3510001		87.8%
3710001		87.5%
61530001		90.0%
66210001		87.9%
67110001		94.4%
<b>zoneRE</b>		
AG00	88.3%	89.6%
AI00	81.2%	83.5%
AR00	82.3%	84.3%
BE00	87.3%	89.0%
BE01	69.9%	70.1%
BE02	86.9%	
BL00	86.9%	88.7%
BS00	85.7%	85.8%
FR00	84.7%	86.6%
GE00	71.0%	72.1%
GEGE	71.9%	72.2%
GL00	89.6%	87.8%
GR00	82.0%	84.2%
JU00	80.9%	84.8%
LU00	87.0%	88.3%
NE00	88.2%	89.5%
NW00	86.6%	85.9%
OW00	81.7%	84.9%
SG00	85.2%	88.8%
SH00	87.5%	88.0%
SO00	87.0%	87.3%
SZ00	85.6%	88.3%
TG00	86.5%	88.4%
TI00	85.6%	87.7%
UR00	85.3%	87.1%
VD00	84.1%	85.0%
VDLA	85.3%	86.6%
VS00	80.6%	83.6%
ZG00	87.4%	89.2%
ZH00	87.0%	88.8%
ZHWI	88.4%	88.8%
ZHZH	82.5%	84.8%

Die Tabelle 7 enthält den Beschrieb bzw. die Notationen der verwendeten Gewichte.

**Tabelle 7** Berechnung der Initialgewichtung der Haushalte

Notation	SAS	Bedeutung
$\tilde{\pi}_{k,P} = \tilde{\pi}_P r_{k,P}$	pikpk_p	für Antwortausfälle korrigierte Ziehungswahrscheinlichkeit.
$r_{k,P}^{sel}$	p_sel_hh_zp	Wahrscheinlichkeit, dass der Haushalt der Zielperson $k$ beibehalten wird.
$\tilde{\pi}_{k,P}^{sel} = \tilde{\pi}_P r_{k,P} r_{k,P}^{sel}$	pikpksk_zp	für Antwortausfälle und Selektion Haushalte korrigierte Einschlusswahrscheinlichkeit Zielperson.
$\tilde{\pi}_{k,H}^{sel} = \min\left(1, \sum_{P \in H} \tilde{\pi}_{k,P}^{sel}\right)$	pikpksk_hh	für Antwortausfälle und Selektion Haushalte korrigierte Einschlusswahrscheinlichkeit Haushalt.
$w_{k,P}^{ini} = [\tilde{\pi}_{k,P}^{sel}]^{-1}$	w_zp_ini_sel	für Antwortausfälle und Selektion korrigiertes Initialgewicht Personen.
$w_{k,H}^{ini} = [\tilde{\pi}_{k,H}^{sel}]^{-1}$	w_hh_ini_sel	für Antwortausfälle und Selektion korrigiertes Initialgewicht Haushalt (Input %calmar2).
$w_{k,ZP}^{ini} = w_{k,P}^{ini} / w_{k,H}^{ini}$	pdk (poidskish)	Gewicht Kish-Zielpersonen. 1 durch die Wahrscheinlichkeit, dass die Zielperson jene ist, welche beobachtet wurde, gegeben der Haushalt ist gezogen (Input %calmar2).

### Bemerkungen:

- Es gibt einige sehr grosse Haushalte, vor allem in Genf, die eine Einschlusswahrscheinlichkeit von 1 haben.
- Einige Haushalte der STATPOP haben eine Einschlusswahrscheinlichkeit von 0. Für jene in der STATPOP per Ende 2014 waren es ungefähr 53'800, welche für die SE2014 nicht gezogen werden konnten. Dies kommt einerseits von Personen, die zwischen dem Datum des Ziehungsrahmens (September) und Dezember in die Schweiz gezogen sind (8800 Haushalte). Andererseits gab es für die Ziehung gesperrte Fälle (indic\_protection=1), da sie in einer alten Erhebung dabei waren (2014: 25'500), oder solche, die im Ziehungsrahmen noch nicht als Privathaushalt definiert waren (2014: 19'000). Zuletzt gibt es noch wenige Privathaushalte, die keine Personen in der Zielpopulation enthielten und nicht gezogen werden konnten.

## 5.2 Kalibrierung

Es folgt die Kalibrierung der Haushaltsgewichte der selektierten Haushalte, die Kalibrierung der Haushaltspersonengewichte und jene der (Kish-ZP) Zielpersonen-Gewichte (im reduzierten Datensatz) auf die Totale gemäss STATPOP per Ende Jahr. Seit der SE2014 wird im SAS-Makro %calmar2 die Funktion benutzt, welche es erlaubt auf zwei Niveaus (Personen und Haushalte) drei Gewichte gleichzeitig/simultan zu bestimmen bzw. zu kalibrieren (siehe [INSEE, 2005](#)). Diese drei kalibrierten Gewichte werden berechnet:

**hh\_gew** Haushaltsgewicht

**hhpers\_gew** Gewicht der Haushaltspersonen

**zp\_gew\_kish** Gewicht der Zielpersonen (Kish)

Folgende Eigenschaften sollten gelten:

- hh\_gew= hhpers\_gew
- zp\_gew\_kish= hh\_gew x Faktor

Der Faktor ist nicht die Anzahl Zielpersonen im Haushalt, sondern das Inverse der Einschliesswahrscheinlichkeit die Kish-Zielperson zu beobachten, gegeben der Haushalt hat geantwortet. Einige Zielpersonen im Haushalt haben eine höhere Wahrscheinlichkeit zu antworten als andere, erhalten daher ein kleineres Gewicht. Es sollen folgende Kalibrierungsgrenzen berücksichtigt werden:

**Haushalte:** Typische HH-Merkmale wie HH-Grösse und HH-Zusammensetzung

**Kish-Zielpersonen:** Alle Grenzen, welche bei der «normalen» Zielpersonengewichtung (Gewichtung für die Zielpersonen-Nettostichprobe) verwendet wurden.

**HH-Personen:** Auch Personengrenzen, jedoch mit zusätzlichen Altersklassen für Kinder.

Die Anzahl Modalitäten sind in %ca1mar2 auf 1000 beschränkt. Wenn man alle wünschenswerten Grenzen eingeben möchte, hätte man ca. 2800 Modalitäten. Ein weiteres Problem war auch, dass es in STATPOP Haushalte gibt, die nur aus Kindern jünger als 15-jährig bestehen, also keine Zielpersonen enthalten. Diese Haushalte können also gar nicht an der SE teilnehmen, daher wurde ihr HH-Typ bei der Kalibrierung zu einem anderen geschlagen (zu compResOfPHHPem=1), damit das Total der Haushaltsgewichte trotzdem der ganzen STATPOP-Haushaltspopulation entspricht.

Es wurden mehrere Ansätze zur Gewichtung versucht, wobei am Schluss folgende Lösung beibehalten wurde. Eine Kalibrierung in zwei Schritten, wobei die Anzahl Grenzen stark reduziert bzw. die Klassen aggregiert wurden.

- 1. Schritt: Integrierte Kalibrierung (gemäss [INSEE, 2005](#)) der Haushalte, der Haushaltspersonen und der Kish-Zielpersonen mit einer reduzierten Anzahl an definierten Grenzen (siehe Tabelle 8). Das ergibt das definitive Gewicht für die Haushalte und Haushaltspersonen, sowie ein provisorisches Gewicht für die Kish-Zielpersonen. Das Gewicht für die Zielpersonen ist provisorisch, da noch nicht alle gewünschten Grenzen ins %ca1mar2 eingeben werden konnten.
- 2. Schritt: Das Resultat der provisorischen Gewichtung der Kish-Zielpersonen aus Schritt 1 wird übernommen und auf alle Personenvariablen kalibriert. Dies ergibt das definitive Gewicht der Kish-Zielpersonen (zp\_gew\_kish).

Diese Prozedur in zwei Schritten erlaubt es nicht, die zweite Eigenschaft oben (zp\_gew\_kish=hh\_gew x Faktor) zu gewährleisten. Dies ist unter den erwähnten Voraussetzungen nicht zu verhindern.

Im ersten Schritt wurden die Variablen in der Tabelle 8 als Grenzen benutzt/definiert.

Dabei sind die Variablen wie folgt definiert:

zoneRE	Schicht (Kantone, erhöhende Regionen und grosse Städte separat)
cl_men=min(taillemen_neu,6)	Haushaltsgrösse (auf max 6 Pers. limitiert)
cl_2=min(cl_men, 4)	Haushaltsgrösse (auf max 4 Pers. limitiert)
zn =substr(zoneRE,1,2)_"00"	zoneRE (aber Halbkantone zusammen)
zn2	zn (plus GL und UR zusammen)

**Tabelle 8** Kalibrierungsvariablen der integrierten Kalibrierung

Niveau	Grenzdatei %calmar2	Kalibrierungs- variablen	Beschrieb
HH-Personen	MARIND=  marges_hhpers_sel	KTVAR1 KTVAR10 KTVAR2 KTVAR3 KTVAR6 KTVAR7	zoneRE X sex zoneRE X altklass_grob X Suisse zoneRE X altklass_grob zoneRE X Suisse zoneRE X sex X altklass_grob zoneRE X sex X Suisse
Kish-Zielpersonen	MARKISH=  marges_hhzpers_sel	CHVAR1 CHVAR2 CHVAR8 KTVAR1	sex X altklass_fein nationalit_fein compResOfPHHPerm zoneRE X sex
Haushalte	MARMEN=  marges_hh_sel	hh_RGVAR0 hh_RGVAR1 hh_RGVAR2 hh_RGVAR3 ab 2015 (statt 0-2) hh_RGVAR4	zoneRE min(cl_men,5) zn X cl_2 zn2 X compresofphperm zoneRE X min(cl_men,5)

Untenstehend folgt der Aufruf von %calmar2 wie er für die Erhebung 2015 definiert wurde. Es wurde also die begrenzte lineare Distanzfunktion (M=4) gewählt, d.h. der Quotient «Gewicht nach Kalibrierung» dividiert durch «Gewicht vor Kalibrierung» wird durch die beiden Grenzen (LO/UP) beschränkt, um eine zu starke Veränderung der Gewichte zu verhindern.

Erster Aufruf %calmar2:

```
%calmar2(DATAMEN =out_hh.se20yy_sel_hh_echprcal,
DATAIND =out_hh.se20yy_sel_hhpers_echprcal,
DATAKISH =out_hh.se20yy_sel_hhzpers_echprcal,
MARMEN =out_hh.se20yy_marges_hh_sel,
MARIND =out_hh.se20yy_marges_hhpers_sel,
MARKISH =out_hh.se20yy_marges_hhzpers_sel,
POIDS =w_hh_ini_sel,
POIDKISH =pdk,
IDENT =householdid,
IDENT2 =vn,
M=4,L0=0.3,UP=12,
MAXITER =99, COLIN=OUI, SEUIL=0.000001,
DATAPOI =out_hh.se20yy_Poids_sortie,
DATAPOI2 =out_hh.se20yy_Poids_hhpers_sortie,
DATAPOI3 =out_hh.se20yy_Poids_hhzpers_sortie,
Poidsfin =w_hh_cal,
Poidskishfin=w_hh_hzp_cal,
MISAJOUR =NON, OBSELI=OUI, EDITPOI=NON);
```

Zweiter Aufruf %calmar2:

Kalibrierung der Zielpersonen auf zusätzliche Variablen, wobei als Inputgewicht jenes aus der integrierten Kalibrierung aus Schritt 1 genommen wird. Zudem wird hier die Methode M=1 (einfache lineare Distanzfunktion ohne Beschränkung), also dieselbe wie bei der Personen-Kalibrierung gewählt.

```
%calmar2(DATAMEN=out_hh.se20yy_sel_zpers_echprcal,
MARMEN =out_hh.se20yy_marges_zpers_sel,
IDENT =vn,
POIDS =w_hh_hzp_cal,
POIDSFIN=w_hh_zp_cal,
M=1,
COLIN =OUI, MISAJOUR=NON,
DATAPOI =out_hh.se20yy_Poids_zpers_sortie);
```

Folgende in der Tabelle 9 aufgeführten SAS-Output-Dateien mit den kalibrierten Gewichten und anderen Variablen werden jedes Jahr erstellt.

**Tabelle 9** Strukturhebung: Berechnete Gewichte und Variablen

SAS-Datei	Gewichtvariable	andere Variablen	
<b>Personen</b>			
se20yy_poids_sortie	w_zp_cal	Zielperson	srph_oid, vn, sas_id, zone, CHVAR1-8, KTVAR1-10.
<b>Haushalte</b>			
se20yy_poids_hh	w_hh_cal	Haushalt	householdid, sas_id, srph_oid, householdyearlyid, (ab SE2015: zone_hh)
	w_hh_zp_cal	Zielperson	
<b>Haushaltspersonen</b>			
se20yy_poids_hhpers	w_hhpers_cal	Haushalt und Haushaltspersonen	vn householdid, SRPH_OID, sas_id, hhpersonnumber, hhnavs13, hhpersyearlyid
Speicherorte (interner Gebrauch) mit yy=Jahr			
Personen: /project/prd/shape/se/20yy/micro/gew/meth_pond/personen/			
Haushalte: /project/prd/shape/se/20yy/micro/gew/meth_pond/haushalte/			

In der Tabelle 10 werden die Verteilungen der einzelnen Gewichte verglichen. Anhand der Quotienten der Gewichte nach und vor der Kalibrierung ( $w\_hh\_cal / w\_hh\_ini\_sel$  und  $w\_hh\_zp\_cal / w\_hh\_hzp\_cal$ ) ist der Einfluss der Kalibrierungen ersichtlich.

**Tabelle 10** Verteilung der Haushaltsgewichte 2014 und 2015

Variable	Anzahl	Min.	p1	p5	p10	p25	Mittelwert	Median	p75	p90	p95	p99	Max.	Summe	Max./Min.
<b>2014</b>															
w_hh_ini_sel	243'236	2.98	4.66	5.68	7.54	8.19	14.74	15.21	17.18	26.39	35.01	38.89	92.41	3'585'616	30.99
pdk	243'236	1.00	1.00	1.00	1.00	1.94	2.22	2.01	2.86	3.84	4.21	5.28	14.27		14.27
w_hh_cal	243'236	1.27	3.74	5.22	6.72	8.22	14.70	13.83	17.41	28.27	34.37	41.28	139.40	3'576'648	110.17
w_hh_cal / w_hh_ini_sel	243'236	0.30	0.53	0.75	0.83	0.92	1.00	0.99	1.06	1.17	1.26	1.61	9.11		30.37
w_hh_hzp_cal	243'236	4.52	11.90	14.85	15.62	16.95	28.08	25.87	34.51	42.52	51.99	76.31	404.11	6'829'610	89.37
w_hh_zp_cal	243'236	2.75	11.64	14.50	15.27	16.99	28.08	25.91	34.71	42.66	52.26	77.52	437.68	6'829'610	159.43
w_hh_zp_cal / w_hh_hzp_cal	243'236	0.18	0.84	0.91	0.94	0.96	1.00	0.99	1.03	1.08	1.12	1.19	1.80		9.74
w_hhpers_cal	636'655	1.27	3.39	4.74	5.59	7.86	12.63	10.28	16.33	19.72	26.36	38.11	139.40	8'041'310	110.17
<b>2015</b>															
w_hh_ini_sel	231'618	2.76	4.65	5.81	7.79	8.65	15.73	16.26	17.67	34.59	36.21	41.63	106.81	3'644'493	38.68
pdk	231'618	1.00	1.00	1.00	1.00	1.94	2.23	2.01	2.87	3.85	4.22	5.30	11.03		11.03
w_hh_cal	231'618	1.39	4.02	5.44	7.17	8.72	15.64	15.37	18.07	31.91	35.37	41.50	215.27	3'623'266	155.06
w_hh_cal / w_hh_ini_sel	231'618	0.30	0.57	0.78	0.85	0.93	1.00	0.99	1.06	1.14	1.23	1.53	9.18		30.61
w_hh_hzp_cal	231'618	4.84	13.32	15.73	16.42	18.19	29.82	31.06	35.83	43.69	51.43	77.31	479.50	6'907'818	99.12
w_hh_zp_cal	231'618	4.35	13.24	15.37	16.23	18.30	29.82	30.30	36.09	43.91	52.06	77.50	545.89	6'907'818	125.49
w_hh_zp_cal / w_hh_hzp_cal	231'618	0.47	0.84	0.91	0.94	0.96	1.00	0.99	1.03	1.08	1.11	1.19	1.75		3.70
w_hhpers_cal	607'645	1.39	3.67	4.88	5.92	8.25	13.38	11.64	16.77	20.36	29.14	38.87	215.27	8'131'033	155.06

## 6 Schätzungen

### 6.1 Punktschätzungen

#### 6.1.1 Die Schätzung von Totalen, Anzahlen, Quotienten

Wir gehen davon aus, dass das Total  $Y$  einer Variablen  $y$  auf Ebene Schweiz geschätzt wird. Der entsprechende Schätzer ist

$$\hat{Y} = \sum_{k \in r} w_k y_k, \quad (1)$$

wobei  $r$  die Menge der antwortenden Personen (oder Haushalte)  $k$  und  $w_k$  das kalibrierte Hochrechnungsgewicht bezeichnet.

Diese Formel ist für stetige wie auch für kategorielle Zielvariablen gültig. Für letztere werden die verschiedenen Ausprägungen mittels Indikatorvariablen beschrieben, welche dann die Rolle der Zielvariablen  $y_k$  einnehmen. Im einem fiktiven Beispiel, wo wir uns für die Anzahl Schweizer/innen ohne Migrationshintergrund und Ausbildung Sekundarstufe 2 interessieren, wäre also  $y_k = I(k \in K_{Bsp})$ . Dabei bezeichnet  $K_{Bsp}$  das Merkmal (die Kategorie) Schweizer/innen ohne Migrationshintergrund und Ausbildung Sek2 und  $I(\cdot)$  die Indikatorfunktion, welche den Wert 1 annimmt, falls die Person dieses Merkmal besitzt und 0 sonst.

Die Schätzung der Anzahl Personen mit diesem Merkmal lässt sich wie folgt schreiben:

$$\hat{N}_{K_{Bsp}} = \sum_{k \in r} w_k I(k \in K_{Bsp}) \quad (2)$$

Der Quotient  $R = \frac{Y}{Q}$  zwischen dem Total der Variablen  $y$  und  $q$  kann wie folgt geschätzt werden

$$\hat{R} = \frac{\hat{Y}}{\hat{Q}} = \frac{\sum_{k \in r} w_k y_k}{\sum_{k \in r} w_k q_k}. \quad (3)$$

Dabei sind  $\hat{Y}$  und  $\hat{Q}$  gewöhnliche Totalschätzungen im Sinne von (1).

Bemerkungen:

- Mittelwerte können als Spezialfall eines Quotienten behandelt werden. Dabei ist die Variable  $y$  im Zähler numerisch und die Variable  $q$  im Nenner eine Indikatorvariable für den interessierenden Untersuchungsbereich.
- Obige Formeln sind sowohl für Personen wie auch für Haushalte gültig.
- Falls nicht das Total auf Ebene Schweiz, sondern ein Total für eine Subpopulation (oder Untersuchungsbereich)  $d$  ( $d$  für Domain) geschätzt werden soll, wird in sämtlichen Berechnungen statt der Variablen  $y_k$  die Variable  $y_{d,k} := y_k I(k \in d)$  verwendet. Beispielsweise wäre die Bevölkerung eines gewissen Gebietes eine solche Subpopulation.

### 6.2 Varianzschätzung und Konfidenzintervall

#### 6.2.1 Einleitung

Varianzen und Konfidenzintervalle von geschätzten Häufigkeiten werden im BFS mit der Software SAS und mit der Prozedur SURVEYFREQ (für Personen und Haushalte) wie auch mit dem

speziell für die SE entwickelten SAS-Makro %Ca1VarSE (nur für Personen programmiert) simultan zu den geschätzten Werten berechnet.

Die Varianzschätzungen leiden unter einem kleinen Bias, denn die Tatsache, dass innerhalb derselben Strukturerhebung nicht mehrere Personen innerhalb des gleichen Haushalts befragt werden können, verhindert eine Varianzschätzung ohne Bias. Es ist auch schwierig eine verwendbare approximative Varianzschätzung zu beschreiben, welche alle Aspekte der Erhebung inklusive der Gewichtung ganz genau berücksichtigt.

Die Varianzschätzung erfolgt bedingt auf die Nettostichprobengrösse in den verschiedenen Zonen  $h$ . Bei den Personen sind die Zonen die Kantone (bzw. grosse Städte, Variable: zone), bei den Haushalten die Kantone/Städte x Haushaltsgrösse (Variable: zone\_hh).

Das heisst, die Zonen werden in der Varianzschätzung als Schichten berücksichtigt und die Menge der antwortenden Personen (Haushalte) werden als Zufallsstichproben innerhalb dieser Schichten behandelt. Streng genommen handelt es sich hierbei um eine etwas vereinfachende Annahme, welche der Realität aber nahekommt. Aus der Quadratwurzel der Varianz ergibt sich die Standardabweichung der Schätzung. Die Standardabweichung bildete die Grundlage zur Ermittlung der Vertrauensintervalle mit einem 95%-Niveau. Dabei verwendet SURVEYFREQ<sup>5</sup>: Schätzwert  $\pm t(97.5, df) \times$  Standardabweichung und %Ca1VarSE die Normalverteilungsapproximation: Schätzwert  $\pm 1.96 \times$  Standardabweichung.

SURVEYFREQ liefert für die geschätzten Häufigkeiten (und Anteile) eine vereinfachte Approximation der Varianzschätzung mittels %Ca1VarSE. Sowohl in SURVEYFREQ wie auch in %Ca1VarSE spielen Population- und Stichprobengrössen in den betrachteten Zonen (Schichten) eine wichtige Rolle. In den folgenden Abschnitten werden diese Methoden erläutert.

## 6.2.2 Schätzung mit SURVEYFREQ (für Personen und Haushalte)

Die SAS Prozedur SURVEYFREQ berücksichtigt den Effekt der Kalibrierung im Rahmen der Varianzschätzungen nicht. Eine Analyse mit SURVEYFREQ ist daher konservativ und geht zugleich im SAS relativ schnell, daher wird diese Schätzung mehrheitlich verwendet.

In der Strukturerhebung wird häufig eine Schätzung entsprechend (2) berechnet. Die Varianz von  $\hat{N}_{K_{Bsp}}$  wird dabei folgendermassen geschätzt:

$$\widehat{\text{Var}}\left(\hat{N}_{K_{Bsp}}\right) = \sum_h \frac{m_h}{m_h - 1} \left(1 - \frac{m_h}{N_h}\right) \sum_{k \in r_h} \left[ w_k I(k \in K_{Bsp}) - \frac{\hat{N}_{K_{Bsp},h}}{m_h} \right]^2. \quad (4)$$

Dabei bezeichnet

- $\hat{N}_{K_{Bsp},h}$  analog zu  $\hat{N}_{K_{Bsp}}$  die geschätzte Anzahl Schweizer/innen ohne Migrationshintergrund mit Ausbildung Sekundarstufe 2, aber in Schicht/Zone  $h$
- $N_h$  das Populationstotal in Schicht/Zone  $h$ .
- $r_h$  die Menge der antwortenden Personen in Schicht resp. Zone  $h$  und  $m_h$  die Anzahl Personen in  $r_h$  (Anzahl vorhandene Antworten in  $h$ ).

<sup>5</sup>t(97.5,df) bezeichnet das 97.5% Quantil der t-Verteilung mit df=(Anzahl Beobachtungen - Anzahl Schichten) Freiheitsgraden, welches von der Prozedur SURVEYFREQ automatisch berechnet wird. Vorausgesetzt, dass die Stichprobengrösse nicht zu klein ist, führt die Verwendung des entsprechenden Quantils der Normalverteilung (1.96) zu sehr ähnlichen Resultaten.

### 6.2.3 Schätzung mit %CalVarSE

Das Hochrechnungsmakro %CalVarSE berücksichtigt den Effekt der Kalibrierung im Rahmen der Varianzschätzungen (gemäss [Deville und Särndal, 1992](#)). Im Folgenden werden diese Schätzungen explizit beschrieben. Dabei werden zwei Fälle unterschieden: die Schätzung eines Totals und die Schätzung eines Quotienten. Zudem wird beschrieben, wie verschiedene Situationen, beispielsweise Schätzungen für Untersuchungsbereiche (Domains), auf diese beiden Fälle zurückgeführt werden können.

#### Varianzschätzung Total

Die Varianzschätzung erfolgt stufenweise:

1. Als Erstes betrachten wir das Regressionsmodell

$$y_k = \sum_{j=1}^p \beta_j x_{j,k} + \epsilon_{y,k}. \quad (5)$$

Dabei ist  $y_k$  der Wert der Zielvariablen für Individuum  $k$  und  $x_{j,k}$  bezeichnen die Werte, der in der Kalibrierung verwendeten Hilfsvariablen für Individuum  $k$ . Im Grunde handelt es sich bei den im Abschnitt Kalibrierung umschriebenen Kalibrierungsvariablen allesamt um kategorielle Variablen. Dabei gibt es zwei Fälle: Variablen,

- die auf Niveau CH kalibriert wurden und solche,
- die auf Niveau Region (Zonen plus Städte Bern, Genf, Lausanne, Winterthur, Zürich) kalibriert wurden.

In den Datensätzen der SE sind die entsprechenden Variablen mit

- CHVar1 - CHVar8 (Kalibrierung auf Niveau CH)
- KtVar1 - KtVar10 (Kalibrierung auf Niveau Region)

bezeichnet. Aufgrund ihrer qualitativen Natur können diese Variablen nicht direkt ins Modell (5) übernommen werden, sondern sie werden zuerst in 1/0 Variablen umcodiert. Innerhalb %CalVarSE wird diese Umcodierung durch die Prozedur GLM automatisch vorgenommen. Dabei wird für jede Ausprägung der Variablen eine separate 1/0 Variable kreiert. Es liegt auf der Hand, dass aus diesem Vorgehen eine grosse Anzahl Indikatorvariablen resultiert. Sie werden mit  $x_1, \dots, x_p$  bezeichnet und in (5) als erklärende Variablen verwendet. Die Regressionsparameter sowie die Residuen in (5) werden basierend auf der Nettostichprobe  $r$  mittels der Methode der gewichteten kleinsten Quadrate (WLS) geschätzt. Für die Gewichtung wird dabei das Hochrechnungsgewicht  $w_k$  (oder optional das für Antwortausfälle korrigierte Initialgewicht) verwendet. Dies führt zu

$$y_k = \sum_{j=1}^p \hat{\beta}_j x_{j,k} + \hat{\epsilon}_{y,k}, k \in r. \quad (6)$$

2. Für jedes Individuum  $k \in r$  hat man aufgrund von (6) ein geschätztes Residuum  $\hat{\epsilon}_{y,k}$ , welches durch die verwendeten Hilfsvariablen nicht erklärt werden kann. Dieser unerklärte Teil wird nun für die Varianzschätzung verwendet:

$$\widehat{\text{Var}}(\hat{Y}) = \sum_h \frac{m_h}{m_h - 1} \left(1 - \frac{m_h}{N_h}\right) \sum_{k \in r_h} \left( w_k \hat{\epsilon}_{y,k} - \frac{\sum_{k \in r_h} w_k \hat{\epsilon}_{y,k}}{m_h} \right)^2. \quad (7)$$

Dabei bezeichnen

- $N_h$  das Total Personen in Zone  $h$
- $r_h, m_h$  die Menge, respektive die Anzahl antwortender Personen in Zone  $h$ .

### Varianzschätzung Quotient

Die Schritte, welche uns zur Varianzschätzung führen, werden analog zum Fall des Totales aufgelistet.

1. Anhand der Variablen  $y_k$  und  $q_k$  sowie den Schätzungen von  $\hat{Q}$  und  $\hat{R}$  gemäss (3) berechnen wir für jede Person  $k \in r$  die linearisierte Variable

$$\hat{z}_k = \frac{1}{\hat{Q}}(y_k - \hat{R}q_k). \quad (8)$$

2. Nun wird analog (6) ein Regressionsmodell für die Zielvariable  $\hat{z}_k$  geschätzt:

$$\hat{z}_k = \sum_{j=1}^p \hat{\beta}_j x_{j,k} + \hat{\epsilon}_{z,k}, k \in r. \quad (9)$$

Hieraus resultieren die Residuen  $\hat{\epsilon}_{z,k}$ , welche im nächsten Schritt für die Varianzschätzung verwendet werden.

3. Varianzschätzung, analog zu (7)

$$\widehat{\text{Var}}(\hat{R}) = \sum_h \frac{m_h}{m_h - 1} \left(1 - \frac{m_h}{N_h}\right) \sum_{k \in r_h} \left( w_k \hat{\epsilon}_{z,k} - \frac{\sum_{k \in r_h} w_k \hat{\epsilon}_{z,k}}{m_h} \right)^2. \quad (10)$$

Bemerkungen:

- Sowohl die Punktschätzung (3) wie auch die Varianzschätzung (10) sind für stetige wie für kategorielle Variablen gültig. Für kategorielle Variablen müssen die verschiedenen Ausprägungen mittels Indikatorvariablen beschrieben werden, welche dann die Rollen der Zielvariablen  $y$  und  $q$  einnehmen und man spricht in diesem Fall eher von Anteilen statt von Quotienten. Wenn wir uns beispielsweise für den Anteil Personen mit Ausbildung Sek2 bei Schweizer/innen ohne Migrationshintergrund (*CHoMig*) interessieren, so betrachten wir die Variablen  $y_k = I(k \in \text{Sek2 und } k \in \text{CHoMig})$  im Zähler und  $q_k = I(k \in \text{CHoMig})$  im Nenner.
- Falls nicht ein Quotient auf Ebene Schweiz, sondern ein Quotient für eine Subpopulation  $d$  ( $d$  für Domain) geschätzt werden soll, werden in sämtlichen Berechnungen statt der Variablen  $y_k$  und  $q_k$  die Variablen  $y_{d,k} := y_k I(k \in d)$  und  $q_{d,k} := q_k I(k \in d)$  verwendet. Dabei ist  $I(k \in d)$  die Indikatorfunktion, welche den Wert 1 annimmt, falls Person  $k$  in  $d$  ist. Andernfalls hat die Funktion den Wert Null. Je Untersuchungsbereich werden also zwei spezifische Variablen erstellt, welche auf der ganzen Population definiert sind und die dann ganz normal auf das Niveau Schweiz hochgerechnet werden können.

**Bemerkung zu Merkmalsausfällen:** Falls eine Person gewisse Fragen unbeantwortet lässt, resultieren Merkmalsausfälle. Dabei ist zwischen kategoriellen und quantitativen Variablen zu unterscheiden:

- Für kategorielle Variablen wurden Merkmalsausfälle in speziellen Kategorien zusammengefasst und stellen somit für die Berechnungen kein spezielles Problem dar.
- Für quantitative Variablen geht %CalVarSE wie folgt vor. Beobachtungen mit Merkmalsausfall in der betreffenden Variablen werden vom interessierenden Untersuchungsbereich ausgeschlossen und in der Hochrechnung nicht berücksichtigt.
  - Dies führt im Falle einer Totalschätzung zu einer tendenziellen Unterschätzung.
  - Falls Mittelwerte oder Quotienten geschätzt werden, ist der Einfluss der Merkmalsausfälle weniger klar.

Der Anteil von Merkmalsausfällen ist generell sehr klein oder gar Null, so dass der Effekt auf die Schätzungen vernachlässigbar sein sollte. Um sicher zu gehen, empfiehlt es sich aber, den Anteil Merkmalsausfälle im interessierenden Untersuchungsbereich im Auge zu behalten.

## 7 Pooling

### 7.1 Einführung

Das BFS hat die Möglichkeit vorgeschlagen und umgesetzt, Strukturhebungsstichproben aus mehreren aufeinanderfolgenden Jahren, derzeit 3 und 5 Jahren, zu kombinieren (zu poolen), um präzise Ergebnisse für Gebiete mit geringerer Grösse zu liefern, als dies allein mit einer Jahrestichprobe möglich ist. Um zu vermeiden, dass die gleiche Person oder derselbe Haushalt in mehreren Strukturhebungen gezogen wird, werden die Stichproben negativ koordiniert gezogen. Diese Koordination der Stichproben hat positive Auswirkungen auf das «Pooling», da sie sicherstellt, dass die Stichproben aus den verschiedenen Jahren des Poolings im Wesentlichen disjunkt sind und somit ein Minimum an redundanten Informationen gesammelt wird. Sie macht es jedoch schwierig, wenn nicht gar unmöglich, sehr genaue Fehlergrenzen zu schätzen (siehe Abschnitt 7.3.2). Dieser Abschnitt beschreibt die verwendete Methodik, die vorgeschlagenen Schätzverfahren und deren Eigenschaften.

### 7.2 Ziele des Poolings

Es gibt viele Erfahrungen, auch im BFS, mit der Kombination von Daten aus verschiedenen oder wiederholten Erhebungen. So werden beispielsweise Stichproben der jährlichen thematischen Erhebungen des BFS aus drei oder vier aufeinanderfolgenden vierteljährlichen Erhebungsrahmen gezogen. Was die «Stichproben der Volkszählungen» betrifft, die in einigen Ländern grosse Erhebungen sind, die die traditionelle Volkszählung ersetzen, so ist dies ebenfalls eine weit verbreitete Praxis. Der Leser könnte am Fall des American Community Survey (siehe [U.S.C.B., 2008](#)) oder der neuen Volkszählung der französischen Bevölkerung (siehe [INSEE, 2005b](#), Kapitel D) interessiert sein.

Der Zweck dieser Kombination oder des «Poolings» besteht darin, durch Erhöhung der Anzahl Beobachtungen genauere Schätzwerte zu berechnen, als jene, die mit jeder Stichprobe einzeln geschätzt werden können. Bei der Verwendung mehrerer Erhebungsrahmen zu unterschiedlichen Zeitpunkten und der Datenerhebung, die sich über einen langen Zeitraum erstreckt, ist

die Wahl dieser Schätzer und der zu schätzenden Parameter jedoch nicht offensichtlich. Tatsächlich können Veränderungen in der Bevölkerung und ihren Merkmalen nicht vernachlässigt werden, wenn die Datenerhebung mehrere Jahre dauert.

Um diese Kombination von Stichproben zu ermöglichen, scheinen zwei Ansätze denkbar zu sein, die zwei verschiedenen Zielen entsprechen. Im Falle des «Poolings» der Stichproben der Strukturhebung hat das BFS den Ansatz gewählt, die Stichproben einfach zu kombinieren. Die resultierenden Schätzer sind Kombinationen der jährlichen Schätzer. Dieser Ansatz hat mehrere Vorteile :

- seine Einfachheit in der Anwendung, weil es nur einfache Operationen auf den bereits berechneten jährlichen Hochrechnungsgewichten erfordert,
- seine Flexibilität, weil das, was für drei aufeinander folgende Jahre entwickelt wird, für jede Dauer und Kombination von Jahren angewendet werden kann, ohne dass grössere Änderungen erforderlich sind,
- seine Interpretierbarkeit, weil das Prinzip und die Zielsetzung bei Summenschätzungen leicht verständlich sind: Wir versuchen, einen Durchschnittswert über einen Zeitraum zu schätzen, und wir verwenden einfach den Durchschnitt der jährlichen Schätzungen über diesen Zeitraum. Die Ergebnisse sind kohärent mit den veröffentlichten jährlichen Schätzungen,
- die Tatsache, dass man kein statistisches und willkürliches Modell der zeitlichen Entwicklungen der Zielgrössen einführen muss.

Andererseits erlaubt dieses Verfahren keine feineren oder genaueren Schätzungen für das Medianjahr des Zeitraums, der durch das Pooling abgedeckt ist. Zudem, auch wenn die Interpretation der Schätzer von Summen in der Regel einfach ist, ist dies nicht unbedingt der Fall für alle erstellten Statistiken, wie Quantile oder Durchschnitte (siehe Abschnitt 7.3.1). Im Allgemeinen erschwert sich die Interpretation der Resultate mit Zunahme der Anzahl gepoolten Jahre. Darum ist es nicht wirklich sinnvoll, zu viele Jahre miteinander zu poolen.

Das andere mögliche Ziel, das für die Strukturhebung nicht weiterverfolgt wurde, besteht darin, Schätzer für das zentrale Jahr des Poolings oder für ein bestimmtes Datum zu erstellen, wobei auch die in den anderen Jahren gesammelten Informationen verwendet werden. Dies erfordert die implizite oder explizite Verwendung von Modellen, die die Bevölkerungsmerkmale anderer Jahre mit denen des Bezugsjahres verbinden. Eine explizite Möglichkeit wäre es, für jede in den anderen Jahren des Poolings befragte Einheit die dem Referenzjahr entsprechenden Merkmale vorherzusagen oder einzusetzen. Einige der Einsetzungen wären deterministisch und betreffen Merkmale, die in den Registern vorhanden sind (wie Alter...) oder aus diesen Beobachtungen berechenbar sind. Die anderen Merkmale sollten mit Hilfe von Modellen vorhergesagt werden, deren Parameter entsprechend den verfügbaren Daten angepasst werden. Eine besondere Schwierigkeit besteht darin, dass es notwendig ist, Einheiten in verschiedenen Zeitpunkten zu beobachten, um die Evolutionen gut schätzen zu können, und dies ist bei der Strukturhebung nahezu ausgeschlossen. Ein einfacherer Weg, bei welchen diese Modellierung implizit bleibt, wäre die Verwendung eines Kalibrierungsverfahrens, um die Hochrechnungsgewichte der verschiedenen Stichproben anzupassen, um eine bestmögliche Übereinstimmung mit den Randverteilungen des Referenzjahres zu erreichen.

Dieser zweite Ansatz wird bei einer Vielzahl der einzelnen Erhebungen beim BFS verwendet: Die Datenerhebung dauert in der Regel mehrere Wochen oder Monate, der Erhebungsrahmen selbst enthält Informationen, die zum Zeitpunkt der Ziehung mehr als einen Monat alt sind. Bei thematischen Erhebungen werden die Stichproben in drei oder vier aufeinanderfolgenden vierteljährlichen Erhebungsrahmen ausgewählt, und die Datenerhebung kann über einen grossen

Teil des Kalenderjahres erfolgen. Diese Streuung, auch wenn sie nicht vollständig kontrolliert wird und im Wesentlichen aus organisatorischen Einschränkungen resultiert, hat den Vorteil, mögliche saisonale Schwankungen der identifizierten Merkmale auszugleichen. Aus Gründen der Interpretierbarkeit werden die Ergebnisse der Erhebung jedoch willkürlich mit einem Stichtag versehen. Dies ist in der Regel der Stichtag für die STATPOP-Bevölkerungsstatistik, der 31. Dezember des Jahres vor der Erhebung. Durch die Kalibrierung auf die STATPOP-Resultate können zu dieser konsistente Erhebungsergebnisse geschätzt werden (siehe [Deville und Särndal, 1992](#)).

### 7.3 Schätzer des Poolings

In diesem Abschnitt wird das «Pooling» von drei Jahren beschrieben. Bei der Ausweitung auf ein Pooling von fünf Jahren wird analog verfahren. Seien  $U_1, U_2$  und  $U_3$  die drei «Populationen» der drei zu betrachtenden Erhebungsjahre,  $\mathbf{Y}_1, \mathbf{Y}_2$  und  $\mathbf{Y}_3$  sind die drei Zielvariablen, welche in diesen Populationen definiert sind und  $\hat{Y}_1, \hat{Y}_2, \hat{Y}_3$  die drei Totalschätzer von  $\mathbf{Y}_1, \mathbf{Y}_2$  und  $\mathbf{Y}_3$  resultierend aus den entsprechenden Strukturserhebungen (siehe Abschnitt 6). Diese drei Schätzer sind gewichtete Summen der Beobachtungen:

$$\hat{Y}_t = \sum_{k \in s_t} w_{k,t} y_{k,t}, \quad t = 1, 2, 3,$$

wobei  $s_t$  die Antwortenden der Strukturserhebung «t»,  $w_{k,t}$  das Hochrechnungsgewicht, das in dieser Erhebung für die Einheit  $k$  berechnet wurde und  $y_{k,t}$  den Wert der erhobenen Zielvariablen ist. Die Gewichtung jeder Erhebung wurde mit dem Ziel berechnet approximativ unverzerrte Schätzer  $\hat{Y}_t$  für das Total  $\mathbf{Y}_t$  unter der Berücksichtigung der zufälligen Stichprobe und den Antwortausfällen zu erhalten.

Der «Pooling-Schätzer» wird definiert mit:

$$\hat{Y}_m = \frac{1}{3}(\hat{Y}_1 + \hat{Y}_2 + \hat{Y}_3) = \frac{1}{3} \left( \sum_{k \in s_1} w_{k,1} y_{k,1} + \sum_{k \in s_2} w_{k,2} y_{k,2} + \sum_{k \in s_3} w_{k,3} y_{k,3} \right). \quad (11)$$

In der Praxis ist der Schätzer (11) der einfache Mittelwert der veröffentlichten jährlichen Schätzer. Er wird durch Aneinanderhängen der drei gewichteten Dateien der Strukturserhebung und Division der jährlichen Hochrechnungsgewichte durch drei erhalten. Die Personen- und Haushaltsdateien, die Gewichte von Personen, Haushalten und Zielpersonen in den ausgewählten Haushalten werden in gleicher Weise verwendet.

#### 7.3.1 Eigenschaften

Der Schätzer (11) ist ein approximativ unverzerrter Schätzer des arithmetischen Mittelwertes:

$$\bar{Y} = \frac{1}{3}(Y_1 + Y_2 + Y_3)$$

der Totale von  $\mathbf{Y}_1, \mathbf{Y}_2$  und  $\mathbf{Y}_3$ . Das bedeutet insbesondere, dass die mittels Pooling erhaltenen Tabellen mit den Populationsgrößen nach Geschlecht, Alter, Kanton, Nationalität etc. die Mittelwerte der Werte aus den jährlichen Tabellen der Periode des Poolings sind.

Die mittels der gepoolten Datei geschätzten Mittelwerte sind hingegen nicht identisch mit dem Mittelwert über die jährlichen Mittelwerte. In der Tat wird, bei Benutzung der gepoolten Daten, der Schätzer für den Mittelwert berechnet mit

$$\widehat{Y}_m = \frac{\widehat{Y}_m}{\widehat{N}_m} = \frac{(\widehat{Y}_1 + \widehat{Y}_2 + \widehat{Y}_3)/3}{(\widehat{N}_1 + \widehat{N}_2 + \widehat{N}_3)/3}, \quad (12)$$

wobei  $\widehat{N}_t$  die geschätzte Grösse der Population oder eines betrachteten Untersuchungsbereichs zum Zeitpunkt  $t$  ist. Der Schätzer (12) ist normalerweise nicht gleich dem arithmetischen Mittelwert der geschätzten jährlichen Mittel  $\widehat{Y}_t/\widehat{N}_t$ ,  $t = 1, 2, 3$ . Sein Erwartungswert ist nur approximativ gleich wie  $(Y_1 + Y_2 + Y_3)/(N_1 + N_2 + N_3)$ .

Die Quantile oder andere nicht lineare Statistiken, welche mit der gepoolten Datei berechnet wurden, entsprechen ihrerseits im Allgemeinen ebenfalls nicht dem arithmetischen Mittel der drei entsprechenden Schätzer der einzelnen Jahre.

Daher muss der Verwendung dieser Daten und der Interpretation der zu schätzenden Parameter besondere Aufmerksamkeit geschenkt werden.

### 7.3.2 Genauigkeit

Die theoretische Varianz des Schätzers (11) erfüllt die Gleichung (13) :

$$\text{Var}(\widehat{Y}_m) = \frac{1}{9} \left[ \text{Var}(\widehat{Y}_1) + \text{Var}(\widehat{Y}_2) + \text{Var}(\widehat{Y}_3) + 2\text{Cov}(\widehat{Y}_1, \widehat{Y}_2) + 2\text{Cov}(\widehat{Y}_1, \widehat{Y}_3) + 2\text{Cov}(\widehat{Y}_2, \widehat{Y}_3) \right], \quad (13)$$

wobei die Terme  $\text{Var}(\cdot)$  und  $\text{Cov}(\cdot, \cdot)$  von der gemeinsamen Wahrscheinlichkeitsverteilung von  $(s_1, s_2, s_3)$  abhängen. Die Stichproben  $s_1$ ,  $s_2$  und  $s_3$  wurden negativ koordiniert und somit nicht unabhängig voneinander gezogen, daher sind die Schätzer  $\widehat{Y}_1$ ,  $\widehat{Y}_2$  und  $\widehat{Y}_3$  ebenso nicht unabhängig und die Kovarianz-Terme in (13) sind nicht Null, ausser in Ausnahmefällen.

Die  $\text{Var}(\widehat{Y}_i)$ -Terme sind gleich den Varianzen der einzelnen Strukturerhebungen. Sie können mit denselben Methoden geschätzt werden, wie sie in den einzelnen jährlichen Strukturerhebungen angewandt werden (siehe Abschnitt 6.2).

Die Kovarianz-Terme hängen von der Überlappung der Stichproben und der Korrelation der Ziel-Variablen zwischen den drei betrachteten Populationen ab. Ihre Schätzung würde erfordern, dass für jedes der drei Erhebungspaare eine ausreichende Anzahl Personen und Haushalte an beiden Erhebungen als Antwortende vorhanden sind. Genau dies wird jedoch durch die Koordination der Stichproben vermieden. Es ist daher nur möglich, eine Annäherung der Varianz des Schätzers (11) durch Vernachlässigung dieser Kovarianz-Terme vorzuschlagen. Der so entstandene approximative Fehler wird in den Anhängen im Abschnitt A untersucht. Er ist im Allgemeinen konservativ, was bedeutet, dass die geschätzten Konfidenzintervalle tendenziell zu gross sind. Für Schätzer bei den Zielpersonen ist der approximative Fehler sehr gering, da die Stichprobenrate der Strukturerhebung relativ gering ist (siehe Anhang A). Er kann jedoch für die Haushaltsschätzungen an Bedeutung gewinnen, wenn wir uns speziell auf grosse Haushalte konzentrieren, die hohe Stichprobenraten aufweisen.

## 7.4 Pooling über 5 Jahre, Entwicklungsmasse

### 7.4.1 Pooling über 5 Jahre

Die Zusammenlegung von fünf aufeinander folgenden Untersuchungsjahren folgt dem gleichen Prinzip wie die Zusammenlegung von drei Jahren. Die Erhöhung der Anzahl der Jahre, die für das Pooling verwendet werden, hat jedoch Auswirkungen auf die Genauigkeitsschätzungen. Nur bei Schätzungen für Zielpersonen ist die Überlappung von Stichproben über fünf aufeinander folgende Jahre in der Regel sehr gering. Selbst bei einem Kanton, der jedes Jahr seine Stichprobe

aufstockt, liegt die kumulierte Stichprobenrate über fünf Jahre bei rund 35%, was eine nahezu vollständige Trennung der Stichproben ermöglicht. Die beobachteten Überlappungen lassen sich in der Regel fast ausschliesslich durch Veränderungen in der Haushaltszusammensetzung erklären. Die Varianz eines Pooling-Schätzers  $\hat{Y}_{12345}$  kann wie folgt ausgedrückt werden

$$\text{Var}(\hat{Y}_{12345}) = \frac{1}{25} \left[ \sum_{t=1}^5 \text{Var}(\hat{Y}_t) + \sum_{t=1}^5 \sum_{r \neq t} \text{Cov}(\hat{Y}_t, \hat{Y}_r) \right].$$

Sie beinhaltet somit 5 Varianz- und 20 Kovarianz-Terme, während es im Falle des dreijährigen Poolings 3 Varianz- und 6 Kovarianz-Terme gab. Es ist daher zu erwarten, dass der Fehler der Varianz-Approximation durch Vernachlässigung der Kovarianzbedingungen grösser als beim dreijährigen Pooling ist.

Bei den Haushaltsschätzungen ist auch zu erwarten, dass der Fehler bei der Genauigkeitsschätzung grösser sein wird als bei einem 3-jährigen Pooling. In einem Kanton mit Aufstockung der Stichprobe sollten fast alle Haushalte mit drei oder mehr Personen über einen Zeitraum von fünf Jahren mindestens einmal befragt werden. So sollte die Varianz der Pooling-Schätzung nahezu Null sein, da die gesamte Population beobachtet wurde. Dieser Tatsache wird man nicht gerecht, wenn man die Kovarianz-Terme vernachlässigt.

#### 7.4.2 Vergleich von sukzessiven Pooling-Resultaten

Eine Charakteristik des Durchschnitts der gepoolten Daten ist seine Fähigkeit, Entwicklungen zu glätten und insbesondere den Einfluss von Maxima oder Minima zu verkleinern. Die durch das Pooling gewonnenen Schätzungen sind daher kein geeignetes Instrument, um Veränderungen genau zu messen. Zum Beispiel, wenn wir  $\hat{Y}_{123}$ , den Schätzer des Totals von  $Y$  gepoolt über die Jahre 1, 2, 3 und  $\hat{Y}_{234}$ , den gepoolten Schätzer über die Jahre 2, 3, 4 betrachten, haben wir

$$\hat{Y}_{234} - \hat{Y}_{123} = \frac{\hat{Y}_4 - \hat{Y}_1}{3}. \quad (14)$$

Der Schätzer (14) ist natürlich nicht ein Schätzer für die Entwicklung der medianen Jahre ( $Y_3 - Y_2$ ), jedoch erlaubt er die mittlere Entwicklung zwischen dem Jahr 1 und dem Jahr 4 zu schätzen. Die Varianz dieses Schätzers ist :

$$\text{Var}(\hat{Y}_{234} - \hat{Y}_{123}) = \frac{1}{9} \left[ \text{Var}(\hat{Y}_4) + \text{Var}(\hat{Y}_1) - 2\text{Cov}(\hat{Y}_4, \hat{Y}_1) \right].$$

Wenn die Stichproben aus den Jahren 1 und 4 disjunkt sind oder eine sehr geringe Überlappung aufweisen, kann diese Varianz mit einem geringen Fehler durch Vernachlässigung des Kovarianz-Terms geschätzt werden. Diese Schätzung wird jedoch im Allgemeinen eher optimistisch sein, d.h. sie wird etwas zu kleine Konfidenzintervalle liefern. Änderungen zwischen Pooling-Schätzern, die keine Totalschätzer sind, zwischen Pooling-Schätzern über verschiedene Zeiträume oder zwischen Pooling-Schätzern über 5 Jahre, müssen von Fall zu Fall untersucht werden.

#### 7.5 Empfehlungen

Die Ergebnisse des Anhangs A zeigen, dass die für das Pooling von drei aufeinanderfolgenden Jahren unter Berücksichtigung der Korrelationen zwischen den Stichproben geschätzten Vertrauensintervalle tatsächlich kleiner sind als die mit dem einfachen Verfahren geschätzten. Der Unterschied ist jedoch nur für die grössten Haushalte in den Kantonen, die ihre Stichproben aufgestockt haben, wesentlich. Angesichts der geringen Gewinne, die im Allgemeinen mit

dem komplexen Verfahren erzielt werden, und da die Daten keine direkte Schätzung ohne sehr einschränkende Hypothesen erlauben, empfehlen wir, das vereinfachte Schätzverfahren zu bevorzugen, wobei wir wissen, dass es konservative Konfidenzintervalle liefert.

In der Praxis können die Schätzverfahren SURVEYFREQ und SURVEYMEANS in SAS als erste Annäherung verwendet werden, indem ein geschichteter Stichprobenplan nach Jahr gekreuzt mit Zone und der Haushaltsgröße (Anzahl der ständigen Wohnbevölkerung ab 15 Jahren) festgelegt wird. Es erscheint sinnvoll, Haushalte ab einer Grösse von 5 und mehr Personen in derselben Klasse zu gruppieren, da es zu wenige grössere Haushalte gibt, um die Streuungen der Variablen korrekt zu schätzen. Darüber hinaus ist die tatsächlich beobachtete Stichprobenrate, die sich aus der Ziehung, den Antwortausfällen und dem Eliminieren von Haushalten ergibt, für diese grossen Haushalte relativ homogen.

## **8 Schlussfolgerungen**

Dieser Methodenbericht beschreibt die aktuelle Methode zur Gewichtung der Stichproben der Strukturhebung. Er liefert auch Informationen über die Schätzung der Genauigkeit der Ergebnisse und die Möglichkeit, Stichproben aus mehreren Erhebungsjahren zu kombinieren. Obwohl die Verfahren zur Ziehung, Datensammlung und Gewichtung der Stichprobe von Jahr zu Jahr im Wesentlichen gleich sind, wurden Anpassungen zur Verbesserung der Qualität der Ergebnisse oder aufgrund veränderter Rahmenbedingungen vorgenommen. Bei der Ziehung der Stichprobe sind die wichtigsten Änderungen: die von den Gemeinden und Kantonen bestellten Aufstockungen und seit der SE2013 der Übergang zu einer zweiphasigen Stichprobenziehung, zuerst Haushalt und dann Zielperson. Hinsichtlich der Gewichtung ist anzumerken, dass ab der SE2014 die Modellierung der Antwortausfälle erweitert und die gleichzeitige Kalibrierung von Haushalts- und Personenstichproben eingeführt wurde. Ab der SE2015 wurde die Modellierung der Antwortausfälle für die Haushalte in bestimmten Gemeinden um einige Variablen erweitert.

## Anhang

### A Approximation der Varianz für das Pooling

#### A.1 Aspekte über die Kovarianzen

Die Studie von vereinfachten Spezialfällen erlaubt es, die Bedeutung der Kovarianz-Terme bei der Schätzung der Genauigkeit des Schätzers (11) zu evaluieren. In einem sehr einfachen Modell ist die Population über den betrachteten Zeitraum konstant und die Stichproben sind disjunkt. Für zwei Stichproben, die nach einem einfachen Stichprobenplan in einer konstanten Population der Grösse  $N$  gezogen werden, ist die Kovarianz zwischen den Horvitz-Thompson-Schätzern, der beiden Stichproben,  $\hat{Y}_1$  und  $\hat{Y}_2$ , (siehe Tam, 1984; Qualité und Tillé, 2008) gleich

$$\text{Cov}(\hat{Y}_1, \hat{Y}_2) = N^2 \left( \frac{n_c}{n_1 n_2} - \frac{1}{N} \right) S_{y_{.,1}y_{.,2}}, \quad (15)$$

wobei  $n_1$  und  $n_2$  die entsprechenden Stichprobengrössen,  $n_c$  die Grösse ihrer Überlappung und

$$S_{y_{.,1}y_{.,2}} = \frac{1}{N-1} \sum_k (y_{k,1} - \bar{y}_{.,1})(y_{k,2} - \bar{y}_{.,2})$$

die korrigierte Kovarianz der während der beiden Erhebungen beobachteten Variablen ist. Falls die Pläne disjunkt sind, erhält man

$$\text{Cov}(\hat{Y}_1, \hat{Y}_2) = -N S_{y_{.,1}y_{.,2}}. \quad (16)$$

Für kleine Stichprobenraten kann die Korrelation vernachlässigt werden. In der Tat ist die Grössenordnung der Jahres-Varianzterme gegeben mit  $N^2/n_1$  und  $N^2/n_2$ , während die Kovarianz in der Grössenordnung von  $N$  ist, also die Varianz multipliziert mit der Stichprobenrate.

Streng genommen erfüllen die Stichproben der Strukturhebung diese Gleichungen nicht genau: ihre Überlappung ist klein, aber nicht immer leer, die Bevölkerung entwickelt sich zwischen den Erhebungen, der Stichprobenplan basiert nicht auf einer einfachen Zufallsstichprobe und die Stichprobenraten können insbesondere für grosse Haushalte relativ hoch sein. Einige dieser Aspekte lassen sich leicht in die theoretische Berechnung integrieren. Die Haushaltsstichproben der Strukturhebung sind Poisson-Stichproben, die aus einer dynamischen Population ausgewählt wurden. Für solche Stichproben ist die Kovarianz zwischen den linearen Total-Schätzern  $\hat{Y}_1$  und  $\hat{Y}_2$  mit festen Gewichten  $w_{.,1}$  und  $w_{.,2}$

$$\begin{aligned} \text{Cov}(\hat{Y}_1, \hat{Y}_2) &= \text{Cov} \left( \sum_{k \in U_1 \setminus U_2} w_{k,1} y_{k,1} I_{k,1} + \sum_{k \in U_1 \cap U_2} w_{k,1} y_{k,1} I_{k,1}, \right. \\ &\quad \left. \sum_{k \in U_1 \cap U_2} w_{k,2} y_{k,2} I_{k,2} + \sum_{k \in U_2 \setminus U_1} w_{k,2} y_{k,2} I_{k,2} \right) \\ &= \text{Cov} \left( \sum_{k \in U_1 \cap U_2} w_{k,1} y_{k,1} I_{k,1}, \sum_{k \in U_1 \cap U_2} w_{k,2} y_{k,2} I_{k,2} \right) \\ &= \sum_{k \in U_1 \cap U_2} w_{k,1} y_{k,1} w_{k,2} y_{k,2} \text{Cov}(I_{k,1}, I_{k,2}) \\ &= \sum_{k \in U_1 \cap U_2} w_{k,1} y_{k,1} w_{k,2} y_{k,2} [\mathbf{E}(I_{k,1} I_{k,2}) - \pi_{k,1} \pi_{k,2}]. \end{aligned}$$

Falls die Pläne disjunkt sind, erhält man

$$\text{Cov}(\widehat{Y}_1, \widehat{Y}_2) = - \sum_{k \in U_1 \cap U_2} w_{k,1} y_{k,1} w_{k,2} y_{k,2} \pi_{k,1} \pi_{k,2}. \quad (17)$$

Wenn darüber hinaus die Bevölkerung  $U$  zwischen den einzelnen Erhebungen vollständig erneuert würde, wäre die Kovarianz gleich Null. Um die durchgeführten Kalibrierungen zu berücksichtigen, können die Variablen  $y_k$  in der Gleichung (17) durch die Residuen der Regression  $y_k$  auf die Kalibrierungsvariablen ersetzt werden.

Diese Formel entspricht nur approximativ dem Fall der Zielpersonen der Strukturhebung, da die Selektionen innerhalb desselben Haushalts nicht unabhängig, sondern ausschliessend sind. Somit passen die Stichproben von Personen aus den Strukturhebungen in keinen der beiden vorgestellten Fälle. Die entsprechenden Formeln liefern uns jedoch relativ gute Approximationen für die Kovarianzen.

Ob in (16) oder (17) muss das Produkt  $y_{k,1} y_{k,2}$  berücksichtigt werden, um die Kovarianz zu schätzen. Bei disjunkten Stichproben beobachten wir jedoch nicht  $y_{k,1}$  und  $y_{k,2}$  an denselben Einheiten, und deshalb kennen wir den Wert von  $y_{k,1} y_{k,2}$  nicht. Da die notwendigen Daten nicht verfügbar sind, könnte nur ein deterministisches und willkürliches Modell diese Informationslücke schliessen, was aber zu Ergebnissen führen würde, die nicht vertretbar sind.

## A.2 Vereinfachte Schätzung der Genauigkeit

In den beiden in der Sektion A.1 untersuchten Fällen und für fast disjunkte Stichproben sind die Kovarianz-Terme in der Gleichung (13) von einer Grössenordnung vergleichbar mit derjenigen der Varianz-Terme multipliziert mit der Stichprobenrate. Solange letztere gering ist, ist der Einfluss von Kovarianzen auf die Gesamtgenauigkeit sehr begrenzt. Dies gilt umso mehr, als die Grösse der Konfidenzintervalle eine Funktion von Standardabweichungen sind, die nicht sehr empfindlich auf kleine Veränderungen der Varianz reagieren. Darüber hinaus, wenn  $y_{.,1}$  und  $y_{.,2}$  positiv korreliert sind, dann ist die Kovarianz negativ, was bei deren Vernachlässigung zu konservativen Konfidenzintervallen führt. Die vorgeschlagene Approximation ist die Berechnung konservativer Konfidenzintervalle unter Vernachlässigung der Kovarianzen zwischen den verschiedenen Erhebungsjahren. Die Varianz von  $\widehat{Y}_m$  wird dann geschätzt bzw. überschätzt mit

$$\widehat{\text{Var}}_a(\widehat{Y}_m) = \frac{1}{9} \left[ \widehat{\text{Var}}(\widehat{Y}_1) + \widehat{\text{Var}}(\widehat{Y}_2) + \widehat{\text{Var}}(\widehat{Y}_3) \right], \quad (18)$$

wobei  $\widehat{\text{Var}}(\widehat{Y}_i)$ ,  $i = 1, 2, 3$  die Varianzschätzer der entsprechenden Strukturhebungen sind. Solange diese Schätzer ohne Bias sind, hat  $\widehat{\text{Var}}_a(\widehat{Y}_m)$  als Erwartungswert

$$\text{Var}_a(\widehat{Y}_m) = \frac{1}{9} \left[ \text{Var}(\widehat{Y}_1) + \text{Var}(\widehat{Y}_2) + \text{Var}(\widehat{Y}_3) \right]. \quad (19)$$

Der Hauptgrund für diese vereinfachte Berechnung ist, dass die für eine genauere, unverzerrte Berechnung notwendigen Daten nicht verfügbar sind und auch in Zukunft nicht zur Verfügung stehen werden.



**4. Welcher Kirche oder Religionsgemeinschaft gehören Sie an?**

1  römisch-katholische      5  evangelisch-freikirchliche\*      9  muslimische\*

2  christkatholische (altkatholische)      6  christlich-orthodoxe\*      10  buddhistische\*

3  evangelisch-reformierte (protestantische)      7  andere christliche Kirche oder Gemeinschaft\*      11  hinduistische\*

4  keine      8  jüdische      12  andere Religionsgemeinschaft\*

\* Bitte präzisieren Sie, um welche Religionsgemeinschaft es sich handelt:

---

**5. Achtung! Frage 5 nur beantworten, wenn Sie in der Schweiz geboren sind.**

Lebten Sie jemals während mindestens eines Jahres ohne Unterbruch ausserhalb der Schweiz?

1  nein      2  ja → In welchem Jahr sind Sie zum letzten Mal zum Wohnen in die Schweiz zurückgekehrt?    (Jahr)

---

**6. Wo wurden Ihre Eltern geboren?**

a) Ihre Mutter      1  in der Schweiz      2  im Ausland      3  unbekannt

b) Ihr Vater      4  in der Schweiz      5  im Ausland      6  unbekannt

---

**7. Achtung! Die Fragen a) und b) nur bei Schweizer Staatsangehörigkeit beantworten.**

a) Haben Sie die Schweizer Staatsangehörigkeit seit Geburt?

1  ja      2  nein → Seit wann haben Sie die Schweizer Staatsangehörigkeit?    (Jahr)

b) Haben Sie aktuell eine zusätzliche Staatsangehörigkeit?

1  nein      2  ja → von welchem Staat? (nur eine Angabe)

2.01  Italien      2.04  Österreich      2.07  Niederlande      2.10  Serbien      2.13  Mazedonien

2.02  Frankreich      2.05  Spanien      2.08  Vereinigtes Königreich      2.11  Bosnien und Herzegowina      2.14  Kosovo

2.03  Deutschland      2.06  Portugal      2.09  Türkei      2.12  Kroatien      2.15  Sri Lanka

2.16  andere:

---

8. Geben Sie in Kolonne a) alle Ihre abgeschlossenen Ausbildungen... (alle erworbenen Abschlüsse ankreuzen)	a) Abgeschlossene Ausbildungen	b) Zurzeit in Ausbildung
9. ... und in Kolonne b) Ihre gegenwärtige Ausbildung an. (nur eine Angabe)		
1 keine	1 <input type="checkbox"/>	1 <input type="checkbox"/>
2 bis maximal 7 Jahre obligatorische Schule	2 <input type="checkbox"/>	2 <input type="checkbox"/>
3 obligatorische Schule (Real-, Sekundar-, Bezirks-, Orientierungsschule, Pro-, Untergymnasium, Sonderschule)	3 <input type="checkbox"/>	3 <input type="checkbox"/>
4 1-jährige Vorlehre, 1-jährige allgemeinbildende Schule, 10. Schuljahr, 1-jährige Berufswahlschule, Haushaltslehrgang, Sprachschule (mindestens 1 Jahr) mit Schlusszertifikat, Sozialjahr, Brückenangebote oder ähnliche Ausbildung	4 <input type="checkbox"/>	4 <input type="checkbox"/>
5 2- bis 3-jährige Diplommittelschule, Verkehrsschule, Fachmittelschule FMS oder ähnliche Ausbildung	5 <input type="checkbox"/>	5 <input type="checkbox"/>
6 Berufliche Grundbildung (Anlehre, 2- bis 4-jährige Berufslehre oder Vollzeitberufsschule, Handelsdiplom, Lehrwerkstätte oder ähnliche Ausbildung)	6 <input type="checkbox"/>	6 <input type="checkbox"/>
7 Gymnasiale Maturität, Lehrkräfte-Seminar (vorbereitende Ausbildung für Lehrkräfte von Kindergarten, Primarschule, Handarbeit, Hauswirtschaft)	7 <input type="checkbox"/>	7 <input type="checkbox"/>
8 Berufs- oder Fachmaturität	8 <input type="checkbox"/>	8 <input type="checkbox"/>
9 Höhere Fach- und Berufsausbildung mit eidg. Fachausweis, eidg. Diplom oder Meisterdiplom, Höhere Fachschule für Technik (Technikerschule TS), für Wirtschaft (HKG) oder ähnliche höhere Fachschule (2 Jahre Voll- oder 3 Jahre Teilzeitstudium)	9 <input type="checkbox"/>	9 <input type="checkbox"/>
10 Höhere Fachschule (Vorgänger von Fachhochschulen, z.B. HTL, HWV, HFG, HFS) inklusive Nachdiplome (3 Jahre Voll- oder 4 Jahre Teilzeitstudium)	10 <input type="checkbox"/>	10 <input type="checkbox"/>
11 Bachelor (Universität, ETH, Fachhochschule, pädagogische Hochschule)	11 <input type="checkbox"/>	11 <input type="checkbox"/>
12 Master, Lizentiat, Diplom, Staatsexamen, Nachdiplom (Universität, ETH, Fachhochschule, pädagogische Hochschule)	12 <input type="checkbox"/>	12 <input type="checkbox"/>
13 Doktorat, Habilitation	13 <input type="checkbox"/>	13 <input type="checkbox"/>



Abbildung 6 Personen Fragebogen, Seite 2

**10. Erlerner Beruf: Welchen Beruf haben Sie im Rahmen Ihrer höchsten abgeschlossenen Ausbildung erlernt?**

Geben Sie nach Möglichkeit die offizielle Bezeichnung an, z.B.: KÖCHIN, POLYMECHANIKER, TREUHANDLER MIT EIDG. FACHAUSWEIS, PFLEGEFACHFRAU FH. Es können auch akademische Titel angegeben werden, z.B.: LIC. IUR., MASTER OF ARTS IN POLITICAL SCIENCE, DR. MED.



**11. Welches ist Ihre gegenwärtige Situation auf dem Arbeitsmarkt? (mehrere Angaben möglich)**

Kreuzen Sie alle zutreffenden Antworten an; zählen Sie auch kleine Gelegenheitsjobs dazu.

Sie sind erwerbstätig, wenn Sie:

- mindestens eine Stunde pro Woche einer bezahlten Arbeit nachgehen,
- oder im Betrieb eines/einer Familienangehörigen unbezahlt arbeiten,
- oder Ihrer Arbeit vorübergehend fernbleiben (Ferien, Krankheit oder bezahlter Mutterschaftsurlaub, Militär-/Zivildienst), ansonsten jedoch unselbstständig oder selbstständig erwerbstätig sind.

**Lehrlinge** kreuzen sowohl «Erwerbstätigkeit» als auch «in Ausbildung» an.

- |                            |  |   |                      |
|----------------------------|--|---|----------------------|
| 1 <input type="checkbox"/> | eine Erwerbstätigkeit (Vollzeit)   | → durchschnittliche Anzahl Stunden pro Woche: | <input type="text"/> |
| 2 <input type="checkbox"/> | eine Erwerbstätigkeit (Teilzeit)   | → durchschnittliche Anzahl Stunden pro Woche: | <input type="text"/> |
| 3 <input type="checkbox"/> | mehrere Erwerbstätigkeiten (Teilzeit)                                    | → durchschnittliche Anzahl Stunden pro Woche: | <input type="text"/> |
| 4 <input type="checkbox"/> | auf Stellensuche (bei der Arbeitslosenversicherung gemeldet oder nicht)  |   |                      |
| 5 <input type="checkbox"/> | in Ausbildung (Schule, Studium, Lehre)                                   |   |                      |
| 6 <input type="checkbox"/> | Hausfrau/-mann   |   |                      |
| 7 <input type="checkbox"/> | invalide oder teilinvalide Person (z.B. IV-Rentner/in)                   |   |                      |
| 8 <input type="checkbox"/> | pensioniert (AHV, andere Pension) oder Rentner/in ausser Invalidenrenten |   |                      |
| 9 <input type="checkbox"/> | andere Situation ohne Erwerbstätigkeit                                   |   |                      |

**Die Fragen 12, 13, 14 richten sich nur an Erwerbstätige, inklusive Lehrlinge.**

**12. Welchen Beruf üben Sie gegenwärtig aus?**

Geben Sie die genaue Bezeichnung an, z.B.: BÜROANGESTELLTER (statt ANGESTELLTER), GERICHTSSCHREIBERIN (statt LIC. IUR.), METALLSCHLEIFER (statt SCHLEIFER), KUNSTMALERIN (statt MALERIN), PRIMARLEHRER (statt UNTERRICHT), FINANZDIREKTORIN (statt DIREKTORIN). Wenn Sie mehrere Berufe ausüben, geben Sie nur den Beruf an, der Ihrer Hauptbeschäftigung entspricht. Die Hauptbeschäftigung ist diejenige Beschäftigung, für die Sie am meisten Zeit aufwenden.



**13. Welches ist Ihre berufliche Stellung? (nur eine Angabe)**

Beziehen Sie sich hier auf Ihre Hauptbeschäftigung, die Sie in Frage 12 angegeben haben.

**Selbstständig oder Firmeneigentümer/in:**

- 1  selbstständig **ohne** Angestellte (der Betrieb hat nicht die Rechtsform einer AG oder einer GmbH)
- 2  selbstständig **mit** Angestellten (der Betrieb hat nicht die Rechtsform einer AG oder einer GmbH)
- 3  Firmeneigentümer/in (AG oder GmbH) **ohne** Angestellte
- 4  Firmeneigentümer/in (AG oder GmbH) **mit** Angestellten
- 5  Mitarbeiter/in im Betrieb eines/einer Familienangehörigen

**Angestellt (inklusive Lehrling):**

- 6  Lehrling
- 7  angestellt als Direktor/in oder Direktionsmitglied
- 8  angestellt **mit** Vorgesetztenfunktion
- 9  angestellt **ohne** Vorgesetztenfunktion
- 10  angestellt in geschützter Werkstatt (Werkstatt für Menschen mit Behinderung oder Suchtproblemen; ausser Betreuungspersonal)

**14. Geben Sie den Firmennamen und die genaue Adresse Ihres Arbeitsgebers an. Falls Sie in einer Filiale arbeiten, Adresse der Filiale angeben. Falls Sie selbstständig sind, Namen und Adresse Ihres Büros/Ihrer Praxis usw. angeben.**

Name der Firma:

Strasse:  Nr.:

PLZ:  Ort:

Kanton:  Land (falls Ausland):



**Abbildung 7** Personen Fragebogen, Seite 3





25. Führen Sie in den folgenden Spalten alle Personen ausser Person ① aus Frage 24 auf und behalten Sie die gleiche Personennummerierung bei.	Person ②	Person ③	Person ④	Person ⑤	Person ⑥
<p>Person ① = Person, die den Personentragebogen ausgefüllt hat.</p> <p>Vorname: _____            Geburtsjahr: (JJJJ) _____            Geburtsmonat: (MM) _____</p>	<p>Person ②</p> <p>Vorname: _____            Geburtsjahr: (JJJJ) _____            Geburtsmonat: (MM) _____</p>	<p>Person ③</p> <p>Vorname: _____            Geburtsjahr: (JJJJ) _____            Geburtsmonat: (MM) _____</p>	<p>Person ④</p> <p>Vorname: _____            Geburtsjahr: (JJJJ) _____            Geburtsmonat: (MM) _____</p>	<p>Person ⑤</p> <p>Vorname: _____            Geburtsjahr: (JJJJ) _____            Geburtsmonat: (MM) _____</p>	<p>Person ⑥</p> <p>Vorname: _____            Geburtsjahr: (JJJJ) _____            Geburtsmonat: (MM) _____</p>
<p>26. Verwandtschaftliche Beziehungen zwischen den Personen, die im Haushalt leben (siehe auch Beispiel nach Frage 24). Antworten Sie für jede erwähnte Person.</p> <p>Ehemann/Ehefrau            Partner/in (eheähnliche Partnerschaft)            Partner/in in eingetragener Partnerschaft (gleichgeschlechtliche Paare)            Sohn oder Tochter (leiblich oder adoptiert)            Stiefsohn oder Stieftochter            Vater oder Mutter            Stiefvater oder Stiefmutter            Bruder oder Schwester            Halbbruder oder Halbschwester            Grossvater oder Grossmutter            Enkel oder Enkelin            andere verwandte Person            andere nicht verwandte Person</p>	<p>Wer ist die Person ② für die Person ... ①? (eine Angabe)</p> <p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/> 11 <input type="checkbox"/> 12 <input type="checkbox"/> 13 <input type="checkbox"/></p>	<p>Wer ist die Person ③ für die Person ... ①; ②? (eine Angabe pro Kolonne)</p> <p>1 <input type="checkbox"/> <input type="checkbox"/> 2 <input type="checkbox"/> <input type="checkbox"/> 3 <input type="checkbox"/> <input type="checkbox"/> 4 <input type="checkbox"/> <input type="checkbox"/> 5 <input type="checkbox"/> <input type="checkbox"/> 6 <input type="checkbox"/> <input type="checkbox"/> 7 <input type="checkbox"/> <input type="checkbox"/> 8 <input type="checkbox"/> <input type="checkbox"/> 9 <input type="checkbox"/> <input type="checkbox"/> 10 <input type="checkbox"/> <input type="checkbox"/> 11 <input type="checkbox"/> <input type="checkbox"/> 12 <input type="checkbox"/> <input type="checkbox"/> 13 <input type="checkbox"/> <input type="checkbox"/></p>	<p>Wer ist die Person ④ für die Person ... ①; ②; ③? (eine Angabe pro Kolonne)</p> <p>1 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 2 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 3 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 4 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 5 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 6 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 7 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 8 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 9 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 10 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 11 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 12 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 13 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>	<p>Wer ist die Person ⑤ für die Person ... ①; ②; ③; ④? (eine Angabe pro Kolonne)</p> <p>1 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 2 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 3 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 4 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 5 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 6 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 7 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 8 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 9 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 10 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 11 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 12 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 13 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>	<p>Wer ist die Person ⑥ für die Person ... ①; ②; ③; ④; ⑤? (eine Angabe pro Kolonne)</p> <p>1 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 2 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 3 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 4 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 5 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 6 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 7 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 8 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 9 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 10 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 11 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 12 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 13 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>
<p>27. Welches ist die Hauptsprache, das heisst die Sprache, in der die Person denkt und die sie am besten beherrscht?</p> <p>Mehrere Angaben möglich für Personen, die in mehreren Sprachen denken und diese sehr gut beherrschen.            Für Kinder, die noch nicht sprechen können, geben Sie die Hauptsprache(n) der Mutter an.</p> <p>Deutsch (oder Schweizerdeutsch)            Französisch (oder Patois Romand)            Italienisch (oder Tessiner/Bündner-Italienischer Dialekt)            Rätoromanisch            Serbisch/Kroatisch            Albanisch            Portugiesisch            Spanisch            Englisch            andere Sprache(n)</p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/></p>
<p>28. Welches ist die höchste abgeschlossene Ausbildung der Person? (nur eine Angabe)</p> <p>keine oder bis 7 Jahre obligatorische Schule            obligatorische Schule (8 oder 9 Jahre), 10. Schuljahr/Vorlehre oder anderes Brückenangebot            Berufliche Grundbildung (Anlehre, 2- bis 4-jährige Lehre oder Vollzeitberufsschule, Handelsdiplom, Lehrwerkstätte oder ähnliche Ausbildung)            Allgemeinbildende Schule (Gymnasiale Maturität, Berufs- oder Fachmaturität, Lehrkräfte-Seminar, Diplom- oder Fachmittelschule, Verkehrsschule oder ähnliche Ausbildung)            Höhere Fach- und Berufsausbildung mit eidg. Fachausweis, Diplom oder Meisterdiplom, Technikerschule, höhere Fachschule, HTL, HWV, HFG, HFS oder ähnliche Ausbildung            Universität, ETH, pädagogische Hochschule, Fachhochschule</p>	<p>Nur für Personen ab 15 Jahren ausfüllen.</p>				
<p>29. Welches ist bei der Person die gegenwärtige Situation auf dem Arbeitsmarkt? (mehrere Angaben möglich)</p> <p>Als erwerbstätig gilt, wer mindestens eine Stunde pro Woche arbeitet.</p> <p>Erwerbstätigkeit (Vollzeit)            Erwerbstätigkeit (Teilzeit)            Anzahl Stunden pro Woche (Teilzeit): → <input type="text"/> Std.            auf Stellensuche (bei der Arbeitslosenversicherung gemeldet oder nicht)            in Ausbildung (Schule, Studium, Lehre)            Hausfrau/-mann            Rentner/in, pensioniert, AHV, IV            andere Situation ohne Erwerbstätigkeit</p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/></p>
<p>30. Welches ist die berufliche Stellung der Person? (nur eine Angabe)</p> <p>Beziehen Sie sich auf die Hauptbeschäftigung der Person.</p> <p>selbständig oder Eigentümer/in der AG oder GmbH            Mitarbeiter/in im Betrieb eines/einer Familienangehörigen            angestellt als Direktor/in oder Direktionsmitglied            angestellt mit Vorgesetztenfunktion            angestellt ohne Vorgesetztenfunktion (inklusive Lehrlinge)</p>	<p>Nur für erwerbstätige Personen (inklusive Lehrlinge) ausfüllen.</p>				
<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/></p>	<p>1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/></p>	

Abbildung 10 Haushalts- und Wohnungsfragebogen, Seite 2

**ANGABEN ZUR WOHSITUATION**

**Wichtig:** Beziehen Sie sich auf die Wohnung an der Adresse, die auf dem **Personenfragebogen auf Seite 1** bedruckt ist oder von Ihnen korrigiert wurde.

**31. Wie viele Zimmer hat die Wohnung, in der Sie leben?**

Als Zimmer gelten Wohnzimmer, Schlafzimmer usw.     (z. B. 3.0 oder 4.5)

**Nicht als Zimmer zu zählen** sind Küchen, Kochnischen, Badezimmer, Toilettenräume, Reduits (Abstellräume), Gänge, Veranden usw. sowie zusätzliche, separate Wohnräume ausserhalb der Wohnung (z. B. Mansarden).

**32. Der Haushalt, zu dem Sie gehören, bewohnt diese Wohnung als...? (nur eine Angabe)**

1  Mieter oder Untermieter

2  Genossenschafter  
Ihr Haushalt ist Mitglied einer Wohnbaugenossenschaft und bewohnt eine Wohnung dieser Genossenschaft.

3  Stockwerk-/Wohnungseigentümer

4  Eigentümer des Hauses  
Zur Erinnerung: Es handelt sich um das Gebäude (Einfamilienhaus/Mehrfamilienhaus), in dem sich die Wohnung befindet.

5  andere Situation: Wohnung wird von einem Verwandten oder Arbeitgeber kostenlos zur Verfügung gestellt, Dienstwohnung (z. B. Abwartwohnung), Nutzniessung, Pächter

**33. Falls Sie bei der vorhergehenden Frage «Mieter» oder «Genossenschafter» angekreuzt haben: Wie viel beträgt die Monatsmiete, die für die Wohnung bezahlt wird (runden Sie bitte auf ganze Frankenbeträge)?**

Die benötigten Angaben finden Sie im Mietvertrag.

**MONATLICHE MIETE** CHF

**NEBENKOSTEN** Sind die Nebenkosten (Heizung, Warmwasser usw.) in der obenstehenden Monatsmiete enthalten?  nein  ja

Wenn ja, geben Sie bitte die monatlichen Nebenkosten an. CHF

**GARAGE(N)/PARKPLATZ (-PLÄTZE)** Ist die Monatsmiete für die Garage/n und Parkplatz/-plätze in der obenstehenden Monatsmiete enthalten?  nein  ja

Wenn ja, geben Sie bitte die Monatsmiete für Garage/n und Parkplatz/-plätze an. CHF

Anzahl Garagen/Parkplätze, die in der obenstehenden Monatsmiete inbegriffen sind.

**Vielen Dank für Ihre Mitarbeit.**



**Abbildung 11** Haushalts- und Wohnungsfragebogen, Seite 3

## Literatur

- [BFS 2008] BFS: Le programme de relevés du recensement fédéral de la population. (2008)
- [BFS 2011] BFS: Le nouveau recensement de la population. (2011)
- [BFS 2016] BFS: Synopsis zu den Schlüsselmerkmalen, BFS-Nummer: do-t-40-vz\_user-syn-SHAPE.02. (2016). – URL <https://www.bfs.admin.ch/bfsstatic/dam/assets/315891/master>
- [Deville und Särndal 1992] DEVILLE, J.-C. ; SÄRNDAL, C.-E.: Calibration estimators in survey sampling. In: *Journal of the American Statistical Association* 87 (1992), S. 376–382
- [INSEE 2005] INSEE: La macro CALMAR2, Redressement d'un échantillon par calage sur marges. (2005)
- [INSEE 2005b] INSEE: Pour comprendre le recensement de la population : les produits du recensement. (2005b)
- [Qualité 2009] QUALITÉ, L.: *Unequal probability sampling and repeated surveys*. Neuchâtel, Suisse, Université de Neuchâtel, Thèse de Doctorat, 2009
- [Qualité und Tillé 2008] QUALITÉ, L. ; TILLÉ, Y.: Variance estimation of changes in repeated surveys and its application to the Swiss survey of value added. In: *Survey Methodology* 34 (2008), Nr. 2, S. 173–181
- [Tam 1984] TAM, S.M.: On covariances from overlapping samples. In: *The American Statistician* 38 (1984), S. 288–289
- [U.S.C.B. 2008] U.S.C.B.: A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know. (2008)

**Methodenberichte der Sektion Statistische Methoden des BFS**  
**Rapports de méthodes de la section méthodes statistiques de l'OFS**  
**Methodology reports published by the FSO's Statistical Methods Section**

- Potterat, J., Qualité, L., Assoulin, D. (2019). Strukturhebung der eidgenössischen Volkszählung. Stichprobenplan und -ziehung, Gewichtung, Schätzverfahren und Pooling 2010-2018. Bestellnummer: 338-0080
- Potterat, J., Assoulin, D., Nicoletti, J.-M. (2019). Beschäftigungsstatistik BESTA. Revision 2015: Stichprobenrahmen, Stichprobenplan und Hochrechnung. Bestellnummer: 338-0079
- Massiani, A. (2017). Estimation de la couverture du nouveau recensement en Suisse en 2012. Numéro de commande : 338-0078
- Nedyalkova, D., Assoulin, D. (2017). Construction of full-time equivalents for the Swiss structural business statistics. Order number: 338-0077
- Ferster, M. (2016). Energieverbrauchsstatistik EVS2014 - Stichprobe, Hochrechnung und Vergleichbarkeit mit der EVS2013. Bestellnummer: 338-0076
- Panchard, C. (2014). Enquête sur la participation aux activités culturelles 2008. Plan de sondage et estimations. Bestellnummer: 338-0073
- Assoulin, D. (2014). Wertschöpfungsstatistik. Revision 2009: Statistische Datenaufbereitung und Hochrechnung. Bestellnummer: 338-0072
- Wilhelm, M. (2014). Echantillonnage boule de neige. La méthode de sondage déterminé par les répondants. Numéro de commande: 338-0071
- Assoulin, D. (2013). Wertschöpfungsstatistik. Revision 2009: Stichprobenrahmen und Stichprobenplan. Bestellnummer: 338-0070
- Ferster, M. (2013). EVS I - Energieverbrauchsstatistik 2002 bis 2007: Stichprobenplan und Hochrechnung. Bestellnummer: 338-0069
- Assoulin, D. (2013). Zusatzhebung 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 échantillonnage. 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 élèves (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

# Publikationsprogramm BFS

**Das Bundesamt für Statistik (BFS) hat als zentrale Statistikstelle des Bundes die Aufgabe, statistische Informationen zur Schweiz breiten Benutzerkreisen zur Verfügung zu stellen. Die Verbreitung geschieht gegliedert nach Themenbereichen und mit verschiedenen Informationsmitteln über mehrere Kanäle.**

## Die statistischen Themenbereiche

- 00 Statistische Grundlagen und Übersichten
- 01 Bevölkerung
- 02 Raum und Umwelt
- 03 Arbeit und Erwerb
- 04 Volkswirtschaft
- 05 Preise
- 06 Industrie und Dienstleistungen
- 07 Land- und Forstwirtschaft
- 08 Energie
- 09 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, regionale und internationale Disparitäten

## Die zentralen Übersichtspublikationen

### Statistisches Jahrbuch der Schweiz



Das vom Bundesamt für Statistik (BFS) herausgegebene Statistische Jahrbuch ist seit 1891 das Standardwerk der Schweizer Statistik. Es fasst die wichtigsten statistischen Ergebnisse zu Bevölkerung, Gesellschaft, Staat, Wirtschaft und Umwelt des Landes zusammen.

### Taschenstatistik der Schweiz



Die Taschenstatistik ist eine attraktive, kurzweilige Zusammenfassung der wichtigsten Zahlen eines Jahres. Die Publikation mit 52 Seiten im praktischen A6/5-Format ist gratis und in fünf Sprachen (Deutsch, Französisch, Italienisch, Rätoromanisch und Englisch) erhältlich.

## Das BFS im Internet – [www.statistik.ch](http://www.statistik.ch)

Das Portal «Statistik Schweiz» bietet Ihnen einen modernen, attraktiven und stets aktuellen Zugang zu allen statistischen Informationen. Gerne weisen wir Sie auf folgende, besonders häufig genutzte Angebote hin.

### Publikationsdatenbank – Publikationen zur vertieften Information

Fast alle vom BFS publizierten Dokumente werden auf dem Portal gratis in elektronischer Form zur Verfügung gestellt. Gedruckte Publikationen können bestellt werden unter der Telefonnummer 058 463 60 60 oder per Mail an [order@bfs.admin.ch](mailto:order@bfs.admin.ch).  
[www.statistik.ch](http://www.statistik.ch) → Statistiken finden → Kataloge und Datenbanken → Publikationen

### NewsMail – Immer auf dem neusten Stand



Thematisch differenzierte E-Mail-Abonnemente mit Hinweisen und Informationen zu aktuellen Ergebnissen und Aktivitäten.  
[www.news-stat.admin.ch](http://www.news-stat.admin.ch)

### STAT-TAB – Die interaktive Statistikdatenbank



Die interaktive Statistikdatenbank bietet einen einfachen und zugleich individuell anpassbaren Zugang zu den statistischen Ergebnissen mit Downloadmöglichkeit in verschiedenen Formaten.  
[www.stattab.bfs.admin.ch](http://www.stattab.bfs.admin.ch)

### Statatlas Schweiz – Regionaldatenbank und interaktive Karten



Mit über 4500 interaktiven thematischen Karten bietet Ihnen der Statistische Atlas der Schweiz einen modernen und permanent verfügbaren Überblick zu spannenden regionalen Fragestellungen aus allen Themenbereichen der Statistik.  
[www.statatlas-schweiz.admin.ch](http://www.statatlas-schweiz.admin.ch)

## Individuelle Auskünfte

### Zentrale Statistik Information

058 463 60 11, [info@bfs.admin.ch](mailto:info@bfs.admin.ch)

Die Strukturerhebung ist ein Element der seit 2010 jährlich stattfindenden eidgenössischen Volkszählung und ergänzt die Informationen aus den Registern mit zusätzlichen Statistiken zur Bevölkerungsstruktur.

Dieser Methodenbericht beschreibt die verwendeten statistischen Methoden rund um den Stichprobenplan, die Hochrechnung sowie das Pooling (Kombinieren mehrerer Erhebungsjahre). Die Stichprobe der Strukturerhebung wird im Stichprobenrahmen für Personen und Haushaltserhebungen gezogen. Die erwünschte nationale Nettostichprobe (verwendbare Antworten) wird auf 200 000 Personen festgelegt. Verschiedene Gebiete wie Kantone oder grössere Städte lassen die Stichprobe in ihrem Gebiet aufstocken. Zusammen mit der Antizipation von erwarteten Antwortausfällen führte dies zu einer Bruttostichprobengrösse von ca. 330 000 Personen.

Es wurden Gewichte für Personen (Personengewichte) und Gewichte für Auswertungen auf Ebene Haushalte (Haushaltsgewichte) entwickelt.

Die Schätzwerte von mehreren Jahren können zusammen kombiniert (gepoolt) werden. Dies hat den Vorteil von genaueren Resultaten oder erlaubt die Analyse für kleinere Untersuchungsbereiche, da die Stichprobengrösse durch das Pooling vergrössert wird.

#### **Online**

[www.statistik.ch](http://www.statistik.ch)

#### **Print**

[www.statistik.ch](http://www.statistik.ch)  
Bundesamt für Statistik  
CH-2010 Neuchâtel  
[order@bfs.admin.ch](mailto:order@bfs.admin.ch)  
Tel. 058 463 60 60

#### **BFS-Nummer**

338-0080

#### **ISBN**

978-3-303-00631-3

---

**Statistik  
zählt für Sie.**

[www.statistik-zaehlt.ch](http://www.statistik-zaehlt.ch)