

00

Statistische Grundlagen und Übersichten

338-0068

Methodenbericht

Zusatzerhebung für die landwirtschaftliche Betriebszählung 2010

Stichprobenplan und Hochrechnung

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Statistik BFS

Neuchâtel, 2013

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

0 Statistische Grundlagen und Übersichten

1 Bevölkerung

2 Raum und Umwelt

3 Arbeit und Erwerb

4 Volkswirtschaft

5 Preise

6 Industrie und Dienstleistungen

7 Land- und Forstwirtschaft

8 Energie

9 Bau- und Wohnungswesen

10 Tourismus

11 Mobilität und Verkehr

12 Geld, Banken, Versicherungen

13 Soziale Sicherheit

14 Gesundheit

15 Bildung und Wissenschaft

16 Kultur, Medien, Informationsgesellschaft, Sport

17 Politik

18 Öffentliche Verwaltung und Finanzen

19 Kriminalität und Strafrecht

20 Wirtschaftliche und soziale Situation der Bevölkerung

21 Nachhaltige Entwicklung und Disparitäten auf regionaler und internationaler Ebene

Methodenbericht

Zusatzerhebung für die landwirtschaftliche Betriebszählung 2010

Stichprobenplan und Hochrechnung

Autor Daniel Assoulin

Herausgeber Bundesamt für Statistik (BFS)

Office fédéral de la statistique (OFS)
Neuchâtel, 2013

Herausgeber: Bundesamt für Statistik (BFS)
Auskunft: Daniel Assoulin, Tel. 032 713 67 51, E-Mail: Daniel.Assoulin@bfs.admin.ch
Bearbeitung: Dienst Statistische Methoden, BFS
Vertrieb: Bundesamt für Statistik, CH-2010 Neuchâtel
Tel. 032 713 60 60 / Fax 032 713 60 61 / E-Mail: order@bfs.admin.ch
Internet: http://www.statistik.admin.ch
Bestellnummer: 338-0068
Preis: gratis
Reihe: Statistik der Schweiz
Fachbereich: 0 Statistische Grundlagen und Übersichten
Originaltext: Deutsch
Titelgrafik: BFS; Konzept: Netthoevel & Gaberthüel, Biel; Foto: © NorthShoreSurfPhotos – Fotolia.com
Grafik/Layout: Sektion DIAM, Prepress/Print
Copyright: BFS, Neuchâtel 2013
Abdruck – ausser für kommerzielle Nutzung –
unter Angabe der Quelle gestattet
ISBN: 978-3-303-00496-8

Inhaltsverzeichnis

1 Einleitung **7**

2 Population und Stichprobenplan **7**
2.1 Präzisionsziele 7
2.2 Stratifizierung 7
2.3 Festlegung der Stichprobengrösse 10
2.3.1 Antwortausfälle 12
2.4 Zu erwartende Präzisionen 13
2.4.1 Der Einfluss der Antwortausfälle auf die geplanten Präzisionen 14

3 Gewichtung und Hochrechnung **15**
3.1 Rücklauf und Gewichtskorrektur für Antwortausfälle 16
3.2 Kalibrierung 17
3.3 Die Schätzung von Totalen und Quotienten 20
3.4 Die Varianz der Schätzungen 21
3.4.1 Geplante versus erreichte Präzision 22

Literatur **25**

Methodenberichte des Dienstes statistische Methoden **27**

Vorwort

Der Dienst Statistische Methoden wurde mit der Erarbeitung von Stichprobenplan und Hochrechnungsmethodik für die Zusatzerhebung zur landwirtschaftlichen Betriebszählung 2010 beauftragt. Der Hauptteil dieser Arbeiten wurde im Zeitraum 2009-2011 durch Daniel Assoulin ausgeführt. Dabei wurde er von Arthur Zesiger und Mamoun Bencheikh Latmani von der damaligen Sektion Unternehmensstruktur (UNS) begleitet. Die Angaben über Resultate der Erhebung in diesem Bericht dienen allein der Illustration der Methoden. Es können sich Abweichungen gegenüber den offiziellen Ergebnissen ergeben, die auf der Wahl der Datenbasis, ihrem Stand, leichten Anpassungen der Methoden und der Darstellung basieren. Nur die offiziell veröffentlichten Ergebnisse sind massgebend. Der Autor dankt Arthur Zesiger sowie Mamoun Bencheikh Latmani für die gute und effiziente Zusammenarbeit. Ein Dank geht auch an

- Jean-Pierre Renfer und Jann Potterat (beide vom Dienst Statistische Methoden) für Ihre fachliche Unterstützung.
- Daniel Kilchmann vom Dienst Statistische Methoden sowie Arthur Zesiger für die aufmerksame Lektüre und hilfreichen Tipps im Rahmen der Erstellung dieses Methodenberichts.

Zusammenfassung

Die Zusatzerhebung der landwirtschaftlichen Betriebszählung ist eine Stichprobenerhebung, die eine Vielzahl Variablen aus verschiedenen Themengebieten umfasst. Der Stichprobenplan beruht auf einer einfachen geschichteten Zufallsstichprobe. Er wurde mit dem Ziel entwickelt, die landwirtschaftliche Nutzfläche sowie den Rindvieh- und Schweinebestand auf Niveau Kanton möglichst effizient mit einem Variationskoeffizienten von 5% schätzen zu können. Zusammen mit einer angenommenen Antwortrate von 60% führte dies auf eine Bruttostichprobe von 16'815 Betrieben. Daraus resultierte ein Rücklauf von 15'169 Fragebogen, die in der Hochrechnung verwendet werden konnten. Die Hochrechnungsgewichte wurden ausgehend von den Stichprobengewichten mittels einer Korrektur für Antwortausfälle und einer Kalibrierung konstruiert. Im Rahmen der Kalibrierung wurden die Gewichte so angepasst, dass die Hochrechnungsresultate verschiedener Hilfsvariablen auf Niveau Kanton mit den entsprechenden Totalen der landwirtschaftlichen Strukturerhebung 2010 übereinstimmen. Die Hochrechnungsresultate haben gezeigt, dass die Qualität der Schätzungen, ausgedrückt mittels Variationskoeffizienten, generell im erwarteten Bereich liegt: Für Zielvariablen, welche mit den im Rahmen von Stichprobenplan und Kalibrierung verwendeten Hilfsvariablen gut korrelieren, liegen die Variationskoeffizienten generell im Bereich 5% oder auch darunter. Für andere Variablen kann die Hochrechnung auf Niveau Kanton zu Variationskoeffizienten führen, die weit über den angestrebten 5% liegen.

1 Einleitung

Der vorliegende Bericht beschreibt die Methoden, welche Stichprobenplan und Hochrechnung der Zusatzerhebung zur landwirtschaftlichen Betriebszählung 2010 (LBZ2010) zugrunde liegen. Die Stichprobenerhebung dient dazu, Informationen zu erfassen, welche in der landwirtschaftlichen Strukturerhebung (Stru) 2010 nicht erhoben werden. Im Rahmen der bilateralen Abkommen hat sich die Schweiz verpflichtet, entsprechende Daten auch an Eurostat zu liefern. Die letzte solche Zusatzerhebung erfolgte im Jahre 2005, war aber betreffend dem Umfang der erhobenen Informationen wesentlich kleiner[1]. Eine, für die methodologischen Fragestellungen wichtige Eigenschaft der Zusatzerhebung 2010 ist, dass sie eine grosse Anzahl Variablen aus verschiedensten Themengebieten umfasst. Stichprobenplan und Hochrechnungsmethodik wurden dahingehend entwickelt, dass möglichst für einen grossen Anteil wichtiger Variablen aus mehreren Themengebieten Resultate mit genügender Präzision auf Stufe Kanton erzielt werden können.

2 Population und Stichprobenplan

In der Population sind alle Landwirtschaftsbetriebe der Schweiz enthalten, die die entsprechenden Normen (z.B. mindestens 1 ha landwirtschaftliche Nutzfläche) der Bundesstatistik erfüllen. Der verwendete Stichprobenrahmen wurde von der Sektion Unternehmensstruktur (UNS) erstellt. Er basiert auf der landwirtschaftlichen Strukturerhebung 2009 und enthält 59'634 Betriebe. Zu jedem Betrieb enthält der Rahmen eine Vielzahl von Angaben, welche für die Ausarbeitung des Stichprobenplanes als Hilfsvariablen dienen können.

2.1 Präzisionsziele

Es wird eine Publikation der Resultate auf Stufe Kanton angestrebt. Der Stichprobenplan soll daher folgendes Kriterium erfüllen: Für wichtige Variablen darf erwartet werden, dass Schätzungen auf Niveau Kanton mit einem Variationskoeffizienten (CV) von höchstens 5% möglich sein werden. Viele wichtige Erhebungsvariablen haben einen starken Zusammenhang mit der landwirtschaftlichen Nutzfläche, dem Rindviehbestand und/oder dem Schweinebestand des Betriebes. Diese drei Grössen wurden in der Betriebsstrukturerhebung 2009 voll erhoben und können darum dazu dienen, den Stichprobenplan zu entwickeln. Die Idee ist, einen Plan auszuarbeiten, mit welchem für jeden Kanton die landwirtschaftliche Nutzfläche, der totale Rindviehbestand sowie der totale Schweinebestand möglichst effizient (d.h. die Präzisionsziele werden mit einer möglichst kleinen Stichprobe erreicht) geschätzt werden kann. Die Annahme ist, dass dann auch für einen Grossteil der übrigen Erhebungsvariablen genügend genaue Schätzungen möglich sein werden. Tabelle 1 gibt eine Übersicht der wichtigsten Terme und Abkürzungen, welche in den folgenden Abschnitten verwendet werden.

2.2 Stratifizierung

Das Präzisionsziel wurde auf Stufe Kanton festgelegt. Aus diesem Grund bilden die 26 Kantone die Primärschichten $stra1$. Als weitere Variable wurde die landwirtschaftliche Nutzfläche verwendet, um die Grössenklassen ln_{kl} gemäss Tabelle 2 zu definieren.

[1] Insbesondere in Bezug auf die Entwicklung des Stichprobenplanes liegen die für die Zusatzerhebung 2010 verwendeten Verfahren näher bei jenen Methoden, wie sie im Rahmen der Zusatzerhebung der landwirtschaftlichen Betriebszählung 2003 verwendet worden sind. Stichprobenplan und Hochrechnungsmethode der Zusatzerhebung für das Jahr 2003 sind in [4] und [5] beschrieben.

Tabelle 1 Einige Abkürzungen und Notationen.

Abkürzung/Term	Erläuterung
ln	Landwirtschaftliche Nutzfläche in Aren
ri	Rindviehbestand
sc	Schweinebestand
$stra1$	Primärschichtung, entspricht den Kantonen
$stra2_{v1}$	Sekundärschichtung basierend auf ln, ri und sc

Tabelle 2 Grössenklassen betreffend der landwirtschaftlichen Nutzfläche.

Aren	$[0, 1000[$	$[1000, 2000[$	$[2000, 3000[$	$[3000, \infty[$
ln_{kl}	1	2	3	4

Für jede Kombination Kanton$\times ln_{kl}$ wurden nun Grenzen für die Anzahl Schweine (lim_{sc}) und die Anzahl Rindvieh (lim_{ri}) festgelegt, ab denen ein Betrieb hinsichtlich des betreffenden Tierbestandes als gross gilt. Die Tatsache, dass Betriebe, welche die Grössenlimite übertreffen, voll erhoben werden sollen, motivierte dazu, die Grössenlimiten mittels dem in [3] präsentierten Kriterium für die Definition von Take-All Schichten festzulegen. Es wird hierzu explizit der Fall einer Kombination Kanton$\times ln_{kl}$ für die Variable x=Rindviehbestand vorgestellt. Betriebe ohne Rindviehbestand ($x = 0$) werden separat behandelt und aus den folgende Betrachtungen ausgeschlossen. Für Betriebe mit positivem Rindviehbestand ($x > 0$) betrachten wir folgende zwei Grössen:

$$\begin{aligned} lim_1 &= \bar{X} + \sqrt{CV^2X^2/N + S_x^2} \\ lim_2 &= \bar{X}_{<lim_1} + \sqrt{CV^2X^2\frac{N-t-1}{(N-t)^2} + S^2_{<lim_1}} \end{aligned}$$

Dabei bezeichnen

- $\bar{X}$, $\bar{X}_{<lim_1}$ mittlerer Rindviehbestand, respektive mittlerer Rindviehbestand in Betrieben mit $x_i < lim_1$.
- S^2, $S^2_{<lim_1}$ Varianz des Rindviehbestandes, respektive Varianz des Rindviehbestandes in Betrieben mit $x_i < lim_1$.
- N,t Anzahl Betriebe, respektive Anzahl Betriebe mit einem Rindviehbestand von $x_i \geq lim_1$
- X bezeichnet den totalen Rindviehbestand.
- CV reflektiert das angepeilte Präzisionsziel für den Variationskoeffizienten der Schätzung und wurde zur effektiven Bestimmung der Limite gleich 0.05 gesetzt.

Schliesslich wird in Abhängigkeit der geschätzten Schiefe der Verteilung von x, sk_x, entweder lim_1 oder lim_2 als Grenze gewählt:

$$Lim_x = \begin{cases} lim_1 & ; sk_x \leq 3 \text{ oder } sk_x = . \\ lim_2 & ; sk_x > 3 \end{cases}$$

Zur Berechnung der Schiefe sind mindestens drei Beobachtungen nötig, weshalb die obige Fallunterscheidung auch den Fall $sk_x = .$ berücksichtigt.

Bemerkung: In (seltenen) Fällen, in welchen Lim_x aufgrund zu weniger Betriebe nicht wie oben beschrieben berechnet werden kann, d.h. lim_1 oder lim_2 ist nicht definiert, wird $lim_x = 1$ gesetzt.

Die Limiten lim_x, berechnet für die Variablen ri und sc, wurden nun dazu verwendet, ri und sc zu kategorisieren (siehe Tabelle 3). Die resultierenden Kategorisierungsvariablen wurden dabei mit ri_{kl} und sc_{kl} bezeichnet.

Tabelle 3 Kategorisierung Schweine- und Rindviehbestand.

ri	ri_{kl}	sc	sc_{kl}	Interpretation
0	0	0	0	keine Tiere
$]0; lim_{ri}[$	1	$]0; lim_{sc}[$	1	mässig Tiere
$[lim_{ri}; \infty[$	2	$[lim_{sc}; \infty[$	2	viele Tiere

Kreuzen wir die Klassen sc_{kl} und ri_{kl}, führt dies zu neun Klassen und zu einer (zu) feinen Stratifizierung. Die neun Kreuzungen wurden darum gemäss Tabelle 4 zu 4 Klassen zusammengefasst. Versuche haben gezeigt, dass diese Zusammenfassung die Effizienz des Stichprobenplanes hinsichtlich der Schätzung des totalen Schweine-, respektive des totalen Rindviehbestandes nicht wesentlich beeinträchtigt. Die resultierende Schichtungsvariable mit vier Ausprägungen wird dabei als ti_{kl} bezeichnet und hat folgende Interpretation.

- $ti_{kl} = 0$: Betrieb hat weder Schweine noch Rindvieh
- $ti_{kl} = 1$: Betrieb hat mässig Rindvieh aber keine Schweine
- $ti_{kl} = 2$: Betrieb hat mässig Schweine und mässig oder kein Rindvieh
- $ti_{kl} = 3$: Betrieb hat entweder viele Schweine oder viel Rindvieh

Tabelle 4 Konstruktion der Grössenvariable ti_{kl} für den Tierbestand.

ri_{kl}	0	0	0	1	1	1	2	2	2
sc_{kl}	0	1	2	0	1	2	0	1	2
ti_{kl}	0	2	3	1	2	3	3	3	3

Die Sekundärschichtung, welche im Gegensatz zur Primärschichtung (Kantone) in erster Linie der Optimierung des Planes gilt, ergibt sich durch die Kreuzung $ln_{kl} \times ti_{kl}$. Sie wird durch die in Tabelle 5 definierte Variable stra2 ausgedrückt.

Tabelle 5 Konstruktion der Sekundärstraten.

ln_{kl}	1	1	1	1	2	2	2	2	3	3	3	3	4	4	4	4
ti_{kl}	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
$stra2$	10	11	12	13	20	21	22	23	30	31	32	33	40	41	42	43

Tabelle 6 enthält nun die Anzahl Betriebe der einzelnen Schichten, $kt \times stra2$, basierend auf dem für die Zusatzerhebung verwendeten Stichprobenrahmen. Die relativ feine Stratifizierung bewirkt teilweise kleine Schichtgrössen. Zu kleine Schichtgrössen können zusammen mit Antwortausfällen zu Problemen bei der Schätzung führen. Es ist vorgesehen, die fett markierten Schichten voll zu erheben und bei Bedarf (zu wenige Antworten in der Nettostichprobe) für die Extrapolation zusammenzufassen. In der Phase der Stichprobenplanung brauchen wir uns

Tabelle 6 Anzahl Betriebe in den konstruierten Schichten

Kanton	stra2																
	10	11	12	13	20	21	22	23	30	31	32	33	40	41	42	43	Total
ZH	830	181	44	**28**	534	746	61	**30**	157	671	55	**37**	88	442	64	**29**	3997
BE	861	2283	429	**65**	472	3670	1112	**117**	133	1497	481	**50**	46	624	183	**58**	12081
LU	419	735	371	**65**	87	1082	1056	**74**	*11*	405	445	**45**	*3*	154	146	**35**	5133
UR	102	224	21	**13**	*14*	205	22	**8**	*2*	31	*3*	**4**	.	*8*	.	**1**	658
SZ	228	346	56	**22**	35	563	85	**15**	*5*	234	30	**19**	*1*	79	18	**6**	1742
OW	86	233	28	**14**	.	269	28	**8**	.	45	*4*	**6**	.	*2*	.	**1**	724
NW	43	127	22	**15**	*7*	179	33	**16**	*1*	38	*7*	**6**	.	*4*	*4*	**3**	505
GL	**48**	**44**	**3**	**6**	**10**	**149**	**8**	**11**	**.**	**101**	**8**	**7**	**.**	**22**	**1**	**4**	**422**
ZG	73	48	19	**6**	24	197	17	**13**	*2*	108	*14*	**13**	.	49	*10*	**10**	603
FR	298	232	75	**27**	168	596	87	**35**	87	662	90	**28**	55	683	109	**27**	3259
SO	279	73	17	**17**	135	257	58	**27**	41	260	67	**26**	20	240	49	**26**	1592
BS	**3**	**.**	**.**	**.**	**.**	**.**	**.**	**1**	**.**	**.**	**.**	**1**	**2**	**2**	**3**	**1**	**13**
BL	182	33	*4*	**7**	89	164	17	**15**	28	192	22	**13**	24	163	26	**17**	996
SH	151	*4*	*3*	**2**	90	24	15	**9**	50	66	*14*	**11**	43	114	30	**19**	645
AR	116	126	27	**21**	*7*	195	62	**17**	*2*	126	40	**6**	*1*	40	18	**5**	809
AI	36	113	44	**14**	*2*	130	76	**21**	*1*	45	29	**9**	.	*9*	*6*	**4**	539
SG	620	552	228	**31**	85	1405	247	**33**	19	815	155	**33**	*4*	267	74	**25**	4593
GR	511	200	*13*	**17**	148	510	18	**7**	33	677	34	**13**	19	475	36	**12**	2723
AG	1017	209	67	**38**	342	612	124	**53**	89	506	125	**37**	48	349	71	**37**	3724
TG	514	136	195	**38**	273	678	94	**32**	79	457	71	**24**	37	283	50	**22**	2983
TI	620	68	*14*	**12**	96	89	23	**20**	24	66	15	**10**	21	67	*14*	**12**	1171
VD	840	82	37	**29**	359	313	24	**16**	295	590	41	**22**	340	1101	69	**46**	4204
VS	2248	536	*5*	**21**	265	370	*9*	**17**	78	171	*7*	**10**	57	173	*6*	**29**	4002
NE	123	22	*5*	**4**	43	79	*2*	**7**	19	148	*12*	**10**	22	411	20	**20**	947
GE	**148**	**2**	**3**	**1**	**59**	**3**	**3**	**1**	**47**	**5**	**2**	**.**	**106**	**32**	**8**	**10**	**430**
JU	118	31	*8*	**7**	22	100	*4*	**10**	29	163	16	**15**	24	482	74	**36**	1139
Total	10514	6640	1738	520	3366	12585	3285	613	1232	8079	1787	455	961	6275	1089	495	59634

deshalb um diese „fetten" Schichten nicht zu kümmern. In Tabelle 6 sind andere Schichten mit weniger als 15 Betrieben kursiv dargestellt [2]. Diese kursiven Schichten wurden mit anderen Schichten zusammengelegt. Dabei erfolgte die Zusammenfassung jeweils innerhalb der gleichen Zeile (Kanton) und innerhalb der gleichen Tierklasse (Endziffer von stra2)[3]. Die Zusammenlegung der Schichten und die daraus resultierende endgültige Variable zur Definition der Sekundärstraten, $stra2_{v1}$, ist in Tabelle 7 dargestellt. Für die weiteren Analysen - insbesondere der Berechnung der Stichprobengrössen - wurden die resultierenden aggregierten Sekundärschichten $stra2_{v1}$ verwendet.

2.3 Festlegung der Stichprobengrösse

Für jede der drei Zielvariablen (ln, ri, sc) wurde mittels einer optimalen Allokation nach Neyman gemäss [2], Seite 105, und der vorgestellten Schichtung ein Stichprobenplan nach folgenden Kriterien erstellt:

- In jeder Schicht wird eine einfache Zufallsstichprobe gezogen.
- Das Total auf Niveau Kanton wird mit einem Variationskoeffzienten von 5% geschätzt.
- Das vorgegebene Präzisionsziel (CV 5%) wird mit einer möglichst kleinen Stichprobengrösse erreicht.

[2] Geht man davon aus, dass die Anzahl Antwortenden in einer Schicht binomialverteilt ist und die Antwortwahrscheinlichkeit 60% (was konservativ ist) beträgt, so ist die Wahrscheinlichkeit in einer vollerhobenen Schicht mit 15 Betrieben fünf oder weniger Beobachtungen zu haben ungefähr 3.4% - also relativ klein. Zu beachten ist zudem, dass der Grossteil der nach der Zusammenlegung resultierenden Schichten wesentlich mehr als 15 Betriebe beinhaltet.

[3] Diese Wahl wurde getroffen, weil entsprechende Tests vermuten liessen, dass insbesondere die Schätzung der Anzahl Schweine sehr empfindlich auf eine Zusammenlegung von Straten mit verschiedenen Ausprägungen von ti_{kl} (Endziffer von stra2) reagieren kann.

Tabelle 7 Zusammenlegung kleiner Schichten.

Kanton	stra2	$\rightarrow$	$stra2_{v1}$
LU	20, 30, 40	$\rightarrow$	20
UR	20, 30	$\rightarrow$	20
UR	22, 32	$\rightarrow$	22
UR	31, 41	$\rightarrow$	31
SZ	20,30, 40	$\rightarrow$	20
OW	22, 32	$\rightarrow$	22
OW	31, 41	$\rightarrow$	31
NW	22, 32, 42	$\rightarrow$	22
NW	20, 30, 10	$\rightarrow$	10
NW	31, 41	$\rightarrow$	31
ZG	32, 42	$\rightarrow$	32
ZG	20, 30	$\rightarrow$	20
BL	12, 22	$\rightarrow$	12
SH	11, 21	$\rightarrow$	11
SH	12, 22, 32	$\rightarrow$	12
AR	10, 20, 30,40	$\rightarrow$	10
AI	20, 30,10	$\rightarrow$	10
AI	31, 41	$\rightarrow$	31
AI	32, 42	$\rightarrow$	32
SG	30,40	$\rightarrow$	30
GR	12,22	$\rightarrow$	12
TI	32,42	$\rightarrow$	32
TI	12, 22	$\rightarrow$	12
VS	12,22,32,42	$\rightarrow$	12
NE	12,22,32	$\rightarrow$	12
JU	12,22,32	$\rightarrow$	12

Dabei wurden noch folgende Eigenheiten berücksichtigt:

- Betriebe mit vielen Tieren, das heisst Sekundärschichten ($stra2_{v1}$) mit Endziffer 3 werden grundsätzlich voll erhoben
- Kantone mit weniger als 500 Betrieben werden grundsätzlich voll erhoben (betrifft BS, GE und GL)

Die Ermittlung der nötigen Stichprobengrössen erfolgte mittels einem durch den Dienst statistische Methoden zu diesem Zweck entwickelten SAS-Makro [4], angewandt auf die Hilfsvariablen sc, ri und ln. Auf diese Weise erhalten wir in einem ersten Schritt drei Stichprobenpläne (für jede Variable einen) und somit für jede Schicht $h \in kt \times stra2_{v1}$ drei Stichprobengrössen. Bezeichnen wir für die Schicht h diese Stichprobengrössen mit $n_h(sc)$, $n_h(ri)$ und $n_h(ln)$, so wurde als eigentliche Stichprobengrösse für die Schicht h, das Maximum aus $n_h(sc)$, $n_h(ri)$, $n_h(ln)$ gewählt, das heisst:

$$n_{net,h} := max(n_h(sc), n_h(ri), n_h(ln)). \quad (1)$$

[4] *alloc_cv_gegeben*, Version 2.34. Als Parameter für die minimale Stichprobengrösse je Schicht wurde 30 gesetzt. Das heisst, Schichten mit weniger als 30 Betrieben wurden voll erhoben. Zudem wurde als obere Grenze für das Stichprobengewicht 40 gesetzt - die minimale Erhebungsrate je Schicht ist somit $\frac{1}{40}$.

Auf diese Weise resultiert ein Stichprobenplan mit 11983 Betrieben, der ohne Antwortausfälle für alle drei Variablen Schätzungen mit einem CV von höchstens 5% garantiert. In Tabelle 8 sind die resultierenden Nettostichprobengrössen nach Kanton aufgelistet und mit n_{net} bezeichnet.

Tabelle 8 Stichprobengrössen nach Kantonen.

kt	n_{brut1}	n_{brut2}	n_{brut3}	n_{net}
ZH	1019	862	756	705
BE	1178	1003	883	825
LU	831	708	625	586
UR	312	275	244	229
SZ	628	536	463	430
OW	342	295	259	243
NW	346	305	264	244
GL	422	422	422	422
ZG	421	375	341	325
FR	1060	887	758	696
SO	864	727	625	578
BS	13	13	13	13
BL	541	474	428	407
SH	488	448	407	385
AR	561	486	419	388
AI	406	358	320	301
SG	908	756	652	603
GR	604	523	469	443
AG	1052	875	756	698
TG	1034	859	739	682
TI	602	524	463	434
VD	945	798	693	642
VS	653	550	474	439
NE	556	479	422	395
GE	430	430	430	430
JU	599	519	465	440
Total	16815	14487	12790	11983

2.3.1 Antwortausfälle

In den letzten beiden Zusatzerhebungen (2003, 2005) wurden glücklicherweise hohe Antwortraten im Bereich 90% erzielt. Gemäss Fachexperten hat sich das Umfeld für die Landwirtschaftsbetriebe wesentlich geändert, was die Bereitschaft zum Ausfüllen des Fragebogens stark reduzieren könnte. Hinzu kommt, dass der aktuelle Fragebogen viel umfangreicher ist als bei den Vorgängererhebungen. Aus diesem Grunde wurden die folgenden alternativen Szenarien zu den Antwortraten θ angenommenen: 75 $\%$ und 60 $\%$. Die Nettostichprobengrössen $n_{net,h}$ werden nun schichtweise mit den Antwortraten θ korrigiert:

$$n_h = \min(\frac{1}{\theta} n_{net,h}, N_h) \tag{2}$$

was in Abhängigkeit des Antwortszenarios $\theta = 60/75/90\%$ zu den in Tabelle 8 dargestellten Bruttostichprobengrössen n_{brut1}, n_{brut2} und n_{brut3} führt. Aus (2) wird klar, dass die Korrektur

Tabelle 9 Schichtweise Stichprobengrössen

Kanton	$stra2_{v1}$																
	10	11	12	13	20	21	22	23	30	31	32	33	40	41	42	43	Total
ZH	88	50	44	28	52	92	50	30	50	122	53	37	88	142	64	29	1019
BE	50	97	50	65	50	160	102	117	50	75	60	50	46	98	50	58	1178
LU	50	52	50	65	50	100	105	74	.	50	55	45	.	50	50	35	831
UR	50	65	21	13	16	70	25	8	.	39	.	4	.	.	.	1	312
SZ	50	67	56	22	41	137	50	15	.	67	30	19	.	50	18	6	628
OW	50	63	28	14	.	93	32	8	.	47	.	6	.	.	.	1	342
NW	50	50	22	15	.	98	44	16	.	42	.	6	.	.	.	3	346
GL	48	44	3	6	10	149	8	11	.	101	8	7	.	22	1	4	422
ZG	50	48	19	6	26	93	17	13	.	53	24	13	.	49	.	10	421
FR	50	50	75	27	50	60	50	35	50	95	50	28	50	293	70	27	1060
SO	58	50	17	17	50	63	50	27	41	92	65	26	20	213	49	26	864
BS	3	.	.	.	.	.	.	1	.	.	.	1	2	2	3	1	13
BL	50	33	21	7	50	50	.	15	28	50	22	13	24	135	26	17	541
SH	50	28	32	2	50	.	.	9	50	50	.	11	43	114	30	19	488
AR	105	50	27	21	.	95	62	17	.	75	40	6	.	40	18	5	561
AI	39	55	44	14	.	82	50	21	.	53	35	9	.	.	.	4	406
SG	60	50	125	31	50	138	50	33	23	107	50	33	.	83	50	25	908
GR	50	50	31	17	50	50	.	7	33	72	34	13	19	130	36	12	604
AG	135	50	50	38	50	87	60	53	50	105	55	37	48	147	50	37	1052
TG	87	50	138	38	50	113	50	32	50	103	50	24	37	140	50	22	1034
TI	160	50	37	12	50	50	.	20	24	60	29	10	21	67	.	12	602
VD	50	50	37	29	50	50	24	16	50	55	41	22	72	295	58	46	945
VS	172	50	27	21	50	57	.	17	50	50	.	10	57	63	.	29	653
NE	50	22	19	4	43	50	.	7	19	50	.	10	22	220	20	20	556
GE	148	2	3	1	59	3	3	1	47	5	2	.	106	32	8	10	430
JU	50	31	28	7	22	50	.	10	29	50	.	15	24	197	50	36	599
Total	1803	1207	1004	520	919	1990	832	613	644	1668	703	455	679	2582	701	495	16815

der Stichprobengrösse durch die Anzahl Betriebe in der betreffenden Schicht, N_h, begrenzt wird. In Abhängigkeit von den Annahmen zu den Antwortausfällen bewegen sich die Bruttostichprobengrössen in einer Grössenordnung zwischen 12000 und 17000. Für den definitiven Stichprobenplan wurde die konservative Variante $\theta = 0.6$ gewählt. Die daraus resultierenden schichtweisen Stichprobengrössen sind in Tabelle 9 dargestellt.

2.4 Zu erwartende Präzisionen

Der in Tabelle 8 vorgestellte Plan für die Nettostichprobe garantiert uns für die Variablen ln, ri, sc Schätzungen mit einem CV von höchstens 5% auf Kantonsebene. Es stellt sich die Frage, wie hoch die geplanten Variationskoeffizienten nun effektiv ausfallen. Die entsprechenden Grössen cv_{ln}, cv_{ri}, cv_{sc} wurden auf Kantonsebene für den Horvitz-Thompson Schätzer berechnet: Für die Berechnungen eines gewissen Kantones, wurde dieser als Untersuchungsbereich d definiert. Weiter bezeichnen wir mit $\hat{Y}_{HT,d}$ den Horvitz-Thompson Schätzer für das Total Y_d der interessierenden Variablen y im Kanton d. Der Variationskoeffizient (CV=Coefficient of Variation) des Schätzers wurde nun aufgrund der Werte im Stichprobenrahmen wie folgt berechnet:

$$\mathbf{CV}(\hat{Y}_{HT,d}) = \frac{\sqrt{\mathbf{Var}(\hat{Y}_{HT,d})}}{Y_d} \quad (3)$$

$$\mathbf{Var}(\hat{Y}_{HT,d}) = \sum_{h=1}^{H} \frac{N_h^2}{n_{net,h}}(1 - \frac{n_{net,h}}{N_h})S^2_{y_d h}, \quad (4)$$

Tabelle 10 Variationskoeffizienten basierend auf der geplanten Nettostichprobengrösse

Kanton	CV in Prozent				
	ln	ri	sc	gvetot	fa
ZH	0.9	2.2	0.8	2.8	3.4
BE	0.9	2.3	4.2	2.1	2.7
LU	1.0	2.3	4.4	2.3	3.1
UR	1.7	2.6	0.0	2.4	4.4
SZ	1.0	2.0	2.7	2.0	3.2
OW	1.5	2.3	0.0	2.1	4.0
NW	1.2	1.9	1.6	2.0	3.3
GL	0.0	0.0	0.0	0.0	0.0
ZG	0.8	1.5	0.0	1.5	3.5
FR	1.0	2.0	2.3	1.9	3.2
SO	0.8	1.8	2.3	2.0	2.9
BS	0.0	0.0	0.0	0.0	0.0
BL	1.0	2.0	0.0	2.7	3.4
SH	0.7	1.5	0.0	1.8	3.0
AR	0.9	1.4	1.1	1.2	3.1
AI	1.0	1.5	2.3	1.7	3.6
SG	1.0	2.0	4.0	2.1	3.3
GR	1.2	2.5	0.0	2.1	3.9
AG	0.9	2.2	3.5	2.7	3.0
TG	0.9	2.1	3.2	2.5	3.1
TI	1.2	1.5	0.0	3.7	5.2
VD	1.4	2.9	1.0	3.3	3.9
VS	2.0	2.8	0.0	5.0	5.6
NE	1.8	2.2	0.0	2.3	3.6
GE	0.0	0.0	0.0	0.0	0.0
JU	1.8	2.5	3.3	2.2	3.2
Median	1.0	2.0	0.9	2.1	3.2

wobei $S^2_{y_d h}$ die Varianz der Variablen y_d bezeichnet, welche für den Betrieb i wie folgt definiert ist:

$$y_{d,i} = \begin{cases} y_i & i \in d \\ 0 & i \notin d \end{cases}$$

Die Resultate sind in Tabelle 10 aufgelistet. Alle Werte bewegen sich im Intervall 0-4.4$\%$. Der Plan wurde auf die Variablen ln, ri, sc ausgerichtet. Weitere Fragen im Fragebogen betreffen die Anzahl Mitarbeiter oder hängen nicht direkt mit der Anzahl Schweine / Rinder sondern eher mit der Anzahl Tiere im Allgemeinen zusammen. Aus diesem Grunde interessierte auch, wie gut anhand des Plans die Variablen fa = Anzahl mitarbeitende Familienmitglieder sowie gvetot = Grossviehäquivalent - ein Grössenmass, das sämtliche Tiere zusammenfasst - geschätzt werden können[5]. Die entsprechenden Variationskoeffizienten des Horvitz-Thompson Schätzers findet man in Tabelle 10 unter cv_{gvetot} und cv_{fa}. Wie erwartet werden darf, sind die Werte zwar etwas höher als bei jenen Variablen, für welche der Plan optimiert wurde, aber sie sind im Grossen und Ganzen mit unserem Präzisionsziel von 5 $\%$ vereinbar. Hingegen ist zu beachten, dass der Fragebogen sehr vielfältig ist, und es auch Variablen geben wird, die weniger gut geschätzt werden können.

2.4.1 Der Einfluss der Antwortausfälle auf die geplanten Präzisionen

Obwohl der Stichprobenplan mit dem Ziel konstruiert wurde, auf Niveau Kanton Schätzungen mit einem CV von höchstens 5$\%$ zu erhalten, ist diese Präzision selbst für jene Variablen, für welche der Plan optimiert wurde, bei der effektiven Erhebung nicht garantiert. Der massgebende Grund ist, dass der Korrektur der Stichprobengrösse für Antwortausfälle durch die Schicht-

[5] Beide Variablen befinden sich im Stichprobenrahmen, was diese Präzisionsbetrachtungen ermöglicht.

Tabelle 11 Variationskoeffizienten von sc basierend auf verschiedenen Antwortszenarien

Kanton	CV in Prozent			
	cv_{net}	cv_{sz1}	cv_{sz2}	cv_{sz3}
ZH	0.8	3.8	6.2	9
BE	4.2	4.2	4.4	4.5
LU	4.4	4.4	4.4	4.5
UR	0	12.4	23.3	33.6
SZ	2.7	6.7	7.8	11.7
OW	0	4.1	7.1	11.1
NW	1.6	4.8	8.1	11.1
GL	0	23.5	24.5	38
ZG	0	4.7	7.2	10.6
FR	2.3	3.1	4.3	5.8
SO	2.3	4.4	6.1	8.6
BS	0	0	18.7	18.7
BL	0	8.1	13	18.2
SH	0	4.9	8.6	13.5
AR	1.1	6.5	11	15.4
AI	2.3	3.5	5.6	8.4
SG	4	4.2	4.4	4.7
GR	0	6	10.3	14.8
AG	3.5	4	4.6	5.6
TG	3.2	3.4	3.8	4.3
TI	0	15.9	30.2	44.3
VD	1	4.7	7.5	10.8
VS	0	16.4	25.3	40.3
NE	0	5.8	12.2	18.8
GE	0	8.1	12.2	19.2
JU	3.3	8.4	13.1	17.8
Mittel	1.4	6.8	10.9	15.5
Median	0.9	4.7	7.9	11.4

grössen N_h Grenzen gesetzt sind, und dass die Anzahl effektiver Antworten unter der berechneten Nettostichprobengrösse (1) liegen kann. Als Illustration sind in Tabelle 11 die Variationskoeffizienten des Horvitz-Thompson Schätzers für die Variable sc berechnet ($cv_{sz1}, cv_{sz2}, cv_{sz3}$), die sich ergäben, falls die verschiedenen betrachteten Antwortszenarien ($\theta = 0.9/0.75/0.6$) effektiv eintreffen würden[6]. Die Werte in Tabelle 11 zeigen, dass die Präzisionen für gewisse Kantone sehr empfindlich auf Antwortausfälle reagieren. Dies kann durch eine Aufblähung der Bruttostichprobengrösse nicht verhindert werden, wie beispielsweise anhand des voll erhobenen Kantons Glarus beobachtet werden kann. Trotz der Vollerhebung, die ohne Antwortausfälle auf einen Variationskoeffizienten von Null ($cvnet = 0\%$) führt, bewirken die Antwortausfälle drastische Erhöhungen des Variationskoeffizienten. Aus diesem Grund muss alles Mögliche getan werden, um Antwortausfälle zu vermeiden.
Schlussfolgerung: Der negative Effekt von Antwortausfällen auf die Präzision der Resultate kann durch eine Erhöhung der Stichprobe nur beschränkt kompensiert werden. Es müssen daher alle Möglichkeiten ausgeschöpft werden, um ein Absinken der Antwortraten zu vermeiden.

3 Gewichtung und Hochrechnung

Ausgangspunkt für die Berechnung des Hochrechnungsgewichts ist das Stichprobengewicht, welches für jeden Betrieb i anhand seiner Schichtzugehörigkeit $i \in h$ wie folgt definiert ist:

$$w_{1,i} := \frac{N_h}{n_n}. \tag{5}$$

[6]Die Berechnung erfolgte analog (3), nur wurden in (4) die Werte von $n_{net,h}$ durch auf ganze Zahlen gerundete Werte von θn_h (erwartete Anzahl Antworten) ersetzt.

Dabei ist n_h die Stichprobengrösse gemäss (2) und N_h die entsprechende Schichtgrösse. Einige weitere für die folgenden Abschnitte wichtigen Terme werden in Tabelle 12 aufgelistet.

Tabelle 12 Abkürzungen und Notationen.

Abkürzung/Term	Erläuterung
$stra2_{v2}$	Grössenklassen, die im Rahmen der Hochrechnung zur Konstruktion von Antworthomogenitätsklassen $kt \times stra2_{v2}$ verwendet werden.
m_h	erzielte Nettostichprobengrösse in $h \in kt \times stra2_{v2}$ (Anzahl Antworten, die für die Hochrechnung zur Verfügung stehen.)
θ_h	Antwortrate in $h \in kt \times stra2_{v2}$
w_1	Stichprobengewicht
w_2	Gewicht nach Korrektur für Antwortausfälle
w_3	Hochrechnungsgewicht nach Kalibrierung

In den folgenden beiden Abschnitten wird beschrieben, wie ausgehend vom Stichprobengewicht das endgültige Hochrechnungsgewicht berechnet wurde.

3.1 Rücklauf und Gewichtskorrektur für Antwortausfälle

Erfreulicherweise erwies sich die bei der Planung zu Grunde gelegte Annahme einer globalen Antwortrate von 0.6 als zu pessimistisch. Insgesamt standen am Ende des Datenaufbereitungsprozesses die Daten für eine Menge r (r=respondents) von 15169 Betriebe zur Verfügung. Verglichen mit der Bruttosichprobengrösse von 16815 ergibt sich also eine globale Antwortrate von rund 90%. Es gilt aber zu beachten, dass die erzielten Antwortraten von Schicht zu Schicht variieren, so dass der Einfluss der Antwortraten je nach betrachteter Schicht unterschiedlich zu beurteilen ist. Die Anzahl eingegangener Antworten je Schicht $kt \times stra2_{v1}$ ist in Tabelle 13 dargestellt. Für die Hochrechnung wurden nun Schichten mit weniger als 5 Beobachtungen und Antwortrate $< 100\%$ mit anderen Schichten des gleichen Kantones zusammengefasst. Die betreffenden Schichten - alle davon wurden vollerhoben - sind in Tabelle 13 kursiv hervorgehoben. Die Zusammenlegung und die dadurch resultierende neue Schichtungsvariable $stra2_{v2}$ ist in Tabelle 14 beschrieben. Dabei gilt für Straten, welche nicht zusammengelegt wurden: $stra2_{v2} = stra2_{v1}$. Für jede aus dieser Zusammenlegung resultierenden Schicht $h \in Kanton \times stra2_{v2}$ wird nun eine Antwortrate

$$\theta_h = \frac{m_h}{n_h}$$

berechnet. Dabei bezeichnet m_h die Anzahl antwortenden Unternehmen in h. Die resultierenden Werte sind in Tabelle 15 dargestellt. Unter der Annahme, dass die Unternehmen in h eine identische Antwortwahrscheinlichkeit besitzen und somit eine Antworthomogenitätsklasse bilden, wird nun ein zweites, an Antwortausfälle angepasstes Gewicht (w_2) berechnet. Für ein beliebiges Unternehmen i, welches geantwortet hat und sich in h befindet, gilt:

$$w_{2,i} := \theta_h^{-1} w_{1,i} \tag{6}$$

Zu beachten ist, dass sich die Gewichte w_2 aufgrund des gewählten Stichprobenplanes und der relativ hohen Antwortquoten in einem recht moderaten Bereich bewegen: das Maximum liegt bei 26.55 ($stra2_{v2} = 11$, Kanton Bern).

Tabelle 13 Anzahl Antworten je Schicht

Kanton	$stra2_{v1}$																
	10	11	12	13	20	21	22	23	30	31	32	33	40	41	42	43	Total
ZH	70	45	37	27	46	81	47	27	46	113	50	35	80	133	63	27	927
BE	40	86	48	58	46	148	99	114	45	70	59	50	40	92	49	54	1098
LU	40	50	46	61	45	93	101	68	.	49	54	41	.	46	48	30	772
UR	45	59	19	11	16	66	23	7	.	39	.	4	.	.	.	1	290
SZ	43	57	52	19	36	129	46	14	.	65	26	18	.	45	18	6	574
OW	38	55	26	10	.	89	32	8	.	46	.	6	.	.	.	1	311
NW	42	45	20	15	.	96	42	16	.	40	.	6	.	.	.	3	325
GL	34	36	3	5	9	135	8	10	.	97	7	6	.	22	1	4	377
ZG	43	39	17	6	25	85	16	12	.	51	23	11	.	42	.	9	379
FR	31	42	59	26	44	56	47	34	42	85	48	24	45	276	67	26	952
SO	39	45	15	16	43	61	47	25	38	91	63	25	20	203	46	26	803
BS	3	.	.	.	.	.	.	1	.	.	.	1	2	2	3	1	13
BL	46	29	20	6	43	48	.	15	20	44	20	13	21	123	22	17	487
SH	38	27	31	2	46	.	.	7	45	46	.	10	40	103	28	18	441
AR	88	46	22	18	.	88	56	15	.	68	37	6	.	37	16	5	502
AI	31	47	40	11	.	71	45	21	.	45	34	*7*	.	.	.	*3*	355
SG	50	44	96	25	46	124	49	26	22	97	48	33	.	78	47	24	809
GR	40	36	30	10	45	46	.	6	32	68	31	13	18	125	36	10	546
AG	110	46	41	36	43	82	58	51	47	99	51	32	43	137	46	32	954
TG	68	39	109	36	46	107	47	29	47	95	49	21	30	133	47	21	924
TI	114	40	30	12	44	45	.	16	20	57	24	9	19	59	.	9	498
VD	35	40	27	25	43	45	24	15	45	54	38	20	70	274	54	43	852
VS	145	41	25	20	49	51	.	16	43	46	.	10	46	60	.	27	579
NE	43	18	17	*3*	34	43	.	*7*	16	46	.	10	21	199	20	17	494
GE	113	2	3	*0*	57	*2*	3	1	43	*5*	2	.	101	27	7	*8*	374
JU	31	25	25	7	22	46	.	9	26	45	.	14	23	177	49	34	533
Total	1420	1039	858	465	828	1837	790	570	577	1561	664	425	619	2393	667	456	15169

Tabelle 14 Zusammenlegung kleiner Schichten im Rahmen der Hochrechnung.

kt	$stra2_{v1}$	$\rightarrow$	$stra2_{v2}$
AI	33, 43	$\rightarrow$	33
NE	13,23	$\rightarrow$	13
GE	13, 43	$\rightarrow$	13
GE	21, 31	$\rightarrow$	31

3.2 Kalibrierung

Die Antwortenden der Zusatzerhebung bilden eine Teilmenge der insgesamt 59065 Betriebe in der landwirtschaftlichen Strukturhebung 2010. Die Strukturerhebung umfasst eine Vielzahl von Variablen rund um den Betrieb, welche dazu verwendet werden können, die Schätzresultate der Zusatzerhebung zu verbessern. Dabei standen folgende Ziele im Vordergrund:

1. Die Hochrechnung der Zusatzerhebung soll Resultate liefern, die zu wichtigen Variablen aus der landwirtschaftlichen Strukturerhebung 2010 auf Niveau Kanton konsistent sind. Implizit bedingt dies, dass die Kalibrierung eine Gewichtsanpassung berechnet, welche auch den leichten Veränderungen in der Populationsstruktur zwischen den Jahren 2009 (Basis für die Stichprobenziehung und somit der Stichprobengewichte) und 2010 Rechnung trägt.

2. Die Anzahl für die Kalibrierung verwendeten Variablen soll so sein, dass die für die Varianzschätzungen benötigten Regressionsparameter noch mit genügender Präzision geschätzt werden können.

3. Die Varianz der Hochrechnungen soll reduziert werden.

Die Erfüllung der gestellten Ziele wurde nun folgendermassen angestrebt:

Tabelle 15 Antwortraten in Antworthomogenitätsklassen.

Kanton	$stra2_{v2}$															
	10	11	12	13	20	21	22	23	30	31	32	33	40	41	42	43
ZH	0.80	0.90	0.84	0.96	0.88	0.88	0.94	0.90	0.92	0.93	0.94	0.95	0.91	0.94	0.98	0.93
BE	0.80	0.89	0.96	0.89	0.92	0.93	0.97	0.97	0.90	0.93	0.98	1.00	0.87	0.94	0.98	0.93
LU	0.80	0.96	0.92	0.94	0.90	0.93	0.96	0.92	.	0.98	0.98	0.91	.	0.92	0.96	0.86
UR	0.90	0.91	0.90	0.85	1.00	0.94	0.92	0.88	.	1.00	.	1.00	.	.	.	1.00
SZ	0.86	0.85	0.93	0.86	0.88	0.94	0.92	0.93	.	0.97	0.87	0.95	.	0.90	1.00	1.00
OW	0.76	0.87	0.93	0.71	.	0.96	1.00	1.00	.	0.98	.	1.00	.	.	.	1.00
NW	0.84	0.90	0.91	1.00	.	0.98	0.95	1.00	.	0.95	.	1.00	.	.	.	1.00
GL	0.71	0.82	1.00	0.83	0.90	0.91	1.00	0.91	.	0.96	0.88	0.86	.	1.00	1.00	1.00
ZG	0.86	0.81	0.89	1.00	0.96	0.91	0.94	0.92	.	0.96	0.96	0.85	.	0.86	.	0.90
FR	0.62	0.84	0.79	0.96	0.88	0.93	0.94	0.97	0.84	0.89	0.96	0.86	0.90	0.94	0.96	0.96
SO	0.67	0.90	0.88	0.94	0.86	0.97	0.94	0.93	0.93	0.99	0.97	0.96	1.00	0.95	0.94	1.00
BS	1.00	.	.	.	.	.	.	1.00	.	.	.	1.00	1.00	1.00	1.00	1.00
BL	0.92	0.88	0.95	0.86	0.86	0.96	.	1.00	0.71	0.88	0.91	1.00	0.88	0.91	0.85	1.00
SH	0.76	0.96	0.97	1.00	0.92	.	.	0.78	0.90	0.92	.	0.91	0.93	0.90	0.93	0.95
AR	0.84	0.92	0.81	0.86	.	0.93	0.90	0.88	.	0.91	0.93	1.00	.	0.93	0.89	1.00
AI	0.79	0.85	0.91	0.79	.	0.87	0.90	1.00	.	0.85	0.97	0.77	.	.	.	.
SG	0.83	0.88	0.77	0.81	0.92	0.90	0.98	0.79	0.96	0.91	0.96	1.00	.	0.94	0.94	0.96
GR	0.80	0.72	0.97	0.59	0.90	0.92	.	0.86	0.97	0.94	0.91	1.00	0.95	0.96	1.00	0.83
AG	0.81	0.92	0.82	0.95	0.86	0.94	0.97	0.96	0.94	0.94	0.93	0.86	0.90	0.93	0.92	0.86
TG	0.78	0.78	0.79	0.95	0.92	0.95	0.94	0.91	0.94	0.92	0.98	0.88	0.81	0.95	0.94	0.95
TI	0.71	0.80	0.81	1.00	0.88	0.90	.	0.80	0.83	0.95	0.83	0.90	0.90	0.88	.	0.75
VD	0.70	0.80	0.73	0.86	0.86	0.90	1.00	0.94	0.90	0.98	0.93	0.91	0.97	0.93	0.93	0.93
VS	0.84	0.82	0.93	0.95	0.98	0.89	.	0.94	0.86	0.92	.	1.00	0.81	0.95	.	0.93
NE	0.86	0.82	0.89	0.91	0.79	0.86	.	.	0.84	0.92	.	1.00	0.95	0.90	1.00	0.85
GE	0.76	1.00	1.00	0.73	0.97	0.88	1.00	1.00	0.91	.	1.00	.	0.95	0.84	0.88	.
JU	0.62	0.81	0.89	1.00	1.00	0.92	.	0.90	0.90	0.90	.	0.93	0.96	0.90	0.98	0.94

- Die Kalibrierung erfolgt grundsätzlich auf Stufe Kanton, einzige Ausnahme bilden die beiden Basler Halbkantone, die aufgrund der geringen Anzahl Betriebe von Basel Stadt zusammengefügt werden. Das heisst, die Gewichtsanpassungen werden für jeden Kanton separat in jenem Sinne vorgenommen, dass die Konsistenz zwischen Resultaten aus der Zusatzerhebung und der Vollerhebung für die verwendeten Kalibrierungsvariablen erfüllt ist. Es folgt dann automatisch, dass dies auch für jedes höhere Aggregat (zum Beispiel Schweiz) der Fall sein wird.

- Die Varianzschätzung basiert auf kantonsweisen Schätzungen von Regressionskoeffizienten, welche den linearen Zusammenhang zwischen den verwendeten Hilfsvariablen und den Zielvariablen beschreiben. Es sollten genügend Beobachtungen da sein, um die Regressionkoeffizienten genügend stabil schätzen zu können. Dabei hängt diese Stabilität vom Verhältnis Anzahl Regressionkoeffizienten vs. Anzahl Beobachtungen ab. Betrachten wir die Grössen der Nettostichproben in Tabelle 13, so beobachten wir die kleinste Fallzahl für den Kanton Uri mit 290 Betrieben. (Basel Stadt spielt aufgrund der Zusammenlegung mit Basel Land keine Rolle). Bei der Anzahl für die Kalibrierung verwendeter Hilfsvariablen wird somit das Verhältnis zu 290 im Auge behalten. Dies impliziert, dass nicht beliebig viele - oder gar alle - Variablen aus der Strukturerhebung für die Kalibrierung verwendet werden. Es wurde schliesslich so vorgegangen, dass für jeden Themenblock zumindest eine Variable in die Kalibrierung einfliesst, für welche Konsistenz zwischen hochgerechneten Resultaten und Werten der Strukturerhebung wichtig ist und welche zumindest für einige wichtige Erhebungsvariablen eine Varianzreduktion der Schätzungen erzielt.

Die letztendlich für die Kalibrierung verwendeten Variablen sind in Tabelle 16 aufgeführt.

Tabelle 16 In der Kalibrierung verwendete Hilfsvariablen.

Variablendefinition gem. Stru2010	Variable	Erläuterungen
ktakt	Kanton des Betriebes	Stellt sicher, dass die hochgerechnete Anzahl Betriebe je Kanton der Anzahl Betriebe in der Strukturerhebung 2010 entspricht
ln	Landwirtschaftliche Nutzfläche	Allgemeine Grössencharakteristik der Betriebe, welche einen Zusammenhang mit einer Vielzahl von Variablen aufweist
X0542 + X0545 + X0546 + X0547 + X0551	Gemüsefläche	Korreliert mit Variablen im Bereich Bewässerung
x0599	offene Ackerfläche	Korrelation zu Variablen wie Düngerausbringung
xbesch	Anzahl Beschäftigte	Wichtige Grössencharakteristik im Zusammenhang zu Variablen betreffend Beschäftigung
GVE_SCHWE_09	Grossvieheinheiten Schweine	Wichtige Grössencharakteristik für Fragen im Zusammenhang mit Stallsystemen für Schweine
GVE_RINDE_09	Grossvieheinheiten Rindvieh	Wichtige Grössencharakteristik für Fragen im Zusammenhang mit Stallsystemen für Rinder und Kühe

Die Auswahl beinhaltet 7 Variablen, was auch im Kanton Uri noch zu einem vernünftigen Verhältnis zwischen Anzahl Beobachtungen und Anzahl Regressionsparameter führt $(\lfloor 290/7 \rfloor = 41)$. Die Kalibrierung wurde mittels vom INSEE [7] entwickelten SAS-Macro CALMAR [8] durchgeführt. Die Kalibrierung liefert für jedes Unternehmen i einen Korrekturfaktor g_i, der vom Gewicht $w_{2,i}$ auf die definitiven Hochrechnungsgewichte $w_{3,i}$ führt:

$$w_{3,i} = g_i w_{2,i}. \tag{7}$$

Tabelle 17 beschreibt die Verteilung der erzielten Gewichte w_3 sowie der Korrekturfaktoren g_i in den 15169 Unternehmen der Nettostichprobe. Es zeigt sich, dass die Korrekturfaktoren für 98% der Betriebe in einem moderaten Bereich zwischen 0.76 und 1.23 liegen und die Gewichte nicht zu stark verändern. Das Maximum und das Minimum zeigen aber, dass es

[7] Institut national de la statistique et des études économiques.

[8] Für die Kalibrierung wurde die Methode Raking Ratio verwendet, womit sichergestellt ist, dass die Gewichte nach Kalibrierung alle positiv sind.

Gewichtsveränderung mit einem Faktor der Grössenordnung 5 geben kann. Generell handelt es sich dabei um Einzelfälle und nicht etwa um ganze Schichten, welche systematisch starke Gewichtsveränderungen erfahren. Dies widerspiegelt sich in Tabelle 18, welche schichtweise mittlere Hochrechnungsgewichte w_3 auflistet. Dabei wurden - entsprechend der Kalibrierung - Basel-Stadt und Basel-Land zusammengefasst.

Tabelle 17 Die Verteilung der definitiven Hochrechnungsgewichte und der g-Faktoren.

Quantile	w_3	g-Faktoren
Max	28.40	5.50
99%	24.67	1.23
95%	12.11	1.09
90%	9.92	1.06
75%	4.37	1.02
50%	2.02	0.99
25%	1.13	0.96
10%	1.02	0.91
5%	0.97	0.87
1%	0.86	0.76
Min	0.22	0.19

3.3 Die Schätzung von Totalen und Quotienten

Die Schätzung des Totales der Untersuchungsvariablen y in der ganzen Population, Y, ergibt sich durch die gewichtete Summe der Variablen y_i der antwortenden Betriebe $i \in r$:

$$\hat{Y} = \sum_{i \in r} w_{3,i} y_i \tag{8}$$

Für die Schätzung des Totales in einem Untersuchungsbereich[9] d , Y_d, verwendet man

$$\hat{Y}_d = \sum_{i \in r} w_{3,i} 1(i \in d) y_i = \sum_{i \in r \cap d} w_{3,i} y_i \tag{9}$$

wobei

$$1(i \in d) = \begin{cases} 1 & i \in d \\ 0 & i \notin d \end{cases}$$

Nebst Totalen interessieren auch Quotienten. Falls Y und Z die Totale zweier Variablen y und z bezeichnen, so ist $Q_{Y,Z} = \frac{Y}{Z}$ der Quotient der beiden Totale. Dieser wird mit

$$\hat{Q}_{Y,Z} = \frac{\hat{Y}}{\hat{Z}} \tag{10}$$

geschätzt. Dabei sind $\hat{Y}$, $\hat{Z}$ analog zu (8) definiert. Die Schätzung eines Quotienten für den Untersuchungsbereich d wird analog konstruiert:

$$\hat{Q}_{Y_d,Z_d} = \frac{\hat{Y}_d}{\hat{Z}_d} \tag{11}$$

Dabei sind Zähler und Nenner Totalschätzungen im Sinne von (9).

[9] Beispielsweise könnte der Untersuchungsbereich einem bestimmten Kanton entsprechen.

Tabelle 18 Schichtweise mittlere Hochrechnungsgewichte

Kanton	$stra2_{v2}$															
	10	11	12	13	20	21	22	23	30	31	32	33	40	41	42	43
ZH	11.81	4.06	1.2	1.08	11.3	9.25	1.3	1.16	3.25	5.95	1.08	1.13	0.96	3.25	0.97	1.15
BE	21.65	26.23	8.86	1.09	10.08	24.53	11.12	1	2.84	21.17	8.09	0.99	1.2	6.8	3.75	1.06
LZ	9.99	13.99	7.79	1.04	2.28	11.62	10.25	1.09	.	8.41	8.45	1.13	.	3.69	3.21	1.2
UR	2.22	3.84	1.1	1.35	0.91	3.01	1.02	1.15	.	0.92	.	1	.	.	.	1.13
SZ	5.43	6.23	1.12	1.22	1.12	4.3	1.81	1.12	.	3.44	1.09	0.98	.	1.62	0.94	0.88
OW	2.32	4.15	1.06	1.33	.	2.99	0.96	1	.	1.03	.	0.94	.	.	.	0.99
NW	1.17	2.8	1.1	1.01	.	1.85	1.08	1.03	.	1.06	.	0.98	.	.	.	1.04
GL	1.44	1.25	1.03	1.19	1.12	1.1	1	1.08	.	1.02	1.11	1.08	.	0.96	0.95	0.93
ZG	1.68	1.22	1.08	0.98	1.01	2.33	1.07	1.08	.	2.14	1.06	1.16	.	1.17	.	1.1
FR	9.57	5.61	1.3	1.09	3.59	10.65	1.86	1.05	1.91	7.68	1.87	1.25	1.06	2.42	1.59	1.04
SO	7.14	1.65	1.14	1.1	3.06	4.22	1.24	1.11	1.03	2.86	1.06	1.08	0.9	1.14	1.03	1.05
BS+BL	3.72	1.13	1.06	1.13	2.04	3.42	.	1.02	1.33	4.38	1.09	1.07	1.05	1.34	1.16	1.08
SH	3.91	1.02	1.05	0.91	1.94	.	.	1.29	1.1	1.42	.	1.08	1.07	1.05	1	1.1
AR	1.47	2.83	1.26	1.18	.	2.24	1.12	1.14	.	1.83	1.06	0.98	.	1.03	1.06	0.93
AI	1.3	2.39	1.14	1.28	.	1.8	1.65	0.95	.	1.18	1.01	1.23	.	.	.	.
SG	12.13	12.7	2.45	1.34	1.66	11.45	5.14	1.32	0.82	8.28	3.26	1.08	.	3.32	1.55	1.26
GR	12.55	5.56	1.02	1.79	3.01	10.92	.	1.16	0.88	9.68	1.09	1.11	0.83	3.61	1	1.29
AG	9.48	4.64	1.69	1.11	7.75	7.57	2.16	1.05	1.94	5.09	2.43	1.15	1.03	2.43	1.57	1.09
TG	7.35	3.46	1.76	1.06	5.56	6.39	2	1.1	1.61	4.87	1.46	1.09	1.13	2.12	1.01	0.97
TI	5.42	1.73	1.23	1.02	2.1	2	.	1.28	1.11	1.18	1.18	1.14	1.01	1.13	.	1.29
VD	23.85	2.09	1.47	1.2	8.01	7	1.02	1.13	6.08	10.99	1.06	1.14	5.02	3.86	1.29	1.17
VS	14.94	12.75	1.09	1.04	5.24	7.21	.	1.07	1.76	3.79	.	1.03	1.22	3.03	.	1.16
NE	2.85	1.22	1.09	1.07	1.24	1.83	.	.	1.17	3.2	.	0.99	1.06	2.04	0.98	1.16
GE	1.37	1.03	1.04	1.27	1.05	1.15	1.02	0.98	1.09	.	0.99	.	1.03	1.12	1.23	.
JU	3.71	1.26	1.1	0.99	0.97	2.17	.	1.11	1.06	3.61	.	1.08	0.93	2.69	1.5	1.06

3.4 Die Varianz der Schätzungen

Unsere Schätzung ist asymptotisch (für grosse Werte von n) gleichwertig zum Generalized Regression Estimator (GREG) und die Varianzschätzung erfolgt analog. Die Kalibrierung erfolgte kantonsweise. Darum wird auch das dem GREG zugrunde liegende Regressionsmodell für jeden Kanton separat betrachtet[10]. Für einen Betrieb i im Kanton c lässt sich dieses Modell wie folgt schreiben:

$$y_i = \sum_{k=1..6} \beta_{c,k} x_{k,i} + \beta_{c,0} + e_i \tag{12}$$

Dabei bezeichnen $x_{k,i}, k = 1, ..., 6$, die Werte der 6 verwendeten Kalibrierungsvariablen und $\beta_{c,k}, k = 1, ..., 6$, die entsprechenden Regressionskoeffizienten. $\beta_{c,0}$ ist der Intercept und e_i das Residuum. Mit dem Intercept tragen wir der Tatsache Rechnung, dass in der Kalibrierung auch die Anzahl Betriebe im Kanton berücksichtigt worden sind. Da die Regressionskoeffizienten nicht bekannt sind, werden sie mittels WLS (Weighted Least Squares), Gewicht=$w_{2,i}$, geschätzt. Die Schätzung führt zu

$$y_i = \sum_{k=1..6} \hat{\beta}_{c,k} x_{k,i} + \hat{\beta}_0 + \hat{e}_i. \tag{13}$$

Die Residuen $\hat{e}$ werden nun für die Varianzschätzung verwendet: Für die Schätzung auf Niveau Gesamtpopulation, $\hat{Y}$ gemäss (8), wird die Varianz durch

$$\widehat{\mathbf{Var}}(\hat{Y}) = \sum_{h \in H} \frac{N_h^2}{m_h}(1 - \frac{m_h}{N_h}) \frac{1}{m_h - 1} \sum_{i=1}^{m_h} (g_i \hat{e}_i - \frac{\sum g_i \hat{e}_i}{m_h})^2 \tag{14}$$

[10] Ausnahme BS und BL die zum Zweck der Kalibrierung aufgrund der geringen Anzahl Betriebe von BS zusammengelegt wurden.

geschätzt. Für die Varianzschätzung im Falle eines Untersuchungsbereiches d, wird in Formel (14) $\hat{e}_i$ durch $1(i \in d)\hat{e}_i$ ersetzt:

$$\widehat{\mathbf{Var}}(\hat{Y}_d) = \sum_{h \in H} \frac{N_h^2}{m_h}(1 - \frac{m_h}{N_h})\frac{1}{m_h - 1}\sum_{i=1}^{m_h}(g_i 1(i \in d)\hat{e}_i - \frac{\sum g_i 1(i \in d)\hat{e}_i}{m_h})^2 \qquad (15)$$

Bemerkung zu (15): In Fällen, in welchen der Untersuchungsbereich nicht einem oder einer Vereinigung mehrerer Kantone entspricht, liefert (15) eine Schätzung basierend auf folgenden Hypothesen:

- der Schätzer in (9) ist identisch zum Schätzer, der aus einer Kalibrierung auf Niveau Kanton$\times$Untersuchungsbereich (Kanton gekreuzt mit dem Untersuchungsbereich) hervorgehen würde.
- die Regressionsgleichung (12) gilt unabhängig davon, ob ein Betrieb im Untersuchungsbereich ist oder nicht. Das heisst, die Zugehörigkeit zum Untersuchungsbereich hat keinen Einfluss auf die Regressionskoeffizienten.

Die Hypothesen sind kaum immer erfüllt und die resultierenden Varianzschätzungen sind somit Approximationen. Es ist aber zu beachten, dass viele wichtige Untersuchungsbereiche aus Vereinigungen von Kantonen hervorgehen (z.B. Niveau Schweiz, Niveau Kanton, Niveau Grossregion) und in diesen Fällen sind die formulierten Bedingungen automatisch erfüllt.

In Anlehnung an [1], Seite 296, sowie (15) läge es nahe, die Varianz des in (11) vorgestellten Schätzers $\hat{Q}_{Y_d,Z_d}$ eines Quotienten im Untersuchungsbereich d wie folgt zu schätzen:

$$\begin{aligned} \widehat{\mathbf{Var}}(\hat{Q}_{Y_d,Z_d}) &= \frac{1}{\hat{Z}_d^2}\sum_{h \in H}\frac{N_h^2}{m_h}(1 - \frac{m_h}{N_h})\frac{1}{m_h - 1}\sum_{i=1}^{m_h}(T_i - \bar{T}(h))^2 \qquad (16)\\ T_i &:= g_i 1(i \in d)\hat{e}_{y,i} - \hat{Q}_{Y_d,Z_d} g_i 1(i \in d)\hat{e}_{z,i} \\ \bar{T}(h) &:= \frac{1}{m_h}\sum_{i=1}^{m_h} T_i \end{aligned}$$

Dabei entsprechen $\hat{e}_{z,i}$ und $\hat{e}_{y,i}$ den geschätzten Residuen, welche man analog zu (13) für die Zielvariablen y_i und z_i erhält. Für die Schätzung von Totalen und die entsprechende Varianzschätzung konnte auf die SAS-Prozedur Surveymeans zurückgegriffen werden. Nun ist es so, dass die Schätzung von Quotienten in Untersuchungsbereichen in der verwendeten SAS-Version (9.2) mittels der Prozedur Surveymeans noch nicht möglich ist. Darum wurde für die Berechnung der Schätzung von $\hat{Q}_{Y_d,Z_d}$ gemäss (11) das SAS-Macro SMSUB verwendet, welches auf der SAS-Homepage unter anderem für diesen Zweck zur Verfügung gestellt wurde. SMSUB liefert nebst dem Schätzer (11) auch eine Varianzschätzung. Diese entspricht gemäss der Dokumentation einer konservativen Approximation, in welcher die Residuen $\hat{e}_{y,i}$, $\hat{e}_{z,i}$ in (16) durch die Werte y_i und z_i ersetzt werden. Der - normalerweise varianzreduzierende - Effekt der Kalibrierung wird also vernachlässigt, weshalb für Quotienten eher von etwas zu konservativen (zu hohen) Varianzschätzungen ausgegangen werden muss.

3.4.1 Geplante versus erreichte Präzision

Die Präzision der Hochrechnungsresultate hängt, wie erwartet werden durfte, von der betrachteten Zielvariablen und dem Untersuchungsbereich ab. Als Illustration sind in Tabelle 19 die Hochrechnungsresultate für Totale der Variablen

- b8021: landwirtschaftliche Nutzfläche im Eigentum (Aren)
- b8431: bewässerbare Fläche (Aren)
- b8612: Anzahl Laufstallplätze für Kühe

auf Niveau Kanton dargestellt. Nebst den Totalen sind auch die Nettostichprobengrösse (n), Standardabweichung (StdDev) sowie der Variationskoeffizient der Schätzung (CV) aufgelistet. Es ist zu beachten, dass es sich bei diesen Ergebnissen um Testhochrechnungen zum Zwecke der Methodenillustration handelt, welche nicht zwingend mit den effektiven Hochrechnungsresultaten gemäss Publikation übereinstimmen. Betrachten wir die Variationskoeffizienten, so fällt folgendes auf:

- Die Hochrechnung der landwirtschaftlichen Nutzfläche im Eigentum liefert auf Niveau Kanton Resultate, welche für sämtliche Kantone einen CV unterhalb der angestrebten $0.05 = 5\%$ aufweisen (Median bei $CV = 2\%$). Das angestrebte Präzisionsziel auf Niveau Kanton wird also für diese Variable übertroffen. Erklärung: Die landwirtschaftliche Nutzfläche im Eigentum hat einen starken Zusammenhang mit der landwirtschaftlichen Nutzfläche, welche eine der Variablen war, nach welcher der Stichprobenplan optimiert wurde und welche auch im Rahmen der Kalibrierung benützt wurde.
- Die Hochrechung der Anzahl Laufstallplätze für Kühe lierfert generell Variationskoeffizienten, welche um die angestrebten 5% herum streuen (Median = 5%). Gewisse Kantone führen also auf einen CV von über 5% (Maximum von 10%), aber generell befinden sich die Genauigkeiten im angestrebten Bereich. Mögliche Erklärung: Der Plan wurde für die Grösse Anzahl Rindvieh optimiert und diese Variable wurde auch im Rahmen der Kalibrierung verwendet. Der Zusammenhang zwischen Anzahl Laufstallplätze für Kühe und den Optimierungsvariablen ist hier etwas schwächer als im Fall der landwirtschaftlichen Nutzfläche, aber immer noch gut.
- Ein Beispiel einer heiklen Variable ist die bewässerbare Fläche. Sie liefert Variationskoeffizienten welche um den Bereich $10\% = 0.1$ (Median =12 %) streuen. Mögliche Erklärung: Die bewässerbare Fläche hat in manchen Kantonen nur wenig mit der Nutzfläche zu tun, nach welcher der Plan optimiert wurde. In solchen Kantonen ist der Plan also für diese Variable nicht optimal, was sich dann in entsprechend hohen Variationskoeffizienten auswirkt. In speziellen Kantonen, wie dem Wallis, wo aufgrund einer bestimmten Bewirtschaftung (Reben) ein starker Zusammenhang zwischen bewässerbarer Fläche und der Nutzfläche vermutet werden darf, ergeben sich aber auch hier tiefere Variationskoeffizienten. Es gibt auch Kantone, für welche die bewässerbare Fläche eine stark untergeordnete Rolle spielt (d.h. nur sehr wenige oder gar keine Betriebe in der Stichprobe haben eine bewässerbare Fläche). In solchen Fällen ist die CV-Schätzung mit Vorsicht zu geniessen (siehe die entsprechende Bemerkung am Schluss des Abschnitts).

Fazit: Für Zielvariablen, welche mit den Optimierungsvariablen für den Plan sowie den Kalibrierungsvariablen gut korrelieren, liegen die Variationskoeffizienten im Bereich 5%, teilweise auch darunter. Für andere Variablen kann die Hochrechnung auf Niveau Kanton zu Variationskoeffizienten führen, die weit über den angestrebten 5% liegen. Dabei ist zu beachten, dass die Korrelation zwischen Variablen von Kanton zu Kanton stark verschieden sein kann.

Bemerkungen zu Fällen mit CV=0: In Tabelle 19 stösst man auf Fälle mit einem Variationskoeffizienten von 0. Dies kann zwei Gründe haben.

1. Der interessierende Kanton ist vollerhoben und sämtliche Betriebe haben geantwortet (Basel-Stadt).

Tabelle 19 Hochrechnung auf Niveau Kanton für drei ausgewählte Variablen.

		b8021			b8431			b8612		
kt	n	b8021	StdDev	cv	b8431	StdDev	cv	b8612	StdDev	cv
ZH	927	3675967	72338	0.02	355868	28789	0.08	32863	1671	0.05
BE	1098	11091110	220634	0.02	1308110	144463	0.11	62942	4232	0.07
LZ	772	5514795	98516	0.02	44103	10278	0.23	45761	2061	0.05
UR	290	383671	9542	0.02	85	3	0.04	1942	191	0.10
SZ	575	1200818	30290	0.03	6366	2746	0.43	8172	545	0.07
OW	311	363131	9502	0.03	90	0	0.00	4523	315	0.07
NW	325	350232	6034	0.02	305	38	0.12	2931	164	0.06
GL	377	269923	2515	0.01	2303	574	0.25	2037	68	0.03
ZG	378	540281	12903	0.02	14031	3896	0.28	7415	245	0.03
FR	952	4097633	75391	0.02	445960	41567	0.09	33164	1389	0.04
SO	802	1664175	24442	0.01	89761	11798	0.13	13972	510	0.04
BS	13	3516	0	0.00	87	0	0.00	184	0	0.00
BL	488	1092864	20128	0.02	46240	4171	0.09	10045	454	0.05
SH	441	828474	7103	0.01	70724	3206	0.05	4458	114	0.03
AR	502	585249	10693	0.02	1357	637	0.47	3965	230	0.06
AI	355	339160	6764	0.02	10	0	0.00	2540	144	0.06
SG	809	3653658	79359	0.02	129446	18400	0.14	34288	1974	0.06
GR	546	2245704	64720	0.03	340251	39720	0.12	21711	1344	0.06
AG	954	3418856	53521	0.02	377406	30615	0.08	28122	1188	0.04
TG	924	2936233	45678	0.02	411179	59433	0.14	32671	1281	0.04
TI	498	415495	11791	0.03	54488	6436	0.12	4324	255	0.06
VD	852	5752629	120995	0.02	944301	100391	0.11	33556	1744	0.05
VS	579	1218682	44856	0.04	1377351	85999	0.06	5326	439	0.08
NE	494	1838930	39312	0.02	105171	23301	0.22	11024	473	0.04
GE	374	416394	6251	0.02	118565	4013	0.03	914	57	0.06
JU	533	2323225	47996	0.02	38045	10761	0.28	19309	628	0.03
Median				0.02			0.11			0.05

2. Die Variable Gemüsefläche wurde im Rahmen der Hochrechnung auf Ebene Kanton als Kalibrierungsvariable verwendet. In gewissen Kantonen tauchte nun die Situation auf, dass in sämtlichen antwortenden Betrieben die bewässerbare Fläche identisch zur Gemüseanbaufläche ist. Diese perfekte Kollinearität führt dann auf eine CV-Schätzung von Null. (Zu beachten ist, dass dieser Fall nur in Kantonen auftaucht, wo die bewässerbare Fläche eine eher marginale Rolle spielt und auf Kantonsebene wohl kaum von grösserem Interesse ist).

Literatur

[1] Särndal, C-E., Svensson, B., Wretman, W. (1992). *Model Assisted Survey Sampling*, Springer Series in Statistics, Springer-Verlag

[2] Cochran, W.G. (1977). *Sampling techniques*, John Wiley & Sons.

[3] Hidiroglou, M.A. (1986). *The construction of a self-representing stratum of large units in survey design.* The American Statistician, 40(1), 27-31.

[4] Potterat, J. (2003). *Landwirtschaftliche Betriebszählung 2003: Stichprobenplan der Zusatzerhebung.* (Bundesamt für Statistik, Methodenbericht 338-0020).

[5] Potterat, J. (2005). *Landwirtschaftliche Betriebszählung 2003: Schätzverfahren für die Zusatzerhebung.* (Bundesamt für Statistik, Methodenbericht 338-0028).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Publikationsprogramm BFS

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

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

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

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

Methodenberichte des Dienstes Statistische Methoden

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

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

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

Die Zusatzerhebung der landwirtschaftlichen Betriebszählung ist eine Stichprobenerhebung, die eine Vielzahl an Variablen aus verschiedenen Themengebieten umfasst. Der Stichprobenplan beruht auf einer einfachen geschichteten Zufallsstichprobe. Er wurde mit dem Ziel entwickelt, die landwirtschaftliche Nutzfläche sowie den Rindvieh- und Schweinebestand auf Niveau Kanton möglichst effizient mit einem Variationskoeffizienten von 5% schätzen zu können. Zusammen mit einer angenommenen Antwortrate von 60% führte dies auf eine Bruttostichprobe von 16'815 Betrieben. Daraus resultierte ein Rücklauf von 15'169 Fragebogen, die in der Hochrechnung verwendet werden konnten. Die Hochrechnungsgewichte wurden ausgehend von den Stichprobengewichten mittels einer Korrektur für Antwortausfälle und einer Kalibrierung konstruiert. Im Rahmen der Kalibrierung wurden die Gewichte so angepasst, dass die Hochrechnungsresultate verschiedener Hilfsvariablen auf Niveau Kanton mit den entsprechenden Totalen der landwirtschaftlichen Strukturerhebung 2010 übereinstimmen. Die Hochrechnungsresultate haben gezeigt, dass die Qualität der Schätzungen, ausgedrückt mittels Variationskoeffizienten, generell im erwarteten Bereich liegt: Für Zielvariablen, welche mit den im Rahmen von Stichprobenplan und Kalibrierung verwendeten Hilfsvariablen gut korrelieren, liegen die Variationskoeffizienten generell im Bereich 5% oder auch darunter. Für andere Variablen kann die Hochrechnung auf Niveau Kanton zu Variationskoeffizienten führen, die weit über den angestrebten 5% liegen.

Bestellnummer
338-0068

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

Preis
gratis

ISBN 978-3-303-00496-8