Methodenbericht

# Vierteljährliche Wohnbaustatistik

0

**Statistische Grundlagen und Übersichten**
**Bases statistiques et produits généraux**
**Basi statistiche e presentazioni generali**

## Stichprobenplan, statistische Datenaufarbeitung und Schätzverfahren 2005

**Office fédéral de la statistique**
**Bundesamt für Statistik**
**Ufficio federale di statistica**
**Uffizi federal da statistica**
**Swiss Federal Statistical Office**

OFS BFS UST

Neuchâtel, 2006

## Statistik der Schweiz

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 Verkehr und Nachrichtenwesen
12 Geld, Banken, Versicherungen
13 Soziale Sicherheit
14 Gesundheit
15 Bildung und Wissenschaft
16 Kultur, 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

## Statistique de la Suisse

La série «Statistique de la Suisse» publiée par l'Office fédéral de la statistique (OFS) couvre les domaines suivants:

0 Bases statistiques et produits généraux
1 Population
2 Espace et environnement
3 Vie active et rémunération du travail
4 Economie nationale
5 Prix
6 Industrie et services
7 Agriculture et sylviculture
8 Energie
9 Construction et logement
10 Tourisme
11 Transports et communications
12 Monnaie, banques, assurances
13 Protection sociale
14 Santé
15 Education et science
16 Culture, société de l'information, sport
17 Politique
18 Administration et finances publiques
19 Criminalité et droit pénal
20 Situation économique et sociale de la population
21 Développement durable et disparités régionales et internationales

**Statistik der Schweiz** | **Statistique de la Suisse**

**Methodenbericht** | **Rapport de méthodes**

# Vierteljährliche Wohnbaustatistik

## Stichprobenplan, statistische Datenaufarbeitung und Schätzverfahren 2005

*Autor* **Daniel Kilchmann**
Bundesamt für Statistik

*Herausgeber* **Bundesamt für Statistik**

OFS BFS UST

**Office fédéral de la statistique**
**Bundesamt für Statistik**
**Ufficio federale di statistica**
**Uffizi federal da statistica**
**Swiss Federal Statistical Office**

Neuchâtel, 2006

**Vorwort**

Die vierteljährliche Wohnbaustatistik erhebt vierteljährlich Daten zur Wohnbautätigkeit bei den Gemeindeverwaltungen. Aufgrund von strukturellen Veränderungen sowie einer neuen Definition der Agglomerationsgemeinden in der Volkszählung 2000 wurde die Vierteljährliche Wohnbaustatistik revidiert. Der Dienst Statistische Methoden (METH) des Bundesamts für Statistik wurde von der Sektion Monetäre Unternehmensstatistik (MON) mit den methodischen Arbeiten beauftragt. Diese wurden durch Daniel Kilchmann durchgeführt. Dabei wurde er von Marlies Henggi, Marc Zimmermann und Kamel Chaouach von der Sektion MON begleitet.

Die Angaben über Resultate der Erhebung in diesem Bericht dienen allein zur Illustration der Methoden. Es können sich Abweichungen gegenüber den offiziellen Ergebnissen ergeben, die auf der Wahl der Datenbasis und ihrem Stand, leichten Anpassungen der Methoden, und der Darstellung basieren. Nur die im offiziellen Bericht veröffentlichten Ergebnisse sind massgebend.

Der Autor dankt Marlies Henggi, Marc Zimmermann und Kamel Chaouach für die sehr gute Zusammenarbeit, sowie Dr. Beat Hulliger vom Dienst Statistische Methoden für die aufmerksame Durchsicht des Berichts.

**Zusammenfassung**

Die vierteljährliche Wohnbaustatistik wurde aufgrund der Einwohnerzahlen und der Definition der Agglomerationsgemeinden der Volkszählung 2000 revidiert. Um den Produktionsablauf zu optimieren, wurde eine maximale Überlappung der neuen Erhebung mit der bisherigen angestrebt. Aufgrund der häufigen Gemeindefusionen und stetigen Änderungen der Einwohnerzahlen, wurden diese Einflüsse in den Erhebungsprozess und die Schätzung miteinbezogen. Es werden Datenkontrollen sowie einfache Methoden der Ausreisserlokalisierung diskutiert. Die Schätzverfahren sind geschichtete Totalschätzer mit Korrekturen für allfällige Antwortausfälle. Die Varianzen der Totalschätzer sind zum Teil gross, so dass die Interpretation von Veränderungen beeinträchtigt wird.

**Schlüsselwörter**

Methodenbericht; Gemeindefusionen; Kish&Scott; Partage des poids; Hidiroglou-Berthelot; Einsetzungen; Grafische Datenkontrollen; Wohnbaustatistik.

**Auskunft:** Daniel Kilchmann, Tel. 032 713 60 84
Daniel.Kilchmann@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
Order@bfs.admin.ch

**Internet:** http://www.statistik.admin.ch

**Bestellnummer:** 338-0035

**Preis:** gratis

**Reihe:** Statistik der Schweiz

**Fachbereich:** 0 Statistische Grundlagen und Übersichten

**Originaltext:** Deutsch

**Grafik/Layout:** BFS

**Copyright:** BFS, Neuchâtel 2006
Abdruck – ausser für kommerzielle Nutzung–
unter Angabe der Quelle gestattet.

**ISBN:** 3-303-00330-0

# Inhaltsverzeichnis

**1 Einführung** **5**

**2 Daten** **5**

**3 Stichprobenplan** **5**
3.1 Schichtung und Zuordnung . . . . . 6
3.2 Ziehung der Stichprobe . . . . . 6

**4 Einwohnerzahlen und Gemeindefusionen** **8**

**5 Datenaufarbeitung** **9**
5.1 Antwortausfälle . . . . . 9
5.2 Kontrollregeln . . . . . 9

**6 Schätzverfahren** **11**

**7 Produktionsablauf** **12**

**Anhang** **14**

**A Schätzverfahren** **14**

**B SAS-Programme** **16**
B.1 Stichprobenplan . . . . . 16
B.2 Schätzverfahren . . . . . 16

**Literatur** **19**

# 1 Einführung

Die vierteljährliche Wohnbaustatistik (TRIM) dient zur Schätzung der Wohnbautätigkeit und deren Veränderung gegenüber dem gleichen Quartal des Vorjahres. Für Gemeinden mit weniger als 5'000 Einwohner wird die Wohnbautätigkeit geschätzt. Gemeinden mit mehr als 5'000 Einwohnern sowie die Agglomerationen der Gemeinden Basel, Bern, Genf, Lausanne und Zürich werden voll erhoben.

Die TRIM wurde letztmals für das Jahr 1994 revidiert. Dabei wurden die Definitionen der Agglomerationen und die Einwohnerzahlen der Gemeinden aus der Volkszählung 1990 (VZ90) übernommen. Die Sektion Monetäre Unternehmensstatistik (MON) des Bundesamts für Statistik (BFS) hat den Dienst statistische Methoden (METH) angefragt, die Erneuerung der Stichprobe der TRIM für das Jahr 2005 methodisch zu begleiten.

METH hat aufgrund der Gemeindestruktur vom 1.4.2005 und der Einwohnerzahlen der Volkszählung 2000 (VZ2000) den Stichprobenplan überarbeitet. Um den Verarbeitungsprozess zu optimieren wurde eine maximale Überlappung der bisherigen Stichprobe mit der neuen Stichprobe angestrebt. MON wurden Prototypen von SAS-Programmen für die Anpassung der Stichprobe bei Gemeindefusionen und für die Hochrechnung geliefert. Der Prototyp beinhaltet zudem eine statistische Datenaufarbeitung und deterministische Kontrollregeln.

Die Schätzverfahren sind geschichtete Totalschätzer mit Korrekturen für allfällige Antwortausfälle. Die Varianzen der Totalschätzer sind zum Teil gross, so dass die Interpretation von Veränderungen beeinträchtigt wird.

Die Medienmitteilung der Wohnbautätigkeit des ersten Quartals 2005 erfolgte Anfangs Juli 2005 auf Basis der neuen Stichprobe und der im vorliegenden Bericht beschriebenen Schätzmethode.

# 2 Daten

Die Grundgesamtheit besteht aus den 2'758 schweizer Gemeinden, Stand 1.4.2005. Die Einwohnerzahlen dieser Gemeinden stammten aus der VZ2000, wobei sie aufgrund von Fusionen und Teilungen durch MON angepasst worden waren. METH wurde eine Datei mit den 2'758 Gemeinden und den angepassten Einwohnerzahlen geliefert.

In der TRIM werden die Anzahl neu erstellter Wohnungen (NW), die Anzahl im Bau befindlicher Wohnungen (IB) und die Anzahl baubewilligter Wohnungen (BB) geschätzt. Die entsprechenden Daten für die Schätzung werden bei den Gemeindeverwaltungen erhobenen.

# 3 Stichprobenplan

Das Gutachten der letzten Revision, [Schwendener, 1993], wurde als methodische Basis für die Erarbeitung des Stichprobenplans verwendet.

Zur Optimierung des Verarbeitungsprozesses wurde angestrebt, einen maximalen Anteil der bestehenden Stichprobe beizubehalten und die Stichprobengrösse bei rund 900 Gemeinden zu belassen. Wenn mehr als 100 neuen Gemeinden in die Stichprobe aufgenommen würden, müsste mit Verzögerungen in der Produktion gerechnet werden.

## 3.1 Schichtung und Zuordnung

Die Gemeinden wurden aufgrund der Einwohnerzahl und der Zugehörigkeit zu den Agglomerationsgemeinden der fünf grössten schweizer Städte geschichtet. Die Agglomerationsgemeinden umfassen alle Gemeinden, die aufgrund der VZ2000 zur Agglomeration von Basel, Bern, Genf, Lausanne und Zürich gezählt wurden.
Die Basis für die Zuweisung der Agglomerationsgemeinden bilden die von MON gelieferten Angaben für den Stand 1.4.2005. Aufgrund dieser Definition der Agglomeration wird in Zukunft die Zugehörigkeit zu den Agglomerationsgemeinden nach Fusionen und Teilungen bestimmt werden, siehe Abschnitt 4.
Die Einwohnerzahlen werden ab 2006 jeweils anfangs Jahr aufgrund der Zahlen der Statistik des jährlichen Bevölkerungsstandes (ESPOP) angepasst. 2005 erfolgte diese Anpassung für das 2. Quartal. Die ESPOP liefert den Gemeindestand des 31.12. des Vorjahres mit den Einwohnerzahlen, die ein weiteres Jahr zurückliegen.

Die Anzahl Schichten wurden von bisher 6 auf 5 verkleinert, da in der bisherigen Stichprobe die Schicht mit der kleinsten Anzahl Einwohner (bis 200) nur 34 Gemeinden umfasste. Durch Fusionen verkleinerte sich die Anzahl Gemeinden dieser Schicht in den letzten Jahren kontinuierlich.
Anhand der *Dalenius*-Methode[1], siehe [Cochran, 1977], wurden die Schichtgrenzen optimal bestimmt, wobei Gemeinden mit mehr als 5'000 Einwohnern und Agglomerationsgemeinden eine voll erhobene Schicht bildeten (siehe Tabelle 1). Die Schichtgrenzen wurden sodann angepasst, wobei 2'000 Einwohner als Schichtgrenze festgelegt wurde, da diese auch in der Publikation vorhanden ist. Die endgültigen Schichtgrenzen können der Tabelle 1 entnommen werden.
Die Stichprobengrösse in den Schichten wurde aufgrund der Schichtgrenzen und anhand der Standardabweichung der neu erstellten Wohnungen der Baustatistik 2003, mit einer *Neyman*-Zuordnung berechnet. Die gesamte Stichprobengrösse beträgt 1'120 Gemeinden, wenn sämtliche publizierten Totalschätzer einen Variationskoeffizienten von weniger als 5% haben sollen. Diese Stichprobengrösse hätte eine Verzögerung in der Produktion zur Folge gehabt. Deshalb wurde entschieden, die Stichprobengrösse bei 900 Gemeinden zu belassen.

## 3.2 Ziehung der Stichprobe

In jeder Schicht wurde eine einfache Zufallsstichprobe mit der Stichprobengrösse $n_h$ gezogen, siehe Tabelle 1. Diese Ziehung war unabhängig von der bisherigen Stichprobe. Rund zwei drittel der Gemeinden der neuen Stichprobe waren bereits in der alten Stichprobe von 1994, vor allem Gemeinden aus der voll erhobenen Schicht. Um die Überlappung

[1] in [Cochran, 1977] *cumulative* $\sqrt{f}$ *rule* genannt.

**Tabelle 1:** Schichtgrenzen nach der *Dalenius*-Methode und die gewählten (BFS). Für die gewählte Schichtung: die Schichtgrösse ($N_h$), die Standardabweichung aus der Baustatistik 03 der neu erstellten Wohnungen ($S_h$), die Stichprobengrössen in den Schichten bei einer Stichprobengrösse von 900 Gemeinden und einer *Neyman*-Zuordnung ($n_h$), das Gewicht ($w_h$), sowie die optimale Stichprobengrösse, wenn sämtliche publizierten Resultate einen Variationskoeffizient $<$ 5% haben sollen ($n_h^*$). 'Nr.' steht für die Schichtnummer.

| | Schichtgrenzen | | | | | | |
|---|---|---|---|---|---|---|---|
| **Nr.** | **Dalenius** | **BFS** | $\mathbf{N_h}$ | $\mathbf{S_h}$ | $\mathbf{n_h}$ | $\mathbf{w_h}$ | $\mathbf{n_h^*}$ |
| 1 | >5'000 + Agglo. | >5'000 + Agglo. | 577 | 92.57 | 577 | 1.00 | 577 |
| 2 | 2'077-4'984 | 2'001-5'000 | 402 | 15.73 | 151 | 2.66 | 164 |
| 3 | 1'031-2'076 | 1'001-2'000 | 443 | 6.90 | 73 | 6.07 | 161 |
| 4 | 437-1'030 | 501-1'000 | 486 | 4.41 | 51 | 9.53 | 113 |
| 5 | <437 | <501 | 850 | 2.34 | 48 | 17.71 | 105 |
| **Total** | | | 2'758 | | 900 | | 1'120 |

der neuen Stichprobe mit der alten Stichprobe zu maximieren, wurde der Algorithmus von *Kish&Scott* [Kisch, Scott, 1971] verwendet. Diese Methode garantiert die Eigenschaften einer Zufallsstichprobe für die resultierende Stichprobe. Auf diese Weise konnte eine rund 90% Überlappung der bisherigen mit der endgültigen Stichprobe erreicht werden, die durch die Anwendung des Algorithmus von *Kish&Scott* resultierte.

Um die zu erwartenden Variationskoeffizienten mit jenen von 1994 zu vergleichen, wurden die Grössenklassen der Publikation als Untersuchungsbereiche verwendet. Die Schätzung für das 4. Quartal 2003 basiert auf der neuen Stichprobe mit der neuen Gewichtung. Fehlende Werte wurden mittels einem linearen Regressionsmodells aufgrund der Daten der Baustatistik 2003 und der vorhandenen Daten für das 4. Quartal 2003 eingesetzt.
Die zu erwartenden Variationskoeffizienten sind für die meisten Schätzer in der gleichen Grössenordnung, wie jene des 2. Quartals 1994, siehe Tabelle 2. Einzig für die Variationskoeffizienten der neu erstellten Wohnungen der Gemeinden zwischen 2'001 und 5'000 Einwohner und der baubewilligten Wohnungen in Gemeinden mit höchstens 2'000 Einwohnern muss für die neue Stichprobe mit deutlich höheren Variationskoeffizienten gerechnet werden. Diese hängen einerseits von der Stichprobengrösse ab und andererseits von der Varianz. Aufgrund mangelnder Zeit konnte nicht untersucht werden, ob sich die Varianzen set 1994 vergrössert haben.

**Tabelle 2:** Zu erwartende Variationskoeffizienten für die geschätzten Totale des 2. Quartals 1994 ($cv_{94,2}$) und des 4. Quartals 2003 ($cv_{03,4}$). Die Werte für 1994 basieren auf der damaligen Stichprobe. Die Werte für 2003 stammen von der neuen Stichprobe mit der neuen Gewichtung; fehlende Werte wurden mittels einem linearen Regressionsmodells eingesetzt.

| **Einwohner** | **NW** | | **BB** | | **IB** | |
|---|---|---|---|---|---|---|
| **Grössenklasse** | $\mathbf{cv_{94,2}}$ | $\mathbf{cv_{03,4}}$ | $\mathbf{cv_{94,2}}$ | $\mathbf{cv_{03,4}}$ | $\mathbf{cv_{94,2}}$ | $\mathbf{cv_{03,4}}$ |
| >10'000 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 5'001-10'000 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2'001-5'000 | 7.01 | 11.62 | 7.16 | 6.93 | 4.52 | 5.11 |
| <2'001 | 12.53 | 10.66 | 12.81 | 16.83 | 8.97 | 10.14 |
| **Total** | 3.38 | 3.64 | 4.56 | 3.44 | 2.11 | 1.98 |

# 4 Einwohnerzahlen und Gemeindefusionen

Vor dem Versand des Fragebogens muss in jedem Quartal die Stichprobe neu bestimmt werden. Eine Datei mit den Gemeindenummern des Vorquartals und den aktuellen Gemeindenummern und -namen muss bereitgestellt werden, damit Gemeindefusionen miteinbezogen werden können.
Der Strukturwandel aufgrund der Einwohnerzahlen wird einmal jährlich anhand der Zahlen der ESPOP[2] vom 31.12. des Vorjahres berücksichtigt, wie im Abschnitt 3.1 bereits erwähnt wurde.
Allgemein gilt, dass Gemeinden, die in die Stichprobe der vierteljährlichen Wohnbaustatistik aufgenommen wurden, auch nach Fusionen oder Änderungen der Einwohnerzahlen immer noch zu dieser Stichprobe dazugehören. Sind in einer Schicht weniger als 40 Gemeinden in der Erhebung, so wird im SAS-Prototypen darauf hingewiesen. Eine Revision sollte in diesem Falle angegangen werden.

**Einwohnerzahl** Gemeinden, die zum Zeitpunkt der Stichprobenziehung weniger als 5'000 Einwohner hatten und danach einen Anstieg über 5'000 Einwohner verzeichnen, werden neu in die Erhebung aufgenommen, sofern sie nicht bereits teilgenommen haben. Diese Gemeinden erhalten das Gewicht 1, da sie in die Schicht der vollerhobenen Gemeinden aufgenommen werden.
Aufgrund der aktuellen Einwohnerzahlen wird die Zuordnung zu den Einwohnergrössenklassen für die Analysen definiert. Die Gewichtung wird der aktuellen Anzahl Gemeinden pro Schicht und Grössenklasse angepasst.

**Schichtung** Die Schichtzugehörigkeit von fusionierten und geteilten Gemeinden ist insofern von Bedeutung, als dass durch die Fusion (Teilung) die Zugehörigkeit zur Schicht 1 vererbt wird. Das heisst, die Teilung einer Agglomerationsgemeinde führt zu zwei Agglomerationsgemeinden. Fusioniert eine Gemeinde mit einer Agglomerationsgemeinde so wird die fusionierte Gemeinde den Agglomerationsgemeinden zugeordnet. Das selbe gilt für Gemeinden mit mehr als 5'000 Einwohnern.

**Gemeindefusionen** Fusionieren Gemeinden, werden bis zur nächsten Aktualisierung der Einwohnerzahlen mit den ESPOP Zahlen, die Einwohnerzahlen der Ursprungsgemeinden summiert, um die Einwohnerzahl der Fusionsgemeinde zu berechnen. Das Gewicht der Fusionsgemeinde wird anhand der Methode der 'Verteilung der Gewichte', c.f. 'Partage des poids' in [Lavallée, 2002], berechnet. Durch Gemeindefusionen entsteht ein Strukturwandel in der Grundgesamtheit. Dieser wird mit einer Nachgewichtung aufgrund der Anzahl Gemeinden pro Schicht und Grössenklasse und der Summe der Gewichte, Rechnung getragen.
Die Methode der 'geteilten Gewichte ist nicht transitiv, siehe [Deville, Lavallé, 2003], und ist somit nur korrekt, solange Gemeinden, die nach dem 1.1.2005 fusioniert haben, sich nicht wieder teilen. Geschieht dies dennoch, so muss mit einem Bias gerechnet werden, sofern die Stichprobe nicht revidiert wird.

[2]Statistik des jährlichen Bevölkerungsstandes.

**Teilung von Gemeinden** Teilt sich eine Gemeinde die in die Stichprobe gezogen wurde, so werden alle Gemeinden nach der Teilung erhoben. Das Gewicht wird vererbt. Die Variablenwerte des vorhergegangenen Jahres werden proportional zur Anzahl Einwohner der neu entstandenen Gemeinden aufgeteilt. Eine Nachgewichtung berücksichtigt die Änderung der Anzahl Gemeinden pro Schicht und Grössenklasse.

# 5 Datenaufarbeitung

Vor der Schätzung wird getestet, ob die Daten vollständig sind und ob sie Kontrollregeln genügen, die verhindern sollen, dass unmögliche Werte in die Schätzung miteinbezogen werden.

## 5.1 Antwortausfälle

Antwortausfällen wird durch Nachgewichtung innerhalb der Schichten und Grössenklassen Rechnung getragen. Gemeinden, die neu in die Stichprobe gezogen wurden weisen für das vorangegangene Jahr keine Werte auf. In der Schätzung dieser Werte werden die neuen Gemeinden als Antwortausfälle betrachtet.

## 5.2 Kontrollregeln

Anhand der angewandten Kontrollregeln wird dem Benutzer einerseits auf dem Bildschirm angezeigt, welche Gemeinden diese verletzen und andererseits werden SAS-Dateien mit den suspekten Gemeinden erstellt. Es wurden keine Aktionen bestimmt, anhand derer Werte abgeändert werden.
Die Resultate der Kontrollregeln sollen dem Benutzer hilfreich sein, um mögliche Fehler zu entdecken. Nach manueller Bearbeitung dieser möglichen Fehler muss das Programm neu gestartet werden und die Kontrollregeln werden von neuem durchgeführt.

**Deterministische Kontrollen** Die deterministischen Kontrollen basieren auf den Regeln, die in [Zambaz, 2002] beschrieben sind, und sind in der Tabelle 3 zusammengefasst. Die Gemeinden, deren Daten die deterministischen Kontrollregeln verletzen, werden im Log-Fenster mit der Regelnummer der verletzten Regel aufgelistet.

**Grafische Kontrollen** Aussergewöhnliche Werte können aus Grafiken gelesen werden, siehe Abbildung 1. Pro Grössenklasse wird für NW pro Einwohner, BB pro Einwohner und IB pro Einwohner eine Grafik erstellt, aus welcher der jeweilige Wert und die Gemeindenummer gelesen werden kann.
Es werden zudem NW gegen BB grafisch dargestellt, um extreme Werte innerhalb der Grössenklassen festzustellen.

Es wurde kein Kriterium festgelegt, um extreme Werte innerhalb dieser Grafiken zu lokalisieren. Es ist somit der Produktionseinheit überlassen, abzuwägen, ob ein Wert extrem ist oder nicht. Extreme Werte sind nicht unbedingt falsch, sondern im Vergleich mit den anderen Werten aussergewöhnlich gross. Es wird empfohlen die Werte der Gemeinden zu

**Tabelle 3:** Deterministische Kontrollregeln, die im SAS-Prototyp einbezogen wurden.

| Nr. | Regel |
|---|---|
| RP2251 | Ist die Anzahl Baubewilligungen während drei aufeinander folgender Quartale gestiegen, so muss die Anzahl neu erstellter Wohnungen innert 7 Quartalen mindestens einmal gegenüber dem Vorquartal steigen. |
| RP2254 | Ist die Anzahl Baubewilligungen während drei aufeinander folgender Quartale gesunken, so muss die Anzahl neu erstellter Wohnungen innert 7 Quartalen mindestens einmal gegenüber dem Vorquartal sinken. |
| RP2256 | Ist die Anzahl neu erstellter Wohnungen grösser als die Differenz zwischen der Anzahl im Bau befindlicher Wohnungen des aktuellen Quartals und des Vorquartals, so sollte dies überprüft werden. |
| MON | Ist die Anzahl neu erstellter Wohnungen grösser als die Anzahl im Bau befindlicher Wohnungen des Vorquartals, so sollte dies überprüft werden. |

überprüfen, die augenscheinlich grösser sind als die Werte der anderen Gemeinden. In der Abbildung 1 wären dies die drei Gemeinden mit den grössten Werten für NW pro Einwohner.

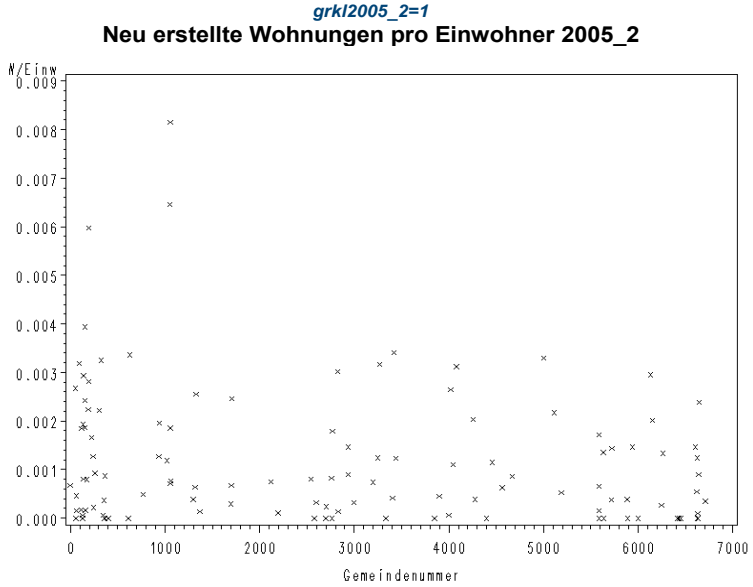



**Abbildung 1:** Grafische Kontrolle der Anzahl neu erstellter Wohnungen pro Einwohner der Gemeinden in der Grössenklasse 1.

***Hidiroglou-Berthelot*-Ausreisser** Aussergewöhnliche Veränderungen gegenüber dem gleichen Quartal des Vorjahres und dem Vorquartal werden mit der Ausreisserlokalisierungsmethode von *Hidiroglou-Berthelot* bestimmt, c.f. [Hidiroglou, Berthelot, 1986]. Dabei wird die relative Veränderung untersucht. Die Parameter der Methode wurden wie

folgt festgelegt: $u = 0.8$, $A = 0.05$ und $c = 50$. Der Benutzer wird am Bildschirm aufgefordert, die Gemeinden, bei denen eine aussergewöhnliche Entwicklung beobachtet werden kann, nochmals zu überprüfen.
Aufgrund der heterogenen Entwicklung von NW, BB und IB hat die Anwendung dieser Methode zum Ziel, die aussergewöhnlichsten Veränderungen zu entdecken. Dies bedeutet jedoch nicht, dass einerseits die Veränderungen, die nicht als aussergewöhnlich erkannt wurden, normal sind, und andererseits, dass die aussergewöhnlichen Veränderungen falsch sind.

# 6 Schätzverfahren

**Totalschätzer** Für die Schätzung der Totale und deren Varianz wird die SAS-Prozedur `proc surveymeans` verwendet. Somit wird, wie bisher, ein einfacher *Horvitz-Thompson* Schätzer unter Einbezug des Stichprobenplans und der Gewichtung verwendet. Die Agglomerationen und die Grössenklassen dienen als Untersuchungsbereiche, siehe Tabellen 5 und 6 im Anhang. Die Werte in den aufgeführten Tabellen entsprechen nicht den veröffentlichten Zahlen, da der Datenaufarbeitungsprozess noch nicht vollständig abgeschlossen war, als diese Tabellen produziert wurden. Erfreulich ist, dass alle Gemeinden geantwortet haben.

In der Tabelle 5 wurden die Schätzer des ersten Quartals 2004 aufgrund der alten Stichprobe und aufgrund der revidierten Stichprobe miteinander verglichen. Die Unterschiede rühren einzig von der neuen Definition der Agglomerationen der VZ2000 her, da alle Agglomerationsgemeinden ein Gewicht 1 haben. Die revidierten Werte fallen im Allgemeinen grösser aus, da in der VZ2000 im Vergleich zur VZ90 mehr Gemeinden zu den Agglomerationen der fünf grössten Städte gezählt werden.

Aus der Tabelle 6 ist ersichtlich, dass die revidierte Statistik gegenüber der alten Statistik vor allem bei den kleinsten Gemeinden weniger Bautätigkeit ausweist. Dies ist auf Fusionen und Bevölkerungszuwachs zurückzuführen. So hatten 1994 2'117 Gemeinden höchstens 2'000 Einwohner und 2005 waren es nur noch 1'779.

**Veränderung** Im Jahr 2005 werden die Werte für das Jahr 2004 aufgrund der neuen Stichprobe neu geschätzt. Dies wurde aufgrund der grossen strukturellen Veränderung der Gemeinden der Schweiz nötig. Gemeinden, die keine Werte für 2004 aufweisen, werden für 2004 wie Antwortausfälle behandelt. Die Veränderung der Schätzwerte von 2004 zu 2005 ist dann die Differenz der beiden Totalschätzer.
Ab 2006 wird die Differenz zu den Schätzern des Vorjahres aufgrund der beiden publizierten Totalschätzer geschätzt. Dies bedeutet, dass bei strukturellen Veränderungen, nicht dieselben Populationen miteinander verglichen werden. Bei der Bereitstellung der neuen Stichprobe werden jeweils Tabellen im SAS-Output ausgewiesen, anhand derer Strukturveränderungen erkennbar sind. Ebenso kann aufgrund der Tabelle der Gewichte abgeschätzt werden, ob der Einfluss der Strukturveränderung gross ist. Es kann aber nicht ausgeschlossen werden, dass ein Teil der ausgewiesenen Veränderung der Bautätigkeit einzig dem Strukturwandel zuzuschreiben ist. Dies gilt vor allem, wenn zusätzliche Ge-

meinden aufgrund von Fusionen oder eines Zuwachses der Einwohnerzahl, in die Erhebung aufgenommen werden.

**Empfehlungen** Da die Schätzer der vierteljährlichen Wohnbaustatistik zum Teil stark von Strukturänderungen beeinflusst werden, wird empfohlen, einen Indikator auszuweisen, der davon möglichst unabhängig ist. Zum Beispiel könnte die Bautätigkeit pro Einwohner verwendet werden.

Aus mangelnder Zeit konnten keine Varianzen für die Veränderung geschätzt werden. Hingegen wurden für die Totalschätzer die Varianzen anhand einer Approximation geschätzt. Unter der Voraussetzung, dass die Überlappung der Stichproben zweier Quartale die miteinander verglichen werden, gross ist, und dass eine positive Korrelation vorhanden ist, kann davon ausgegangen werden, dass die Standardabweichung der Veränderung kleiner ist, als die Wurzel der Summe der Varianzen der beiden Totalschätzer: $\hat{\sigma}(\Delta) \leq \sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}$. Wir können diese Schranke für eine konservative Regel benützen: Wenn eine Veränderung $\Delta$ kleiner ist als $\hat{\sigma}(\Delta)$, dann nehmen wir an, dass die Veränderung nicht signifikant ist. Aufgrund dieser Regel können zum Beispiel in der Tabelle 6 die Veränderungen von NW der Grössenklassen 3 und 4 nicht interpretiert werden.
Bei vollständig erhobenen Schichten sollte berücksichtigt werden, dass sämtliche Angaben immer eine Unsicherheit aufweisen. Es wird angeraten, in diesen Schichten eine Faustregel anzuwenden, wonach Veränderungen, die kleiner sind als die Wurzel des grösseren Totalschätzers, nicht interpretiert werden sollten.

# 7 Produktionsablauf

**Stichprobe bestimmen** Von der Sektion MON wird vierteljährlich eine Datei bereit gestellt, welche die Gemeindefusionen und -teilungen der vorangegangenen drei Monate beinhaltet. Dabei wird angegeben, welche Gemeinden (Gemeindenummern) vor der Fusion/Teilung existierten und wie sie sich und die Einwohnerzahlen entwickelt haben. Aufgrund dieser Daten wird die Grundgesamtheit und die Stichprobe aktualisiert und die Gewichtung angepasst.
Einmal jährlich werden zusätzlich die Einwohnerzahlen aufgrund der Daten der ESPOP aktualisiert.

**Hochrechnung** Nach der Befragung der Gemeinden der aktualisierten Stichprobe und nachdem sämtliche möglichen Anstrengungen unternommen wurden, um die Antwortausfälle zu vermeiden, exportiert MON die Daten des Rücklaufs aus der Datenbank in eine ASCII-Datei. Der von METH bereitgestellte Prototyp verarbeitet die Daten in mehreren Schritten, die in der Tabelle 4 aufgeführt sind. Während des Programmablaufs wird der Status des Ablaufs ins SAS Log-Fenster geschrieben. Falls es zu einem Abbruch des Programms kommt, wird die Begründung ebenfalls ausgewiesen. Das Log-Fenster muss für die Qualitätskontrolle analysiert werden und sollte archiviert werden.
Tabellen für die Publikation erscheinen einerseits im 'html' Format auf dem Bildschirm und werden andererseits in ASCII-Dateien exportiert. Diese müssen von MON noch in die gewünschte Form gebracht werden.

**Tabelle 4:** Programmschritte der SAS-Programme 'Stichprobe_bestimmen.sas' und 'Hochrechung.sas'.

| Schritt | Stichprobe bestimmen |
|---|---|
| 1 | Initialisierung |
| 2 | Einlesen der aktuellen Grundgesamtheit |
| 3 | Einlesen des Einwohnerzahlen |
| 4 | Erneuerung der Stichprobe |
| 5 | Abschluss des Programms |

| Schritt | Hochrechnung |
|---|---|
| 1 | Initialisierung |
| 2 | Einlesen des Rücklaufs |
| 3 | Bearbeiten der Antwortausfälle |
| 4 | Konsistenzkontrollen |
| 5 | Hochrechnung |
| 6 | Bereitstellen der Tabellen für die Publikation |
| 7 | Abschluss des Programms |

# Anhang

# A Schätzverfahren

**Tabelle 5:** Provisorische Schätzung für das 1. Quartal 2005 innerhalb der Agglomerationsgemeinden der VZ2000 ($\hat{Y}^r_{05,1}$), für das 1. Quartal 2004 innerhalb der Agglomerationsgemeinde der VZ2000 ($\hat{Y}^r_{04,1}$) und für das 1. Quartal 2004 innerhalb der Agglomerationsgemeinden der VZ90 ($\hat{Y}^a_{04,1}$). Der Variationskoeffizient *cv* ist 0, da die Agglomerationsgemeinden voll erhoben werden.

| **Variable** | **Agglo** | $\hat{Y}^r_{05,1}$ | *cv* | $\hat{Y}^a_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^a_{04,1}$ | $\hat{Y}^r_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^r_{04,1}$ |
|---|---|---|---|---|---|---|---|
| NW | Basel | 264 | 0 | 256 | 8 | 347 | -83 |
| | Bern | 307 | 0 | 244 | 63 | 247 | 60 |
| | Genf | 364 | 0 | 291 | 73 | 347 | 17 |
| | Lausanne | 519 | 0 | 326 | 193 | 328 | 191 |
| | Zürich | 1'447 | 0 | 1'108 | 339 | 1'213 | 234 |
| | **Total** | 2'901 | 0 | 2'225 | 676 | 2'482 | 419 |

| **Variable** | **Agglo** | $\hat{Y}^r_{05,1}$ | *cv* | $\hat{Y}^a_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^a_{04,1}$ | $\hat{Y}^r_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^r_{04,1}$ |
|---|---|---|---|---|---|---|---|
| IB | Basel | 2'605 | 0 | 1'551 | 1'054 | 1'984 | 621 |
| | Bern | 1'853 | 0 | 1'578 | 275 | 1'696 | 157 |
| | Genf | 3'329 | 0 | 2'774 | 555 | 3'222 | 107 |
| | Lausanne | 2'593 | 0 | 2'692 | -99 | 2'749 | -156 |
| | Zürich | 9'540 | 0 | 7'640 | 1'900 | 8'934 | 606 |
| | **Total** | 19'920 | 0 | 16'235 | 3'685 | 18'585 | 1'335 |

| **Variable** | **Agglo** | $\hat{Y}^r_{05,1}$ | *cv* | $\hat{Y}^a_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^a_{04,1}$ | $\hat{Y}^r_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^r_{04,1}$ |
|---|---|---|---|---|---|---|---|
| BB | Basel | 382 | 0 | 406 | -24 | 463 | -81 |
| | Bern | 249 | 0 | 248 | 1 | 251 | -2 |
| | Genf | 489 | 0 | 501 | -12 | 451 | 38 |
| | Lausanne | 461 | 0 | 424 | 37 | 429 | 32 |
| | Zürich | 2'267 | 0 | 1'792 | 475 | 1'943 | 324 |
| | **Total** | 3'848 | 0 | 3'371 | 477 | 3'537 | 311 |

**Tabelle 6:** Provisorische Schätzungen nach Grössenklasse (grkl) aufgrund der neuen Stichprobe für das 1. Quartal 2005 ($\hat{Y}^r_{05,1}$) und das 1. Quartal 2004 ($\hat{Y}^r_{04,1}$). Die Schätzung für das 1. Quartal 2004 aufgrund der alten Stichprobe ist mit $\hat{Y}^a_{04,1}$ bezeichnet. Der Variationskoeffizient $cv$ bezieht sich auf $\hat{Y}^r_{05,1}$, so wie die untere und obere Grenze des Vertrauensintervalls $L^r_{05,1}$ und $U^r_{05,1}$. Diese Grenzen wurden auch für $\hat{Y}^r_{04,1}$ aufgeführt ($L^r_{04,1}$, $U^r_{04,1}$).

| **Variable** | **grkl** | $\hat{Y}^r_{05,1}$ | $\hat{\sigma}$ | $cv$ | $L^r_{05,1}$ | $U^r_{05,1}$ | $\hat{Y}^a_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^a_{04,1}$ | $\hat{Y}^r_{04,1}$ | $L^r_{04,1}$ | $U^r_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^r_{04,1}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 3'140 | 0.00 | 0.00 | 3'140 | 3'140 | 1'878 | 1'262 | 2'185 | 1'972 | 2'398 | 955 |
| NW | 2 | 1'430 | 0.00 | 0.00 | 1'430 | 1'430 | 1'457 | -27 | 1'438 | 1'308 | 1'568 | -8 |
| | 3 | 1'790 | 118.49 | 6.62 | 1'558 | 2'022 | 1'564 | 226 | 1'594 | 1'232 | 1'956 | 196 |
| | 4 | 1'667 | 215.65 | 12.94 | 1'244 | 2'089 | 1'670 | -3 | 1'296 | 871 | 1'720 | 371 |
| | **Total** | 8'027 | 246.06 | 3.07 | 7'544 | 8'509 | 6'569 | 1'458 | 6'513 | 5'910 | 7'117 | 1'514 |

| **Variable** | **grkl** | $\hat{Y}^r_{05,1}$ | $\hat{\sigma}$ | $cv$ | $L^r_{05,1}$ | $U^r_{05,1}$ | $\hat{Y}^a_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^a_{04,1}$ | $\hat{Y}^r_{04,1}$ | $L^r_{04,1}$ | $U^r_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^r_{04,1}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 20'516 | 0.00 | 0.00 | 20'516 | 20'516 | 17'427 | 3'089 | 18'591 | 16'815 | 20'367 | 1'925 |
| IB | 2 | 9'145 | 0.00 | 0.00 | 9'145 | 9'145 | 8'883 | 262 | 9'186 | 8'551 | 9'821 | -41 |
| | 3 | 9'541 | 519.84 | 5.45 | 8'522 | 10'560 | 9'722 | -181 | 9'344 | 8'205 | 10'482 | 197 |
| | 4 | 7'174 | 653.19 | 9.10 | 5'894 | 8'454 | 8'025 | -851 | 7'527 | 5'886 | 9'168 | -353 |
| | **Total** | 46376 | 834.80 | 1.80 | 44'740 | 48'012 | 44'057 | 2'319 | 44'648 | 41'998 | 47'297 | 1'728 |

| **Variable** | **grkl** | $\hat{Y}^r_{05,1}$ | $\hat{\sigma}$ | $cv$ | $L^r_{05,1}$ | $U^r_{05,1}$ | $\hat{Y}^a_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^a_{04,1}$ | $\hat{Y}^r_{04,1}$ | $L^r_{04,1}$ | $U^r_{04,1}$ | $\hat{Y}^r_{05,1} - \hat{Y}^r_{04,1}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 3'944 | 0.00 | 0.00 | 3'944 | 3'944 | 4'260 | -316 | 4'436 | 4'001 | 4'871 | -492 |
| BB | 2 | 2'490 | 0.00 | 0.00 | 2'490 | 2'490 | 1'974 | 516 | 2'019 | 1'841 | 2'197 | 471 |
| | 3 | 2'732 | 182.43 | 6.68 | 2'374 | 3'089 | 2'749 | -17 | 2'728 | 2'291 | 3'164 | 4 |
| | 4 | 2'492 | 526.09 | 21.11 | 1'461 | 3'523 | 2'557 | -65 | 2'388 | 1'740 | 3'036 | 104 |
| | **Total** | 11'658 | 556.82 | 4.78 | 10'566 | 12'749 | 11'540 | 118 | 11'570 | 10'674 | 12'467 | 88 |

# B SAS-Programme

## B.1 Stichprobenplan

Das Hauptprogramm für die Erarbeitung des Stichprobenplans heisst 'stichprobenplan.sas'. Darin werden die nötigen Macros aufgerufen. Danach werden die Programme 'trim.sas' und 'ziehung310305.sas' aufgerufen, in welchen die Daten eingelesen werden, die Schichtung und Zuordnung bestimmt werden und die neue Stichprobe mit der maximalen Überlappung gezogen wird.

```
/*************************************************************/
/*                                                           */
/* Dieses Programm dient zur Erarbeitung des Stichprobenplans    */
/* der revidierten vierteljährlichen Wohnbaustatistik TRIM 2005. */
/*                                                           */
/* Autor: Daniel Kilchmann, METH                             */
/* Erstellt: 22.03.2005                                      */
/* Letzte Änderung: 06.01.2006                               */
/*                                                           */
/*************************************************************/

libname trim0 '/app/home/bfs02/u3902/projets/trim/data';
libname trim1 '/app/home/bfs02/u3902/projets/trim/res';
libname vz00 '/project/vz2000/MainDataSets';

/* Einbinden der Macros */
%include '/app/home/bfs02/u3902/projets/trim/progs/stichprobenplan_final/trim_macro.sas';

/* Einlsesen, Schichtung, Dalenius, Neyman, etc. */
%include '/app/home/bfs02/u3902/projets/trim/progs/stichprobenplan_final/trim.sas';

/* Ziehung der Stichprobe am 31.03.2005: trim1.stichp05_2_final */
%include '/app/home/bfs02/u3902/projets/trim/progs/stichprobenplan_final/ziehung310305.sas';
```

## B.2 Schätzverfahren

Die hauptsächlichen SAS-Programme heissen 'Stichprobe_bestimmen.sas' und 'Hochrechnung.sas'. Damit diese beiden Programme funktionieren, muss die entsprechende Verzeichnisstruktur bestehen, siehe Programme. Zudem wird die SAS-Datei 'hochrechnungs_macro.sas' verwendet und die Datei 'data_handling_macros.sas'.

### Stichprobe_bestimmen.sas

```
/*##########################################################################################*/
/*                                                                                          */
/* Prototyp für die Bestimmung der Stichprobe der vierteljährlichen Wohnbaustatistik.       */
/* Es wird davon ausgegangen, dass die Stichprobe 'stichp2005_1' im SAS-Format im Ver-      */
/* zeichnis 'Stichprobe_verz' (s. Initialisierung) existiert. In diesem Verzeichnis wird    */
/* die SAS-Datei 'stichp<Jahr>_<Quartal>' erstellt, welche alle Gemeinden, die im Quar-     */
/* tal 'Quartal' des Jahres 'Jahr' in der Stichprobe sind. Diese Datei enthält ebenso       */
/* die Schichtung und die Grössenklassen, sowie die Zugehörigkeit zu einer Agglomeration    */
/* und die Gewichtung.                                                                      */
/*                                                                                          */
/* Autor: Daniel Kilchmann, METH.                                                           */
/* Version: 1.0                                                                             */
/* Erstellt: 15.05.2005                                                                     */
/* Letzte Änderung: 16.06.2005                                                              */
/*                                                                                          */
/*##########################################################################################*/
```

```
/* Einbinden der Macros für die Bearbeitung und Schätzungen.                          */
%include '/project/bau/TRIM HOCHRECHNUNG/Programme/hochrechnungs_macro.sas';

/* Setzen der libnames und Initialisierung von Variablen.                             */
/* 'Basispfad' ist der UNIX-Pfad, der zu den untenstehenden Verzeichnissen führt.     */
/* 'Gemeinden_verz' bezeichnet das UNIX-Verzeichnis das zur Textdatei der Gemeindedaten */
/*      führt. Am gleichen Ort werden auch die entsprechenden SAS-Dateien hingeschrieben*/
/* 'Einwohner_verz' bezeichnet das UNIX-Verzeichnis das zur Textdatei der Einwohnerzah- */
/*      len führt. Gleichen Orts werden auch die entsprechenden SAS-Dateien gespeichert.*/
/* 'Ruecklauf_verz' bezeichnet das UNIX-Verzeichnis, das zur Textdatei des Ruecklaufs   */
/*      führt. Die Daten des Ruecklaufs sind nach dem Einlesen auch hier im SAS-Format. */
/* 'Stichprobe_verz' bezeichnet das UNIX-Verzeichnis das zur SAS-Datei der Stichprobe   */
/*      führt, die innerhalb der Aufarbeitung kreiert wird.                           */
/* 'Plaus_verz' ist das UNIX-Verzeichnis, mit Daten mit fehlenden Werten, Ausreissern,  */
/*      etc. und den dazugehörigen Analysen.                                          */
/* 'Schaetzung_verz' bezeichnet das UNIX-Verzeichnis für die SAS-Dateien der Schätzung. */
/* 'Publikation_verz' ist das UNIX-Verzeichnis unter 'Schaetzung_verz', das die Publi-  */
/*      kationstabellen in HTML-Format beinhaltet.                                    */
/* 'Archiv_verz' bezeichnet das UNIX-Verzeichnis für die Archiv-Dateien. Innerhalb von  */
/*      diesem Verzeichnis muss die gleiche Verzeichnisstruktur wie oben existeieren.   */
/* 'Jahr' bezeichnet das aktuelle Jahr.                                               */
/* 'Quartal' bezeichnet das Quartal, für welches die Schätzung gemacht werden soll.     */
%Initialisierung(Basispfad=/project/bau/TRIM HOCHRECHNUNG/,
                 Gemeinden_verz=Gemeinden,
                 Einwohner_verz=Einwohnerzahl,
                 Ruecklauf_verz=Ruecklauf,
                 Stichprobe_verz=Stichprobe,
                 Plaus_verz=Plaus,
                 Schaetzung_verz=Schaetzung,
                 Publikation_verz=Publikation,
                 Archiv_verz=Archiv,
                 Jahr=2005,
                 Quartal=2);

/* Die Gemeinde-Textdatei des aktuellen Jahres wird*/
/* eingelesen.                                     */
%Einlesen_gmde;

/* Liest die Einwohnerzahl aus der Textdatei ein.          */
/* Aktualisierung der Einwohnerzahl, die ab dem Quartal    */
/* 'Einw_quart' und dem Jahr 'Einw_jahr' gültig sind.      */
%Einwohner_aktualisieren(Einw_jahr=2004,Einw_quart=4);

/* Macro anhand dessen die Stichprobe bestimmt wird, indem */
/* Fusionen und Teilungen miteinbezogen wurden.            */
/* Zudem werden die Gewichte aufgrund von Fusionen und     */
/* Teilungen angepasst.                                    */
%Erneuerung_stichprobe;

/* Ende der Bestimmung der Stichprobe.              */
%Trim_stichprobe_best_ende;

/* Ende. */
```

## Hochrechnung.sas

```
/*##############################################################################################*/
/* Prototyp für die Bearbeitung und Hochrechnung der vierteljährlichen Wohnbaustatistik.*/
/* In diesem Programm wird davon ausgegangen, dass zuvor das Programm                   */
/* 'Stichprobe_bestimmen.sas' mit den gleichen Prametern in der Initialisierung ausge-  */
/* führt wurde.                                                                         */
/*                                                                                      */
/* Autor: Daniel Kilchmann, METH.                                                       */
/* Version: 1.0                                                                         */
/* Erstellt: 15.05.2005                                                                 */
/* Letzte Änderung: 08.06.2005                                                          */
/*##############################################################################################*/

/* Einbinden der Macros für die Bearbeitung und Schätzungen.                            */
%include '/project/bau/TRIM HOCHRECHNUNG/Programme/hochrechnungs_macro.sas';

/* Setzen der libnames und Initialisierung von Variablen.                               */
/* 'Basispfad' ist der UNIX-Pfad, der zu den untenstehenden Verzeichnissen führt.       */
/* 'Gemeinden_verz' bezeichnet das UNIX-Verzeichnis der zur Textdatei der Gemeindedaten */
/*      führt. Am gleichen Ort werden auch die entsprechenden SAS-Dateien hingeschrieben*/
/* 'Einwohner_verz' bezeichnet das UNIX-Verzeichnis der zur Textdatei der Einwohnerzah- */
/*      len führt. Gleichen Orts werden auch die entsprechenden SAS-Dateien gespeichert.*/
/* 'Ruecklauf_verz' bezeichnet das UNIX-Verzeichnis, der zur Textdatei des Ruecklaufs   */
/*      führt. Die Daten des Ruecklaufs sind nach dem Einlesen auch hier im SAS-Format. */
/* 'Stichprobe_verz' bezeichnet das UNIX-Verzeichnis der zur SAS-Datei der Stichprobe   */
/*      führt, die innerhalb der Aufarbeitung kreiert wird.                             */
/* 'Plaus_verz' ist das UNIX-Verzeichnis, mit Daten mit fehlenden Werten, Ausreissern,  */
/*      etc. und den dazugehörigen Analysen.                                            */
/* 'Schaetzung_verz' bezeichnet das UNIX-Verzeichnis für die SAS-Dateien der Schätzung. */
/* 'Publikation_verz' ist das UNIX-Verzeichnis unter 'Schaetzung_verz', das die Publi-  */
/*      kationstabellen in HTML-Format beinhaltet.                                      */
/* 'Archiv_verz' bezeichnet das UNIX-Verzeichnis für die Archiv-Dateien. Innerhalb von  */
/*      diesem Verzeichnis muss die gleiche Verzeichnisstruktur wie oben existeieren.   */
/* 'Jahr' bezeichnet das aktuelle Jahr.                                                 */
/* 'Quartal' bezeichnet das Quartal, für welches die Schätzung gemacht werden soll.     */
%Initialisierung(Basispfad=/project/bau/TRIM HOCHRECHNUNG/,
                 Gemeinden_verz=Gemeinden,
                 Einwohner_verz=Einwohnerzahl,
                 Ruecklauf_verz=Ruecklauf,
                 Stichprobe_verz=Stichprobe,
                 Plaus_verz=Plaus,
                 Schaetzung_verz=Schaetzung,
                 Publikation_verz=Publikation,
                 Archiv_verz=Archiv,
                 Jahr=2005,
                 Quartal=2);

/* Der Rücklauf wird aus der Textdatei eingelesen. */
/* Er muss aus Daten des aktuellen Jahres und des  */
/* gleichen Quartals des Vorjahres bestehen.       */
%Einlesen_ruecklauf;

/* Antwortausfälle werden behandelt.               */
%Trim_antwortausfaelle;

/* Konsistenzkontrollen der Daten.                 */
%Trim_konsistenz;

/* Hochrechnung der Daten.                         */
%Trim_hochrechnung;

/* Die Publikationstabellen werden erstellt.       */
%Trim_publikation;

/* Ende der Bearbeitung/Hochrechnung.              */
%Trim_ende;

/* Ende. */
```

# Literatur


[Cochran, 1977] Cochran, W. G. (1977). Sampling Techniques. 3rd ed., Wiley & Sons, New York.

[Deville, Lavallé, 2003] Deville, J.-C., Lavallé, P. (2003). Indirect Sampling: The Foundations of the Generalized Weight Share Method. *Proceedings of the ICRASS'02*, LRSP, Carleton University, Ottawa.

[Hidiroglou, Berthelot, 1986] Hidiroglou, M. A., Berthelot, J.-M. (1986). Statistical Editing and Imputation for Periodic Business Surveys. *Survey Methodology*, Vol. 12, No. 1, 73-83.

[Kisch, Scott, 1971] Kish, L., Scott, A. (1971). Retaining Units After Changing Strata and Probabilities. *Journal of the American Statistical Association*, Vol. 66, No. 335, 461-470.

[Lavallée, 2002] Lavallée, P. (2002). Le sondage indirect, Éditions de l'Université de Bruxelles, Bruxelles.

[Ruckstuhl, 2004] Ruckstuhl, A., (2004). Einführung in die robusten Schätzmethoden. *Unterlage Nachdiplomkurs*, Zürcher Hochschule Winterthur, Winterthur.

[Särndal, Swensson, Wretman, 1997] Särndal, C.-E., Swensson, B. and Wretman, J. (1997). Model assisted Survey Sampling, Springer, New York.

[Schwendener, 1993] Schwendener, P. (1993). Bestimmung einer Zufallsstichprobe für die vierteljährliche Wohnbaustatistik. *Gutachten im Auftrag des BFS*, Basel.

[Sirken, 1972] Sirken, M. G. (1972). Stratified Sample Surveys with Multiplicity. *Journal of the American Statistical Association*, Vol. 67, No. 337, 224-227.

[Zambaz, 2002] Zambaz, J.-M. (2002). Informatiklösung TRIM, Spécification détaillée: Règles de qualité. *Internes Dokument*, BFS, Neuchâtel.

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


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). Erhebung über Forschung und Entwicklung in der schweizerischen Privatwirtschaft 2000. Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. 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

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

# Programme des publications de l'OFS

En sa qualité de service central de statistique de la Confédération, l'Office fédéral de la statistique (OFS) a pour tâche de rendre les informations statistiques accessibles à un large public.

L'information statistique est diffusée par domaine (cf. verso de la première page de couverture); elle emprunte diverses voies:

| *Diffusionsmittel* | *Kontakt*<br>*N° à composer* | *Moyen de diffusion* |
|---|---|---|
| Individuelle Auskünfte | 032 713 60 11<br>info@bfs.admin.ch | Service de renseignements individuels |
| Das BFS im Internet | www.statistik.admin.ch | L'OFS sur Internet |
| Medienmitteilungen zur raschen Information der Öffentlichkeit über die neusten Ergebnisse | www.news-stat.admin.ch | Communiqués de presse: information rapide concernant les résultats les plus récents |
| Publikationen zur vertieften Information (zum Teil auch als Diskette/CD-Rom) | 032 713 60 60<br>order@bfs.admin.ch | Publications: information approfondie (certaines sont disponibles sur disquette/CD-Rom) |
| Online-Datenbank | 032 713 60 86<br>www.statweb.admin.ch | Banque de données (accessible en ligne) |

Nähere Angaben zu den verschiedenen Diffusionsmitteln liefert das laufend nachgeführte Publikationsverzeichnis im Internet unter der Adresse www.statistik.admin.ch→Aktuell→Publikationen.

La liste des publications, mise à jour régulièrement, donne davantage de détails sur les divers moyens de diffusion. Elle se trouve sur Internet à l'adresse www.statistique.admin.ch →Actualités→Publications.

# Methodenberichte des Dienstes Statistische Methoden
# Rapports de méthodes du Service de méthodes statistiques
# Methodological reports of the Statistical Methods Unit

Die Methodenberichte beschreiben die mathematischen und statistischen Methoden, die den Resultaten und Analysen der öffentlichen Statistik zu Grunde 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 sicher zu stellen, 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.

Les rapports de méthodes décrivent les méthodes mathématiques et statistiques à la base des résultats et des analyses de la statistique publique. Ils présentent également l'évaluation et le développement de nouvelles méthodes en vue d'une application future. Ces publications visent d'une part à documenter les méthodes utilisées ou envisagées dans un souci de transparence et de rigueur scientifique, et d'autre part à favoriser la collaboration avec le monde scientifique et universitaire.

Les résultats numériques présentés dans les rapports de méthodes illustrent les concepts mathématiques décrits, mais ne sont pas des résultats officiels des enquêtes concernées. De même, les méthodes réellement appliquées peuvent différer légèrement de celles décrites dans ces rapports.

Les rapports de méthodes sont disponibles sous forme électronique sur le site internet de l'OFS.

Die vierteljährliche Wohnbaustatistik wurde aufgrund der Einwohnerzahlen und der Definition der Agglomerationsgemeinden der Volkszählung 2000 revidiert. Um den Produktionsablauf zu optimieren, wurde eine maximale Überlappung der neuen Erhebung mit der bisherigen angestrebt. Aufgrund der häufigen Gemeindefusionen und stetigen Änderungen der Einwohnerzahlen, wurden diese Einflüsse in den Erhebungsprozess und die Schätzung miteinbezogen. Es werden Datenkontrollen sowie einfache Methoden der Ausreisserlokalisierung diskutiert. Die Schätzverfahren sind geschichtete Totalschätzer mit Korrekturen für allfällige Antwortausfälle. Die Varianzen der Totalschätzer sind zum Teil gross, so dass die Interpretation von Veränderungen beeinträchtigt wird.

Bestellnummer: 338-0035

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

Preis: gratis

ISBN 3-303-00330-0