

Gewichte

Nicolás Albacete, Pirmin Fessler, Peter Lindner

(Abteilung für volkswirtschaftliche Analysen, OeNB)

Technischer Workshop zum HFCS in Österreich

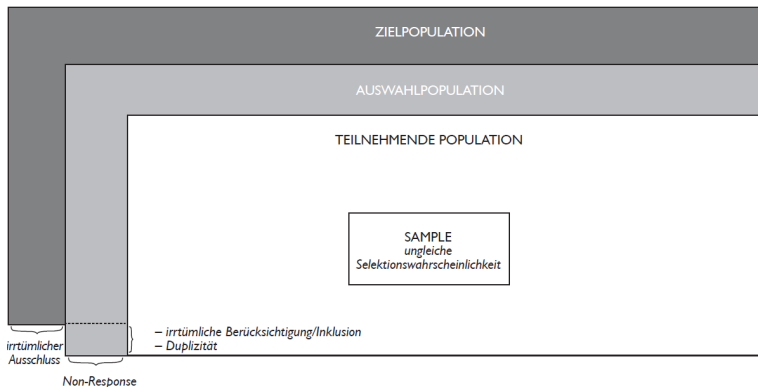
25. und 26.2.2013

Outline

- 1 Fehlerhafte Repräsentativität der Zielpopulation in der Stichprobe
- 2 Wiederherstellung der Repräsentativität durch Gewichtung
- 3 Erstellung der HFCS-Gewichte
 - Design-Gewichte
 - Nonresponse-Gewichte
 - Poststratifizierungs-Gewichte
 - Finale Gewichte
- 4 Ausgewählte Ergebnisse

FEHLERHAFTE REPRÄSENTATIVITÄT DER ZIELPOPULATION IN DER STICHPROBE ↪ Darstellung

Fehlerhafte Repräsentativität der Grundgesamtheit in der Stichprobe



Quelle: Biemer und Christ (2008) (adaptiert).

FEHLERHAFTE REPRÄSENTATIVITÄT DER ZIELPOPULATION IN DER STICHPROBE \leftrightarrow Ungleiche Auswahlwahrscheinlichkeiten

- Private Haushalte in kleinen PSUs haben höhere Auswahlwahrscheinlichkeit als jene in großen PSUs
- Anteil von Wiener Haushalten überproportional hoch
- Unequal Probability Sampling Bias
- Design-Gewichte korrigieren solche Verzerrungen

FEHLERHAFTE REPRÄSENTATIVITÄT DER ZIELPOPULATION IN DER STICHPROBE \leftrightarrow Vollständiger Antwortausfall

- Nur ein Teil der in der Bruttostichprobe enthaltenen Haushalte sind auch tatsächlich bereit, an der Erhebung teilzunehmen
- Bestimmte Gruppen privater Haushalte weisen eine niedrigere Wahrscheinlichkeit auf, am HFCS teilzunehmen, als andere
- Daher sind Schätzungen für die gesamte Auswahlpopulation hinsichtlich dieser Gruppenmerkmale verzerrt, obwohl sie es für die teilnehmende Population nicht sind (nonresponse bias)
- Non-Response-Gewichte korrigieren diesen Bias

FEHLERHAFTE REPRÄSENTATIVITÄT DER ZIELPOPULATION IN DER STICHPROBE↔ Unvollkommenheiten in der Auswahlpopulation

- Irrtümlicher Ausschluss (erroneous exclusion) von privaten Haushalten, z.B. in WGs
- Irrtümliche Berücksichtigung (erroneous inclusion), z.B. Firmen, Pflegeinstitutionen
- Mehrfacherfassung (frame multiplicity), z.B. Pendler
- Frame bias wird durch Design-Gewichten (im Fall irrtümlicher Berücksichtigung oder Mehrfacherfassung) oder durch Poststratifizierungsgewichten (im Fall irrtümlichen Ausschlusses) reduziert

FEHLERHAFTE REPRÄSENTATIVITÄT DER ZIELPOPULATION IN DER STICHPROBE ↔

Ländervergleich

Table 5.1 Response behaviour indicators in the HFCS

Country	Gross sample size	Net sample size	Response rate	Weighted response rate	Refusal rate	Cooperation rate	Contact rate	Eligibility rate
Belgium	11,376	2,364	21.8%	n.a.	57.6%	27.2%	80.1%	95.4%
Germany	20,501	3,565	18.7%	n.a.	69.7%	21.1%	94.2%	92.9%
Greece	6,354	2,971	47.2%	48.7%	46.4%	47.8%	98.7%	99.1%
Spain	11,782	6,197	56.7%*	n.a.	34.8%	58.4%	97.2%	92.6%
France	24,289	15,006	69.0%	69.6%	30.0%	69.0%	100%	89.8%
Italy	15,592	7,951	52.1%*	53.2%	37.8%	57.8%	90.2%	97.8%
Cyprus	3,938	1,237	31.4%	32.4%	56.6%	35.7%	88.0%	100%
Luxembourg	5,000	950	20.0%	19.3%	63.7%	21.0%	95.5%	94.9%
Malta	3,000	843	29.9%	30.4%	34.1%	44.3%	67.5%	94.0%
Netherlands	2,263	1,301	57.5%*	n.a.	42.5%	57.5%	100%	100%
Austria	4,436	2,380	55.7%	56.4%	39.6%	56.7%	98.1%	96.3%
Portugal	8,000	4,404	64.1%	59.0%	10.3%	80.2%	79.9%	85.9%
Slovenia	965	343	36.4%	35.6%	45.9%	41.6%	87.5%	97.8%
Slovakia	2,000	2,057	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Finland	13,525	10,989	82.2%*	85.0%	11.1%	86.2%	95.4%	98.7%

In France and Portugal, survey participation is compulsory for households.

* Response rates for the whole sample; more comparable response rates are the response rates for households interviewed for the first time, which are 40.3% in Spain, 35.0% in Italy and 70.1% in Finland. This figure is not available for the Netherlands.

WIEDERHERSTELLUNG DER REPRÄSENTATIVITÄT DURCH GEWICHTUNG

- Ziel ist die Berechnung eines finalen Survey-Gewichts für jeden Haushalt i , w_i , das für jene Haushalte, die in der Stichprobe im Vergleich zur Zielpopulation überrepräsentiert sind, relativ klein ist und für jene, die unterrepräsentiert sind, relativ groß ist
- Drei Arten von Gewichten: Design-Gewichte (w_{Di}), Nonresponse-Gewichte (w_{NRi}), Poststratifizierungs-Gewichte (w_{PSi})
- Das finale Survey-Gewicht w_i ergibt sich aus dem Produkt dieser drei Gewichte: $w_i = w_{Di} \times w_{NRi} \times w_{PSi}$

HFCS-Design-Gewichte nach Bundesland

	Mittleres Design-Gewicht	Median-Design-Gewicht	Kleinstes Design-Gewicht	Größtes Design-Gewicht
Vorarlberg	895	849	0	1.560
Tirol	872	879	0	1.852
Salzburg	905	969	0	1.712
Oberösterreich	844	857	0	2.394
Kärnten	874	864	0	2.229
Steiermark	902	756	0	3.271
Burgenland	916	1.050	0	1.379
Niederösterreich	931	882	0	2.437
Wien	747	693	0	2.554
Insgesamt	851	833	0	3.271

Quelle: HFCS Austria 2010, OeNB.

NONRESPONSE-GEWICHTE \leftrightarrow Logit-Schätzung

Schätzungen der Response-Propensität auf Basis eines Logit-Regressionsmodells

Kovariaten	Koeffizienten
Schlechter Zustand der Wohnung	-0,260 (0,233)
Wohnsitz auf dem Land	-0,077 (0,0837)
Wohnsitz in Wien	-1,091*** (0,0924)
Weiblicher Interviewer	-0,138* (0,0738)
Alter des Interviewers	-0,005 (0,00359)
Interviewer mit Ausbildung auf Universitätsniveau	-0,550*** (0,0906)
Erfahrung des Interviewers (in Monaten)	0,00172*** (0,000512)
Zahl der Kontaktaufnahmen des Interviewers	-0,022 (0,0376)
Konstante	0,883*** (0,197)
Beobachtungen	4,273

Quelle: HFCS Austria 2010, OeNB.

Anmerkung: Angabe von robusten Standardfehlern in Klammer. *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$

HFCS-Non-Response-Gewichte nach Response-Propensität

Vorhergesagte Response-Propensität (Quintile, in %)	Gewicht
Erstes Quintil (0 bis 40)	2,837
Zweites Quintil (40 bis 61)	2,110
Drittes Quintil (61 bis 64)	1,580
Viertes Quintil (64 bis 67)	1,471
Fünftes Quintil (67 bis 100)	1,457

Quelle: HFCS Austria 2010, OeNB.

POSTSTRATIFIZIERUNGS-GEWICHTE

HFCS-Post-Stratifizierungs-Gewichte für Gemeinde- und Haushaltsgröße

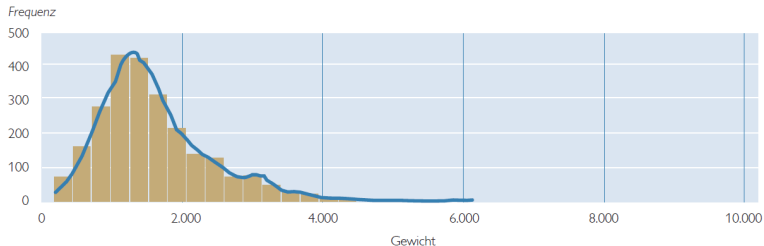
Gemeindegröße (Anzahl der Einwohner)	Haushaltsgröße (Anzahl der Personen)	
	1 bis 4	5 oder mehr
0 bis 2.000	0,974	1,133
2.001 bis 3.000	1,169	1,055
3.001 bis 5.000	0,963	1,447
5.001 bis 10.000	1,053	1,813
10.001 bis 20.000	0,915	2,012
20.001 bis 1.000.000	0,911	1,913
≥ 1.000.001	0,977	1,249

Quelle: HFCS Austria 2010, OeNB.

FINALE GEWICHTE

Minimum=169, Maximum=9.054, Median=1.429, Mittelwert=1.586

Verteilung der endgültigen HFCS-Gewichte



Quelle: HFCS Österreich 2010, OeNB.

Vergleich gewichteter und ungewichteter Mittelwerte von HFCS-Variablen

	Mittelwert	
	ungewichtet	gewichtet
Haushaltsgröße (Anzahl der Personen)	2,11	2,13
<i>in EUR</i>		
Geschätztes monatliches Haushaltsnettoeinkommen	2.300	2.323
Wert des Hauptwohnsitzes	252.686	255.407

Quelle: HFCS Austria 2010, OeNB.