Eine kurze Einleitung in das Sozio-Oekonomische Panel (SOEP)

Teil 4

Gewichtung im SOEP

4 Gewichtung

Empirische Sozialforschung

- Ziel sind Aussagen über größere Grundgesamtheiten
- Inferenz auf Basis kleiner Untersuchungsgruppe (Stichproben)
- Analyseverfahren unterstellen oft einfach Zufallsstichproben

Problem:

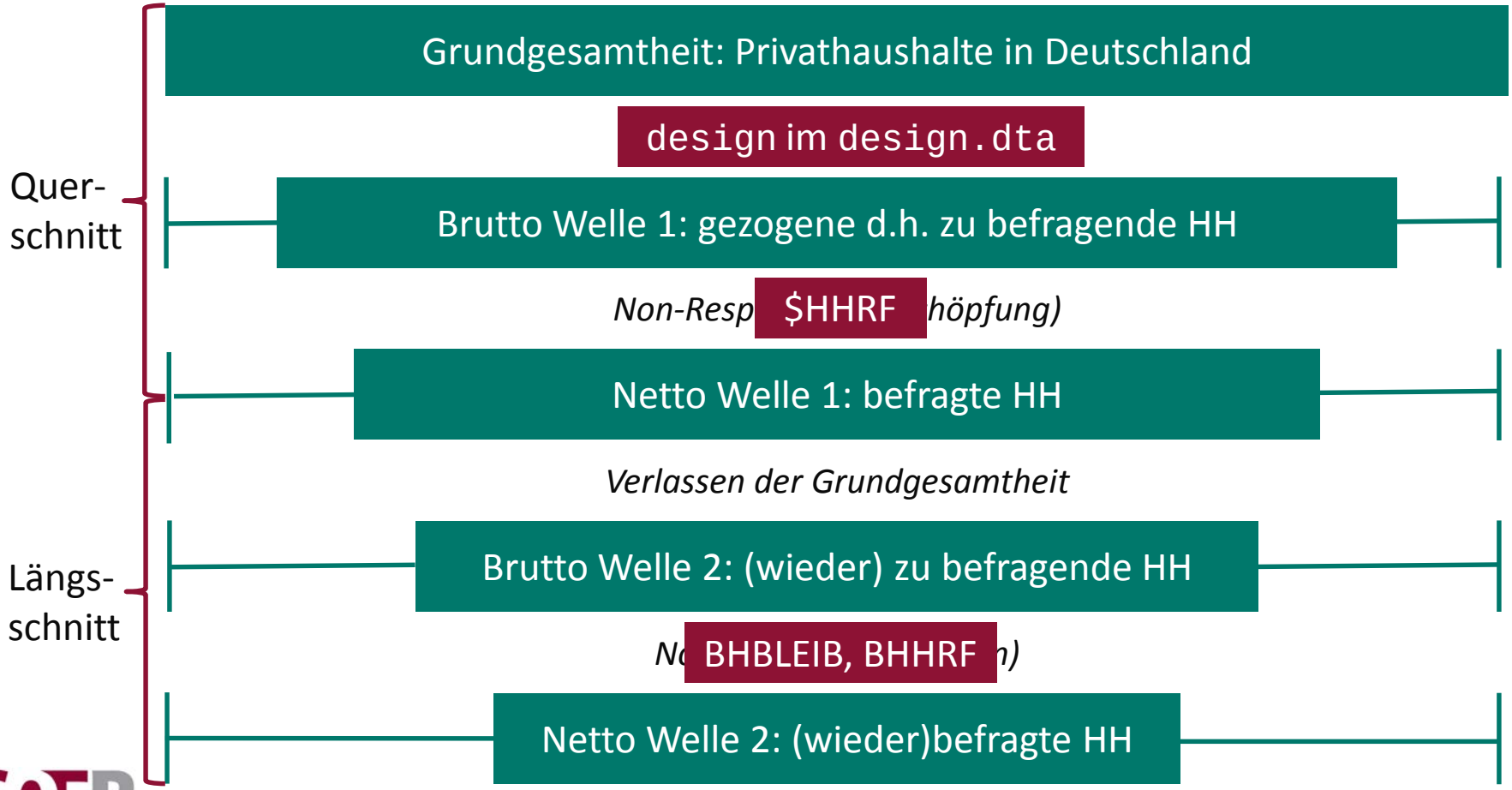
- Alternative Stichprobendesigns haben zwar viele Vorteile können aber zu einer *selection by design* führen
- Gezogene Untersuchungseinheiten können die Teilnahme an der Befragung verweigern (*self-selection*)

→ Ignorieren von Stichprobenselektivität kann Parameter-schätzungen verzerren

Mögliche Lösung: Gewichtung*

4 Gewichtung

Stichprobenentwicklung (vom Brutto zum Netto)



Gewichtung - Design

Stichprobenentwicklung (von der GG zum Brutto)

4 Gewichtung - Design

Grundgesamtheit → Brutto 1: Stichprobendesign `design`

- Disproportionales sampling design etwa nach Regionen und Ziehungs-einheiten (Migranten, Alleinerziehende) → *selection by design*
- Designgewichtung gleicht diese unterschiedlichen, durch den Forscher initiierten Auswahlwahrscheinlichkeiten aus
- Stichprobendesigngewicht = Kehrwert (Inverse π_i^{-1})

  der Ziehungswahrscheinlichkeit (Bsp. $\pi_1 = \frac{n_1}{N_1}$ und $\pi_2 = \frac{n_2}{N_2}$,)

  →Wenn alle Haushalte dieselbe Wahrscheinlichkeit haben Teil der „Stichprobe“ zu werden = Designgewicht eine Konstante (Bsp. Sample K – Aufstockung 2012)

Kroh, M., Siegers, R., & Kühne, S. (2015). Gewichtung und Integration von Auffrischungsstichproben am Beispiel des Sozio-oekonomischen Panels (SOEP). In *Nonresponse Bias* (pp. 409-444). Springer Fachmedien Wiesbaden.

Gewichtung - Design

Grundgesamtheit → Brutto 1: Stichprobendesign `design`

Grundgesamtheit

	West	Ost	Gesamt
Kein Migrationshintergrund	640.000	185.000	825.000
Migrationshintergrund	160.000	200.000	175.000
Gesamt	800.000	200.000	1.000.000

Proportional geschichtete Stichprobenziehung

→ Ziehungswahrscheinlichkeit Nicht-Migranten West: 640/640.000 = 1/1.000

	West	Ost	Gesamt
Kein Migrationshintergrund	640	185	825
Migrationshintergrund	160	200	175
Gesamt	800	200	1.000

Disproportional geschichtete Stichprobenziehung

→ Ziehungswahrscheinlichkeit Nicht-Migranten West: 250/640.000 = 1/2.560

	West	Ost	Gesamt
Kein Migrationshintergrund	250	250	500
Migrationshintergrund	250	250	500
Gesamt	500	500	1.000

→ Stichprobendesigngewicht für Nicht-Migranten in West: π_i^{-1} 2.560

4 Gewichtung - Design

Grundgesamtheit → Brutto 1: Stichprobendesign `design`

Schichtung (`strat`)

- Keine Stratifizierung in A, C, E und H
- Stratifizierung nach Zuwanderer(-gruppen) in B, D, F, J und M1, M2, M3
- Stratifizierung nach Einkommen und Region (Ost/West) in G sowie Einkommen in L2
- Stratifizierung nach Familientypen in L1, L2, L3 (sowie Einkommen in L2)
- Integration von Sub-Stichproben in gesamtes SOEP Form von Stratifizierung

Ziehungswahrscheinlichkeiten (`design`)

- Hohe Ziehungs'wkt (=geringe Gewichte) NBL (C und G)
- Hohe Ziehungs'wkt (=geringe Gewichte) Hocheinkommen ab 2002 (G) und Familien mit geringem Einkommen (L2)
- Hohe Ziehungs'wkt (=geringe Gewichte) Familientypen ab 2010 (L1, L2, L3)
- Hohe Ziehungs'wkt (=geringe Gewichte) Zuwanderer (B, D, F, J und M1, M2, M3)

Gewichtung - Design

Grundgesamtheit → Brutto 1: Stichprobendesign `design`

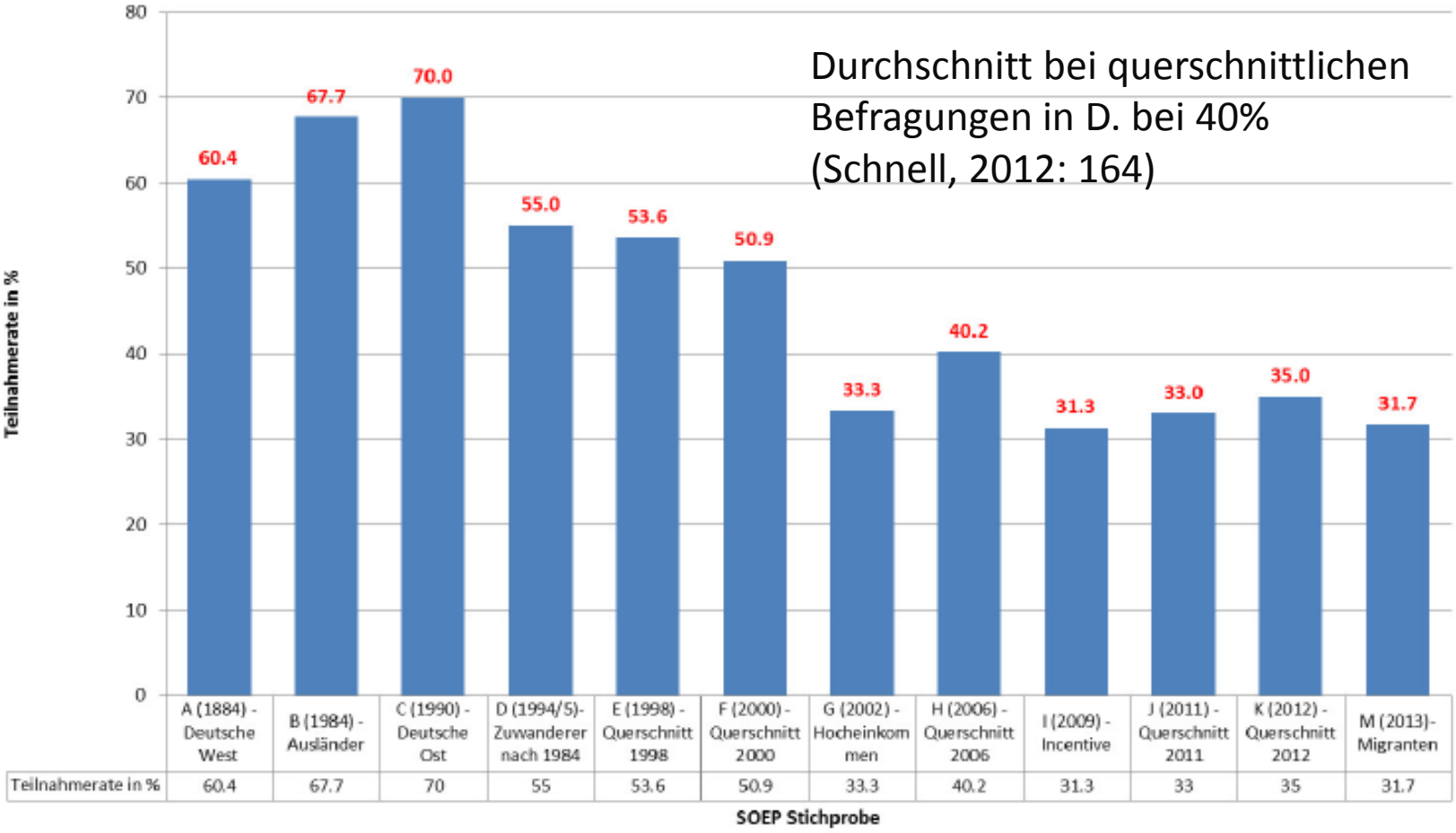
Stichprobenart	min(design)	max(design)
[1] A 1984 Ausgangs-Sample (D-West)	3344	3344
[2] B 1984 Migration (bis 1983, D-West)	68	1697
[3] C 1990 Ausgangs-Sample (D-Ost)	1900	1900
[4] D 1994/5 Migration (1984-92/94 D-Wes	2946	3693
[5] E 1998 Aufstockung	19081	19081
[6] F 2000 Aufstockung	1980	3519
[7] G 2002 Hoch-Einkommen	490	956
[8] H 2006 Aufstockung	10432	10432
[9] I 2009 Innovation Sample	3357	26293
[10] J 2011 Aufstockung	1513	6120
[11] K 2012 Aufstockung	8982	8982
[12] L1 2010 Geburtskohorten (2007-2009)	218	754
[13] L2 2010 Familientypen	722	1290
[14] L3 2011 Familientypen	1394	2496
[15] M1 2013 Migration (1995-2010)	31	957
Total	31	26293

4 Gewichtung - Ausschöpfung

Stichprobenentwicklung (vom Brutto zum Netto)

Gewichtung - Ausschöpfung

Ausschöpfungsquote nach Samples (Welle 1)



Durchschnitt bei querschnittlichen Befragungen in D. bei 40% (Schnell, 2012: 164)

Schnell, R. (2012). *Survey-Interviews. Methoden standardisierter Befragungen*. Berlin: Springer.

Gewichtung - Ausschöpfung

Brutto 1 → Netto 1: non-response, Ausschöpfung

- Nicht alle gezogenen Untersuchungseinheiten nehmen teil, dafür muss kontrolliert werden damit die Stichprobe dennoch repräsentativ für die GG ist
- Schätzung der Responsewahrscheinlichkeit zur Erstellung von Non-Response Gewichten $\hat{\phi}_i^{-1}$ mithilfe einer erweiterten „Inverse Selection Probability“ Gewichtung
  - Erfordert Informationen über (gezogene) Teilnehmer und Nicht-Teilnehmer
- **Lösung**: **a)** Nonresponse-Gewichte werden anhand der gezogenen Brutto-Stichprobe geschätzt und **b)** an bekannte (Rand-) Verteilungen der Grundgesamtheit angepasst

4 Gewichtung - Ausschöpfung

Brutto 1 → Netto 1: non-response, Ausschöpfung

$HHRF

a) **Ausfallanalyse** auf Basis von Informationen zu Brutto 1
   - Interviewerangaben zum Wohnumfeld und Haushalt
   - Befragungsinformationen aus Screening
   - Regionalinformationen zu Kreisen und zur Nachbarschaft (microm)

b) **Randanpassung:**
   - Abweichungen von bekannten Verteilungen in der GG mithilfe des Ranking-Verfahrens (Referenz: Mikrozensus)
   - HH Level: Bundesländer, Gemeindegrößenklasse, Hausbesitz, Haushaltsgröße
   - Personen Level: Altersgruppen , Geschlecht, Nationalität

Kroh, M., Siegers, R., & Kühne, S. (2015). Gewichtung und Integration von Auffrischungsstichproben am Beispiel des Sozio-oekonomischen Panels (SOEP). In *Nonresponse Bias* (pp. 409-444). Springer Fachmedien Wiesbaden.

Gewichtung - Ausschöpfung

Brutto 1 → Netto 1: non-response, Ausschöpfung

$HHRF

Gewichte, die die Stichprobeneinheit auf Summen der GG „hochrechnen“, werden im SOEP Hochrechnungsfaktoren ($HHRF und $PHRF) genannt

Wo sind die HRFs der Welle 1 abgelegt?

- AHHRF für A und B (1984)
- LHHRF für D (1995)
- QHHRFF für F (2000)
- BAHHRFL1 für L1 (2010)
- BBHHRFL3 für L3 (2011)
- BCHHRFK für K (2012)
- GHHRF für C (1990)
- OHHRFE für E (1998)
- SHHRFG für G (2002)
- WHHRFH für H (2006)
- BAHHRFL2 für L2 (2010)
- BBHHRFJ für J (2011)
- BDHHRFM für M1 (2013) im Datensatz HHRF

… entsprechendes gilt für die Personen HRFs im Datensatz PHRF

→ Gewichte können individuell angepasst werden:

1. Herausrechnen der Design-Informationen: AHHRF/design
2. Herausrechnen der querschnittlichen Post-Stratifizierung: XHHRF + YHBLEIB = YHHRFNEU statt YHHRF

4 Gewichtung mit Stata – Beispiel Bildungsabschlüsse

```
. tab casmin84
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	1,067	8.88	8.88
(1b) general elementary school	2,975	24.76	33.64
(1c) basic vocational qualification	4,319	35.94	69.58
(2b) intermediate general qualificatio	591	4.92	74.49
(2a) intermediate vocational	1,580	13.15	87.64
(2c_gen) general maturity certificate	322	2.68	90.32
(2c_voc) vocational maturity certifica	370	3.08	93.40
(3a) lower tertiary education	252	2.10	95.50
(3b) higher tertiary education	541	4.50	100.00
Total	12,017	100.00	

$HHRF

```
. tab casmin84 [aweight=aphrf]
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	300.830687	2.50	2.50
(1b) general elementary school	2,943.2102	24.49	27.00
(1c) basic vocational qualification	4,745.196	39.49	66.48
(2b) intermediate general qualificatio	566.33219	4.71	71.20
(2a) intermediate vocational	1,782.657	14.83	86.03
(2c_gen) general maturity certificate	353.460538	2.94	88.97
(2c_voc) vocational maturity certifica	447.426998	3.72	92.69
(3a) lower tertiary education	301.084046	2.51	95.20
(3b) higher tertiary education	576.80228	4.80	100.00
Total	12,017	100.00	

`aweight` geeignet zur Berechnung von Anteilen, Mittelwerten, etc.
→ Standardfehler nicht robust (ggf. `pweight`)

4 Gewichtung mit Stata – Beispiel Bildungsabschlüsse

```
. tab casmin84 [fweight=round(aphrf)]
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	1,235,350	2.50	2.50
(1b) general elementary school	12,086,105	24.49	27.00
(1c) basic vocational qualification	19,485,818	39.49	66.48
(2b) intermediate general qualificatio	2,325,593	4.71	71.20
(2a) intermediate vocational	7,320,331	14.83	86.03
(2c_gen) general maturity certificate	1,451,471	2.94	88.97
(2c_voc) vocational maturity certifica	1,837,338	3.72	92.69
(3a) lower tertiary education	1,236,389	2.51	95.20
(3b) higher tertiary education	2,368,604	4.80	100.00
Total	49,346,999	100.00	

```
. tab casmin84 [iweight=aphrf]
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	1,235,338	2.50	2.50
(1b) general elementary school	12086069.5	24.49	27.00
(1c) basic vocational qualification	19485787.4	39.49	66.48
(2b) intermediate general qualificatio	2,325,600	4.71	71.20
(2a) intermediate vocational	7,320,346	14.83	86.03
(2c_gen) general maturity certificate	1,451,459	2.94	88.97
(2c_voc) vocational maturity certifica	1,837,325	3.72	92.69
(3a) lower tertiary education	1,236,379	2.51	95.20
(3b) higher tertiary education	2,368,594.8	4.80	100.00
Total	49346898.8	100.00	

$HHRF

Bei der Berechnung von Totals (Bsp: Anzahl an Personen in der Grundgesamtheit mit Merkmal x) ist die `iweight` Option geeignet oder `fweight` (genauer: `[fw=round(weight)]`).

→ Jedoch Standardfehler berechnet als hätte das SOEP wesentlich mehr Fälle

4 Gewichtung mit Stata – Beispiel Bildungsabschlüsse

```
. tab casmin84 [fweight=round(aphrf)]
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	1,235,350	2.50	2.50
(1b) general elementary school	12,086,105	24.49	27.00
(1c) basic vocational qualification	19,485,818	39.49	66.48
(2b) intermediate general qualificatio	2,325,593	4.71	71.20
(2a) intermediate vocational	7,320,331	14.83	86.03
(2c_gen) general maturity certificate	1,451,471	2.94	88.97
(2c_voc) vocational maturity certifica	1,837,338	3.72	92.69
(3a) lower tertiary education	1,236,389	2.51	95.20
(3b) higher tertiary education	2,368,604	4.80	100.00
Total	49,346,999	100.00	

```
. tab casmin84 [iweight=aphrf]
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	1,235,338	2.50	2.50
(1b) general elementary school	12086069.5	24.49	27.00
(1c) basic vocational qualification	19485787.4	39.49	66.48
(2b) intermediate general qualificatio	2,325,600	4.71	71.20
(2a) intermediate vocational	7,320,346	14.83	86.03
(2c_gen) general maturity certificate	1,451,459	2.94	88.97
(2c_voc) vocational maturity certifica	1,837,325	3.72	92.69
(3a) lower tertiary education	1,236,379	2.51	95.20
(3b) higher tertiary education	2,368,594.8	4.80	100.00
Total	49346898.8	100.00	

$HHRF

Bei der Berechnung von Totals (Bsp: Anzahl an Personen in der Grundgesamtheit mit Merkmal x) ist die `iweight` Option geeignet oder `fweight` (genauer: `[fw=round(weight)]`).

→ Jedoch Standardfehler berechnet als hätte das SOEP wesentlich mehr Fälle

4 Gewichtung mit Stata – Beispiel Bildungsabschlüsse

```
. tab casmin84 [fweight=round(aphrf)]
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	1,235,350	2.50	2.50
(1b) general elementary school	12,086,105	24.49	27.00
(1c) basic vocational qualification	19,485,818	39.49	66.48
(2b) intermediate general qualificatio	2,325,593	4.71	71.20
(2a) intermediate vocational	7,320,331	14.83	86.03
(2c_gen) general maturity certificate	1,451,471	2.94	88.97
(2c_voc) vocational maturity certifica	1,837,338	3.72	92.69
(3a) lower tertiary education	1,236,389	2.51	95.20
(3b) higher tertiary education	2,368,604	4.80	100.00
Total	49,346,999	100.00	

```
. tab casmin84 [iweight=aphrf]
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	1,235,338	2.50	2.50
(1b) general elementary school	12086069.5	24.49	27.00
(1c) basic vocational qualification	19485787.4	39.49	66.48
(2b) intermediate general qualificatio	2,325,600	4.71	71.20
(2a) intermediate vocational	7,320,346	14.83	86.03
(2c_gen) general maturity certificate	1,451,459	2.94	88.97
(2c_voc) vocational maturity certifica	1,837,325	3.72	92.69
(3a) lower tertiary education	1,236,379	2.51	95.20
(3b) higher tertiary education	2,368,594.8	4.80	100.00
Total	49346898.8	100.00	

$HHRF

Bei der Berechnung von Totals (Bsp: Anzahl an Personen in der Grundgesamtheit mit Merkmal x) ist die `iweight` Option geeignet oder `fweight` (genauer: `[fw=round(weight)]`).

→ Jedoch Standardfehler berechnet als hätte das SOEP wesentlich mehr Fälle

4 Gewichtung mit Stata – Beispiel Bildungsabschlüsse

```
. tab casmin84 [fweight=round(aphrf)]
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	1,235,350	2.50	2.50
(1b) general elementary school	12,086,105	24.49	27.00
(1c) basic vocational qualification	19,485,818	39.49	66.48
(2b) intermediate general qualificatio	2,325,593	4.71	71.20
(2a) intermediate vocational	7,320,331	14.83	86.03
(2c_gen) general maturity certificate	1,451,471	2.94	88.97
(2c_voc) vocational maturity certifica	1,837,338	3.72	92.69
(3a) lower tertiary education	1,236,389	2.51	95.20
(3b) higher tertiary education	2,368,604	4.80	100.00
Total	49,346,999	100.00	

```
. tab casmin84 [iweight=aphrf]
```

CASMIN-Klassifikation	Freq.	Percent	Cum.
(1a) inadequately completed	1,235,338	2.50	2.50
(1b) general elementary school	12086069.5	24.49	27.00
(1c) basic vocational qualification	19485787.4	39.49	66.48
(2b) intermediate general qualificatio	2,325,600	4.71	71.20
(2a) intermediate vocational	7,320,346	14.83	86.03
(2c_gen) general maturity certificate	1,451,459	2.94	88.97
(2c_voc) vocational maturity certifica	1,837,325	3.72	92.69
(3a) lower tertiary education	1,236,379	2.51	95.20
(3b) higher tertiary education	2,368,594.8	4.80	100.00
Total	49346898.8	100.00	

$HHRF

Bei der Berechnung von Totals (Bsp: Anzahl an Personen in der Grundgesamtheit mit Merkmal x) ist die `iweight` Option geeignet oder `fweight` (genauer: `[fw=round(weight)]`).

→ Jedoch Standardfehler berechnet als hätte das SOEP wesentlich mehr Fälle

4 Gewichtung – Brutto Welle 2

Stichprobenentwicklung (vom Brutto zum Netto)

4 Gewichtung – Brutto Welle 2

Netto 1 → Brutto 2: Verlassen der Grundgesamtheit

Wer soll in der Zweiten Welle wieder befragt werden?

- HH der Welle 1 die weiterhin zur GG gehören (d.h. in Deutschland gemeldet sind)
- Abspaltungen von HH der Welle 1 die zur GG gehören

Wer soll in der zweiten Welle nicht mehr befragt werden?

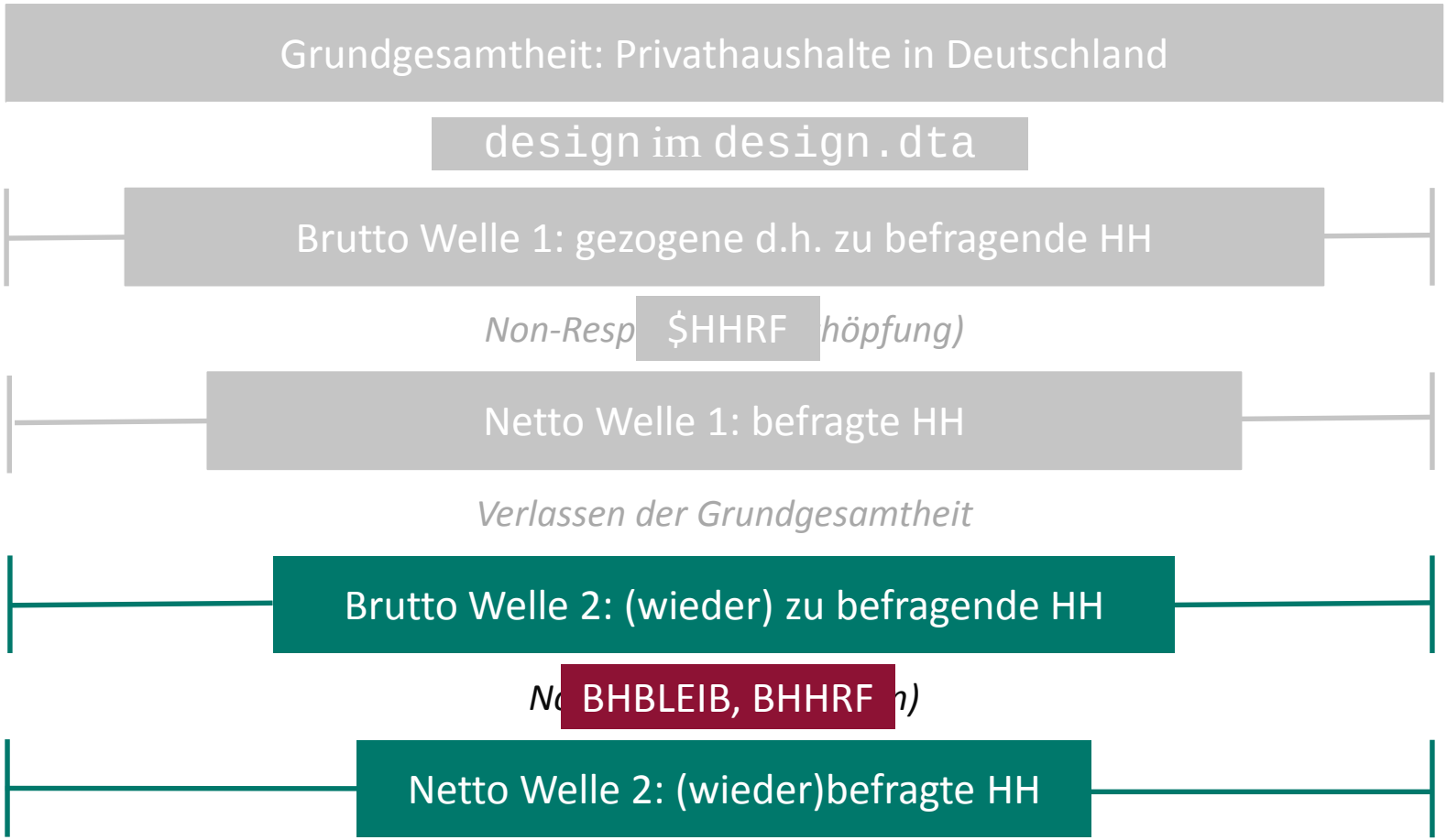
- HH der Welle 1 die nicht mehr zur GG gehören (Tod, Ausland)

Was ist mit Zugängen zur GG zwischen Welle 1 und Welle 2?

- Auffrischungen auf der Haushaltsebene (vgl E, F, H, J, K, und insbes. C, D, M1, M2, M3)
- Zugänge in bestehende HH (Geburt, Zuzug)

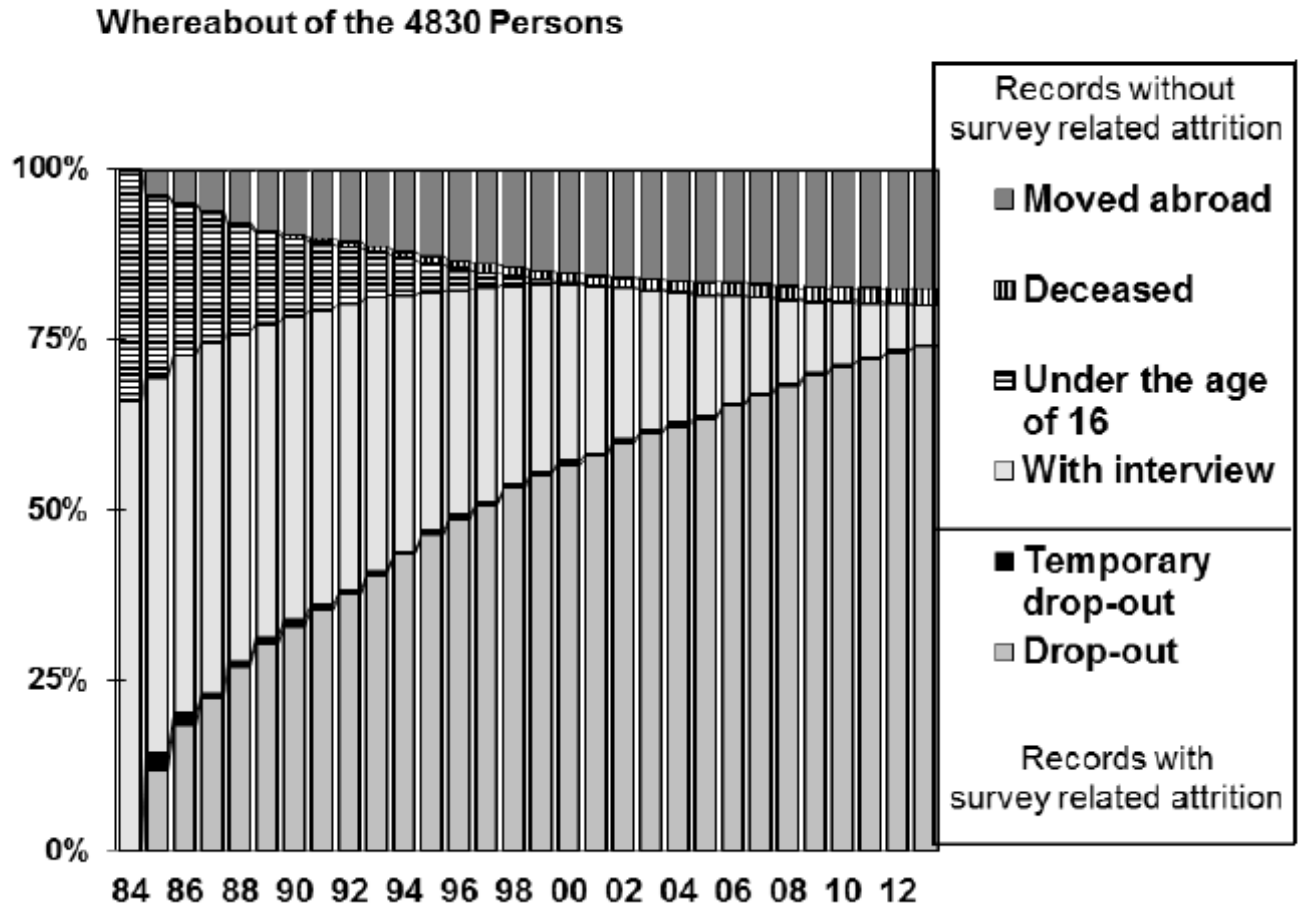
4 Gewichtung – Panel Attrition

Stichprobenentwicklung (vom Brutto zum Netto)



Gewichtung – Panel Attrition

Verbleib der Personen aus Stichprobe B, 1984–2013



4 Gewichtung – Panel Attrition

Brutto 2→ Netto 2: non-response, attrition

$HBLEIB

1. Ausfallanalyse auf Basis von Infos zu Brutto 2 aus t-1
   - Merkmale des Interviews (Mode, Dauer, Interviewerwechsel, …)
   - Merkmale des HH (Vermögen, HH-Größe)
   - Aggregation von Personenmerkmalen über HH (Migrationshintergrund, Arbeitslosigkeit, ..)
2. Schätzung der Kontakt‘wkt und Response‘wkt
   - Modell zur Bestimmung der Wkt, den HH wieder zu kontaktieren (Adressermittlung).
   - Modell zur Bestimmung der Wkt, dass HH zur Wiederbefragung bereit ist (gegen der KK konnte kontaktiert werden)

$$\$HBLEIB = (\text{Kontakt‘wkt} * \text{Response‘wkt})^{-1}$$

Gewichtung – Panel Attrition

Bestimmung der Response‘wkt nach Stichproben in 2010

Table 6a: **Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2010.**

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H	Sample
Intercept	0.26 (0.13)**	1.26 (0.13)***	0.74 (0.16)***	0.04 (0.33) n.s.	1.23 (0.08)***	0.17 (0.13) n.s.	0.29 (0.34) n.s.	-0.37 (0.26) n.s.	-0.26 (0.13)**
Interview Characteristics									
Freshmen			-0.79 (0.34)**						
Original Sample Member		0.50 (0.15)***	0.18 (0.08)**			0.36 (0.06)***			
Moving Out							1.22 (0.50)**		
HH Move	-0.28 (0.11)**								
New HH				-1.59 (0.52)***					
Partial Unit Nonresponse	-0.32 (0.09)***		-0.40 (0.13)***			-0.39 (0.08)***		-0.41 (0.15)***	-0.34 (0.09)**
Temporary Drop-Out		-1.22 (0.40)***			-1.08 (0.33)***				
Email Disclosed	0.15 (0.07)**					0.15 (0.06)**		¤	0.38 (0.10)**
Phone Disclosed	0.44 (0.10)***		0.78 (0.12)***	1.21 (0.33)***		0.73 (0.10)***	0.90 (0.27)***	0.79 (0.22)***	0.64 (0.11)**
Change in Interviewer	-1.03 (0.08)***	-1.12 (0.18)***	-0.94 (0.09)***	-1.24 (0.26)***	-0.74 (0.14)***	-1.16 (0.07)***	-0.95 (0.16)***	-1.16 (0.12)***	-0.70 (0.11)**
Short Interview	0.17 (0.06)***						0.39 (0.15)**		
CAPI	0.25 (0.08)***								
SAQ			0.45 (0.10)***					-0.57 (0.17)***	
Change in Interview Mode		-0.41 (0.20)**		-1.50 (0.35)***		-0.33 (0.08)***	-0.50 (0.21)**		
Mother-Child-Questionnaire								0.69 (0.33)**	
Temp. Related HH			-0.63 (0.18)***			-0.67 (0.16)***			
Refusal Related HH			-0.60 (0.15)***		-0.76 (0.38)**				
Interviewer Related HH	0.13 (0.06)**			0.70 (0.29)**		0.20 (0.07)***			

Note. *** p < 0.01; ** p < 0.05; * p < 0.10; standard errors in parentheses.

4 Gewichtung – Panel Attrition

Brutto 2→ Netto 2: non-response, attrition

$HHRF

1. Bestimmung der Rohgewichte

AHHRF x BHBLEIB = BHHRF

2. Post-Stratifizierung: Abweichungen von Verteilungen der Grundgesamtheit (Referenz: Mikrozensus)
   - HH-Ebene: Bundesland, Gemeindegröße, Haushaltsgröße, Eigentümer/Mieter, Migration, HH-Typen
   - Personenebene: Alter, Geschlecht Migration, HH-Typen
   - Summen der Gewichte werden (iterativ) den Summen in der GG angepasst.

4 Gewichtung

Unterschied $BLEIB und $HRF

- BLEIBs stellen die Stichprobensummen des Samples t-1 wieder her
- HRFs rechnen die Stichprobe auf die GG hoch

Wo sind die BLEIB‘s der Welle 2+ abgelegt?

- BHBLEIB bis BFHBLEIB in HHRF
- … entsprechendes gilt für die Personen-BLEIB‘s in PHRF

Wo sind die HRFs der Welle 1 abgelegt?

- BHHRF bis BFHHRF in HHRF
- … entsprechendes gilt für die Personen-HRFs in PHRF

Kombinierbarkeit der Gewichte

- Querschnittsgewicht $ = **Produkt** aus Ziehungs‘wkt und Response‘wkt in Welle 1 und allen Response‘wkts bis $
- Modulares Prinzip
  1. Ziehungsdesign → Design
  2. Ausfallanalyse Welle 1
  3. Post-Stratifizierung Welle 1 → AHHRF
  4. Ausfallanalyse Welle 2 → BHBLEIB
  5. Post-Stratifizierung welle 2 → BHHRF
  6. …
  7. Ausfallanalyse Welle 32 → BFHBLEIB
  8. Post-Stratifizierung Welle 32 → BFHHRF

4 Gewichtung

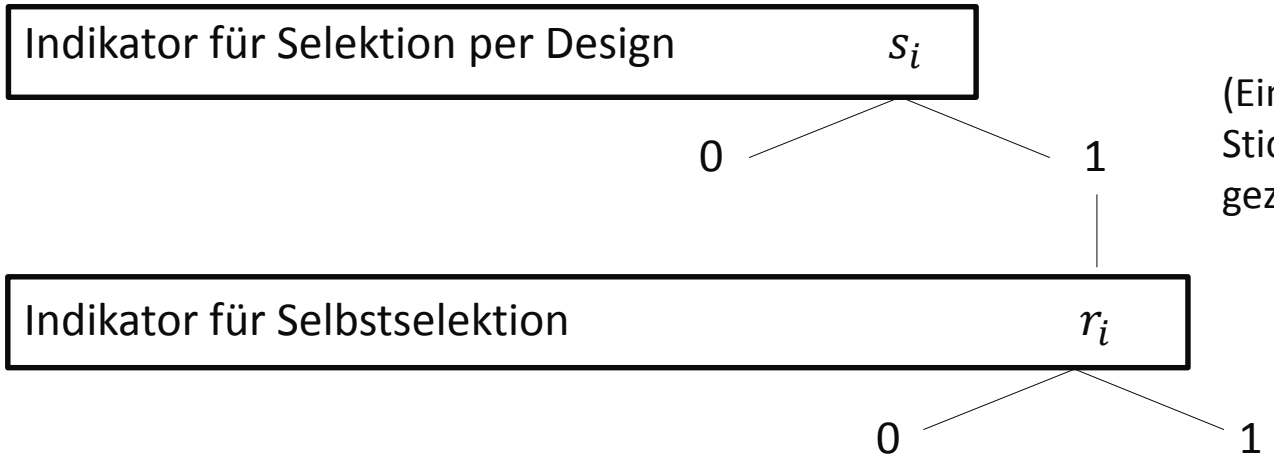
Modulares Prinzip → Möglichkeit der individuellen Manipulation der Gewichte

- Herausrechnen der Design-Informationen:
  AHHRF/DESIGN
- Herausrechnen der querschnittl. Post-Stratifizierung:
  BCHHRF x BDHBLEIB = BDHHRFNEU statt BDHHRF

Backup - Slides

Erweiterte Inverse Selection Probability Gewichtung

Erstellung der Non-Response Gewichte



Beobachtungswahrscheinlichkeit der Einheit i: c_i

$$\Pr(c_i = 1) = \Pr(s_i = 1) \times \Pr(r_i = 1 | s_i = 1) = \left[\pi_i \times \hat{\phi}_i\right]^{-1}$$

Wobei: $\Pr(s_i = 1)$ bekannt ist,

$\Pr(r_i = 1 | s_i = 1)$ unbekannt ist und geschätzt werden muss

Thank you for your attention.

DIW Berlin — Deutsches Institut für Wirtschaftsforschung e.V.
Mohrenstraße 58, 10117 Berlin
www.diw.de

Editor
Sandra Bohmann