

Preliminary Estimate of the Integrated Cross-sectional Weighting (Sub-samples A - G) for 2003

Note

- All processes are only applicable for private households (POP=1 or POP=2); in the case of institutional households, only their dropout behaviour is taken into consideration, i.e. it is not aligned to external structure information.
- Weighting is always carried out separately for new and old federal states.
- In principle, the standard weighting for the sub-samples A-F is always carried out in the first step. High income sub-sample G is then integrated, with respect to its proportion. This proportion is the result of external information on the distribution of private households according to their income class (micro census).

Household Weights

1. The population universe is comprised of all private households in Germany in 2003.
2. In order to calculate the cross-sectional weights for wave 20 (T/2003), all SOEP households who agreed to a household interview and
 - received a household weight in wave 19 or
 - temporarily dropped out in wave 19 or
 - were interviewed for the first time in wave 20were taken into account.

E.g. the snowball ("Schneeball") households in wave D, for whom all weights were reduced to zero, are once again not taken into account.

3. In the first step, we calculated the initial weights for all samples, based on the weights for wave S (2002), by estimating the drop-out probabilities (see Kroh/Spieß 2004).
4. As usual the initial weights for samples A-E and F were marginally aligned to the key data of the 2003 micro census.
 - Household size (5 categories)
 - Nationality (German / non-German head of the household)
 - Gender of the head of the household
 - Age of the household members (17 categories)
5. In contrast to the calculation of the integrated weights for the first wave S (2002), due to the previously calculated dropout probabilities, the structure of the weighting for the sub-sample A-F and the high income sub-sample G was already available for the

second wave T (2003). The only task remaining is the normalisation with regards to the sample sizes and the adherence to the key figures. To do so, the sub-samples A-F and the high income sub-sample G were normalised to their previous year proportions.

6. These normalised initial weights underwent a repeated marginal alignment, which took into account the marginal values of the income categories (albeit on the basis of the 2002 micro census)

up to 3835 €
3835 - 5112 €
5113 + €

as well as the standardised key sums. Furthermore, the shares of sub-samples A-F vs. G which have been provided are kept constant.

Individual Weights

1. The population universe is comprised of the legitimately resident population in private households, who have their main residence in Germany in the year 2003.
2. Every individual is assigned a so-called initial weight. For the sub-samples A-F, this consisted of the individual weight that had been determined by the dropout analysis and marginally aligned to the 17 age categories in 2003.
3. The marginally aligned individual weights and the estimated individual weights for G form new basic values for the alignment to the age distribution. The proportions of the sub-samples A-F and G were defined by the integrated household weights which had previously been established.
4. The normalised individual weights were aligned towards the age structure of 2003.

For information only: integration of sub-sample G in wave 19 (S/2002)

1. This was based on the final weights for A-F and the isolated weights for wave S:
2. Three partial-groups/income categories were established in A-F and G.
 - 1 0 - 3834 € (inc. no answer)
 - 2 3835 - 5112 €
 - 3 5113+ €
3. The next step consisted of establishing the weighting sums and structure of the three partial-groups in sub-samples A-F and in G.
4. Next, it was necessary to establish a principle weighting factor for sub-samples A-F and G; i.e. with which proportions should the partial-groups for the two sub-samples be entered into the integrated weighting A-G. The number of realised cases in the partial-groups was used as the benchmark. Therefore, there were 2*3 weighting figures (samples * income categories) available. These weighting figures were then distributed accordingly. The key figure was now technically

correct, although naturally this was no longer the case for characteristics that were yet to be aligned. Therefore, ...

5. the weights from 4. formed the initial values for a repeated marginal alignment to the established structure. The proportions provided for A-F and G were additionally used as a restriction in order to avoid a mix-up of the sub-samples.
6. Individual weights were initially equated with the household weights. The alignment to the age structure of the population was then carried out as before.

Some key data
Households (inc. institutions)

Variable	Obs	Mean	Std. Dev.	Min	Max
hhnr	11910	280298.6	233365.3	27	824984
hhnrakt	11910	288600.2	225514.2	27	825514
thhrf	11910	3308.706	3678.073	0	38486
thhrff	11910	3301.746	6212.566	0	78905
thhrfae	11910	3314.392	6065.723	0	69974
thhrfag	11910	3309.288	3525.59	21	38627

-> tabulation of privat by tsampreg

privat	current sample region 2003		Total
	W Germany	E Germany	
0	83	22	105
1	9,000	2,805	11,805
Total	9,083	2,827	11,910

privat	current sample region 2003		Total
	W Germany	E Germany	
0	356,705	104,905	461,610
	77.27	22.73	100.00
	1.11	1.43	1.17
1	31730062	7215013	38945075
	81.47	18.53	100.00
	98.89	98.57	98.83
Total	32086767	7319918	39406685
	81.42	18.58	100.00
	100.00	100.00	100.00

. tab2 privat tsampreg [w=**thhrfag**], col row

	current sample region 2003		Total
privat	W Germany	E Germany	
0	363,657	104,905	468,562
	77.61	22.39	100.00
	1.13	1.43	1.19
1	31730034	7215025	38945059
	81.47	18.53	100.00
	98.87	98.57	98.81
Total	32093691	7319930	39413621
	81.43	18.57	100.00
	100.00	100.00	100.00

. tab2 privat tsampreg [w=**thhrfae**], col row

	current sample region 2003		Total
privat	W Germany	E Germany	
0	390,768	138,602	529,370
	73.82	26.18	100.00
	1.22	1.88	1.34
1	31730051	7214984	38945035
	81.47	18.53	100.00
	98.78	98.12	98.66
Total	32120819	7353586	39474405
	81.37	18.63	100.00
	100.00	100.00	100.00

. tab2 privat tsampreg [w=**thhrff**], col row

	current sample region 2003		Total
privat	W Germany	E Germany	
0	315,074	63,722	378,796
	83.18	16.82	100.00
	0.98	0.88	0.96
1	31730000	7215004	38945004
	81.47	18.53	100.00
	99.02	99.12	99.04
Total	32045074	7278726	39323800
	81.49	18.51	100.00
	100.00	100.00	100.00

Only Private Households From Here On!

```
. tab2 thgr tsampreg
```

```
-> tabulation of thgr by tsampreg
```

size of household	current sample region 2003		Total
	W Germany	E Germany	
1	2,205	706	2,911
2	3,239	1,091	4,330
3	1,500	520	2,020
4	1,426	377	1,803
5	630	111	741
Total	9,000	2,805	11,805

```
. tab2 thgr tsampreg [w=thhrf ], col row  
(frequency weights assumed)
```

```
-> tabulation of thgr by tsampreg
```

size of household	current sample region 2003		Total
	W Germany	E Germany	
1	11790990	2635000	14425990
	81.73	18.27	100.00
	37.16	36.52	37.04
2	10622003	2547990	13169993
	80.65	19.35	100.00
	33.48	35.32	33.82
3	4279996	1181997	5461993
	78.36	21.64	100.00
	13.49	16.38	14.02
4	3582001	687,006	4269007
	83.91	16.09	100.00
	11.29	9.52	10.96
5	1455004	163,001	1618005
	89.93	10.07	100.00
	4.59	2.26	4.15
Total	31729994	7214994	38944988
	81.47	18.53	100.00
	100.00	100.00	100.00

. tab2 thgr tsampreg [w=thhrfag], col row

size of household	current sample region 2003		Total
	W Germany	E Germany	
1	11791018	2635004	14426022
	81.73	18.27	100.00
	37.16	36.52	37.04
2	10622004	2547998	13170002
	80.65	19.35	100.00
	33.48	35.32	33.82
3	4279998	1182013	5462011
	78.36	21.64	100.00
	13.49	16.38	14.02
4	3582005	687,000	4269005
	83.91	16.09	100.00
	11.29	9.52	10.96
5	1455008	162,999	1618007
	89.93	10.07	100.00
	4.59	2.26	4.15
Total	31730033	7215014	38945047
	81.47	18.53	100.00
	100.00	100.00	100.00

. tab2 thgr tsampreg [w=thhrfae], col row

size of household	current sample region 2003		Total
	W Germany	E Germany	
1	11791018	2634994	14426012
	81.73	18.27	100.00
	37.16	36.52	37.04
2	10621999	2548002	13170001
	80.65	19.35	100.00
	33.48	35.32	33.82
3	4280009	1182012	5462021
	78.36	21.64	100.00
	13.49	16.38	14.02
4	3582017	687,006	4269023
	83.91	16.09	100.00
	11.29	9.52	10.96
5	1455003	162,999	1618002
	89.93	10.07	100.00
	4.59	2.26	4.15
Total	31730046	7215013	38945059
	81.47	18.53	100.00
	100.00	100.00	100.00

. tab2 thgr tsampreg [w=thrrff], col row

size of household	current sample region 2003		Total
	W Germany	E Germany	
1	11791009	2634996	14426005
	81.73	18.27	100.00
	37.16	36.52	37.04
2	10622012	2548002	13170014
	80.65	19.35	100.00
	33.48	35.32	33.82
3	4280007	1182000	5462007
	78.36	21.64	100.00
	13.49	16.38	14.02
4	3582004	686,999	4269003
	83.91	16.09	100.00
	11.29	9.52	10.96
5	1455005	162,998	1618003
	89.93	10.07	100.00
	4.59	2.26	4.15
Total	31730037	7214995	38945032
	81.47	18.53	100.00
	100.00	100.00	100.00

. tab2 hsample tsampreg

sample type	current sample region 2003		Total
	W Germany	E Germany	
a germans, west	3,008	27	3,035
b foreigners, west	735	1	736
c germans, east	187	1,607	1,794
d immigrants 1984-93	241	3	244
e supplement 1998	601	137	738
f supplement 2000	3,412	938	4,350
g high income	816	92	908
Total	9,000	2,805	11,805

```
. tab2 hsample tsampreg [w=thhrf ], col row
```

sample type	current sample region 2003		Total
	W Germany	E Germany	
a germans, west	12760434	148,609	12909043
	98.85	1.15	100.00
	40.22	2.06	33.15
b foreigners, west	750,110	1,176	751,286
	99.84	0.16	100.00
	2.36	0.02	1.93
c germans, east	511,899	3136532	3648431
	14.03	85.97	100.00
	1.61	43.47	9.37
d immigrants 1984-93	640,643	3,067	643,710
	99.52	0.48	100.00
	2.02	0.04	1.65
e supplement 1998	2788435	678,864	3467299
	80.42	19.58	100.00
	8.79	9.41	8.90
f supplement 2000	14278541	3246765	17525306
	81.47	18.53	100.00
	45.00	45.00	45.00
Total	31730062	7215013	38945075
	81.47	18.53	100.00
	100.00	100.00	100.00

. tab2 hsample tsampreg [w=thhrfag], col row

sample type	current sample region 2003		Total
	W Germany	E Germany	
a germans, west	11913849	141,335	12055184
	98.83	1.17	100.00
	37.55	1.96	30.95
b foreigners, west	723,404	1,092	724,496
	99.85	0.15	100.00
	2.28	0.02	1.86
c germans, east	486,091	3064874	3550965
	13.69	86.31	100.00
	1.53	42.48	9.12
d immigrants 1984-93	618,652	2,955	621,607
	99.52	0.48	100.00
	1.95	0.04	1.60
e supplement 1998	2630098	663,047	3293145
	79.87	20.13	100.00
	8.29	9.19	8.46
f supplement 2000	13559256	3179501	16738757
	81.01	18.99	100.00
	42.73	44.07	42.98
g high income	1798684	162,221	1960905
	91.73	8.27	100.00
	5.67	2.25	5.04
Total	31730034	7215025	38945059
	81.47	18.53	100.00
	100.00	100.00	100.00

. tab2 hsample tsampreg [w=thhrfae], col row

sample type	current sample region 2003		Total
	W Germany	E Germany	
a germans, west	23200796	270,199	23470995
	98.85	1.15	100.00
	73.12	3.74	60.27
b foreigners, west	1363830	2,137	1365967
	99.84	0.16	100.00
	4.30	0.03	3.51
c germans, east	930,731	5702784	6633515
	14.03	85.97	100.00
	2.93	79.04	17.03
d immigrants 1984-93	1164824	5,577	1170401
	99.52	0.48	100.00
	3.67	0.08	3.01
e supplement 1998	5069870	1234287	6304157
	80.42	19.58	100.00
	15.98	17.11	16.19
Total	31730051	7214984	38945035
	81.47	18.53	100.00
	100.00	100.00	100.00

. tab2 hsample tsampreg [w=thhrff], col row

sample type	current sample region 2003		Total
	W Germany	E Germany	
f supplement 2000	31730000	7215004	38945004
	81.47	18.53	100.00
	100.00	100.00	100.00
Total	31730000	7215004	38945004
	81.47	18.53	100.00
	100.00	100.00	100.00

Individuals(inc. individuals in institutional households)

Variable	Obs	Mean	Std. Dev.	Min	Max
hhnr	27985	290089	243427.4	27	824984
persnr	27985	2959275	2388738	201	8254402
tphrf	27985	2935.503	3462.408	0	40655
tphrff	27985	2934.288	5786.511	0	80525
tphrfae	27985	2936.496	5596.091	0	73918
tphrfag	27985	2935.505	3152.292	0	39923

. generate privat=1

. tab2 privat tsampreg

private	current sample region 2003		Total
	W Germany	E Germany	
0	111	30	141
1	21,503	6,341	27,844
Total	21,614	6,371	27,985

. tab2 privat tsampreg [w=**tphrf**], col row

private	current sample region 2003		Total
	W Germany	E Germany	
0	448,852	117,150	566,002
	79.30	20.70	100.00
	0.67	0.79	0.69
1	66912053	14671996	81584049
	82.02	17.98	100.00
	99.33	99.21	99.31
Total	67360905	14789146	82150051
	82.00	18.00	100.00
	100.00	100.00	100.00

. tab2 privat tsampreg [w=**tphrfag**], col row

	current sample region 2003		Total
private	W Germany	E Germany	
0	448,852	117,150	566,002
	79.30	20.70	100.00
	0.67	0.79	0.69
1	66912126	14671990	81584116
	82.02	17.98	100.00
	99.33	99.21	99.31
Total	67360978	14789140	82150118
	82.00	18.00	100.00
	100.00	100.00	100.00

. tab2 privat tsampreg [w=**tphrfae**], col row

	current sample region 2003		Total
private	W Germany	E Germany	
0	440,393	153,402	593,795
	74.17	25.83	100.00
	0.65	1.03	0.72
1	66912031	14672005	81584036
	82.02	17.98	100.00
	99.35	98.97	99.28
Total	67352424	14825407	82177831
	81.96	18.04	100.00
	100.00	100.00	100.00

. tab2 privat tsampreg [w=**tphrff**], col row
(frequency weights assumed)

	current sample region 2003		Total
private	W Germany	E Germany	
0	459,179	72,842	532,021
	86.31	13.69	100.00
	0.68	0.49	0.65
1	66912003	14672025	81584028
	82.02	17.98	100.00
	99.32	99.51	99.35
Total	67371182	14744867	82116049
	82.04	17.96	100.00
	100.00	100.00	100.00

Only individuals in private households in their main residence from here on!

Population membership	current sample region		Total
	2003	2003	
	W Germany	E Germany	
private_hh., german	19,124	6,334	25,458
private_hh., foreign	2,379	7	2,386
Total	21,503	6,341	27,844

. tab2 tpop tsampreg [w=**tphrf**], col row

Population membership	current sample region		Total
	2003	2003	
	W Germany	E Germany	
private_hh., german	60232444	14586380	74818824
	80.50	19.50	100.00
	90.02	99.42	91.71
private_hh., foreign	6679609	85,616	6765225
	98.73	1.27	100.00
	9.98	0.58	8.29
Total	66912053	14671996	81584049
	82.02	17.98	100.00
	100.00	100.00	100.00

. tab2 tpop tsampreg [w=**tphrfag**], col row

Population membership	current sample region		Total
	2003	2003	
	W Germany	E Germany	
private_hh., german	60610432	14589679	75200111
	80.60	19.40	100.00
	90.58	99.44	92.17
private_hh., foreign	6301694	82,311	6384005
	98.71	1.29	100.00
	9.42	0.56	7.83
Total	66912126	14671990	81584116
	82.02	17.98	100.00
	100.00	100.00	100.00

. tab2 tpop tsampreg [w=**tphrfae**], col row

Population membership 2003	current sample region 2003		Total
	W Germany	E Germany	
	private_hh., german	60171439	
	80.51	19.49	100.00
	89.93	99.25	91.60
private_hh., foreign	6740592	110,233	6850825
	98.39	1.61	100.00
	10.07	0.75	8.40
Total	66912031	14672005	81584036
	82.02	17.98	100.00
	100.00	100.00	100.00

. tab2 tpop tsampreg [w=**tphrff**], col row

Population membership 2003	current sample region 2003		Total
	W Germany	E Germany	
	private_hh., german	60306980	
	80.49	19.51	100.00
	90.13	99.62	91.84
private_hh., foreign	6605023	55,527	6660550
	99.17	0.83	100.00
	9.87	0.38	8.16
Total	66912003	14672025	81584028
	82.02	17.98	100.00
	100.00	100.00	100.00

. tab2 psample tsampreg

sample type	current sample region 2003		Total
	W Germany	E Germany	
	a germans (west)	6,799	
b foreigners (west)	2,108	1	2,109
c germans (east)	392	3,723	4,115
d immigrants 1984-93	756	12	768
e supplement 1998	1,318	289	1,607
f supplement 2000	7,886	1,976	9,862
g high income	2,244	282	2,526
Total	21,503	6,341	27,844

```
. tab2 psample tsampreg [w=tphrf ], col row
```

sample type	current sample region 2003		Total
	W Germany	E Germany	
a germans (west)	25415270	253,865	25669135
	99.01	0.99	100.00
	37.98	1.73	31.46
b foreigners (west)	2883732	3,637	2887369
	99.87	0.13	100.00
	4.31	0.02	3.54
c germans (east)	1067445	6403211	7470656
	14.29	85.71	100.00
	1.60	43.64	9.16
d immigrants 1984-93	1732901	12,694	1745595
	99.27	0.73	100.00
	2.59	0.09	2.14
e supplement 1998	5702314	1396202	7098516
	80.33	19.67	100.00
	8.52	9.52	8.70
f supplement 2000	30110391	6602387	36712778
	82.02	17.98	100.00
	45.00	45.00	45.00
Total	66912053	14671996	81584049
	82.02	17.98	100.00
	100.00	100.00	100.00

. tab2 psample tsampreg [w=**tphrfag**], col row

sample type	current sample region 2003		Total
	W Germany	E Germany	
a germans (west)	23689609	247,124	23936733
	98.97	1.03	100.00
	35.40	1.68	29.34
b foreigners (west)	2675953	3,570	2679523
	99.87	0.13	100.00
	4.00	0.02	3.28
c germans (east)	993,119	6213030	7206149
	13.78	86.22	100.00
	1.48	42.35	8.83
d immigrants 1984-93	1608793	12,383	1621176
	99.24	0.76	100.00
	2.40	0.08	1.99
e supplement 1998	5294412	1350989	6645401
	79.67	20.33	100.00
	7.91	9.21	8.15
f supplement 2000	28032415	6403979	34436394
	81.40	18.60	100.00
	41.89	43.65	42.21
g high income	4617825	440,915	5058740
	91.28	8.72	100.00
	6.90	3.01	6.20
Total	66912126	14671990	81584116
	82.02	17.98	100.00
	100.00	100.00	100.00

. tab2 psample tsampreg [w=**tphrfae**], col row

sample type	current sample region 2003		Total
	W Germany	E Germany	
a germans (west)	46209514	461,576	46671090
	99.01	0.99	100.00
	69.06	3.15	57.21
b foreigners (west)	5243128	6,613	5249741
	99.87	0.13	100.00
	7.84	0.05	6.43
c germans (east)	1940806	11642185	13582991
	14.29	85.71	100.00
	2.90	79.35	16.65
d immigrants 1984-93	3150726	23,078	3173804
	99.27	0.73	100.00
	4.71	0.16	3.89
e supplement 1998	10367857	2538553	12906410
	80.33	19.67	100.00
	15.49	17.30	15.82
Total	66912031	14672005	81584036
	82.02	17.98	100.00
	100.00	100.00	100.00

. tab2 psample tsampreg [w=**tphrff**], col row

sample type	current sample region 2003		Total
	W Germany	E Germany	
f supplement 2000	66912003	14672025	81584028
	82.02	17.98	100.00
	100.00	100.00	100.00
Total	66912003	14672025	81584028
	82.02	17.98	100.00
	100.00	100.00	100.00