

Research Notes

28

Martin Spieß and
Markus Pannenberg

Documentation of Sample Sizes and
Panel Attrition in the German Socio
Economic Panel (GSOEP)
(1984 until 2002)

Berlin, September 2003



DIW Berlin

German Institute
for Economic Research

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

DIW Berlin

German Institute
for Economic Research

Königin-Luise-Str. 5
14195 Berlin,
Germany

Phone +49-30-897 89-0

Fax +49-30-897 89-200

www.diw.de

ISSN 1619-4551

**Documentation of Sample Sizes
and Panel Attrition in the
German Socio Economic Panel (GSOEP)
(1984 until 2002)**

by

Martin Spiess* and Markus Pannenberg

Berlin, September 2003

Deutsches Institut für Wirtschaftsforschung, Berlin
Königin-Luise-Str. 5, 14195 Berlin
Phone: +49-30-89789- 0
Fax: +49-30-89789- 200
Internet: <http://www.diw-berlin.de>

* DIW Berlin, Longitudinal Data and Microanalysis. Homepage: <http://www.diw.de/gsoep/> ; resp. <http://www.diw.de/programme/jsp/MA.jsp?uid=mspiess&language=en> ; e-mail: mspiess@diw.de; phone: +49-30-897-89-602.
I would like to thank Daniel Wachtlin for excellent research assistance.

Contents

1	Development of sample sizes	2
1.1	Development of the number of successful interviews by cross-section	2
1.2	Longitudinal development of losses due to panel attrition	14
1.3	Entrants by birth or move-ins and their participation behavior	20
2	Losses due to unsuccessful follow-up	21
2.1	Drop-out rates by mobility behavior	21
2.2	Definition of the regressors for a Logit analysis	24
2.3	Estimated coefficients of the Logit model	25
3	Losses due to refusals	31
3.1	Drop-out rates by different household characteristics	31
3.2	Definition of the regressors for a Logit analysis	33
3.3	Estimated coefficients of the Logit model	36
4	References	47

1 Development of sample sizes

General comment: The sample sizes of the English public use version of the GSOEP and the German DIW version differ by approximately five percent. The exclusion of 5 percent of the original data from the GSOEP was necessary to meet the requirements of the German data protection laws. Technically, this was done by dropping randomly 5 percent of the original wave 1 households. All persons and households which stem from these root households are excluded from the English public use version. Hence the difference in sample sizes is not always exactly 5 percent. The sample sizes documented below refer to the original DIW data base.

With respect to the development of sample sizes our focus is on:

- Comparison of the number of successful interviews by cross-section.
- Longitudinal development of panel attrition.
- Entrants by birth or move-ins and their participation behavior.

1.1 Development of the number of successful interviews by cross-section

The following figures display the number of successful interviews considering different aspects:

Figure 1 Comparison for individuals and households (subsamples A and B), waves 1 to 19 (1984 – 2002).

Figure 2 Comparison between subsamples A and B on the individual level, waves 1 to 18 (1984–2002).

Figure 3 Comparison for individuals and households (subsample C), waves 1 to 12, (1990–2002)

Figure 4 Comparison between the subsamples A, B and C on the individual level, waves 1 to 13.

Figure 5 Comparison for individuals and households in Subsample D, waves 1 to 8, (1995–2002).

Figure 6 Comparison for individuals and households in Subsample E, waves 1 to 5, (1998–2002).

Figure 7 Comparison for individuals and households in Subsample F, waves 1 to 3, (2000–2002).

Figure 1: Comparison of successful interviews with persons and households (subsample A and B), waves 1 to 19.

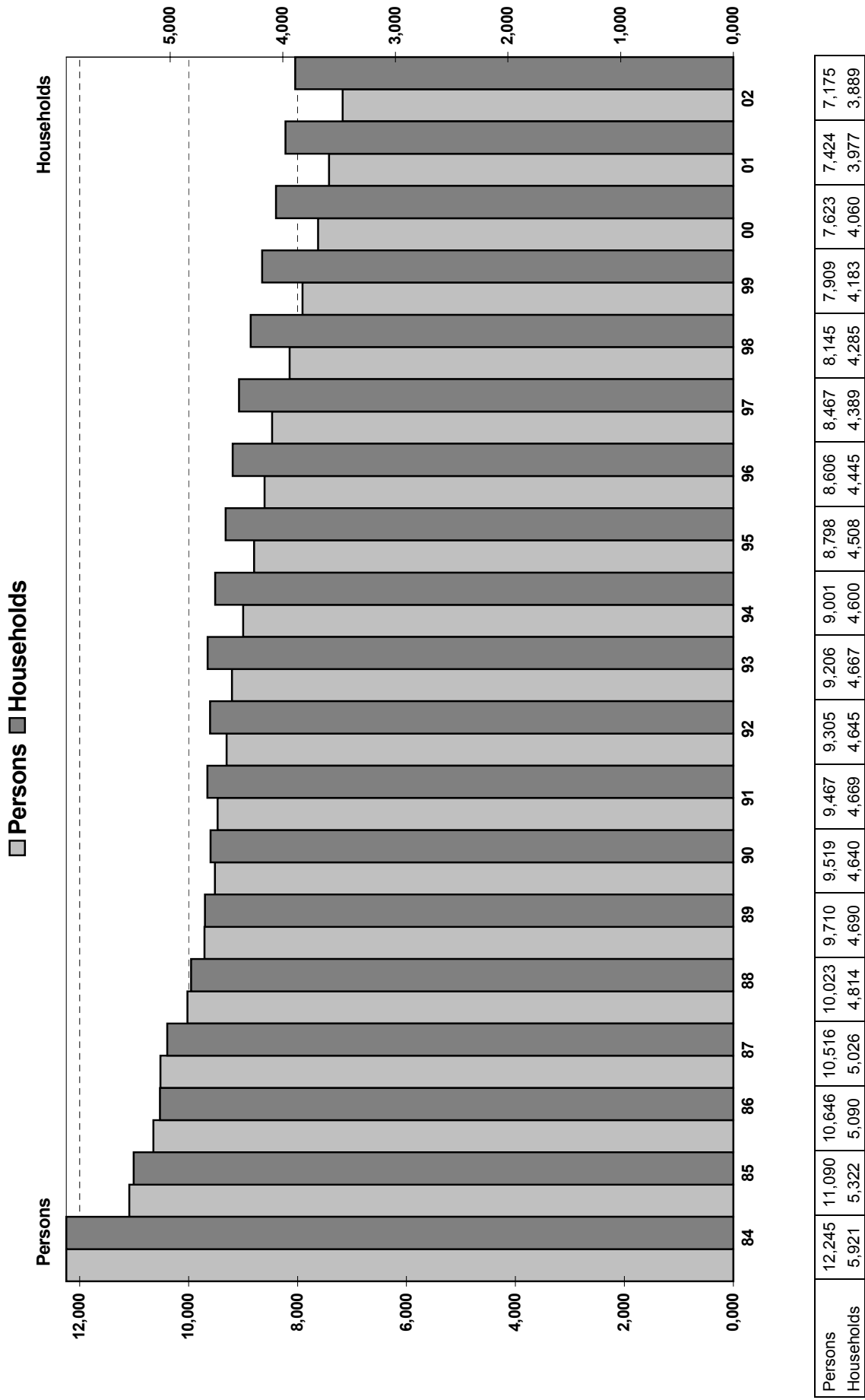


Figure 2: Comparison of successful interviews between subsamples A and B (individual level), waves 1 to 19.

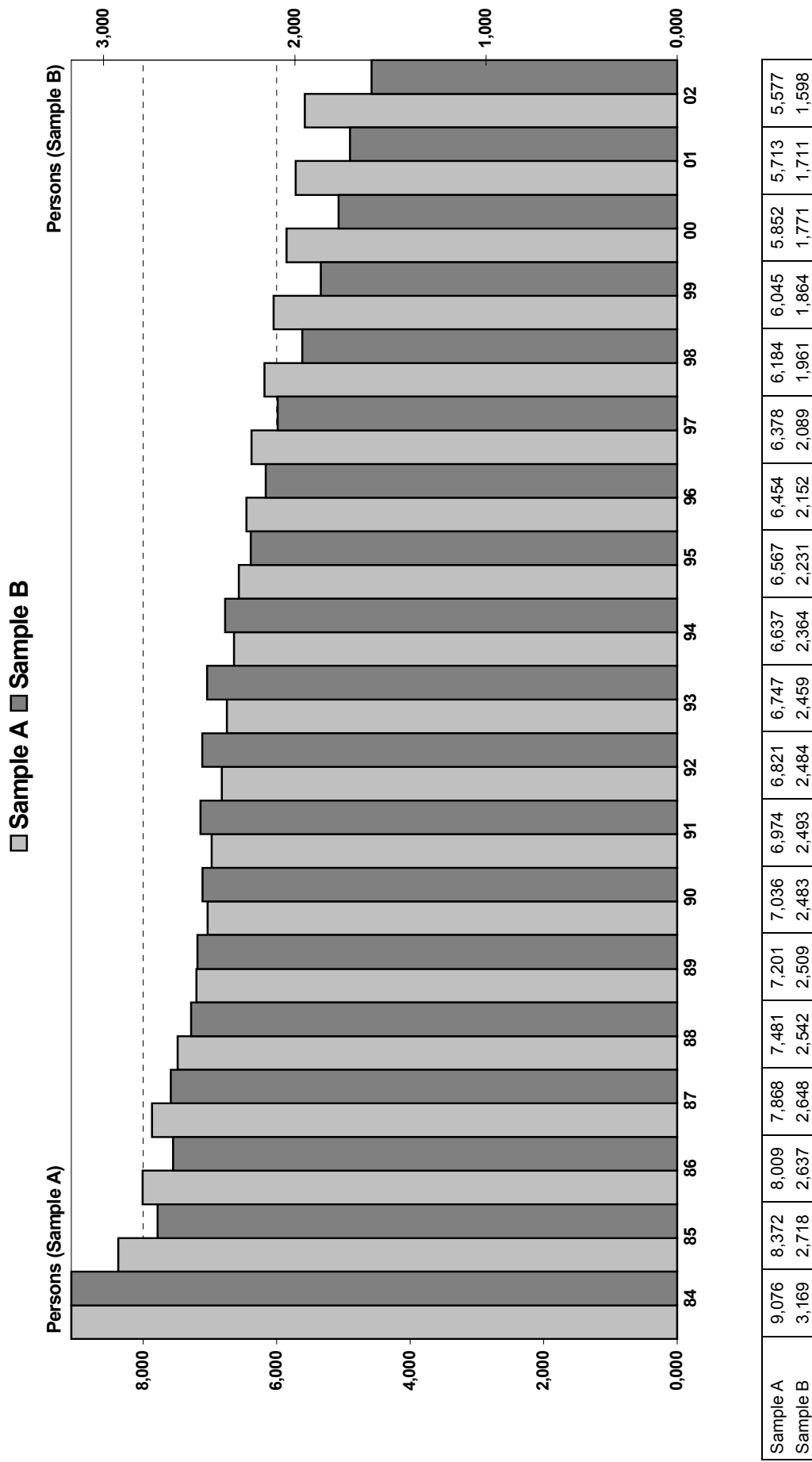


Figure 3: Comparison of successful interviews with persons and households (subsample C), waves 1 to 13.

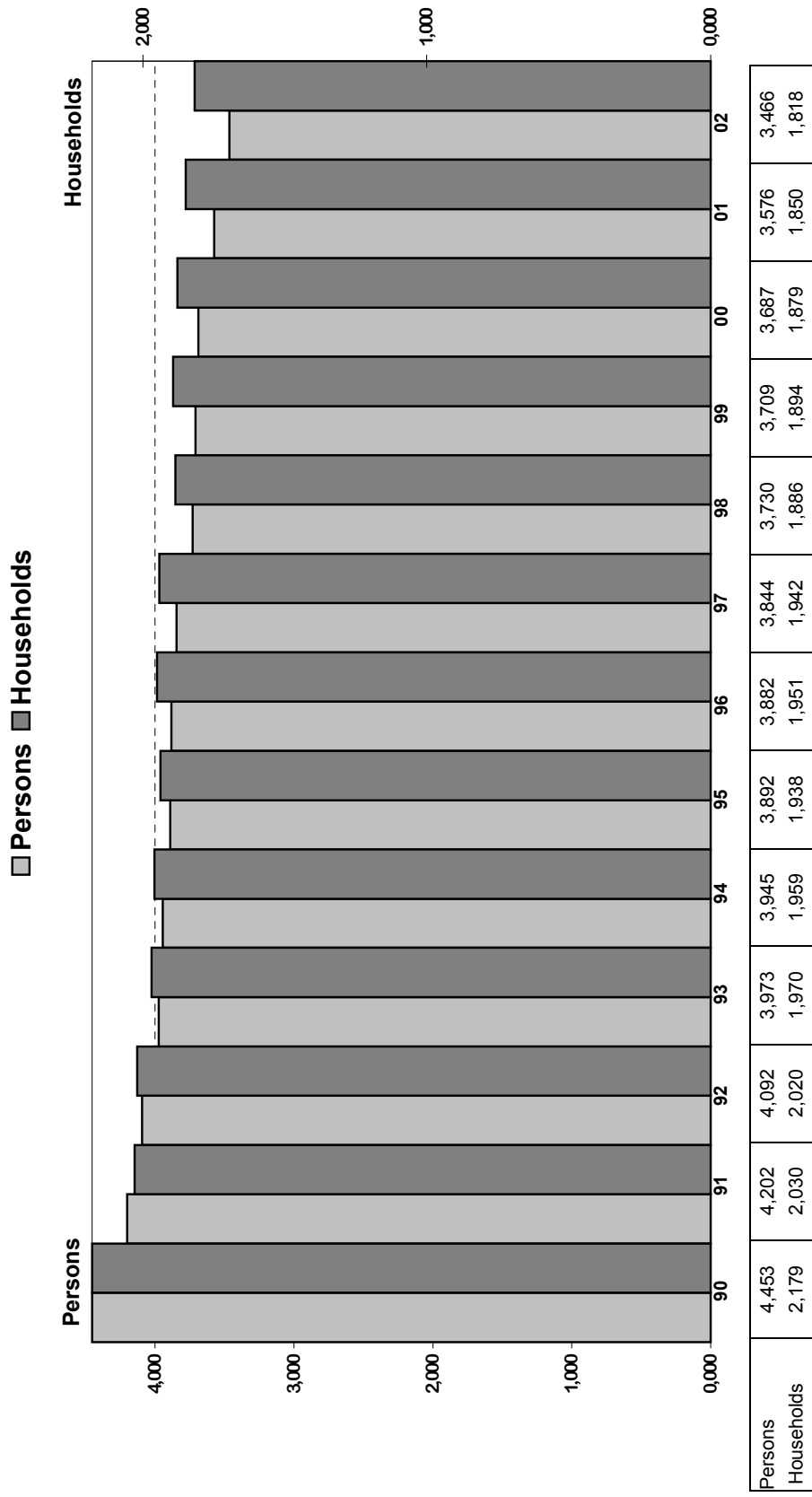


Figure 4: Comparison of successful interviews between subsamples A and B vs. subsample C (individuals), waves 1 to 13.

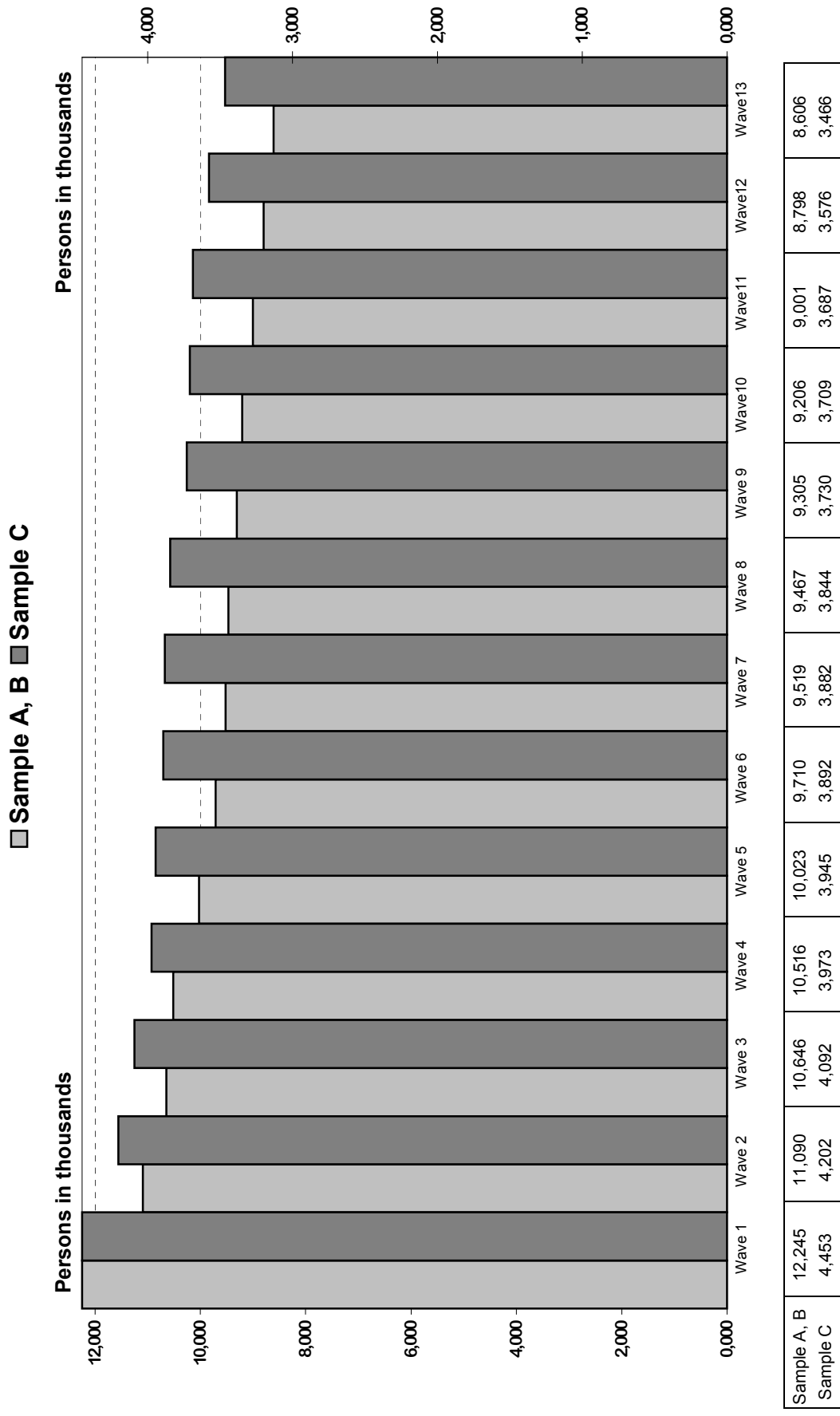
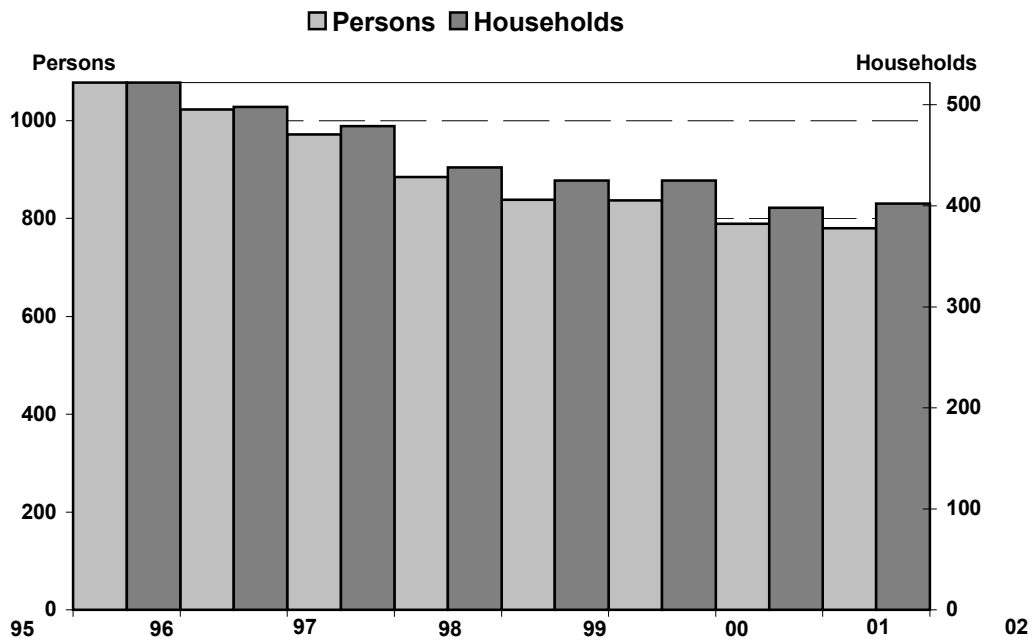
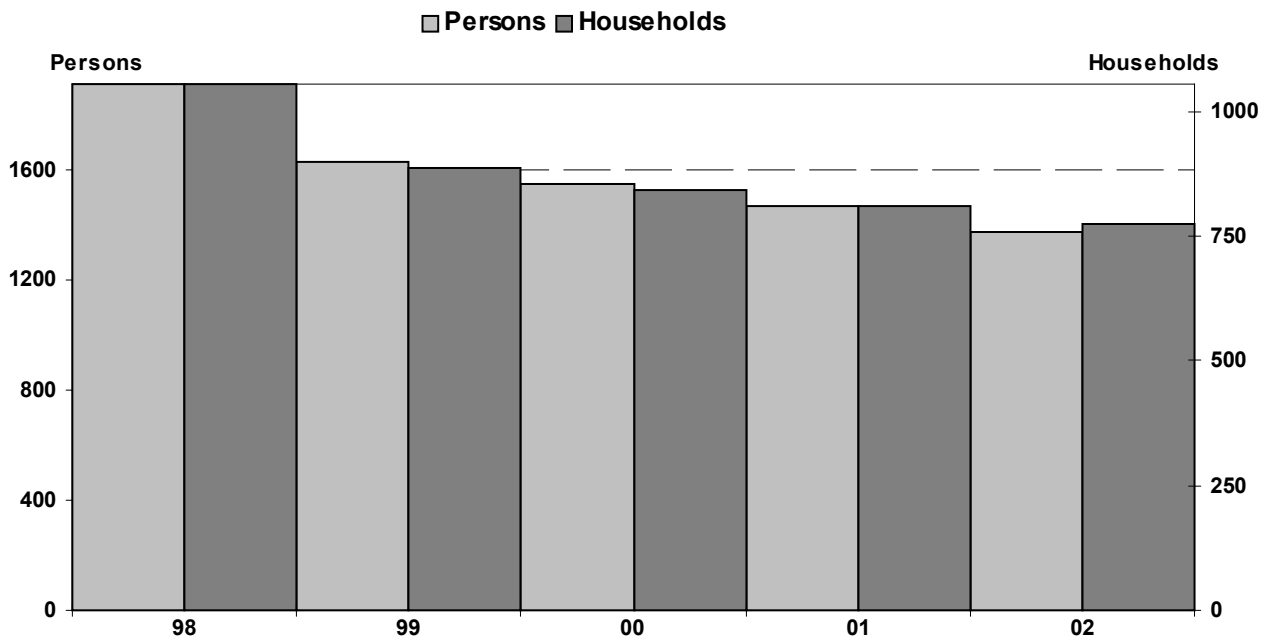


Figure 5: Comparison of successful interviews with individuals and households (subsample D), waves 1 to 8.



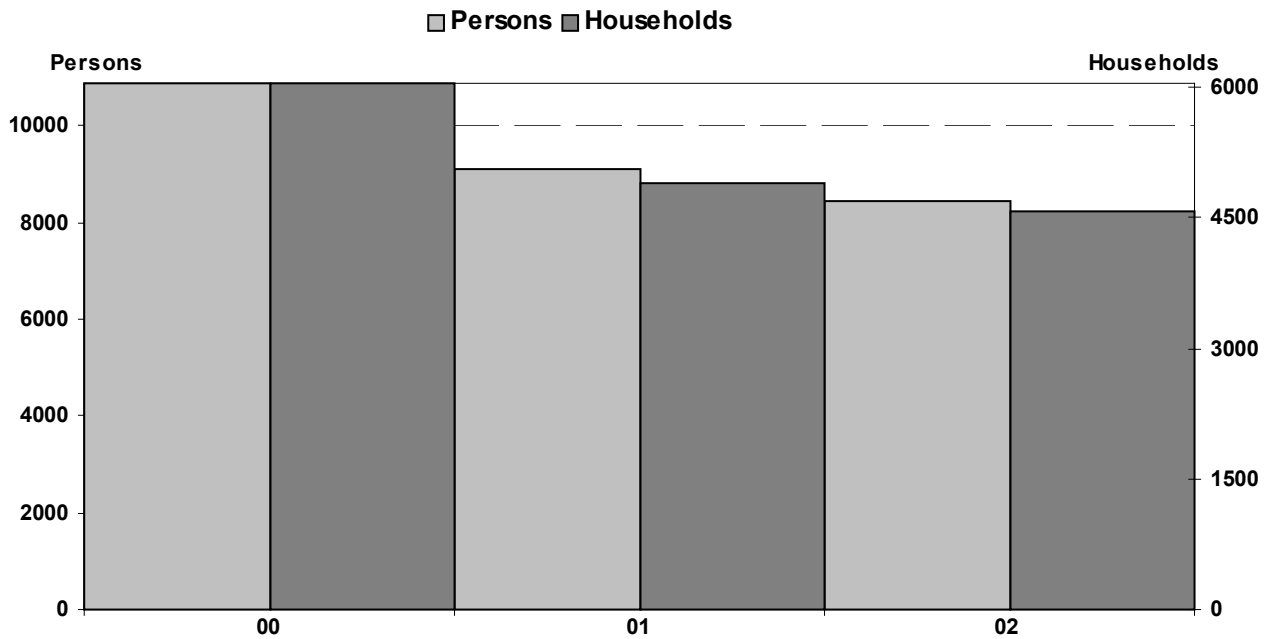
Persons	1078	1023	972	885	838	837	789	780
Households	522	498	479	441	425	425	398	402

Figure 6: Comparison of successful interviews with individuals and households (subsample E), waves 1 to 5.



Persons	1910	1629	1549	1464	1373
Households	1056	886	842	811	773

Figure 7: **Comparison of successful interviews with individuals and households (subsample F), waves 1 to 3.**



Persons	10890	9098	8427
Households	6052	4911	4586

Due to the individual regional mobility the power of the initial subsample indicator to predict the actual sampling region vanishes in course of time.

Table 1a displays the actual sampling region of the GSOEP households since 1990 for subsample A, B and C.

Table 1b shows the same information for the immigrant sample D since 1995.

Table 1c displays current sample regions for subsample E since 1998.

Table 1d displays current sample regions for subsample F in 2000.

Table 1a: **Development of sample sizes (sample A, B, C) by sampling region and institutional status 1990 to 2002.** n = Number of successful interviews, N = Estimated population total in thousands. Population margins for the number of households and individuals living in private households by sampling region are taken from the German microcensus. Because of the different definitorial concepts the figures for the institutional population are not comparable to the microcensus.

Survey year		Sampling region							
		West				East			
		Sample A+B		Sample C		Sample C		Sample A+B	
		1*	2*	1*	2*	1*	2*	1*	2*
Households									
1990	n	4592	48	-	-	2158	21	-	-
	N	28176	399	-	-	6769	90	-	-
1991	n	4620	49	22	-	1988	20	-	-
	N	28475	395	110	-	6672	109	-	-
1992	n	4598	46	58	3	1946	13	1	-
	N	28743	388	300	19	6655	78	2	-
1993	n	4609	53	78	5	1878	9	5	-
	N	29085	442	411	30	6678	56	56	-
1994	n	4545	47	93	5	1850	11	8	-
	N	29420	444	487	23	6667	83	121	-
1995	n	4451	45	111	3	1814	10	12	-
	N	28339	449	558	10	6620	83	165	-
1996	n	4383	48	118	3	1820	10	14	-
	N	28562	546	680	8	6641	79	150	-
1997	n	4316	54	128	3	1797	14	19	-
	N	28631	582	727	8	6606	117	219	-
1998	n	4212	51	125	3	1742	16	22	-
	N	24058	592	610	8	5562	136	213	-
1999	n	4111	49	139	5	1735	15	23	-
	N	24420	590	722	17	5548	111	217	-
2000	n	3986	51	146	7	1710	16	23	-
	N	13424	298	437	12	3063	82	140	-
2001	n	3906	46	161	6	1666	17	25	-
	N	13447	318	468	12	3080	100	162	-
2002	n	3820	40	175	4	1624	15	28	1
	N	12516	227	371	8	2881	81	128	5
Persons (including children)									
1990	n	12151	59	-	-	6014	30	-	-
	N	62365	445	-	-	16313	120	-	-
1991	n	12100	61	44	-	5613	26	-	-
	N	62988	439	221	-	15807	129	-	-
1992	n	11884	58	133	3	5331	18	2	-
	N	63400	439	601	11	15620	92	3	-
1993	n	11726	63	182	5	5078	11	7	-
	N	63938	466	837	25	15477	62	68	-
1994	n	11468	55	225	5	4938	13	11	-
	N	64046	454	1082	16	15296	89	164	-
1995	n	11194	54	277	3	4769	12	23	-
	N	60339	478	1269	10	15055	91	304	-
1996	n	10952	55	291	3	4670	12	29	-
	N	60583	591	1457	8	14965	87	302	-
1997	n	10742	61	311	3	4526	21	32	-
	N	60714	623	1563	8	14881	137	351	-

Table 1a: continued

Survey year		Sampling region							
		West				East			
		Sample A+B		Sample C		Sample C		Sample A+B	
		1*	2*	1*	2*	1*	2*	1*	2*
		Persons (including children)							
1998	n	10315	63	291	3	4349	24	41	-
	N	50699	778	1284	8	12417	162	389	-
1999	n	10027	60	321	5	4244	23	42	-
	N	51352	752	1526	17	12238	138	357	-
2000	n	9639	64	339	7	4143	24	49	-
	N	28158	395	919	12	6626	92	274	-
2001	n	9461	59	358	6	3976	26	56	-
	N	28311	413	954	12	6510	116	300	-
2002	n	9110	46	355	4	3789	24	61	1
	N	28374	346	985	14	6476	116	290	14
		1*: Private households							
		2*: Institutionalized population							

Table 1b: **Development of sample sizes by sampling region and institutional status 1995 to 2002 for Sample D.** n = Number of successful interviews with weighting factor greater than zero (**hrf* > 0). N = estimated population total in thousands.

Survey year		Sampling region							
		West				East			
		Standard Weights		D-specific Weights		Standard Weights		D-specific Weights	
		1*	2*	1*	2*	1*	2*	1*	2*
Households									
1995	n	307	13	362	14	2	-	2	-
	N	1247	88	1875	96	9	-	9	-
1996	n	291	7	347	8	4	-	4	-
	N	1230	54	1931	63	18	-	22	-
1997	n	278	4	327	4	4	-	5	-
	N	1251	32	1890	27	22	-	32	-
1998	n	253	4	295	4	2	-	3	-
	N	1017	42	1874	33	11	-	28	-
1999	n	246	4	282	4	2	-	4	-
	N	1048	22	1927	27	10	-	36	-
2000	n	242	4	278	4	3	-	7	-
	N	596	13	1986	29	8	-	59	-
2001	n	227	4	263	4	3	-	6	-
	N	572	13	1991	30	10	-	58	-
2002	n	237	4	273	4	3	-	8	-
	N	620	7	2240	27	10	-	73	-
Persons (including children)									
1995	n	977	30	1139	32	6	-	6	-
	N	3794	194	5773	211	26	-	27	-
1996	n	909	12	1068	14	9	-	9	-
	N	3665	96	5724	114	37	-	49	-
1997	n	857	11	1006	11	6	-	9	-
	N	3675	91	5632	82	31	-	53	-
1998	n	759	9	884	9	4	-	7	-
	N	2940	98	5380	80	18	-	65	-
1999	n	715	11	826	11	4	-	9	-
	N	2917	72	5397	86	23	-	87	-
2000	n	688	11	791	11	6	-	15	-
	N	1629	43	5385	93	18	-	131	-
2001	n	639	11	738	11	6	-	13	-
	N	1559	45	5337	96	23	-	133	-
2002	n	636	8	735	8	6	-	17	-
	N	1631	10	5785	66	22	-	166	-
1*: Private households									
2*: Institutionalized population									

Table 1c: **Development of sample sizes by sampling region and institutional status from 1998 to 2002 for Sample E.**

n = Number of successful interviews, N = Estimated population total in thousands.

Survey Year		Sampling region			
		West		East	
		1*	2*	1*	2*
Households					
1998	n	861	1	194	-
	N	4951	7	1110	-
1999	n	712	4	170	-
	N	4632	52	1196	-
2000	n	673	6	162	1
	N	2618	46	682	7
2001	n	650	8	151	2
	N	2728	58	684	14
2002	n	619	7	146	1
	N	2473	54	620	7
Persons (including children)					
1998	n	1956	3	417	-
	N	11008	20	2355	-
1999	n	1657	6	372	-
	N	10287	71	2509	-
2000	n	1548	8	353	2
	N	5762	57	1367	14
2001	n	1468	11	331	2
	N	5763	81	1386	14
2002	n	1362	9	309	1
	N	5786	77	1367	14
1*: Private households					
2*: Institutionalized population					

Table 1d: **Sample sizes by sampling region and institutional status for Sample F from 2000 to 2002.**

n = Number of successful interviews, N = Estimated population total in thousands.

Survey Year		Sampling region			
		West		East	
		1*	2*	1*	2*
Households					
2000	n	4829	45	1176	2
	N	13970	346	3185	15
2001	n	3881	29	997	4
	N	14085	276	3220	25
2002	n	3607	31	943	5
	N	12354	238	2941	28
Persons (including children)					
2000	n	11223	52	2599	9
	N	29837	399	6779	69

Table 1d: continued

Survey Year		Sampling region			
		West		East	
		1	2	1	2
Persons (including children)					
2001	n	9265	37	2202	6
	N	29935	338	6725	33
2002	n	8462	40	2029	7
	N	29935	384	6671	41
1*: Private households					
2*: Institutionalized population					

Considering the estimated population for sample A and B since 1995 (West) at the household and the personal level, we have to take into account that beginning with wave 12 (1995), the A and B weights are reduced to reflect the fact that immigrants are contained now in sample D (see Rendtel/Pannenberg/Daschke 1997 for details). Moreover since 1998 the estimates for samples A, B, C and D are reduced due to the incorporation of sample E (see Spiess/Rendtel 2000 for details). Since 2000 the estimates for samples A, B, C, D and E are reduced due to the incorporation of sample F (see Spiess 2001).

1.2 Longitudinal development of losses due to panel attrition

The following figures display the development of the number of losses due to panel attrition:

- Figure 8:** All first wave persons of subsamples A and B. Whereabout until wave 19.
- Figure 9:** All first wave persons of subsample A. Whereabout until wave 19.
- Figure 10:** All first wave persons of subsample B. Whereabout until wave 19.
- Figure 11:** All first wave persons of subsample C. Whereabout until wave 13.
- Figure 12:** All first wave persons of subsample D. Whereabout until wave 8.
- Figure 13:** All first wave persons of subsample E. Whereabout until wave 5.
- Figure 14:** All first wave persons of subsample F. Whereabout until wave 3.
- Figure 15:** All first wave persons (A, B, C). Comparison of the development until wave 13.
- Figure 16:** All first wave persons (A, B, C, D). Comparison of the development until wave 8.
- Figure 17:** All first wave persons (A, B, C, D, E). Comparison of the development until wave 5.
- Figure 18:** All first wave persons (A, B, C, D, E, F). Comparison of the development until Wave 3.
- Figure 19:** Entrants by birth or move-in and their participation behavior (subsamples A, B).

The figures in the center display the percentage of records that are without survey related attrition until the corresponding wave. These percentages may be taken as an indicator for panel stability.

Figure 8: **All first wave persons (subsample A+B). Development until wave 19.**

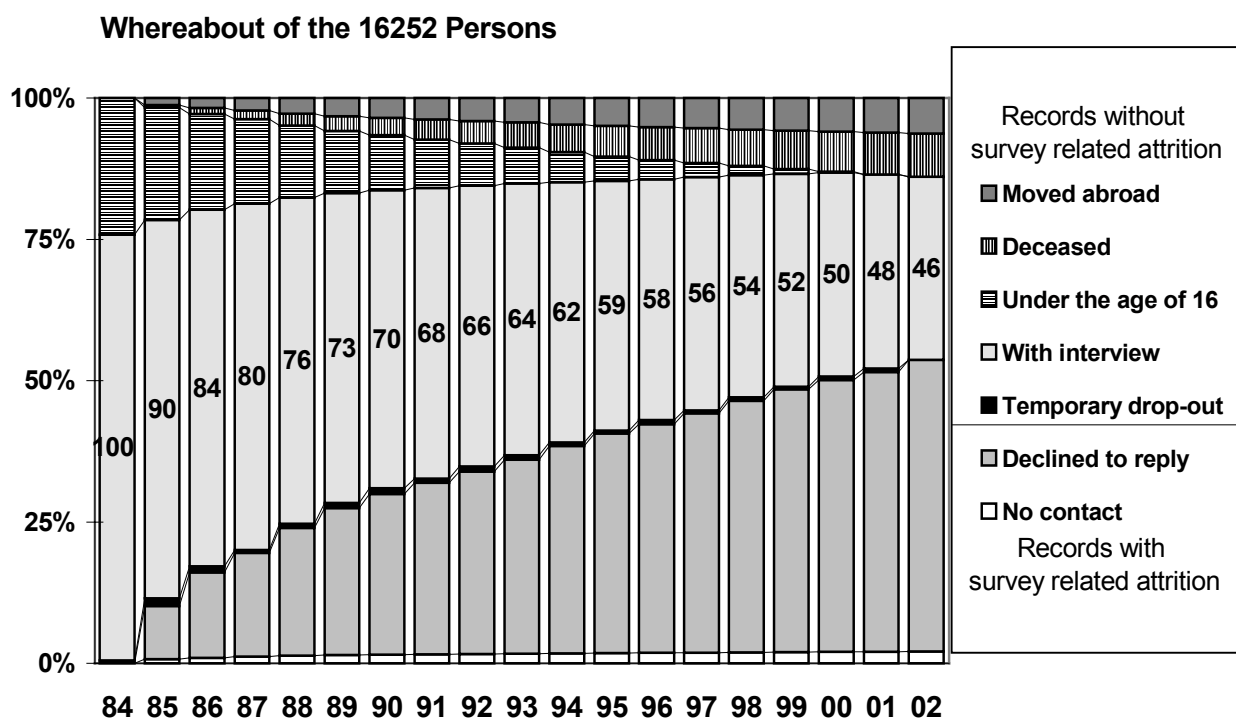


Figure 9: All first wave persons (subsample A). Development until wave 19.

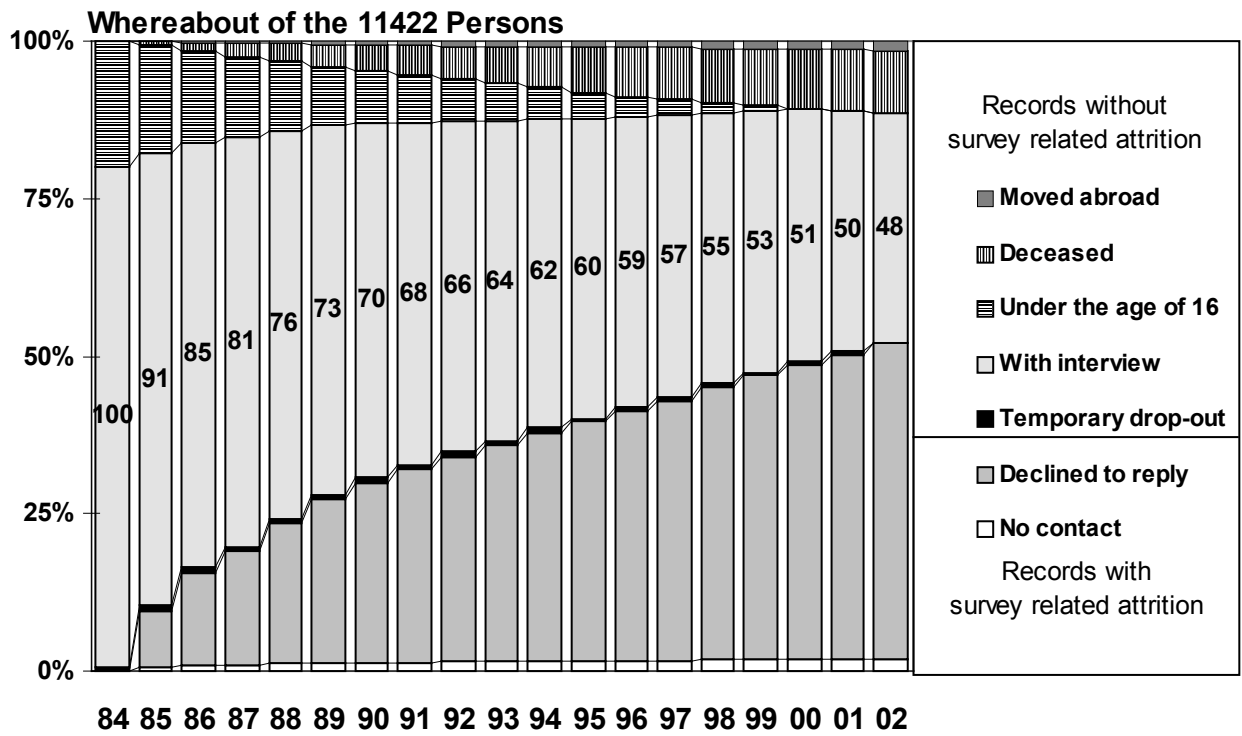


Figure 10: All first wave persons (subsample B). Development until wave 19.

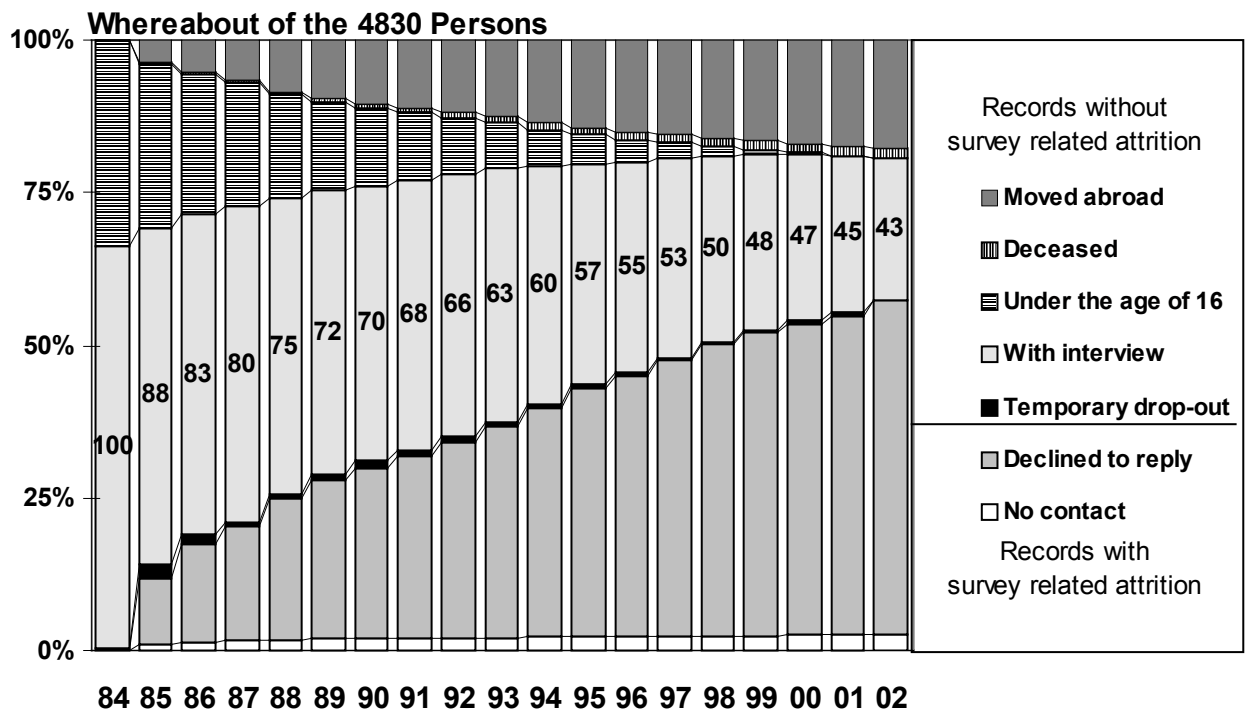


Figure 11: All first wave persons (subsample C). Development until wave 13.

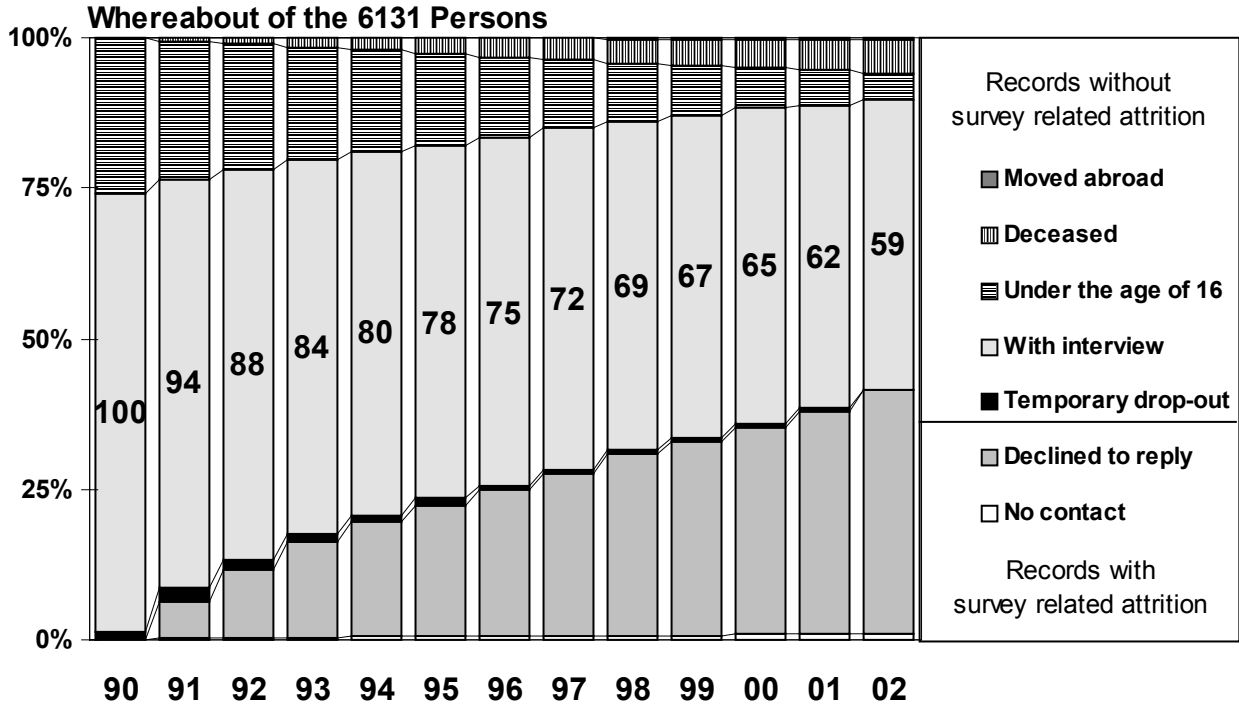


Figure 12: All first wave persons (subsample D). Development until wave 8.

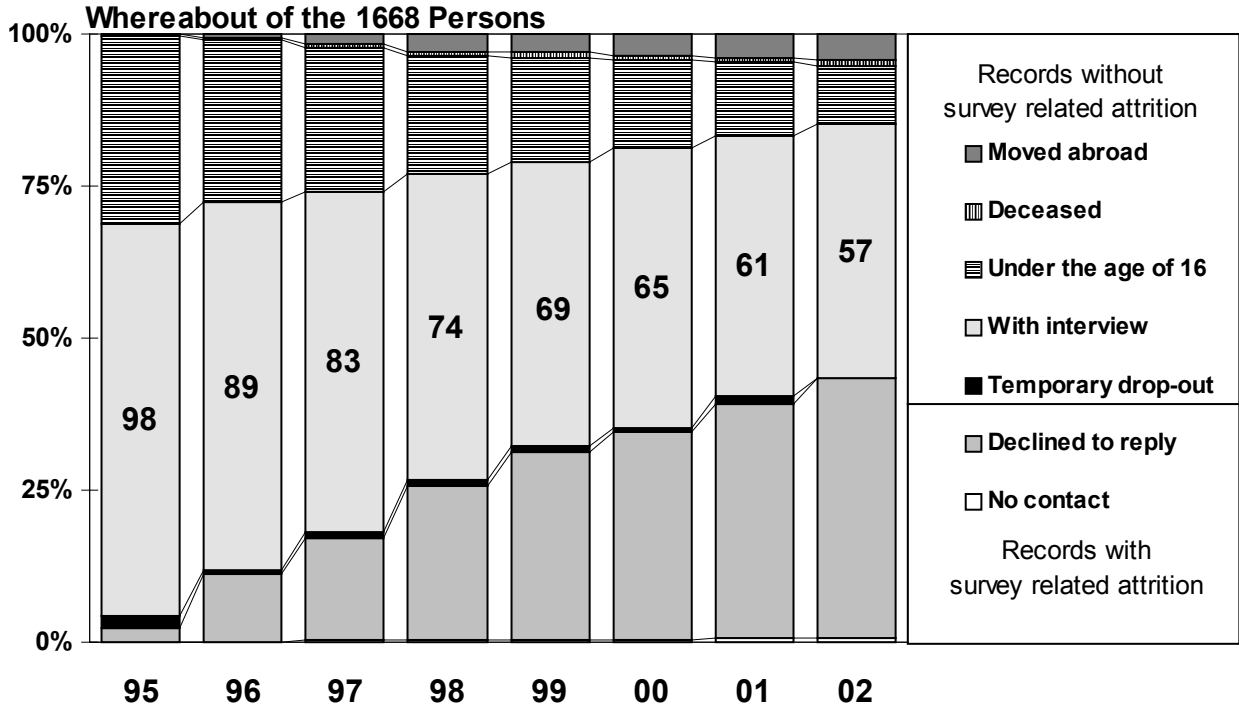


Figure 13: All first wave persons (subsample E). Development until wave 5.

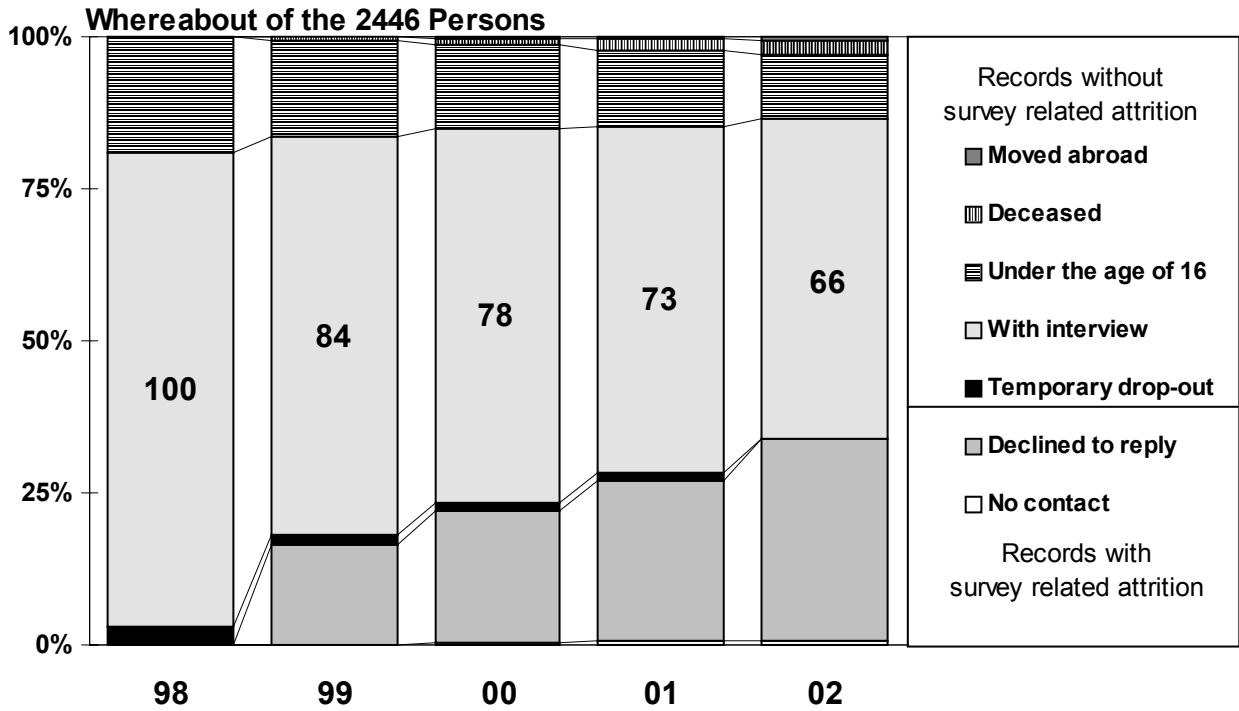


Figure 14: All first wave persons (subsample F). Development until wave 3.



Figure 15: All first wave persons (A, B, C). Comparison of the development until wave 13.

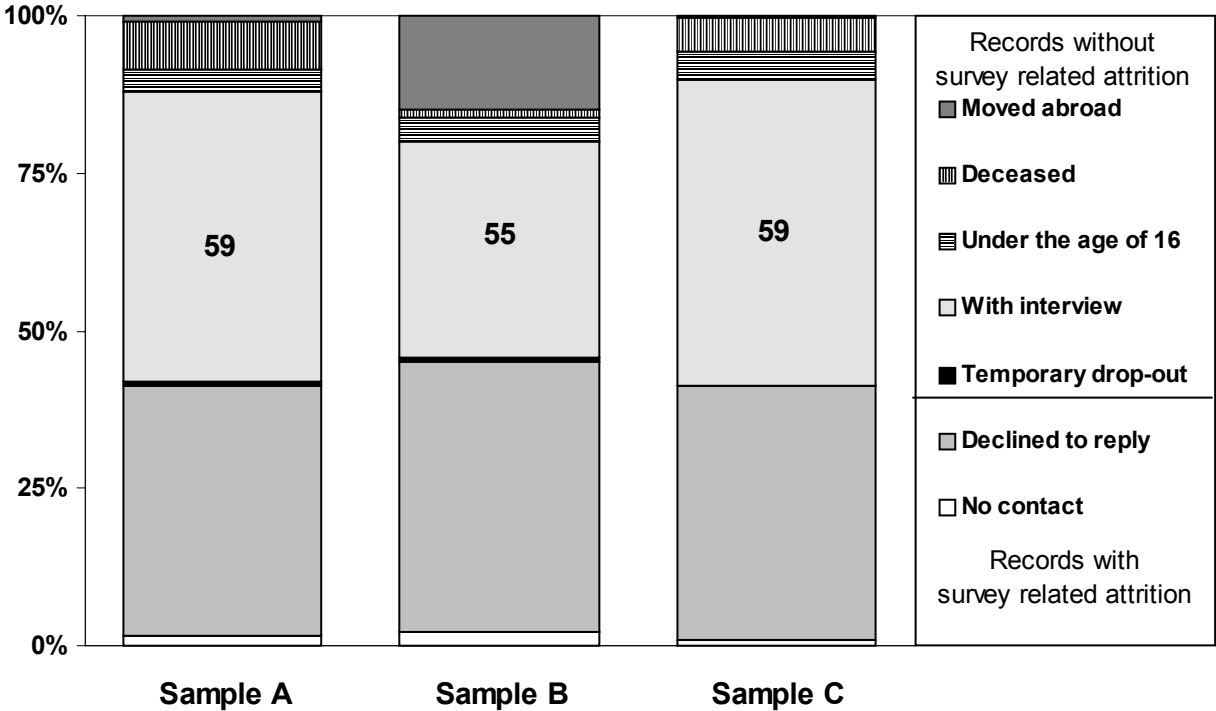


Figure 16: All first wave persons (A, B, C, D). Comparison of the development until wave 8.

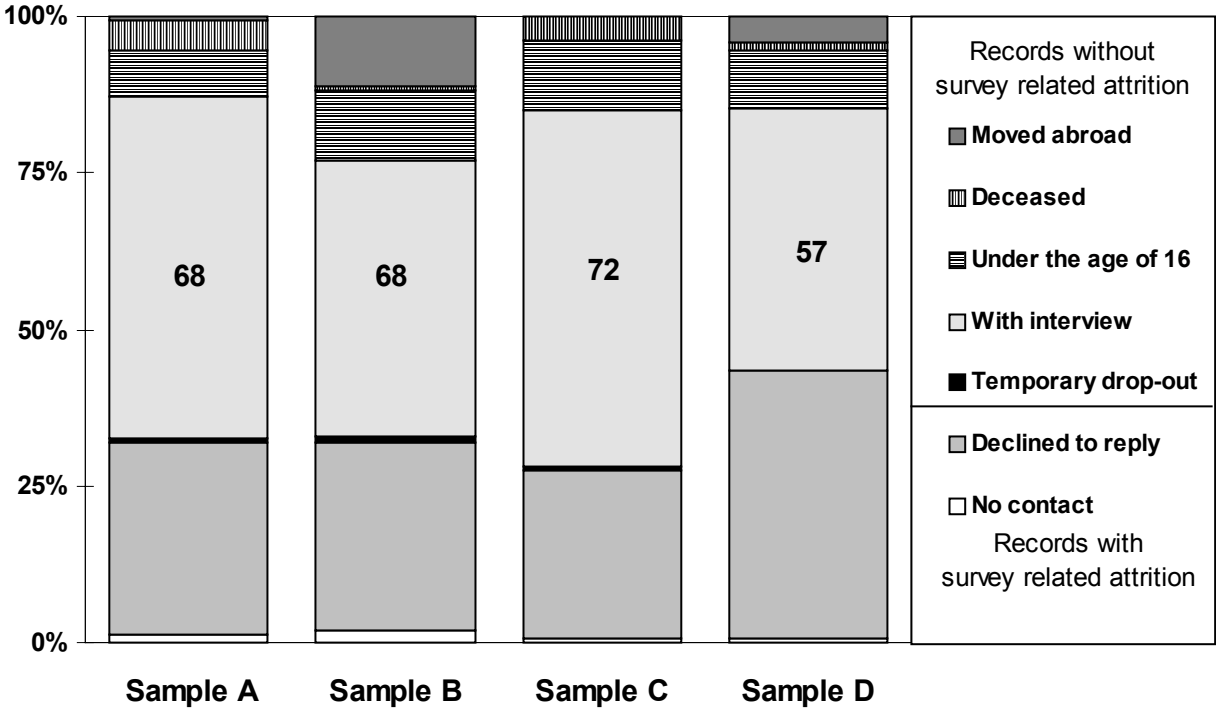


Figure 17: All first wave persons (A, B, C, D, E). Comparison of the development until wave 5.

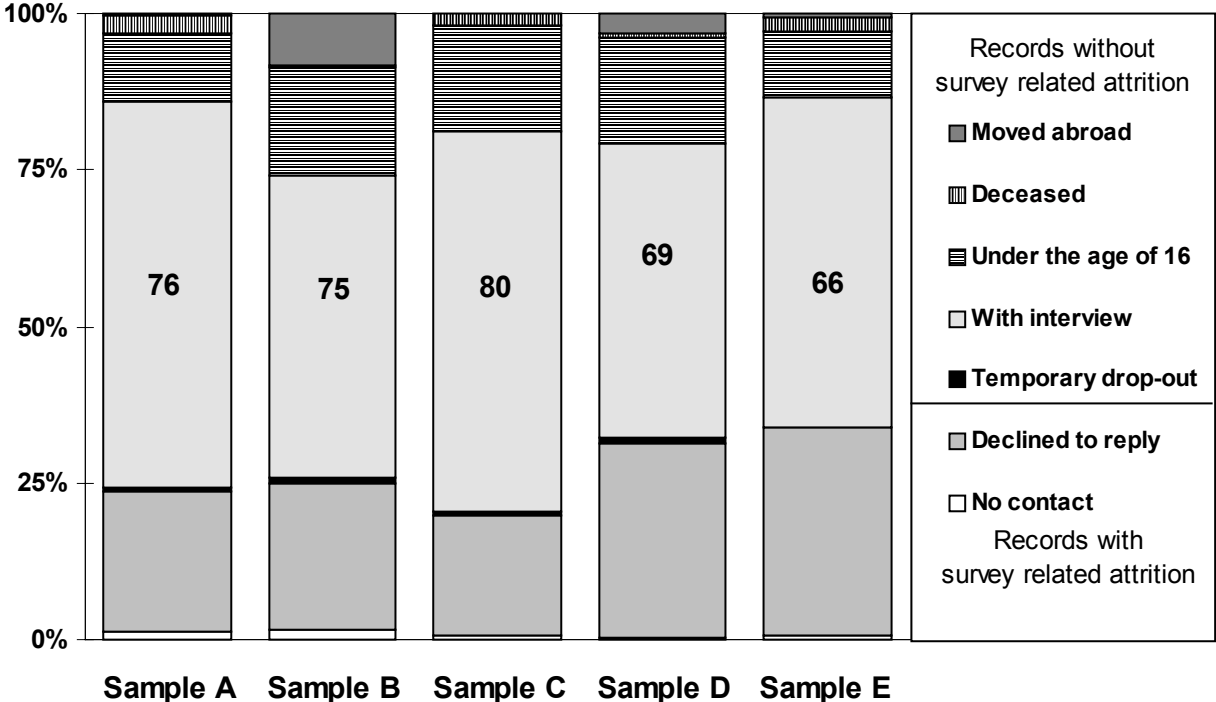
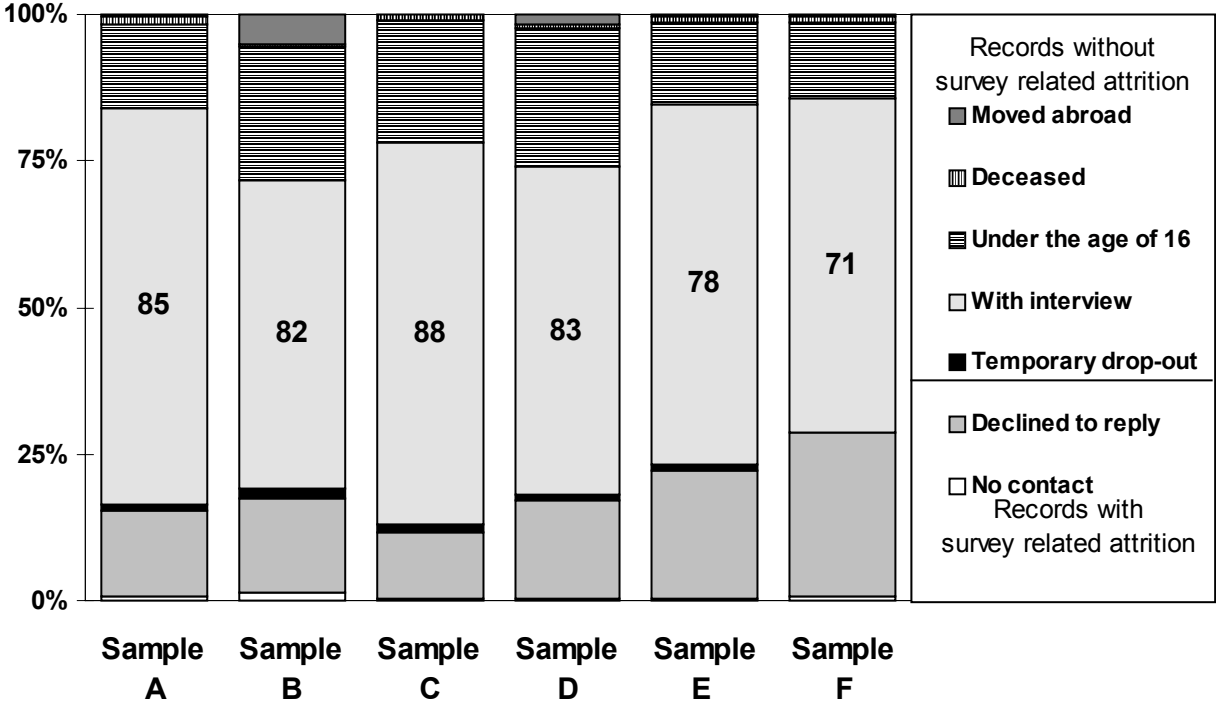
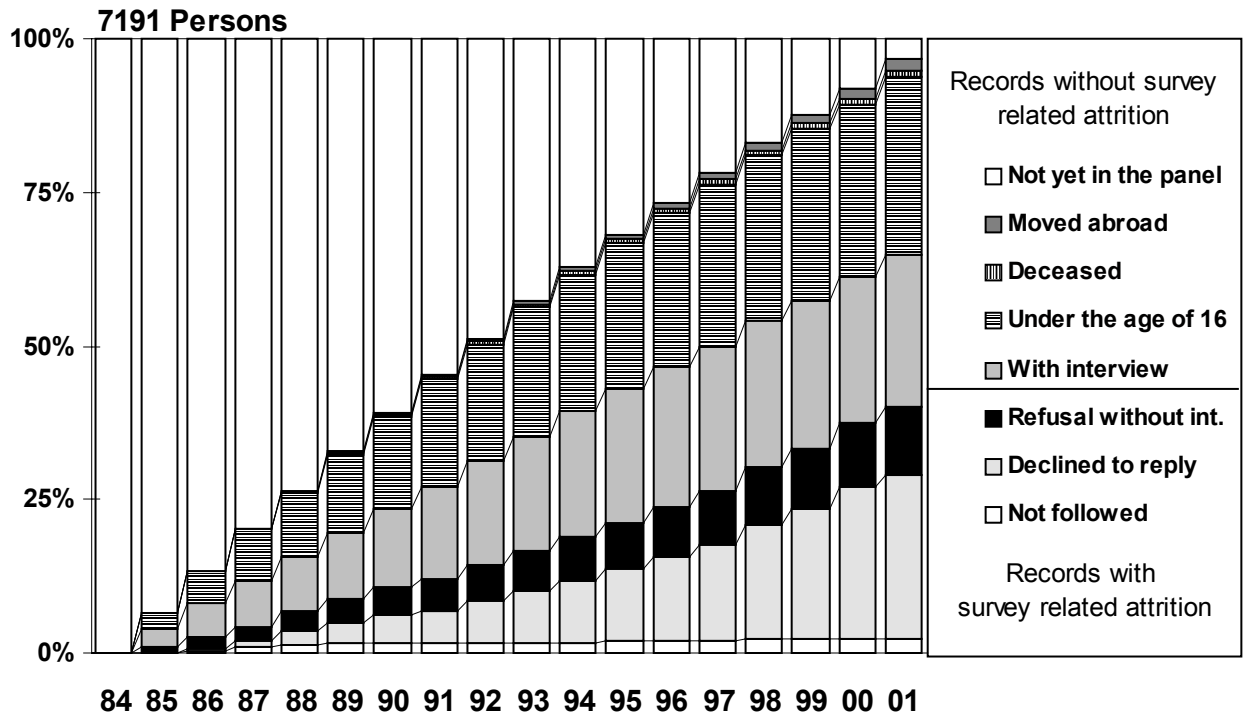


Figure 18: All first wave persons (A, B, C, D, E,F). Comparison of the development until wave 3.



1.3 Entrants by birth or move-ins and their participation behavior

Figure 19: Entrants by birth or move-in and their participation behavior (subsamples A, B).



2 Losses due to unsuccessful follow-up

In each panel wave it is necessary to re-contact the households of the preceeding wave. Therefore we have to check wether:

- the household still lives at the old address,
- the entire household has moved,
- all household members deceased,
- all household members left the sampling area,
- all household members returned into an existing panel household.

2.1 Drop-out rates by mobility behavior

Table 2 to 6 display the success of the field work with respect to recontacting of households for Sample A, B, C, D,E and F. The drop-out rates refer to all households of the previous wave that still exist in the sampling area plus split-off households. A contact is regarded to be successfully established if the interviewer recorded an interview or a refusal in the address protocol. Moreover, if the household members returned into an existing panel household, this is classified as a successful follow-up.

Table 2: **Drop-out rates due to unsuccessful follow-up in the GSOEP subsamples A and B.**

N = Number of households to be recontacted; % = percentage of households without contact.

Wave	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
	Total																		
N	6051	5814	5465	5342	5156	5044	5029	5006	5049	5008	4900	4817	4733	4695	4616	4495	4371	4290	
%	1.9	1.4	1.0	0.9	0.9	0.9	0.5	0.4	0.9	0.8	0.6	0.4	0.5	0.6	0.5	0.4	0.5	0.4	
	Households without move																		
N	5413	5039	4808	4683	4545	4472	4448	4447	4395	4359	4292	4178	4153	4022	3965	3927	3807	3749	
%	0.8	0.4	0.1	0.1	0.2	0.0	0.04	0.0	0.02	0.1	0.07	0.02	0.05	0.0	0.05	0.00	0.01	0.0	
	Moved multi-person households																		
N	298	307	272	274	228	186	197	195	231	239	264	301	249	281	265	236	242	240	
%	7.4	3.6	4.0	5.5	0.5	1.6	0.5	0.5	0.9	0.0	1.9	1.7	0.8	1.1	1.1	2.5	1.24	1.67	
	Moved single-person households																		
N	119	180	142	143	126	122	94	90	105	146	127	120	121	157	159	117	143	121	
%	21.0	14.4	7.7	5.6	4.7	5.7	1.1	0.0	7.6	6.2	0.8	0.0	0.8	3.2	0.6	1.7	3.5	0.8	
	Split-off households																		
N	221	295	242	242	246	263	290	273	317	264	217	218	210	235	227	215	179	180	
%	11.7	8.4	10.4	7.4	11.8	12.9	7.6	7.3	10.7	9.9	9.2	6.9	8.6	8.5	6.6	5.1	5.6	7.2	

Table 3: **Drop-out rates due to unsuccessful follow-up in the GSOEP subsample C.**

N = Number of households to be recontacted;

% = percentage of households without contact.

Wave	2	3	4	5	6	7	8	9	10	11	12	13
	Total											
N	2246	2304	2227	2136	2113	2104	2091	2081	2041	2028	2036	2010
%	1.5	0.5	0.9	0.6	0.4	0.5	0.5	0.6	0.3	0.4	0.3	0.5
	Households without move											
N	2062	2043	2021	1904	1862	1796	1771	1732	1750	2028	1740	1702
%	0.0	0.05	0.05	0.0	0.1	0.0	0.1	0.0	0.06	0.4	0.0	0.06
	Moved multi-person households											
N	81	106	82	92	119	142	153	175	132	161	132	133
%	11.1	0.0	3.7	2.2	0.0	1.4	0.0	0.6	0.0	0.6	1.5	0.8
	Moved single-person households											
N	21	43	14	39	30	45	60	64	56	63	52	65
%	14.3	9.3	0.0	2.6	3.3	4.4	1.7	1.6	0.0	0.0	0.0	0.0
	Split-off households											
N	82	112	110	104	102	121	107	110	103	107	112	110
%	25.6	6.3	13.6	8.6	6.9	5.8	8.4	10.0	5.8	5.6	3.6	6.4

Table 4: **Drop-out rates due to unsuccessful follow-up in the GSOEP subsample D.**

N = Number of households to be recontacted;

% = percentage of households without contact.

Characteristic	Wave 2		Wave 3		Wave 4		Wave 5		Wave 6		Wave 7		Wave 8	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Total	544	0.4	542	0.7	498	0.6	529	0.9	467	0.2	454	0.9	450	0.2
Households without move	431	0.0	424	0.0	394	0.0	409	0.0	396	0.0	381	0.0	370	0.0
Moved multi- person households	74	0.0	65	0.0	60	1.7	65	3.1	41	0.0	43	2.3	38	0.0
Moved single- person households	16	6.3	16	6.3	15	6.7	18	5.6	7	0.0	11	27.3	13	0.0
Split-off households	23	4.4	37	8.1	29	3.5	37	5.4	23	4.4	19	0.0	29	3.5

Table 5: **Drop-out rates due to unsuccessful follow-up in the GSOEP subsample E.**

N = Number of households to be recontacted;

% = percentage of households without contact.

Characteristic	Wave 2		Wave 3		Wave 4		Wave 5	
	N	%	N	%	N	%	N	%
Total	1100	0.5	968	0.8	922	0.87	875	0.57
Households without move	996	0.0	869	0.1	814	0.1	775	0.0
Moved multi- person households	36	0.0	40	7.5	33	3.0	41	0.0
Moved single- person households	32	3.1	19	0.0	25	4.0	25	8.0
Split-off households	36	11.1	40	10.0	50	10.0	34	8.8

Table 6: **Drop-out rates due to unsuccessful follow-up in the GSOEP subsample F.**

N = Number of households to be recontacted;

% = percentage of households without contact.

Characteristic	Wave 2		Wave 3	
	N	%	N	%
Total	6172	1.0	5451	0.5
Households without move	5557	0.0	4915	0.0
Moved multi- person households	275	7.6	208	0.0
Moved single- person households	176	10.8	127	3.2
Split-off households	164	13.4	201	10.5

2.2 Definition of the regressors for a Logit analysis

The estimation of the probability that a household is lost by unsuccessful follow-up is done by means of a Logit model with the following characteristics:

Characteristic	Abbreviation	Code	Values
Moved	MOVE	1	household, not moved
		2	Moved multi-person household
		3	Moved single-person household
		4	Split-off household
Large City	LARGE	0	Else
		1	More than 100 thousand inhabitants
Household size	SIZE	1	Single-person household
		2	2 person household
		3	3 person household
		4	4 or more persons household
Single-person Household	SINGLE	0	Else
		1	Single-person household
Type of house	TYP	1	Single house or rural area
		2	Multi storey house
		3	Else
Split-off household	SPLIT	1	Moved multi-person household
		2	Moved single-person household
		3	Split-off household

2.3 Estimated coefficients of the Logit model

The regressors defined in the previous section were employed in a Logit analysis. The model estimates the probability $P_c = (\text{contact} = \text{no})$. For the computation of the GSOEP weighting schemes only model specifications with all regressors being significant were used. The specification is:

$$\ln \frac{P_{c,i}}{1 - P_{c,i}} = \text{const} + X'_{i}\beta$$

Thus, positive estimated parameters indicate an increased drop-out rate compared to the sample average.

Table 6 uses a simple symbolic notation for the models and their estimated parameters. Here „+,, means the addition of a main effect, an „*,, indicates an interaction term. Variable 1 (Variable 2 = c) symbolizes a conditional main effect which is linked to cases where variable 2 = c. The estimated coefficients are displayed under the model equation. The notation uses the convention: variable (value 1: coefficient 1/value 2: coefficient 1/...).

The estimated drop out rates due to unsuccessful follow-up may be easily calculated from table 6. For example: In wave 2, subsample A, we find for a multiple-person household, that moved (MOVE=2) from a large city (LARGE=1) the logit value $-2.87+0.24+ 0.11=-2.52$. Thus we get

$$\text{Pr}(\text{contact} = \text{no}) = \frac{e^{-2.52}}{1 + e^{-2.52}} = 0.074.$$

Table 7: **The estimates of a Logit model for the probability of a drop-out due to unsuccessful follow-up in the GSOEP.**
Description of coefficients: variable (value 1: coefficient 1/value 2: coefficient 1/...).

Subsample A (West-Germans)	
Wave	Model and coefficients
2	Model = CONST + LARGE + MOVE CONST (-2.87), LARGE (0: -0.24/1: 0.24) MOVE (1: -2.52 / 2: 0.11 / 3: 1.53 / 4: 0.84)
3	Model = CONST + LARGE + MOVE CONST (-3.62), LARGE (0: -0.36 / 1: 0.36), MOVE (1: -1.79 / 2: -0.49 / 3: 1.48 / 4: 0.80)
4	Model = CONST + MOVE CONST (-3.42), MOVE (1: -3.01 / 2: 0.78 / 3: 0.98 / 4: 1.25)
5	Model = CONST + MOVE + SINGLE (MOVE) CONST (-3.76), MOVE (1: -3.09 / 2,3: 1.34 / 4: 1.75) SINGLE (MOVE = 1) (0: -1.35 / 1: 1.35) SINGLE (MOVE = 2,3) 0: -0.28 / 1: 0.28) SINGLE (MOVE = 4) (0: -0.63 / 1: 0.63)
6	Model = CONST + MOVE + SINGLE (MOVE) CONST (-3.48), MOVE (1: -2.33 / 2,3: 0.64 / 4: 1.69) SINGLE (MOVE = 1) (0: -0.75 / 1: 0.75) SINGLE (MOVE =2,3) (0: -0.76 / 1: 0.76) SINGLE (MOVE= 4) (0: -0.26 / 1: 0.26)
7*	Model = CONST + LARGE + SPLIT CONST (-2.97), LARGE (0: -0.39 / 1: 0.39), SPLIT (1: -1.10 / 2: -0.07 / 3: 1.17)
8	Model = CONST + MOVE CONST (-5.03) MOVE 1: -2.79 / 2: -0.24 / 3: 0.50 / 4: 2.53)
9	Pr (contact = no) = 0 if MOVE = 1,2,3 / =0.06 if MOVE =4
10	Model = CONST + LARGE + MOVE CONST (-4.44), LARGE (0: -0.44 / 1: 0.44), MOVE (1: -3.65 / 2: 0.10 / 3: 1.12 / 4: 2.42)
11	Model = CONST + SINGLE + MOVE CONST (-6.01), SINGLE (0: -1.06 / 1: 1.06) MOVE (1: -0.99 / 2: -5.13 / 3: 1.84 / 4: 4.28)

Table 7: continued (1)

12	Model = CONST + SINGLE + MOVE CONST (-4.61), SINGLE (0: -0.72 / 1: 0.72) MOVE (1: -2.68 / 2: 0.78 / 3: -0.83 / 4: 2.73)
13	Model = CONST + MOVE CONST (-6.89) MOVE (1: -1.21 / 2: 2.30 / 3: -5.31 / 4: 4.22)
14	Model = CONST + MOVE + SINGLE CONST (-6.95) SINGLE (0: -0.73 / 1: 0.73) MOVE (1: -9.09 / 2: 2.56 / 3: 1.62 / 4: 4.91)
15	Model = CONST + MOVE + SINGLE CONST (-3.97) MOVE (1,2,3: -2.15 / 4: 2.15) SINGLE (0: -0.76 / 1: 0.76)
16	Model = CONST + MOVE CONST (-4.82) MOVE (1,2,3: -2.23 / 4: 2.23)
17	Model = CONST + MOVE CONST (-4.64) MOVE (1,2,3: -1.70 / 4: 1.70)
18	Model = CONST + MOVE CONST (-4.44) MOVE (1,2,3: -1.73 / 4: 1.73)
19	Model = CONST + MOVE CONST (-5.1) MOVE (1,2,3: -2.29 / 4: 2.29)
* In wave 7 all households that did not move were successfully re-contacted. The drop-out analysis is therefore based only on households with an observed move.	
Subsample B (Foreigners)	
2	Model = CONST + LARGE + MOVE + SIZE CONST (-2.28), LARGE (0: -0.50 / 1: 0.50), MOVE (1: -1.66 / 2: 0.69 / 3: -0.07 / 4: 1.04) SIZE (1: 1.23 / 2: 0.26 / 3: -0.82 / 4: -0.67)

Table 7: continued (2)

3	Model = CONST + LARGE + MOVE CONST (-2.65), LARGE (0: -0.72 / 1: 0.72), MOVE (1: -3.06 / 2: 0.16 / 3: 1.64 / 4: 1.26)
4	CONST (-3.34), MOVE (1: -3.60 / 2: -0.46 / 3: 2.19 / 4: 1.87)
5	like Subsample A
6	like Subsample A
7*	Model = CONST + LARGE + SPLIT + TYPE CONST (-2.93), LARGE (0: 0.64 / 1: -0.64), SPLIT (1: -1.65 / 2: 0.58 / 3: 1.07), TYPE (1: -0.73 / 2: 1.32 / 3: -0.59)
8	Like Subsample A
9	Pr (contact = no) = 0 if MOVE = 1,2,3 / = 0.10 if MOVE = 4
10	Model = CONST + LARGE + MOVE CONST (-7.98), LARGE (0: -0.81 / 1: 0.81), MOVE (1: -7.63 / 2: -4.69 / 3: 6.50 / 4: 5.82)
11	Model = CONST + SINGLE + MOVE CONST (-5.39), SINGLE (0: -1.5 / 1: 1.54), MOVE (1: -1.19 / 2: -4.26 / 3: 2.07 / 4: 3.39)
12	Model = CONST + MOVE CONST (-5.34), MOVE (1: -1.52 / 2: 2.21 / 3: -3.86 / 4: 3.17)
13	Model = CONST + MOVE CONST (-8.32), MOVE (1: -7.08 / 2: 4.83 / 3: -3.61 / 4: 5.86)
14	Model = CONST + MOVE CONST (-5.69), MOVE (1: -0.40 / 2: 1.31 / 3: -4.51 / 4: 3.60)
15	Model = CONST + MOVE CONST (-4.72), MOVE (1,2,3: -2.14 / 4: 2.14)
16	Model = CONST + SINGLE + MOVE CONST (-3.90) SINGLE (0: -0.93 / 1: 0.93) MOVE (1,2,3: -1.47 / 4: 1.47)
17	Model = CONST + MOVE CONST (-4.47) MOVE (1,2,3: -1.62 / 4: 1.62)
18	Model = CONST + MOVE CONST (-4.42) MOVE (1,2,3: -1.2 / 4: 1.2)

Table 7: continued (3)

19	Model = CONST + MOVE CONST (-3.77) MOVE (1,2,3: -1.86 / 4: 1.86)
* In wave 7 all households that did not move were successfully re-contacted. The drop-out analysis is therefore based only on households with an observed move.	
Subsample C (East-Germans)	
Wave	Model and coefficients
2	Pr(contact=no) = MOVE (1: 0.0 / 2: 0.11 / 3: 0.14 / 4: 0.25)
3	Pr(contact=no) = MOVE (1,2: 0.0 / 3: 0.09 / 4: 0.07)
4	Pr(contact=no) = MOVE (1: 0.0 / 2: 0.04 / 3: 0.0 / 4: 0.14)
5	Pr(contact=no) = MOVE (1: 0.0 / 2: 0.02 / 3: 0.03 / 4: 0.09)
6	Pr(contact=no) = MOVE (1: 0.0 / 2: 0.0 / 3: 0.03 / 4: 0.07)
7	Pr(contact=no) = MOVE (1: 0.0 / 2: 0.01 / 3: 0.04 / 4: 0.06)
8	Pr(contact=no) = MOVE (1: 0.0 / 2: 0.0 / 3: 0.02 / 4: 0.08)
9	Model = CONST + MOVE + SIZE CONST (-4.80) MOVE (1,2,3: -2.55 / 4: 2.55) SIZE (1,2: -0,96 / 3,4: 0.96)
10	Model = CONST + MOVE + SINGLE CONST (-4.80) MOVE (1,2,3: -2.61 / 4: 2.61) SINGLE (0: -1.00 / 1: 1.00)
11	Model = CONST + MOVE CONST (-5.19) MOVE (1,2,3: -2.37 / 4: 2.37)
12	Model = CONST + MOVE CONST (-5.08) MOVE (1,2,3: -1.79 / 4: 1.79)
13	Model = CONST + MOVE CONST (-4.77) MOVE (1,2,3: -2.08 / 4: 2.08)
Subsample D (Immigrants)	
Wave	Model and coefficients*
2	Pr(contact=no) = MOVE (1: 0.0 / 2: 0.0 / 3: 0.07 / 4: 0.05)
3	Pr(contact=no) = MOVE (1: 0.0 / 2: 0.0 / 3: 0.08 / 4: 0.08)
4	Pr(contact=no) = MOVE (1: 0.0 / 2: 0.04 / 3: 0.08 / 4: 0.04)

Table 7: continued (4)

5	Model = CONST + MOVE CONST (-4.24) MOVE (1,2,3: -1.46 / 4: 1.46)
6	Pr(contact=no)=0 / all households successfully followed-up.
7	Model = CONST + MOVE CONST (-8.43) MOVE (1,2,3: 2.78 / 4: -2.78)
* excluding households with *hhrfd ≤ 0.	
Subsample E (Refreshment)	
Wave	Model and coefficients
2	Model = CONST + MOVE CONST (-4.52) MOVE (1,2,3: -2.44 / 4: 2.44)
3	Model = CONST + MOVE CONST (-3.96) MOVE (1,2: -2.80 / 3: 1.04 / 4: 1.76)
4	Model = CONST + MOVE + LARGE CONST (-4.00) MOVE (1,2,3: -1.80 / 4: 1.80) LARGE (0: -0.99 / 1: 0.99)
5	Model = CONST + MOVE CONST (-4.19) MOVE (1,2,3: -1.85 / 4: 1.85)
Subsample F (Innovation)	
Wave	Model and coefficients
2*	Model = CONST + SPLIT CONST (-2.16) SPLIT (1: -0.34 / 2: 0.05 / 3: 0.29)
3	Model = CONST + MOVE + SINGLE CONST (-4.30) MOVE (1,2,3: -2.35 / 4: 2.35) SINGLE (0: 0.52 / 1: -0.52)
* In wave 2 all households that did not move were successfully re-contacted. The drop-out analysis is therefore based only on households with an observed move.	

3 Losses due to refusals

3.1 Drop-out rates by different household characteristics

Table 8 displays the drop-out rates due to refusal by different household characteristics. In general the characteristics are drawn from the previous wave. However, the survey related characteristics refer to the actual sampling wave.

The individual attributes refer to the head of the household in the previous wave. However, for split-off households the attributes refer to the person that moved from the panel household (in case of several persons that moved from a panel household: the person first mentioned in the address protocol).

No differences were made between various reasons for the refusal like explicit denial or refusal because of lack of time, bad health conditions, etc..

Considering the type of interview (CAPI vs. PAPI) we do not observe any significant impact of the type of interview itself or of interactions of the interview type with other household characteristics in our subsequent logit estimates for all subsamples.

Table 8: **Comparison of drop-out rates between subsamples A/B, C, D, E and F. Current wave.**

N = Number of households to be recontacted;

% = percentage of households without interview.

Characteristic		Sample				
		AB	C	D	E	F
All recontacted households	N	4265	1996	286	868	5420
	%	8.82	8.92	6.64	10.94	15.39
Temporary dropouts previous wave	N	187	73	-	40	400
	%	57.22	63.01	-	65.00	69.00
		Households with participation in previous wave				
		Gender				
Male	N	2527	928	188	525	3322
	%	6.17	6.25	8.51	9.14	11.23
Female	N	1551	995	98	303	1698
	%	7.29	7.44	3.06	6.93	10.90
		Type of household				
Old household not moved	N	3602	1655	238	742	4575
	%	5.44	5.08	5.04	7.41	10.05
Old household moved	N	334	185	32	62	313
	%	8.38	10.81	15.63	11.29	16.61

Table 8: continued

Characteristic		Sample				
		AB	C	D	E	F
Split-off household	N	142	83	16	24	132
	%	31.69	33.73	12.50	29.17	34.85
		Interviewer has changed				
Yes	N	405	125	28	86	650
	%	25.43	29.60	39.29	39.53	44.92
No	N	3319	1528	247	711	4370
	%	4.07	3.93	2.83	4.22	6.09
Special cases	N	354	270	11	31	-
	%	8.76	12.96	9.09	16.13	-
		CAPI-Interview				
Yes	N	1006	273	102	516	1908
	%	4.87	8.42	5.88	5.04	10.80
No	N	3072	1650	184	312	3112
	%	7.16	6.61	7.07	13.78	11.31
		Telephone number available for Infratest				
No	N	279	184	14	49	343
	%	20.79	22.28	14.29	18.37	20.70
Yes	N	3799	1739	272	779	4677
	%	5.55	5.23	6.25	7.70	10.41
		E-mail-address available				
Yes	N	596	224	36	124	727
	%	3.52	3.57	5.56	4.84	5.23
No	N	3482	1699	250	704	4293
	%	7.12	7.30	6.80	8.95	12.11

3.2 Definition of the regressors for a Logit analysis

The characteristics used in the descriptive statistics in the preceding section were employed in a Logit analysis of the refusal rate. However, we use only model specifications where all included regressors are significantly different from zero. The definition of the regressors is given in the list below:

Characteristic	Abbreviation	Code	Values
Age of the head Of household	ALTHV	1	Older than 75 years
		2	65-74 years
		3	55-64 years
		4	35-54 years
		5	25-34 years
		6	Younger than 25 years
Gender of the head	SEX	0	Male
		1	Female
Typ of the household	HTYP	1	Old household without move
		2	Old household moved
		3	Split-off household
Change of Interviewer	INTW	0	No change
		1	Change since previous wave
		2	Not regular interviewer number
Number of interviews	BETREUUNG		Number of interviews with the interviewer of the present wave
Starting from the Beginning	BEGINN	0	Else
		1	Heads participation since wave 1
Person moving out	AUSZUG	0	Else
		1	A respondent left the household since the previous wave
Separation of a couple	PAAR	0	Else
		1	The head or the spouse (cohabitor) of the previous wave left the household
Interaction of household Type and separation of The couple	TYP	0	HTyp = 1,2 and Paar = 1
		1	HTyp = 1 and Paar = 0
		2	HTyp = 2 and Paar = 0
		3	HTyp = 3 and Paar = 0
		4	HTyp = 3 and Paar = 1
East-Berlin	OSTB	0	Else
		1	household is located in East-Berlin
Marital status	FAMSTD	1	Married living together
		2	Married living separately
		3	Single
		4	Divorced
		5	Widowed

List: continued (1)

Characteristic	Abbreviation	Code	Values
Household income East, quantiles	INCE**	0	No
		1	Yes
Jobless	ALOS	0	Else
		1	Head is jobless
Loss of job (subjective notion)	VERLUST	0	Else
		1	Loss expected or probable
Occupational status Of the head	STATUSH	0	Else
		1	High status
Social aid	SOZH	0	Else
		1	Household is recipient of social aid payments
Household income West-Germany	EINKW	1	Income not reported
		2	≤ 2000 DM
		3	2000 - 4000 DM
		4	≥ 4000 DM
Household income (East-Germany)	EINKO	1	Income not reported
		2	≤ 800 DM
		3	800-1200 DM
		4	1200-1800 DM
		5	1800-2500 DM
		6	≥ 2500 DM
Income quartiles	EINKQU	0	Income not reported
		1	≤ 25 %
		2	50 %
		3	75 %
		4	≥ 75 %
Household income Not reported	KAEINK	0	Else
		1	Income not reported
Balance of assets Not reported	KAVB	0	Else
		1	Balance not reported in wave 5
Number of different Kinds of assets in the Households	ANZASSET	0	No assets reported
		1	Number = 0
		2	Number = 5 (Maximum)
		3	Else
Firm assets	BETRIEB	0	Else
		1	Household owns firm assets
Savings reported As one kind of assets	SPAR	0	No
		1	Yes
Household migrated From East to West Germany	OSTWEST	0	No
		1	Yes
Member of D-Subsamples	MIGRANT	1	Subsample D1
		2	Subsample D2
Telephone	TELEPHON	0	No
		1	Yes

List: continued (2)

Characteristic	Abbreviation	Code	Values
Subtenant	UNTMIE TE	0	No
		1	Yes
Apprenticeship	APPRENT	0	No
		1	Yes
University	UNI	0	No
		1	Yes
Change of interview Type (t-2 => t-1)	INTWTYPE	0	No
		1	Yes
Satisfaction with life	NSAT	0	No
		1	Yes
Housing subsidy	WOHNGELD	0	No
		1	Yes
Type of interview	INTWART	0	Written form
		1	Verbal form
Household size	SIZE	1	Single-person household
		2	2 person household
		3	3 person household
		4	4 or more persons household
Telephone number available for Infratest	TELINFRA	0	No
		1	Yes
E-mail-address available for Infratest	EMAIL	0	No
		1	Yes
Head of household: state of health	HANDICAP	0	No
		1	Yes
Head of household: nationality	FOREIGNHV	0	No
		1	Yes
Sample Region	REGIONF	0	West Germany
		1	East Germany
Person moving in	MOVE_IN	0	Else
		1	A respondent moved into the new household since previous wave
Head of household: Health state	SICK	0	Else
		1	Strongly limits everyday life
Computer assisted interview	CAPI	0	Else
		1	Computer assisted personal interview
Residential area	RESID	0	Else
		1	Rural
Change of interview type (PAPI (t-2) => CAPI (t-1))	PAPCAP	0	Else
		1	PAPI => CAPI

3.3 Estimated coefficients of the Logit model

The regressors defined above were used in a multiple Logit analysis. The model estimates the probability $P_R = P(\text{Response} = \text{no})$. For the computation of the GSOEP weighting schemes only model specifications with all regressors being significant were used. The specification is:

$$\ln \frac{P_{R,i}}{1 - P_{R,i}} = \text{const} + X'_i \beta$$

Thus, positive estimated parameters indicate an increased drop-out rate compared to the sample average.

Table 12 uses a simple symbolic notation for models and their estimated parameters. Here „+„ means the addition of a main effect, an „*„ indicates an interaction term. Variable 1 (Variable 2 = c) symbolizes a conditional main effect which is linked to cases where variable 2 = c. The estimated coefficients are displayed under the model equation. The notation uses the convention: variable (value 1: coefficient 1/value 2: coefficient 1...).

The estimated drop-out rates due to refusals may be easily calculated from the estimated parameters displayed in table 12. For example: In wave 2, subsample A, we find for a household with no change of the interviewer (INTW = 0) and age of the head between 35 and 74 years (ALTHV = 2,3,4) and the reported household income below 2000 DM (EINKW = 2), which did not move (HTYP = 1) the logit value $-1.53 - 0.25 + 0.03 - 0.68 + 0.12 = -2.31$.

$$\text{Thus we have } \Pr(\text{Response}=\text{no}) = \frac{e^{-2.31}}{1 + e^{-2.31}} = 0.09.$$

Table 9: **The estimates of a Logit model for the probability of a drop-out due to refusal in the GSOEP. Representation of coefficients: variable (value 1: coefficient 1/value 2: coefficient 1/...).**

Subsample A (West-Germans)	
Wave	Model and coefficients
2	Model = CONST + INTW + ALTHV + HTYP + EINKW CONST (-1.53), INTW (0: -0.25 / 1: 0.25), ALTHV (1: 0.66 / 2,3,4: 0.03 / 5: -0.39 / 6: -0.30), HTYP (1: -0.68 / 2: -0.19 / 3: 0.87), EINKW (1: 0.61 / 2: 0.12 / 3: -0.35 / 4: -0.38)
3	Model = CONST + INTW + ALTHV + INTW * ALTHV + HTYP + ALOS + KAEINK CONST (-1.22), INTW (0: -0.39 / 1: 0.39), ALTHV * (INTW=0) (1: -0.13 / 2: -0.11 / 3,4: -0.39 / 5: 0.26 / 6: 0.37), ALTHV * (INTW=1) (1: 0.13 / 2: 0.11 / 3,4: 0.39 / 5: -0.26 / 6: -0.37), ALTHV (1: 0.59 / 2: 0.16 / 3,4: -0.06 / 5: -0.53 / 6: -0.16) HTYP (1: -0.52 / 2: 0.10 / 3: 0.42), ALOS (0: -0.21 / 1: 0.21), KAEINK (0: -0.39 / 1: 0.39)
4	Model = CONST + ALTHV + INTW (ALTHV) + HTYP + KAEINK CONST (-1.83), INTW (ALTHV = 1) (0: -0.44 / 1: 0.44), INTW (ALTHV =2) (0: -0.74 / 1: 0.74), INTW (ALTHV =3,4) (0: -0.59 / 1: 0.59), INTW (ALTHV =5) (0: -0.41 / 1: 0.41), INTW (ALTHV =6) (0: -0.32 / 1: 0.32), ALTHV (1: 0.21 / 2: -0.38 / 3,4: -0.24 / 5: 0.06 / 6: 0.35), HTYP (1: -0.46 / 2: 0.28 / 3: 0.18), KAEINK (0: -0.39 / 1: 0.39)
5	Model = CONST + BETREUUNG + ALTHV (INTW =1) + HTYP + KAEINK + ANZASSET CONST (-1.60), BETREUUNG (1: 1.15 / 2: 0.41 / 3: 0.18 / 4: -0.71 / 5: -1.03), ALTHV (INTW = 1) (1,2: 0.52 / 3,4,5: -0.11 / 6: -0.40), HTYP (1: -0.49 / 2: 0.11 / 3: 0.38), KAEINK (0: -0.45 / 1: 0.45), ANZASSET (0,2,3: -0.38 / 1: 0.38)
6	Model = CONST + BETREUUNG + ALTHV (INTW = 1) + HTYP + KAEINK + KAVB + BETRIEB CONST (-2.44), BETREUUNG (1: 0.75 / 2: 0.58 / 3: 0.21 / 4: -0.59 / 5: -0.43 / 6: -0.52), ALTHV (INTW = 1) (1,2: 0.26 / 3,4,5: 0.05 / 6: -0.31), HTYP (1: -0.32 / 2: -0.04 / 3: 0.37), KAEINK (0: -0.26 / 1: 0.26), BETRIEB (0: 0.41 / 1: -0.41)
7	Model = CONST + HTYP + INTW (HTYP) + KAEINK + STATUSH CONST (-1.34), INTW (HTYP = 1) (0: -0.75 / 1: 0.75), INTW (HTYP = 2) (0: -0.56 / 1: 0.56), INTW (HTYP = 3) (0: -0.12 / 1: 0.12), HTYP (1: -0.66 / 2: -0.24 / 3: 0.90), KAEINK (0: -0.58 / 1: 0.58) STATUSH (0: -0.30 / 1: 0.30)

Table 9: continued (1)

Subsample A (West-Germans)	
Wave	Model and coefficients
8	Model = CONST + INTW + HTYP + KAEINK + ANZAHL CONST (-1.15), INTW (=: -0.55 / 1: 0.55), HTYP (1: -0.83 / 2: -0.14 / 3: 0.97), KAEINK (0: -0.57 / 1: 0.57), ANZAHL (1: -0.08 / 2: 0.70 / 3: -0.62)
9	Model = CONST + INTW (BEGINN) + BEGINN (ALTHV) + HTYP + AUSZUG (HTYP=1) + KAEINK + ANZASSET + SEX CONST (-1.31), INTW (BEGINN = 0) (0: -0.17 / 1: 0.17), INTW (BEGINN = 1) (0: -0.68 / 1: 0.68), BEGINN (ALTHV = 1) (0: -0.09 / 1: 0.09), BEGINN (ALTHV = 2) (0: 0.70 / 1: -0.70), BEGINN (ALTHV = 3) (0: 1.20 / 1: -1.20), BEGINN (ALTHV = 4) (0: 0.49 / 1: -0.49), BEGINN (ALTHV = 5) (0: 0.48 / 1: -0.48), BEGINN (ALTHV = 6) (0: 0.10 / 1: -0.10), HTYP (1: -0.53 / 2: 0.07 / 3: 0.46), AUSZUG (HTYP=1) (0: -0.47 / 1: 0.47), KAEINK (0: -0.25 / 1: 0.25), ANZASSET (0,2,3: -0.29 / 1: 0.29), SEX (0: 0.15 / 1: -0.15)
10	Model = CONST + HTYP + BEGINN (HTYP) + INTW (HTYP) + PAAR (HTYP=1) + ALTHV (HTYP=1) CONST (-1.89), HTYP (1: -0.12 / 2: -0.39 / 3: 0.51), INTW (HTYP=1) (0: -0.95 / 1: 0.08 / 2: 0.88), INTW (HTYP=2) (0: -0.24 / 1: -0.06 / 2: 0.30), INTW (HTYP=3) (0: 0.16 / 1: -0.47 / 2: 0.31), BEGINN (HTYP=1) (0: 0.43 / 1: -0.43), BEGINN (HTYP=2) (0: 0.21 / 1: -0.21), BEGINN (HTYP=3) (0: -0.07 / 1: 0.07), PAAR (HTYP=1) (0: -0.58 / 1: 0.58), ALTHV (HTYP=1) (1: 0.41 / 2: -0.26 / 3: -0.08 / 4: -0.50 / 5: 0.01 / 6: 0.42)
11	Model = CONST + HYTP + BEGINN + INTW + KAEINK + TELEPHON (INTW=1) CONST (-1.68) HTYP (1: -0.39 / 2: -0.09 / 3: 0.48) BEGINN (0: 0.27 / 1: -0.27) INTW (0: -0.63 / 1: -0.10 / 2: 0.73) KAEINK (0: -0.35 / 1: 0.35) TELEPHON (INTW=1) (0: 0.49 / 1: -0.49)
12	Model = CONST + HTYP + INTW + ALTHV (HTYP = 1) CONST (-1.92) HTYP (1: -0.36 / 2: -0.52 / 3: 0.88) INTW (0: -1.10 / 1: 0.03 / 2: 1.07) ALTHV (HTYP =1) (1: 0.57 / 2,3,4,5,6: -0.57)

Table 9: continued (2)

Subsample A (West-Germans)	
Wave	Model and Coefficients
13	<p>Model = CONST + HTYP + INTW + BEGINN + ALTHV + KAEINK + PAAR (HTYP = 1)</p> <p>CONST (-1.42) HTYP (1: -0.39) / 2: -0.23 / 3: 0.62) INTW: (0: -0.75 / 1: 0.25 / 2: 0.5) BEGINN (0: 0.35 / 1: -0.35) ALTHV (1: 0.61 / 2: -0.29 / 3: 0.42 / 4: -0.20 / 5: -0.15 / 6: -0.39) KAEINK (0: -0.26 / 1: 0.26) PAAR (HTYP = 1) (0: -0.64 / 1: 0.64)</p>
14	<p>Model = CONST + HTYP + INTW + EINKQU + BETREUUNG + PAAR (HTYP = 1) + INTWTYPE + NSAT</p> <p>CONST (-0.48) HTYP (1: -0.79 / 2: -0.05 / 3: 0.84) INTW (0: -0.59 / 1: -0.11 / 2: 0.70) EINKQU (0: 0.60 / 1: -0.08 / 2,3,4: -0.52) BETREUUNG (0[=3]: 0.5 / 1[=4]: 0.5) PAAR (HTYP = 1) (0: -0.41 / 1: 0.41) INTWTYPE (0: -0.42 / 1: 0.42) NSAT (0: -0.38 / 1: 0.38)</p>
15	<p>Model = CONST + INTW + HTYP + APPRENT + INTWTYPE + INTWART + TELINFRA</p> <p>CONST (-1.80) INTW (0: -1.20 / 1: 0.03 / 2: 1.17) HTYP (1: -0.62 / 2: -0.32 / 3: 0.94) SEX (0: -0.19 / 1: 0.19) APPRENT (0: -0.21 / 1: 0.21) INTWTYPE (0: -0.30 / 1: 0.30) INTWART (0: 0.59 / 1: -0.59) TELIINFRA (0: 0.32 / 1: -0.32)</p>
16	<p>Model = CONST + INTW + HTYP + PAAR + INTWTYPE + TELINFRA</p> <p>CONST (-1.25) INTW (0: -0.78 / 1: 0.08 / 2: 0.7) HTYP (1: -0.59 / 2: -0.27 / 3: 0.86) PAAR (0: -0.30 / 1: 0.30) INTWTYPE (0: -0.31 / 1: 0.31) TELIINFRA (0: 0.33 / 1: -0.33)</p>
17	<p>Model = CONST + BEGINN + INTW + HTYP + UNTMIETE + TELINFRA</p> <p>CONST (-1.35) BEGINN (0: 0.23 / 1: -0.23) INTW (0: -0.87 / 1: 0.39 / 2: 0.48) HTYP (1: -0.63 / 2: -0.15 / 3: 0.78) UNTMIETE (0: -0.33 / 1: 0.33) TELIINFRA (0: 0.25 / 1: -0.25)</p>

Table 9: continued (3)

Subsample A (West-Germans)	
Wave	Model and Coefficients
18	Model = CONST + BEGINN + INTW + HTYP + UNTMIETE + TELINFRA + INTWTYPE + PAAR CONST (-0.06) BEGINN (0: 0.16 / 1: -0.16) INTW (0: -1.28 / 1: 0.70 / 2: 0.58) HTYP (1: -0.33 / 2: -0.32 / 3: 0.65) UNTMIETE (0: -0.51 / 1: 0.51) TELINFRA (0: 0.55 / 1: -0.55) INTWTYPE (0: -0.44 / 1: 0.44) PAAR (0: -0.58 / 1: 0.58)
19	Model = CONST + BETREUUNG + INTW + HTYP + MOVE_IN + INTWART + TELINFRA + INTWTYPE + EMAIL + SOZH + TYPCHANGE CONST (-1.69) BETREUUNG (-0.064) INTW (0: -2.66 / 1: 2.46 / 2: 0.20) HTYP (1: -0.24 / 2: 0.04 / 3: 0.20) MOVE_IN (0: 3.14 / 1: -3.14) INTWART (0: -1.09 / 1: 1.09) TELINFRA (0: 1.11 / 1: -1.11) INTWTYPE (0: -0.59 / 1: 0.59) EMAIL (0: 0.72 / 1: -0.72) SOZH (0: -0.96 / 1: 0.96) TYPCHANGE (0: 1.61 / 1: -1.61)
Subsample B (Foreigners)	
2	Model = CONST + INTW + HTYP CONST (-1.96), INTW (0: -0.55 / 1: 0.55) HTYP (1: -0.03 / 2: -0.58 / 3: 0.62)
3	Model = CONST + SEX + HTYP CONST (-1.60), SEX (0: -0.31 / 1: 0.31), HTYP (1,2: -0.46 / 3: 0.46)
4	Model = CONST + INTW (ALTHV) + HTYP + EINKW CONST (-1.69), INTW (ALTHV =1,2,3) (0: -0.47 / 1: 0.47), INTW (ALTHV =4) (0: -0.73 / 1: 0.73), INTW (ALTHV =5) (0: -0.60 / 1: 0.60), INTW (ALTHV =6) (0: -0.26 / 1: 0.26), HTYP (1: -0.34 / 2: 0.46 / 3: -0.12), EINKW (1: 0.75 / 2: 0.10 / 3: -0.85)
5	Model = CONST + BETREUUNG + HTYP + KAEINK CONST (-1.87), BETREUUNG (1: 1.26 / 2: 0.14 / 3: -0.21 / 4: -0.70 / 5: -0.50), HTYP (1: -0.47 / 2: 0.89 / 3: -0.42), KAEINK (0: -0.43 / 1: 0.43)
6	Model = CONST + BETREUUNG + HTYP + KAEINK CONST (-1.89), BETREUUNG (1: 0.83 / 2: 0.37 / 3: -0.31 / 4: -0.55 / 5: 0.04 / 6: -0.37), HTYP (1: -0.41 / 2: 0.22 / 3: 0.19), KAEINK (0: -0.54 / 1: 0.54)

Table 9: continued (4)

Subsample B (Foreigners)	
Wave	Model and Coefficients
7	Model = CONST + HTYP + INTW (HTYP) + KAEINK CONST (-1.50), INTW (HTYP=1) (0: -0.55 / 1: 0.55), INTW (HTYP=2) (0: -0.98 / 1: 0.98), INTW (HTYP=3) (0: -1.06 / 1: 1.06), HTYP (1: -0.50 / 2: -0.88 / 3: 1.38), KAEINK (0: -0.66 / 1: 0.66)
8	Model = CONST + INTW + HTYP CONST (-2.05), INTW (0: -0.48 / 1: 0.48), HTYP (1: -0.85 / 2: 0.22 / 3: 0.63)
9	Model = CONST + INTW + BEGINN + TYP + ALTHV+ KAEINK + ANZASSET + SOZH CONST (-1.79), INTW (0: -0.50 / 1: 0.50), BEGINN (0: 0.39 / 1: -0.39), TYP (0: 0.16 / 1: -0.59 / 2: -1.90 / 3: -0.03 / 4: 2.36), ALTHV 1,2,3: 0.28 / 4: -0.10 / 5: -0.65 / 6: 0.47), KAEINK (0: -0.66 / 1: 0.66), ANZASSET (0,2,3: -0.53 / 1: 0.53), SOZH (0: 0.73 / 1: -0.73)
10	Model = CONST + HTYP + PAAR + ALTHV + INTW (ALTHV) CONST (-1.58) HTYP (1: -0.44 / 2: -0.11 / 3: 0.55), PAAR (0: -0.63 / 1: 0.63), ALTHV (1,2,3: -0.79 / 4: -0.04 / 5: 0.77 / 6: 0.06), INTW (ALTHV = 4) (0: -1.11 / 1: -0.10 / 2: 1.21), INTW (ALTHV = 5) (0: -0.79 / 1: -0.22 / 2: 1.01)
11	Model = CONST + BEGINN + HTYP + INTW + ANZASSET + FAMSTD CONST (-1.43), INTW (0: -0.69 / 1: 0.01 / 2: 0.70), BEGINN (0: 0.33 / 1: -0.33), HTYP (1,2: -0.48 / 3: 0.48), ANZASSET (0,2,3: -0.31 / 1: 0.31), FAMSTD (1: 0.25 / 2,3,4,5: -0.25)
12	Model = CONST + HTYP + INTW + PAAR + ALTHV CONST (-0.88) HTYP (1: -0.97 / 2: 0.36 / 3: 0.61) INTW (0: -0.67 / 1: -0.45 / 2: 1.12) PAAR (0: -0.84 / 1: 0.84) ALTHV (1,2,3: -0.35 / 4,5,6: 0.35)
13	Model = CONST + HTYP + INTW CONST (-1.73) HTYP (1: -0.66 / 2: 0.12 / 3: 0.54) INTW (0: -0.82 / 1: -0.02 / 2: 0.84).

Table 9: continued (5)

Subsample B (Foreigners)	
Wave	Model and Coefficients
14	Model = CONST + INTW + PAAR + HTYP + SOZH + UNTMIETE + BEGINN CONST (-0.14) INTW (0: -1.13 / 1: 0.49 / 2: 0.64) PAAR (0: -0.86 / 1: 0.86) HTYP (1: -0.42 / 2: -0.20 / 3: 0.62) SOZH (0: -0.39 / 1: 0.39) UNTMIETE (0: -0.38 / 1: 0.38) BEGINN (0: 0.36 / 1: -0.36)
15	Model = CONST + HTYP + INTW + APPRENT CONST (-1.5) HTYP (1: -0.58 / 2: -0.0 / 3: 0.58) INTW (0: -0.76 / 1: 0.21 / 2: 0.55) APPRENT (0: -0.22 / 1: 0.22)
16	Model = CONST + INTW + HTYP + FAMSTD + ANZASSET + PAAR CONST (-0.58) INTW (0: -0.92 / 1: 0.17 / 2: 0.75) HTYP (1: -0.28 / 2: -0.30 / 3: 0.58) FAMSTD (1,2,4,5: -0.34 / 3: 0.34) ANZASSET (0: -0.29 / 1,2,3 : 0.29) PAAR (0: -0.83 / 1: 0.83)
17	Model = CONST +INTW + HTYP + KAEINK + INTWTYPE + WOHN GELD CONST (-0.13) INTW (0: -0.90 / 1: 0.22 / 2: 0.68) HTYP (1: -0.74 / 2: -0.57 / 3: -1.31) KAEINK (0: -0.51 / 1: 0.51) INTWTYPE (0: -0.57 / 1: 0.57) WOHN GELD (0: -0.44 / 1: 0.44)
18	Model = CONST + INTW + HTYP + INTWART + UNTMIETE + TELINFRA CONST (-0.01) INTW (0: -0.70 / 1: 1.24 / 2: -0.54) HTYP (1: -0.83 / 2: -0.42 / 3: 1.25) INTWART (0: -0.91 / 1: 0.91) UNTMIETE (0: -0.89 / 1: 0.89) TELINFRA (0: 0.39 / 1: -0.39)
19	Model = CONST + INTW + HTYP + (PAAR=1)*(HTYP=1) + INTWART CONST (-3.25) INTW (0: -2.23 / 1: 2.24 / 2: -0.01) HTYP (1: -1.7 / 2: -0.04 / 3: 1.74) (PAAR=1*HTYP=1) (0: -1.94 / 1: 1.94) INTWART (0: -1.50 / 1: 1.50)

Table 9: continued (6)

Subsample C (East-Germans)	
Wave	Model and Coefficients
2	Model = CONST + HTYP + INTW + ALTHV + EINKO + VERLUST + OSTB CONST (-0.91), INTW (0: -0.47 / 1: -0.04 / 2: 0.51), ALTHV (1: 0.41 / 2,3,4,5,6: -0.41), HTYP (1,2: -0.84 / 3: 0.84), EINKO (1: 0.24 / 2: 0.44 / 3: 0.12 / 4: 0.00 / 5: -0.37 / 6: -0.44), VERLUST (0: -0.17 / 1: 0.17), OSTB (0: -0.29 / 1: 0.29)
3	Model = CONST + HTYP + INTW (HTYP) + ALTHV + SPAR CONST (-1.36), HTYP (1: -0.39 / 2: 0.08 / 3: 0.31), INTW (HTYP=1) (0: -0.28 / 1,2: 0.28), INTW (HTYP=2) (0: 0.42 / 1,2: -0.42), INTW (HTYP=3) (0: -0.36 / 1,2: 0.36), ALTHV (1: 0.02 / 2,3,4: -0.38 / 5: -0.20 / 6: 0.56), SPAR (0: 0.35 / 1: -0.35)
4	Model = CONST + HTYP + INTW + ALTHV + KAEINK + FAMSTD CONST (0: -0.62), HTYP (1: -0.47 / 2: 0.25 / 3: 0.22), INTW (0: -0.78 / 1: -0.04 / 2: 0.82), ALTHV (1: 0.47 / 2,3,4,5,6: -0.47), KAEINK (0: -0.54 / 1: 0.54), FAMSTD (1: -0.12 / 2: 1.13 / 3: 0.24 / 4: -0.73 / 5: -0.51),
5	Model = CONST + HTYP + INTW + KAEINK + VANZAHL + VERLUST CONST (-0.82), HTYP (1: -0.45 / 2: -0.32 / 3: 0.77), INTW (0: -0.67 / 1: -0.18 / 2: 0.84), KAEINK (0: -0.49 / 1: 0.49), VANZAHL (-0.32), VERLUST (0: -0.20 / 1: 0.20)
6	Model = CONST + HTYP + KAEINK + INTW (OSTWEST = 0) + BEGINN (OSTWEST = 0) CONST (-1.33); HTYP (1: -0.65 / 2: -0.32 / 3: 0.97); KAEINK (0: -0.66 / 1: 0.66); INTW (OSTWEST = 0) (0: -0.46 / 1: -0.31 / 2: 0.77); BEGINN (OSTWEST = 0) (0: 0.31 / 1: -0.31)
7	Model = CONST + HTYP + INTW + ALTHV + PAAR + EINKO CONST (-2.12) HTYP (1: -0.39 / 2: -0.35 / 3: 0.74) INTW (0: -0.68 / 1: 0.19 / 2: 0.49) ALTHV (1,2: 0.12 / 3: -1.25 / 4,5: 0.33 / 6: 0.8) PAAR (0: -0.42 / 1: 0.42) EINKO (1,2,3,4,5: 0.32 / 6: -0.32)

Table 9: continued (7)

Subsample C (East Germans)	
Wave	Model and Coefficients
8	Model = CONST + INTW + HTYP + ALTHV + BEGINN + UNTMIETE CONST (-1.83) INTW (0: -0.53 / 1: 0.33 / 2: 0.2) HTYP (1: -0.55 / 2: -0.30 / 3: 0.85) ALTHV(1,2,5,6: 0.29 / 3,4: -0.29) BEGINN (0: 0.27 / 1: -0.27) UNTMIETE (0: -0.34 / 1: 0.34)
9	Model = CONST + HTYP + INTW + ALTHV + TELINFRA + ANZASSET CONST (-0.76) HTYP (1: -0.77 / 2: 0.23 / 3: 0.54) INTW (0: -0.94 / 1: 0.31 / 2: 0.62) ALTHV (2,3,4,5,6: -0.54 / 1: 0.54) TELINFRA (0: 0.47 / 1: -0.47) ANZASSET (0,1: 0.35 / 2,3: -0.35)
10	Model = CONST + INTW + HTYP + PAAR + SEX + TELINFRA + INTWTYPE CONST (-1.29) INTW (0: -0.70 / 1: 0.22 / 2: 0.48) HTYP (1: -0.71 / 2: -0.11 / 3: 0.82) PAAR (0: -0.41 / 1: 0.41) SEX (0: 0.36 / 1: -0.36) TELINFRA (0: 0.30 / 1: 0.30) INTWTYPE (0: -0.45 / 1: 0.45)
11	Model = CONST + INTW + HTYP + TELINFRA + ALTHV + SIZE CONST (-1.31) INTW (0: -0.78 / 1: 0.34 / 2: 0.44) HTYP (1: -0.84 / 2: 0.02 / 3: 0.82) TELINFRA (0: 0.38 / 1: -0.38) ALTHV (2,3,4,5,6: -0.44 / 1: 0.44) SIZE (1,3,4: -0.25 / 2: 0.25)
12	Model = CONST + INTW + HTYP + TELINFRA CONST (-1.29) INTW (0: -1.30 / 1: 0.67 / 2: 0.63) HTYP (1: -0.57 / 2: -0.14 / 3: 0.71) TELINFRA (0: 0.63 / 1: -0.63)
13	Model = CONST + INTW + HTYP + TELINFRA + (EINKQU=4) + SICK + BEGINN + (SIZE=1) CONST (-1.54) INTW (0: -3.16 / 1: 1.79 / 2: 1.37) HTYP (1: -1.14 / 2: 0.14 / 3: 1.00) TELINFRA (0: 1.38 / 1: -1.38) (EINKQU=4) (0: 0.68 / 1: -0.68) SICK (0: -0.64 / 1: 0.64) BEGINN (0: 0.67 / 1: -0.67) (SIZE=1) (0: 0.53 / 1: -0.53)

Table 9: continued (8)

Subsample D (Immigrants)	
Wave	Model and Coefficients
2	Model = CONST + HTYP (MIGRANT = 1) + ALTHV (MIGRANT = 2) + KAEINK CONST (-1.08) HTYP (MIGRANT = 1) (0: -1.41 / 1,2: 1.41) ALTHV (MIGRANT = 2) (1,2,3: -0.93 / 4,5,6: 0.93) KAEINK ((0: -0.72 / 1: 0.72)
3	Model = CONST + HTYP CONST (-2.02) HTYP (1: -0.23 / 2: -0.81 / 3: 1.04)
4	Model = CONST + HTYP + TELINFRA + INTWTYPE CONST (-0.76) HTYP (1: -0.49 / 1: -0.34 / 2: 0.83) TELINFRA (0: 0.50 / 1: -0.50) INTWTYPE (0: -0.70 / 1: 0.70)
5	Model = CONST + INTW + PAAR CONST (-0.57) INTW (0: -1.02 / 1: 0.37 / 2: 0.65) PAAR (0: -1.38 / 1: 1.38)
6	Model= CONST + HTYP + ANZASSET CONST (-4.57) HTYP (1: 1.74 / 2: -6.25 / 3: 4.51) ANZASSET (2,3: -0.63 / 0,1: 0.63)
7	Model = CONST + INTW + KAEINK + BEGINN + TELINFRA CONST (0.06) INTW (0: -0.85 / 1: -0.10 / 2: 0.95) KAEINK (0: -0.98 / 1: 0.98) BEGINN (0: 0.88 / 1: -0.88) TELINFRA (0: 0.86 / 1: -0.86)
8	Model = CONST + INTW + BEGINN + TELINFRA + SEX CONST (-1.31) INTW (0: -4.29 / 1: 3.75 / 2: 0.54) BEGINN (0: 2.34 / 1: -2.34) TELINFRA (0: 1.96 / 1: -1.96) SEX (0: -1.69 / 1: 1.69)
Subsample E (Refreshment)	
2	Model = CONST + HTYP + APPRENT + ANZASSET + FAMSTD CONST (-1.54) HTYP (0: -0.88 / 1: -0.17 / 2: 1.05) APPRENT (0: -1.75 / 1: 1.75) ANZASSET (0,1: 0.35 / 2,3: -0.35) FAMSTD (1,2,3,4: -0.44 / 5: 0.44)
3	Model = CONST + HTYP + INTW + SIZE CONST (-1.39) HTYP (1: -1.02 / 2: 0.12 / 3: 0.90) INTW (0: -0.55 / 1: 0.50 / 2: 0.05) SIZE (1,2,3: 0.38 / 4: -0.38)
4	Model = CONST + HTYP + INTW + PAAR + EMAIL + SIZE CONST (-1.95) HTYP (1: -0.61 / 2: 0.04 / 3: 0.57) INTW (0: -1.31 / 1: 0.17 / 2: 1.14) PAAR (0: -0.65 / 1: 0.65) SIZE (1,2,3: 0.47 / 4: -0.47) EMAIL (0: 1.14 / 1: -1.14)

Subsample E (Refreshment)	
Wave	Model and Coefficients
5	<p>Model = CONST + CAPI + INTW + ALTHV + (INTW=1)*(SIZE=2) + (INTW=1)*(RESID=1)</p> <p>CONST (-3.36) CAPI (0: 1.34 / 1: -1.34) INTW (0: -2.74 / 1: 2.03 / 2: 0.71) ALTHV (1,2,3: -1.07 / 4,5,6: 1.07) (INTW=1)*(SIZE=2) (0: -2.36 / 1: 2.36) (INTW=1)*(RESID=1) (0: -1.64 / 1: 1.64)</p>
Subsample F (Innovation)	
2	<p>Model = CONST + HTPY + INTW + PAAR + KAEINK + APPRENT + UNI + ANZASSET + TELINFRA + EMAIL + HANDICAP + INTWART + FOREIGNHV + REGIONF</p> <p>CONST (0.10) HTYP (1: -0.24 / 2: -0.01 / 3: 0.25) INTW (0: -0.54 / 1: 1.16 / 2: -0.62) PAAR (0: -0.31 / 1: 0.31) KAEINK (0: -0.31 / 1: 0.31) APPRENT (0: 0.13 / 1: -0.13) UNI (0: 0.12 / 1: -0.12) ANZASSET (0,1,3: -0.32 / 2: 0.32) TELINFRA (0: 0.42 / 1: -0.42) EMAIL (0: 0.59 / 1: -0.59) HANDICAP (0: -0.13 / 1: 0.13) INTWART (0: -0.52 / 1: 0.52) FOREIGNHV (0: -0.24 / 1: 0.24) REGIONF (0: 0.13 / 1: -0.13)</p>
3	<p>Model = CONST + BEGINN + (INTW=1) + INTWART + FOREIGNHV + EMAIL + SICK + PAPCAP + (CAPI=1)*(HTYP=1) + RESID + BEGINN*(INTW=1) + KAEINK*(ANZASSET=1)</p> <p>CONST (-1.40) BEGINN (0: 1.59 / 1: -1.59) (INTW=1) (0: -1.27 / 1: 1.27) INTWART (0: -1.74 / 1: 1.74) FOREIGNHV (0: -0.81 / 1: 0.81) EMAIL (0: 0.78 / 1: -0.78) SICK (0: -0.69 / 1: 0.69) PAPCAP (0: -0.6 / 1: 0.6) (CAPI=1)*(HTYP=1) (0: 0.48 / 1: -0.48) RESID (0: -0.25 / 1: 0.25) BEGINN*(INTW=1) (0: -1.45 / 1: 1.45) KAEINK*(ANZASSET=1) (0: -1.29 / 1: 1.29)</p>

4 References

- Arbeitsgemeinschaft ADM-Stichproben und Bureau Wendt* 1994: Das ADM-Stichproben - System (Stand 1993). In: Gabler, S.; Hoffmeyer-Zlotnik, J.; Krebs, D. (Hg): Gewichtung in der Umfragepraxis. Westdeutscher Verlag, Opladen, S. 188-203.
- GSOEP Group* 2001: The German Socio-Economic Panel (GSOEP) after more than 15 Years. In: Vierteljahrshefte zur Wirtschaftsforschung, 2001/1.
- Haisken-DeNew, John and Joachim Frick (eds.)* 2001: Desktop Companion to the German Socio Economic Panel (GSOEP), Version 5, Berlin.
- Rendtel, Ulrich* 1991: Die Schätzung von Populationswerten in Panelerhebungen. Allgemeines Statistisches Archiv, 75, S. 225-244.
- Rendtel, Ulrich* 1995: Panelausfälle und Panelrepräsentativität, Campus Verlag, Frankfurt/M.-New York.
- Rendtel, Ulrich, Markus Pannenberg and Stefan Daschke* 1997: Die Gewichtung der Zuwanderer-Stichprobe des Sozio-ökonomischen Panels (SOEP), in: Vierteljahrshefte zur Wirtschaftsforschung, Heft 2, S. 271-286.
- Schupp, Jürgen and Gert Wagner* 1995: Die Zuwandererstichprobe des Sozio-oekonomischen Panels (SOEP). Vierteljahrshefte zur Wirtschaftsforschung, Heft 1.
- Spiess, Martin; Rendtel, Ulrich* 2000: Combining an ongoing panel with a new cross-sectional sample, DIW Discussion Paper No. 198.
- Spiess, Martin* 2001: Derivation of design weights: The case of the German Socio-Economic Panel (GSOEP), DIW Research Notes No. 5.