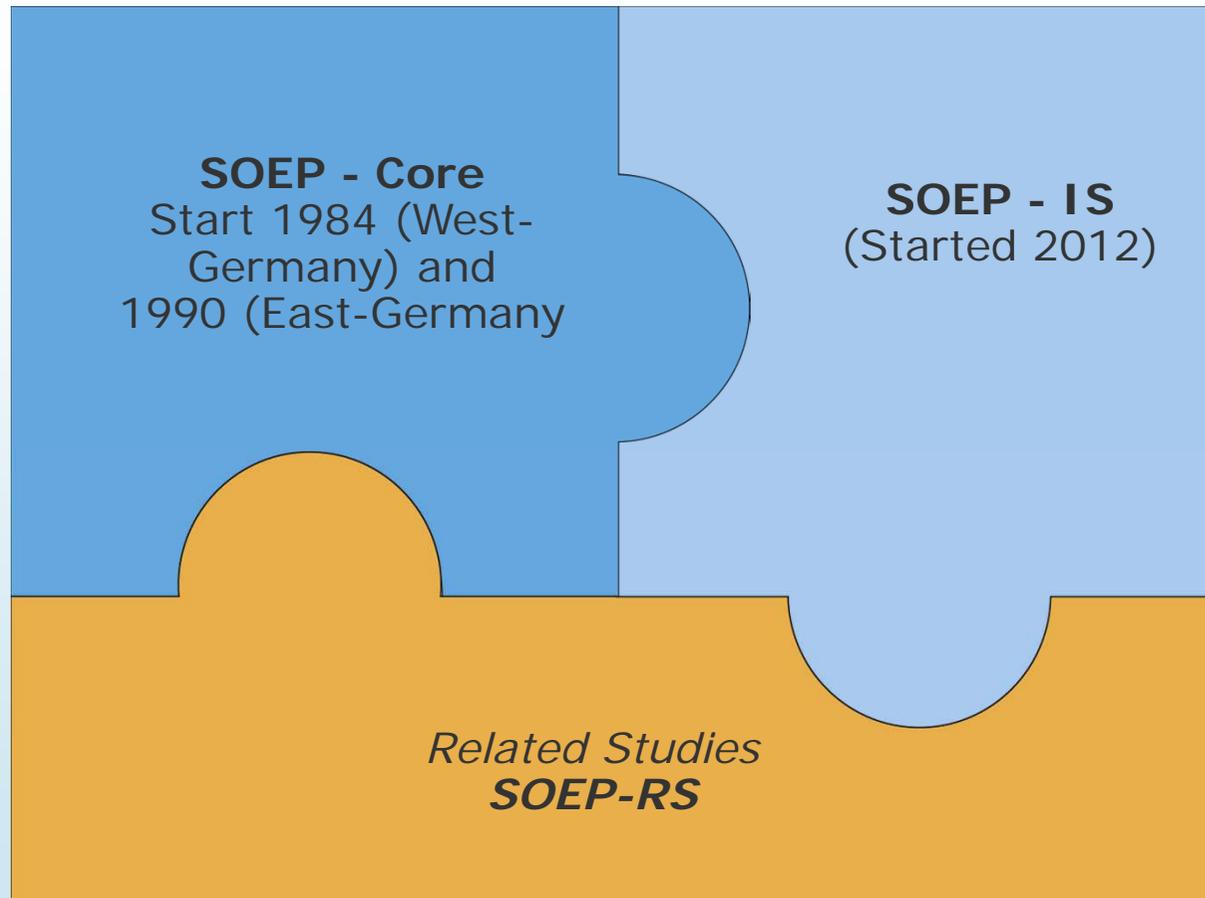


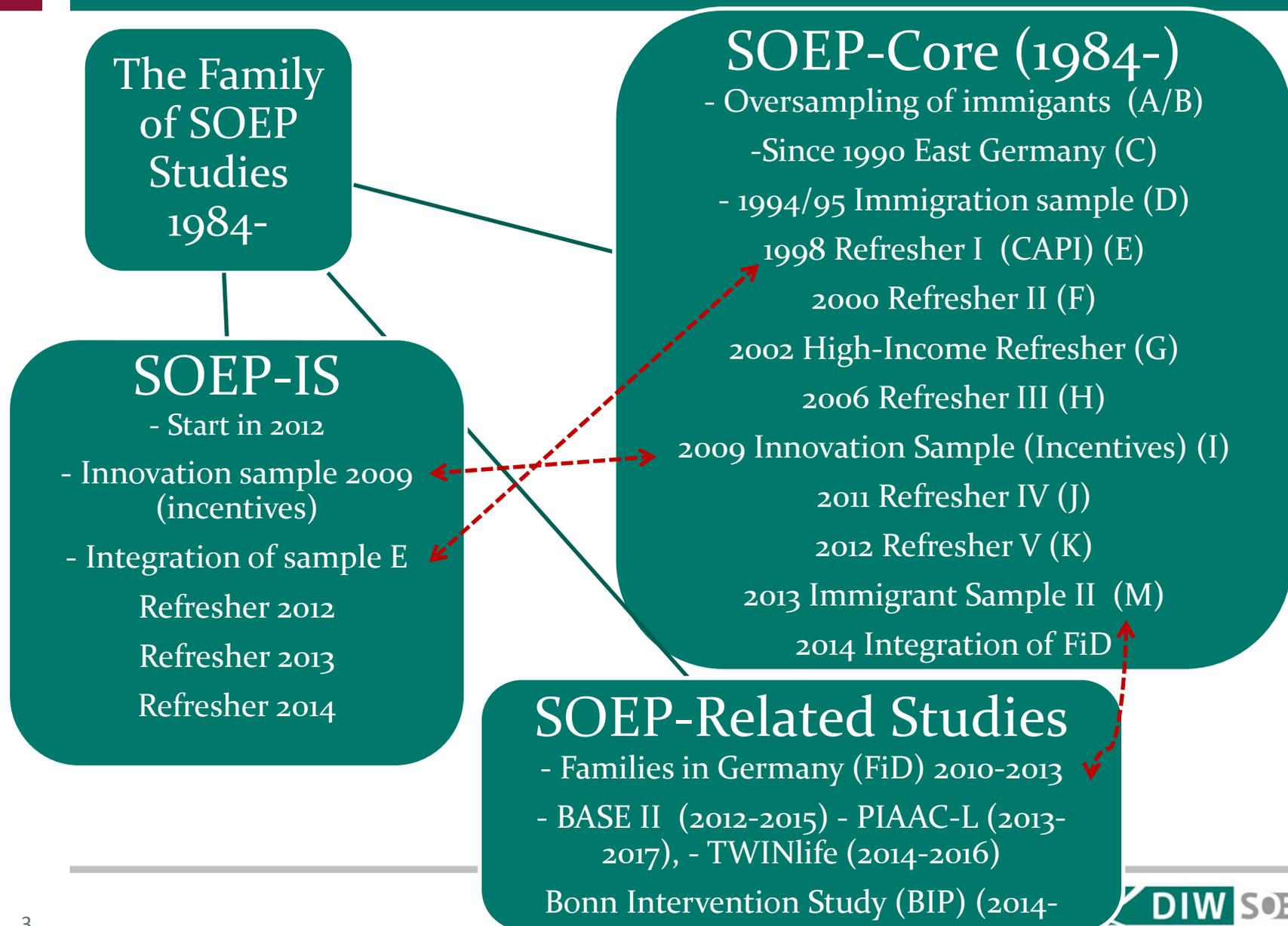
Changing from CAPI to CAWI in an ongoing household panel – experiences from the German Socio-Economic Panel (SOEP)

Jürgen Schupp and Denise Saßenroth

The Research Infrastructure SOEP at the German Institute for Economic Research (DIW-Berlin), Germany

Presentation at the “6th Conference of the European Survey Research Association (ESRA)”, 13th-17th July 2015 in Reykjavik, Iceland





Background FiD

- Systematic evaluation of policies for children and families. Initiated by the Ministry of Family Affairs (BMFSFJ) and the Ministry of Finance (BMF)
- Challenge: not enough data for specific groups
“Without additional data only a limited number of policies regarding families and children can be evaluated.” (Feasibility study, December 2008)



- Cooperative development of questionnaire contents:
 - SOEP comparability
 - Additional information for evaluation purposes
 - First data collection in March 2010
 - Restricted delivery of first waves to contracted evaluation teams

- Challenge: relevant groups for the evaluation are in the SOEP, but numbers are low
- Goal to increase the number of cases for...
 - Single parents (SP)
 - Large families (LF)
 - Low-income families (LI)
 - Families with (very) young children (including migrants)
- Sampling Frame
 - Register for (very) young children
 - Screening for SP, LF, LI

- Face-to-face CAPI interviews (except parent questionnaire)
- Mode switch not allowed in later waves
- New incentive scheme specifically for this sample:
 - €5 for HH+1st Person, €5 for each additional individual or youth questionnaire
 - If household complete: +€5/child
 - In-kind incentives targeted at the children: balloons, washcloths, bibs, reflectors, pencils, erasers, ...
 - Panel care: letters for Children's Day on November 20

Integration of FiD into LiD (SOEP-Core)

2010-2012: FiD financed by external funds (BMF & BMFSFJ)

- 66 – 80 % identical questions towards SOEP-Core; ministry influence
- data generation and restricted distribution to evaluating teams

2013: FiD financed by external funds (BMFSFJ)

- questions moving towards coreSOEP; ministry influence remains
- continued data generation and distribution

2014: end of FiD as related study - financed by SOEP funds

- runs parallel as LiD (SOEP-Core), with identical questionnaires
- Split of mode for screening and cohort samples of FiD
- ***Quasi-experimental conversion of screening sample to CAWI***

2015: First joint data distribution with SOEP-Core

- “old” FiD-waves (2010-2013) included in the SOEP-Core distribution

The Mode Experiment in FiD

Due to budget constraints we had the choice

- (Randomly) drop households in FiD-Sample
- look for cost reduction

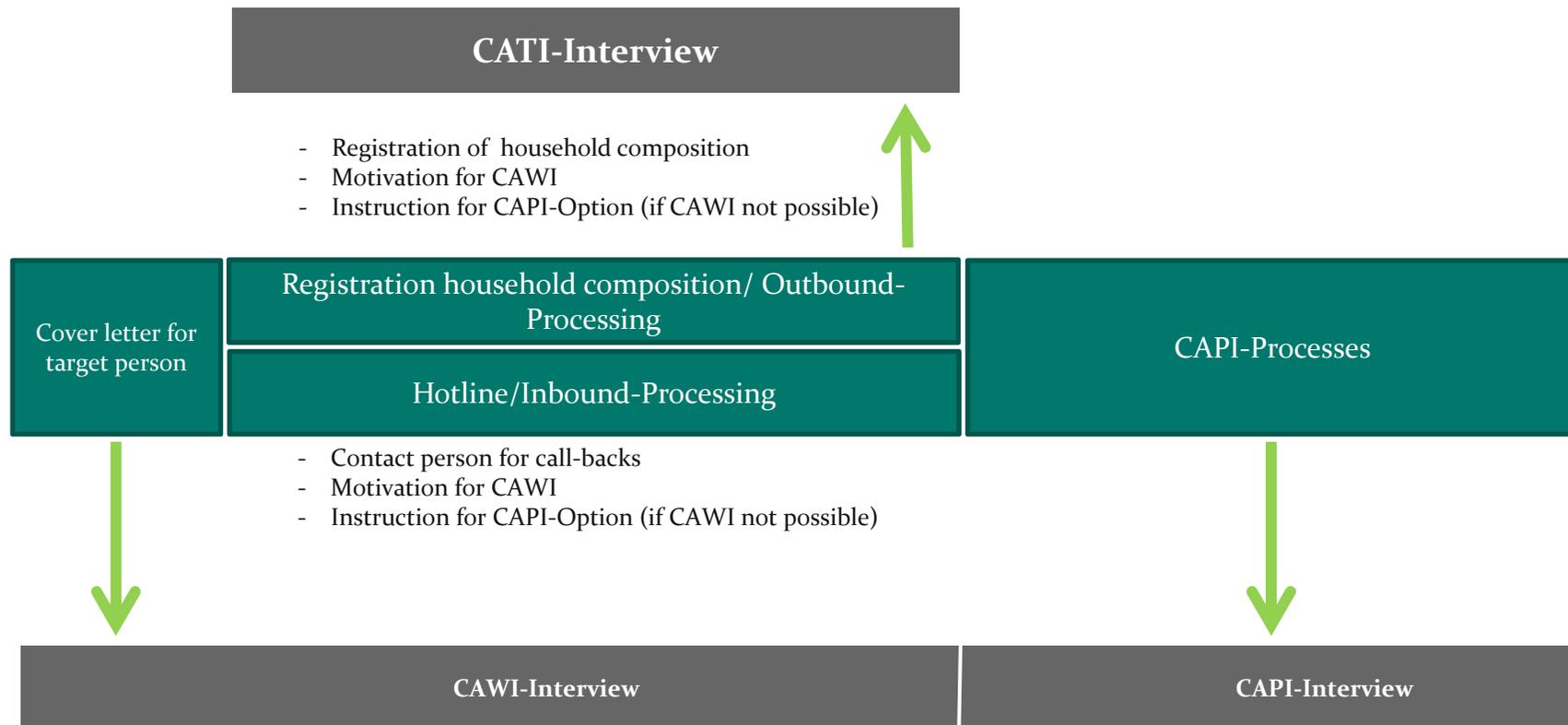
Continuity of cohort samples

- Letter of new funding for the study and conversion in SOEP-Core
- CAPI mode only – in majority of households same interviewers as last year

Screening Samples - Interviewer Assisted (IA) CAWI-CAPI multi-mode

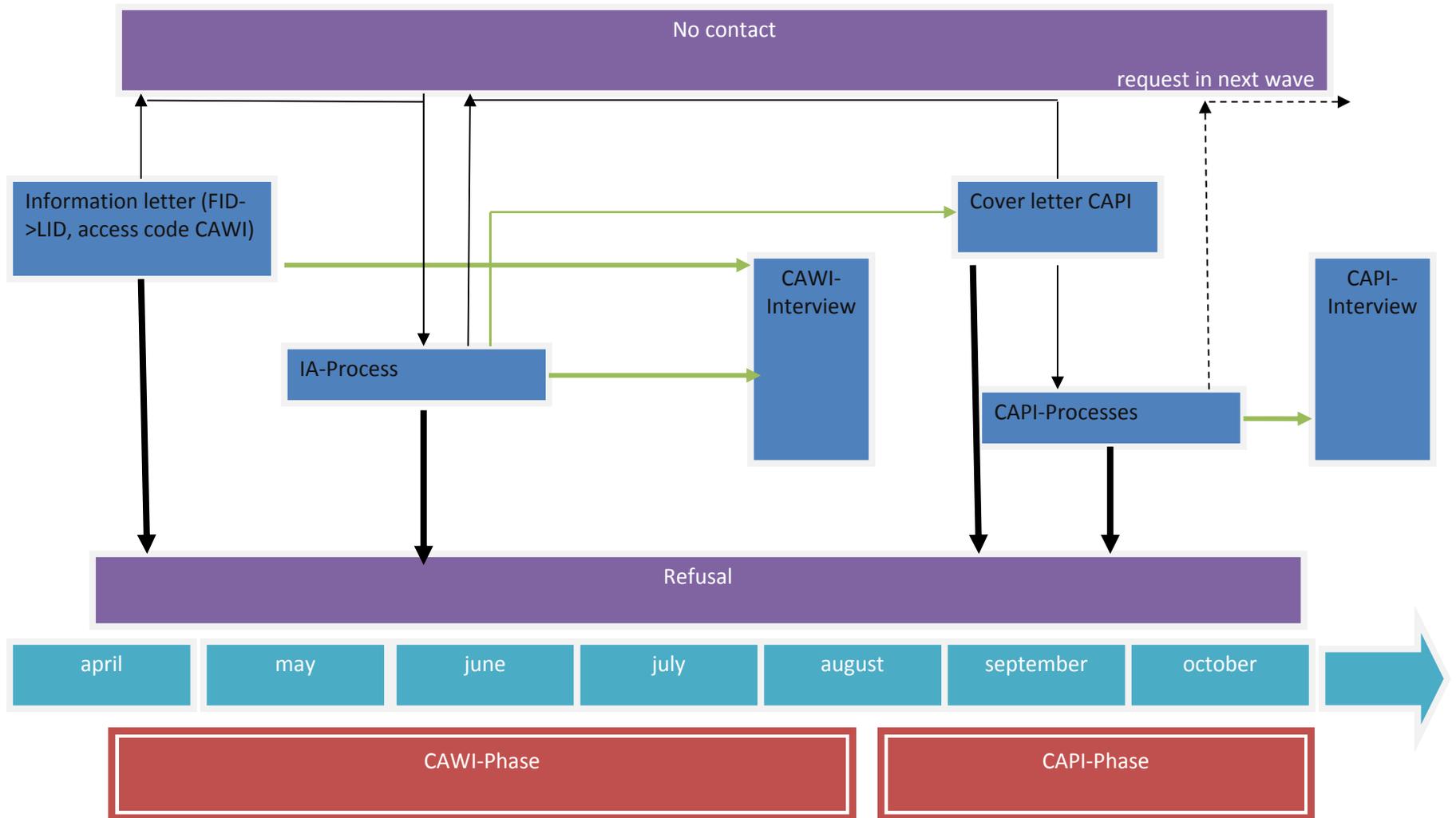
- Estimated reduction of survey expenses per realized household by nearly **30 %**
- complete integration of all households of FiD into SOEP
- New insights for survey research
- Estimate of the field institute for screening

Implementation Design



4

Implementation of IA-CAWI-CAPI mode



4

Current knowledge of conversion to CAWI

- Great **potential for cost savings** due to the conversion to CAWI (Jäckle, Lynn & Burton 2013)
- **Lower response rates** and **higher PUNR** in the mixed-mode design (Jäckle, Lynn & Burton 2013)
- **Higher item nonresponse** in mixed-mode design (Jäckle, Lynn & Burton 2013)
- CAWI non-respondents **have higher nonresponse rates** in subsequent CAPI mode (Collins & Mitchell 2014)
- Lower motivation to participate in CAPI mode if at least one household member had already participated in CAWI (Collins & Mitchell 2014)
- Respondents are more likely to respond in the web mode if email contact was established by survey agency (Cernat & Lynn 2014)
- Nonresponse in web surveys is strongly related to computer literacy and the intensity of computer and Internet use (Callegaro, Lozar Manfreda & Vehovar 2015)

4

History and Starting Point for fieldwork 2014

	2011				2013				
	gross	net	PUNR	Response Rate	gross	net	PUNR	Response Rate	Cumulated RR 2011-13
Low-income families (LI)	1.080	883	7,2	81,8	993	820	6,6	82,6	75,9
Single Parents (SP)	1.163	998	3,0	85,8	1.024	840	7,3	82,0	72,2
Large Families (LF)	1.198	1.001	9,9	83,6	1.089	901	10,1	82,7	75,2
Screening total	3.441	2.882	6,7	83,8	3.106	2.561	8,1	82,5	74,4
Register - birth cohort 2007	524	408	9,3	77,9	447	338	7,4	75,6	64,5
Register - birth cohort 2008	546	419	8,6	76,7	448	353	8,2	78,8	64,7
Register - birth cohort 2009	510	405	9,4	79,4	422	330	3,0	78,2	64,7
Register - birth cohort 2010	530	415	8,7	78,3	443	341	4,4	77,0	64,3
Register for (very) young children total	2.110	1.647	9,0	78,1	1.760	1.362	5,8	77,4	64,5
Screening and Register total	5.551	4.529	7,5	81,6	4.866	3.923	7,3	80,6	70,7

5

Overall Result of Fieldwork 2014 (FiD-LiD)

	2014				
	gross	net	PUNR	Response Rate	Cumulated RR 2011-14
Low-income families (LI)	918	648	15,1	70,6	60,0
Single Parents (SP)	929	649	10,9	69,9	55,8
Large Families (LF)	1.021	718	19,2	70,3	59,9
Screening total	2.868	2.015	15,2	70,3	58,6
Register - birth cohort 2007	377	308	5,5	81,7	58,8
Register - birth cohort 2008	400	323	6,2	80,8	59,2
Register - birth cohort 2009	357	308	3,6	86,3	60,4
Register - birth cohort 2010	384	308	5,8	80,2	58,1
Register for (very) young children total	1.518	1.247	5,3	82,1	59,1
Screening and Register total	4.386	3.262	9,5	74,4	58,8

5

Participation 2014 by subsamples (household level)

	Response	Refusal	Ineligibles	Total
Cohort 2007	308	63	6	377
Cohort 2008	323	62	15	400
Cohort 2009	308	40	9	357
Cohort 2010	308	51	25	384
Low-income families	648	235	35	918
Single parents	649	236	44	929
Large families	718	270	33	1,021
Total	3,262	957	167	4,386

4

Determinants of Willingness in Wave 2014

	Low-income	single parents	Large families
Information 2014			
urban	-0.57*** (0.19)	0.14 (0.19)	-0.33* (0.19)
New address	-0.19 (0.23)	-0.10 (0.21)	-1.07*** (0.29)
Email address known	0.73*** (0.17)	0.90*** (0.17)	0.73*** (0.17)
Same interviewer 1 wave, ref: 0	-0.06 (0.41)	0.01 (0.39)	-0.80* (0.42)
2 waves	0.22 (0.35)	0.51* (0.26)	-0.52* (0.31)
3 waves	0.26 (0.24)	0.29 (0.26)	-0.54* (0.31)
Int. contacts 3+	0.42** (0.21)	0.12 (0.20)	0.40* (0.22)
Adults in hh: 2, ref: 1	0.79*** (0.24)	0.08 (0.19)	0.72*** (0.28)
3+	0.57* (0.31)	0.47 (0.43)	0.61** (0.30)
mean age of adults in hh	0.01 (0.01)	0.05*** (0.01)	0.04*** (0.01)
Internet in hh, ref: no internet	0.29 (0.27)	0.12 (0.29)	0.11 (0.35)
No information	-3.05*** (1.17)	-0.84 (0.64)	-2.91*** (0.81)
Information 2013			
Low income quartile, ref: INR	-1.02 (1.11)	0.23 (0.53)	0.48 (0.63)
Median income	-1.50 (1.11)	-0.11 (0.53)	0.14 (0.56)
High income q.	-1.42 (1.17)	0.24 (0.62)	-0.08 (0.56)
Questionnaires 2+	0.25 (0.24)	0.12 (0.18)	-0.25 (0.30)
Constant	0.76 (1.33)	-2.13** (0.91)	0.63 (1.15)
R ²	0.12	0.08	0.18
N	883	885	988

4

Mode of data collection 2014 by subsamples

	Total gross *	CAWI gross	CAWI net	CAWI PUNR %	CAWI RR %	CAPI gross	CAPI gross to total gross %	CAPI Netto	CAPI PUNR	CAPI RR	Total net	Total PUNR %	Total RR %
Low-income families (LI)	918	899	330	20,9	36,7	508	55,3	318	9,1	62,6	648	15,1	70,6
Single Parents (SP)	929	917	355	12,1	38,7	473	50,9	294	9,5	62,2	649	10,9	69,9
Large Families (LF)	1.021	1.006	437	20,1	43,4	445	43,6	281	17,8	63,1	718	19,2	70,3
Screening total	<u>2.868</u>	<u>2.822</u>	<u>1.122</u>	17,8	39,8	1.426	49,7	893	12,0	62,6	2.015	15,2	70,3

* Including new households identified in CAPI line

5

Determinants of CAWI mode for CAWI & CAPI respondents 2014

	Low-income	single parents	Large families
urban	0.02 (0.18)	0.01 (0.20)	-0.04 (0.19)
New address	-0.26 (0.25)	-0.40 (0.26)	-2.13*** (0.48)
Email address known	1.00*** (0.18)	0.83*** (0.19)	1.22*** (0.18)
Same interviewer 1 wave, ref: 0	0.71 (0.48)	-0.24 (0.46)	-0.75* (0.43)
2 waves	-0.10 (0.36)	-0.46 (0.30)	-0.20 (0.29)
3 waves	-0.01 (0.25)	-0.27 (0.30)	-0.39 (0.29)
Adults in hh: 2, ref: 1	-0.43 (0.27)	-0.04 (0.21)	0.18 (0.32)
3+	-0.25 (0.34)	-0.25 (0.45)	-0.05 (0.35)
mean age of adults in hh	-0.01 (0.01)	-0.01 (0.01)	-0.00 (0.02)
Internet some users, ref: all user	-0.31* (0.18)	-0.22 (0.21)	-0.14 (0.17)
No internet	-2.03*** (0.57)	-1.76*** (0.51)	-0.93 (0.62)
Low income quartile, ref: INR	-0.61 (0.49)	0.26 (0.36)	-0.34 (0.55)
Median income	-0.05 (0.49)	0.36 (0.36)	-0.07 (0.48)
High income q.	0.61 (0.61)	0.91* (0.48)	0.60 (0.48)
Questionnaires 2+	0.02 (0.25)	-0.07 (0.19)	0.10 (0.29)
Constant	0.65 (0.85)	0.46 (0.82)	1.91* (1.06)
R ²	0.09	0.06	0.13
N	648	649	718

5

Results of the CATI-Screening (in %)

	Pro CAWI	Contra CAWI	No information	Total
Low-income families	64.0	10.2	25.8	648
Single parents	70.1	11.3	18.6	649
Large families	76.9	8.1	15.0	718
Total	70.6 1,422	9.8 197	19.7 396	2,015

5

Determinants of CAWI participation for persons with CAWI acceptance

	Low-income	single parents	Large families
urban	0.15 (0.24)	-0.25 (0.27)	0.08 (0.22)
New address	-0.09 (0.25)	-0.63* (0.35)	-2.71*** (0.81)
Email address known	1.22*** (0.24)	0.60** (0.23)	1.15*** (0.22)
Same interviewer 1 wave, ref: 0	0.38 (0.60)	-0.00 (0.63)	-0.88* (0.53)
2 waves	-0.11 (0.47)	-0.59 (0.38)	-0.51 (0.39)
3 waves	0.18 (0.33)	-0.60 (0.38)	-0.79*** (0.39)
Adults in hh: 2, ref: 1	-0.39 (0.35)	0.08 (0.26)	0.07 (0.46)
3+	-0.51 (0.43)	-0.08 (0.57)	-0.37 (0.50)
mean age of adults in hh	-0.01 (0.02)	-0.01 (0.01)	0.01 (0.02)
Internet some users, ref: all user	-0.75*** (0.25)	0.05 (0.27)	-0.03 (0.21)
No internet	-2.18** (0.90)	-1.69** (0.73)	-0.28 (0.97)
Low income quartile, ref: INR	-0.48 (0.63)	0.34 (0.47)	-0.07 (0.70)
Median income	0.21 (0.63)	0.42 (0.47)	-0.20 (0.58)
High income q.	0.54 (0.76)	1.06* (0.63)	0.25 (0.58)
Questionnaires 2+	0.21 (0.32)	-0.24 (0.23)	0.35 (0.36)
Constant	0.72 (1.11)	1.60 (1.03)	2.70* (1.43)
R ²	0.10	0.05	0.10
N	415	455	552

Overall Nonresponse

- Overall nonresponse does not show systematic patterns
- Differences between subsamples and waves
- Constant finding: higher participation probability if email address is registered by Infratest
- long interviewer binding positive related to participation
 - Effect turns negative for participation 2014 among CAWI-subsamples!

Participation in CAWI mode

- Higher probability for households where all members are using the internet and if the email address is known by Infratest
- Decreased probability for households without internet access.

Summary

- Conversion from interviewer administered longitudinal household panel to CAWI can reduce costs of up to 30% with a response optimized strategy
- Successful implementation of the CAWI mode in an ongoing household panel study
 - High overall motivation to participate in the CAWI mode
 - Higher Partial Unit Nonresponse (about 10% points)
 - Lower actual response rate than expected
 - About 78 % of gross sample agreed to participate, but “only” 58 % participated at the end in CAWI-mode
 - Low acceptance of conversion to CAWI among interviewers 
- Conversion at an earlier stage WITHOUT already high interviewer loyalty might result in higher response rates in CAWI-line

Conclusion

- Minor concern: overall nonresponse
 - can be resolved with weighting adjustments
- Major concern: complete shift to CAWI mode
 - Remarkably lower response rates
 - Increasing risk of biased data
- Solution: sequential CAWI-CAPI mode is mandatory
- Recommendation: implementation of CAWI at an early stage of the panel, before interviewer binding is strong

References

- Callegaro, Mario, Katja Lozar Manfreda & Vasja Vehovar (2015): “Web Survey Methodology”. Sage publications.
- Cernat, Alexandru & Peter Lynn (2014): “The role of email addresses and email contact in encouraging web response in a mixed mode design”. *Understanding Society Working Paper Series No. 2014 – 10*.
- Collins, Debbie & Martin Mitchell (2014): “Role of mode in respondents’ decisions to participate in IP5: Findings from a qualitative follow-up study.” *Understanding Society Working Paper Series No. 2014 – 03*.
- Jäckle, Annette, Peter Lynn & Jon Burton (2013): “Going Online with a Face-to-Face Household Panel: Initial Results from an Experiment on the Understanding Society Innovation Panel”. *Understanding Society Working Paper Series No. 2013 – 03*.
- Schröder, Mathis, Rainer Siegers & C. Katharina Spieß (2013a): "Familien in Deutschland" - FiD - Enhancing Research on Families in Germany. *SOEPpapers on Multidisciplinary Panel Data Research No. 556*.
- Mathis Schröder, Rainer Siegers, C. Katharina Spieß (2013b): "Familien in Deutschland" *Schmollers Jahrbuch ; 133(4), 595-606*

Vielen Dank



Research Infrastructure Socio-Economic Panel Study (SOEP)
at DIW Berlin — German Institute
for Economic Research e.V.
Mohrenstraße 58, 10117 Berlin
www.diw.de/soep

jschupp@diw.de,
dsassenroth@diw.de



Backup

a

Internet Access 2013 by subsamples (household level)

Sample	No Internet	Internet	Total
Low income	10.5%	89.5%	820
Single parent	8.3%	91.7%	840
Children 3+	6.2%	93.8%	901
Cohort 2007	6.8 %	93.2 %	338
Cohort 2008	10.5 %	89.5 %	353
Cohort 2009	8.2 %	91.8 %	330
Cohort 2010	10.3 %	89.7 %	341
Total	8.5 %	91.5 %	3,923

d

Participation 2013 by subsamples (household level)

	Response	Refusal	Ineligibles	Total
Cohort 2007	338	106	3	447
Cohort 2008	353	90	5	448
Cohort 2009	330	87	5	422
Cohort 2010	341	96	6	443
Low-income families	820	159	14	993
Single parents	840	173	11	1,024
Large families	901	170	18	1,089
Total	3,923	881	62	4,866

Determinants of Willingness in Wave 2013

	cohort 2007	cohort 2008	cohort 2009	cohort 2010	Low-income	single parents	Large families
Information 2014							
urban	0.19 (0.30)	-0.10 (0.36)	-0.21 (0.37)	0.02 (0.34)	-0.10 (0.21)	0.02 (0.23)	-0.09 (0.22)
New address	-0.06 (0.19)	0.49** (0.20)	0.30 (0.21)	0.09 (0.25)	0.10 (0.12)	0.15 (0.14)	0.29*** (0.11)
Email address known	0.01 (0.27)	0.42 (0.29)	0.67** (0.30)	0.56** (0.28)	0.41** (0.20)	0.19 (0.20)	0.58*** (0.20)
Same interviewer 1 wave, ref: 0	-0.31 (0.59)	0.39 (0.68)	-0.30 (0.65)	1.51 (0.93)	0.03 (0.46)	-0.24 (0.43)	-0.04 (0.39)
2 waves	-0.74 (0.56)	0.02 (0.58)	-0.60 (0.61)	-0.05 (0.53)	0.35 (0.40)	0.15 (0.31)	0.81*** (0.28)
3 waves	0.39 (0.41)	0.24 (0.42)	0.26 (0.46)	0.40 (0.44)	0.40 (0.26)	0.21 (0.32)	0.71** (0.28)
Int. contacts 3+	0.22 (0.27)	-0.35 (0.29)	-0.38 (0.31)	-0.45 (0.31)	0.03 (0.19)	-0.68*** (0.21)	-0.06 (0.20)
Adults: 2, ref: 1	-0.06 (0.52)	0.32 (0.58)	0.75 (0.49)	0.11 (0.51)	-0.44 (0.29)	-0.30 (0.25)	-0.20 (0.38)
3+	0.43 (0.79)	1.65 (1.30)	1.02 (0.92)	-1.04 (0.80)	-0.78** (0.35)	-0.43 (0.68)	-0.20 (0.41)
mean age adults	-0.01 (0.02)	0.00 (0.03)	0.03 (0.03)	-0.01 (0.03)	0.03*** (0.01)	0.01 (0.01)	0.02 (0.02)
Internet in hh ref: no internet	-0.37 (0.60)	0.17 (0.51)	0.60 (0.46)	0.35 (0.46)	-0.48 (0.40)	-0.65 (0.45)	0.54 (0.40)
No information	-2.55** (1.07)	-2.31** (1.23)	-2.49* (1.44)	-1.51 (1.48)	-1.85** (0.76)	-2.32*** (0.85)	-0.63 (0.69)
Information 2013							
Low income quartile, ref: INR	0.03 (0.87)	1.02 (0.94)	-0.03 (1.17)	1.37* (0.78)	0.51 (0.59)	1.32* (0.71)	1.82*** (0.60)
Median income	0.05 (0.83)	0.80 (0.89)	-0.43 (1.11)	1.15 (0.70)	0.72 (0.60)	1.07 (0.72)	1.86*** (0.53)
High income q.	-0.37 (0.84)	0.69 (0.90)	0.20 (1.18)	1.28* (0.75)	0.31 (0.85)	0.50 (0.78)	1.48*** (0.53)
Questionnaires 2+	0.66 (0.50)	-0.46 (0.60)	0.01 (0.84)	-0.56 (1.22)	0.57** (0.28)	-0.50* (0.28)	-0.05 (0.23)
Constant	0.78 (1.53)	0.75 (1.67)	-0.16 (2.35)	1.31 (2.69)	-0.41 (0.99)	1.89 (1.15)	-1.68 (1.06)
R ²	0.20	0.19	0.25	0.19	0.15	0.19	0.19
N	444	443	417	437	979	1,013	1,071

Determinants of Willingness in Wave 2014

	cohort 2007	cohort 2008	cohort 2009	cohort 2010	Low-income	single parents	Large families
Information 2014							
urban	-0.57 (0.40)	-0.58 (0.42)	-0.48 (0.48)	0.13 (0.41)	-0.57*** (0.19)	0.14 (0.19)	-0.33* (0.19)
New address	0.71* (0.42)	0.37 (0.46)	0.22 (0.48)	-0.41 (0.32)	-0.19 (0.23)	-0.10 (0.21)	-1.07*** (0.29)
Email a. known	0.20 (0.31)	0.85*** (0.32)	0.54 (0.38)	0.19 (0.33)	0.73*** (0.17)	0.90*** (0.17)	0.73*** (0.17)
Same interviewer 1 wave, ref: 0	-1.24 (0.77)	-0.03 (0.71)	0.72 (0.81)	-0.03 (0.99)	-0.06 (0.41)	0.01 (0.39)	-0.80* (0.42)
2 waves	0.27 (0.91)	0.26 (0.64)	-0.05 (0.70)	-0.60 (0.68)	0.22 (0.35)	0.51* (0.26)	-0.52* (0.31)
3 waves	-0.66 (0.61)	0.88** (0.45)	1.04** (0.50)	-0.07 (0.63)	0.26 (0.24)	0.29 (0.26)	-0.54* (0.31)
Int. contacts 3+	0.01 (0.33)	-0.19 (0.34)	0.25 (0.39)	0.22 (0.34)	0.42** (0.21)	0.12 (0.20)	0.40* (0.22)
Adults: 2, ref: 1	-0.10 (0.50)	1.53*** (0.48)	-1.03* (0.60)	-0.56 (0.60)	0.79*** (0.24)	0.08 (0.19)	0.72*** (0.28)
3+	0.41 (0.90)	0.44 (0.76)	-0.12 (1.27)	-0.63 (0.90)	0.57* (0.31)	0.47 (0.43)	0.61** (0.30)
mean age adults	0.01 (0.03)	-0.00 (0.03)	-0.05 (0.04)	0.03 (0.03)	0.01 (0.01)	0.05*** (0.01)	0.04*** (0.01)
Internet in hh ref: no internet	0.10 (0.75)	-0.32 (0.55)	-	-	0.29 (0.27)	0.12 (0.29)	0.11 (0.35)
No information	-1.33 (1.13)	-1.92 (1.23)	-	-	-3.05*** (1.17)	-0.84 (0.64)	-2.91*** (0.81)
Information 2013							
Low income quartile, ref: INR	0.90 (0.86)	1.46 (1.13)	-	1.84*** (0.66)	-1.02 (1.11)	0.23 (0.53)	0.48 (0.63)
Median income	0.36 (0.70)	0.23 (0.98)	-	1.44*** (0.46)	-1.50 (1.11)	-0.11 (0.53)	0.14 (0.56)
High income q.	0.59 (0.72)	0.58 (0.98)	-	1.19** (0.51)	-1.42 (1.17)	0.24 (0.62)	-0.08 (0.56)
Questionnaires 2+	0.89 (1.02)	-0.10 (0.58)	2.16*** (0.45)	-	0.25 (0.24)	0.12 (0.18)	-0.25 (0.30)
Constant	-0.43 (2.19)	-0.13 (2.01)	0.25 (1.83)	0.33 (1.62)	0.76 (1.33)	-2.13** (0.91)	0.63 (1.15)
R ²	0.14	0.15	0.15	0.07	0.12	0.08	0.18
N	371	385	348	359	883	885	988

Sampling (cont.)

- Screening sample: 4-step procedure
 1. Pre-selection of households from Infratest “omnibus” surveys (multipurpose CATI and F2F)
 2. CATI-Screening in pre-selected households for target population of single parents, large, or low income families
 - Three thresholds for LI: below €1,500 ; below €2,000; below €2,500
 3. Asking for participation if eligible (during CATI)
 4. Gross sample of target households willing to participate
- Register sample:
 - Based on community registers (EMA), draw households with children born between 2007 and 2010

- Screening 2010:
 - Single parents (500)
 - Large families (500)
 - Low income families (1000)
- Screening 2011
 - Single parents (500)
 - Large families (500)
- Cohorts 2010
 - Cohorts from 2007-2010 (500 each)

Initial Response Rates Screening (FiD)

	Screening 2010		Screening 2011	
Initial gross sample	13,653		8,400	
No contact/cooperation	51.2%		42.5%	
In target population	27.1%		15.6%	
Field gross sample	23.1%	3,149	13.3%	1,113
Participation	18.3%	79.4%	11.0%	83.0%
Usable cases	16.6%	71.9%	10.9%	82.2%