SHARE-RV:
Linking SHARE with data from records of the German Pension Fund

Christin Czaplicki and Julie Korbmacher
Workshop „Linking survey and social security data"
23.11.2012, Berlin

SHARE

- Survey of Health, Ageing and Retirement in Europe
  - Multidisciplinary
  - Population 50+
  - Face-to-face
  - bi-annual, longitudinal

The Survey of Health, Ageing and Retirement in Europe (SHARE)

Wave 1, 2004

➢ SE, DK, NL, DE, BE, FR, CH, AT, ES, IT, GR
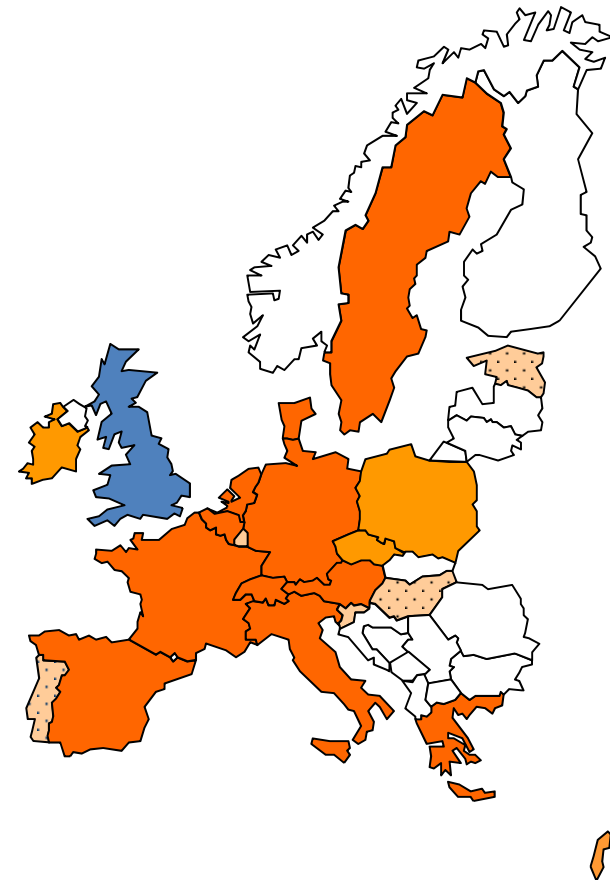
Wave 2, 2006

➢ + IE, CZ, PL, IL

Wave 3, 2008

➢ Retrospective life histories (SHARELIFE)
DE: Pilot study on record linkage

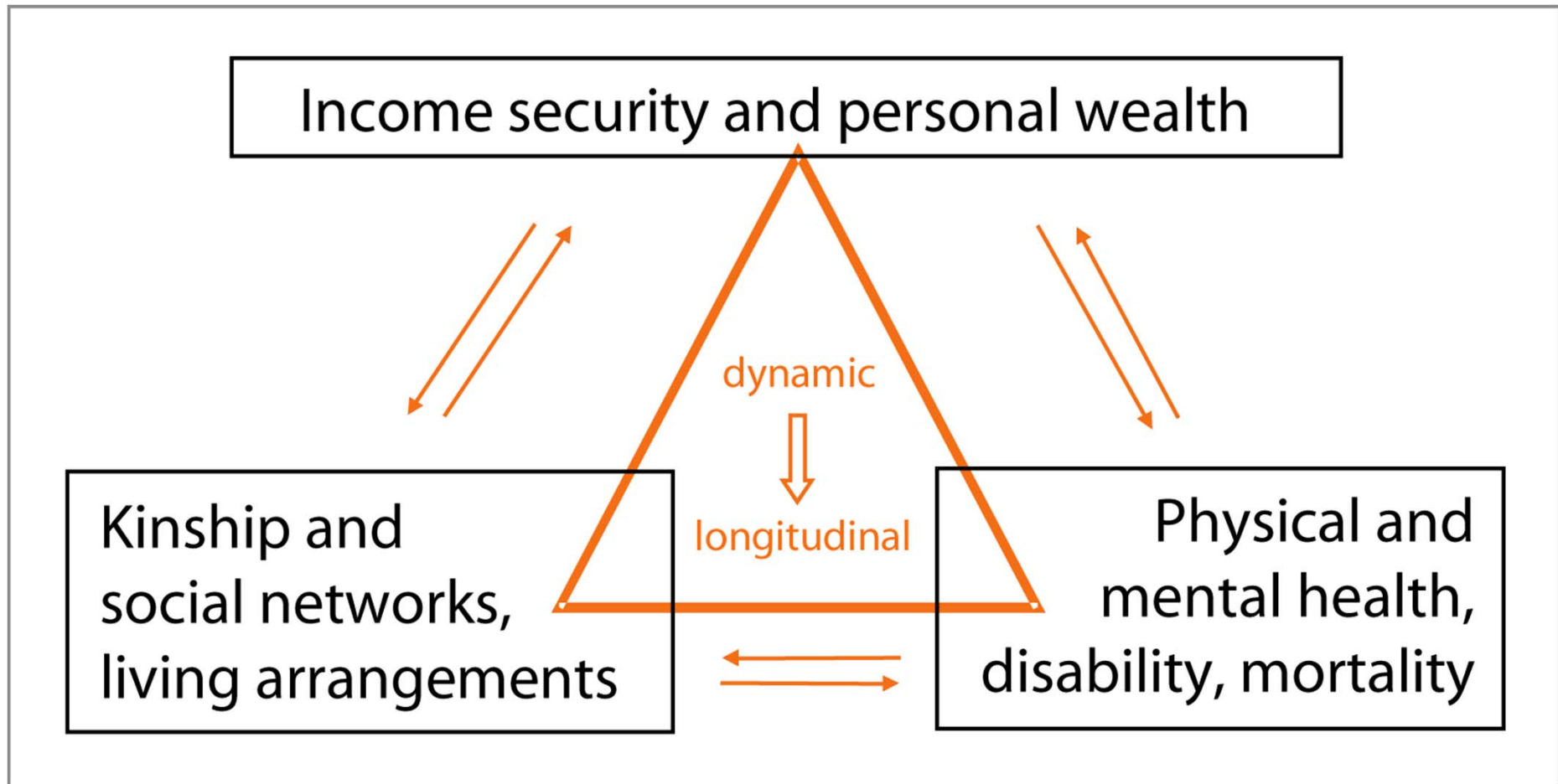**Wave 4, 2010**

➢ + PT, LUX , SV, HU, EE

expansion of the sample

social network module

DE: record linkage, biomarkers, nonresponse experiment

The Survey of Health, Ageing and Retirement in Europe (SHARE)

What do we link?

- ## Survey data (SHARE-Germany)

- ## **Administrative records of the German Pension Fund (DRV)**

  - Research Data Center
    - Longitudinal data of the insured population (VSKT)
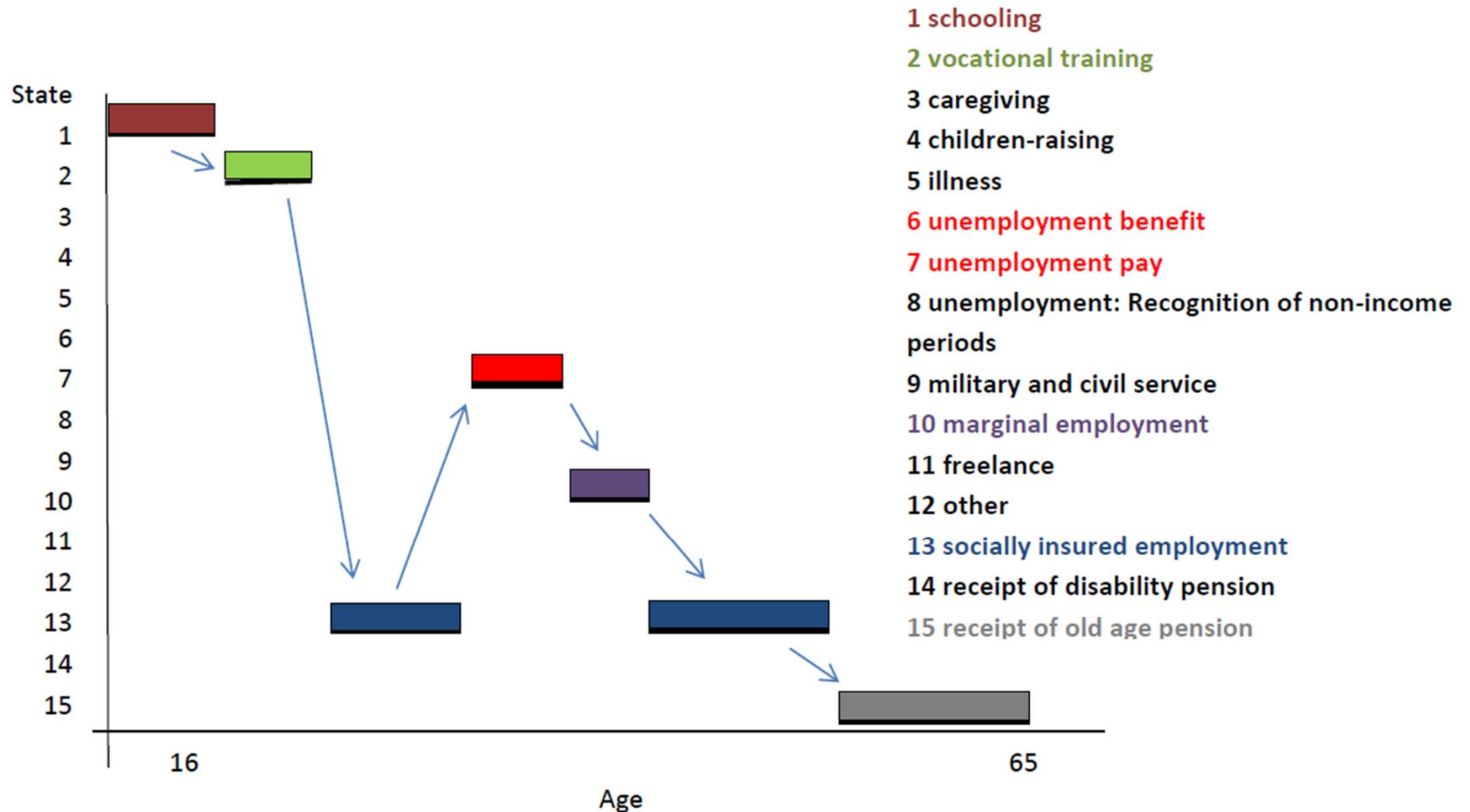   - Pension data (RTBN)

Persons included in the state pension scheme

- people who have at least one day of insured activity

- Insured activities are:

  – Schooling, university, vocational training, socially insured employment, military or civil service, child rearing, unemployment, caregiving, receipt of disability or old age pension

- Not included are following occupations:

  – civil servants, farmers, professionals (e. g. lawyers, medical doctors) who have different old age protection schemes

- No information available on spouse/household

Example for an employment biography in the administrative data

1 schooling
2 vocational training
3 caregiving
4 children-raising
5 illness
6 unemployment benefit
7 unemployment pay
8 unemployment: Recognition of non-income periods
9 military and civil service
10 marginal employment
11 freelance
12 other
13 socially insured employment
14 receipt of disability pension
15 receipt of old age pension

State

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15

16 ... 65

Age

Adapted from Stegmann, 2011: FDZ-RV: Daten, Datenzugang und Service. Neunter Workshop des FDZ-RV.

Pension data: (retirees)

- Year and month of the first payment

- ∑ earning points

- Amount of the pension paid by the DRV
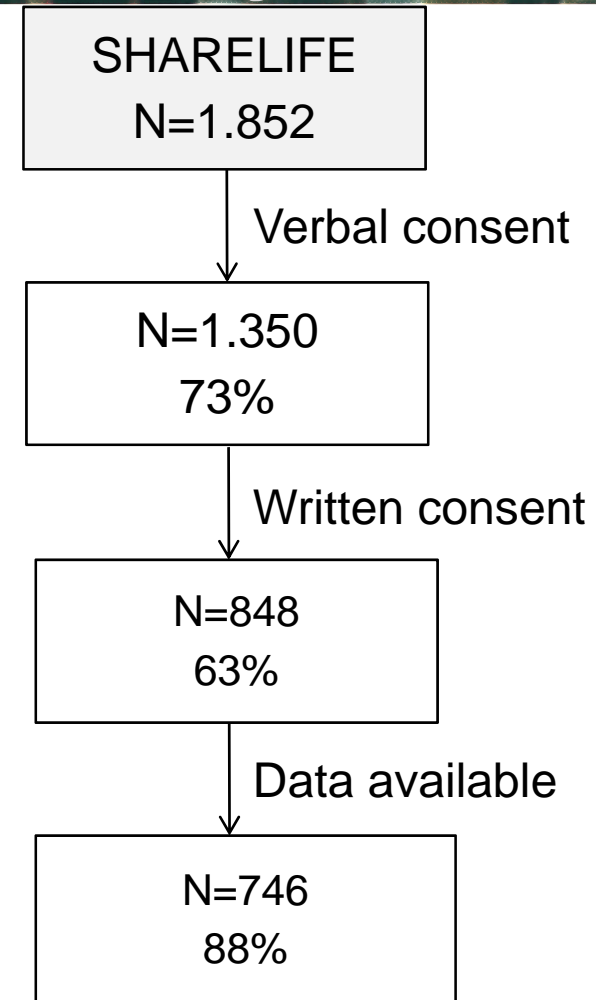
How do we link?

- Survey data (SHARE-Germany)
- Administrative records of the German Pension Fund (DRV)
  - Research Data Center
    - Longitudinal data of the insured population (VSKT)
    - Pension data (RTBN)
- Identification of data from the records of the FDZ-RV via Social Security Number (SSN)
- Written consent

SHARELIFE: Pilot Study

Summary:

- Gap between verbal and written consent

- Empty records at DRV reduce number of cases

SHARELIFE
N=1.852

Verbal consent

N=1.350
73%

Written consent

N=848
63%

Data available

N=746
88%

SHARELIFE: Release 1.0

when linking survey and administrative data in SHARE

- Identifier
- Respondents' consent

Social Security Number

1) Identifier

- Social Security Number (SSN)

Do respondents provide their SSN?

SSN provided	all	W3	W4
Yes	75.3%	88.2%	60.8%
No	24.7%	11.8%	39.2%

Identification successful in 90% → 97,5% of all consent forms

Social Security Number

1) Identifier

- possible to collect in a survey

- some additional work to identify respondents if SSN missing

- 97,5% could be identified in the records of the German Pension Fund

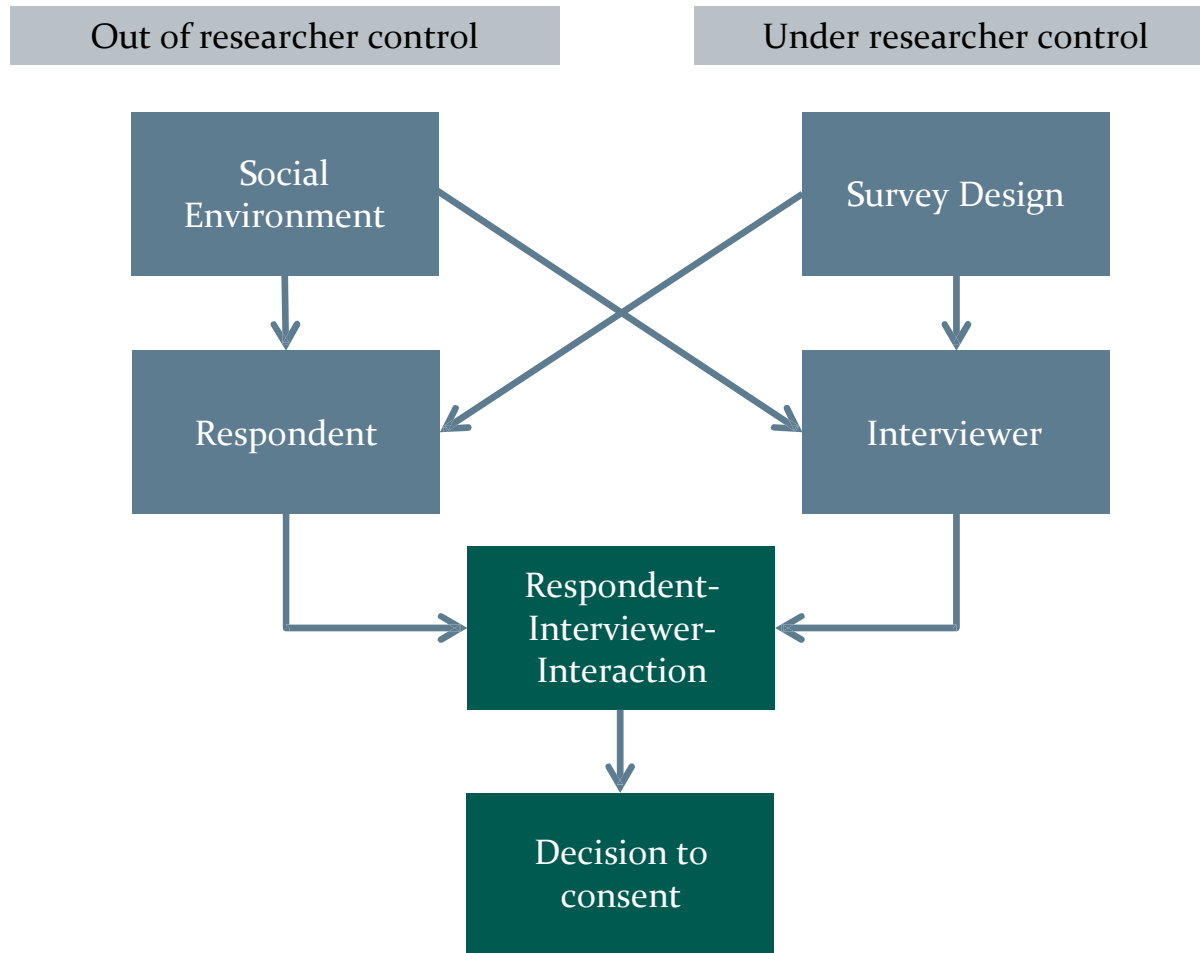2) Consent to record linkage

- Decreases the number of cases
- Systematic differences between consenters and non –consenters -> consent bias
- Determinants of consent (verbal consent)

Out of researcher control

Under researcher control

Social Environment

Survey Design

Respondent

Interviewer

Respondent-Interviewer-Interaction

Decision to consent

Adapted from: conceptual framework for survey cooperation developed by Groves & Couper, 1998

Respondent
- Demographics
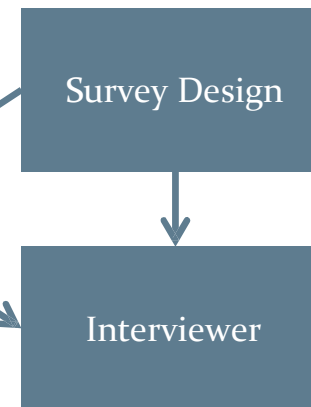- Employment status
- Household information

Respondent-Interviewer-Interaction
- Interviewer known
- Respondents comprehension
- Time
- Missing rate
- Interviewer experience
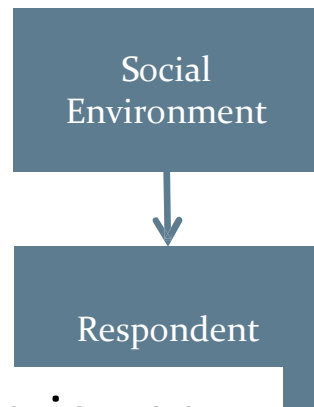
Interviewer
- Demographics
- Time
- Rounding

Out of researcher control

Under researcher control

Social Environment

Survey Design

Respondent

Interviewer

Respondent-Interviewer-Interaction

Decision to consent
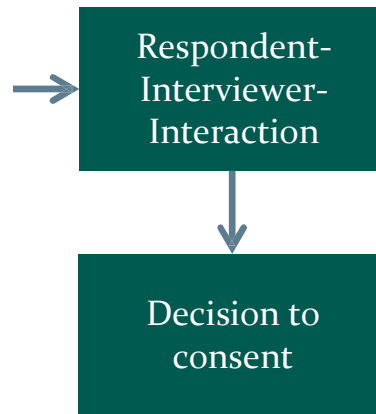
· survey cooperation developed by Groves & Couper, 1998

Determinants of consent

- Mulitlevel logistic regression

- Respondents nested within interviewers

- Significant effects on all 3 areas

- Respondent (age, living with partner, ever divorced, GDR, urban area, income)

- Interview situation (RES comprehension, missing rate, interviewers experience)

- Interviewer (age, rounding)

The role of the interviewer

- Intra Class Correation (ICC): proportion of the variance coming from the interviewer level compared to the overall variance

- 0-Model: ICC= 55%

- Final Model: ICC= 35%

- → very large interviewer influence

Administrative data are a useful extension to survey data

Room for improvement:

- More effort in interviewer training
- Make it easy for respondents
  - Huge gap between verbal and written consent
- Ask again

Thank you for your attention!

Questions or comments?