

# Far Off or Close by?

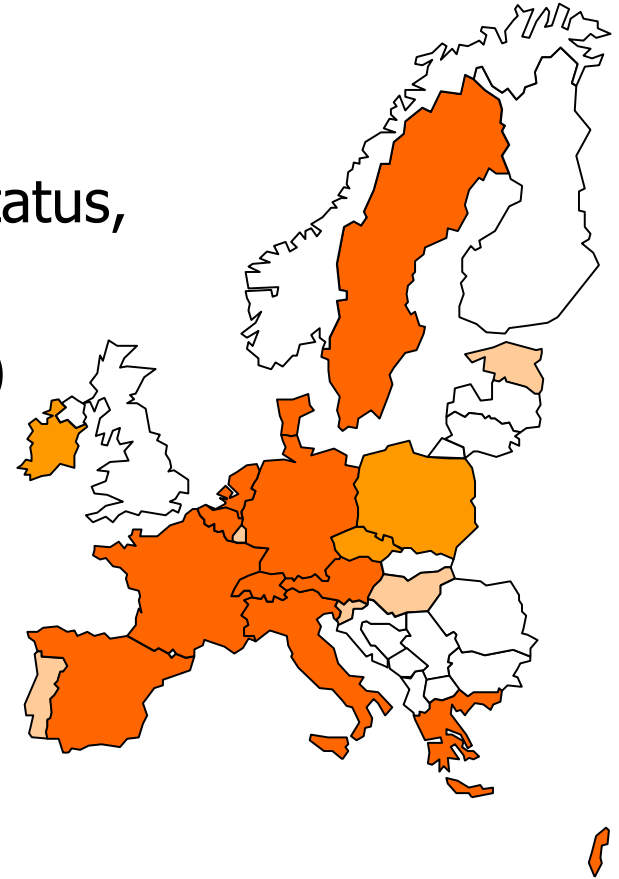
## Validating German SHARE data with Administrative Records

Julie Korbmacher  
Mathis Schröder

ITSEW July, 2010

# The Survey of Health, Ageing and Retirement in Europe

- ▶ **SHARE:**
  - ▶ Multidisciplinary
  - ▶ Micro data on health, socio-economic status, social and family networks
  - ▶ bi-annual, longitudinal (wave 1 in 2004)
  - ▶ More than 45.000 respondents (aged 50+) in more than 15 European countries
- ▶ **SHARELIFE:**
  - ▶ Retrospective life histories in wave 3



- ▶ Combination of SHARELIFE data with data from the German Pension Fund (DRV):
  - ▶ Administrative process data
  - ▶ Working history
  - ▶ German SHARE sample:
    - ▶ N= 1.940
    - ▶ Linkage via Social Security Number
    - ▶ Written consent during the interview

# The two steps of consent

---

- ▶ **Step 1:** Verbal consent to link the SHARE data with the DRV data (at the end of CAPI)
  - ▶ 1.400 consent (72%)
  - ▶ 418 refuse (22%)
  - ▶ 122 DK/NA (6%)
  
- ▶ **Step 2:** Sending back the signed letter of consent conditional on step 1
  - ▶ 898 consent (64%)
  - ▶ (849) linkable (60%)

# Determinants of consent

	consent 1		consent 2   consent 1	
	Coef	s.e.	Coef	s.e.
Women	-0.038	(0.13)	0.212	(0.12)
Age at interview	-0.020*	(0.01)	0.019*	(0.01)
Middle education	0.196	(0.18)	0.231	(0.19)
High education	0.255	(0.21)	0.344	(0.22)
Single with child	0.465	(0.39)	-0.179	(0.43)
Couple without children	0.804***	(0.16)	0.022	(0.18)
Couple with children	0.835***	(0.25)	0.042	(0.25)
other	0.631	(0.34)	-0.344	(0.36)

Significance: \*\*\*=1%; \*\*=5%; \*=10%

# Determinants of consent

	consent 1		consent 2   consent 1	
Income 1. quartile	-0.890***	(0.22)	-0.653***	(0.20)
Income 2. quartile	-0.547*	(0.22)	0.197	(0.20)
Income 3. quartile	-0.191	(0.21)	-0.136	(0.17)
Income missing	-1.149***	(0.22)	-1.193***	(0.21)
Income non-response	-0.642***	(0.17)	-0.406*	(0.20)
Still working	-0.484**	(0.17)	-0.310	(0.17)
Duration of interview	0.013***	(0.00)	-0.001	(0.00)
GDR (before 1989)	-0.579***	(0.14)	0.612***	(0.15)
_cons	2.039**	(0.68)	-0.858	(0.69)

Significance: \*\*\*=1%; \*\*=5%; \*=10%

# Determinants of consent

---

- ▶ Interviewer effects: multi level analysis
  - ▶ Highest level: interviewer
  - ▶ Lowest level: respondent
- ▶ Significant effects on the interviewer level!
  - ▶ Differences to the initial model:
    - ▶ GDR
      - ▶ Interviewer effect vs. GDR effect ?

# Determinants of consent

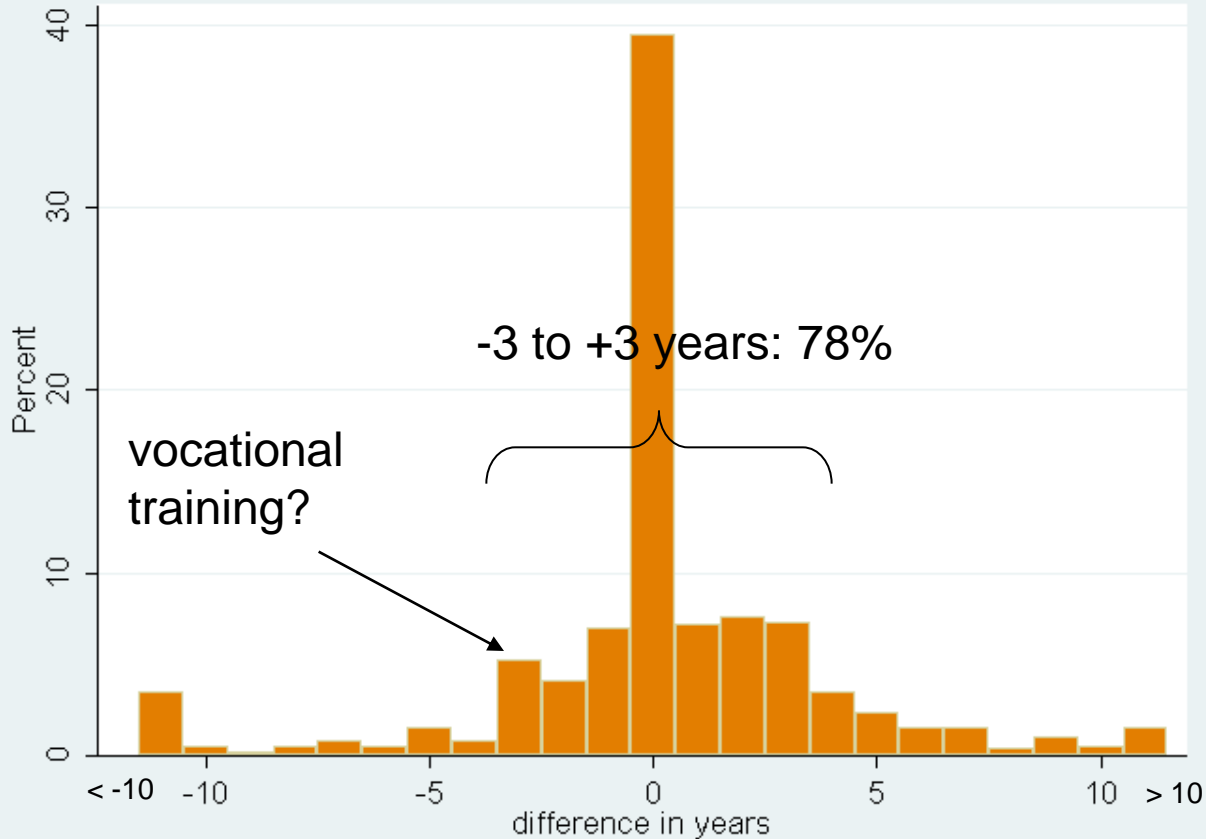
---

- ▶ No significant association:
  - ▶ Respondent's health (chronic diseases & self perceived health)
  - ▶ Trust in other people
  - ▶ Other people present during the interview
  - ▶ Problems with the interview

- ▶ Interviewer effects
  - ▶ Characteristics of the interviewer, e.g.
    - ▶ Education and age
    - ▶ Quality related indicators
- ▶ Process behind sending the consent letter back
  - ▶ Step 1 consent but no step 2 consent
    - ▶ Burden of bringing the letter to the post office
    - ▶ Social Security Number
    - ▶ => results of wave 4

- ▶ Year of first job
  - ▶ SHARELIFE: retrospective, job by job (year of beginning and end)
  - ▶ DRV: working status since age 14 (monthly)
- ▶ Comparison of the year of the first spell (SHARELIFE) & year of the first transition into employment

# SHARELIFE vs. DRV



## Which factors are responsible for the deviations?

- Personal characteristics
- Job characteristics:
  - number of different job spells
  - time passed since first job
  - kind of the first job

- ▶ 1) unit non-response
  - ▶ Identification of determinants
    - ▶ Age
    - ▶ HH-composition & income
    - ▶ Previous interview
    - ▶ Working status
  - ▶ The role of the interviewer
- ▶ 2) measurement error (retrospective data)
  - ▶ Evaluation of the error
  - ▶ Identification of determinants