

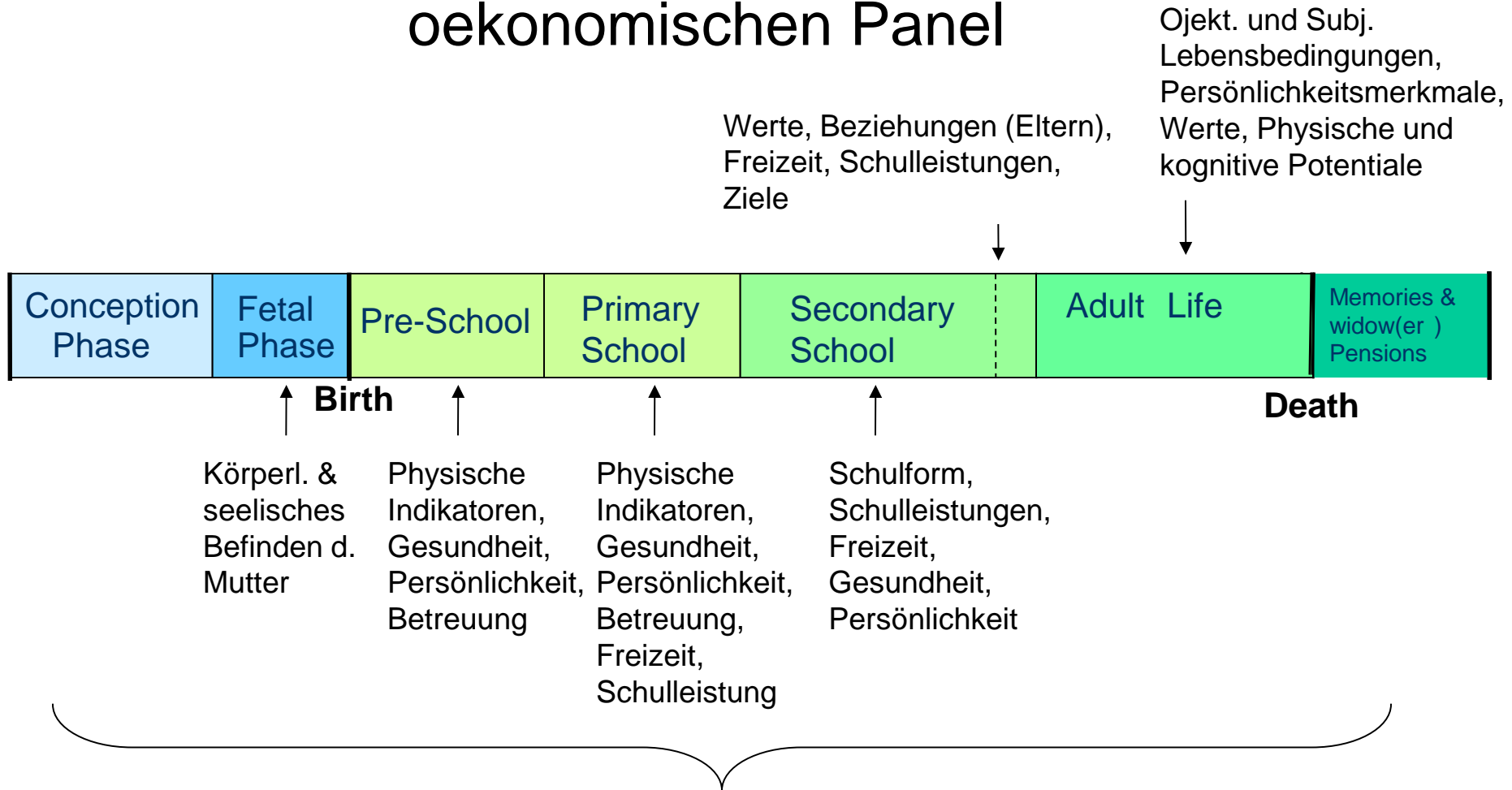
# Einführung in das SOEP

**Berlin 2014**

---

Marco Giesselmann, Jan Goebel  
(DIW Berlin, SOEP)

# Übersicht: Informationen im Sozio- oekonomischen Panel



Informationen über Lebensbedingungen der Eltern (Partner) / Haushaltskontext  
Regionaler Kontext

## *Überblick*

1. Was sind Panel-Daten?
  2. Was ist das SOEP?
    - a. Themen
    - b. Fragebögen
    - c. Zusammenfassung
  3. Datenzugang
-

## Was sind Paneldaten?

- ***Paneldaten: Wiederholte Messungen mehrerer Einheiten zu jeweils identischen Zeitpunkten***
  - ***Abgrenzungen: ...***
-

# *Datentypen*

## *A) Cross-Section Data*

	2000	2001	2002	2003	2004	2005	2006	2007
Unit 1				X				
Unit 2				X				
Unit 3				X				

---





# *Datentypen*

## *D) Pooled Cross-Section Data*

	2000	2001	2002	2003	2004	2005	2006	2007
Unit 1				X				
Unit 2					X			
Unit 3						X		

---



## *Vorteile von Paneldaten*

- ⇒ Verknüpfung von Informationen aus verschiedenen Lebensphasen
  - ⇒ Identifikation von Übergängen/Ereignissen sowie deren Auswirkungen
  - ⇒ Kontrolle zeitkonstanter Störvariablen durch die Längsschnittperspektive (→ Annäherung an Kausalität)
  - ⇒ (Trendanalysen)
  - ⇒ Dekomposition von Trends
-

## *Vorteile von Paneldaten, insb SOEP: Illustrationen*

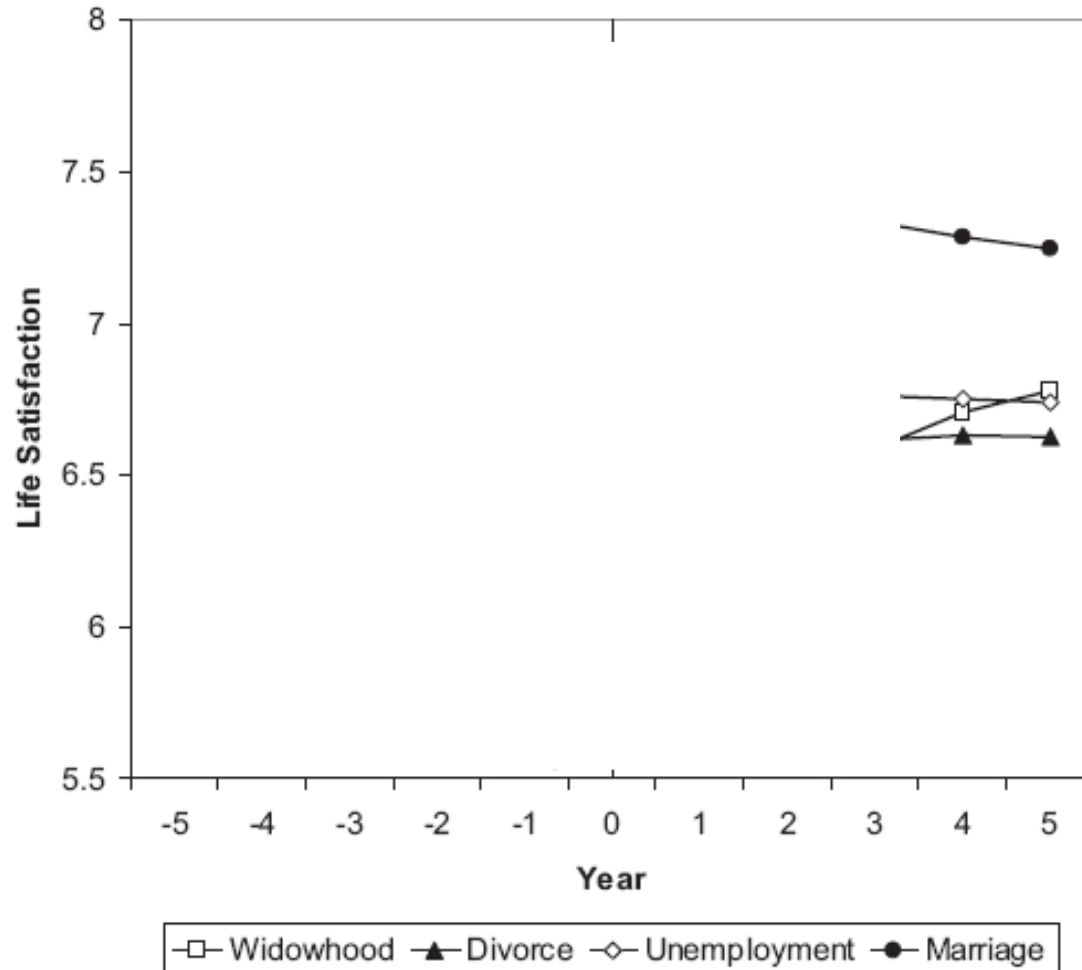
### **Analysen zum Einfluß von Ereignissen:**

Beispiel: Wie wird die Lebenszufriedenheit durch sozio-demografische Veränderungen beeinflusst?

(Diener et al. 2006)

---

**Figure 2**  
*Adaptation to Good and Bad Events*

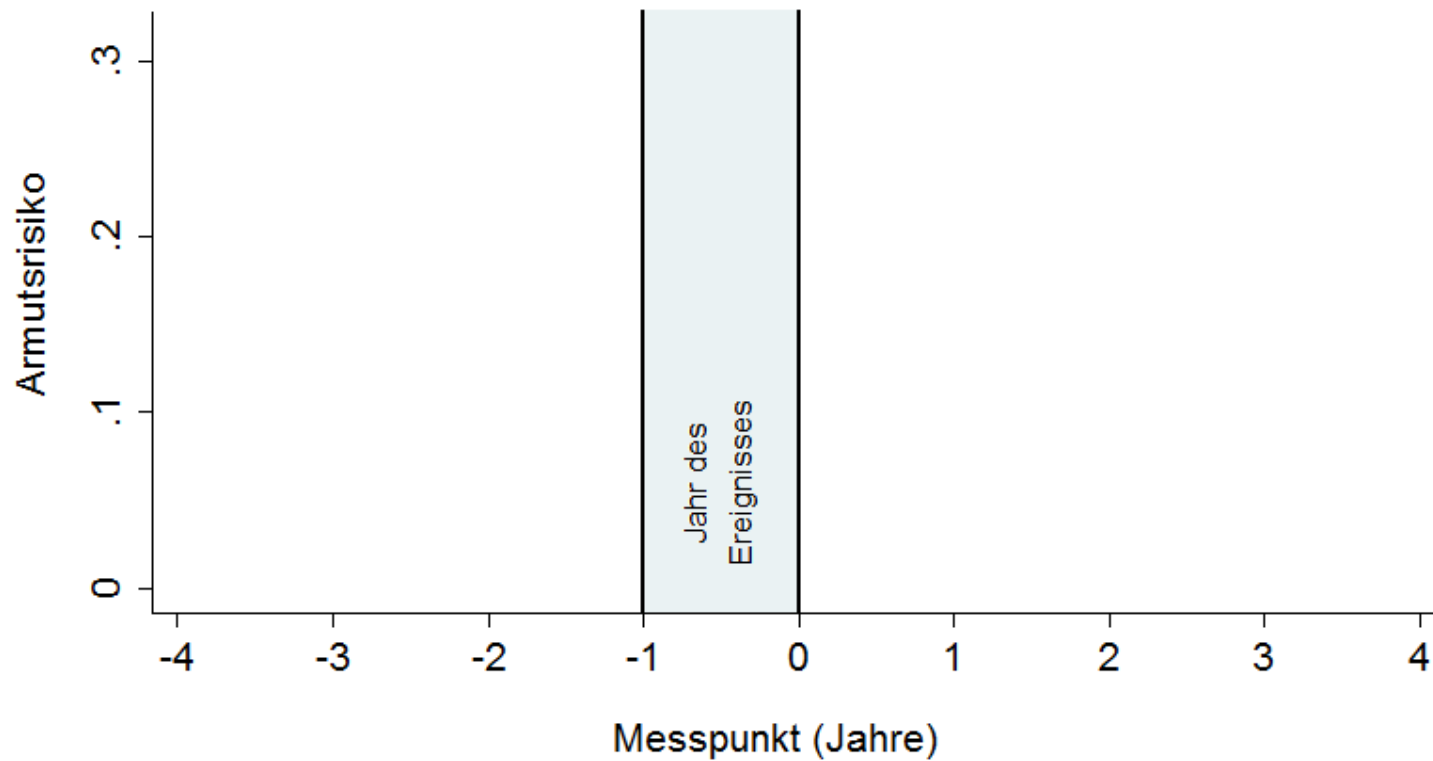


## *Vorteile von Paneldaten, insb SOEP: Illustrationen*

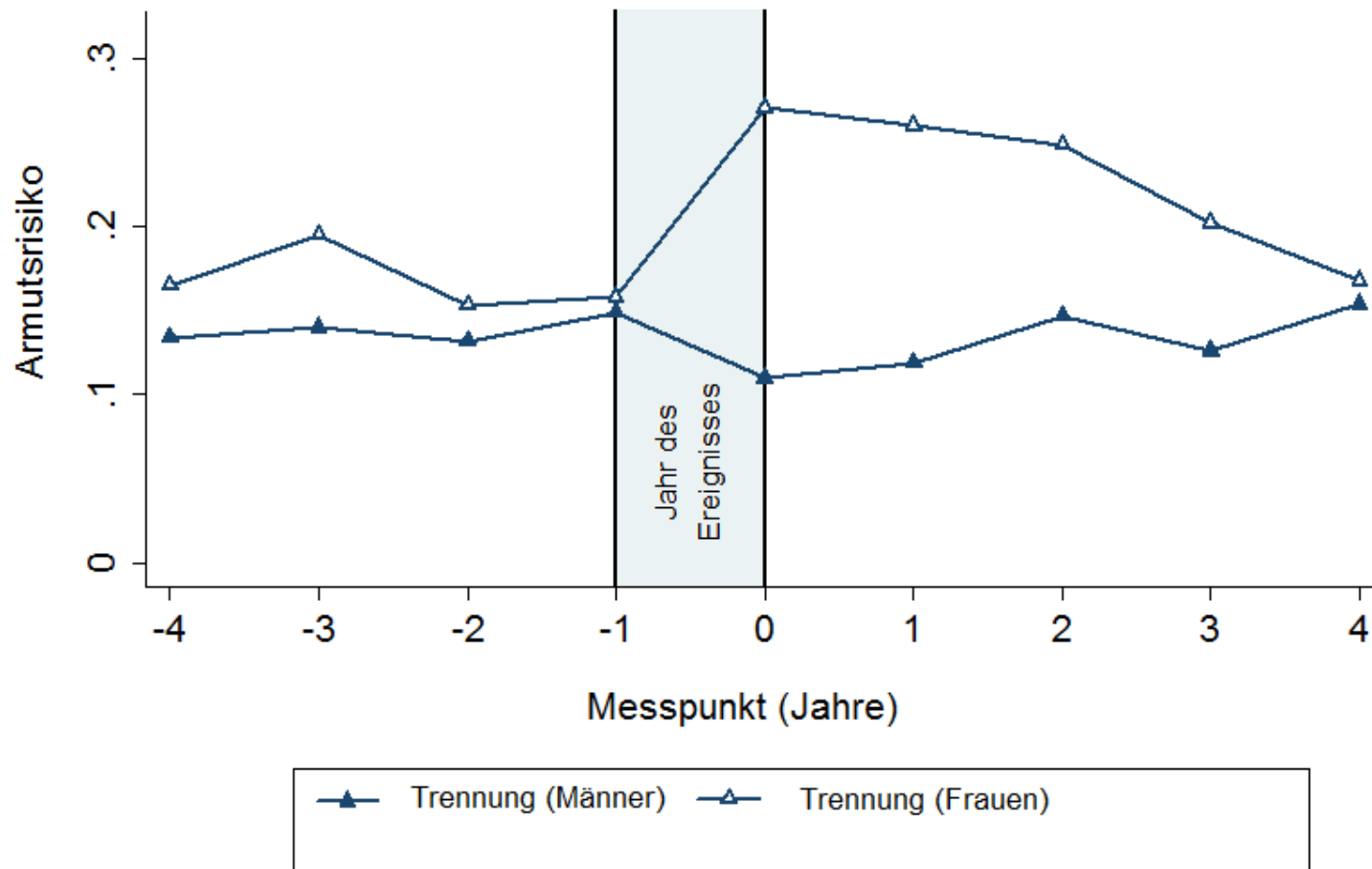
**Analysen zum Einfluß von Ereignissen: Ein weiteres Beispiel**

---

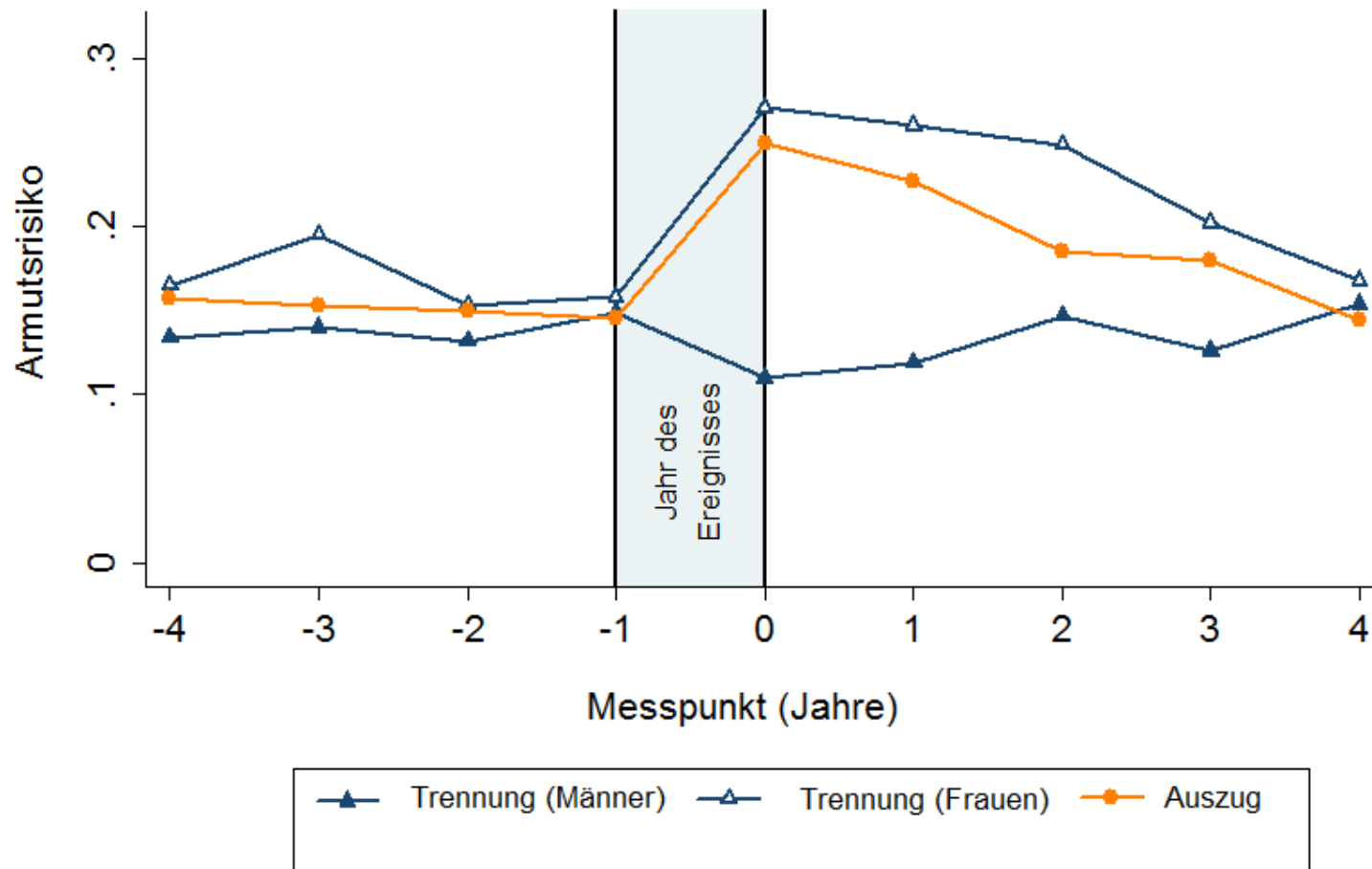
## Critical life events and poverty dynamics



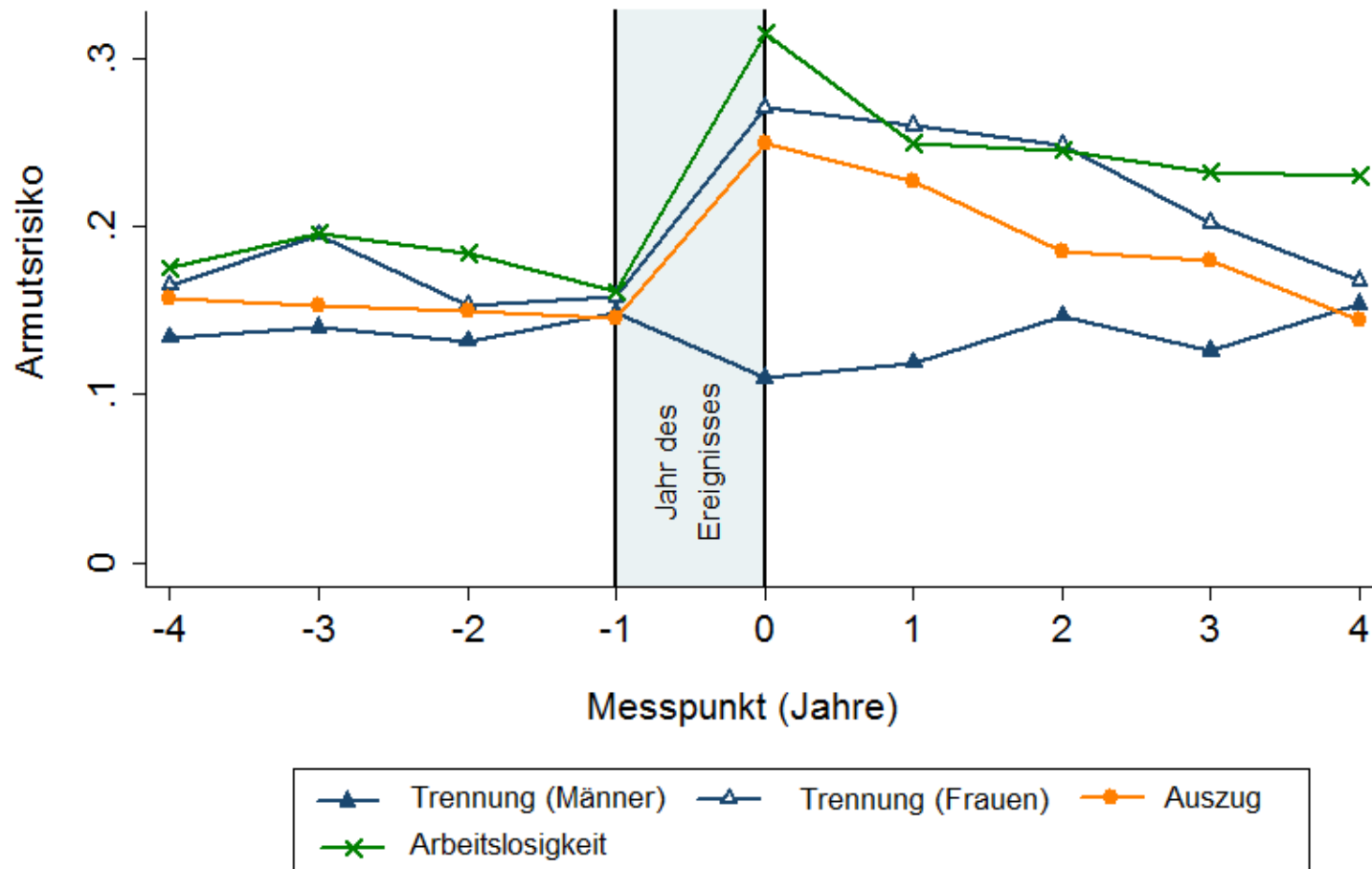
## Critical life events and poverty dynamics



## Critical life events and poverty dynamics

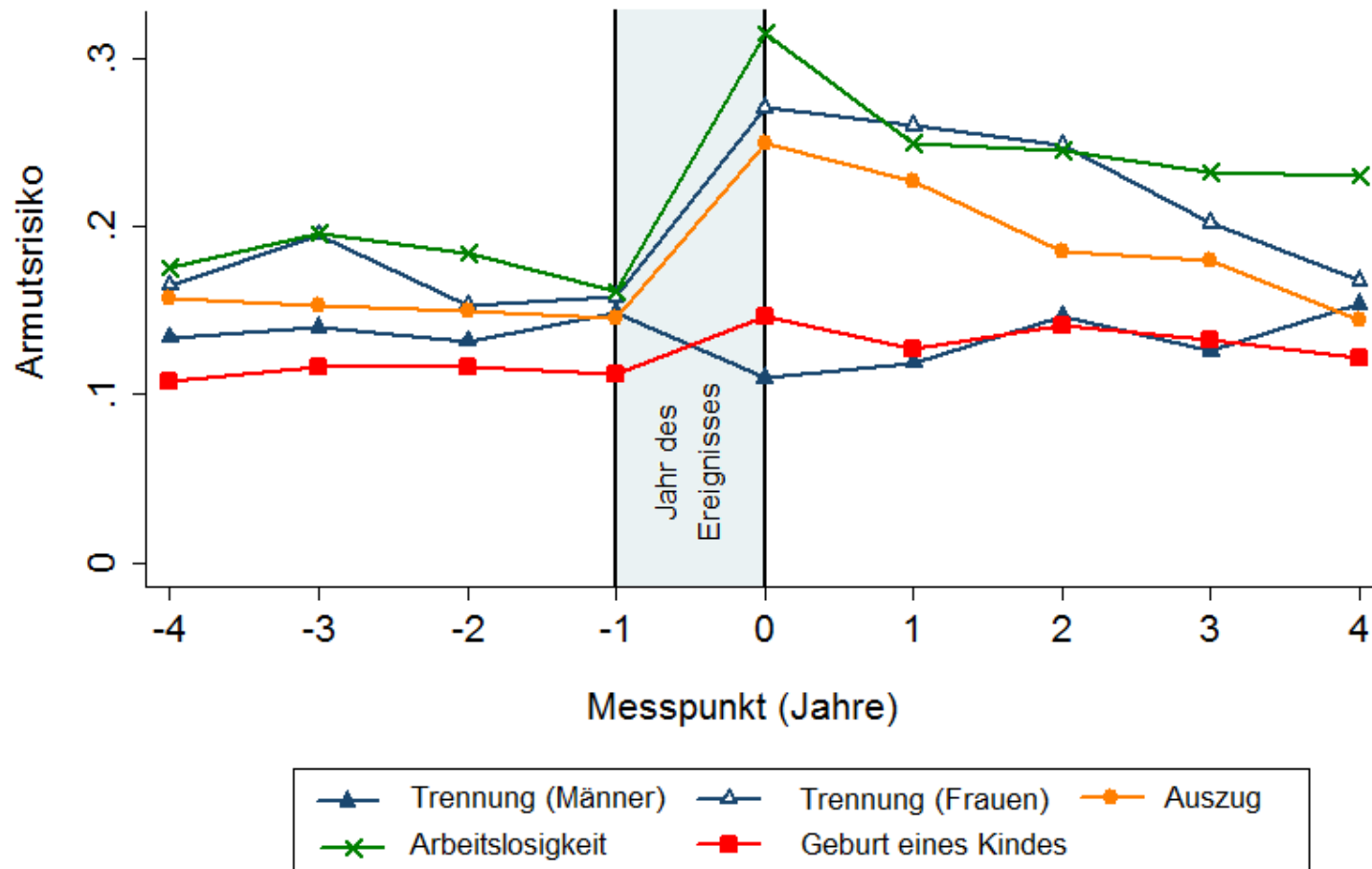


## Critical life events and poverty dynamics





## Critical life events and poverty dynamics



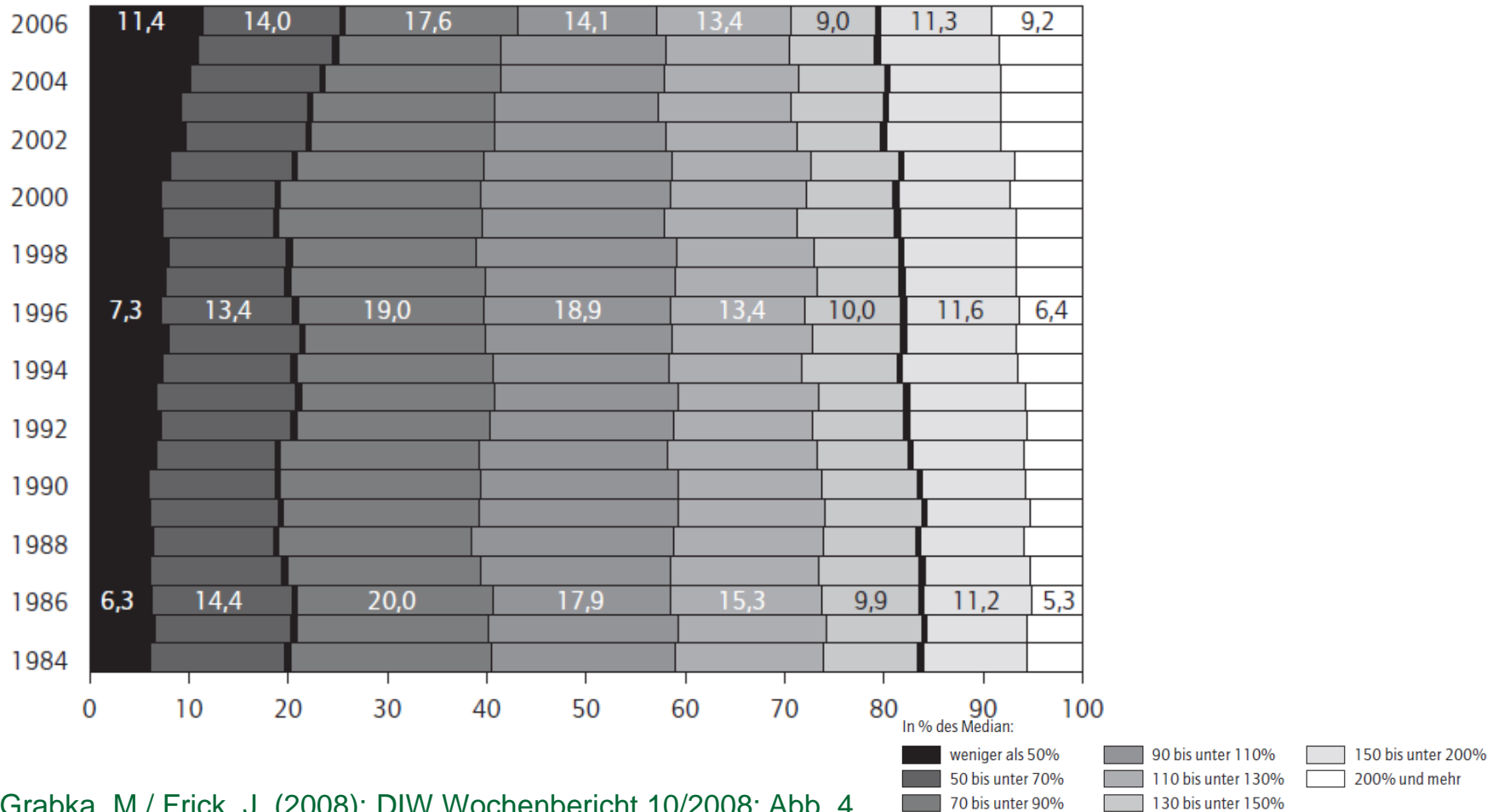
## *Vorteile von Paneldaten, insb SOEP: Illustrationen*

**Verteilungsanalysen: Ein Beispiel**

---

# Einkommenschichtung in Deutschland<sup>1</sup> 1984 bis 2006

In Prozent, Einkommensposition auf Basis des jahresspezifischen Median



Grabka, M./ Frick, J. (2008): DIW Wochenbericht 10/2008: Abb. 4.

## *Mögliche Nachteile von Paneldaten*

- ⇒ Verzerrungen aufgrund von Panelmortalität
  - ⇒ Paneleffekte
  - ⇒ Rigidität des Erhebungsinstrumentes
-

## *Überblick*

1. Was sind Panel-Daten?
  2. Was ist das SOEP?
    - a. Themen
    - b. Fragebögen
    - c. Zusammenfassung
  3. Datenzugang
-

# Was ist das SOEP?

- Repräsentative Wiederholungsbefragung privater Haushalte in Deutschland (seit 1984, aktuell ~12.000 HH).
- Theoriegeleiteter und interdisziplinären Mix von subjektiven und objektiven Daten
- Teil der sozial- und wirtschaftswissenschaftlichen Infrastruktur
- Der SOEP-Datensatz wird an Universitäten und unabhängigen Forschungseinrichtungen im In- und Ausland für Forschung und Lehre weitergegeben.

# Wer wird befragt?

- Mit allen Haushaltsmitgliedern ab 17 Jahren wird ein persönliches Interview durchgeführt.
- Es werden Proxy-Information zu den in den Haushalten lebenden Kindern erhoben.
- Weiterverfolgungskonzept:
  - Befragung der „Kinder“ ab dem Alter von 17 Jahren,
  - Personen, die in einen SOEP-Haushalt ziehen,
  - Personen, die einen neuen Haushalt gründen oder in einen anderen Haushalt ziehen.
- Regelmäßige Auffrischungstichproben, um Bevölkerungsveränderungen abbilden zu können.

# Geschichte der Studie

- Idee: während Forschungsaufenthalt von H.J. Krupp in den 60er Jahren an der University of Wisconsin, Madison
- 1981/1982: Beginn der Studie (Projektantrag)
- 1984: erste Erhebungswelle als Teilprojekt B-5 (Sozio-ökonomisches Panel) des Sonderforschungsbereichs 3 "Mikroanalytische Grundlagen der Gesellschaftspolitik"
- Seit 1989: Förderung gemäß Beschluss der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung



## Geschichte der Studie (Fortsetzung)

- 1994: Empfehlung des Wissenschaftsrates zur Aufnahme des SOEP in die Leibniz-Gemeinschaft (WGL)
- Seit 2003: institutionelle Förderung des SOEP als „unabhängige Abteilung“ im DIW Berlin durch die Bundes-Länder-Finanzierung als Serviceeinrichtung der WGL
- Mitarbeit im „Interdisziplinären Verbund Serviceeinrichtungen“ (IVS) der WGL

## *SOEP Basics*

### Was wird gemessen ?

- ⇒ Jährlich: Repräsentative Daten zu Personen, Haushalten, Familien
  - ⇒ Objektive und subjektive Maße
  - ⇒ Multidisziplinäre Anlage
-

# Zeitdimension

- *Frage zu einem Zeitpunkt (Gegenwart)*  
*z.B. derzeitiger Erwerbsstatus oder Zufriedenheit*
- *Einzelne retrospektive Frage zu Ereignissen (Vergangenheit)*  
*z.B. wie oft haben Sie Ihren Job in den letzten Jahren gewechselt?*
- *Retrospektiv erfragte „Ereignishistorie“ seit Alter 15 (Vergangenheit)*  
*z.B. Erwerbstätigkeit oder Familienstand*
- *Monatlicher Kalender zu Einkommen und Erwerbstätigkeit (Vergangenheit)*  
*z.B. Erwerbsstatus von Januar bis Dezember des letzten Jahres*
- *Fragen zu einer Zeitperiode (Vergangenheit)*  
*z.B. Änderung im Familien- oder Bildungsstand seit letztem Interview*
- *Fragen zu zukünftigen Erwartungen (Zukunft)*  
*z.B. Zufriedenheit in 5 Jahren oder Erwartungen bez. Beruf*

## *Überblick*

1. Was sind Panel-Daten?
  2. Was ist das SOEP?
    - a. Themen
    - b. Fragebögen
    - c. Zusammenfassung
  3. Datenzugang
-

# *Inhalt*

## Zentrale Fragen

- ⇒ Demografie
  - ⇒ Ausbildung und Qualifikation
  - ⇒ Arbeitsmarktposition
  - ⇒ Einkommen
  - ⇒ Wohnsituation
  - ⇒ Werte, Ziele und Partizipation
  - ⇒ Lebenszufriedenheit
-

**In jährlich wechselnden *Schwerpunktt*hemen werden darüber hinaus Informationen bereitgestellt, z.B. über:**

- Soziale Netzwerke (1991, 1996, 2001, 2006, 2011)
- Freizeitaktivitäten (lange Fragebatterie 1990, 1995, 1998, 2003, 2008, 2013)
- Weiterbildung und Qualifizierung (1989, 1993, 2000, 2004, 2008)
- Vermögen (2002, 2007, 2012)
- Umweltverhalten (1998, 2003, 2015)

## **Herausforderung: Messung des „Unbeobachtbaren“**

**... und die Erweiterung der Meßinstrumente jenseits  
standardisierter Fragebögen**

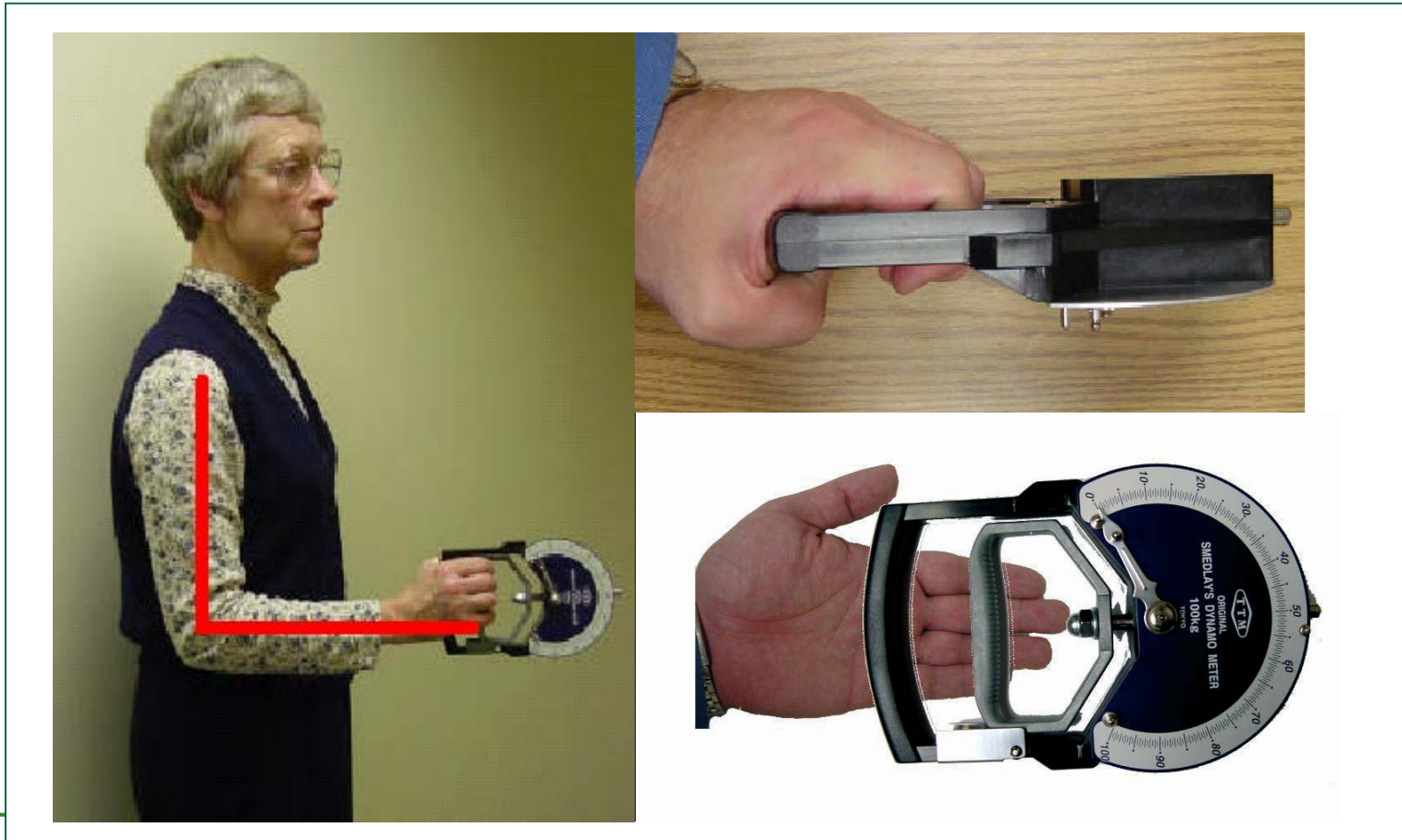
- **Objektive Messung physischer Eigenschaften**
  - **Messung kognitiver Fähigkeiten**
  - **Messung von Persönlichkeitseigenschaften**
-

# Physis / Gesundheitseigenschaften

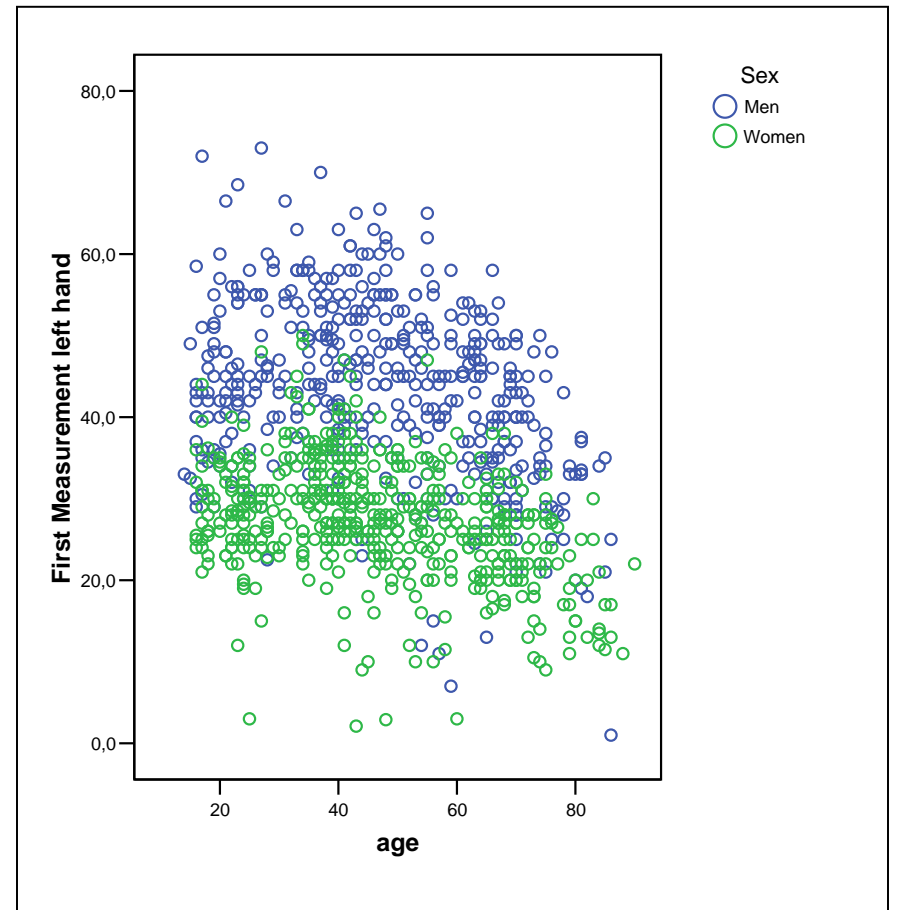
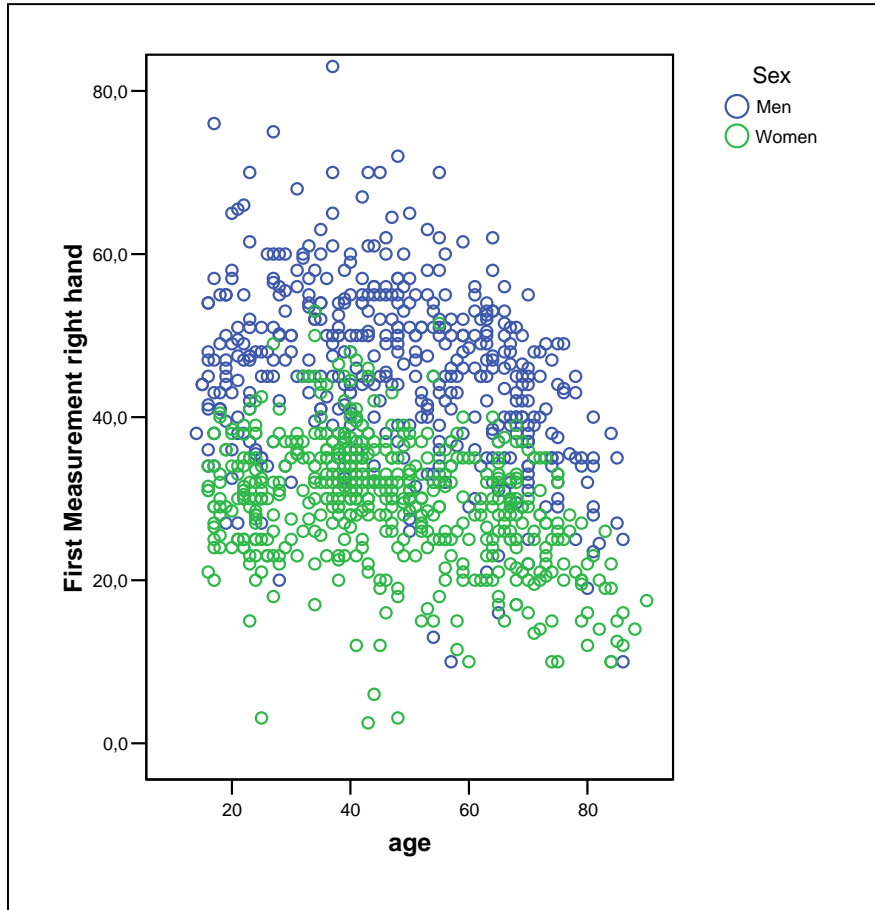
- ✓ **Einfache Abfrage seit 1984 (Selbsteinschätzung, Behinderung, Krankheitstage, Krankenhausaufenthalte, Arztbesuche)**
  - ✓ **Verbesserte Maße seit 2002**
    - ✓ Physische und mentale Gesundheitsindikatoren SF12 (seit 2000 in 2-Jahres-Wellen)
    - ✓ Körpergröße und Gewicht (seit 2000 in 2-Jahres-Wellen)
    - ✓ Rauchverhalten (2001, 2002, 2004, 2006, 2008, 2010, 2012)
    - ✓ Alkoholkonsum (2006, 2008, 2010)
    - ✓ Greifkrafttest (2006, 2008)
-



# Measuring Grip Strength in SOEP



# Grip Strength by Age and Sex



## Kognitive Fähigkeiten (seit 2006)

- 2 Kurztests (90 sek.) für Erwachsene (2006)
  - Wiederholung in 2012  
(plus dritter Kurztest Wörter kennen)
- Erweiterte Tests (30 min) für Jugendliche (DJ)

**Paper:**

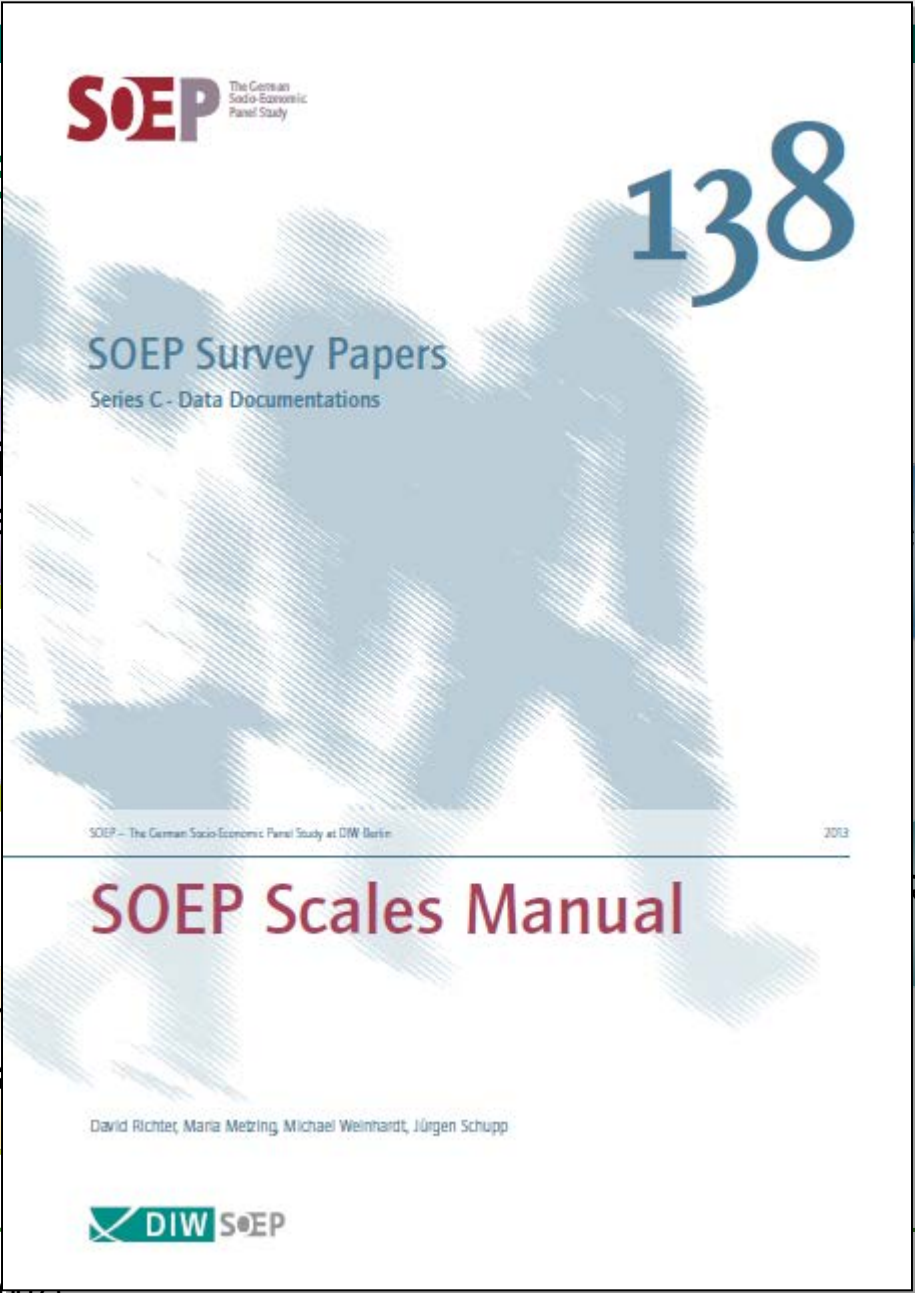
*Anger & Heineck (2009): Do Smart Parents Raise Smart Children?: The Intergenerational Transmission of Cognitive Abilities, SOEPpapers #156*

# Psychologis

Personalit  
„Big Five

Trust

Time Preferen  
Impulsiveness, Pa



ous of Control

Risk Aversion

## *Überblick*

1. Was sind Panel-Daten?
  2. Was ist das SOEP?
    - a. Themen
    - b. Erhebung
    - c. Zusammenfassung
  3. Datenzugang
-

## SOEP-Erhebungsinstrumente

(\$ = A-Z, BA-BZ, A=1984, B=1985, ...)

### Standardinstrumente

- ⇒ Haushaltsfragebogen (→ \$H – Files)
- ⇒ Personenfragebogen für jedes Haushaltsmitglied ab 16 bzw. 17 Jahre (→ \$P - Files)

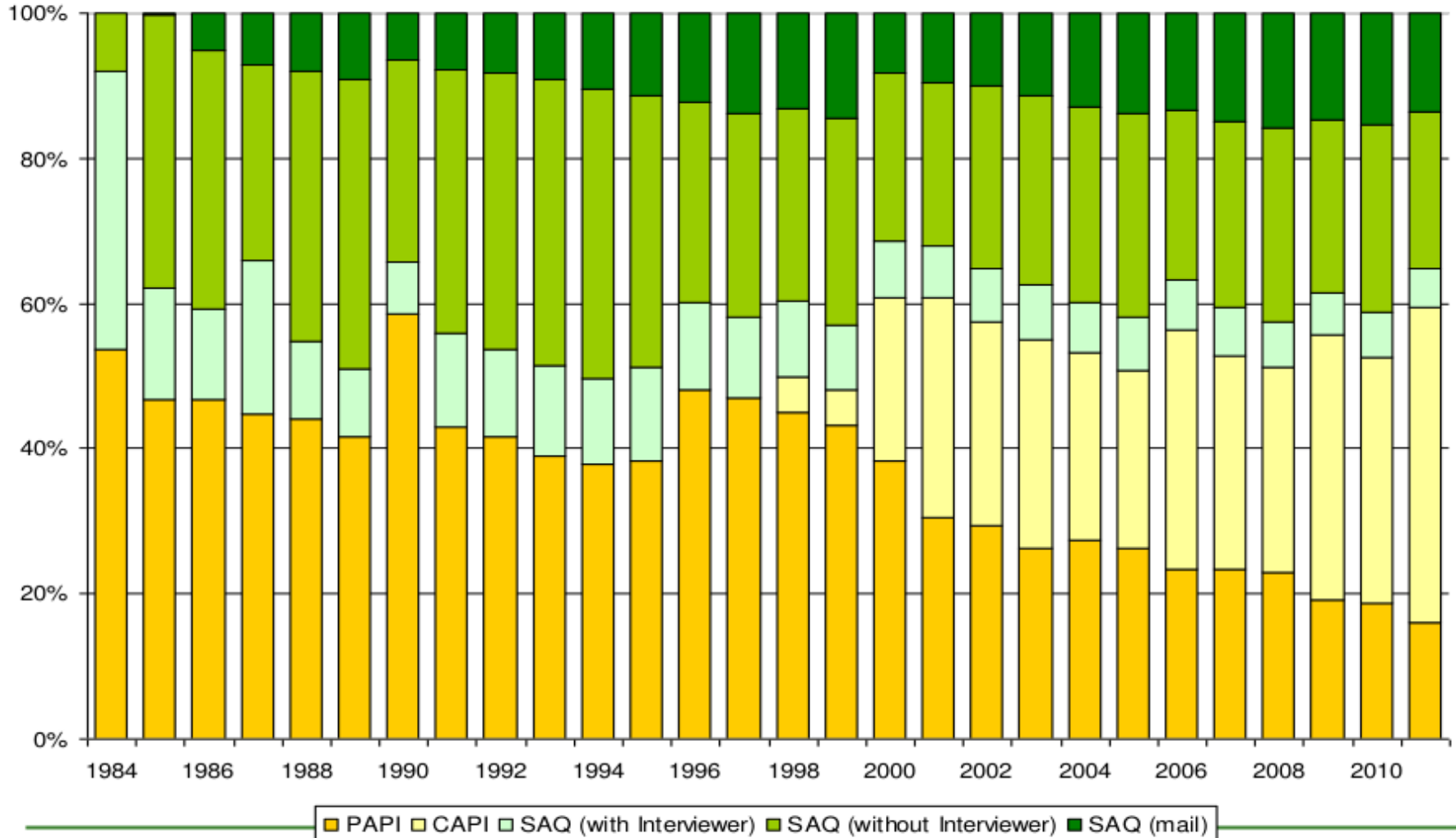
### Biografieinterviews

- ⇒ Zusatzfragebogen Lebenslauf (Erstbefragte 18+) (→ Biosoc, Bio...)
- ⇒ Zusatzfragebogen Jugend (Erstbefragte 16/17 seit 2000) (→ Bioage 17, \$page17)

### Teilgruppenspezifische Instrumente

- ⇒ Zusatzfragebögen Mutter & Kind (Neugeborene seit 2003, 2-3-jährige seit 2005, 5-6-jährige seit 2008, 7-8-jährige seit 2010) (→ Bioage1, Bioage...)
  - ⇒ Zusatzfragebogen Lücke (Temporäre Ausfälle) (→ \$Pluecke)
  - ⇒ Zusatzfragebogen „Die verstorbene Person“ (seit 2009)
-

# Entwicklung der Erhebungsmethodik



# Entwicklung der Stichprobengröße

## ■ *Personen steigen aus bei:*

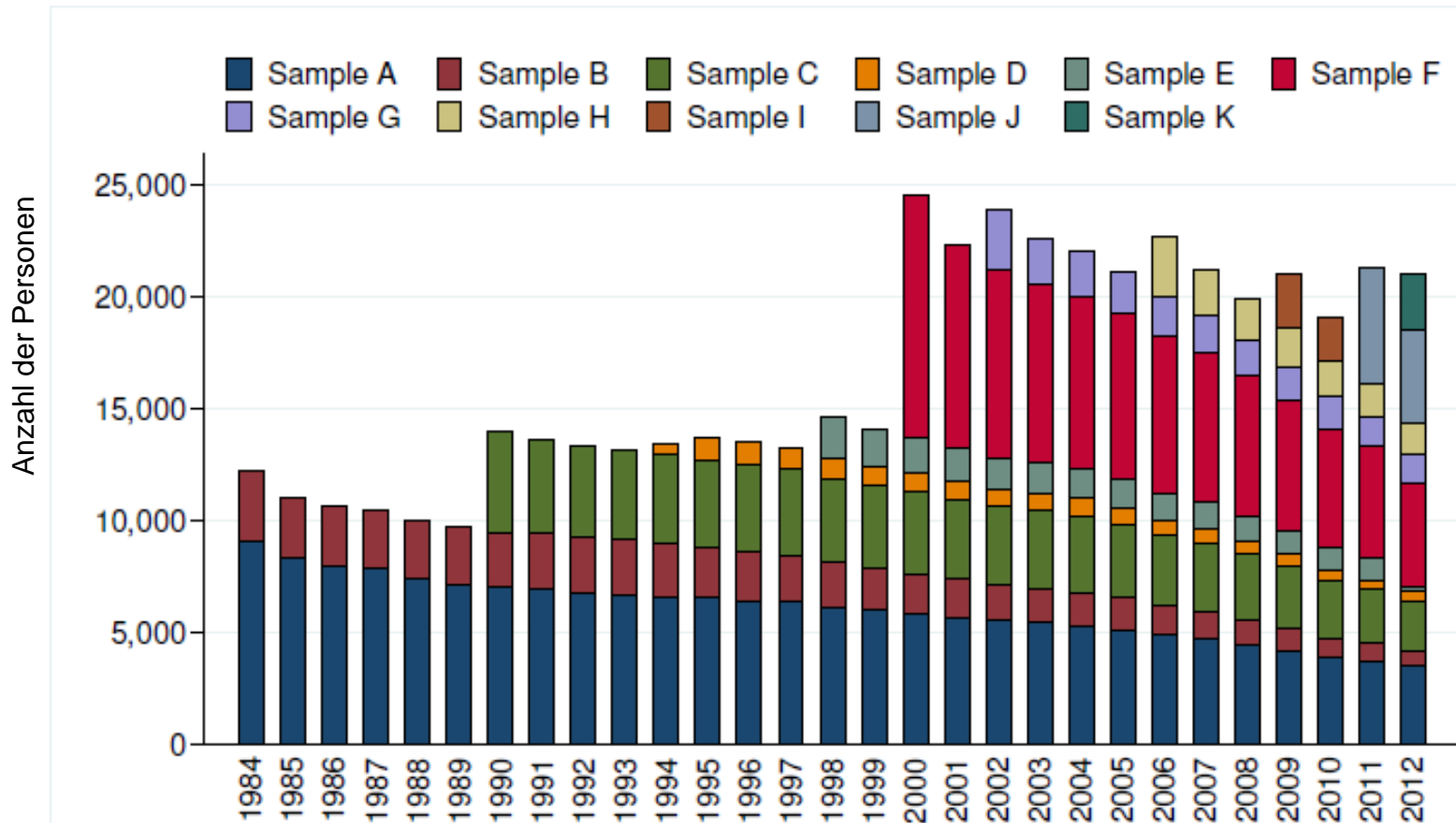
- Tod
- Umzug ins Ausland
- Verweigerung
- Kontaktabbruch

## ■ *Personen steigen ein bei:*

- Geburt
  - Einzug in einen SOEP-Haushalt (aus Deutschland oder dem Ausland)
  - Im Alter von 17 Jahren (vorher Kinder in SOEP-Haushalt ohne persönlichen Fragebogen), Kalenderjahr, in dem Personen 17 Jahre alt werden
  - Haushaltsneugründungen von mindestens einem „alten“ SOEP-Haushaltsmitglied
  - Auffrischung
-



# Querschnittsperspektive: Erfolgreich interviewte Personen

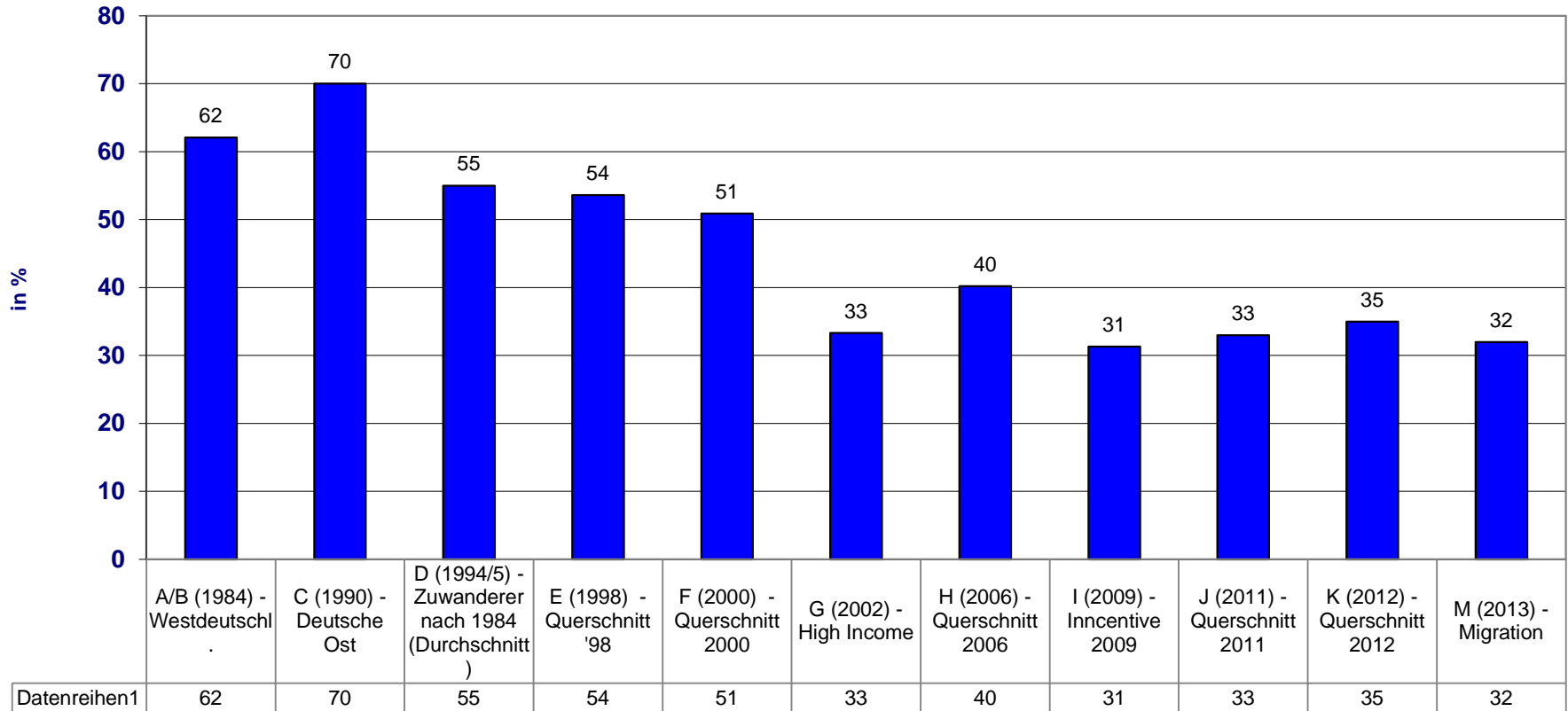


# Stichproben

→ Siehe auch Variable `psample`  
im Datensatz `ppfad`

- **1984: Sample A** Deutsche (West)
- **1984: Sample B** Ausländer (West)
- **1990: Sample C** Deutsche (Ost)
- **1994: Sample D** Zuwanderer (West), 1994/1995
- **(1998: Sample E** Innovation (CAPI))
- **2000: Sample F** Ergänzung (Verdoppelung der Stichprobe)
- **2002: Sample G** Hocheinkommensbezieher
- **2006: Sample H** Ergänzung
- **(2009: Sample I** Innovation (Incentivierung))
- **2011: Sample J** Ergänzung
- **2012: Sample K** Ergänzung
- **2013: Sample M** Migrationsstichprobe

# Ausschöpfung (Haushalte, erste Welle)



Source: SOEP.

## *Überblick*

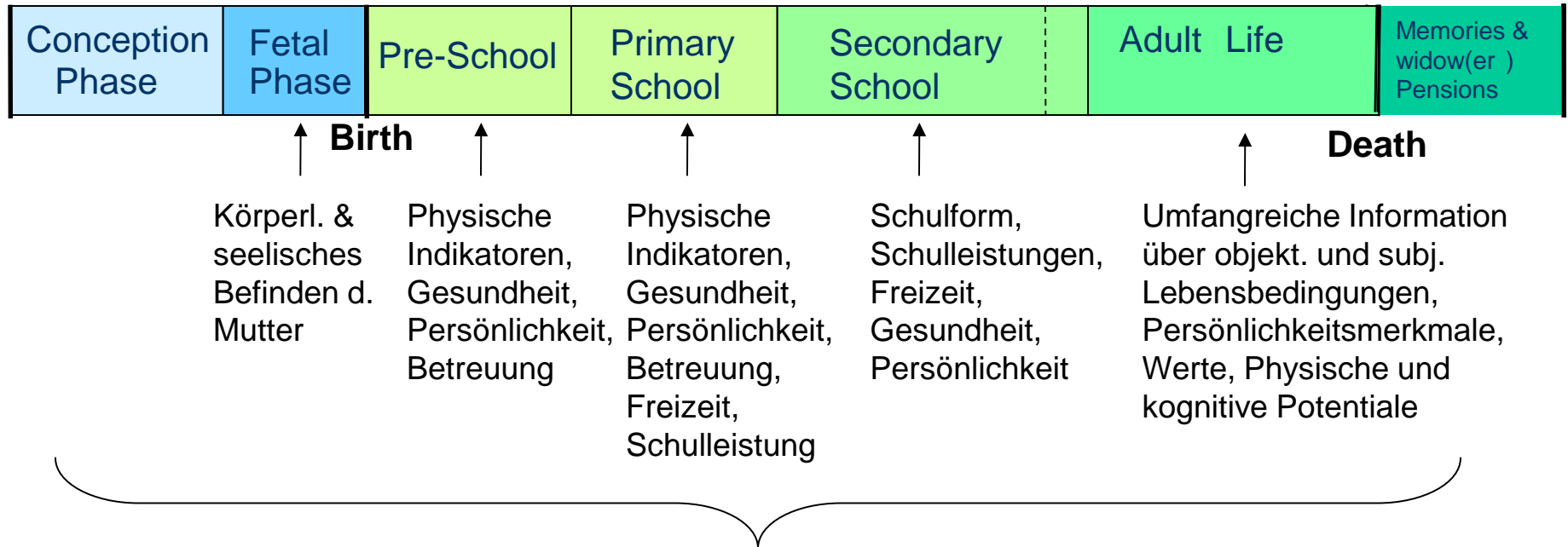
1. Was sind Panel-Daten?
  2. Was ist das SOEP?
    - a. Themen
    - b. Erhebung
    - c. Zusammenfassung
  3. Datenzugang
-

# Inhalte des SOEP

## Lebensverlauf und verfügbare Informationen

Ojekt. und Subj.  
Lebensbedingungen,  
Persönlichkeitsmerkmale,  
Werte, Physische und  
kognitive Potentiale

Werte, Beziehungen (Eltern),  
Freizeit, Schulleistungen,  
Ziele



Informationen über Lebensbedingungen der Eltern (Partner) / Haushaltskontext  
Regionaler Kontext

## *Überblick*

1. Was sind Panel-Daten?
  2. Was ist das SOEP?
    - a. Themen
    - b. Erhebung
    - c. Zusammenfassung
  3. Datenzugang
-

# Datenzugang

- Der anonymisierte Mikrodatensatz des SOEP wird Universitäten und anderen Forschungseinrichtungen für **Forschung und Lehre kostenfrei** zur Verfügung gestellt.
- Für die **Datennutzung** ist aus Gründen des Datenschutzes der Abschluss eines **Datenweitergabevertrages** mit dem DIW Berlin notwendig.
- Nach Abschluss können die Daten über einen personalisierten Link mit individuellem Passwort verschlüsselt herunter geladen werden.

# Datenzugang (Fortsetzung)

- Der Antrag auf Datennutzung kann unter:  
[http://www.diw.de/de/diw\\_02.c.222836.de/datenzugang.html](http://www.diw.de/de/diw_02.c.222836.de/datenzugang.html)  
heruntergeladen werden.

SOEPhotline

Ansprechpartnerin: Michaela Engelmann

Tel.: +49 30 89789-292, Fax.: +49 30 89789-109

soepmail@diw.de



# Regional- daten

Ebene	Verfügbar ab:	Datenzugang	Datenschutzrechtliche Zugangs-voraussetzung
Bundesländer	1984	SOEP-Standard-Datensatz (Scientific Use File)	Datenweitergabevertrag
Gemeindegrößenklassen (z. B. nach Boustedt)	1984	SOEP-Standard-Datensatz mit speziellem Passwort	erweiterter Datenweitergabevertrag zur Nutzung der Gemeindegrößenklassendaten + Datenschutzkonzept
Raumordnungsregionen (Geocodes)	1985	SOEP-Standard-Datensatz plus SOEP-Geocode-CD-ROM	erweiterter Datenweitergabevertrag zur Nutzung der Geocodes + erweitertes Datenschutzkonzept
Kreiskennziffern	1985	SOEPremote: Fernzugang zu den SOEP-Daten zur Nutzung von Regionaldaten auf Kreisebene oder am Forschungsdatenzentrum SOEP	erweiterter Datenweitergabevertrag zur Nutzung von SOEPremote + Antrag zur Nutzung von SOEPremote
Gemeindekennziffern Postleitzahlen Microm Nachbarschaftsdaten	2000 1993 2000	Nutzung der Daten ausschließlich im Forschungsdatenzentrum des SOEP	nur nach persönlicher Absprache im Rahmen des SOEP-Gästeprogramms

Vielen Dank für Ihre  
Aufmerksamkeit.  
Fragen?

---

# Cognitive Abilities (since 2006)

- 2 Ultra short tests (90 sec) for adult respondents (replication in 2012 currently in planning stage)
- Full ability measures (about 30 min) for teenagers, only

**Paper:**

*Anger & Heineck (2009): Do Smart Parents Raise Smart Children?: The Intergenerational Transmission of Cognitive Abilities, SOEPpapers #156*

---

# Measuring Adults' Cognitive Abilities

---

## Measures of cognitive ability:

2 ultrashort tests (Lang, 2005)

---

### *Symbol Correspondence test* (fluid mechanics):

developed after the Symbol-digit-modalities-test (SDMT) by Smith (1973, 1995)

Respondents have 90 seconds to assign as many correct signs as possible to the consecutively displayed digits, while the appropriate assignment code is visible to them.

This test measures fluid mechanics, which refers to general and largely innate ability.

---

### *Word fluency test* (crystallized pragmatics):

developed after the Animal-naming-task (Lindenberger and Baltes, 1995)

Respondents have 90 seconds to name as many different animals as possible.

This test measures crystallized pragmatics, which includes all cognitive abilities that consist of accumulated knowledge and skills acquired in the past.

• ) + † 7 V ( ¯ †  
1 2 3 4 5 6 7 8 9

# Measurement of Teenagers' General Cognitive Abilities (DJ) (since 2006)

## *Objectives*

- *Gain insight in skill formation processes among young people (skill acquisition)*
- *Contribution to interdisciplinary research efforts on social disparities in educational attainment and labor market success*

## *Intelligence-Structure-Test*

- *I-S-T 2000 R (Amthauer et al. 2001)*

## *Dimensions*

- *Verbal potential (analogies)*
- *Numerical potentials (number sequences)*
- *Figural potentials (matrices/pictorial material)*
- *Reasoning (= sum score)*

## *Survey*

- *Test take about 30 minutes*
- *Participation rate about 90%*
- *Additional incentive: stop-watch*

### **Paper:**

*Solga, H.; Stern, E., B. v. Rosenblatt; Schupp, J.; G G. Wagner (2006): The Measurement and Importance of General Reasoning Potentials in Schools and Labor Markets. DIW Research Notes 2006-10. DIW Berlin.*