

---

# SOEP Innovation Sample

David Richter

Lausanne, October 7, 2014

1. Scope & Structure of SOEP-IS
2. Data Collection 2012
3. Data Collection 2013 & 2014  
Subjective Well Being -Triangulation in SOEP-IS
4. Data Collection 2015  
Saliva Collection to Assess Epigenetics
5. Call for Proposals for Data Collection 2015

- Representative sample of private households.
- Adequate sample size for high-quality data analysis.
- Longitudinal data from existing SOEP subsamples.
- Innovative procedures that go beyond the classic survey components (IAT, DRM, experiments).
- Tool to test new survey methodologies beyond the limits of the SOEP pretests.
- Cross-references and linkages to the topics and content of the main SOEP survey.

- Two existing subsamples of the SOEP:
  - Sample E – first surveyed in 1998, originally 1,056 households.
  - Sample I – first surveyed in 2009, originally 1,531 households.
  - existing longitudinal data will expand the potential for analysis of innovative questions.
- Supplementary Sample in 2012: 1,000 hh.
- Supplementary Sample in 2013: 1,000 hh.
- Supplementary Sample in 2014: 1,000 hh.
- Total Sample in 2014: > 5,500 respondents in > 3,500 hh.

- Integrated questionnaire (household, all adult household members, biography, information on children).
- Set of core questions, 45 % of the main SOEP
- Additional topics are added, i.e., sets of questions that are asked only at specific intervals.
- The inclusion additional topics is oriented toward the rotation scheme of the main SOEP as well as toward the innovative modules in the particular survey wave.
- Overview of innovation modules and additional topics from SOEP-Core in the questionnaires is provided on our website (<http://www.diw.de/soep-is>).

## Evaluation of feasibility by SOEP survey management:

- If too burdensome, not feasible, or if other data sources are available → rejection.
- If very short and related to wave-specific focus of SOEP-IS → acceptance without review or SC decision.
- If the proposal has been externally reviewed and is of general importance for SOEP (e.g., comparison of national cohort panel and SOEP) → acceptance without review or SC decision.
- All other proposals → decision after review by expert in the respective field.
- SC decides on the main experimental issues and innovativeness of the proposal.

- All SOEP data are freely available to **all** researchers (*scientific open access*).
- This also applies to:
  - all of the different sets of questions and experiments submitted by external applicants.
  - data collected with funding from external sources.
- Data will be made publicly available **no later than one year** after the SOEP-IS applicant has received the data.

- Data available to user community since end of April 2014.
  - Main questions and additional topics of 2012.
  - User-friendly generated SOEP-variables of 2012.
  - Including innovative modules of 2011 (Gender-IAT, justice assessment, loneliness – 12 month embargo ended).
- Data are provided in long-format, only.
- New challenges for user-friendly documentation.
  - Questionnaires with links to variables and data sets.
  - Codebooks with additional information for all data sets.
  - Full inclusion in new SOEP info.
  - CAPI-screenshots are provided on request.



**Q53** Thinking back on the last four weeks, please state how often you have experienced each of the following feelings very rarely, rarely, occasionally, often, or very often. How often have you felt...

|                           | No answer                                      | Very rarely | Rarely | Occa-<br>sionally | Often | Very often |
|---------------------------|--|-------------|--------|-------------------|-------|------------|
| ...angry?                 | -1   | 1           | 2      | 3                 | 4     | 5          |
| ...worried?               | -1   | 1           | 2      | 3                 | 4     | 5          |
| ...happy?                 | -1   | 1           | 2      | 3                 | 4     | 5          |
| ...sad?                   | -1   | 1           | 2      | 3                 | 4     | 5          |
| Q53:pgef1    p    plh0184 | Frequency of being angry in the last 4 weeks   |             |        |                   |       |            |
| Q53:pgef2    p    plh0185 | Frequency of being worried in the last 4 weeks |             |        |                   |       |            |
| Q53:pgef3    p    plh0186 | Frequency of being happy in the last 4 weeks   |             |        |                   |       |            |
| Q53:pgef4    p    plh0187 | Frequency of being sad in the last 4 weeks     |             |        |                   |       |            |

SOEP-IS 2012—KID: Pooled Dataset on Children

## 4 Variables on day care arrangements (var names starting with KD\_)

Variables KD\_PUBL - KD\_AMTP summarize information about day care arrangements. With regard to wave O to BB (1998-2011), this information is taken from the SOEP \$KID-LONG data file whereas no information is available for the survey years 2011 and 2012 of SOEP-IS. More detailed information can be found in SOEP KIDLONG documentation

### kd\_publ – Public Institution

|    |  |      |
|----|--|------|
| 1  | [1] Public Institution                             | 375  |
| 2  | [2] Public School (2011)                           | 68   |
| 3  | [3] Public Provider (2011)                         | 13   |
| -1 | [-1] No Answer                                     | 0    |
| -2 | [-2] Does not apply                                | 142  |
| -3 | [-3] Answer improbable                             | 0    |
| -4 | [-4] inadmissible multiple answer                  | 0    |
| -5 | [-5] not contained in questionnaire                | 3893 |
| -6 | [-6] questionnaire version with modified filtering | 0    |

Variable KD\_PUBL summarizes information collected in the years 2002, 2005, 2007 and 2011. This information KS\_PUBL is not available for the survey years 2011 and 2012 of SOEP-IS.

SOEPinfo v.2 beta ▾ Studies Topics Publications

SOEPinfo v.2 beta



DDI on Rails  
Understanding Data

DDI on Rails, designed and built by Marcel Hebing, proudly sponsored by the German Socio-economic Panel (SOEP).

# Search

Study

**SOEP Core study**

x

Analysis unit

**individual level**

x

Conceptual dataset

Original data (net)

15

Generated data

2

Period

**2012**

x

17 results.

**Current Life Satisfaction [bcj101]**

bcj101 ↑

Variable in study: **soep-core** | dataset: **bcpage17** | period: **2012** | analysis unit: **p** | concept: **Item\_5423**

**Satisfaction With Health [bcp0101]**

bcp0101 ↑

Variable in study: **soep-core** | dataset: **bcp** | period: **2012** | analysis unit: **p** | concept: **pzuf01**

**satisfaction with sleep [bcp0102]**

bcp0102 ↑

Variable in study: **soep-core** | dataset: **bcp** | period: **2012** | analysis unit: **p** | concept: **pzuf20**

**Satisfaction With Work [bcp0103]**

bcp0103 ↑

Variable in study: **soep-core** | dataset: **bcp** | period: **2012** | analysis unit: **p** | concept: **pzuf02**

**Satisfaction With Housework [bcp0104]**

bcp0104 ↑

Variable in study: **soep-core** | dataset: **bcp** | period: **2012** | analysis unit: **p** | concept: **pzuf03**

**Satisfaction With Household Income [bcp0105]**

bcp0105 ↑

Variable in study: **soep-core** | dataset: **bcp** | period: **2012** | analysis unit: **p** | concept: **pzuf04**

**Satisfaction With Personal Income [bcp0106]**

bcp0106 ↑

# Demo basket SOEP-Core: R-Script

## Demo basket SOEP-Core

Basket overview

Variables by concepts

Available datasets

Available variables

Script: SoepStata

Script: SoepSPSS

Script: SoepR

List of variables as CSV

## Settings

### Input path

data/

### Output path

out/

### Analysis unit

Individual

### Balanced sample design

balanced

### Age group

All Sample Members

### Private households

Private households or

Rescript

```
/* ### LOCAL VARIABLE ### */
```

```
global MY_PATH_IN "data/"
global MY_PATH_OUT "out/"
global MY_FILE_OUT ${MY_OUT_PATH}new.dta
global MY_FILE_LOG ${MY_OUT_PATH}new.log
```

```
log using "${MY_FILE_LOG}", text replace
set more off
```

```
/* ### NOT PROCESSED ###
--> gpost: gp5501e gp5502e
*/
```

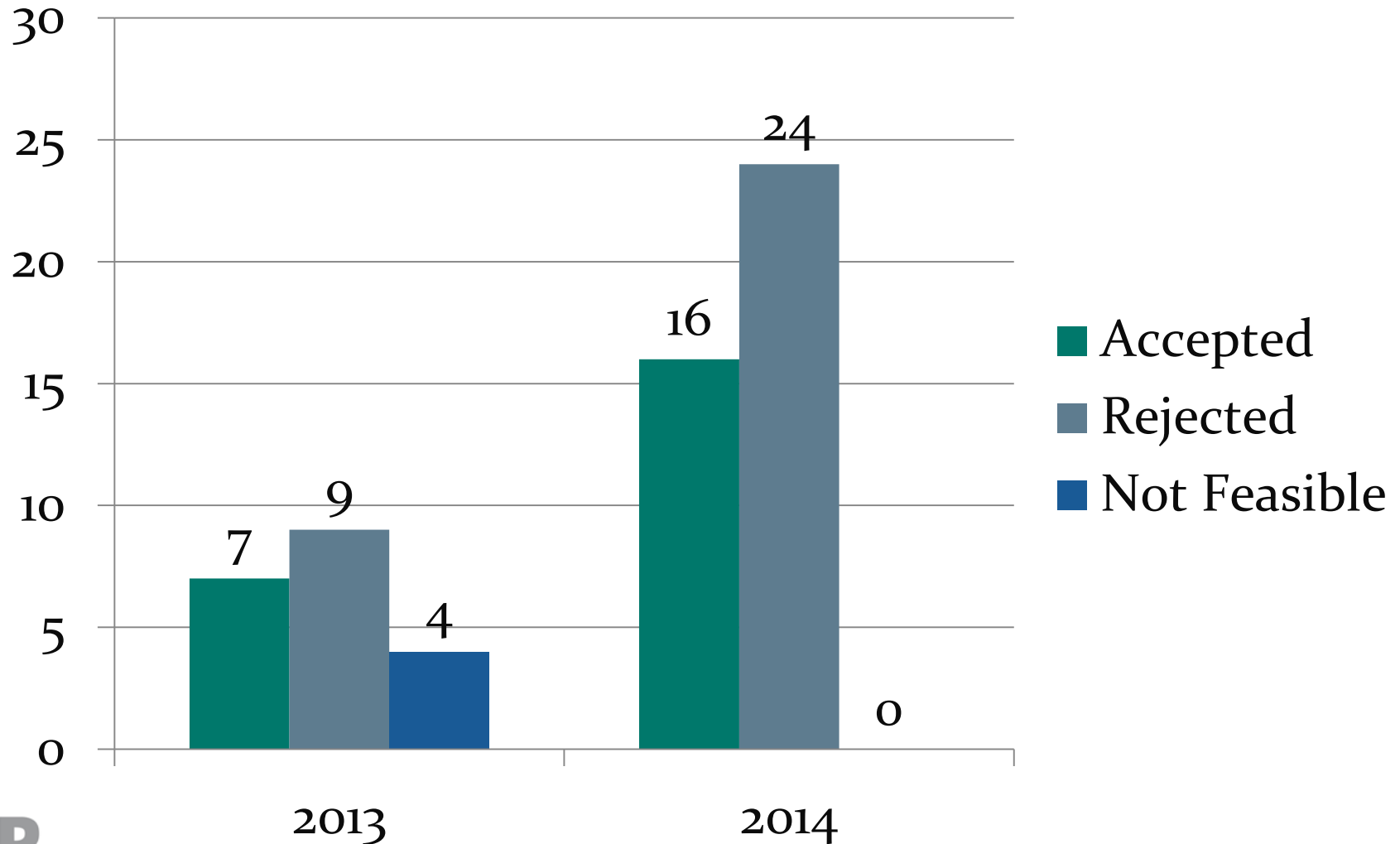
```
/* ### LOAD [H|P]PFAD ### */
```

```
use hhnr persnr sex gebjahr psample ahnr anetto apop
bhnr bnetto bpop chnr cnetto cpop dhnr dnetto dpop
ehnr enetto epop fhnr fnetto fpop ghnr gnetto gpop
hhnr hnetto hpop ihnr inetto ipop jhnr jnetto jpop
khnr knetto kpop lhnr lnetto lpop mhnr mnetto mpop
nhnr nnetto npop ohnr onetto opop phnr pnetto ppop
qhnr qnetto qpop rhnr rnetto rpop shnr snetto spop
thnr tnetto tpop uhnr unetto upop vhnr vnetto vpop
whnr wnetto wpop xhnr xnetto xpop yhnr ynetto ypop
zhnr znetto zpop bahnr banetto bapop bbhnr bnetto
```

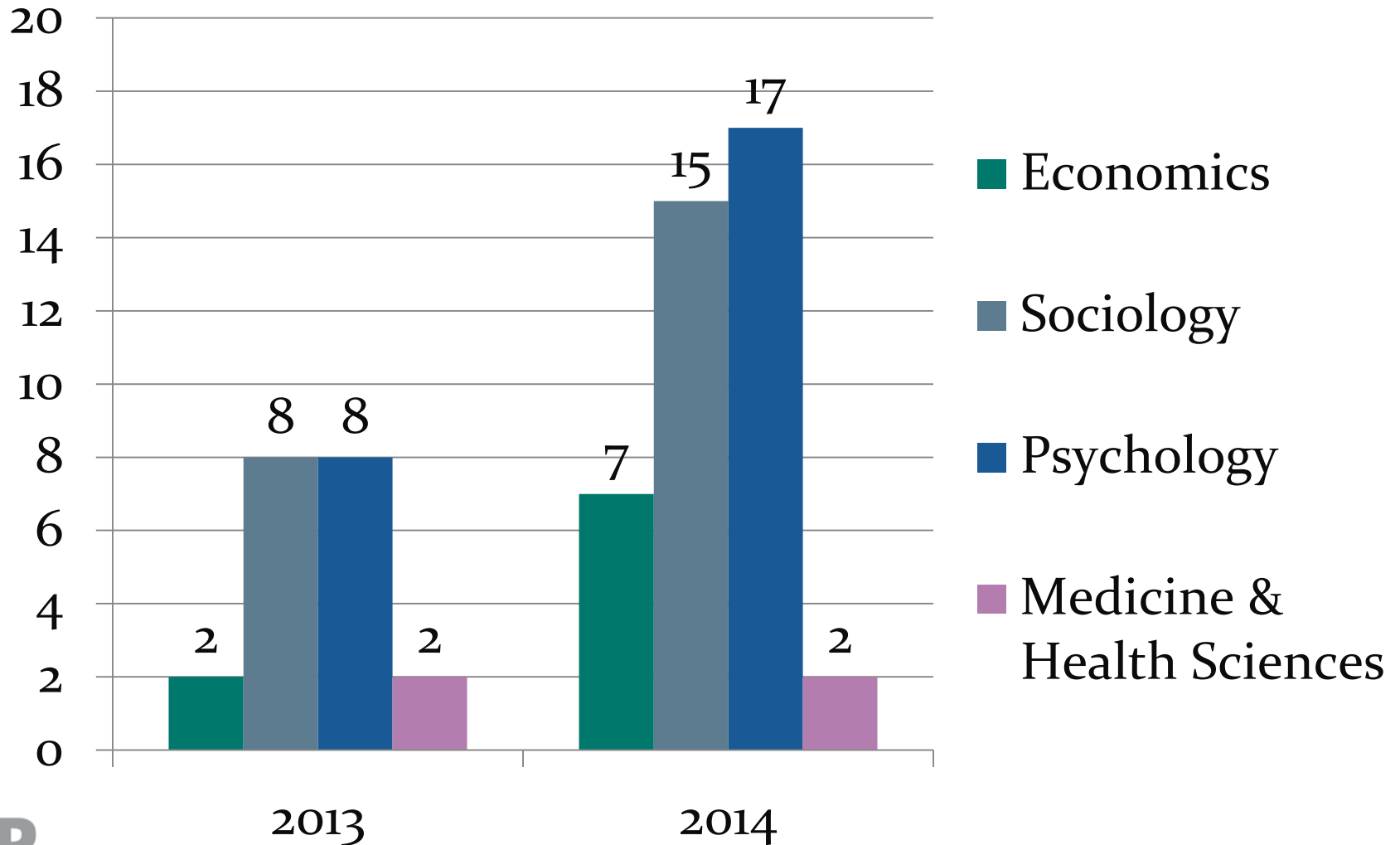
## Compare category labels

| Variable:             | rp51       | sp52a      | tp7004     | up51       | vp63       | wp52       | xp65       | yp61       | zp63       | bap52      | bbp65      | bcp52      |
|-----------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Dataset:              | rp         | sp         | tp         | up         | vp         | wp         | xp         | yp         | zp         | bap        | bbp        | bcp        |
| Period:               | 2001       | 2002       | 2003       | 2004       | 2005       | 2006       | 2007       | 2008       | 2009       | 2010       | 2011       | 2012       |
| [x] answer improbable | -3 (0)     | -3 (0)     | -3 (0)     | -3 (0)     | -3 (0)     | -3 (0)     | -3 (0)     | -3 (0)     | -3 (0)     | -3 (0)     | -3 (0)     | -3 (0)     |
| [x] does not apply    | -2 (20037) | -2 (21335) | -2 (20165) | -2 (19630) | -2 (18807) | -2 (19905) | -2 (18512) | -2 (17436) | -2 (18376) | -2 (16594) | -2 (18646) | -2 (18310) |
| [x] no answer         | -1 (78)    | -1 (98)    | -1 (73)    | -1 (52)    | -1 (55)    | -1 (54)    | -1 (51)    | -1 (48)    | -1 (55)    | -1 (51)    | -1 (44)    | -1 (33)    |
| [x] yes               | 1 (843)    | 1 (844)    | 1 (1125)   |            |            |            |            |            |            |            |            |            |
| [x] no                | 2 (1393)   | 2 (1615)   | 2 (1248)   | 3 (1267)   | 3 (1208)   | 3 (1292)   | 3 (1256)   | 3 (1147)   | 3 (1229)   | 3 (1185)   | 3 (1257)   | 3 (1282)   |
| [x] yes mini-job      |            |            |            | 1 (897)    | 1 (863)    | 1 (946)    | 1 (885)    | 1 (867)    | 1 (941)    | 1 (886)    | 1 (915)    | 1 (966)    |
| [x] yes midi job      |            |            |            | 2 (173)    | 2 (172)    | 2 (161)    | 2 (182)    | 2 (186)    | 2 (191)    | 2 (197)    | 2 (207)    | 2 (215)    |

# 3 Submissions to SOEP-IS for survey-years 2013 & 2014



# Submissions to SOEP-IS for survey-years 2013 & 2014





# Samples and Questions – SOEP-IS 2013

| Module                   | Supplementary<br>Sample 2013<br>N = 1,000 hh | Longitudinal<br>Samples (I+E)<br>N = 1,100 hh | Supplementary<br>Sample 2012<br>N = 750 hh |
|--------------------------|--|---|--|
| Factorial Survey         |  | X   | X  |
| Individual Job<br>Tasks  | X  |   |  |
| Conspiracy<br>Mentality  |  | X   |  |
| Narcissism               |  | X   |  |
| Physical Activity        | X  |   | X  |
| Mobility and<br>Identity |  | X   |  |
| Sleep                    |  |   | X  |
| DRM                      |  | X   |  |

# Samples and Questions – SOEP-IS 2014

| Module              | SS 2014<br>N = 1,000 hh | I+E<br>N = 1,100 hh | SS 2012<br>N = 750 hh | SS 2013<br>N = 900 hh |
|---------------------|-------------------------|---------------------|-----------------------|-----------------------|
| Investment          |                         |                     | X                     |                       |
| Desc. vs. Exp.      |                         |                     |                       | X                     |
| Survey Error        | X                       |                     |                       |                       |
| IB                  | X                       |                     |                       |                       |
| Life Events         | X                       | X                   | X                     | X                     |
| Redistribution      | X                       | X                   | X                     | X                     |
| CHAOS               |                         | X                   | X                     | X                     |
| Overconfidence      |                         |                     |                       | X                     |
| Justice Sensitivity |                         | X                   |                       |                       |
| Lottery Play        | X                       | X                   | X                     | X                     |
| Question Format     |                         |                     | X                     |                       |
| Flourishing State   |                         | X                   | X                     | X                     |
| Ideal Level of LS   |                         | X                   | X                     | X                     |
| DRM                 |                         | X                   |                       |                       |

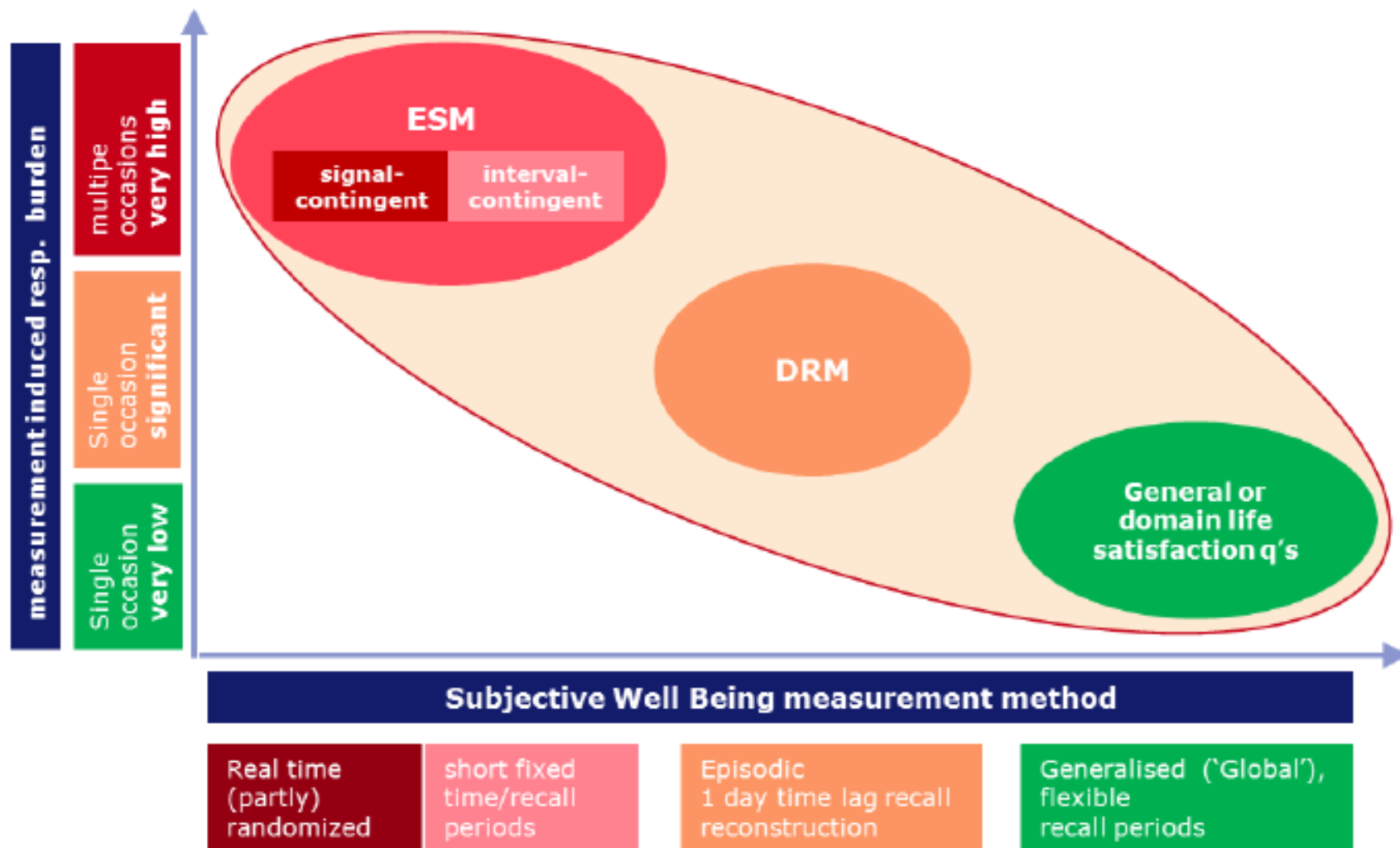
- Funded by NIH-project of Richard E. Lucas.
- Reconstruction of previous work day.  
[adaption of DRM by Kahneman et al. 2004 and a replication of 2009 PSID's supplement on Disability and Use of Time (DUST)]
  - List of 22 activities, two activities per episode possible.
  - General rating of pleasantness for every episode.
  - Additional ratings of emotions for 3 random episodes.
- Data collection in the SOEP Innovation Sample.
  - CAPI, Interviewer fills in questionnaire.
  - Pretest of procedure in SOEP-IS supplementary sample 2012.
  - Application of procedure to longitudinal SOEP-IS in 2012.
  - Measurement of long-term stability in 2013, 2014, 2015.

## Subjective Well Being – ESM in SOEP-IS

- Realized by the 25 to 30 most experienced interviewers.
- About 500 households in big German cities will be contacted.
- 225 individuals are expected to finish the data collection.

| Day     | T1   | T2                   | T3    | T4    | T5    | T6    | T7    | T8    | T9   |
|---------|------|----------------------|-------|-------|-------|-------|-------|-------|------|
| Measure | DRM1 | ESM <sub>t2-t8</sub> |       |       |       |       |       |       | DRM2 |
| Mode    | F2F  | Mobil                |       |       |       |       |       |       | F2F  |
| Length  | 15   | 7*2,5                | 7*2,5 | 7*2,5 | 7*2,5 | 7*2,5 | 7*2,5 | 7*2,5 | 15   |

- TNS Infratest will buy 111 smartphones for this year's data collection and further use in the SOEP-IS or other SOEP-projects.



## Saliva Collection in 2015 to Assess Epigenetics

- Project of Roland Weierstall & Elisa Helms (Uni Konstanz) in cooperation with Thomas Elbert (Uni Konstanz) & Isabelle Mansuy (Uni Zürich).
- Background: Influence of stress on epigenetics.
- Sample size: about 400 individuals.
- External funding is already secured.
- Project is approved by ethics committee of University of Konstanz.
- Interviewers will provide background information and drop off sampling kits.



## Saliva Collection in 2015 to Assess Epigenetics

- Participants will collect saliva samples themselves.





## Call for Proposals for Data Collection 2015

- Call for 2015 was sent out with SOEPnewsletter and is distributed to different list-servers and journals (announcement sections)
- Fieldwork will start in September and end in December 2015.
- Researchers should contact SOEP Survey Management with a project outline by November 30, 2014, at the latest
- Deadline for submissions: December 31, 2014
- Decision and feedback to applicant by end of April 2015

Thank you very much for your attention

---



**DIW Berlin — German Institute  
For Economic Research**

Mohrenstraße 58, 10117 Berlin, Germany  
[www.diw.de](http://www.diw.de)