

# gesis

Leibniz Institute  
for the Social Sciences



## IPD Meta-Analysis of Complex Survey-Based Data

**Anna-Carolina Haensch, Bernd Weiß**  
**Research Synthesis Conference**

## Motivation: The research project HaSpaD

- Harmonizing and synthesizing partnership histories from different research data infrastructures (HaSpaD)

**DJI**  
Deutsches  
Jugendinstitut



**GGP**

**ALBUS**



  
**SHARE**  
Survey of Health, Ageing  
and Retirement in Europe  
50+ in Europe

**SOEP** Das Sozio-  
oekonomische  
Panel

## Motivation

- **Goal:** To allow researchers to investigate previously unanswered questions regarding relationship stability from a historical and life-course perspective
- One of the research puzzles on the way: **How do we deal with complex survey data and their survey weights when conducting an individual participant data (IPD) meta-analysis (MA)?**

- Survey programmes and survey weights:
  - **pairfam**: poststratification weights
  - **ALLBUS**: household sample, transformation weights+design weights
  - **Lebenslaufstudie (Life History Study)**: self-weighting random sample
  - **SOEP**: design weights+poststratification weights
  - **Familiensurvey**: post-stratification weights
  - **Family and Fertility Survey**: design weights
  - **Mannheim Divorce Study**: design weights
  - **Generations & Gender Survey**: Design weights + post-stratification weights
  - **SHARElife**: Design weights + post-stratification weights



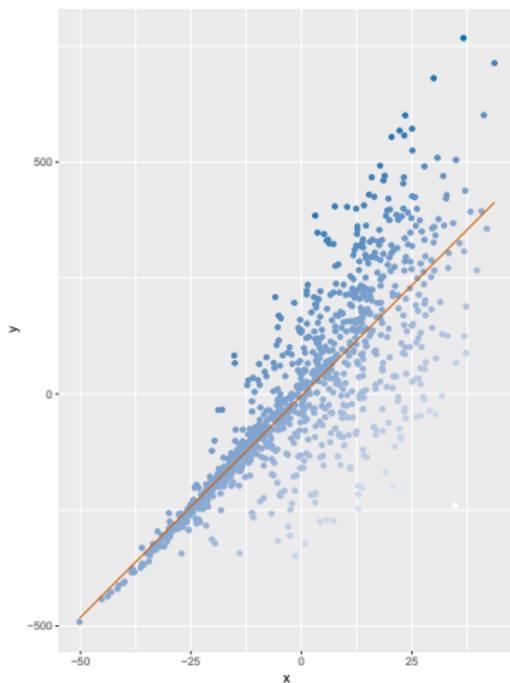
## Research Questions

- When do we have to use survey weights in MA of regression coefficients?
- Do weights have to be transformed when we pool data? If yes, how and when?
- Which meta-analytical approach is better suited: one-stage or two-stage IPD MA?

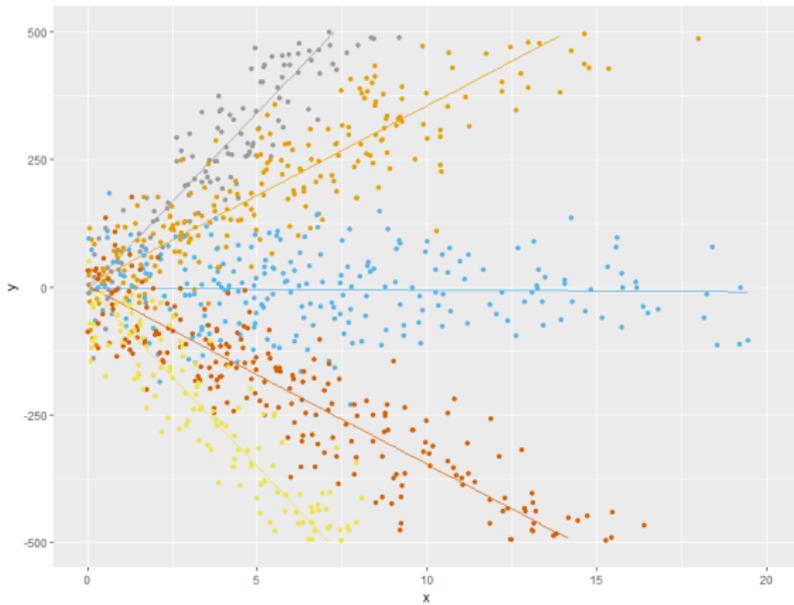
## Literature

- Little literature on survey weighted IPD meta-analysis of regression coefficients with pooled complex data
- More on survey weighted linear regression with a single dataset
  - DuMouchel/Duncan (1983), Pfeffermann (2009) and Solon, Haider and Wooldridge (2013) provide guidance on **when to use survey weighted regressions instead of unweighted regressions**
  - However, researchers seldom take into account the complex survey design in their regression analysis even though neglecting it can lead to substantial **analytical error** (West, Sakshaug, Aurelien 2016)

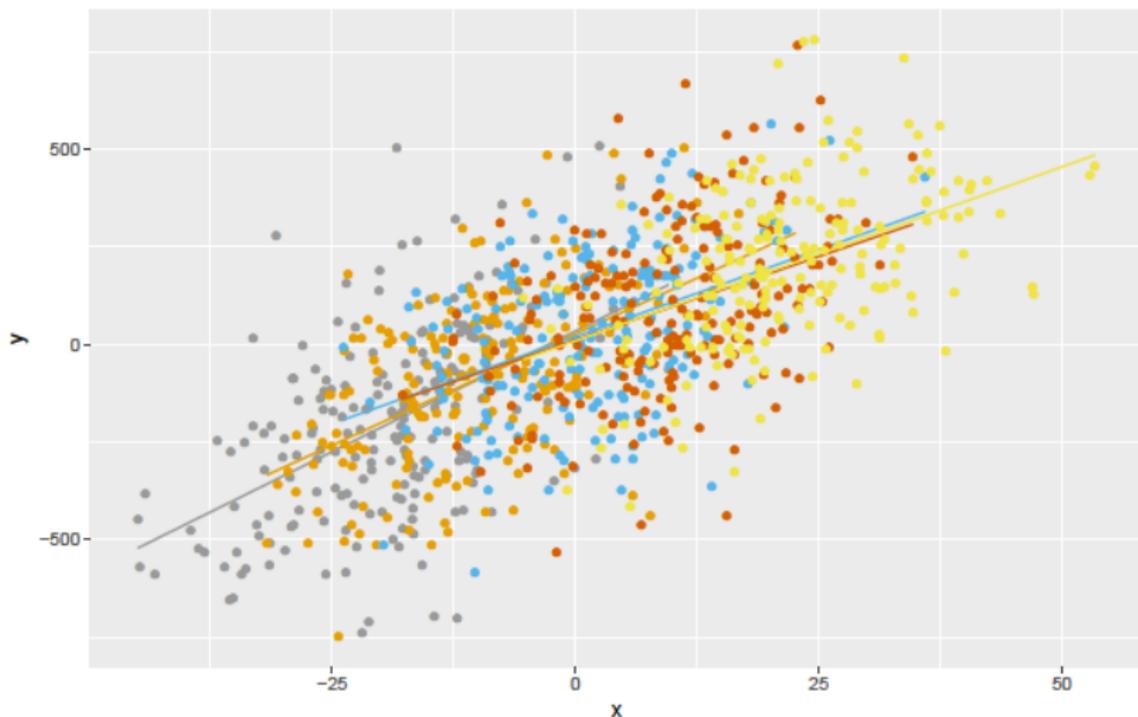
# Endogenous sampling



## Heterogeneity of effects



## Ignorable survey weights



## Literature

- Little literature on survey weighted IPD meta-analysis of regression coefficients with pooled complex data
- More on survey weighted linear regression with a single dataset
  - DuMouchel/Duncan (1983), Pfeffermann (2009) and Solon, Haider and Wooldridge (2013) provide guidance on **when to use survey weighted regressions instead of unweighted regressions**
  - However, researchers seldom take into account the complex survey design in their regression analysis even though neglecting it can lead to substantial **analytical error** (West, Sakshaug, Aurelien 2016)

## Results (1/3)

- Possible bias reduction in IPD MA through survey weighting in case of
  - endogenous selection
  - heterogeneity of effects models
- Whether survey weights are needed should be judged depending on the research question and the model at hand and survey by survey

## Research Questions

- When do we have to use survey weights in MA of regression coefficients?
- Do weights have to be transformed when we pool data? If yes, how and when?
- Which meta-analytical approach is better suited: one-stage or two-stage IPD MA?

## Results (2/3)

- Transformation of survey weights may be needed, e.g.,
  - to ensure that the survey weights can be interpreted in the usual sense of oversampling
  - if one wants to include random effects into a one-stage meta-analysis
    - Different possibilities: Scale the weights so that the transformed weights sum to the effective sample size or the sample size (Asparouhov 2006)

## Research Questions

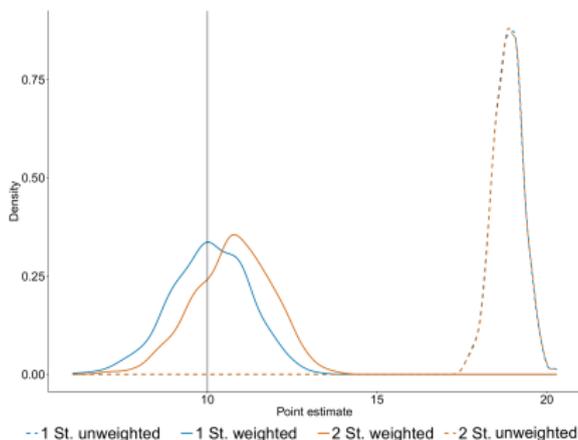
- When do we have to use survey weights in MA of regression coefficients?
- Do weights have to be transformed when we pool data? If yes, how and when?
- Which meta-analytical approach is better suited: one-stage or two-stage IPD MA?

## One-stage and two-stage IPD meta-analysis

- Two-stage IPD meta-analysis
  - Stage 1: Analysis of each survey separately to obtain survey-specific effect sizes
  - Stage 2: Combining the effect sizes by calculating a weighted average (often inverse error-variance based)
- One-stage IPD meta-analysis
  - Only one combined dataset: We combine surveys and then analyze the combined dataset

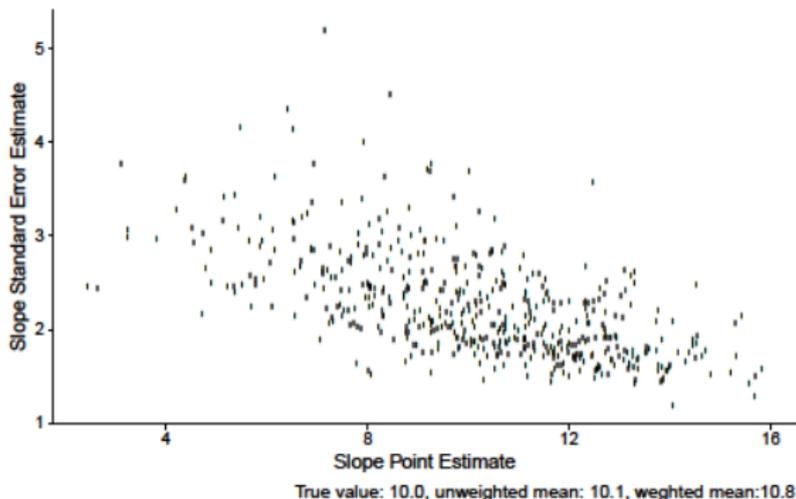
## Results (3/3)

Survey weighting can introduce bias in case of two-stage IPD meta-analysis



Density plot of slope point estimates. Surveys drawn with strata-specific sampling probability, slopes varying by strata. 1000 MC reps.

## Results (3/3)



Scatter plot of slope point estimates against standard error estimates for 500 simulated surveys. Surveys drawn with strata-specific sampling probability, slopes varying by strata.

## Conclusion

General "rules" of survey-weighted regression also apply to the IPD meta-analytical case

- Survey weights are needed/useful if informative about error terms, decide survey by survey if needed
- Survey weights cannot always be used in raw (inverse probability) form, especially in case of pooled surveys (one-stage meta-analysis)

Incorporating survey weights can change standard error estimates and can **bias two-stage meta-analytical estimates** due to the assumption of known variances

Thank you for your attention!  
Contact: [anna-carolina.haensch@gesis.org](mailto:anna-carolina.haensch@gesis.org)

## References

- Asparouhov, T. (2006). General Multi-Level Modeling with Sampling Weights. *Communications in Statistics - Theory and Methods*, 35(3):439–460.
- DuMouchel, W. H. and Duncan, G. J. (1983). Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples. *Journal of the American Statistical Association*, 78(383):535–543.
- Pfeffermann, D. and Sverchkov, M. (1999). Parametric and semi-parametric estimation of regression models fitted to survey data. *Sankhya: The Indian Journal of Statistics, Series B (1960-2002)*, 61(1):166–186.

# References

- Solon, G., Haider, S. J., and Wooldridge, J. (2013). What Are We Weighting For? Working Paper 18859, National Bureau of Economic Research.
- West, B. T., Sakshaug, J. W., and Aurelien, G. A. S. (2016). How big of a problem is analytic error in secondary analyses of survey data? PLOS ONE, 11(6):1–29.