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)
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?
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 and 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.
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.

True value: 10.0, unweighted mean: 10.1, weighted mean: 10.8
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
References


References
