Session 3	**: Monday, June 15th, 2009 (13:00 – 15:00 hrs)**
Chair	**: Frauke Kreuter**
Presenters	**: Yehuda Dayan, Gerrit Müller, Trena Ezzati-Rice, Lundquist**
Format	**: Questions following each presentation. General comments at the end.**

A Structured Approach to Inference from Internet Panels (*Yehuda Dayan*)

- Response propensity models do not work well when you have propensities close to zero. How do you deal with this issue? (Paul Biemer)
    - We use propensity strata, no the predicted response propensity $(\hat{p}_i)$ directly. There're approximately 40 cases out of 45,000 with predicted response propensities equal to zero. It's not a large problem (*Yehuda Dayan*).
- Why are you interested in modeling the response propensity in stages for purposes of adjustment? I understand that you usually want to model the stages separately to understand better the mechanisms behind each process, but this is not your purpose (Andy Peytchev)
    - We have tried to model the response propensity in a single stage before and it did not work. We're hopping that modeling in 3 stages can work better (*Yehuda Dayan*).
- My hunch is that your model will fit better if you model it in a single stage. If you do 3 stages the combined errors will be larger than in a single stage. The purpose of doing separate stages follows the Groves and Couper framework where you're interested in modeling two different processes (contact and cooperation) (Matthias Schonlau).
- Using 3 stages is useful to understand the influences on each stage, but this modeling strategy might not be "efficient" to get good model fit. (N.N.)
- If you use the same covariates in the 3 models you can just fit a single model. But if you have different covariates or covariates at different levels for different stages then it makes snese to use the different models (Andy Peytchev).

Nonresponse and Measurement Error in Employment Research (Frauke Kreuter, *Gerrit Müller* and Mark Trappmann)

- Was missing data a serious problem in your analysis? (N.N.)
    - The percentage of missing data was small - around 3-5%. We did complete case analysis. We don't think this was a serious problem in our analysis (*Gerrit Müller*).
- This analysis was inspired by an old paper by Cannell and colleagues. They looked at reports of number of days stayed at hospitals and they found larger errors as the period of retrospective recall increased. Respondents interviewed later in the field period face this challenge (Frauke Kreuter).
- So, are you saying that if we aim for 100% response you get more biased estimates? Is that a fair conclusion from your analysis? (N.N.)
    - Yes (Frauke Kreuter).

- There's something else. Even if people express interest (in the survey) they are bad reporters! Old people are the worst reporters. The Consumer Expenditure Survey is a tedious survey. They (old people) are interested but they don't know what a "good job" is (Clyde Tucker).
- So what to do finally? Would you suggest dropping these cases (fielded later and with error)? (N.N.)
    - No, it does not make sense to drop cases. One could model this problem in substantive analysis (*Gerrit Müller*)
    - One could try to improve the questionnaire (e.g. ask for a moving reporting period not a fixed date), but not drop the cases (Frauke Kreuter).

Evaluation of Respondents' Reporting of Medical Events and Relationship to Response Propensities in the Medical Expenditure Survey (*Trena Ezzati-Rice*, Frederick Rohde, and Robert Baskin)

- Does MEPS oversample people with high number of medical events? (N.N.)
    - Not currently, but we're looking into it (*Trena Ezzati-Rice*).
- Did you take into account the complex design in your analysis? (N.N.)
    - No, but we could do it in the future (*Trena Ezzati-Rice*).
    - Well, you may get different results if you consider the complex design.
- What do you think about how useful are those response propensity models if they explain so little (less than 10% of the variance)? (N.N.)
    - This is a criticism we receive a lot. We have pressure to look into this models because we see these huge drops in the reports of medical events as the individuals stays in the panel (*Trena Ezzati-Rice*).
- Maybe you wan to use CHART or Regression Trees or other nonparametric strategy available. Maybe the influences are nonlinear and your logistic regression models are not able to capture them. (N.N.)
    - We have used regression trees in BLS, but they are not easy to present to the audience. In our field people don't like them! (Clyde Tucker)
- One last suggestion. You may want to use the "number of attempts to 1st contact" instead of the "number of attempts to interview". We have found that the former is more predictive of contact. (N.N.)

Choosing the Number of Call Attempts to Minimize the Nonresponse Bias under a Response Homogeneity Groups (RHG) Approach (Annica Isaksson and *Peter Lundquist*)

- No questions asked.

General Comments

- It's important to consider the interplay between *nonresponse* and *measurement error*. If we use survey data to get estimates of nonresponse bias we might be including measurement error because we're not looking at true values. This is a general critique to those studies looking at propensity models (sample data) to do nonresponse bias analysis (Paul Biemer).
    - Well – this is exactly what *Gerrit* looked at in his presentation. He had the true values from the administrative records, so he was able to decompose the error into the nonresponse and the measurement error part (Frauke Kreuter).
- It seems to me that nonresponse bias is like a mith … we're still looking for it! (Stephen Horn).
- Another problem is that we don't have a theory that incorporates both bias and variance – like for confidence intervals for example. The Neyman framework assumes zero bias. We don't have theory to include a non-zero bias in our confidence intervals (Paul Biemer).
- There's also the issue of "statistician variance". You ask 5 different statisticians and you'll get 5 different results! (Paul Biemer).