Item Sum: A New Technique for Asking Quantitative Sensitive Questions

gesis Colloquium
Mannheim, May 3rd, 2016

Mark Trappmann (IAB / Universität Bamberg)
Ivar Krumpal (Universität of Leipzig)
Antje Kirchner (University of Nebraska at Lincoln)
Ben Jann (Universität Bern)
Research Objectives
Research Questions

- Develop a privacy preserving survey-based technique to measure quantitative sensitive characteristics and derive estimators
- Apply to measurement of the amount of undeclared work in Germany
- Compare with standard direct questioning (‘more-is-better’ assumption)
Sensitive Questions
Asking Sensitive Questions in Surveys

- “A question is sensitive when it asks for a socially undesirable answer, when it asks in effect, that the respondent admits he or she has violated a social norm.” (Tourangeau and Yan, 2007: 860)

- Non-random responses distort estimates due to (Tourangeau and Smith, 1996: 276)
  - partial nonresponse (break-offs),
  - (item-)nonresponse (refusal),
  - misreporting (here: systematic underreporting)
Techniques for Asking Sensitive Questions

- **Wording**
  - load questions / “forgiving wording”
  - ask about long periods / distant past
  - paraphrase

- **Mode**
  - self administration

- **Dejeopardizing Techniques**
  - Randomized Response Technique (Warner 1965)
  - Item Count Technique (Smith et al. 1975, Droitcour et al. 1991)
  - Crosswise Model (Yu et al. 2008, Jann et al. 2012)
The Item Count Technique

- Randomize respondents to two subsamples
  - Group 1: Short list (SL) of n innocuous questions
  - Group 2: Long list (LL): Same as SL plus sensitive item
- Ask for number of items in the list that apply
- Mean difference between group between groups provides prevalence estimate of sensitive item
- Expectation of higher estimates confirmed in several experimental studies (overview Holbrook and Krosnick 2010a)
- Moderate cognitive burden
- Recently shown to outperforms RRT (Holbrook and Krosnick 2010a, b)
- Until recently never generalized to quantitative items
The Item Sum Technique
The Item Sum Technique: Basic Idea

- Generalization of the IST
- Same randomization to SL and LL
- Ask for the sum of values of a list instead of the number of items that apply
Let S be the sensitive item of interest and C be the non-sensitive control item. Observed is:

\[ Y_i = \begin{cases} 
S_i + C_i & \text{if } i \in LL \text{ (Long List)} \\
C_i & \text{if } i \in SL \text{ (Short List)} 
\end{cases} \]

Estimates of the unbiased population mean of S are calculated as the difference in the expected values of the two lists (assuming the two samples are unbiased):

\[ \hat{\mu} = \bar{Y}_{LL} - \bar{Y}_{SL} \]

The sampling variance of the mean estimate of S is given as:

\[ \hat{V}[\hat{\mu}] = \hat{V}[\bar{Y}_{LL}] + \hat{V}[\bar{Y}_{SL}] \]
Application: Measurement of Undeclared Work
Measurement of Undeclared Work

- Definition: Paid work hidden from (tax) authorities
- Punishable by law
- Obviously sensitive topic: norm violation punishable by law
  - socially undesirable
  - fear of disclosure
- In spite of this: No nationwide studies in Germany using dejeopardizing techniques (Boockmann et al. 2010)
- Wide range of prevalence estimates from existing studies
Our Study (Design)

- Nationwide telephone survey conducted in 2010
  - 3,211 interviews
  - Sample 1: People aged 18-70 who were employed in December 2009 (RR1: 16.3%) drawn from registers at FEA
  - Sample 2: UB II recipients in June 2010 (RR1: 18.8%) drawn from registers at FEA

- Randomization to DQ and IST
  - Randomization within both subsamples
  - ~ one third DQ (n=1,145), ~ two thirds IST (n=2,066)
  - Within IST: 50% LL and 50% SL
  - About 15% noncompliance related to previous RRT experiment: Responded in DQ mode, but unrelated to randomization within IST
### Implementation of the IST

<table>
<thead>
<tr>
<th>Group LL (Long List)</th>
<th>Group SL (Short List)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C1:</strong> How many hours did you watch TV last week?</td>
<td><strong>C1:</strong> How many hours did you watch TV last week?</td>
</tr>
<tr>
<td><strong>S1:</strong> How many hours per week do you usually engage in undeclared work?</td>
<td></td>
</tr>
</tbody>
</table>

Please sum up the answer to both questions, please, do not report individual answers.

<table>
<thead>
<tr>
<th>Group LL (Long List)</th>
<th>Group SL (Short List)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C2:</strong> How much do you pay per month for your apartment or house?</td>
<td><strong>C2:</strong> How much do you pay per month for your apartment or house?</td>
</tr>
<tr>
<td><strong>S2:</strong> How high are your usual earnings per month engaging in undeclared work?</td>
<td></td>
</tr>
</tbody>
</table>

Please sum up the answer to both questions, please, do not report individual answers.
Results
### Item Sum: Mean Estimates
(isreg: Jann 2013)

<table>
<thead>
<tr>
<th>Hours of undeclared work (per week)</th>
<th>Employees</th>
<th>Benefit recipients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct questioning (DQ)</td>
<td>0.07 (0.03)</td>
<td>0.14 (0.06)</td>
</tr>
<tr>
<td>Assigned to DQ</td>
<td>0.07 (0.04)</td>
<td>0.19 (0.08)</td>
</tr>
<tr>
<td>Opted for DQ</td>
<td>0.05 (0.05)</td>
<td>0.00 (-)</td>
</tr>
<tr>
<td>Item sum technique (IST)</td>
<td>0.85 (0.70)</td>
<td>-0.17 (1.06)</td>
</tr>
</tbody>
</table>
## Item Sum: Mean Estimates
(isreg: Jann 2013)

<table>
<thead>
<tr>
<th>Hours of undeclared work (per week)</th>
<th>Employees</th>
<th>Benefit recipients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct questioning (DQ)</td>
<td>0.07</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Assigned to DQ</td>
<td>0.07</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Opted for DQ</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(-)</td>
</tr>
<tr>
<td>Item sum technique (IST)</td>
<td>0.85</td>
<td>-0.17</td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(1.06)</td>
</tr>
</tbody>
</table>
## Item Sum: Mean Estimates
*(isreg: Jann 2013)*

<table>
<thead>
<tr>
<th></th>
<th>Hours of undeclared work (per week)</th>
<th>Earnings from undeclared work (per month)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Employees</td>
<td>Benefit recipients</td>
</tr>
<tr>
<td>Direct questioning (DQ)</td>
<td>0.07</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Assigned to DQ</td>
<td>0.07</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Opted for DQ</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(-)</td>
</tr>
<tr>
<td>Item sum technique (IST)</td>
<td>0.85</td>
<td>-0.17</td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(1.06)</td>
</tr>
</tbody>
</table>
An extension: Item Sum Regression Models
(isreg: Jann 2013)

<table>
<thead>
<tr>
<th></th>
<th>Hours of undeclared work</th>
<th>Earnings from undeclared work</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Naive</td>
<td>ITT</td>
</tr>
<tr>
<td>IST</td>
<td>0.78</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(0.62)</td>
</tr>
<tr>
<td>IST × Benefit recipients</td>
<td>-1.09</td>
<td>-1.02</td>
</tr>
<tr>
<td></td>
<td>(1.27)</td>
<td>(1.04)</td>
</tr>
<tr>
<td>Benefit recipients</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.07*</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>N</td>
<td>3199</td>
<td>3072</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses (ITT/IV: bootstrap standard errors); ITT: Intention-to-treat estimates; IV: Instrumental variables estimates

* p < 0.05, ** p < 0.01
Summary and Discussion
Summary

- Generalizing the ICT we have presented a new privacy preserving technique for quantitative sensitive items and applied it in a study on undeclared work
  - no randomizing device
  - low cognitive effort
  - implementation easily possible (interviewer and self administered)

- We have derived point estimators and regression estimators for IST variables

- Results indicate that the item sum technique (IST) might be a fruitful approach for eliciting a higher extent of socially undesirable behaviour compared to direct questioning (‘more-is-better’ assumption)
Discussion

- Unknown why IST produced higher estimates for earnings but not for hours: Further research needed:
  - Not a power issue
  - Choice of innocuous item? Replication desirable
  - Test “no design effect” assumption (Blair and Imai 2012) with two innocuous items

- Large standard errors of point estimates and regression coefficients. Trade-off with privacy protection: Find innocuous items …
  - … whose variance is overestimated by respondents or that have small variance in comparison to their mean
  - … that are negatively correlated to the sensitive item
  - … that can be explained well by other variables in the survey
Thank You

mark.trappmann@iab.de

This work has been published as:

References


References


Appendix: Estimators
Recall

\[ Y_i = \begin{cases} S_i + C_i & \text{if } L_i = 1 \text{ (long-list sample)} \\ C_i & \text{else (short-list sample)} \end{cases} \]

Assume that \( S \) and \( C \) can be modeled as

\[
S_i = X'_i \beta + v_i, \quad E(v_i) = 0 \\
C_i = Z'_i \gamma + v_i, \quad E(v_i) = 0
\]

It follows that

\[ Y_i = L_i S_i + C_i \]

\[ = L_i (X'_i \beta + v_i) + (Z'_i \gamma + v_i) \]

\[ = L_i X'_i \beta + Z'_i \gamma + \varepsilon_i \quad \text{with } \varepsilon_i = L_i v_i + v_i, \quad E(\varepsilon_i) = 0 \]

Where \( \hat{\beta} \), is an estimate of \( E(S) \), and \( \hat{\gamma} \), is an estimate of \( E(C) \)
Let $D$ be an indicator for the direct-questioning sample

$$D_i = 1 \text{ (direct-questioning sample and else 0)}$$

Let

$$Y_i = S_i^* = X_i'\beta^* + v_i, \text{ if } D_i = 1$$

where $\beta^*$ is a coefficient vector including social-desirability bias,

then

$$Y_i = L_iX_i'\beta + (1 - D_i)Z_i'\gamma + D_iX_i'\beta^* + \epsilon_i$$

or, equivalently,

$$Y_i = L_iX_i'(\beta - \beta^*) + (1 - D_i)Z_i'\gamma + (D_i + L_i)X_i'\beta^* + \epsilon_i$$

$L$: coefficient attached to $L$ provides IST effect
$(1 - D)$: mean of the non-sensitive item $C$
$(D + L)$: mean of the sensitive variable based on direct questioning

Regressing $Y$ on $L$ (i.e. if no covariates are taken into account)
For intention-to-treat estimate (ITT) of the effect of the IST in two stages:

1. Fit $C_i = Z_i' \gamma + \nu_i$ using the observations from the short-list group (realized).

2. Residualize the (realized) long-list observations using the least-squares estimate $\hat{\gamma}$ and fit

$$\tilde{Y}_i = \tilde{L}_i X_i' (\beta - \beta^*) + X_i' \beta^* + \varepsilon_i \quad \text{with} \quad \tilde{Y}_i = \begin{cases} Y_i - Z_i \hat{\gamma} & \text{if } L_i = 1 \\ Y_i & \text{else (i.e. DQ)} \end{cases}$$

based on the respondents who were initially assigned to the long-list sample or to direct questioning.

$\tilde{L}$ as an indicator for initial assignment to the long-list sample.
Estimators: IST cont.

Improve on the ITT by instrumenting $L_i$ with $\tilde{L}_i$ based on a two-stage least squares procedure in the second step:

2. $\tilde{Y}_i = L_i X_i' (\beta - \beta^*) + X_i' \beta^* + \epsilon_i$

i.e. using $L_i$ (realized) instead of $\tilde{L}_i$ (assigned) in the equation

=> instrumental variables (IV) estimate of $(\beta - \beta^*)$ provides consistent estimate of the local average treatment effect (LATE) of the IST.