JUST TAKING THE GIFT OR RETURNING THE FAVOR?

A Meta-Analysis on the Effects of Incentives for Survey Participation

Andreas Schneck and Katrin Auspurg

Rational Choice Sociology: Theory and Empirical Applications 2013 Workshop at Venice International University, San Servolo
Motivation

• Problem of declining response rates over time (for empirical evidence on decreasing response rates see Aust & Schröder 2009; De Leeuw & De Heer 2002; Groves 2011; Schnell 1997)

• Especially low response rate in web surveys (Shih & Xitao Fan 2008)

  ➢ Increased risk of nonresponse bias

Methods to increase response rates:

• Careful survey design: total (Dillman 1978) or tailored design (Dillman 2007): incentives, reminder, personalization (e.g. hand signature), etc.

  ➢ We focus on incentives in self-administered surveys
Side note: incentive terminology

- **Conditional**: on completion of survey; after survey participation
- **Unconditional**: with survey request; before survey participation
- **Monetary**: cash or check incentive
- **Nonmonetary**: items, lottery incentives (in this study also monetary lotteries)
Theory I

• **Norms of reciprocity** (Gouldner 1960; Mauss 1967)
  – Norm to repay gift (unconditional)
  ➢ In general no sanctioning possible – no “loss of face” (Mauss 1967: 41)

• **Exchange theory** (Blau 1967)
  – Focus on possible future interactions (future gains)
  – Unconditional incentive “symbol of trust” (Dillman 1978: 16)
  – Social exchange (unconditional incentive - diffuse obligation) or economic exchange (conditional incentive - payment)
  ➢ Most surveys only one-shot interaction – no future interactions
Theory II

- **Strict RC** pure utility maximizing actors: take incentive – but refuse participation to avoid opportunity costs
  - Surveys: low profit and low cost situation
    - Do only participate in case of conditional incentives

- **Bounded rationality** (Simon 1983)
  - Situations in which the actor isn’t aware of all potential costs and benefits
    - Use of simple decision heuristics (e.g.: ignore requests from strangers)
Theory III

- **Leverage salience theory** (Groves et al. 2000)
  - Leverage (preference set)
  - Salience (trigger preference by making survey attribute salient)

  - Incentives can’t convert “hard-core” nonrespondents, but unstable nonrespondents (unconditional = more salient)

  (Groves et al. 2000: 300)
Hypotheses

• H₁: The higher the incentive, the higher the odds of response (effect with declining rate)
• H₂.₁: Unconditional incentives are more effective than conditional incentives
• H₂.₂: Conditional incentives are more effective than unconditional incentives
• H₃: Monetary incentives are more effective than nonmonetary incentives
• H₄: The combination of monetary and unconditional incentives is even more effective
Incentives as a central aspect to enhance survey participation. (e.g. Armstrong 1975; Edwards et al. 2009)

- Unconditional & monetary incentives more effective
- Relationship between incentive-value and odds of response unclear (linear, curvilinear)
- Incentives effective also in telephone (Singer et al. 2000) and face-to-face surveys
Research gap

• Most studies focus on mean effect sizes and bivariate subgroup analyses only
• No analysis of the heterogeneity of incentive conditions
• No in-depth theoretical explanation of incentive-mechanisms

➤ What are conditions of incentives to be effective under control of study characteristics?
Data

• (Hopefully) all published English and German language incentive experiments (Deadline March 2013)

• Inclusion criteria:
  – Self-administered survey
  – Non-incentive control group
  – Report on number of participants & nonrespondents
  – Description of incentive (incentive amount or incentive value)
Data

• Extensive literature search
  – relevant meta-analyses (e.g. the Cochrane Review: Edwards et al. 2009)*

• Coded effect size (ES) → Odds Ratio (OR)
  – Log(OR) unbound, thus better than Risk Ratio (biased if high control group risk) or Risk Difference (RD)
  – but lower interpretability

Dataset (meeting inclusion criteria):
133 publications/ 175 studies/ 320 trials

*Special thanks to Phil Edwards for the provision of his dataset (Edwards et al., 2002)!
Methods

Meta-Analysis (MA)

- Weighted mean effect size
  - Problematic if high degree of heterogeneity

- Problem of MAs “statistical fruit salad” (Brüderl 2004); problem similar to omitted variable bias (c.f. Greene 2012: 219)

- Control for heterogeneity by Meta Regression Analyses.
  For all non-statisticians: we are trying to disentangle the fruits!
Methods

Meta Regression Analysis (MRA)

• Also possible in a common OLS framework

\[ ES_i = \beta_0 + \beta_x M_i + \varepsilon_i \]

• Problem of heteroskedasticity

➢ WLS (weighted least squares) (Stanley & Doucouliagos 2013a: 12)
  – Inverse variance weighted

• Problem of dependent ES (one control-, mult. test-groups)

➢ Multilevel models: fixed- (FE-ML) random-effects (RE-ML)
Methods

Advantages of WLS-MRA

• Better coverage and less biased as models typically used in psychology or medicine, especially in case of heterogeneity) (Stanley & Doucouliagos 2013a; Stanley & Doucouliagos 2013b)

• Good implementation in statistical packages due to the relation to “normal” OLS
  (e.g. in Stata: regress AV UV [aweight=invVar]) (c.f. MAER-Net)
Publication bias

• “Publication of research findings based on the nature and direction of the research results“ (Dickersin 2005: 13)

• Often triggered by significance thresholds (1/ 5/ 10%)

  ➢ Biased MRA (similar to nonresponse bias in surveys)

• **MRA identification method** (Stanley 2008)

  \[ ES_i = \beta_0 + \beta_1 SE_i + \beta_x M_i + \varepsilon_i \]

  – \( \beta_0 \) Precision-Effect-Test (PET) – any genuine effect of treatment?
  – \( \beta_1 \) Funnel-Asymmetry-Test (FAT) – any publication bias?
  – **Correction**: PET with squared standard Error \((SE_i)^2\); PEESE)
Descriptive results

incentive value (kernel-density-plot)

inflation-adjusted incentive value 2011 (N=320)

red line 25% & 75% percentile, green line median
WLS-MRA

Model with clustered SEs; controls: country of survey, highest lottery incentive, netto sample, surveyed population, study topic, randomisation, survey mode, trial year, reminder not displayed
Illustration of effect sizes – RD model

Models with clustered SEs; controls not displayed
Publication bias test

- Marginal significant FAT (but in the other direction as supposed (high SE - high effect)
- Small study effect?
- Significant PET – true overall effect

WLS with clustered SEs; controls not displayed
Publication bias correction

WLS with clustered SEs; controls not displayed
Multilevel implementation

WLS with clustered SEs; Controls not displayed; Multilevel necessary \( F(174, 138) = 2.94 \), random effects unbiased \( \chi^2(6) = 6.92 \), thus FE-ML not displayed.

\( WLS + \text{multilevel models} \)

- incentive-value (per 5$)
- squared incentive-value (per 5$)
- unconditional monetary interaction
- ucond. mon.

\( -0.2 \) \( 0 \) \( 0.2 \) \( 0.4 \) \( 0.6 \)

\( \text{WLS} \quad \text{RE-ML} \)
The effect of the incentive-value

functional form of incentive values

controls not displayed
Hypotheses revisited

• $H_1$ (+) the more US$ the better (effect with declining marginal rate: higher effect per US$ if low incentive)

• $H_{2.1}$ (+) unconditional incentives better

• $H_{2.2}$ (−) conditional incentives better

• $H_3$ (+) monetary incentives slightly better

• $H_4$ (+) combination of both strategies best (except WLS)
Main limitation

• **Nonresponse bias** is threatening the validity of survey results (c.f. Groves 2009: 59)

\[ \bar{y}_r - \bar{y}_s = \frac{m_s}{n_s} (\bar{y}_r - \bar{y}_m) \]

  – Differences between respondents (r) and nonrespondents (m) matter
  – High nonresponse rates increase those potential differences

➢ Response rates are only half of the story
Discussion and outlook

- Strict RC not confirmed, but applicable if extended by the model of bounded rationality and the leverage salience theory
- Norms of reciprocity one possible mechanism besides ext. RC
- Exchange theory does not fit to one-shot situations

Future work:
- Disentangle ext. RC and norms of reciprocity (e.g. potential survey participation in a factorial survey experiment)
- Include better nonresponse bias and data quality indicators
- Tackle also issues of efficiency beside effectivity
- Exchange theory better testable in panel incentive experiments (Fumagalli et al. 2013)
Thanks a lot for your attention!
Literature

Literature


• MAER-Net. "Guidelines for the Meta-Analysis of Economics Research." in http://www.hendrix.edu/uploadedFiles/Departments_and_Programs/Business_and_Economics/AMAES/Be%20Rigorous%283%29.pdf (last access: 08.11.2013), edited by Hendrix College.


• Stanley, TD, and Hristos Doucouliagos. 2013b. "Neither Fixed nor Random: Weighted Least Squares Meta-Analysis." Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.
Appendix I

Data problems

• solved
  – Inflation adjusted incentive amount/ value (by CPI)
  – Continuity correction (+0.5) to make OR computation feasible
  – Multi-level structure due to dependent effects sizes (on control group)

• unsolved
  – Missing study information (e.g. study sponsor)
  – Overestimation of the real inflation using the CPI by approximately 1.1% per year (Boskin et al. 1998:11)
## Appendix II

### Incentive modes

<table>
<thead>
<tr>
<th></th>
<th>monetary</th>
<th>Time of payment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>conditional</td>
<td>unconditional</td>
</tr>
<tr>
<td>Nonmonetary</td>
<td>76</td>
<td>40</td>
</tr>
<tr>
<td>monetary</td>
<td>24</td>
<td>179</td>
</tr>
</tbody>
</table>
Appendix III

incentive value over time

inflation-adjusted incentive value 2011 (N=200)

year of survey
unconditional conditional trendline
Appendix IV

incentive value over time

inflation-adjusted incentive value 2011 (N=320)
### Appendix V

#### Robustness checks

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>WLS-FAT-PET</th>
<th>FE-ML</th>
<th>RE-ML</th>
</tr>
</thead>
<tbody>
<tr>
<td>logOR</td>
<td>se</td>
<td>logOR</td>
<td>se</td>
</tr>
<tr>
<td>incentive-value (per $5$)</td>
<td>0.0996***</td>
<td>0.122***</td>
<td>0.115***</td>
</tr>
<tr>
<td>squared incentive-value (per $5$)</td>
<td>-0.00266***</td>
<td>0.00285***</td>
<td>-0.00280***</td>
</tr>
<tr>
<td>unconditional</td>
<td>0.335***</td>
<td>0.153</td>
<td>0.257***</td>
</tr>
<tr>
<td>monetary</td>
<td>0.241**</td>
<td>0.159</td>
<td>0.175**</td>
</tr>
<tr>
<td>interaction ucond. mon.</td>
<td>0.117</td>
<td>0.442***</td>
<td>0.295***</td>
</tr>
<tr>
<td>country: Europe (reference: Northern America)</td>
<td>0.00251</td>
<td>0.0641</td>
<td>-0.00479</td>
</tr>
<tr>
<td>country: Australia/ Oceania</td>
<td>-0.0707</td>
<td>0.0785</td>
<td>-0.0515</td>
</tr>
<tr>
<td>country: Asia</td>
<td>0.190**</td>
<td>0.0963</td>
<td>0.234</td>
</tr>
<tr>
<td>highest lottery incentive</td>
<td>0.000410***</td>
<td>0.000557**</td>
<td>0.000384***</td>
</tr>
<tr>
<td>adjusted sample</td>
<td>0.0186</td>
<td>0.0517</td>
<td>-0.0425</td>
</tr>
<tr>
<td>pop: health (reference: general)</td>
<td>-0.137</td>
<td>0.0928</td>
<td>-0.109</td>
</tr>
<tr>
<td>pop: customers</td>
<td>-0.139</td>
<td>0.123</td>
<td>-0.170</td>
</tr>
<tr>
<td>pop: education</td>
<td>0.0157</td>
<td>0.0702</td>
<td>0.0686</td>
</tr>
<tr>
<td>pop: others</td>
<td>0.0111</td>
<td>0.0723</td>
<td>0.0884</td>
</tr>
<tr>
<td>top: social (reference: market research)</td>
<td>-0.0312</td>
<td>0.0616</td>
<td>0.0193</td>
</tr>
<tr>
<td>top: health</td>
<td>0.0344</td>
<td>0.0772</td>
<td>0.0702</td>
</tr>
<tr>
<td>top: others</td>
<td>-0.160*</td>
<td>0.0962</td>
<td>-0.225*</td>
</tr>
<tr>
<td>qual: unclear (reference: nonrandom)</td>
<td>-0.243*</td>
<td>0.125</td>
<td>-0.264*</td>
</tr>
<tr>
<td>qual: random</td>
<td>-0.166</td>
<td>0.121</td>
<td>-0.190</td>
</tr>
<tr>
<td>internet</td>
<td>0.255***</td>
<td>0.0928</td>
<td>0.215*</td>
</tr>
<tr>
<td>year of study</td>
<td>-0.00747**</td>
<td>0.00341</td>
<td>-0.00675***</td>
</tr>
<tr>
<td>reminder</td>
<td>-0.0214</td>
<td>0.0200</td>
<td>-0.0157</td>
</tr>
<tr>
<td>SE (FAT)</td>
<td>-0.730*</td>
<td>0.370</td>
<td>-0.677*</td>
</tr>
<tr>
<td>constant (PET)</td>
<td>0.736***</td>
<td>0.236</td>
<td>0.681***</td>
</tr>
</tbody>
</table>

**Observations** 296  296  296  
**R-squared** 0.552  0.865  
**Number of q_StudyID** 157  157  

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
## Appendix VI

<table>
<thead>
<tr>
<th>Robustness checks</th>
<th>WLS-FAT-PET</th>
<th>+ page length</th>
<th>outlier robust</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VARIABLES</strong></td>
<td>logOR</td>
<td>se</td>
<td>logOR</td>
</tr>
<tr>
<td>incentive-value (per 5$)</td>
<td>0.0996***</td>
<td>(0.0312)</td>
<td>0.0786**</td>
</tr>
<tr>
<td>squared incentive-value (per 5$)</td>
<td>-0.00266***</td>
<td>(0.000847)</td>
<td>-0.00193**</td>
</tr>
<tr>
<td>unconditional</td>
<td>0.335***</td>
<td>(0.0807)</td>
<td>0.393***</td>
</tr>
<tr>
<td>monetary</td>
<td>0.241**</td>
<td>(0.119)</td>
<td>0.369***</td>
</tr>
<tr>
<td>interaction ucond. mon.</td>
<td>0.117</td>
<td>(0.133)</td>
<td>0.109</td>
</tr>
<tr>
<td>country: Europe (reference: Northern America)</td>
<td>0.00251</td>
<td>(0.0641)</td>
<td>-0.0317</td>
</tr>
<tr>
<td>country: Australia/Oceania</td>
<td>-0.0707</td>
<td>(0.0785)</td>
<td>0.106</td>
</tr>
<tr>
<td>country: Asia</td>
<td>0.190**</td>
<td>(0.0963)</td>
<td>0.120</td>
</tr>
<tr>
<td>highest lottery incentive</td>
<td>0.000410**</td>
<td>(0.000185)</td>
<td>0.000428</td>
</tr>
<tr>
<td>adjusted sample</td>
<td>0.0186</td>
<td>(0.0517)</td>
<td>0.0251</td>
</tr>
<tr>
<td>pop: health (reference: general)</td>
<td>-0.137</td>
<td>(0.0928)</td>
<td>-0.198**</td>
</tr>
<tr>
<td>pop: customers</td>
<td>-0.139</td>
<td>(0.123)</td>
<td>-0.00134</td>
</tr>
<tr>
<td>pop: education</td>
<td>0.0157</td>
<td>(0.0702)</td>
<td>0.189</td>
</tr>
<tr>
<td>pop: others</td>
<td>0.0111</td>
<td>(0.0723)</td>
<td>0.0323</td>
</tr>
<tr>
<td>top: social (reference: market research)</td>
<td>-0.0312</td>
<td>(0.0616)</td>
<td>0.196*</td>
</tr>
<tr>
<td>top: health</td>
<td>0.0344</td>
<td>(0.0772)</td>
<td>0.162*</td>
</tr>
<tr>
<td>top: others</td>
<td>-0.160*</td>
<td>(0.0962)</td>
<td>-0.0455</td>
</tr>
<tr>
<td>qual: unclear (reference: nonrandom)</td>
<td>-0.243*</td>
<td>(0.125)</td>
<td>-0.319*</td>
</tr>
<tr>
<td>qual: random</td>
<td>-0.166</td>
<td>(0.121)</td>
<td>-0.238</td>
</tr>
<tr>
<td>internet</td>
<td>0.255***</td>
<td>(0.0928)</td>
<td>0.110</td>
</tr>
<tr>
<td>year of study</td>
<td>-0.00747***</td>
<td>(0.00341)</td>
<td>-8.64e-05</td>
</tr>
<tr>
<td>reminder</td>
<td>-0.0214</td>
<td>(0.0200)</td>
<td>-0.0403**</td>
</tr>
<tr>
<td>SE (FAT)</td>
<td>-0.730*</td>
<td>(0.370)</td>
<td>-0.392</td>
</tr>
<tr>
<td>page length (questionaire)</td>
<td>-0.0109***</td>
<td>(0.00387)</td>
<td>0.294</td>
</tr>
<tr>
<td>constant (PET)</td>
<td>0.736***</td>
<td>(0.236)</td>
<td>0.294</td>
</tr>
<tr>
<td>Observations</td>
<td>296</td>
<td>179</td>
<td>294</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.552</td>
<td>0.537</td>
<td>0.562</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1