Plan

1. Introduction, wasted information, previous attempts to deal with dks et cetera.

In our attempt to measure political involvement and attitudes we’ve been looking to correct various sources of error measurement. The first is each individual question framing (Turangeau et al); second, survey effects (question order, panel effects etc); the third is the sampling & non response problem. This paper is about a fourth source of potential error measurement: respondent fatigue. Unlike most previous sources of error concerns, which look at various individual aspects of a survey, this paper looks at the gestalt respondent survey behavior, constructs a direct behavioral measure of respondent survey reaction, and uses it to understand how people answer questions in surveys.

1. Some lit review (about the psychology of survey response), the pre-wave measurement; some possible hypotheses:
   1. for example: correlated with interest in the survey;
   2. correlated with “interest in pols” type of variables – there is a general category (vague opinions)
2. Analysis
   1. the surveys

The four surveys used in this paper share two characteristics. They are internet based, and therefore rely on respondent self administering the survey. They are also nationwide panel surveys before and after national elections in each country. In each case, the post-election wave contains a question asking respondent whether they voted or not.

The surveys are different in many other important respects, aside from the fact of them being run in four different countries. For one, the American and the Canadian are both the result of the collaboration with the same firm, YouGov Polimetrix. On the other hand, The UK survey was run with XXX, as a regular election survey. The German survey is a pilot for a larger project, run with XXX. The more important difference is the easiness with which respondents could declare their inability to answer questions. In the American and Canadian case, the DK option is available for only XX% and XX% of the pre-election questions respectively, including knowledge questions. In the UK survey, this option is available for xx% of the questions, and in the German survey, respondents could say DK virtually for every question (this answer was available for XX of the questions).

In terms of data collection periods and number of respondents, there is significant variation among the four surveys. The ANES data was collected in several waves over the course of 12 months. In this paper I use the data from wave one, collected in January 2008 and wave 11, collected immediately in the aftermath of the general election in November 2008. The two wave sample has a total of 1322 respondents. The Canadian survey was conducted between October 2008 and May 2009 in British Columbia and Quebec, two provinces who were hosting provincial elections. A federal election also took place on October 14, 2008 and the post-wave turnout question refers to respondents’ presence at the polls for the federal elections. Given the difference in timing between the Quebec post-wave (conducted in November 2008) and the BC post-election wave (April and May 2009), I only retain data from Quebec (N=1334). The C-CAP British survey was conducted between April 2010 and June 2010, and the available sample is N=792XX. The German survey was conducted in 2010 to coincide with the May 9 federal elections in the land of North-Rhine Westphalia. The data were collected immediately prior and after the election in the same month, resulting in 786 completed pre-post election responses.

* 1. The Respondent fatigue variable

The respondent fatigue variable relies in each case on respondent survey behavior in the pre-election wave. Given the differences in questionnaire in each case, the variables used to measure respondent fatigue vary, and the list of questions involved in each survey is available in the appendix. As a general rule however, the variable captures only the number of times a person chose the “don’t know” option, and does not include the number of times the person skipped a question. Moreover, to obtain a more conservative estimate of fatigue, the RF variable does not take into account the respondent’s admitted lack of knowledge with regard to any of the political information questions or to any question asking for precise information (for example, income, or, more politically relevant information, whether they had been contacted by any party in the course of the campaign). To prevent any contamination with other important political attitudes, it also does not include questions about political interest or intention to vote. In essence then, the RF variable measures the frequency of “don’t know”s in response to questions probing for general political opinions, ratings and preferences.

The construction of the variable follows a simple algorithm. For each question a new binary variable is created, ranging from 1 (if the respondent she didn’t know the answer to the question) to 0 (if the respondent chose any other option). The RF measure is the sum of all binary variables, ranging from 0 to the maximum number n of “Don’t know”s by an individual in the survey.

For weighting purposes, the variable is transformed by adding 1 (which means that it runs from 1 to n+1), and by dividing the sum to match the real number of cases as closely as possible.[[1]](#footnote-1) Table 1 presents the distribution of the RF variables, and RF weights in each case.

[Table 1 here]

* 1. correlations
  2. analyses
     1. as IV
     2. as weighing – iw weights are just as pw weights, but with robust standard errors; however, the model errors do not seem correlated with the IVs which means that non-robust errors can be used
     3. Understanding the difference in estimates when used as an IV and when used as a weight is crucial.

\*\*\*here: look at late Nov 15 analyses – some testing of what the weight does as opposed to the IV; basically the IV leaves the predicted Y (almost) intact, but lowers the coefficient on the IV of interest. The weight lowers the predicted Y, but boosts a bit the coefficient (and to a larger extent the std error) on the IV of interest.

1. Conclusions
   1. Why this is important – the organic nature
   2. Improve estimation of important variables
   3. Use either as a weight or as a IV depending on the variation in the sample.

Table 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  | N | Mean | Std. Dev. | Min | Max |
| Germany (2010) |  |  |  |  |  |
| Respondent fatigue | 783 | 5.643678 | 10.38773 | 0 | 67 |
| Respondent fatigue weight | 783 | .9969356 | 1.558759 | .1500578 | 10.20393 |
|  |  |  |  |  |  |
| UK (2010) |  |  |  |  |  |
| Respondent fatigue | 762 | 1.192913 | 2.459833 | 0 | 27 |
| Respondent fatigue weight | 762 | 1.034296 | 1.160189 | .4716536 | 13.2063 |
|  |  |  |  |  |  |
| Canada (2009) |  |  |  |  |  |
| Respondent fatigue | 1334 | .1116942 | .694182 | 0 | 9 |
| Respondent fatigue weight | 1334 | 1 | .6244361 | .899528 | 8.99528 |
|  |  |  |  |  |  |
| US (2008) |  |  |  |  |  |
| Respondent fatigue |  |  |  |  |  |
| Respondent fatigue weight |  |  |  |  |  |

\*\*\*Must recalculate the RFs for US – I’ve checked and the missing cases are legitimate since they reflect instances where the individual was not asked the question.

1. Here, I could divide the thing by the mean, but I would lose cases. Checked this on the UK data. [↑](#footnote-ref-1)