LORE
Laboratory of
Opinion Research

UNIVERSITY OF GOTHENBURG

LORE working paper 2013:1

Field work, survey completion times and data quality in Citizen Panel 4 - 2012

Christensen, L.
Martinsson, J.

Title:
LORE working paper 2013:1.
Field work, survey completion times and data quality in Citizen Panel 4 - 2012

LORE
Laboratory of
Opinion Research
University of Gothenburg

University of Gothenburg
Sweden
Box 100, S-405 30
Gothenburg

Title: Field work, survey completion times and data quality in Citizen Panel 4 – 2012

LORE working paper 2013:1

Series editor: Johan Martinsson

Authors:
Love Christensen, B.A.
Research assistant
Department of Political Science
University of Gothenburg
love.christensen@gu.se

Johan Martinsson, Ph.D.
Research Director of LORE
Department of Political Science
University of Gothenburg
johan.martinsson@pol.gu.se

ABSTRACT

This report examines the inflow over the period of field work of a web survey, the amount of time it takes a respondent to complete the survey (duration) and a few data quality indicators.

The first part shows us a very large part of the total completed questionnaires are received during the very first day of field work, close to 50 percent. However, after a field work period extended to three weeks, twice as many completed surveys are received in total. It is also made clear that the time of day when most people tend to answer their questionnaires is between 8 am and 11 am. These hours account for almost 40 percent of all completed questionnaires. We also discover a high amount of variation when it comes to the time respondents spend filling out the web questionnaire. On average, this particular survey took 25 minutes to complete. When it comes to predictors of survey duration it is found that people who answer late at night or who are highly interested in politics spend more time on the survey than others.

When it comes to data quality the respondents of the Citizen Panel 4 seem be straight-lining to a fairly high extent. As many as one out of four respondents had straight-lined at least one out of seven question batteries. We also confirm a curve-linear relationship between time spent on survey and data quality. Those filling out the questionnaire just slightly slower than the average respondent are those who are least likely to be straight-lining.

FIELD WORK, SURVEY COMPLETION TIMES AND DATA QUALITY IN CITIZEN PANEL 4 – 2012

One of the advantages of working with internet surveys, compared to for example postal surveys, is the possibility to acquire precise information about when a respondent started filling out the survey and when the respondent completed it. This feature of internet surveys enables us to study the field work period in detail and whether there are systematic differences between groups of respondents answering the surveys at different hours or using different amounts of time to complete the surveys.

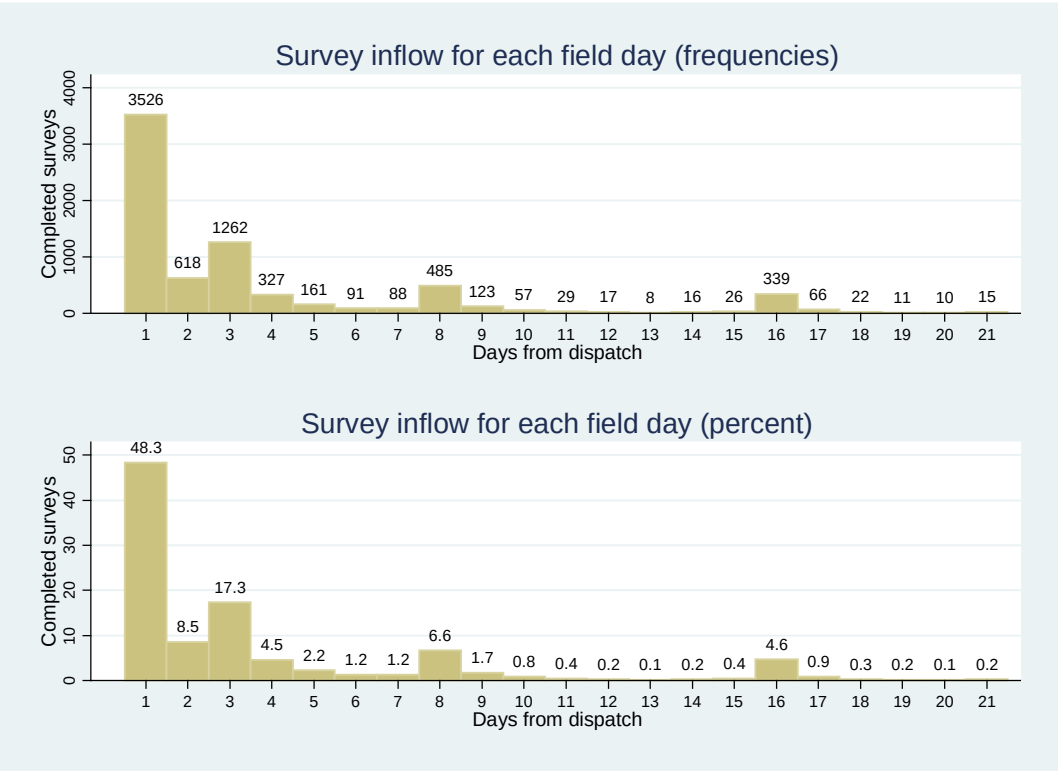
The first part of this report examines the completion rate of a large scale internet survey by the University of Gothenburg during the three weeks of fieldwork. The second part focuses on – survey duration – the amount of time a respondent uses to complete the survey, what explains variations in survey duration and if it is correlated with indicators of data quality. The survey analyzed in this report is the fourth wave of the Citizen Panel run by the Laboratory of Opinion Research (LORE) at University of Gothenburg, Sweden.

Completion rates at different stages of the field work

The field period for the Citizen Panel 4 lasted from the 26th of March to the 16th of April (i.e. three weeks), and collected a total of 7 297 completed surveys. The total number of people that was sent a link to the survey through an e-mail was 11 597, of which 244 bounced and could not be delivered. This leaves us with 11 353 people that received the survey, and thus a completion rate of 64 percent. Two things are worth noting, however: Firstly, the Citizen Panel does not use any incentives for the respondents, and secondly, the Citizen Panel is a fairly comprehensive and demanding survey.

The e-mails containing the survey links were sent out at 6 am on Monday the 26th of March. Figure 1 below displays which day, after the survey dispatch, the respondents complete the survey. What is striking about figure 1, is that about half of the respondents (more precisely, 48.3 %) completed the survey within the first field day. We can also quite clearly discern the effect of the three reminders that were sent out to the respondents on day three (Wednesday week 1), day eight (Monday week 2) and day sixteen (Tuesday week 3). The reminders seemingly have quite a substantial effect on the completion rate.

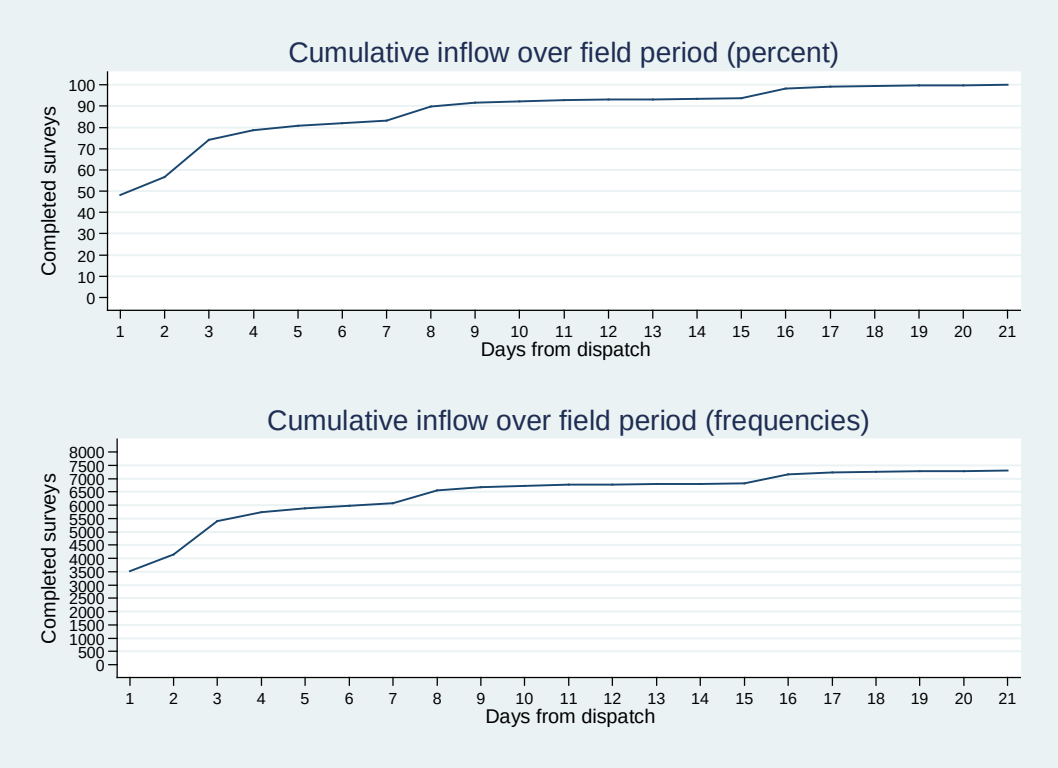
Figure 1. Survey inflow per field day



***Comment*: The upper part of the graph displays the number of completed questionnaires for each day during the three week period of field work. The lower part displays the percentage of the total number of completed questionnaires that was received during field day.**

For an inverse take on the same data, figure 2 shows the cumulative inflow over the field period. Figure 2 makes it clear that although a majority of the respondents answer within the first week, approximately 2300 respondents completes the survey in the second or third week. We also notice a typical example of the survey law of diminishing returns, as the effect of reminders on the inflow steadily decreases with the number of reminders. Even the third and last reminder, however, does yield a significant increase in the number of completed questionnaires on the two next days.

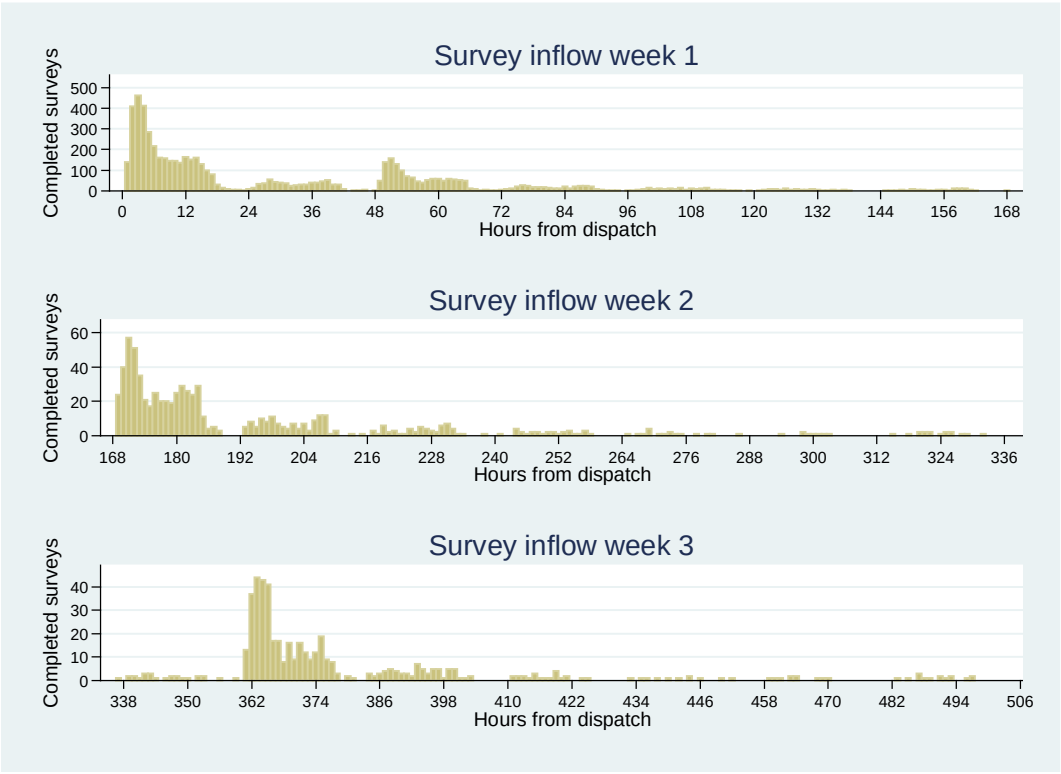
Figure 2. Cumulative survey inflow over field period



***Comment*: The upper part displays the cumulative percentage of the total number of completed questionnaires for each day of the three week period of field work, while the lower part displays the cumulative number of completed questionnaires achieved for each day.**

The next step is to examine in further detail, hour by hour, the flow of completed questionnaires. In figure 3 the number of completed surveys hour by hour (in units of hours from the time of dispatch) during the first, second and final week of field work is shown. Figure 3 makes it evident that the very first hours after dispatch are crucial for the completion rate. The strongest response is received until lunchtime on the first day. As was noted above, the flow of completed surveys is declining very quickly after dispatch, and each reminder. Approximately three days after the second reminder in the beginning of week 2, we observe that the flow of completed questionnaires has almost completely stopped until the third and last reminder in week 3, which produce a temporary boost in completed questionnaires during two days.

Figure 3. Survey inflow by week

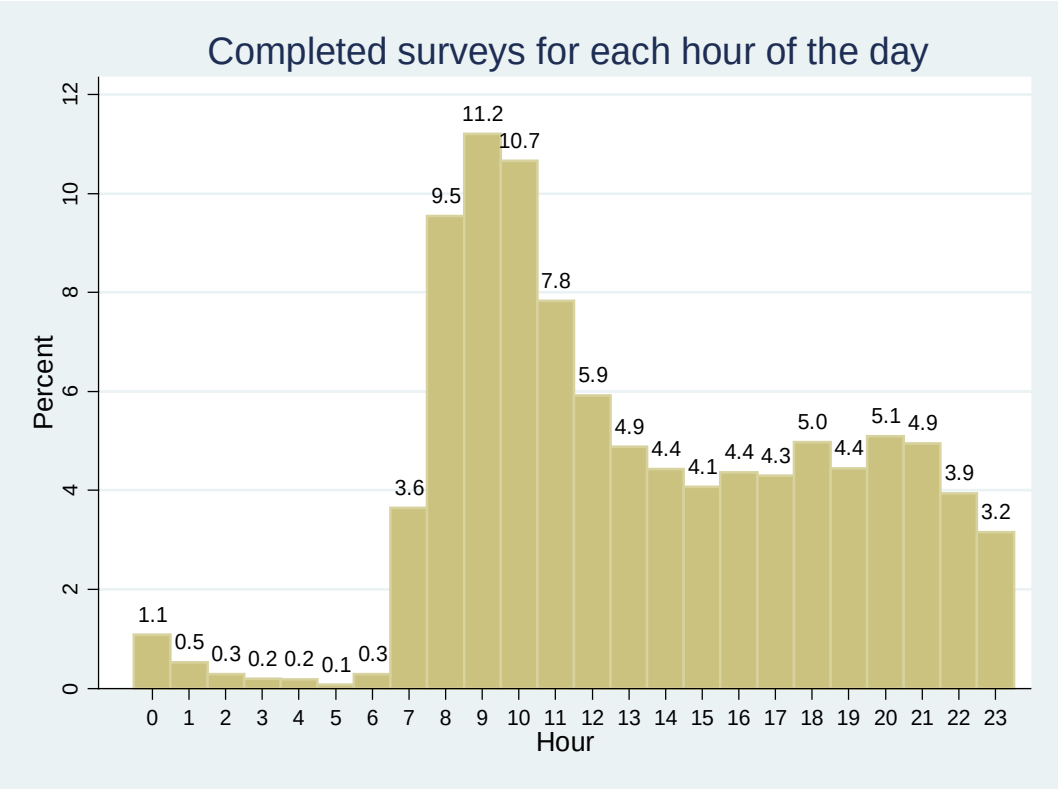


***Comment*: This graph displays the number of completed questionnaires hour by hour for the three week period of field work.**

Lastly this section of the report examines at what time of day respondents completed their surveys. Figure 4 shows the share of all completed surveys that are completed at different hours, averaged over the entire period of field work.

As seen in figure 4, it is evident that most respondents take their survey between 8 and 11 am, presumably indicating that respondents tend to take the survey during their first hours at work. Subsequently, it decreases steadily until about 6 pm, where it starts rising a little again. After 10 pm we notice slowly declining numbers again. At midnight however, the flow of responses stops suddenly, although not completely.

Figure 4. Completed surveys for each hour of the day (percent of total replies)



Comment: This graph illustrates at what time of day the respondents answered the survey. The bars indicate the percentage of the total number of the questionnaires received that was completed at that hour. For example: 1.1 percent completed their questionnaire between midnight and 1 am, while 0.5 percent completed their questionnaire between 1 am and 2 am.

Survey duration

How much time do our respondents actually spend taking the survey? As we can see in table 1 below, this varies greatly. The respondent who took the longest in filling out the survey spent approximately 3 821 minutes (more than two and a half days) filling out the survey. However, this is obviously the case of somebody who has started filling out the survey, stopped and completed it much later. Such extreme durations do not reflect the actual time necessary for respondents to complete the survey. Therefore, we have excluded outliers from the following analyses and focus only on those respondents who spent 90 minutes or less filling out the survey. This reduces our N with 182 respondents from 7297 to 7115 (thus including 97.5% of all respondents). In addition, the standard deviation decreases substantially with approximately 46 minutes from 60 minutes to 14. Hence, the dispersion of the data is much smaller after the exclusion of outliers, but with only a very small decrease in numbers of respondents. This operation also decreases the mean with 3.8 minutes from 28.82 minutes to 25.02 minutes and the median changes from 22 to 21 minutes.

Table 1. Survey duration in Citizen Panel 4, descriptive statistics (minutes)

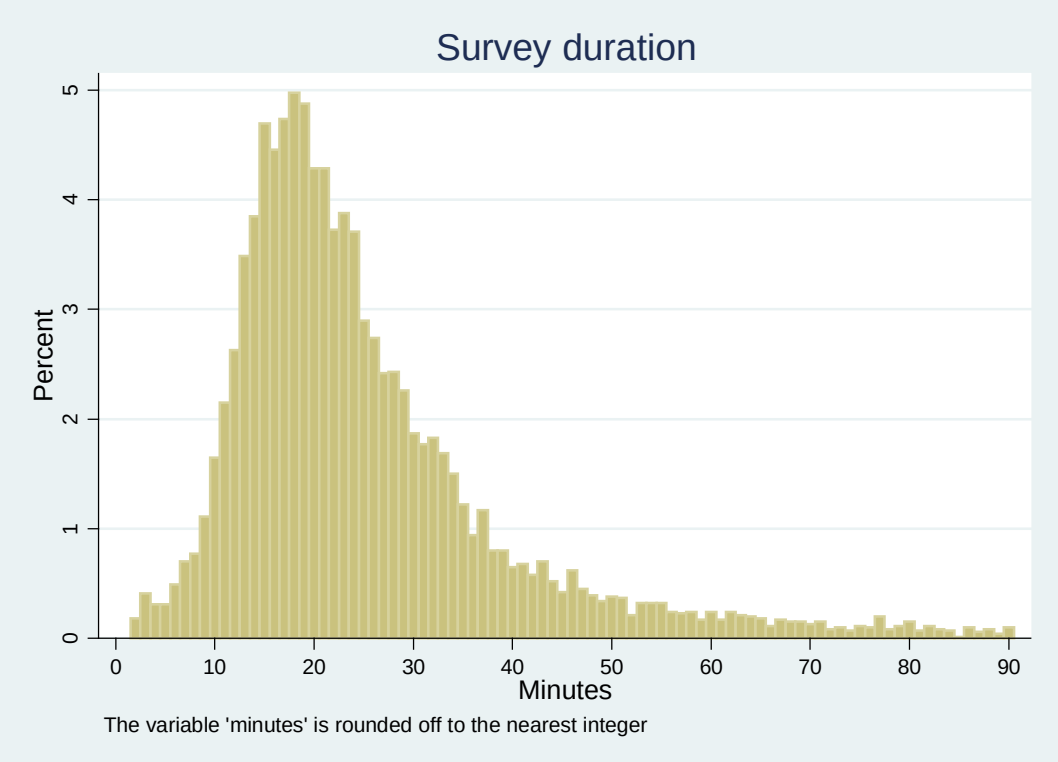
Variable	N	Median	Mean	Std. Dev.	Min	Max
Minutes	7 297	22	28.82	60.44	2	3 821
Minutes, outliers excluded	7 115	21	25.02	14.04	2	90

***Comment*: This table contains summary statistics for survey duration, measured as the number of minutes a respondent spent answering the questionnaire. Outliers were defined as those who spend more than 90 minutes completing the questionnaire.**

Generally speaking, respondents spend very different amounts of time on completing the survey. We should be aware, however, that this particular survey was not identical for all respondents. The survey included several survey embedded experiments that were assigned to sub-sets of the sample. Therefore, part of this variation derives from actual differences in survey length. Although we wish to emphasize that all versions of the survey were designed to be of about equal length.

The large variation in time spent on completing the survey is further illustrated in Figure 5 below (outliers excluded). The vast majority of respondents spend about 10-40 minutes on the survey (within one standard deviation from the mean), which seems quite reasonable, although still undeniably a fairly comprehensive internet survey.

Figure 5. Survey completion times in Citizen Panel 4



***Comment*: This graph displays the distribution of survey duration, measured as how many minutes a respondent spent answering the questionnaire. The numbers are percentages of all respondents.**

The large variation in survey duration displayed in Figure 5 leads to questions about both the sources and the consequences of this variation. Who are the respondents that answer hastily and slowly respectively, and does this affect the quality of the data? Do night owls exhibit lower response quality than larks? This is what we investigate below.

Causes of variation in survey duration

The first step is to investigate whether the time of the day affect survey completion times. The respondents are grouped into four different categories, depending on in which quarter of the day they completed the survey. Table 2 displays the results. One group in particular distinguishes itself from the rest: those respondents who completed the survey in the first quarter of the day (ie. 00-06 am) take approximately four minutes longer to complete the survey compared to other groups. The difference between this group and the other groups is the only one that is statistically significant.

Table 2. Survey duration and time of day (ANOVA)

Quarter	Mean	N	ΔQ1-Q2	ΔQ1-Q3	ΔQ1-Q4	ΔQ2-Q3	ΔQ2-Q4	ΔQ3-Q4
1 (00:00-05:59)	28.58	170	-0.364	-3.28786	-4.01175	0.352147	-0.37175	-0.7239
2 (06:00-11:59)	24.94	3 150	0.007	0.022	0.002	1	1	0.657
3 (12:00-17:59)	25.29	2 040						
4(18:00-23:59)	24.57	1 937						

Comment: **The respondents were grouped into four different quarters according to what time of day they answered (Q1 to Q4 respectively). Each group's survey duration mean was then compared to all other groups to test for any significant differences with a bonferroni post hoc-test. The mean differences between the groups are display in the table with corresponding p-values below.**

The next step is to examine whether at what day during the field period the respondents completed the survey affect the survey duration. More precisely, we categorize the respondents according to how many reminders they received before answering the survey. At most, the respondent can be reminded three times during the field period. The results from the ANOVA displayed in table 3 clearly show that the respondent's survey completion time diminishes with the number of reminders. There are no significant differences between the ones who answered without being reminded or those respondents that just where reminded one time. However, there are significant differences between those two groups and the two other groups, who received two or three reminders. The respondents who received two or three reminders completed the survey approximately two minutes faster than the other respondents.

Table 3. Survey duration and number of reminders received (ANOVA)

Reminders	Mean	N	ΔR0-R1	ΔR0-R2	ΔR0-R3	ΔR1-R2	ΔR1-R3	ΔR2-R3
0	25.36	4 144	-0.15	-1.63	-2	-1.48	-1.85	-0.37
1	25.21	1 929	1	0.023	0.025	0.092	0.073	1
2	23.73	761						
3	23.36	463						

Comment: **The respondents were grouped into three different groups according to how many reminders they received. Each group's survey duration mean was then compared to all other groups to test for any significant differences with a bonferroni post hoc-test. The mean differences between the groups are display in the table with corresponding p-values below.**

These analyses thus suggest that those who answer in the morning or earlier in the field period spend more time and perhaps think more carefully about the questions than others. Reasonably, this group is more highly motivated to take the survey.

Finally, the effect of political interest on survey completion time is also examined. As expected, those who are very interested in politics do spend a bit more time on filling out the survey, as presented in table 4 below. However, none of the differences between the four groups are significant, despite a quite substantial difference in mean survey completion time. This is due to the fact, that the Citizen Panel is a self-recruited panel, with participants that, generally, are highly interested in politics. This causes the N for the group with very low interest in politics to become quite small, thereby rendering significant mean differences difficult to obtain.

Table 4. Survey duration and political interest (ANOVA)

Political interest	Mean	N	Δvery int - fairly int	Δvery int - not p. int	Δvery int - not int	Δfairly int - not p. int	Δfairly int - not int	Δ not p. int - not int
Very interested	25.50	3 082	-0.74	-1.41	-3.76	-0.67	-3.02	-2.35
Fairly interested	24.76	2 702	0.281	0.258	1	1	1	1
Fairly uninterested	24.08	466						
Not interested	21.74	24						

Comment: **Political interest was measured as "How interested in politics are you?" The response categories are: "very interested", "fairly interested", "not particularly interested", "not interested at all".**

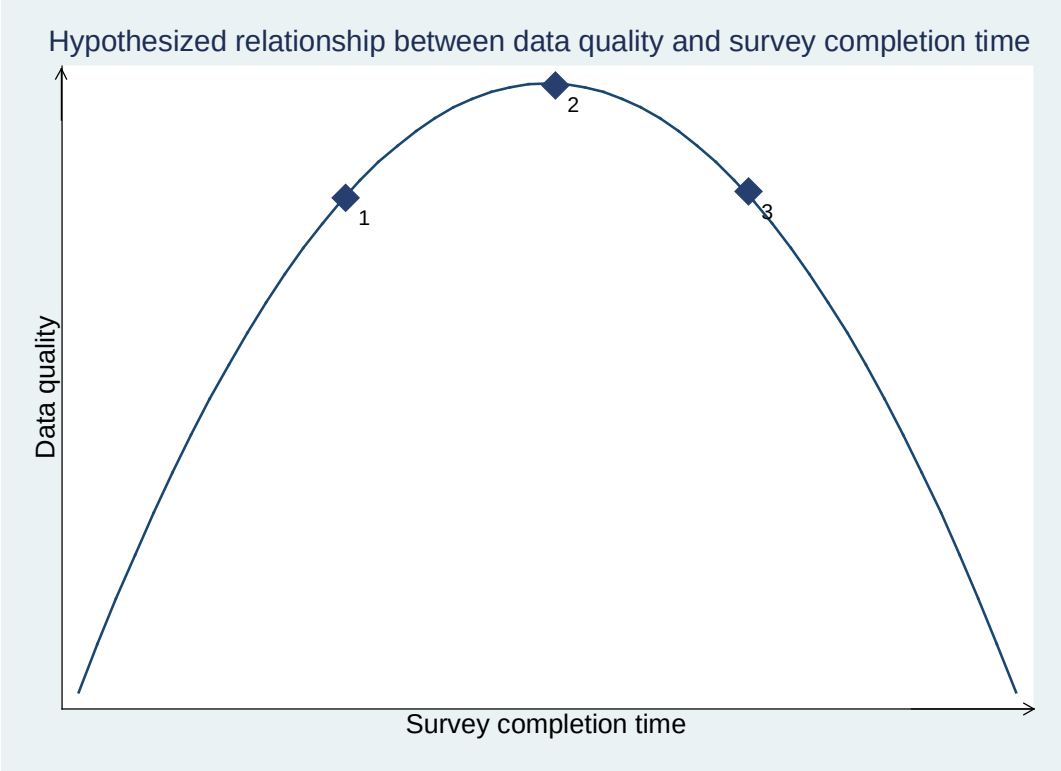
Data quality

So far, we have offered a description of the data and the field work as well as examined the survey completion time for different groups of respondents. However, the heart of the matter is how this affects the quality of the data. Do respondents with a shorter survey completion time produce data of lower quality?

The relationship between survey completion time and data quality is hypothesized to be curve linear, as illustrated in figure 6. Turning to Kroznick's theoretical framework of optimizing and satisficing, every respondent can be placed on a scale ranging from strong satisficers to optimizers, where optimizers represent the survey researcher's ideal respondent providing carefully considered answers to correctly interpreted questions and

satisficers represent respondents who try not to find the optimal answer to researcher's questions, but a satisfying one (Krosnick 1991, 1999; Krosnick et al 1996).

Figure 6. Model of the relationship between data quality and survey completion time



Comment: Illustration of hypothesized relationship between data quality and survey completion time (the amount of time a respondents uses to answer the questionnaire).

Point 1 in figure 6, is an example of a satisficer: a respondent who invests less time in answering the survey question than the average respondent, because he/she settles with a passable answer, which takes less time. This decreases the survey completion time as well as the quality of the data. Point 2 represents the optimizer, who laboriously and correctly interprets and answers the questions at a steady pace. The optimizer is often thought of as an able respondent, with relatively high cognitive ability and high motivation to complete the survey compared to satisficers. Finally, in point 3, we find the respondent who is not satisficing, but who lacks the cognitive ability of the optimizer and therefore has a harder time interpreting and answering the questions. This increases the survey completion time and affects the survey data quality negatively at the same time.

In this study we use two indicators intended to measure specific aspects of survey data quality: item missing and straight-lining. Item missing is measured as the percentage of missing answers on 72 questions (that all respondents who took survey received) in the Citizen Panel 4. Our measurement of straight-lining is slightly more complicated. First, seven question batteries were chosen. If the respondent provided the same answer to all questions in a battery (for an example "very high trust"), a battery dummy was coded as 1.

All these dummies were then added together to produce two measures of straight-lining. Straight-lining 1 was coded as 1 if the respondent had straight-lined any of the seven batteries, whilst Straight-lining 2 was coded as 1 if the respondent straight-lined two or more batteries. Table 5 presents summary statistics of these two indicators of data quality.

Table 5. Straight-lining and item missing in Citizen Panel 4, descriptive statistics

Variable	N	Mean	Std. Dev.	Min	Max
Straight-lining 1	7 297	0.28	0.45	0	1
Straight-lining 2	7 297	0.06	0.24	0	1
Item missing	7 297	0.13	0.06	0	0.93

***Comment*: Item missing is the percentage out of 72 items that respondents did not answer. Straight-lining 1 indicates that the respondent has straight-lined one or more out of seven batteries, while straight-lining 2 indicates that the respondent has straight-lined two or more of the seven batteries.**

The average share of item missing for our respondents is thirteen percent, which is generally speaking fairly high. This corresponds to about 9 items of the 72 questions included in this analysis. Our two measures of straight-lining differ substantially. The more generous measure tells us that 28 percent of our respondents straight-lined at least one of the seven question batteries we analyzed. This is a high number, but on the other hand, it is sometimes reasonable to give the same answer to all questions in fairly short batteries. Our more strict measure of straight-lining on the other hand, shows us that six percent straight-lined at least two of the seven batteries, a safer indicator of poor data quality.

Table 6 below contains details of how many respondents that straight-line and how many of the question batteries they straight-line. The vast majority of the respondents (72.40 %) do not straight-line at all, but approximately 2000 respondents straight-line at least one question battery. Worth noticing is that very few respondents straight-line three or more batteries.

Table 6. Straight-lining respondents in Citizen Panel 4 2012

Straight-lined batteries	Frequency	Percent	Cumulative percent
0	5 283	72.40	72.40
1	1 578	21.63	94.02
2	352	4.82	99.12
3	68	0.93	99.90
4	14	0.19	99.97
5	1	0.01	99.99
7	1	0.01	100.00

***Comment*: The table shows how many out of seven batteries that were straight-lined by the respondents. Straight-lining is defined as providing the same answer to all items in a question battery.**

Table 7 displays the details of the questions batteries included in the analysis of straight-lining. The battery with the highest percent of respondents straight-lining it is battery 183 (perceived authority disaster handling capacity for six different public authorities) with 12.5%, whilst only 0.2% of the respondents straight-line battery 20 (policy proposals). Presumably, relatively many respondents felt they had no opinion on the issue of authority disaster handling capacities or any knowledge of these authorities that could help them give an answer.

Table 7. Straight-lined batteries details

Question battery	Battery theme	Number of items	Percent straight-lining	Frequency straight-lining	Std. Dev.
Battery 9	Institutional trust	8	4.7	345	0.212
Battery 20	Policy proposals	8	0.2	14	0.044
Battery 22	Anxiety about individual problems	6	10.4	760	0.305
Battery 23	Anxiety about societal problems	6	2.7	199	0.163
Battery 180	Social media exposure	6	3.1	224	0.173
Battery 182	Social media participation	6	1.4	100	0.116
Battery 183	Perceived authority disaster handling capacity	6	12.5	912	0.331

***Comment*: The table shows the percent of all respondents that straight-lined the seven different batteries included in this study. Straight-lining is defined as providing the same answer to all items in a question battery.**

In the following, we analyze whether data quality is connected to the amount of time a respondent spends on completing the survey. We do this mainly by comparing five-minute intervals. Thus, respondents in interval 1 spend between one and five minutes taking the survey, while those in interval 2 spend six to ten minutes, and those in interval

18 spend between 86 and 90 minutes. In this analysis we have also excluded outliers, defined as respondents with a survey completion time of more than 90 minutes. Table 8 presents descriptive statistics for three indicators of data quality for groups of respondents with different survey duration.

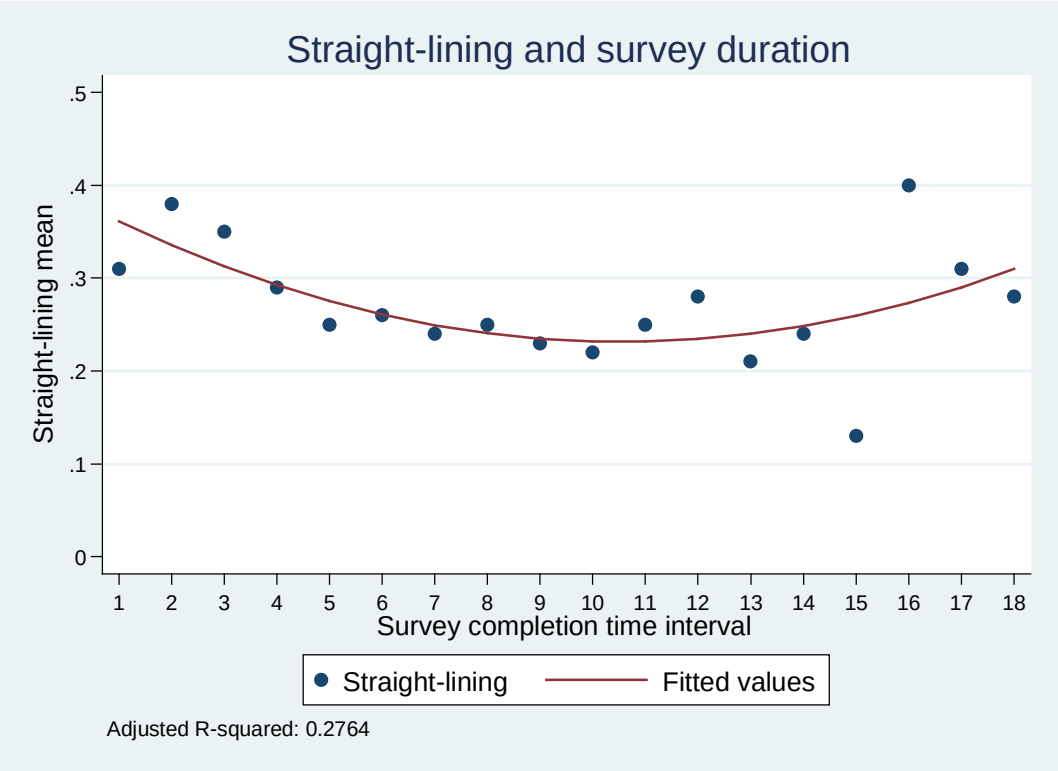
Table 8. Item missing and straight-lining mean grouped over five minute intervals.

	Item missing		Straight-lining1		Straight-lining 2	
Interval	Mean	N	Mean	N	Mean	N
1	0.129	64	0.31	64	0.05	64
2	0.140	241	0.38	241	0.10	241
3	0.133	979	0.35	979	0.09	979
4	0.127	1 689	0.29	1 689	0.06	1 689
5	0.126	1 415	0.25	1 415	0.05	1 415
6	0.123	907	0.26	907	0.06	907
7	0.126	616	0.24	616	0.04	616
8	0.128	351	0.25	351	0.03	351
9	0.120	222	0.23	222	0.05	222
10	0.137	158	0.22	158	0.03	158
11	0.128	114	0.25	114	0.04	114
12	0.133	85	0.28	85	0.05	85
13	0.124	75	0.21	75	0.04	75
14	0.128	55	0.24	55	0.04	55
15	0.134	38	0.13	38	0.00	38
16	0.122	43	0.40	43	0.09	43
17	0.126	35	0.31	35	0.09	35
18	0.116	28	0.28	28	0.04	28

Comment: **The time intervals each represents five minutes, thus running from 5 minutes or less, to 86-90 minutes at most. Item missing is the percentage out of 72 items that respondents did not answer. Straight-lining 1 indicates that the respondent has straight-lined one or more out of seven batteries, while straight-lining 2 indicates that the respondent has straight-lined two or more of the seven batteries.**

These results indicate that there is indeed find a curve linear-relationship between straight-lining and survey duration time. The relationship is illustrated more clearly in figure 7 (straight-lining 2). The relationship between straight-lining (straight-lining 2) and survey completion time is fairly strong with an adjusted R^2 of .28. This analysis indicates that the highest quality of answers (the lowest amount of straight-lining) is provided by the respondents taking 20 to seventy minutes in filling out the survey.

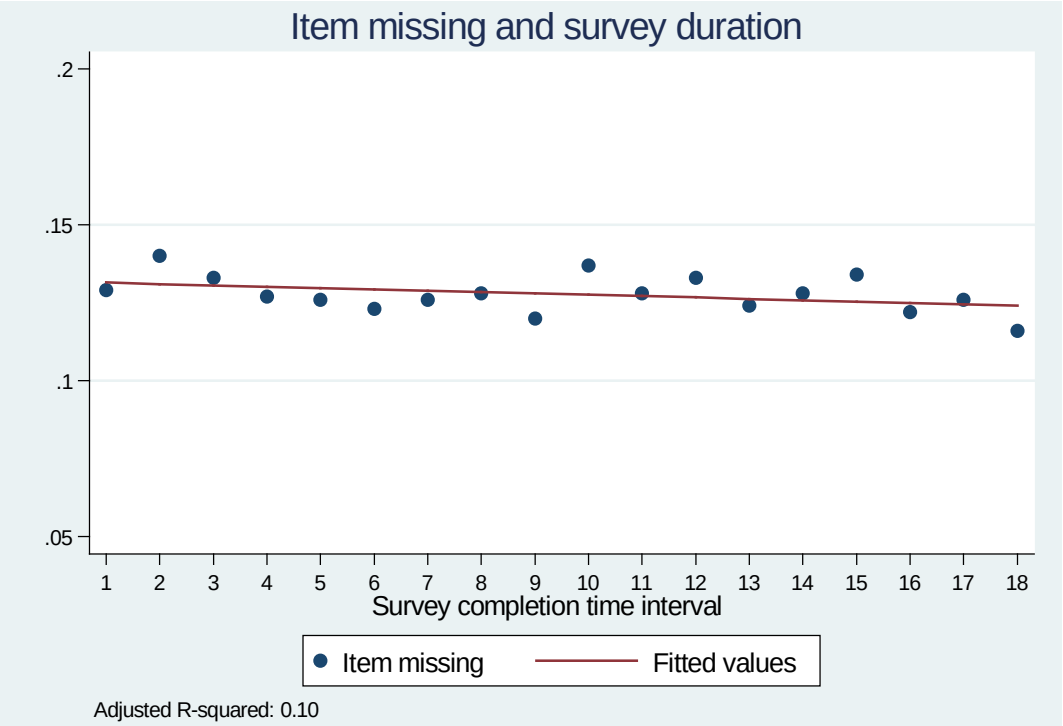
Figure 7. The relationship between straight-lining and survey completion time.



Comment: **Each time interval represents five minutes. The vertical axis measures the average out of 7 batteries that respondents have straight-lined. The relationship was examined with an OLS logistic regression. Three models were tested, one linear and two curve-linear (one quadratic and one cubic). The first curve-linear model received a significantly better fit than both other models.**

For item missing, however, our hypothesis do not receive any support as the relationship between survey completion time and item missing is negative and linear, indicating that the longer a respondent takes to complete the survey, the lower the respondent's percent of item missing will be. Still, the relationship is not very strong, with an adjusted R^2 of .10 only. This is illustrated in figure 8.

Figure 8. Correlation between item missing and survey completion time.



Comment: The time intervals represent five minutes each. The vertical axis represents the average percentage out of 72 items that respondents did not answer.

Summary and concluding discussion

This report has examined the inflow over the period of field work of a web survey, the amount of time it takes a respondent to complete the survey (duration) and a few data quality indicators. The first part shows us a very large part of the total completed questionnaires are actually received during the very first day of field work, close to 50 percent. However, after a field work period extended to three weeks, twice as many completed surveys are received in total. It is also made clear that the time of day when most people tend to answer their questionnaires is between 7 am and 11 am. These hours account for almost 40 percent of all completed questionnaires. We also discover a high amount of variation when it comes to the time respondents spend filling out the web questionnaire. On average, this particular survey took 25 minutes to complete. We also find that survey duration it is found that people who answer late at night or who are highly interested in politics spend more time on the survey than others.

When it comes to data quality the respondents of the Citizen Panel 4 seem be straight-lining to a fairly high extent. As many as one out of four respondents had straight-lined at least one out of seven question batteries. As hypothesized, we also confirm a curve-linear relationship between time spent on survey and data quality. Those filling out the questionnaire just slightly quicker than the average respondent are those who are least likely to be straight-lining.

References

Krosnick, J. A. (1991) Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5, 213-236.

Krosnick, J. A. (1999) Survey research. *Annual Review of Psychology, 50*, 537-567.

Krosnick, J. A., Narayan, S., & Smith, W. R. (1996) Satisficing in surveys: Initial evidence. *New Directions for Evaluation, 70*, 29-44

Svensk titel: Fältarbete, inflöde, tidsåtgång och datakvalitet i Medborgarpanelen 4 - 2012.

LORE working paper 2013:1

Seriens redaktör: Johan Martinsson

Författare:

Love Christensen, Fil. Kand.
Forskningsassistent
Statsvetenskapliga institutionen
Göteborgs universitet
love.christensen@gu.se

Johan Martinsson, Fil. Dr.
Forskningsledare LORE
Statsvetenskapliga institutionen
Göteborgs universitet
johan.martinsson@pol.gu.se

SAMMANFATTNING

Denna rapport undersöker en webbenkäts inflöde under dess fältperiod, tiden det tar för en respondent att avsluta enkäten (survey duration), samt ett antal datakvalitetsindikatorer.

Det första avsnittet visar att en stor del, nästan 50 %, av alla besvarade enkäter mottas under fältarbetets första dag. Efter tre veckor har dock totalt dubbelt så många avslutade enkäter mottagits. Det är också tydligt att de flesta respondenterna tenderar att besvara sina enkäter mellan klockan 08.00 och 11.00. Detta tidsintervall svarar för nästan 40 % av alla avslutade enkäter. Genomsnittstiden för att besvara enkäten var 25 minuter. När möjliga prediktorer för svarstid undersöktes framkom att människor som besvarar enkäten nattetid samt respondenter som uppger sig vara mycket intresserad av politik tillbringar mer tid med att besvara enkäten jämfört med andra respondenter.

Gällande datakvalitet tycks respondenterna i Medborgarpanelen 4 ägna sig åt ”automatiskt rätlinjesvarande” (eng. straight-lining) i tämligen hög utsträckning. Så många som en av fyra respondenter rätlinjesvarade minst ett av sju frågebatterier. Vi fastställer också ett kurvlinjärt samband mellan svarstid och datakvalitet. De som fyller i enkäten något långsammare än genomsnittsrespondenten är de som är minst sannolika att rätlinjesvara.

The Laboratory of Opinion Research (LORE) is an academic web survey center located at the Department of Political Science at the University of Gothenburg. LORE was established in 2010 as part of an initiative to strengthen multidisciplinary research on opinion and democracy. The objective of the Laboratory of Opinion Research is to facilitate for social scientists to conduct web survey experiments, collect panel data, and to contribute to methodological development. For more information, please contact us at:

info@lore.gu.se