Measuring expectations of inflation:

Effects of survey mode, wording, and opportunities to revise

Journal of Economic Psychology, in press

Wändi Bruine de Bruin[1][*]

Wilbert van der Klaauw[2]

Maarten van Rooij[3]

Federica Teppa[3]

Klaas de Vos[4]

[1] Leeds University Business School and Carnegie Mellon University

[2] Federal Reserve Bank of New York

[3] De Nederlandsche Bank (Dutch Central Bank)

[4] CentERdata

[*] Correspondence to: Wändi Bruine de Bruin, Leeds University Business School, Centre for Decision Research, Maurice Keyworth Building, Leeds LS2 9JT, United Kingdom, w.bruinedebruin@leeds.ac.uk (email)

ABSTRACT

Several national surveys aim to elicit consumers' inflation expectations. Median expectations tend to track objective inflation estimates over time, although responses display large dispersion. Medians also tend to differ between surveys, possibly reflecting survey design differences. Using a nationally representative Dutch sample, we evaluate the importance of three survey design features in explaining observed differences: mode (face-to-face vs. web), question wording ('prices in general' vs. 'inflation'), and the explicit opportunity to revise responses. We examine effects on item non-responses, revisions, reported inflation expectations and their deviation from the CPI inflation rate. We discuss implications of our findings for survey design.

Keywords: Consumer surveys, inflation expectations, mode effects, question wording

JEL codes: E31, D84

PsycInfo codes: 2229, 2260, 3920

1. INTRODUCTION

Worldwide, governments, central banks, research and financial institutions are investing resources to survey consumers about their inflation expectations. Central banks that conduct national consumer surveys about inflation expectations include the Federal Reserve Bank of New York, the European Central Bank, the Bank of Australia, the Bank of Canada, and the Swedish Riksbank (Armantier et al., 2013a; Cunningham et al., 2010). Together with market-based measures and surveys of businesses and professional forecasters, consumer surveys of inflation expectations are used by central banks to inform monetary policy decisions (Cunningham et al., 2010; Gali, 2008). That practice is supported by the finding that, despite imperfect price perceptions (see Ranyard et al., 2008, for a review), consumers' inflation expectations generally track objective estimates of realized inflation (Ang et al., 2007; Bryan & Venkatu, 2001; Hafer & Hein, 1985; Thomas, 1999), and inform consumers' actual financial behaviors in incentivized experiments (Armantier et al., 2015). Indeed, inflation expectations are relevant to people's decisions about savings, investments, purchasing durable goods, and wage negotiations – which, in turn, affect economic activity and realized inflation.

The U.S. is one of the few countries where multiple consumer surveys of inflation expectations run simultaneously, including the Conference Board's Consumer Confidence Survey, the University of Michigan's Surveys of Consumers, and the Federal Reserve Bank of New York's Survey of Consumer Expectations. These surveys show seemingly systematic variation in the reported medians and dispersion of year-ahead inflation expectations (Armantier et al., 2013a). Relatively little is known about why these surveys yield different results. Sampling procedures differ between surveys, but findings are weighted to match focal demographic characteristics of the U.S. adult population. Reported medians may also differ between surveys as a result of variations in their treatment of outliers (e.g., Curtin, 1996).

This paper focuses on the yet untested possibility that differences in survey design features may also contribute to discrepancies between the findings of consumer surveys. One survey design feature that differs across US consumer surveys of inflation expectations is administration mode. The University of Michigan's Surveys of Consumers uses phone interviews (Ludvigson, 2004), the Conference Board's Consumer Confidence Survey sends paper-and-pencil questionnaires through the mail, (Ludvigson, 2004), and the New York Fed's Survey of Consumer Expectations runs online (Armantier et al. 2013b, 2016).

A second survey design feature that differs between consumer surveys of inflation expectations is question wording. The University of Michigan's Surveys of Consumers ask about expectations for changes in "prices in general" (Curtin, 1996), and the Conference Board's Consumer Confidence Survey asks about changes in "prices" (The Conference Board, personal communication, 2015), while the New York Fed's Survey of Consumer Expectations asks about expectations for "inflation" (Armantier et al., 2013c).

A third feature that varies between surveys is whether or not participants receive an explicit opportunity to re-think and revise their answers. While surveys often allow participants to change their answers, it is typically not explicitly encouraged. However, the University of Michigan's Surveys of Consumers provide respondents who have given an inflation expectation over 5% the explicit opportunity to re-think and revise their answer by asking "Let me make sure I have that correct. You said that you expect prices to go up during the next 12 months by [x%]. Is that correct?" (Curtin, 1996). In contrast, the Federal Reserve Bank of New York's Survey of Consumer Expectations and The Conference Board's Consumer Confidence Survey do not provide this explicit revision opportunity. Due to variations in recruitment and (unpublished) outlier treatment, it has not been possible to identify the contribution of these survey design features to reported survey results.

Below, we review survey design research on administration mode, question wording, and opportunities to revise. For each of these three survey design features, we discuss the potential relevance to survey-based measures of consumers' inflation expectations. We then introduce our study, which examines the effect of these three survey design features on response patterns in a consumer survey about inflation expectations.

1.1. Effects of administration mode: Interviewer vs. web

Traditional survey research involved face-to-face or telephone interviews, which tended to yield similar findings (Shuy 2002). The few studies that did find differences between these interviewer-administered modes suggest that face-to-face interviews may yield data of better quality (Aquilino, 1994; Holbrook et al., 2003). More recently, internet surveys have become more common, raising potential concerns about mode differences. However, most changes in responding may not be due to the new mode but rather to a re-design of questions that is implemented to fit the new mode (Dillman & Christian, 2005).

A main difference between face-to-face and web surveys is the presence of an interviewer. One potential benefit of involving interviewers is that they can motivate participants and explain confusing questions (Conrad & Schober, 2000). Yet, interviewers may also influence answers in unwanted ways (Groves & Magilavy, 1986). The presence of an interviewer may lead participants to edit their answers (Chang & Krosnick, 2009; Dillman & Christian, 2005; Kreuter et al., 2008; Parks et al., 2006; Toureangeau & Yan, 2007). Indeed, participants are less likely to report socially undesirable behaviors when talking to an interviewer than when answering a self-administered survey (Dillman & Christian, 2005).

Web surveys are self-administered, allowing participants to answer questions in private and at their own pace. Web surveys therefore reduce concerns about socially desirable responding (Chang & Krosnick, 2009; Schonlau, et al., 2003; Taylor, 2000). However, there

may be concerns about non-representative internet coverage (Fricker et al., 2005). Inexperienced internet users are less likely to answer web surveys, as compared to postal surveys (Kwak & Radler, 2002; Sax et al., 2003). To recruit individuals without internet access, some panels have offered free internet, Web TVs, and training, thus alleviating concerns about sample representativeness (Schonlau et al., 2009).

Ultimately, the decision about survey mode should be based on the relative costs and benefits. As people are becoming more computer literate and gain internet access, it will become increasingly feasible to use the internet for conducting consumer surveys. Basic broadband was already available to all EU citizens in 2013, allowing for 100% internet coverage in every member state (European Union, 2015a). Faster internet is available to 98% of households in the Netherlands, where we conducted our study (European Union, 2015b). At the same time, response rates on telephone surveys appear to be falling (Curtin et al, 2005).

Yet, even among the growing numbers of internet users, mode differences could still occur. There have been no studies of mode effects in consumer surveys about inflation expectations, but studies in other domains have extensively compared telephone vs. face-to-face interviews, and web vs. mail surveys (Couper, 2011). Comparisons of self-administered online surveys with interviewer-administered surveys are relatively uncommon, but have suggested some mode differences (Couper, 2011; Fricker et al., 2005). We highlight two findings that may contribute to mode effects in surveys about inflation expectations.

First, online surveys may promote higher item response rates than interviewer-administered modes. It is possible that this is due to internet surveys having automated prompts to discourage the skipping of answers (Fricker et al., 2005; Link & Mokdad, 2005), highlighting the importance of using comparable survey designs in studies of mode differences (Dillman & Christian, 2005). If online surveys promote higher item response

rates, then even individuals who are uncertain about future inflation may provide a response, rather than choosing to skip the question. Uncertain participants tend to give higher responses (Bruine de Bruin et al., 2011a). Therefore, higher inflation expectations may be more frequently seen with online surveys than with other modes.

Second, as noted, research on mode effects has repeatedly shown that the presence of an interviewer may increase socially desirable responses (Chang & Krosnick, 2009; Kreuter et al., 2008; Toureangeau & Yan, 2007). In the context of inflation expectations, American studies have often treated responses over 5% as very high, because the US Consumer Price Index (CPI) has not been over 5% since 1990 (Bruine de Bruin et al., 2010; 2011b; 2012). If respondents are aware of the actual CPI inflation rate and of the expectations held by others, they may feel more pressure to report lower inflation expectations in the presence of an interviewer.

1.2. Effects of question wording: 'Prices in general' vs. 'inflation'

The survey design literature recommends simple question wordings, because they lead to fewer skipped questions (Bassili & Scott, 1996; Knäuper, Belli, Hill, & Herzog, 1997; Yan & Toureangaeu, 2008). However, seemingly irrelevant wording changes can influence people's answers (Bruine de Bruin, 2011; Glaser et al., 2007; Holleman, 1999; Loftus & Palmer, 1974; Rasinski, 1989). As a notable example, questions about inflation expectations can ask about 'prices in general' or 'inflation' (Bruine de Bruin et al., 2012). Although the 'inflation' wording may seem more difficult (Bruine de Bruin et al., 2012), people tend to be familiar with the term (Leiser & Drori, 2005; Svenson & Nilsson, 1986; Williamson & Wearing, 1996).

More importantly, asking about 'inflation' often yields lower and less dispersed expectations than asking about 'prices in general' (Bruine de Bruin et al., 2012). One reason

is that while some participants recognize that 'prices in general' refers to overall inflation, others think of their personal price experiences (Bruine de Bruin et al., 2010; 2012). Those who think of inflation tend to give lower responses than those who think of personal price experiences (Bruine de Bruin et al., 2010). Because individuals may purchase different products, it is reasonable to observe some dispersion in reported inflation expectations (Hafer & Hein, 1985; Ranyard et al., 2008). However, thoughts about personal price experiences tend to focus on large price changes (Bruine de Bruin et al., 2012; Bruine de Bruin et al., 2011b), which are especially salient (Christandl et al., 2011; Del Missier et al., 2016; Greitemeyer, et al., 2005; Ranyard et al., 2008, in press). Indeed, according to sampling theories of judgment and decision making, expectations about the future are often based on 'mental sampling' of past experiences from memory, with extreme experiences being easier to remember than average ones (Fiedler & Juslin, 2005; Stewart et al., 2006). As a result, participants may report expectations for 'prices in general' that seem extreme, as compared to the CPI inflation rate or other indices of overall inflation.

1.3. Effects of the opportunity to revise responses

The survey design literature has suggested that survey respondents will provide answers to questions about unfamiliar or even fictitious topics (Bishop et al., 1980; de Best-Waldhober et al., 2009). Especially people with lower levels of education may feel pressure to respond (Bishop et al., 1980, 1986). Participants who are more uncertain about their inflation expectations tend to have lower levels of education, and give responses that are more dispersed and variable over time (Bruine de Bruin et al., 2011a). They are also more likely to give responses that are seemingly high (Bruine de Bruin et al., 2010, 2011a). Perhaps as a result, the University of Michigan's Surveys of Consumers interview protocol (Curtin, 1996) requires that inflation expectations over 5% are followed up with "Let me

make sure I have that correct. You said that you expect prices to go up during the next 12 months by [x%]. Is that correct?" The likely reason for the 5% cut-off may be that the U.S.' CPI has not been over 5% for decades (Bruine de Bruin et al., 2012). If this follow-up prompt suggests to participants that their high response is seen as wrong, then they may revise their answer to a lower number more in line with the actual CPI inflation rate.

Although test takers' common intuitions are that test scores will be hurt by switching answers, research on the self-administered Graduate Records Exam (GRE) has suggested that test takers who change their answer are more likely to switch from an incorrect answer to a correct answer than the other way around (Kruger et al., 2005; Liu et al. 2015). Test takers with lower levels of ability tend to make more revisions, suggesting that they are more uncertain about their answers (Liu et al. 2015).

However, the GRE is self-administered online or on paper. It is possible that participants will feel less inclined to revise their answers in the presence of an interviewer. As noted, people seek to make a positive impression on an interviewer (Chang & Krosnick, 2009; Parks et al., 2006). The presence (vs. absence) of another person can lead to so-called 'defensive bolstering' or amplification of commitment to previously expressed beliefs (Lerner & Tetlock, 1999). If so, participants may be more inclined to revise their answers in a self-administered than in an interviewer-administered survey mode.

1.4. Research questions

Here, we report on the first experiment that systematically tested the separate and combined effects of three crucial survey design features on reported inflation expectations: (a) web vs. face-to-face administration mode, (b) 'prices in general' vs. 'inflation' question wording and (c) a revision prompt. Participants from a national Dutch sample were randomly assigned to receive our survey in one of the two modes and one of the two question wordings.

After reporting initial responses, all participants received the explicit opportunity to make revisions. Our aim was to examine effects on the observed (1) item non-response rates for the expectations question; (2) percent of participants revising their reported expectations after receiving the revision prompt; (3) the central tendency of responses and (4) the deviation of responses from the CPI inflation rate observed for the relevant period. Of special interest was the role of participants' educational attainment, because this is a demographic characteristic that tends to be related to reporting inflation expectations that are higher, more dispersed and expressed with more uncertainty (Bruine de Bruin et al., 2010, 2011a; Bryan & Venkatu, 2001).

2. MATERIAL AND METHODS

2.1. Timing of study

Our study was conducted in the Netherlands in April 2014. Over 2013, the overall CPI inflation rate had been 2.5%, which signified no change from 2012 (Statistics Netherlands, 2015). During 2014, the CPI inflation rate dropped to 1.0%, the lowest the Netherlands had experienced in more than 25 years (Statistics Netherlands, 2015). A press release by the Netherlands' Central Bureau of Statistics published on 10 April 2014 already noted this marked decrease (Statistics Netherlands, 2014). Because the overall inflation rate had not been over 5% since it reached 6.0% in 1982, and not been over 3% since it reached 3.4% in 2002 (Statistics Netherlands, 2015), responses over 5% may seem relatively high (as in Bruine de Bruin et al., 2010; 2011b; 2012).

2.2. National sample

Participants were recruited from the Longitudinal Internet Studies for the Social Sciences (LISS) panel, which is conducted by CentERdata at the University of Tilburg

(www.lissdata.nl). The panel reflects a true probability sample of households drawn from the national population registers. If needed, households were provided with a computer and internet connection. Their demographic information was collected at their entry to the panel, and updated regularly. After joining the panel, participants received monthly invitations to complete online surveys, including the one presented in this paper.

As recommended, we recruited all participants for our survey experiment in the same way (Fricker et al., 2005). In February 2014, 4310 panel members received an electronic invitation that asked whether they would be willing to complete our questions in a self-administered web survey or face-to-face interview, with the mode to be determined at random. Specifically, they were told: "An international team of researchers will conduct a study in the LISS panel to examine how people make financial decisions. For this study we are looking for people who, for this one occasion, are also willing to talk to an interviewer in their home. About half of those who are interested will be randomly selected for an interview at home. If you are prepared to participate and you are selected for an interview at home, we will give your phone number to TNS-NIPO.[1] The TNS-NIPO interviewer will make an appointment with you and visit you at the scheduled time to conduct the survey. The survey will take 20 minutes. Content will cover (dealing with) finances, lifestyle, and expectations for the future. The compensation is €15. Following LISS panel custom, this amount will be added to your bank account." They then indicated whether they would participate, with response options being "yes," "no," or "maybe but I want more information." In March 2014, "maybe" responders received more information and a help desk phone number. Subsequently, they received a second request to indicate whether or not they would be willing to participate.

Of the 4310 panel members who were originally contacted, 3392 responded 'yes' or 'no' to our invitations (78.7%). Of those 3392, 1539 agreed to participate (45.4%). Among

the 4310 who were invited, the 1539 who agreed to participate were similarly likely to report low income, defined as below the median of €1651 after taxes per month, (50.8% vs. 52.3%), $\chi(1)=.89$, p=.35. Compared to those who did not sign up, the 1539 who did were slightly less likely to lack a college education (60.6% vs. 64.4%), $\chi(1)=6.07$, $p=.01$, be somewhat older ($M=54.2$, $SD=16.4$ vs. $M=51.7$, $SD=16.0$), $t(4308)=-4.87$, $p<.001$, more likely to be male (53.7% vs. 45.4%), $\chi(1)=26.88$, $p<.001$.

2.3. Online survey procedure

The 1539 who signed up were randomly assigned to the web ($n=769$) or to the face-to-face mode ($n=770$). Questions were worded and designed in the same way for each mode, so that any differences in responses would reflect mode differences rather than question differences (Dillman & Christian 2005). The web mode was administered through the LISS panel. The face-to-face mode was administered by trained interviewers from TNS-NIPO. For the purpose of another mode effects study, participants were asked to self-report financial behaviors and outcomes. Within each mode, participants were also randomly assigned to receiving questions about 'prices in general' or 'rate of inflation,' which were analyzed here. Appendix A shows the 'prices in general' wording, which was adapted from the University of Michigan's Surveys of Consumers interview protocol (Curtin, 1996). The first question asked "Do you think that prices in general will increase, decrease, or stay the same over the next 12 months?" The next question depended on participants' answers. Participants who answered "increase" or "decrease" were asked to indicate a percentage for that change. Participants who answered "the same" were asked to clarify whether they meant that prices would go up at the same rate, or would not go up during the next 12 months. Those who answered "increase at the same rate" were asked by what percent. All participants then received the opportunity to revise their answer by being asked "I would like to make sure I understood

your answer. You said that you expect prices to go [up/down] by [x] percent. Is that correct? An answer of "no" would then trigger a repeat of the original question.

Following the survey experiment by Bruine de Bruin et al. (2012), questions about 'inflation' followed the same pattern, so that the only difference with questions about 'prices in general' pertained to the wording. Appendix B shows the full set of questions for the 'inflation' wording. Participants were first asked whether they thought there would be "inflation, deflation (the opposite of inflation), or neither" during the next 12 months. Participants who answered "inflation" or "deflation" were asked to indicate a percent change. Participants who answered "neither" were asked to clarify whether they thought the inflation rate would be the same, or be zero, over the next 12 months. Those who answered "the same rate" were asked to indicate a percent change.

All participants received the opportunity to revise their answer, even though the University of Michigan's Surveys of Consumers interview protocol only offers it to respondents who give answers over 5% (Curtin, 1996): "I would like to make sure that I understood your answer. You said that you expect [inflation/deflation] to be [x] percent during the next 12 months. Is that correct?" We recorded initial as well as revised answers. As a result, the effect of the opportunity to revise could be examined by comparing responses provided initially and after the revision prompt, while also allowing the examination of revisions made by respondents who initially gave responses over 5% (vs. not).

2.4. Effects of mode on survey participation rates

Table 1 shows that random assignment was successful, such that there were no statistically significant differences between those who were assigned to the web mode and those who were assigned to the face-to-face mode, in terms of college education, income, gender, and age (each $p>.05$). More importantly, there was no significant mode difference in

the survey participation rates among those who had indicated agreement to participate. As seen in Table 1, out of the individuals who answered 'yes' to our invitation, the percent who actually started the survey was similar in the web and the face-to-face mode. Moreover, the individuals who ended up participating in the two modes showed no significant differences in terms of their education, or the other demographic measures for income, gender and age (p>.05).

Our subsequent analyses did not use statistical weights to correct for the small deviations in demographic composition between the overall contacted sample (n=4310) and the sample of individuals who participated *and* responded to the expectations question in each of the four conditions (web vs. face-to-face mode x 'prices in general' vs. 'inflation' wording). When applying such weights, we found no effect on the reported medians, and only small effects on reported means and standard deviations (results available from the authors upon request). Hence, the overall pattern of results and conclusions were unaffected by non-participation and item non-response rates.

3. RESULTS

3.1. Effects of mode and wording on item non-responses for the expectation question.

Participants were most likely to give no response to the 'inflation' question in the 'face-to-face' mode (Table 2, left panel). Logistic regressions confirmed that there were significantly more non-responses for the 'inflation' wording and the face-to-face mode, both before and after taking into account demographic variables (Table 3, left panel). However, the logistic regressions could not compute interactions with mode or wording due to two cells having 0% non-responses, and one cell showing only .3% non-responses (Table 2, left panel). There was no significant relationship of item non-responses with whether or not participants had a college education (Table 3, left panel).

3.2. Wording and mode effects on revisions made

Participants were most likely to make revisions to the 'inflation' question in the web mode (Table 2, right panel). Logistic regressions showed that revisions were significantly more likely with the 'inflation' wording and the web mode, both before and after taking into account demographic variables (Table 3, right panel). Interactions with mode and wording could not be computed due to the 0% revisions being observed in one cell and .6% in another (Table 2, right panel). Additionally, participants without a college education were significantly more likely than those with a college education to make revisions (4.7% vs. 1.7%; Table 3, right panel). Because the University of Michigan's Surveys of Consumers ask only participants who give initial expectations of over 5% to consider a revision, we confirmed that those who gave initial expectations over 5% were more likely to revise their responses than participants who gave lower initial expectations (19.4% vs. 2.7%; Table 3, right panel).

3.3. Effects of mode, wording, and revision opportunity on the central tendency of reported expectations.

Table 4 shows the medians and means of participants' expectations, as reported before and after they were given the opportunity to revise, by question wording and survey mode. Because associated distributions were not normal (Figure 1), Table 4 also reports the between-subjects Mann–Whitney (M–W) test and the paired-sample Wilcoxon (W) test, each of which provides a non-parametric alternative to the equivalent *t*-test for examining main effects on mean responses (Siegel & Castellan, 1988). In these analyses, the following findings emerged about the three survey design features we examined. First, mode effects systematically emerged in the means of initial and final responses for every question

wording, such that they were consistently higher on the web than face to face. Second, previously reported question wording effects on reported expectations (Bruine de Bruin et al., 2012) were seen for both initial and final responses in the face-to-face mode, and for final responses in the web mode, but not for initial responses in the web mode (Table 4). Third, final expectations were lower than initial expectations across both question wordings in the web mode, with no differences observed for either question wording in the face-to-face mode.

Next, we conducted a repeated-measures Analysis of Variance that allowed for the examination of both main effects and interactions, with a within-subjects variable for response type (initial vs. after the revision prompt), and between-subjects variables for mode (web vs. face-to-face) and wording ('prices in general' vs. 'inflation'). Across responses given before and after the revision prompt, we found a significant effect of mode, $F(1, 1369)=17.85$, $p<.001$, such that reported expectations were larger on the web than in the face-to-face mode ($M=2.36$, $SD=4.56$ vs. $M=1.50$, $SD=3.02$). Across these response types, there was no main effect of wording, $F(1, 1369)=1.77$, $p=.18$, and no significant interaction between mode and wording, $F(1, 1369)=1.97$, $p=.16$. However, final responses given after the revision prompt were significantly lower than those given initially ($M=1.81$, $SD=3.31$ vs. $M=2.07$, $SD=5.18$), $F(1, 1369)=5.67$. $p=.02$, with a significant interaction between response type and mode, $F(1, 1369)=5.81$, $p=.02$, indicating that final responses were larger than initial responses on the web ($M=2.10$, $SD=3.55$ vs. $M=2.63$, $SD=6.62$) as compared to the face-to-face mode, where no observable change occurred between revisions and initial responses ($M=1.50$, $SD=3.02$ vs. $M=1.50$, $SD=3.02$). We found no significant interaction between response type and question wording, $F(1, 1369)=3.66$, $p=.06$. However, a three-way interaction between response type, mode and wording, $F(1, 1369)=3.78$, $p=.05$, indicated that

final responses were much larger for the 'inflation' wording on the web than for any other wording-by-mode condition (Table 4).

Table 5 shows that the reported mode effects held in separate linear regressions on initial expectations and on final expectations, before and after taking into account demographic differences. Wording effects emerged in final expectations, in models with and without demographics, such that the 'prices in general' wording yielded higher responses (Table 4). There was no significant interaction between mode and wording (p>.05 with or without demographics; not shown). In the models that included demographics (Table 5), we found that participants without (vs. with) a college education gave significantly higher expectations before the revision prompt (M=2.40, SD=6.35 vs. M=1.54, SD=2.30), with that gap being less pronounced in final expectations provided after the revision prompt (M=1.99, SD=3.81 vs. M=1.51, SD=2.28). There were no other significant two-way or three-way interactions between mode, wording, and whether or not participants had completed a college education (p>.05; not shown).

3.4. Effects of mode, wording, and revision opportunity on the deviation of reported expectations from the CPI inflation rate

Table 4 shows two measures of dispersion, including standard deviations of reported expectations and the mean absolute deviation from 1%, which was the Netherlands' CPI inflation rate over 2014 (Statistics Netherlands, 2015). We focus on the latter, because such mean absolute deviation from a specific value is less likely than standard deviations to be affected by the skewness or the outliers of the distribution (Conover et al., 1981). Moreover, this measure summarizes across deviations observed for individual responses, which can be used in tests of group differences (Conover et al., 1981; see also Bruine de Bruin et al., 2011b). In the Online Supplemental Materials, we present additional analyses on the mean

absolute deviation from the median, and the percent of responses over 5%, which yield similar patterns of results.

Mann-Whitney tests were conducted on absolute deviations from the CPI inflation rate of 1%, allowing for examination of main effects of the three focal survey design features (Table 4). First, the web mode evoked more dispersion than the face-to-face mode, for both initial and final expectations and both question wordings. Second, we found more dispersion for 'prices in general' than for 'inflation' wording in initial and final responses within each mode, with one exception. That is, the opposite pattern was seen in initial expectations as reported in the web mode, with dispersion being higher for 'inflation' than for 'prices in general.' Third, the revision prompt reduced the dispersion of expectations for both question wordings in the web mode, but not in the face-to-face mode for either question.

Next, we conducted a repeated-measures Analysis of Variance that allowed for the examination of both main effects and interactions on dispersion from the CPI inflation rate, with a within-subjects variable representing response type (initial vs. after the revision prompt), and between subject-variables for mode (web vs. face-to-face) and wording ('prices in general' vs. 'inflation'). Across responses given before and after the revision prompt, we found a significant effect of mode, $F(1, 1369)=13.72$, $p<.001$, such that responses were more dispersed on the web than face to face ($M=2.04$, $SD=4.30$ vs. $M=1.32$, $SD=2.76$). Across these response types, there was no main effect of wording, $F(1, 1369)=.10$, $p=.75$, and no significant interaction between mode and wording, $F(1, 1369)=.59$, p=.44. However, final responses given after the revision prompt were significantly less dispersed than initial responses ($M=1.55$, $SD=3.03$ vs. $M=1.81$, $SD=4.97$), $F(1, 1369)=5.72$, $p=.02$, with a significant interaction between response type and mode, $F(1, 1369)=5.79$, $p=.02$, indicating that dispersion was lower when comparing revisions to initial responses on the web ($M=1.78$, $SD=3.26$ vs. $M=2.30$, $SD=6.41$) as compared to the face-to-face mode, where dispersion was

similar for final and initial responses (M=1.32, SD=2.76 vs. M=1.32, SD=2.76). There was no significant interaction between response type and question wording, $F(1, 1369)$=3.66, p=.06. However, a three-way interaction between response type, mode, and wording, $F(1, 1369)$=3.72, p=.05, indicated that final responses showed less dispersion as compared to initial responses for 'inflation' wording on the web than for any other wording-by-mode condition (Table 4).

Table 5 also shows a significant mode effect in linear regressions on absolute deviation from the CPI inflation rate, as seen in initial and final responses, before and after taking into account demographic differences. Unlike previous studies (e.g., Bruine de Bruin et al., 2012), we found no significant effect of question wording on the dispersion of either the initial or final expectations. We also found no significant interaction effects between mode and wording, in dispersion observed before or after the revision prompt (p>.05 with or without demographics; not shown).

After adding demographics to the models (Table 5), we found that deviations from the CPI inflation rate were significantly larger among participants without (vs. with) a college education, both in initial responses (M=2.20, SD=6.12, M=1.21, SD=2.03) and final expectations (M=1.78, SD=3.51 vs. M=1.19, SD=2.01). In initial responses, we found a significant interaction between college education and mode (B=1.26, t=2.28, p=.02), such that individuals without a college education gave relatively higher responses on the web than face to face (M=2.94, SD=8.04 vs. M=1.47, SD=3.10) as compared to individuals with a college education (M=1.32, SD=1.96 vs. M=1.09, SD=2.10). This interaction was reduced after the revision prompt (B=.81, t=1.72, p=.09) due to relatively more similar gaps between web and face-to-face responses for individuals without a college education (M=2.10, SD=3.87 vs. M=1.47, SD=3.10) and those with a college education (M=1.29, SD=1.91 vs. M=1.09, SD=2.10).

There were no other significant interactions between mode, wording, and whether or not participants had completed a college education ($p>.05$; not shown).

4. DISCUSSION

A growing number of national household surveys in different countries have been following consumers' inflation expectations. Median responses tend to be in line with realized inflation estimates (Ang, Bekaert, & Wei, 2007; Hafer & Hein, 1985; Thomas, 1999; Christensen, Van Els, & Van Rooij, 2006). However, responses have revealed relatively large disagreement between respondents and positively skewed distributions (Bates & Gabor, 1986; Bruine de Bruin et al., 2010; Bryan & Venkatu, 2001). The central tendency and dispersion of reported findings have also tended to differ between surveys, perhaps because of differences in their survey design features.

Here, we presented the findings of what we believe to be the first study that systematically tested for effects of the administration mode (web vs. face-to-face) on reported expectations. We also examined the effect of additional survey design features that vary across consumer surveys about inflation expectations: Question wording ('prices in general' vs. 'inflation') and the opportunity to revise responses. To avoid the need to adjust for sample differences, we followed Fricker et al.'s (2005) recommendation to recruit all participants from the same sample. Specifically, we selected members of a Dutch internet panel who consented to receiving our questions in either mode, and who ended up being equally likely to participate in either mode. They were randomly assigned to mode and question wording. To avoid the criticism that mode effects often occur due to different question designs being used in different modes (Dillman & Christian, 2005), we used the same question designs in each mode. We report on four main findings.

Our first main finding is that item non-response rates were low in both modes for both question wordings, but slightly higher for the 'inflation' question in the face-to-face mode. Previous research found that item non-response rates were similar for both question wordings, on a web survey conducted with RAND's American Life Panel, where participants were used to receiving regular questions about inflation and were discouraged from skipping questions by automatic prompts to please provide an answer (Bruine de Bruin et al. 2012). Yet, that study did report that participants found the 'inflation' wording somewhat more difficult than the 'prices in general' wording (Bruine de Bruin et al., 2012). It is possible that participants in our survey were slightly less willing to answer the more difficult 'inflation' question in front of an interviewer, due to concerns about giving the wrong answer. Yet, missing responses were not more likely among participants who had no college education, who tend to feel more uncertainty about what inflation expectations to report (Bruine de Bruin et al., 2011a). It is also possible that interviewers were more permissive about non-responses for the 'inflation' question, due to perceiving the 'inflation' question as more difficult. Indeed, even trained interviewers can inadvertently influence participants' motivation to provide an answer (Conrad & Schober, 2000), and make mistakes when deviating from standard protocols (Groves & Magilavy, 1986).

Second, we found systematic mode effects on reported expectations and their deviation from the Netherlands' 2014 CPI inflation rate of 1% (Statistics Netherlands, 2015). Specifically, the face-to-face mode resulted in somewhat lower reported expectations than the web mode, which were also more accurate in terms of showing less deviation from realized CPI inflation (Table 4). There were no significant mode differences in participation rates or the percent of participants with a college education (Table 1) that could have explained this result. However, participants were significantly more likely to skip the 'inflation' question in the face-to-face mode, as compared to other combinations of wording and mode. Previous

research has suggested that participants perceive the 'inflation' wording as somewhat harder than the 'prices in general' wording (Bruine de Bruin et al., 2012), which may make uncertain participants more concerned about giving a 'wrong' answer in front of an interviewer. Skipping the question provides a strategy for uncertain participants to comfortably opt out. Uncertain participants who do answer the question tend to report higher and more dispersed expectations (Bruine de Bruin et al., 2011a), perhaps explaining differences in responses between 'inflation' questions in the face-to-face mode, as compared to the other conditions. Thus, survey design conditions that encourage non-responses may not represent the full range of participants' expectations, leading to artificially lower and less dispersed responses. Future research should therefore examine the usefulness of prompts to discourage the skipping of answers, which are already common in web surveys (e.g., Fricker et al., 2005; Link & Mokdad, 2005).

Additionally, the mode difference in the central tendency and dispersion of expectations may also reflect the finding from the survey design literature that the presence of an interviewer, as in the face-to-face mode, may increase socially desirable responses. If participants were aware of realized inflation and the inflation expectations of most others, the presence of an interviewer may have reduced their use of extreme responses. However, the reported mode effects on deviation from the CPI inflation rate were somewhat more pronounced for participants without (vs. with) a college education, who may have been less likely to know the CPI inflation rate. It is therefore also possible that interviewers gave implicit or explicit cues about the appropriateness of reporting high inflation expectations, especially to participants without a college education who seemed more uncertain.

Third, we found that revisions were most likely to be made for the 'inflation' question in the web survey, resulting in lower and less dispersed expectations after the revision prompt. Possibly, this reflects the 'inflation' question being perceived as somewhat more

difficult (Bruine de Bruin et al., 2012), and being answered with more uncertainty. Participants may have also been less willing to make revisions in front of an interviewer as a result of what psychologists refer to as 'defensive bolstering' (Lerner & Keltner, 1999). Indeed, people express stronger beliefs when they feel that another person might be judging them (Lerner & Keltner, 1999). Across modes, revisions were most likely to be made by participants without a college education and who gave initial expectations over 5%, who also tend to feel the most uncertain about which inflation expectations to report (Bruine de Bruin et al., 2011a).[2] Yet, revisions were also made by participants with a college education and those giving responses below 5%. From a survey design perspective, if the opportunity to revise is provided, it is therefore important to (a) give *all* participants the opportunity to revise rather than just a sub-set, so that all responses are comparable in the sense of being generated through the same survey design; (b) report on expectations reported before and after the opportunity to revise, so as to understand the effect of this survey design feature on responses.

Fourth, we found question wording effects on expectations that emerged in non-parametric Mann-Whitney tests, but not in parametric Analyses of Variance, which may be due to the skewness of response distributions. Where wording effects emerged, participants had higher median and mean responses, as well as greater dispersion in expectations for "prices in general" than in expectations for "inflation". Previous web studies had also found that the 'prices in general' wording lead to more disagreement about how to interpret the question, with some participants focusing on personal experiences with extreme price increases, while others thinking of the 'inflation rate' (Bruine de Bruin et al., 2012). Our findings indicate that this effect of question wording is similar in web and face-to-face modes, suggesting that the presence of the interviewer did not help promote a more consistent interpretation of the two question wordings.

Like any study, ours had limitations. As noted, our study was conducted with a Dutch internet panel. Internet panels tend to have relatively high response rates, perhaps due to members having an ongoing relationship with the research organization. Response rates would likely have been lower, and varied more between modes, if participants had been recruited through other means. Second, our study was conducted at a time of historically low inflation (Statistics Netherlands, 2014, 2015). Effects of administration mode, and perhaps especially question wording, may be more pronounced when the actual inflation rate is higher, and when the variability in price changes is higher. Indeed, the 'prices in general' wording tends to focus participants' attention on higher and more prices (which are more salient), while the 'inflation' wording highlights overall inflation estimates (Bruine de Bruin et al., 2012). Thus, future research should examine how the reported effects of survey design features change as prices vary over time.

Overall, our findings suggest that mode and wording differences may influence responses to consumer survey questions about inflation. The 'inflation' wording, as asked in the face-to-face mode, produced expectations that were the lowest and closest to the CPI inflation rate. Although these findings suggest that asking 'inflation' questions in the face-to-face mode will yield accurate responses, we are unable to conclude that these responses also had more predictive validity. To test for predictive validity, it is crucial to examine whether reported expectations correspond with actual behavior (Armantier et al., 2015). One possibility is that web-based and interviewer-assessed expectations are both relevant, but for different types of behaviors. Expectations expressed in front of an interviewer may possibly be more indicative of expectations that are used in decisions made in the presence of others. Expectations expressed on a web survey may possibly be more indicative of expectations that are used in decisions that are made alone. Similarly, it has previously been argued that questions about 'inflation' may be better at assessing long-term macro-economic

expectations, and questions about 'prices in general' may be better at assessing consumers' purchasing decisions (Bruine de Bruin et al., 2012). If so, both administration modes and question wordings may be of relevance to economists and policy makers. Which specific survey design features are implemented, should therefore depend on the research and policy goals the findings are meant to inform.

Our paper suggests that systematic differences in administration mode, question wording, and opportunities to revise could significantly affect the comparability of survey findings. In public opinion research, it has also been noted that variations in survey design systematically affect participants' responses (e.g., Dillman et al. 2009). Thus, systematic studies like the one presented here are needed to understand the effect of survey design features on responses to existing consumer surveys. To permit more informative comparisons across related surveys, it is of course important that details about mode, wording, the opportunity to revise, and other survey design features are made publicly available.

Our findings are also relevant for institutions considering a change in survey design. Some organizations may be considering a switch from in-person interviews to web-based interviews, in light of cost-effectiveness goals. The choice of mode should also be motivated by potential effects on participation rates, item non-responses, as well as responses themselves (Couper 2011, Dillman et al. 2009).

5. FOOTNOTES

1 Taylor Nelson Sofres - Nederlands Instituut voor Publieke Opinie. TNS-NIPO is a reputable survey company that is well-known in the Netherlands.

2 In our previous web surveys with RAND's American Life Panel, we found little to no use of the revision prompt. It is possible that those participants were more certain about their inflation expectations, due to being asked financial and inflation questions more regularly (Armantier et al., 2016). It is also possible that those participants have learned to anticipate reminder prompts, because they get them every time they try to skip a question

7. REFERENCES

Ang, A, Bekaert, G, & Wei, M. (2007). Do Macro variables, asset markets or surveys forecast inflation better? *Journal of Monetary Economics, 54,* 1163–1212.

Armantier, O., Bruine de Bruin, W., Topa, G., van der Klaauw, W., & Zafar, B. (2015). Inflation expectations and behavior: Do survey respondents act on their beliefs? *International Economic Review, 56,* 505-536

Armantier O., Bruine de Bruin W., Potter S., Topa G., van der Klaauw W., & Zafar B. (2013a). Measuring inflation expectations. *Annual Review of Economics, 5,* 273-301.

Armantier, O., Topa, G., van der Klaauw, W., & Zafar, B. (2013b). Introducing the FRBNY Survey of Consumer Expectations: Survey goals, design and content. http://libertystreeteconomics.newyorkfed.org/2013/12/introducing-the-frbny-survey-of-consumer-expectations-survey-goals-design-and-content.html#.Vgo-sKSFOP8. Last accessed 8 December 2016.

Armantier, O., Topa, G., van der Klaauw, W., & Zafar, B. (2016). An overview of the Survey of Consumer Expectations. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2871857. Last accessed 8 December 2016.

Aquilino, W.S. 1994. Interview mode effects in surveys of drug and alcohol use: A field experiment. *Public Opinion Quarterly, 58,* 210-240.

Bassili, J.N., & Scott, B.S. (1996). Response latency as a signal to question problems in survey research. *Public Opinion Quarterly, 60,* 390-399.

Bates, J.M., & Gabor, A. 1986. Price perception in creeping inflation: Report on an enquiry. *Journal of Economic Psychology, 7,* 269–290.

de Best-Waldhober, M., Daamen, D., & Faaij, A. (2009). Informed and uninformed public

opinions on CO_2 capture and storage technologies in the Netherlands. *International*

Journal of Greenhouse Gas Control, 3, 322-332.

Bishop, G.F., Oldendick, R.W., Tuchfarber, A.J., & Bennett, S.E. 1989. Pseudo-opinions on

public affairs. *Public Opinion Quarterly, 44,* 198-209.

Bishop, G.F., Tuchfarber, A.J., & Oldendick, R.W. 1986. Opinions on fictitious issues: The

pressure to answer survey questions. *Public Opinion Quarterly, 50,* 240-250.

Bruine de Bruin, W. (2011). Framing effects in survey design: How respondents make sense

of the questions we ask. In: Keren, G. Ed. *Perspectives on Framing* (pp. 303-324).

London, UK: Taylor Francis.

Bruine de Bruin, W., van der Klaauw, W., Topa, G., Downs, J.S., Fischhoff, B., & Armantier,

O. (2012). The effect of question wording on consumers' reported inflation

expectations. *Journal of Economic Psychology, 4,* 749-757.

Bruine de Bruin, W., Manski, C.F., Topa, G., & van der Klaauw, W. (2011a). Measuring

consumer uncertainty about future inflation. *Journal of Applied Econometrics, 26,* 454-

478.

Bruine de Bruin, W., van der Klaauw, W., & Topa, G. (2011b). Inflation expectations: The

biasing effect of thoughts about specific prices. *Journal of Economic Psychology, 32,*

834-845.

Bruine de Bruin, W., van der Klaauw, W., Downs, J.S., Fischhoff, B., Topa, G., & Armantier,

O. (2010). Expectations of inflation: The role of financial literacy and demographic

variables. *Journal of Consumer Affairs, 44,* 381-402.

Bryan, M.F., & Venkatu, G. (2001). The demographics of inflation opinion surveys. *Federal*

Reserve Bank of Cleveland Economic Commentary, 10, 1–4.

Chang, L., & Krosnick, J.A. (2009). National surveys via RDD telephone interviewing versus the internet comparing sample representativeness and response quality. *Public Opinion Quarterly, 73,* 641-678.

Christandl, F., Fetchenhauer, D., & Hoelzl, E. (2011). Price perception and confirmation bias in the context of a VAT increase. *Journal of Economic Psychology, 32*, 131–141.

Christensen, C.A., van Els, P.J.A., & van Rooij, M.C.J. (2006). Dutch households' perceptions of economic growth and inflation. *De Economist, 154,* 277-294.

Conover, W.J., Johnson, M.E., & Johnson, M.M. (1981). A comparative study of tests for homogeneity of variances. *Technometrics, 23,* 351-361.

Conrad, F.G., & Schober, M.F. (2000). Clarifying question meaning in a household telephone survey. *Public Opinion Quarterly, 64,* 1-28.

Couper, M. (2011). The future of modes of data collection. *Public Opinion Quarterly, 75,* 889-908.

Cunningham, R., Desroches, B., & Santor, E. (2010). Inflation expectations and the conduct of monetary policy: A review of recent evidence and experience. *Bank of Canada Review, Spring,* 13-25.

Curtin, R.T. (1996). Procedure to Estimate Price Expectations. Ann Arbor, MI: University of Michigan Survey Research Center.

Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly, 69,* 87-98.

Del Missier, F., Ranyard, R., & Bonini, N. (2016). Perceived inflation: The role of product accessibility and attitudes towards inflation. *Journal of Economic Psychology, 56,* 97-106.

Dillman, D.A., & Christian, L.M. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods, 17,* 30-52.

Dillman, D.A., Smyth, J.D., & Christian, L.M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method.* Third Edition. Hoboken, NJ: Wiley.

European Union (2015a). Digital Agenda for Europe: Broadband strategy and policy. https://ec.europa.eu/digital-agenda/en/broadband-strategy-policy Last accessed 8 December 2016.

European Union (2015b). Digital Agenda for Europe: Netherlands. https://ec.europa.eu/digital-agenda/en/scoreboard/netherlands Last accessed 8 December 2016.

Gali, J. (2008). *Monetary policy, inflation, and the business cycle: An introduction to the new Keynesian framework.* Princeton NJ: Princeton University Press.

Fricker, S., Galesic, M., Tourangeau, R., & Yan, T. (2005). An experimental comparison of Web and telephone surveys. *Public Opinion Quarterly, 69,* 370–392.

Glaser, M., Langer, T., Reynders, J., & Weber, M. (2007). Framing effects in stock market expectations. *Review of Finance, 11,* 325-357.

Greitemeyer, T., Schulz-Hardt, S., Traut-Matteusch, E., & Frey, D. (2005). The influence of price trend perceptions: Why the Euro seems to make life more expensive? *Journal of Economic Psychology, 26,* 541-548.

Groves, R.M., & Magilavy, LJ. (1986). Measuring and explaining interviewer effects in centralized telephone surveys. *Public Opinion Quarterly, 50,* 251-260.

Hafer, R.W., & Hein, S.E. (1985). On the accuracy of time-series, interest rate, and survey forecasts of inflation. *The Journal of Business, 58,* 377–398.

Heerwegh, D., & Loosveld, G. (2008). Face-to-face versus web surveying in a high internet-coverage population. *Public Opinion Quarterly, 72,* 836-846.

Holbrook, A.L., Green, M.C., & Krosnick, J.A. (2003). Telephone vs. face-to-face interviewing of national probability samples with long questionnaires: Comparisons of

respondent satisficing and social desirability response bias. *Public Opinion Quarterly, 67,* 79-125.

Holleman, B. (1999). Wording effects in survey research using meta-analysis to explain the forbid/allow asymmetry. *Journal of Quantitative Linguistics, 6,* 29–40.

Knäuper, B., Belli, R.F., Hill, D.H., & Herzog, AR. (1997). Question difficulty and respondents' cognitive ability: The effect on data quality. *Public Opinion Quarterly, 13,*181-199.

Kwak, N., & Radler, B. (2002). A comparison between mail and web surveys: Response pattern, respondent profile, and data quality. *Journal of Official Statistics, 18,* 257-274.

Kreuter, F., Presser, S., & Toureangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: The effect of mode and question sensitivity. *Public Opinion Quarterly, 72,*847-865.

Kruger, J., Wirtz, D., & Miller, D.T. (2005). Counterfactual thinking and the first instinct fallacy. *Journal of Personality and Social Psychology, 88,* 725-735.

Leiser, D., & Driori, S. (2005). Naïve understanding of inflation. *The Journal of Socio-Economics, 34,* 179–198.

Lerner, J.S., & Tetlock, P.E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 126,* 255-275.

Link, M.W., & Mokdad, A.H. (2005). Effects of survey mode on self-reports of adult alcohol consumption: a comparison of mail, Web and telephone approaches. *Journal of Studies on Alcohol, 66,* 239–245.

Liu. O,L,, Bridgeman, B., Lixiong, G., Xu, J., & Kong, N. (2015). Investigation of response changes in the GRE revised general test. *Educational and Psychological Measurement,* 1-19.

Loftus, E.F., & Palmer J.C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 5,* 585–589.

Ludvigson, S.C. (2004). Consumer confidence and consumer spending. *Journal of Economic Perspectives, 18,* 29-50.

Parks, K.A., Pardi, A.M., Bradizza, C.M. (2006). Collecting data on alcohol use and alcohol-related victimization: a comparison of telephone and Web-based survey methods. *Journal of Studies on Alcohol, 67,* 318–323.

Ranyard, R., Del Missier, F., Bonini, N., Duxbury, D., & Summers, B. (2008). Perceptions and expectations of price changes and inflation: A review and conceptual framework. *Journal of Economic Psychology, 29,* 378–400.

Ranyard, R., Del Missier, F., Bonini, N., & Pietroni, D. (in press). The citizen's judgments of prices and inflation. In: R. Ranyard (Ed.) *Economic Psychology: The science of economic mental life and behaviour.* London, UK: Wiley Blackwell.

Rasinski, K.A. (1989). The effect of question wording on public support for government spending. *Public Opinion Quarterly, 53,* 388-394.

Sax, L.J., Gilmartin, S.K., & Bryant, A.N. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education, 44,* 409-432.

Schonlau, M., Asch, B.J., & Du, C. (2003). Web survey as part of a mixed-mode strategy for populations that cannot be contacted by email. *Social Science Computer Review, 21,* 218-222.

Schonlau, M., van Soest, A., Kapteyn, A., & Couper, M. (2009). Selection bias in Web surveys and the use of propensity scores. *Sociological Methods and Research, 373,* 291-318.

Shih, T.H., & Fan, X. (2008). Comparing response rates from web and mail surveys: A meta-analysis. *Field Methods, 3,* 249-271.

Shuy, R.W. (2002). In-person versus telephone interviewing. In: Gubrium J.G., & Holstein, J.A. *Handbook of Interview Research: Context and Method* (pp 537–555). Sage Publications, Thousand Oaks, CA.

Siegel, S., & Castellan, N.J. 1988. *Nonparametric statistics for the behavioral sciences 2nd ed.*. Boston, MA: McGraw-Hill.

Statistics Netherlands (2014). Inflatie zakt naar 0.8%. http://www.cbs.nl/nl-NL/menu/themas/prijzen/publicaties/artikelen/archief/2014/2014-022-pb.htm. Last accessed 8 December 2016.

Statistics Netherlands (2015). Inflation; CPI, since 1963. http://statline.cbs.nl/Statweb/publication/?DM=SLEN&PA=70936eng&D1=a&D2=12,25,38,51,64,77,90,103,116,129,142,155,168,181,194,207,220,233,246,259,272,285,298,311,324,337,350,363,376,389,402,415,428,441,454,467,480,493,506,519,532,545,558,571,584,597,610,623,636,649,662,671-684&LA=EN&STB=G1,T&CHARTTYPE=3&VW=T Last accessed 8 December, 2016.

Svenson, O., & Nilsson, G. (1986). Mental economics: Subjective representations of factors related to expected inflation. *Journal of Economic Psychology, 7,* 327–349.

Taylor, H. (2000). Does internet research work? *International Journal of Market Research, 42,* 51-63.

Thomas, L.B. Jr. (1999). Survey measures of expected US inflation. *Journal of Economic Perspectives, 13,* 125–144.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133,* 859-883.

Williamson, M.R., & Wearing, A.J. (1996). Lay people's cognitive models of the economy. *Journal of Economic Psychology, 17,* 3–38.

Yan, T., & Tourangeau, R. (2008). Fast times and easy questions: The effects of age, experience and question complexity on web survey response times. *Applied Cognitive Psychology, 22,* 51-68.

Table 1: Demographic characteristics of signed-up and participating sample.

	Signed-up sample			Participating sample		
	Web mode (*N*=769)	Face to face mode (*N*=770)	Test of difference	Web mode (*N*=690)	Face to face mode (*N*=701)	Test of difference
Percent participated	89.7%	91.0%	$\chi(1)$=.76 *p*=.38	-	-	-
Percent no college	59.4%	61.0%	$\chi(1)$=.42 *p*=.52	60.4%	61.5%	$\chi(1)$=.16, *p*=.69
Percent low income [a]	50.1%	49.8%	$\chi(1)$=.02 *p*=.90	50.1%	49.1%	$\chi(1)$=.14, *p*=.71
Percent female	46.0%	46.6%	$\chi(1)$=.05 *p*=.82	45.9%	46.2%	$\chi(1)$=.01, *p*=.92
Mean age (*SD*)	54.1 (16.9)	54.3 (16.2)	*t*(1537)=-.18 *p*=.86	55.5 (16.1)	54.8 (16.8)	*t*(1389)=.80, *p*=.43

[a] Low income was defined as below the median of €1651 in income after taxes per month.

Table 2: Percent of participants providing no response and making revision (vs. not).

Wording	Percent providing no response			Percent making revision (if initial response provided)		
	Web (N=327)	Face to face (N=368)	Test of mode difference	Web (N=363)	Face to face (N=333)	Test of mode difference
Prices in general	.3%	.0%	$\chi(1)=1.13, p=.29$	3.4%	.0%	$\chi(1)=19.00, p<.001$
Inflation	.0%	5.1%	$\chi(1)=12.62, p<.001$	9.6%	.6%	$\chi(1)=26.61, p<.001$
Test of wording difference	$\chi(1)=1.11, p=.29$	$\chi(1)=19.25, p<.01$	-	$\chi(1)=10.83, p<.01$	$\chi(1)=2.34, p=.13$	-

Table 3: Regression analyses on providing item non-response and making revision (vs. not).

	Item non-response (N=1354)		Revision (N=1337)	
Web mode (vs. face to face)	.06** (.01, .42)	.05** (.01, .41)	22.77*** (5.49, 94.37)	20.94*** (5.02, 87.31)
Prices in general wording (vs. inflation)	.06** (.01, .42)	.05** (.01, .40)	.31*** (.16, .62)	.30*** (.14, .57)
No college education	-	.51 (.17, 1.48)	-	2.80* (1.20, 6.50)
Low income[a]	-	3.42 (1.03, 11.40)	-	.93 (.44, 1.97)
Female	-	1.03 (.36, 2.94)	-	1.46 (.73, 2.93)
Age	-	.98 (.95, 1.01)	-	1.00 (.98, 1.02)
Initial response over 5%	-	-	-	4.73** (2.07, 10.80)
R^2	.19	.24	.17	.25

Note: Presented numbers represent odds ratios (and associated 95% confidence intervals). Some cell counts were too low to compute interactions between mode, wording, and college education.

* $p<.05$; ** $p<.01$; *** $p<.001$

[a] Low income was defined as below the median of €1651 in income after taxes per month.

Table 4: Descriptive statistics for reported expectations by question wording and administration mode.

	Initial (before revision prompt)			Final (after revision prompt)		
	Median	Mean(SD)	Mean (SD) absolute deviation from CPI inflation rate	Median	Mean (SD)	Mean (SD) absolute deviation from CPI inflation rate
Prices in general						
Face-to-face mode	1.50	1.77[i] (3.20)	1.42[i] (2.96)	1.50	1.77[i] (3.20)	1.42[i] (2.96)
Web mode	2.00	2.41[fa] (3.53)	2.04[f] (3.20)	2.00	2.31[fi] (3.36)	1.94[fi] (3.04)
Inflation						
Face-to-face mode	1.00	1.19 (2.78)	1.21 (2.51)	1.00	1.20 (2.78)	1.21 (2.51)
Web mode	1.50	2.82[fa] (8.48)	2.53[fp] (8.30)	1.50	1.92[f] (3.70)	1.63[f] (3.44)

Note: M-W tests applied to between-group differences in means, and Wilcoxon tests to within-subject differences in means; Significance ($p<.05$) is noted for the following differences: [f]=significantly higher than for face-to-face mode; [i]=significantly higher than for inflation question; [p]=significantly higher than 'prices in general' question [a]=significantly higher than after prompt

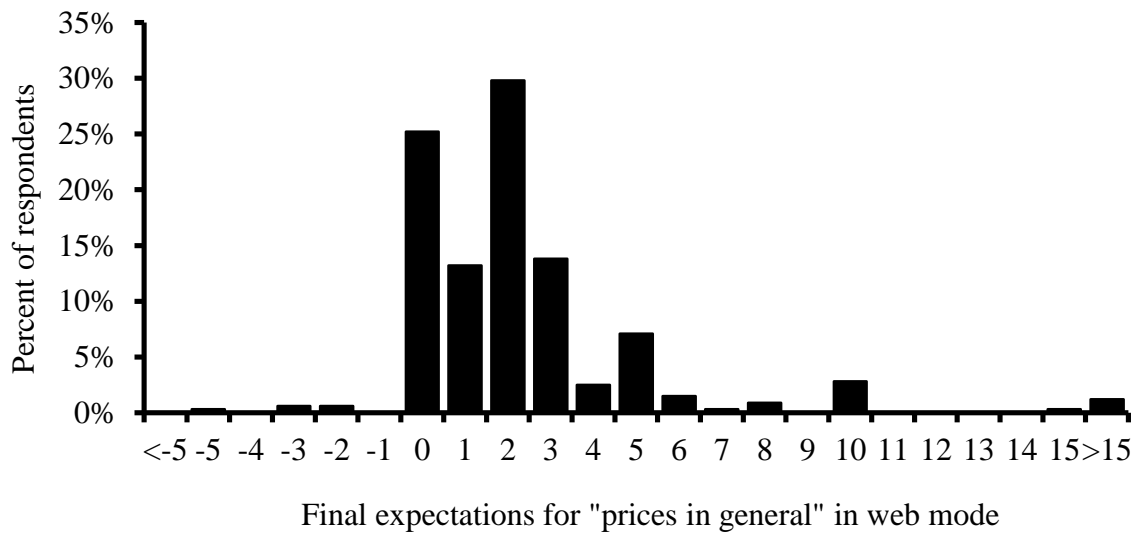Table 5: Regression analyses on expectations and dispersion.

	Initial (before revision prompt)				Final (after revision prompt)			
	Expectations (B)		Absolute deviation from CPI inflation rate (B)		Expectations (B)		Absolute deviation from CPI inflation rate (B)	
Web mode (vs. face to face)	1.15***	1.16***	.98***	.99***	.64***	.65***	.48**	.48**
Prices in general wording (vs. inflation)	.08	.03	-.14	-.20	.49**	.46**	.27	.23
No college education	-	.72*	-	.77*	-	.34	-	.38*
Low income[a]	-	.21	-	.37	-	.20	-	.35
Female	-	.58	-	.70*	-	.32	-	.45**
Age	-	.00	-	-.01	-	.00	-	.00
R^2	.01	.02	.01	.03	.01	.02	.01	.03

Note: For each model, N=1337. B is an unstandardized estimate in linear regression predicting continuous variable. $^*p<.05$; $^{**}p<.01$; $^{***}p<.001$. There were no significant two-way or three-way interactions between mode, wording, and whether or not participants had completed a college education ($p>.05$), except for a significant interaction between college education and mode in deviation of initial responses from the CPI inflation rate (B=1.26, t=2.28, p=.02).
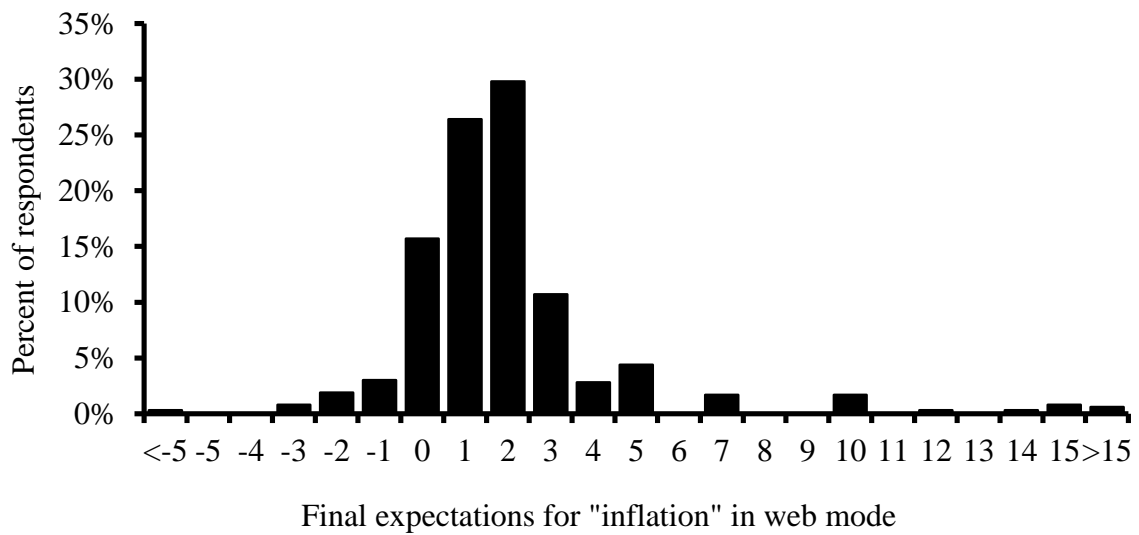[a] Low income was defined as below the median of €1651 in income after taxes per month.

Figure 1: Final expectations reported for (A) the 'prices in general' question in the web mode, (B) the 'inflation' question in the web mode, (C) the 'prices in general' question in the face-to-face mode, and (D) the 'inflation' question in the face-to-face mode

(A)

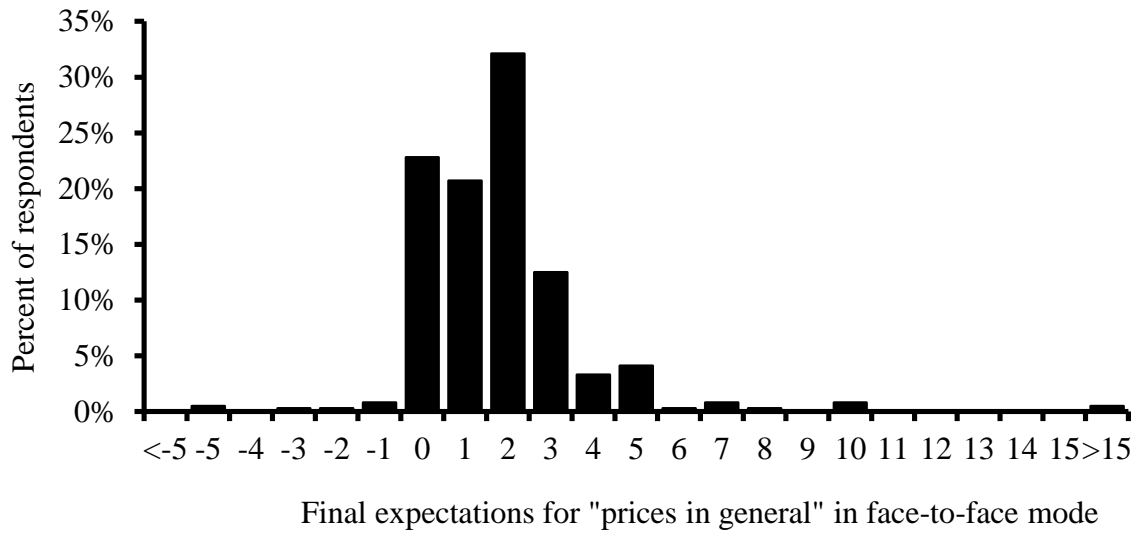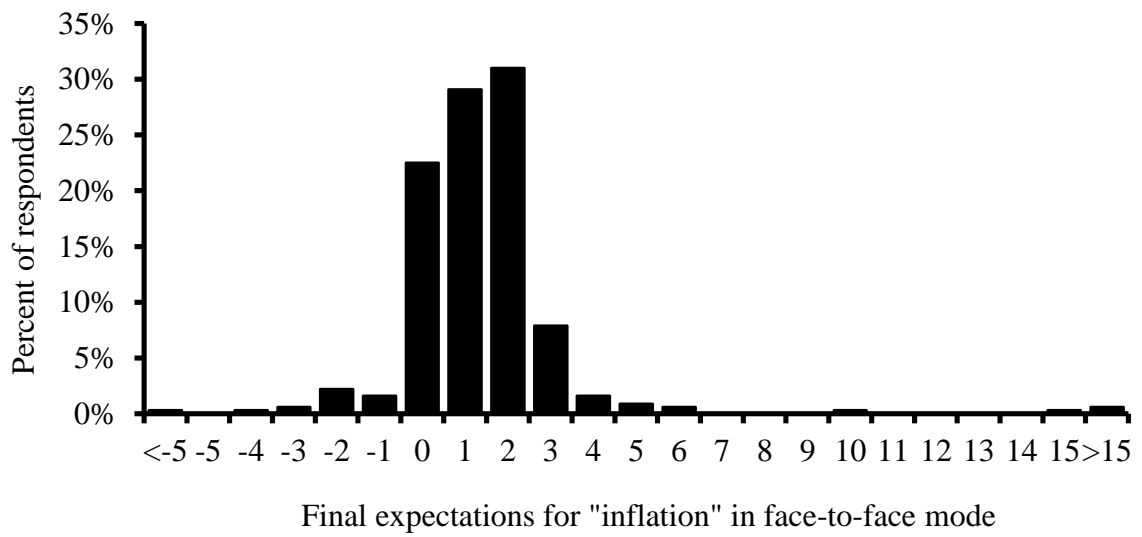

Final expectations for "prices in general" in web mode

(B)

Final expectations for "inflation" in web mode

(C)

Final expectations for "prices in general" in face-to-face mode

(D)

Final expectations for "inflation" in face-to-face mode

APPENDIX A: Protocol for 'prices in general' question wording

---

Q1a Do you think that, during the next 12 months, prices in general will go up, or go down, or stay where they are now?
___ Go up
___ Go down
___ Stay the same

[If Q1a response is ''Stay the same'']
Q1b Do you mean that prices will go up at the same rate as now, or that prices in general will not go up during the next 12 months?
___ Prices will go up at the same rate
___ Prices will not go up

[If Q1a response is ''Go up'' or Q1b response is ''Prices will go up at same rate,'' ask Q2-3 about prices going up. If Q1a response is ''Go down'' ask Q2-3 about prices going down]
Q2 By what percent do you expect prices to go [up/down] on the average, during the next 12 months?
___ percent

Q3a I would like to make sure that I understood your answer. You said that you expect prices to go [up/down] during the next 12 months by [x] percent. Is that correct?
___ Yes
___ No

[If Q3a response is ''Yes'']
Q3b By what percent do you expect prices to go [up/down] on the average, during the next 12 months?
___ percent

[For every non-response to Q1-Q3]
You skipped the question [that asked about …]. Don't you know it or do you not want to answer it?

Note: Adapted from the University of Michigan's Surveys of Consumers interview protocol (Curtin, 1996).

APPENDIX B: Protocol for 'inflation' question wording

Q1a Do you think that, during the next 12 months, there will be inflation, deflation (the opposite of inflation) or neither?
___ Inflation
___ Deflation (the opposite of inflation)
___ Neither

[If Q1a response is ''Neither'']
Q1b Do you mean that the inflation rate will be the same rate as now, or that inflation during the next 12 months will be 0 (zero)?
___ The inflation rate will be the same as now
___ The inflation will be 0 (zero)

[If Q1a response is ''Inflation'' or Q1b response is ''The inflation rate will be the same as now,'' ask Q2-3 about inflation. If Q1a response is ''deflation'' ask Q2-3 about deflation]
Q2 What percent do you expect [inflation/deflation] to be during the next 12 months?
___ percent

Q3a I would like to make sure that I understood your answer. You said that you expect [inflation/deflation] to be [x] percent during the next 12 months. Is that correct?
___ Yes
___ No

[If Q3a response is ''Yes'']
Q3b What percent do you expect [inflation/deflation] to be during the next 12 months?
___ percent

[For every non-response to Q1-Q3]
You skipped the question [that asked about …]. Don't you know it or do you not want to answer it?

Note: Equivalent to 'prices in general' protocol (Appendix A), which was adapted from the University of Michigan's Surveys of Consumers interview protocol (Curtin, 1996).

ONLINE SUPPLEMENTAL MATERIALS

S.1. Effects of mode, wording, and revision opportunity on deviation from the median

Table S1 shows the mean absolute deviation from the median, which is less likely than standard deviations to be affected by the skewness of the distribution and allows for individual-level analyses (Conover et al., 1981). Mann-Whitney tests were conducted on absolute deviations from the median, allowing for examination of main effects of the three focal survey design features. First, the web mode evoked more dispersion than the face-to-face mode, for both initial and revised expectations with both question wordings. Second, we found more dispersion for 'prices in general' than for 'inflation' wording in initial and revised responses within both modes – except for the initial responses in the web mode, which showed the opposite pattern. Third, the revision prompt reduced the dispersion of expectations across question wordings in the web mode, but not in the face-to-face mode ($p>.05$).

Next, we conducted a repeated-measures Analysis of Variance that allowed for the examination of both main effects and interactions on dispersion from the median, with a within-subjects variable representing response type (initial vs. after the revision prompt), and between subject-variables for mode (web vs. face-to-face) and wording ('prices in general' vs. 'inflation'). Across responses given before and after the revision prompt, we found a significant effect of mode, $F(1, 1369)=11.49$, $p<.001$, such that responses were more dispersed on the web than face to face ($M=1.94$, $SD=4.19$ vs. $M=1.29$, $SD=2.73$). Across these response types, there was no main effect of wording, $F(1, 1369)=.01$, $p=.95$, and no significant interaction between mode and wording, $F(1, 1369)=.83$, p=.36. However, responses given after the revision prompt were significantly less dispersed than initial responses ($M=1.49$, $SD=2.94$ vs. $M=1.75$, $SD=4.89$), $F(1, 1369)=.5.46$. $p=.02$, with a significant interaction between response type and mode, $F(1, 1369)=5.54$, $p=.02$, indicating

that dispersion was lower when comparing revisions to initial responses on the web (M=1.69, SD=3.13 vs. M=2.19, SD=6.32) as compared to the face-to-face mode, where dispersion remained the unchanged across revised and initial responses (M=1.29, SD=2.73 vs. M=1.29, SD=2.73). There was no significant interaction between response type and question wording, $F(1, 1369)$=3.56, p=.06, nor a significant three-way interaction between response type, mode and wording $F(1, 1369)$=3.62, p=.06 (Table S1).

A significant mode effect also emerged in separate linear regressions on absolute deviation from the median, as seen in initial and final responses, before and after taking into account demographic differences (Table S2). Unlike previous research (Bruine de Bruin et al., 2012), we found no significant effect of question wording on the dispersion of either the initial or final expectations. We also found no significant interaction effects between mode and wording, in dispersion observed before or after the revision prompt (p>.05 with and without demographics; not shown).

In the models that included demographics, we found that participants without (vs. with) a college education gave more dispersed expectations both before the revision prompt (M=2.14, SD=6.03 vs. M=1.12, SD=1.98) and after (M=1.74, SD= 3.41 vs. M=1.11, SD=1.96). Before the revision prompt, there was a significant interaction between mode and college education (B=1.33, t=2.44, p=.02), such that individuals without a college education gave relatively more dispersed initial responses on the web than face to face (M=2.86, SD=7.92 vs. M=1.44, SD=3.06) as compared to individuals with a college education (M=1.18, SD=1.88 vs. M=1.06, SD=2.08). This interaction was not seen after the revision prompt (B=.52, t=.1.61, p=.11) due to relatively more similar gaps between web and face-to-face responses for individuals without a college education (M=2.04, SD=3.71 vs. M=1.44, SD=3.06) and those with a college education (M=1.16, SD=1.84 vs. M=1.06, SD=2.08).

There were no other significant two-way or three-way interactions between mode, wording, and whether or not participants had completed a college education ($p>.05$; not shown).

S.2. Effects of mode, wording, and revision opportunity on the percent of expectations over 5%

Table S1 shows the percent of participants who reported expectations that were higher than '5%'. Following previous work (Bruine de Bruin et al., 2010, 2011b, 2012), we used this cut-off, because the Michigan Survey of Consumers treats participants who provide expectations above 5% and provides them with the opportunity to revise (Curtin, 1996). As compared to previous studies that were conducted at times of higher realized inflation (Bruine de Bruin et al. 2010, 2012), the percent of reported expectations over 5% were relatively low. Table S1 suggests the following insights regarding the three survey design features of interest. First, significant mode effects emerged for each question wording, such that responses over 5% were more likely to be used in the web mode than in the face-to-face mode, in responses provided both before and after the revision prompt – with the exception of revised 'inflation' responses. Second, unlike Bruine de Bruin et al. (2012) we found no effect of wording on the use of responses over 5%. Third, being presented with the opportunity to revise only significantly reduced the percent of expectations over 5% for the web administration of the 'inflation' question.

Table S2 shows a significant administration mode effect on the likelihood of reporting expectations over 5%, in logistic regressions that took into account wording effects and demographic differences. Reporting expectations over 5% was more likely in the web mode than in the face-to-face mode, both before and after the revision prompt. There was no significant wording effect on reporting initial expectations over 5%. There was no significant

interaction between mode and wording, before or after adding demographics ($p>.05$; not shown).

In models that included demographics, participants without (vs. with) a college education were more likely to report initial expectations over 5% (6.3% vs. 1.7%) and final expectations over 5% (5.5% vs. 1.5%). There were no significant two-way or three-way interactions between mode, wording, and whether or not participants had completed a college education ($p>.05$; not shown).

Table S1: Descriptive statistics for reported expectations by question wording and administration mode.

	Initial (before revision prompt)			Final (after revision prompt)		
	Median	Mean (*SD*) absolute deviation from median	Percent reporting expectation over 5%	Median	Mean (*SD*) absolute deviation from median	Percent reporting expectation over 5%
Prices in general						
Face-to-face mode	1.50	1.37[i] (2.90)	2.7%	1.50	1.37[i] (2.90)	2.7%
Web mode	2.00	1.89[fa] (3.01)	7.7%[f]	2.00	1.80[fi] (2.85)	7.1%[f]
Inflation						
Face-to-face mode	1.00	1.21 (2.51)	1.9%	1.00	1.21 (2.51)	1.9%
Web mode	1.50	2.47[fpa] (8.22)	5.8%[fa]	1.50	1.59[f] (3.36)	4.1%

Note: M-W tests applied to between-group differences in means, and χ^2 tests to group differences in percentages, Wilcoxon tests to within-subject differences in means and percentages; Significance ($p<.05$) is noted for the following differences: [f]=significantly higher than for face-to-face mode; [i]=significantly higher than for inflation question; [p]=significantly higher than 'prices in general' question [a]=significantly higher than after prompt

Table S2: Regression analyses on absolute deviation from median and whether or not expectations were over 5%.

	Initial (before revision prompt)				Final (after revision prompt)			
	Absolute deviation from median (*B*)		Expectation over 5% (*OR, 95% CI*)		Absolute deviation from median (*B*)		Expectation over 5% (*OR, 95% CI*)	
Web mode (vs. face to face)	.90**	.90**	3.12*** (1.71, 5.66)	3.25*** (1.77, 5.99)	.41*	.41**	2.62** (1.42, 4.82)	2.68** (1.44, 4.99)
Prices in general wording (vs. inflation)	-.21	-.28	1.43 (.84, 2.44)	1.39 (.81, 2.40)	.18	.14	1.68 (.95, 2.97)	1.64 (.91, 2.94)
No college education	-	.80**	-	3.01** (1.34, 6.72)	-	.41*	-	2.43* (1.08, 5.49)
Low income[a]	-	.36	-	2.25* (1.07, 4.73)	-	.34	-	2.74* (1.22, 6.13)
Female	-	.77**	-	2.64** (1.37, 5.09)	-	.51**	-	2.28* (1.15, 4.50)
Age	-	-.01	-	.98* (.97, 1.00)	-	-.01	-	.98* (.97, 1.00)
R^2	.01	.03	.04	.16	.01	.01		.03

Note: For each model, *N*=1337. *B* is an unstandardized estimate in linear regression predicting continuous variable; *OR*=Odds Ratio in logistic regression predicting dichotomous variable. *p<.05; **p<.01; ***p<.001. Significance denotes difference from 0 (*B*) or 1 (*OR*). There were no significant two-way or three-way interactions between mode, wording, and whether or not participants had completed a college education (p>.05), except for a significant interaction between college education and mode in deviation of initial responses from the median (*B*=1.33, *t*=2.44, *p*=.02).
[a] Low income was defined as below the median of €1651 in income after taxes per month.