

Question-Order Effects on Judgements under Uncertainty[†]

Rafael Novella
Institute of Education, University College London
r.novella@ucl.ac.uk

Oxford Department of International Development, University of Oxford

Ericka G. Rascón Ramirez[‡]

Business School, Middlesex University London
E.Ramirez@mdx.ac.uk

Centro de Investigación y Docencia Económicas

December 20, 2023

Abstract

Data on judgements under uncertainty (e.g., expectations and risk preferences) have been increasingly used to predict and understand people’s decision making. However, little is known about how sensitive the elicitation of such measures is to questionnaire design. Using a survey experiment conducted in Chile, we assess order effects on the elicitation of expectations and risk preference measures. To do so, we randomly alter the position of questionnaire modules. Our findings suggest that expectations about future work, school, and wellbeing-related outcomes, as well as risk preferences, are sensitive to order effects. When people are asked about their past experiences immediately before the elicitation of such measures, more pessimistic expectations about future outcomes are reported, as well as higher levels of risk aversion. We also observed that positioning non-cognitive skill questions immediately before the elicitation of expectations and risk preferences generates data with lower variance suggesting more reliable measurements. Indeed, using these measures in prediction models of schooling and labour market outcomes, we find higher predictive power of future schooling and labour market outcomes when our measures are preceded by non-cognitive skills questions.

JEL classification: C93, D91, J24

Keywords: question-order, expectations, risk preferences, priming, belief elicitation

[†] We thank Marta Favara, Laura Fumagalli, Thomas Martin, Andreas Murr, and Graciana Rucci for helpful comments. We also thank Gabriela Munares for her work on a preliminary version of this paper and Belen Conde for excellent research assistance. Data collection was funded by the Inter-American Development Bank (IDB).

[‡] Corresponding author.

1. Introduction

Judgments under uncertainty (i.e., subjective expectations on future outcomes and risk preferences) have increasingly being used in observational and experimental studies to better understand decision-making. Expectations data have been used to understand and predict retirement age (van der Klaauw and Wolpin, 2008), sexual behaviour (Delavande, 2008; de Paula et al., 2014), schooling decisions (Arcidiacono et al., 2012; Attanasio and Kaufmann, 2014); and risk preferences to explain how people make decisions given their attitudes towards risk such as career and schooling decisions (Weiss, 1972; Belzil and Leonardi, 2013), entrepreneurship (Cramer et al., 2002), technology adoption (Liu, 2013), financial portfolios (Schubert et al., 1999), among others. However, to the best of our knowledge, there is limited experimental evidence on how sensitive these measures are when respondents are primed by prior information when collecting such data – an exception is Brañas-Garza et al (2022) where the authors look at anchoring effects when eliciting educational expectations. Understanding how sensitive data collection on judgments under uncertainty measures is to questionnaire order allows us to: (i) assess the comparability of questionnaires when they follow different module-order and (ii) assess which type of elicitation better predicts future outcomes.¹

This study addresses two main questions: (i) *can the elicitation of judgements under uncertainty (i.e., expectations and risk aversion) vary depending on the information that has been asked prior to their collection?* and (ii) *can the predictive power of this elicitation differ depending on the information collected prior to their collection?* To address these questions, we designed a survey experiment where we randomised the order of the questionnaire module located before our measures of expectations and risk aversion.² We use data collected for an impact evaluation in Chile that aimed at evaluating the impact of a job training programme on labour market outcomes (Doerr and Novella, 2020). Baseline data (2013) were used to check balance before our experiment; the first follow-up data (2014-2015) to implement our survey experiment; and the second follow-up data (2016) to assess which order, randomly assigned in the first follow-up, better predicts future outcomes reported in the second follow-up.³

The implementation of our survey experiment consisted of randomly assigning two questionnaires (Q1 and Q2) to survey respondents. Both questionnaires collected the same modules, but their position in the questionnaire differed. The first block of socioeconomic (SES) modules preserved the same position in both questionnaires. However, Q1 collects data on past experiences with labour market and schooling outcomes immediately before the module containing our measures of risk preferences and expectations about future work, school, and wellbeing; in contrast, Q2 collects data on non-cognitive skills immediately before these measures. Our survey design allows us to evaluate whether priming with past experiences or non-cognitive skills⁴ differently affects data collection on measures of judgments under uncertainty (subjective

¹ Similar studies have also explored the effect of framing on risk preference elicitation (Espinosa and Ezquerro, 2022), as well as how sensitive such measures are to visual and contextual aids (Estepa-Mohedano and Espinosa, 2023).

² We define a questionnaire module as a set of questions that belong to the same theme.

³ We also provide balance tests using sociodemographic and economic characteristics in Follow-Up 1 (2014-2015) that were not subject to our survey experiment.

⁴ Non-cognitive skills are broadly defined as personality traits, attitudes, and motivations that influence behaviour and are valued in the labour market, in school, and in many other domains (Heckman and Kautz, 2012). The non-cognitive skills considered in this paper consists of three types of measures. First, the Personal and Social Competence (CPS) is a non-cognitive test containing 44 items measuring six skills: leadership; conflict management; self-esteem (not

expectations and risk preferences). Both modules on past experiences (labour market and education) and non-cognitive skills are now common blocks of questions appearing in observational and experimental studies aiming to explain and predict labour market and schooling outcomes.⁵

To answer our questions of interest, we first test whether the priming generated by our design affects the mean and variance of expectations and risk preference measures; then we test whether this priming also affects their predictive power. We also provide robustness checks including placebo tests using observed variables that are not subject to uncertainty but were located after the two modules that were subject to changes in order. We finalise analysing order effects by gender, education, non-cognitive skills, cognitive skills, and past experiences in the labour market.

Our results show that information preceding the collection of expectations and risk aversion may affect the elicitation of the latter. Asking about past experiences in the labour market and school immediately before the expectations questions results in more pessimistic expectations, whereas asking non-cognitive skills before them results in more optimistic ones. Similarly, having past experiences questions immediately before risk preference measures results in higher levels of risk aversion than for those who were asked about non-cognitive skills before risk aversion. Our placebo tests reveal that only the elicitation of judgments under uncertainty is subject to priming; the rest of the module questionnaires are not statistically different between Q1 and Q2. Based on our prediction using the second follow-up data of the impact evaluation study, we found that the expectation data collected immediately after the non-cognitive skills tests better predict future outcomes (smaller Root Mean Square Error - RMSE) and has lower spread (smaller variance). Incorporating these findings in the design of survey questionnaires will help researchers to evaluate to what extent priming effects may affect data collection on judgements under uncertainty and if any, what the direction of such effect might be.

Our study contributes to the literature on the elicitation of judgments under uncertainty (Dave et al., 2010; Delavande et. al., 2011; Kerwin and Ordaz Reynoso, 2021), as well as to the literature on question-order effects on belief formation (Hogarth and Einhorn, 1992; Tubbs et al., 1993; Schlag et al., 2015). Identifying order-effects will help survey design, impact evaluation, and data collection experts to be aware of the possible lack of comparability between survey instruments when order is altered. Besides, our design also allows identifying which type of question-order results in measures that better predict future labour market and schooling outcomes.

The rest of the paper is organized as follows: Section 2 describes priming effects in survey data and the theoretical expectations of our study; Section 3 describes our experimental design and data; Section 4 presents our main findings; Section 5 presents robustness checks; and, Section 6 concludes and presents some recommendations.

included in the next dimension); interpersonal skills; organization and tidiness; and, empathy and communication skills (Brea, 2011). The second is based on the Rosenberg Self-Esteem Scale (RSES), which consists of 10 items about the person's overall evaluation of their own worth (Rosenberg and Pearlman, 1978). The final measure is Grit, which is a 14-item scale measuring perseverance, passion to achieve long-term goals, and consistency (Duckworth and Quinn, 2009). CPS and RSES were elicited by using a Likert scale from 1 to 4 ("Nothing like me", "Little like me", "Like me" and "Very Similar to me"). The Grit measure uses a Likert scale from 1 to 5 ("Nothing like me", "Little like me", "Something similar to me", "Like me" and "Very Similar to me"). In Chile, the Rosenberg Self-Esteem Scale was validated by Rojas-Barahona, Zegers and Förster (2009).

⁵ Our questionnaires have 18 modules: household characteristics, education, employment, incomes, household assets, non-cognitive skills, educational trajectory, training programme experience, employment trajectory, health and social security, intertemporal discounting, access to financial services, expectations, cognitive skills, risk preferences, parental background, risky behaviours, and individual/household shocks.

2. Question-order and Judgements under Uncertainty

Scholars have used surveys as a tool to explore how question-order may affect respondents' answers (Bradburn and Mason, 1964). These studies primarily focus on understanding to what extent questions located early in a questionnaire affect answers in subsequent questions. In this paper, we focus on analysing whether priming with questions about past experiences or non-cognitive skills affects answers in measures of judgments under uncertainty (subjective expectations and risk preferences). The importance of exploring how sensitive measures of judgments under uncertainty are to order-effects is threefold. First, surveys collecting identical data, but following different question order may not be comparable. Although survey questionnaires may preserve the same design and wording over time, changes in the position of survey modules may affect responses through the context outlined by earlier questions. Second, if question order matters, prediction estimates may differ between questionnaires using the same predictors and outcomes, even when the same questions have been asked. Third, if responses are subject to question order effects, the position of questionnaire modules can be subject to manipulation by data users with specific response interests.

Although all responses in questionnaires are subject to measurement error, those aiming at collecting data on observed or known outcomes (e.g., age, gender, income, number of children, marital status, skills) are less likely to be affected by order effects. When asking about observed/known outcomes, respondents rely primarily on memory or retrieval of past experiences, but not on the assessment of possible scenarios that have not happened yet. Conversely, survey responses about uncertain outcomes are more likely to be affected by mood, priming through previous responses, and framing of survey questions. Mood and priming effects can be induced during the interview through reviving joyful memories (Kavanagh and Bower, 1985; O'Hagan et al., 2006; Israel et al., 2014) or by question order (Molden, 2014; Loginova, 2009).

Eliciting beliefs about uncertain outcomes requires the respondent to go through cognitive processes involving various stages (Strack and Martin, 1987): first, the respondent reads or listens the question and then she interprets it; second, she creates an opinion – which in our case entails evaluating the chances of observing an outcome in the future or evaluating the level of risk of a hypothetical situation; and finally, she chooses an answer or formulates one. During the two initial stages, a fundamental cognitive process is triggered: retrieving relevant information (e.g., prior beliefs, previous opinions, experiences, emotions) from long-term memory. Because individuals have limited capacity to retrieve long-term memory information (or have imperfect memory), they are more likely to access only the most recent or the most remarkable information they have in their mind. This is also known as the satisficing principle (Simon, 1996).⁶

When collecting data through surveys, respondents are subject to priming from questions located at the beginning of the questionnaire which results in the most recent information they have retrieved from their memory. Prior activation of a trait category (i.e., stored information such as particular behaviours, appearance, intentions) makes more accessible such information and increases the likelihood of relating subsequent questions to the activated category (Higgings et al., 1977).

We allude to the concept of priming used by social psychologists, which is concerned about how cues derived from particular social situations can influence people's responses even when

⁶ Additionally, Dohmen et al. (2018) shows that cognitive skills is associated with risk-taking: they tend to be positively correlated with avoidance of harmful risky situations, but they tend to be negatively correlated with risk aversion in advantageous situations.

participants are not aware of such influence. The failure to recognise the implications of first stimuli or exposures is what makes critical priming effects (Molden, 2014; Aimone et al., 2019). Initial studies on social priming have shown that people utilize activated social knowledge in their judgments even when such activation arises from unrelated and irrelevant sources, and after a substantial delay following the initial activation or stimuli (Molden, 2014). Studies suggesting that priming could affect social behaviours and impressions have been carried out by social psychologists (Dijksterhuis and Bargh, 2001). Our study complements this literature by analysing priming effects in the elicitation of decisions under uncertainty using univariate tests (differences in mean and variance) and the predicted power of the judgements under uncertainty that have been subject to question order.

In our study, we are not able to distinguish whether representativeness or availability of information is the main driver of question order, as suggested by Tversky and Kahneman (1974). However, we know that in both types of questionnaires (Q1 and Q2), respondents may retrieve information from memories and experiences to then report their expectations and risk preferences, as suggested by Higgins and King (1981).

3. Data collection and Experimental Design

Our survey experiment is part of a large-scale impact evaluation that took place in Chile from 2012 to 2016. In 2012, the Chilean National Training and Employment Service (SENCE) and the Inter-American Development Bank (IDB) decided to evaluate the effect of the social programme *Formación para el Trabajo* (FOTRAB, Training for the Job). This programme was one of the widely spread job training programmes in the country which targets men and women aged between 18 and 65 living in vulnerable conditions.⁷

The baseline survey was collected between March and July 2013, the first follow-up between November 2014 and January 2015; and the second one between March and June 2016. Our survey experiment took place in the first follow-up data and our prediction analysis uses the second follow-up. We use baseline data to check the statistical balance between individuals randomly assigned to complete either of the two questionnaires. Individuals surveyed in the first follow-up had on average 12 years of education – as expected from the selection criteria, where 65% of the sample was women, 44% was household head, and the average age was 34 years old.

Table 1 shows all questionnaire modules and their respective number of questions. The last column indicates whether the module is self-applied by the respondent or not. Data collection was done using either computer-assisted personal interviewing (CAPI) or pen-and-paper personal interviewing (PAPI) for both self-applied and non-self-applied modules. Most of our data collection was primarily done using CAPI (97%) and the rest using PAPI questionnaires.⁸ The total duration of our follow-up surveys was on average 1 hour and 15 minutes.

⁷ Participants should have less than tertiary education and live in vulnerable conditions to be eligible. Individuals aged 16 and 17 were able to apply to the FOTRAB programme conditional on having parental responsibilities. To be identified as an individual living in vulnerable conditions, applicants had to live in households belonging to the first two quintiles of the scoring tool known as *Ficha de Protección Social (FPS)*, which was used to target social programmes at the national level until 2015.

⁸ The main reasons to use paper instead of the tablet was due to problems with the tablet (37%), security reasons (28%), issues with the battery of tablet (14%), among other reasons.

Table 1: Survey Modules: First Follow-up

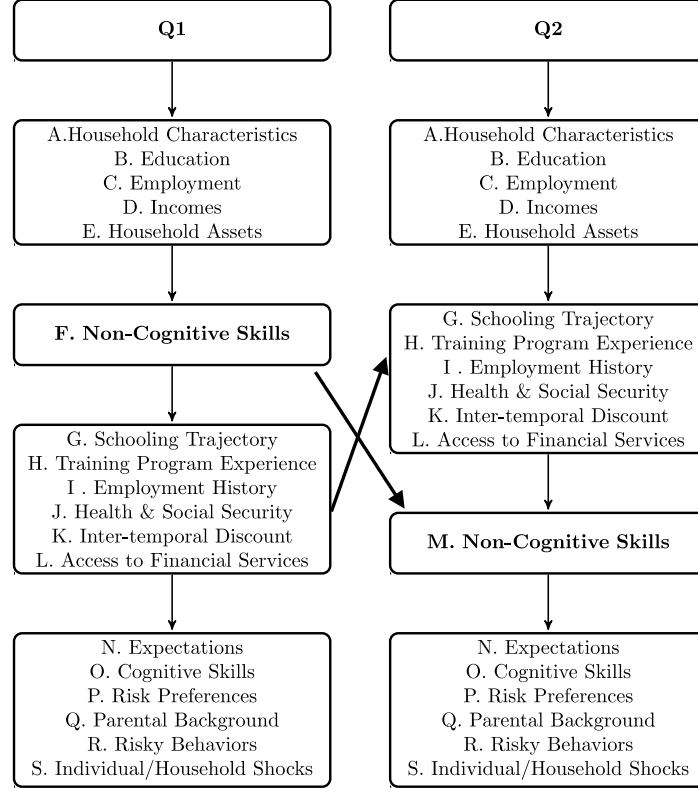
Module	Module Name	No. of Questions	Self-Applied
A	Household Characteristics*	8	No
B	Education	6	No
C	Employment**	9	No
D	Incomes	12	No
E	Household Assets	3	No
F/M	Non-Cognitive Skills	68	Yes
G	Schooling Trajectory	6	No
H	Training Program Experience	25	No
I	Employment History	30	No
J	Health & Social Security	13	No
K	Inter-temporal Discount	3	No
L	Access to Financial Services	7	No
N	Expectations	19	No
O	Cognitive Skills	12	No
P	Risk Preferences	3	No
Q	Parental Background	16	No
R	Risky Behaviours	15	Yes
S	Individual/Household Shocks	9	No
Total		264	—

* All household members answered these questions.

** Only household members older than 12 years old answer these questions.

To test question order effects, two different questionnaires were designed where they uniquely differ in the order of questionnaire modules. Questionnaires Q1 and Q2 were randomly allocated among survey respondents. Figure 1 shows the modules included in each questionnaire (Q1 and Q2) and how the order of the survey modules varies between them. Q1 and Q2 have the first five (A - E) and the last four (N - S) modules located in the same position. In Q1, the block of modules of past experiences (G-L) is located immediately before the elicitation of expectations (N) and risk preferences (P), whereas in Q2 the non-cognitive skill module appears immediately before these measures.

Figure 1: Questionnaire design



The random allocation of Q1 and Q2 allows us to estimate the causal effect of altering the order of past experiences and non-cognitive skills on measures of judgements under uncertainty (i.e., expectations and risk preferences) located in the last boxes (N and P modules) shown in Figure 1. Our experiment is in line with the literature on how randomization of questions can be used to guarantee unbiased estimates of order effects (Glennerster and Takavarasha, 2013).

To collect expectations data, we considered the likelihood of observing specific outcomes in the next five years. A Likert scale using five possible values, ranging from not very likely to very likely, was used to elicit expectations on schooling, employment, wealth, crime, and issues related with family, friends, and health. A z-score was constructed for each measure of expectations using the mean and standard deviation of each variable. The majority of expectation questions asked for a prediction of a specific outcome for the next five years. Twenty per cent of all expectation variables were measured using a 10 and 20-year time frame. Despite the differences in time frames, z-scores allow comparability across expectations variables and all questions were identical in both Q1 and Q2 questionnaires. Here we enlist a representative question for each type of expectations. Expectations on *schooling outcomes* include (one question): what is the likelihood that you would have completed a higher level of education in five years? On *employment* (three questions), what is the likelihood that you would have your own business in five years? On *wealth* (six questions), what is the likelihood that you would live in a better neighbourhood, or have the house/car that you desire in five years? On *crime* (one question), what is the likelihood that you would have committed a crime in five years? On *family and friends* (seven questions), what is the

likelihood that your aspired personal/family plans are accomplished in five years? And on *health-related issues* (one question), what is the likelihood that you would have been infected of AIDS in five years?⁹ We also created an aggregate measure of expectations where we add up all expectation z-scores.

Finally, to collect risk preference measures, we use three hypothetical scenarios, with two alternatives each. The first choice always offers a fixed payment forever with no risk and the second, a lower and upper bound of a possible payment with 50 per cent probability to observe each scenario. The lower bound is always lower than the option of offering a fixed payment. The respondent selects one option per scenario, see Table A1 showing choices, payments and probabilities for each event. We identified a risk-averse individual (=1) if she was classified as very high or high-risk aversion given her choices and risk-lover (=0) if the individual presented a very low or low-risk aversion.¹⁰ According to our definition of risk preference, no risk-neutral individuals are identified in our sample. A wide range of risk preference measures have been suggested by lab experimentalists using monetary incentives and lotteries involving multiple choices, see Camerer and Hogarth (1999) for a review. In this study, we use a simple hypothetical measure that is used in the Social Protection Survey (EPS), which is the oldest longitudinal survey in Chile having 16,000 respondents across the country.¹¹

4. Results

As discussed in previous sections, our survey experiment consisted of randomly allocating two types of questionnaires (Q1 and Q2) to survey respondents. These questionnaires only differ in the order of past experiences and non-cognitive skill modules prior to the elicitation of expectations and risk preferences measures, our outcomes of interest. After randomly allocating individuals to

⁹ For *employment-related* expectations, the other two questions are: what is the likelihood that you would have the job that you desire in five years? and what is the likelihood that your professional/employment aspirations or plans are accomplished in five years? For *wealth-related expectations*, the other five questions are: what is the likelihood that you will live in Chile in five years?; in 20 years, what is the likelihood of having a better life than the one your parents have?; what is the likelihood of your children having a better life than yours in five years? The last two items on *wealth expectations* are not based on likelihood measures, however, they follow an identical ranking as the Likert scale (1-5): In five years, if 1 represents the poorest person in Chile and 5 the richest, in which position would you be situated? Using the same scale, where will you be situated in 10 years? For *family and friends-related* expectations, the other six questions are: what is the likelihood that your current group of friends will be kept in five years?; what is the likelihood of having a problem with a family member in five years (reversed scale to be comparable with the rest of expectations measures); what is the likelihood that someone close to you has been infected of AIDS in five years? (reversed scale); what is the likelihood that someone close to you has died as a result of a violent event in five years? (reversed scale). The last two items on *family and friends expectations* are not based on likelihood measures, however, they follow an identical ranking as the Likert scale (1-5): In five years, if 1 corresponds to the least respected person in Chile and 5 the most respected, in which position would you be situated? Using the same scale, where will you be situated in 10 years?

¹⁰ The three choices offer a hypothetical sure payment of \$300,000 Chilean pesos and uncertain payment with expected value of \$400,000. All uncertain payments maintain the same probabilities and the lower payment increases by 50,000 in each scenario (i.e., the upper payment decreases by the same amount) maintaining constant the expected value. A person is risk-averse if she chooses the sure payment over the uncertain outcome. This measure does not capture different levels of risk aversion as in the standard approach suggested by Holt and Laury (2002), but follows the same rationale as their measure.

¹¹ There is growing evidence from lab and field experiments suggesting no significant differences in risk preference profiles when using hypothetical and incentivised risk preference measures (Wiseman and Levin, 1996; Kühberger et al., 2002). Branas-Garza et al. (2021) show that incentivised and hypothetical measures do not differ in consistency and the number of safe choices when using lotteries in the lab (Spain) and in the field (Honduras and Nigeria).

Q1 and Q2, we compared the pre-treatment characteristics of both samples using the baseline questionnaire. Table A.2 of Appendix shows the statistical balance of baseline characteristics. We observe balance for sociodemographic characteristics, non-cognitive skill, and expectations measures; risk aversion was not elicited in the baseline survey. We also provide in Table A3 balance tests for the set of sociodemographic characteristics located prior to the modules that were subject to the survey experiment.

Next, we address our two main questions: (i) are there order-effects on the elicitation of expectations and risk preference measures? and (ii) is the predictive power of these measures the same, no matter the information that precedes their elicitation? For the first question, we first analyse whether there are order effects on past experiences and non-cognitive skills as they are the modules that were asked prior to the elicitation of our outcomes of interest. Then, we look into order effects on expectations and risk preferences. For the second main question, we test the predictive power of expectations and risk aversion using prediction models with the second follow-up data. To finalise our analysis, we discuss placebo tests based on known (certain) variables located in the same module as our measures of judgements under uncertainty. We also provide a subgroup analysis by gender and education at the end of this section.

Are there order-effects in outcomes involving uncertainty?

We start testing whether past experiences and non-cognitive skills were subject to order effects. Using a representative selection of variables from the past experiences of schooling and employment modules (G-L), we test whether eliciting non-cognitive skills before this module affects the responses of schooling trajectory, training programme experience, employment history, self-assessment of health, inter-temporal discounting, and access to financial services. For these modules, the respondent does not require making any inference about the events/outcomes asked, but only relies on her memory to respond. Thus, we expect that data collection on past experiences will not be affected by the position of the non-cognitive skill module.

Table 2. Report on Past Experiences

	Means		Standard Deviations		Observations		Difference p-value
	Q1	Q2	Q1	Q2	Q1	Q2	
Edu: Participated in Educational Programme	0.02	0.03	0.15	0.17	2238	2258	0.24
Edu: Subsidised Secondary School	0.20	0.20	0.40	0.40	2219	2250	0.60
Edu: Private Secondary School	0.03	0.02	0.18	0.15	2219	2250	0.11
Training: Applied or Participated in programme	0.97	0.98	0.17	0.15	2246	2266	0.18
Training: Score by the respondent	-0.02	0.02	1.05	0.95	1254	1237	0.23
Work: Looking for job in the last month	0.45	0.45	0.50	0.50	710	661	0.92
Work: Would have accepted a paid job	0.31	0.27	0.46	0.44	379	357	0.18
Work: Are you an employer?	0.14	0.14	0.35	0.34	2246	2266	0.87
Health: Self-assessment of health	5.93	5.87	1.02	1.09	2245	2265	0.05
Health: Pregnancy in the last year	0.12	0.10	0.33	0.31	1175	1183	0.14
Health: Pregnancy of your partner in the last year	0.15	0.15	0.36	0.36	492	475	0.96
Intertemporal Discount: Score	0.43	0.44	0.50	0.50	2196	2237	0.73
Financial: Have applied for a credit in the last 12 months	0.15	0.14	0.36	0.35	2243	2259	0.41
Financial: Have stopped paying a debt for more than 30 days	0.37	0.38	0.48	0.49	2242	2256	0.41

Note: Discrete variables taking 0 and 1 values are reported in this table with exception of training score, health self-assessment, and intertemporal discount score. Differences are evaluated using standard t-tests.

Table 3. Elicitation of Non-Cognitive Skills

	Means		Standard Deviations		Difference p-value
	Q1	Q2	Q1	Q2	
Non-Cognitive	176.00	175.41	27.82	27.78	0.47
Non-Cognitive (PCA)	0.03	-0.03	3.79	3.77	0.65
CPS	96.83	96.54	19.16	19.12	0.61
Rosenberg Self-Esteem	21.94	21.92	6.55	6.56	0.90
Grit	57.22	56.95	8.74	8.88	0.29

Note: Scores are reported in this table. The total number of observations is 4,512 where questionnaire Q1 has 2,246 and Q2 2,266. Differences are evaluated using standard t-tests.

As shown in Table 2, the report of past experiences is not affected by order effects. Similarly, we test whether non-cognitive skills are subject to order effects. Table 3 shows there are none. Despite having non-cognitive skills administered after five modules (A-E) in Q1 and after 11 modules (A-L) in Q2, non-cognitive skills primarily reflect personality traits that are presumably known by the respondent and are not affected by the possible priming effects generated in Q2. Comparing the means of a global measure of non-cognitive skills (sum of measures and principal component analysis index) and each of the three dimensions encompassing our non-cognitive skill measure,

Table 3 shows no significant differences between questionnaires. The magnitude of the differences is also trivial. Similar to the report of past experiences shown in Table 2, the elicitation of non-cognitive skills is not affected by order effects (i.e., no differences in means).¹²

Moving on to the test for the subjective expectations and risk preference measures, our main outcomes of interest, Table 4 shows significant differences in all, but two measures of expectations – health and crime expectations. Having the modules (G-L) of past experiences immediately before the module of expectations and risk preference in Q1 shows more pessimistic expectations and higher risk-aversion than those elicited in Q2. We find differences in means in expectations and risk preferences. Also, we observe a difference in the spread of the measures (in most of the cases the variance is smaller in the expectations data generated by Q2).¹³

Table 4. Elicitation of Expectations and Risk Preference

Measure	Means		Standard Deviations		Difference p-value
	Q1	Q2	Q1	Q2	
Education	-0.06	0.06	1.05	0.95	0.00
Work	-0.06	0.06	1.05	0.94	0.00
Family & Friends	-0.03	0.03	0.99	1.01	0.08
Health	0.00	0.00	0.98	1.02	0.96
Crime	-0.01	0.01	0.98	1.02	0.49
Wealth	-0.06	0.06	1.03	0.97	0.00
Expectations (all)	-0.07	0.07	1.04	0.96	0.00
Risk preferences	0.54	0.51	0.50	0.50	0.06

Note: Z-scores are reported in this table. Higher values of expectations denote optimistic beliefs and higher values of risk preference denote higher risk aversion. The total number of observations is 4,512 where questionnaire Q1 has 2,246 and Q2 2,266. Differences are evaluated using standard t-tests.

To explore some of the possible drivers of the observed order effects, we estimated the association between expectations and past experiences/non-cognitive skills using OLS regressions. Also, we estimated such associations for our measure of risk preferences using linear probability models (LPM). We use as dependent variables expectations (all) and risk preferences; as regressors we use a subset of variables representative of the past experience modules (G-L) and all the items considered in the non-cognitive skills module (F module in Q1 and M module in Q2). For each regression, we have one regressor where we test equality of coefficients under Q1 and Q2. We then ran a joint hypothesis test using the same aggregation described in Table 2 for past experiences and in footnote 4 for non-cognitive skills. Table 5 shows Chi-2 and p-values corresponding to the joint tests of equality of coefficients between Q1 and Q2.

We observe that the main differences are likely to be explained by two main sources: the association between expectations and self-esteem, as well as between expectations and past experiences with health-related events. No discrepancies in the associations between past experiences/non-cognitive skills and risk preferences are found. In Table 5, the question modules

¹² No correlational differences are found when comparing the correlation of non-cognitive skills and past experiences in Q1 and Q2. Multivariate tests on correlations were used for this analysis. These tests are not shown in the tables but are available upon request.

¹³ No differences in correlations were observed when comparing the correlation of non-cognitive skills and expectations/risk preference in Q1 and Q2. This result is not shown in Table 4, but are available upon request.

are positioned in the table based on the order of appearance in the questionnaires (for past experiences, the education module is collected first, and the financial module is collected at the end; for non-cognitive skills, empathy is collected first and grit at the end). For the non-cognitive skills items, we provide a general overview of the location of the majority of items encompassing each block of questions as some of them are spread along the non-cognitive skill module. Our table shows that our main findings are likely to be driven by the priming generated by the questions collected prior to expectations and risk preferences, and not by fatigue effects. If the latter was the case, we would primarily observe significant differences in associations between our variables of interest (expectations and risk preferences) and the latest questions asked in past experiences (financial) and the latest questions in non-cognitive skills (grit).

In the next section we analyse the prediction power of our measures, followed by a discussion of placebo tests and sub-group analyses.

Table 5. Joint tests of differences in OLS/LPM coefficients of expectations and risk preferences on past experiences and non-cognitive skills

Coefficients from Q1 vs Coefficients from Q2

Modules	Expectations		Risk Preferences	
	Chi-2	p-value	Chi-2	p-value
Past experiences				
Education	1.25	0.74	0.58	0.90
Training	0.23	0.89	0.60	0.74
Work	2.98	0.40	1.17	0.76
Health	27.09	0.00	1.33	0.72
Intertemporal Preferences	0.00	0.97	0.15	0.70
Financial	2.93	0.23	0.29	0.86
Non-Cognitive Skills				
Empathy	14.62	0.26	8.05	0.78
Conflict Management	6.13	0.80	14.02	0.17
Self-Esteem*	8.01	0.05	4.76	0.19
Organisation and Tidiness	6.13	0.52	2.51	0.93
Leadership	9.21	0.24	3.55	0.83
Interpersonal Skills	9.05	0.11	5.76	0.33
Rosenberg Self-Esteem	16.41	0.09	9.74	0.46
Grit	13.78	0.47	5.91	0.97

Are there differences in predictive power given the change in order?

In the previous section, we showed that expectation and risk preference data can be sensitive to question-order. We also showed that past experiences and non-cognitive skills were not sensitive to this change in order. In this section, we analyse whether this change in order changed the predictive power of expectations and risk aversion data. Despite not having perfect correspondence between observed outcomes collected in the second follow-up and expectations/risk preference measures in the first follow-up, we can test whether the measures collected in Q1 are better or worse predictors of future outcomes than those collected in Q2. This is in line with the research led by scholars who use expectations data to predict future outcomes (Leiter et al., 2018). Our prediction exercise helps us: (i) test the validity of both Q1 and Q2 measures; and (ii) inform data collection experts and users about which measure might be more adequate when the purpose of the questionnaire aims to predict future behaviours and outcomes.

Table 6 shows the RMSE of OLS regressions having as dependent variable schooling (attendance and years of schooling), labour-related (employment status, hours worked, and number of jobs), and well-being (health, societal self-valuations) outcomes collected in the second follow-up and explained by expectations and risk preference measures collected in the first follow-up.¹⁴ We have selected the prediction of such measures given that they are common variables explored by social scientists. The questionnaire for the second follow-up followed the same order as Q1 in the first follow-up.

Table 6. Prediction power of Behavioural Outcomes in Follow-up 2: Q1 vs Q2

Dependent Variable	Root Mean Square Error (RMSE)		Observations	
	Q1	Q2	Q1	Q2
Attending school	0.31	0.30	1,994	2,082
Years of schooling	1.96	1.81	1,994	2,081
Employer	0.36	0.35	1,994	2,082
Employee	0.44	0.44	1,994	2,082
Weekly hours in main job	15.72	15.11	1,430	1,503
Number of jobs	0.25	0.24	1,436	1,504
Self-assessment of health	1.08	1.16	1,994	2,081
Feeling valued by society	0.49	0.49	1,883	1,993
Believes he/she can receive help from friends if needed	0.49	0.50	1,925	2,024
Feeling part of the community in the workplace	0.49	0.49	1,706	1,783

Note: This table shows the RMSE of OLS regressions of the corresponding dependent variable collected in the second follow-up. The covariates used in these prediction models are expectations, risk preference and non-cognitive measures collected in the first follow-up.

Comparing the RMSE of Q1 and Q2, we observe that in all but two regressions, Q2 measures of judgments under uncertainty better predict schooling and labour market outcomes than Q1. The largest differences in RMSE are found for the outcomes of years of schooling (Q2 presents a 7.6% RMSE lower than Q1), weekly hours in the main job (3.9% RMSE lower than Q1), and being an employer (3.2% lower than Q1). For self-assessment of health and believe he/she can receive help from friends if needed, we obtain the opposite result, where the RMSE is lower in Q1 than in Q2 (7.4% and 0.7%, respectively).

We conclude that the measures of expectations and risk preference under Q2 (i.e., having the module of non-cognitive skills immediately before the elicitation of expectations and risk preferences) result in better predictors of future schooling and labour market outcomes. Our results are likely to be driven by the measures of self-esteem in the non-cognitive skills module and of past experiences with health-related events. Because our findings show the combined effect of the change in both past experiences and non-cognitive skill questions, further research is needed to

¹⁴ To assess how well our regression models forecast outcomes in the second follow-up, we use the standard measure of RMSE. This measure corresponds to the standard deviation of the forecast errors without any degrees of freedom adjustment. When comparing models, we would select the model with the smallest RMSE, see Wooldridge (2016), Chapter 18. This is in line with the criterion of minimizing the sum of square errors from which OLS regression is defined.

identify which trait may tend to prime more the elicitation of judgements under uncertainty, as well as help such data become a better predictor of future outcomes.

5. Robustness checks: Placebo Tests and Sub-group Analysis

In this section, we discuss two checks: (i) placebo tests using variables which are already observed and are located close to our expectations and risk preference measures; and (ii) sub-group analysis. The first robustness check tests the validity of our argument regarding the malleability of survey responses when these involve the elicitation of outcomes that are uncertain to the respondent. Table 7 presents a placebo test. Here we use long-term observed outcomes (cognitive skills and parental schooling and employment background) and short/medium-term observed behaviour or events (risky behaviour and negative shocks). These outcomes are subject to the same order effects as the expectation and risk preference modules – see Figure 1 in section 3. For these observed outcomes, the respondent does not require the processing of information to make inference about an uncertain outcome. Table 7 confirms that for those measures where the respondent does not require an assessment of uncertain outcomes, the individual is not subject to order effects.

Table 7. Placebo tests

Measure	Means		Standard Deviations		Observations		Difference p-value
	Q1	Q2	Q1	Q2	Q1	Q2	
Raven Test (% correct)	54.28	53.85	20.51	20.29	2049	2072	0.50
Father's years of schooling	8.33	8.31	3.91	3.93	1695	1749	0.87
Father's employment when respondent was under 15	0.93	0.94	0.26	0.25	1864	1915	0.29
Mother's years of schooling	7.92	8.01	3.90	3.84	2012	2050	0.47
Mother's employment when respondent was under 15	0.56	0.55	0.50	0.50	2155	2187	0.70
Risky behaviour (index)	0.20	0.21	0.16	0.17	2233	2248	0.17
Negative shocks (index)	1.18	1.23	1.18	1.20	2246	2266	0.20

Note: Differences are evaluated using standard t-tests. The index of risky behaviour considers information about being engaged in some type of anti-social behaviour such as being involved in a fight, receiving physical threats, consumption of drugs/alcohol, among others. The index of negative shocks considers events such as losing a job, illness, burglary, decrease in agricultural yield, among others.

In sum, our findings provide evidence on: (i) differences in elicited expectations and risk preferences measures when the information preceding such elicitation is different; and (ii) this effect is only present on measures under uncertainty. Because the module of cognitive skills is located between the modules of expectations and risk preference, and there are no significant effects between Q1 and Q2 in such measures, we claim that the effect we observe in such measures is derived from the nature of the elicitation of uncertain outcomes, which might be more sensitive to the activation of categories and emotional mood related to personality traits. This conclusion is also supported by the evidence of non-significant differences between Q1 and Q2 when collecting data on past experiences (G-L) and non-cognitive skills (M).

Finally, we explore which type of respondent is more susceptible to be subject to order effects. Tables A.4-A7 of the Appendix shows OLS/LPM regression models for each measure of expectations and risk preferences interacting a dummy variable to denote questionnaire order with gender, education, employment, cognitive, and non-cognitive skills. To ease the interpretation of our interactions, Table A4 presents interactions with gender and education; regressions shown in

Tables A5-A7 maintain the same controls as in Table A4 but interact the dummy variable of order with employment variables (A5), intelligence, spatial ability, and numeracy/literacy skills (A6) and non-cognitive skills (A7). Overall, we do not find systematic differences between females and males across expectation measures (Table A4). For crime-related expectations, we observe that respondents with higher levels of education tend to be more sensitive to order effects than those respondents with no education or primary schooling. In Table A6, there is some evidence on individuals ranking in the lowest tercile of the fluid intelligence score (Raven test) being more sensitive to order effects when reporting their risk preferences. Overall, we find little evidence to assert that certain sub-populations are more sensitive to order effects than others.

6. Discussion

Data collection on expectations and risk preference measures in household surveys has become a common practice in observational and experimental studies to better understand and predict labour market outcomes, schooling achievement, retirement decisions, among others. This paper studies how the elicitation of such measures is affected by question-order and whether question-order affects the predictive power of expectations and risk preference measures.

We find that placing past experience questions immediately before the elicitation of expectations and risk preferences prompt more pessimistic expectations about work, schooling and well-being, as well as higher levels of risk aversion. Conversely, having non-cognitive skills immediately before the elicitation of such measures derives into more optimistic expectations and higher risk-seeking attitudes. Hence, the location of non-cognitive skill and past experience modules when eliciting subjective expectation and risk preference data is not irrelevant and has to be considered when analysing and predicting future outcomes related to schooling, employment and well-being.

Our results also suggest that changes in the order of module questionnaires can also change the predictive power of expectations and risk preference data to forecast future outcomes. In our study, having the non-cognitive skills module immediately before such measures improves the predictive power of our measures of judgments under uncertainty. The reason is that the variance of these measures is also smaller in this case, which means their reliability increased.

There are a few limitations of our study which are worth mentioning. First, given that our treatment consisted of moving a whole module instead of single questions, we are not able to disentangle whether the estimated effects would remain if only single questions would have been moved. Second, we are not able to identify which question order makes quicker the elicitation of our measures as we do not have response times per module.

To better understand how question-order may also affect another type of psychological and judgement under uncertainty measures, further research is needed in this field. We recommend to lab and field researchers to pilot survey questionnaires before large-scale data collection takes place to assess how sensitive the measures of judgment under uncertainty are to question-order.

References

- Aimone, J., North, C. and Rentschler, L. (2019). "Priming the jury by asking for Donations: An empirical and experimental study". *Journal of Economic Behavior & Organization*, Volume 160, April 2019, 158-167.
- Arcidiacono, P., Hotz, V. J. and Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1), 3-16.
- Attanasio, O. P. and Kaufmann, K. M. (2014). "Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender". *Journal of Development Economics*, 109, 203-216.
- Belzil, C. and Leonardi, M. (2013). "Risk aversion and schooling decisions". *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, Vol. 111-112. 35-70.
- Bradburn, N. M. and Mason, W. M. (1964). "The effect of question order on responses". *Journal of Marketing Research*, 1, 57-61.
- Brañas-Garza, P., Ciacci, R., & Rascon-Ramírez, E. (2022). Anchors matter: Eliciting maternal expectations on educational outcomes. *Journal of Economic Psychology*, 90, 102510.
- Brañas-Garza, P., Estepa-Mohedano, L., Jorrat, D., Orozco, V., & Rascon-Ramirez, E. (2021). To pay or not to pay: Measuring risk preferences in lab and field. *Judgment and Decision Making*, 16(5), 1290-1313.
- Brea, M. (2011). "Informe de la interpretación de las escalas CPS, Rosenberg y Grit y propuestas de revisión para mejorar la confiabilidad". Washington, D.C.: World Bank Group. <http://documents.worldbank.org/curated/en/418161467995809251/Informe-de-la-interpretacion-de-las-escalas-CPS-Rosenberg-y-grit-y-propuestas-de-revision-para-mejorar-la-confiabilidad>.
- Camerer, C. F. and Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty*, 19(1):7-42.
- Cramer, J. S., Hartog, J., Jonker, N. and Van Praag, C. M. (2002). "Low risk aversion encourages the choice for entrepreneurship: an empirical test of altruism". *Journal of Economic Behavior & Organization*, 48(1), 29-36.
- Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better?. *Journal of Risk and Uncertainty*, 41(3), 219-243.

- Delavande, A. (2008). “Pill, patch, or shot? Subjective expectations and birth control choice”. *International Economic Review*, 49(3), 999-1042.
- Delavande, A., Giné, X., & McKenzie, D. (2011). Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design?. *Journal of Applied Econometrics*, 26(3), 479-497.
- Doerr, A. and Novella, R. (2020). “The Long-Term Effects of Job Training on Labor Market and Skills Outcomes in Chile”. IDB Working Paper Series No. IDB-WP-1156.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2018). “On the relationship between cognitive ability and risk preference”. *Journal of Economic Perspectives*, 32(2), 115-134.
- Duckworth, A. L., Peterson, C., Matthews, M. D. and Kelly, D. R. (2007). “Grit: perseverance and passion for long-term goals”. *Journal of personality and social psychology*, 92(6), 1087.
- Duckworth, A. L. and Quinn, P. D. (2009) “Development and Validation of the Short Grit Scale (Grit-S)”, *Journal of Personality Assessment*, 91: 2, 166-174.
- Espinosa, M. P., and Ezquerro, L. (2022). The effect of domain and framing on elicited risk aversion. *PLoS One*, 17(9): Article e0267696.
- Estepa-Mohedano, L. and Espinosa, M. P. (2023). Comparing risk elicitation in lotteries with visual or contextual aids. *Journal of Behavioral and Experimental Economics*, 103: 101974.
- Glennerster, R. and Takavarasha, K. (2013). “Running randomized evaluations. A practical guide”. Princeton University Press.
- Heckman, J. and Kautz, T. (2012). “Hard evidence on soft skills”. *Labour Economics* 19 (2012) 451–464.
- Higgins, E. T. and King, G. A. (1981). Accessibility of social constructs: Information processing consequences of individual and contextual variability. In: N. Cantor & JF Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 69-121).
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive psychology*, 24(1), 1-55.
- Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *The American Economic Review*, 92(5):1644–1655.
- Israel, A., Rosenboim, M. and Shavit, T. (2014). “Using priming manipulations to affect time preferences and risk aversion: An experimental study”. *Journal of Behavioral and Experimental Economics*, 53, 36-43.

- Kavanagh, D. J. and Bower, G. H. (1985). "Mood and self-efficacy: Impact of joy and sadness on perceived capabilities". *Cognitive Therapy and Research*, 9(5), 507-525.
- Kerwin, J. T., & Ordaz Reynoso, N. (2021). You Know What I Know: Interviewer Knowledge Effects in Subjective Expectation Elicitation. *Demography*, 58(1), 1-29.
- Kühberger, A., Schulte-Mecklenbeck, M., and Perner, J. (2002). Framing decisions: Hypothetical and real. *Organizational Behavior and Human Decision Processes*, 89(2):1162–1175.
- Leiter, D., Murr, A., Rascon, E. and Stegmaier, M. (2018). "Social networks and citizen election forecasting: The more friends the better". *International Journal of Forecasting*, 34(2), 235-248.
- Liu, E. M. (2013). Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China. *Review of Economics and Statistics*, 95(4):1386–1403.
- Loginova, O. (2009). "Exposure order effects and advertising competition". *Journal of Economic Behavior & Organization*, 71(2), 528-538.
- Molden, D. C. (2014). "Understanding priming effects in social psychology: What is "social priming" and how does it occur?". *Social Cognition*, 32(Supplement), 1-11.
- O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., and Rakow, T. (2006). "Uncertain judgements: eliciting experts' probabilities". John Wiley & Sons.
- Paula, Á. D., Shapira, G., and Todd, P. E. (2014). "How beliefs about HIV status affect risky behaviors: Evidence from Malawi". *Journal of Applied Econometrics*, 29(6), 944-964.
- Rojas-Barahona, C.; Zegers, B. and Förster, C. E. (2009). "La escala de autoestima de Rosenberg: Validación para Chile en una muestra de jóvenes adultos, adultos y adultos mayores". *Rev. Méd. Chile*, 137: 791-800.
- Rosenberg, M. and Pearlin, L. I. (1978). Social class and self-esteem among children and adults. *American Journal of Sociology*, 84(1), 53–77.
- Schlag, K. H., Tremewan, J., & Van der Wee, J. J. (2015). A penny for your thoughts: A survey of methods for eliciting beliefs. *Experimental Economics*, 18(3), 457-490.
- Schubert, R., Brown, M., Gysler, M. and Brachinger, H. W. (1999). Financial decision-making: are women really more risk-averse? *The American Economic Review*, 89(2):381–385.
- Simon, H. (1986). "Alternative visions of rationality". In H. Arkes & K. Hammond (Eds), *Judgment and Decision Making*, pp. 3-35. Cambridge: Cambridge University Press.

- Strack, F. and Martin, L. L. (1987). "Thinking, judging, and communicating: A process account of context effects in attitude surveys". In H.-J. Hippiert, N. Schwarz, & S. Sudman (eds.), *Social information processing and survey methodology*. New York: Springer-Verlag.
- Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.
- Tubbs, R. M., Gaeth, G. J., Levin, I. P., & Van Osdol, L. A. (1993). Order effects in belief updating with consistent and inconsistent evidence. *Journal of Behavioral Decision Making*, 6(4), 257-269.
- Van der Klaauw, W. and Wolpin, K. I. (2008). Social security and the retirement and savings behavior of low-income households. *Journal of econometrics*, 145(1-2), 21-42.
- Weiss, Y. (1972). The risk element in occupational and educational choices. *Journal of Political Economy*, 80(6):1203–1213.
- Wiseman, D. B. and Levin, I. P. (1996). Comparing risky decision making under conditions of real and hypothetical consequences. *Organizational behavior and human decision processes*, 66(3):241–250.
- Wooldridge, J. M. (2016). "Introductory Econometrics: A modern approach". Nelson Education.

Appendix

Table A1. Choices, Payments and Probability of Event

Choices	Payments	Probability
1	A: \$300,000 Chilean pesos	A: 100%
	B: \$700,000 or \$100,000 Chilean pesos	B: 50%
2	A: \$300,000 Chilean pesos	A: 100%
	B: \$650,000 or \$150,000 Chilean pesos	B: 50%
3	A: \$300,000 Chilean pesos	A: 100%
	B: \$600,000 or \$200,000 Chilean pesos	B: 50%

Table A2 Balance test - Pre-treatment characteristics (Baseline)

Variables	Means		Standard Deviations		Observations		Difference p-value
	Q1	Q2	Q1	Q2	Q1	Q2	
Fotrab beneficiary (=1)	0.51	0.52	0.50	0.50	2246	2266	0.75
Age	32.69	32.70	10.14	10.34	2053	2075	0.96
Male (=1)	0.35	0.34	0.48	0.48	2246	2266	0.80
Years of schooling	11.48	11.57	1.93	1.78	2052	2075	0.13
Household size	4.20	4.26	1.66	1.69	2053	2075	0.28
Unemployed	0.17	0.18	0.38	0.39	2048	2065	0.35
Monthly labour income	166,026	170,350	106,547	111,081	1088	1025	0.36
Lost job (=1, if any family member)	0.47	0.48	0.50	0.50	2050	2074	0.63
Lost agricultural income (=1, if any family member)	0.50	0.50	0.50	0.50	2050	2074	0.66
Index of negative shocks (max value = 7)	1.66	1.66	1.33	1.32	2050	2074	0.93
Big five: Agreeableness (max value = 7)	4.39	4.44	2.08	2.10	2050	2074	0.41
Big five: Conscientiousness (max value = 7)	5.52	5.51	1.85	1.86	2050	2074	0.88
Big five: Extraversion (max value = 7)	3.64	3.71	2.15	2.13	2050	2074	0.31
Big five: Neuroticism/Emotional stability (max value = 7)	2.39	2.44	1.63	1.66	2050	2074	0.33
Big five: Openness to experience (max value = 7)	6.16	6.11	1.39	1.45	2050	2074	0.32
Rosenberg: Self-Esteem score	33.96	34.03	4.93	4.94	2050	2074	0.66
Literacy score	0.88	0.88	0.18	0.17	2037	2064	0.68
Numeracy score	0.76	0.75	0.33	0.34	2047	2071	0.47
Spatial ability score	0.78	0.76	0.41	0.43	2050	2074	0.16
Raven test score	54.28	53.85	20.51	20.29	2049	2072	0.50
Intertemporal preference (=1, if preferring present)	0.37	0.35	0.48	0.48	2035	2066	0.21

Note: Negative shocks refer to the number of events experienced by any household member including: decrease in agricultural income or production, job loss, severe illness, divorce/abandonment, burglary, damages to the dwelling and death of a household member or relative. Differences are evaluated using standard t-tests.

Table A3 Balance test - Pre-treatment characteristics (Follow-Up 1)

Variables	Means		Standard Deviations		Observations		Difference p-value
	Q1	Q2	Q1	Q2	Q1	Q2	
Age	33.64	33.50	10.24	10.31	2246	2266	0.66
Years of schooling	11.65	11.72	1.88	1.77	2246	2266	0.21
Household size	4.12	4.18	1.68	1.72	2246	2266	0.23
Unemployed	0.08	0.09	0.28	0.28	2243	2262	0.69
Monthly labour income	239,501	234,805	133,751	130,624	1520	1552	0.32

Note: These variables were collected in Follow-Up 1 in the first module of the questionnaire where the order of questions was identical. Fotrab beneficiary and sex variables were only collected at baseline.

Table A4. Heterogeneous Effects on Expectations and Risk Aversion**Gender and Education**

OLS regression coefficients

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Expectations							Risk Aversion
	Education	Work	Family & Friends	Health	Crime	Wealth	All	
Questionnaire 2 (Q2)	0.0460 (0.114)	0.020 (0.302)	0.042 (0.105)	-0.220** (0.107)	-0.242** (0.112)	0.178 (0.115)	0.071 (0.115)	-0.039 (0.053)
Male (=1)	0.0886* (0.0482)	-0.189 (0.126)	-0.155*** (0.048)	0.063 (0.051)	0.121** (0.049)	-0.257*** (0.049)	-0.170*** (0.049)	-0.101*** (0.023)
Edu: Secondary incomplete	0.0612 (0.1000)	-0.097 (0.259)	0.168* (0.098)	-0.090 (0.123)	-0.214* (0.114)	0.280*** (0.107)	0.165 (0.103)	0.008 (0.046)
Edu: Secondary complete	0.0296 (0.0825)	-0.518** (0.217)	0.096 (0.078)	-0.179* (0.101)	-0.228** (0.099)	0.227** (0.088)	0.038 (0.086)	0.002 (0.038)
Edu:Tertiary	0.141 (0.103)	-0.320 (0.262)	0.042 (0.098)	-0.138 (0.119)	-0.270** (0.107)	0.150 (0.110)	0.033 (0.104)	-0.095** (0.048)
Q2*Male	-0.0867 (0.0650)	0.124 (0.169)	-0.009 (0.067)	-0.120* (0.065)	-0.064 (0.069)	0.109 (0.067)	0.030 (0.066)	0.053 (0.033)
Q2*Edu: Secondary Incomplete	0.178 (0.140)	0.191 (0.368)	0.005 (0.137)	0.232 (0.144)	0.338** (0.146)	-0.144 (0.144)	0.047 (0.144)	0.015 (0.066)
Q2*Edu: Secondary Complete	0.0951 (0.119)	0.300 (0.312)	0.017 (0.111)	0.303*** (0.112)	0.299** (0.121)	-0.141 (0.120)	0.054 (0.120)	0.004 (0.055)
Q2*Edu: Tertiary	0.146 (0.143)	0.042 (0.373)	0.020 (0.139)	0.186 (0.134)	0.318** (0.141)	0.034 (0.146)	0.096 (0.144)	-0.053 (0.069)
Constant	-0.145* (0.0793)	12.863*** (0.209)	-0.060 (0.074)	0.124 (0.099)	0.151 (0.093)	-0.168** (0.085)	-0.063 (0.082)	0.587*** (0.036)
Observations	4,127	4,127	4,127	4,127	4,127	4,127	4,127	4,072
R-squared	0.008	0.007	0.008	0.003	0.005	0.016	0.012	0.014

Note: Questionnaire 2 (Q2) corresponds to the questionnaire where the most recent information prior to the collection of expectations and risk aversion measures is on non-cognitive skills. Questionnaire 1 (Q1) corresponds to the

reference category in this regression. In Q1, the most recent information prior to the collection of expectations and risk aversion is on past experiences with labour market and schooling outcomes. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5. Heterogeneous Effects on Expectations and Risk Aversion
Unemployment and Loss of Income/Job
 OLS regression coefficients

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Expectations							Risk Aversion
	Education	Work	Family & Friends	Health	Crime	Wealth	All	
Questionnaire 2 (Q2)	0.123** (0.0541)	0.244* (0.139)	0.091* (0.051)	0.002 (0.048)	0.039 (0.047)	0.066 (0.054)	0.130** (0.052)	-0.027 (0.026)
Male (=1)	0.0388 (0.0325)	-0.129 (0.085)	-0.163*** (0.034)	0.006 -0.033	0.092*** (0.035)	-0.208*** (0.033)	-0.160*** (0.033)	-0.077*** (0.016)
Edu: Secondary incomplete	0.152** (0.0703)	-0.005 (0.184)	0.173** (0.069)	0.012 (0.075)	-0.058 (0.075)	0.217*** (0.072)	0.191*** (0.072)	0.016 (0.033)
Edu: Secondary complete	0.0702 (0.0597)	-0.348** (0.157)	0.110** (0.056)	-0.045 (0.059)	-0.094 (0.063)	0.168*** (0.061)	0.072 (0.061)	0.006 (0.028)
Edu:Tertiary	0.207*** (0.0720)	-0.262 (0.188)	0.060 (0.070)	-0.064 (0.070)	-0.125* (0.073)	0.179** (0.074)	0.091 (0.073)	-0.117*** (0.034)
Have lost job or income	0.0780 (0.0495)	-0.077 (0.127)	0.074 (0.045)	0.042 (0.045)	0.060 (0.043)	-0.004 (0.047)	0.042 (0.048)	-0.003 (0.023)
Unemployed	0.0933 (0.0615)	-0.313* (0.172)	-0.086 (0.059)	0.017 (0.060)	-0.026 (0.054)	0.000 (0.058)	-0.064 (0.063)	-0.016 (0.030)
Q2*Have lost job or income	0.00761 (0.0665)	0.066 (0.172)	-0.080 (0.065)	-0.032 (0.063)	-0.049 (0.063)	0.062 (0.066)	-0.004 (0.066)	-0.009 (0.033)
Q2*Unemployed	-0.0429 (0.0816)	0.034 (0.224)	0.068 (0.083)	0.050 (0.087)	0.048 (0.081)	0.006 (0.080)	0.035 (0.083)	0.055 (0.041)
Constant	-0.242*** (0.0681)	12.816*** (0.178)	-0.103 (0.063)	0.005 (0.069)	0.002 (0.070)	-0.141** (0.070)	-0.116* (0.069)	0.582*** (0.031)
Observations	4,108	4,108	4,108	4,108	4,108	4,108	4,108	4,053
R-squared	0.010	0.009	0.009	0.001	0.004	0.015	0.013	0.013

Note: Questionnaire 2 (Q2) corresponds to the questionnaire where the most recent information prior to the collection of expectations and risk aversion measures is on non-cognitive skills. Questionnaire 1 (Q1) corresponds to the reference category in this regression. In Q1, the most recent information prior to the collection of expectations and risk aversion is on past experiences with labour market and schooling outcomes. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6. Heterogeneous Effects on Expectations and Risk Aversion
Skills, Fluid Intelligence and Spatial Ability
 OLS regression coefficients

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Expectations							Risk Aversion
	Education	Work	Family & Friends	Health	Crime	Wealth	All	
Questionnaire 2 (Q2)	0.219 (0.192)	0.337 (0.520)	-0.070 (0.197)	-0.217 (0.183)	-0.299 (0.189)	0.157 (0.193)	0.089 (0.191)	-0.049 (0.096)
Male (=1)	0.0399 (0.0329)	-0.119 (0.085)	-0.153*** (0.034)	0.010 (0.033)	0.093*** (0.035)	-0.196*** (0.034)	-0.149*** (0.033)	-0.075*** (0.017)
Edu: Secondary incomplete	0.133* (0.0706)	0.033 (0.186)	0.189*** (0.069)	0.025 (0.076)	-0.052 (0.077)	0.226*** (0.073)	0.205*** (0.073)	0.011 (0.034)
Edu: Secondary complete	0.0576 (0.0604)	-0.299* (0.159)	0.128** (0.057)	-0.027 (0.061)	-0.087 (0.065)	0.189*** (0.062)	0.094 (0.062)	0.003 (0.028)
Edu:Tertiary	0.192*** (0.0731)	-0.190 (0.190)	0.084 (0.071)	-0.040 (0.073)	-0.112 (0.075)	0.210*** (0.075)	0.124* (0.073)	-0.129*** (0.035)
Numeracy Score	-0.0476 (0.0766)	0.013 (0.195)	-0.018 (0.071)	-0.014 (0.070)	-0.051 (0.075)	-0.101 (0.072)	-0.063 (0.072)	-0.067* (0.036)
Literacy Score	-0.0668 (0.140)	-0.153 (0.379)	0.024 (0.137)	-0.008 (0.153)	-0.074 (0.155)	0.051 (0.142)	-0.010 (0.144)	0.065 (0.067)
Raven test: lowest tercile	0.0356 (0.0612)	0.231 (0.160)	0.101* (0.059)	0.110** (0.054)	0.078 (0.051)	0.121** (0.060)	0.148** (0.061)	-0.035 (0.029)
Raven test: middle tercile	-0.0184 (0.0624)	0.148 (0.161)	0.058 (0.057)	0.007 (0.049)	0.068 (0.051)	0.033 (0.059)	0.063 (0.061)	-0.042 (0.030)
Spatial test: Correct (1 question)	0.0961 (0.0616)	-0.213 (0.150)	-0.095* (0.056)	-0.029 (0.057)	-0.020 (0.058)	0.014 (0.058)	-0.050 (0.060)	0.049* (0.028)
Q2*Numeracy Score	0.0615 (0.102)	0.166 (0.265)	-0.004 (0.101)	-0.066 (0.096)	0.018 (0.102)	0.112 (0.102)	0.079 (0.100)	0.114** (0.050)
Q2*L iteracy Score	-0.0159 (0.191)	-0.029 (0.512)	0.094 (0.195)	0.276 (0.182)	0.293 (0.191)	-0.089 (0.192)	0.037 (0.192)	-0.105 (0.095)
Q2*Raven test: lowest tercile	-0.123 (0.0821)	-0.151 (0.213)	-0.075 (0.083)	-0.097 (0.079)	-0.043 (0.077)	-0.072 (0.081)	-0.119 (0.081)	0.081** (0.041)
Q2*Raven test: middle tercile	-0.000870 (0.0834)	-0.216 (0.218)	-0.066 (0.082)	-0.016 (0.073)	-0.008 (0.079)	0.035 (0.081)	-0.046 (0.082)	0.068 (0.042)
Q2*Spatial test: Correct (1 question)	-0.0956 (0.0806)	-0.030 (0.201)	0.122 (0.078)	0.077 (0.078)	0.081 (0.076)	-0.048 (0.080)	0.020 (0.080)	-0.026 (0.038)
Constant	-0.153 (0.150)	12.814*** (0.414)	-0.086 (0.145)	0.009 (0.169)	0.096 (0.165)	-0.207 (0.154)	-0.111 (0.155)	0.563*** (0.072)
Observations	4,092	4,092	4,092	4,092	4,092	4,092	4,092	4,038
R-squared	0.009	0.010	0.010	0.004	0.005	0.017	0.015	0.017

Note: Questionnaire 2 (Q2) corresponds to the questionnaire where the most recent information prior to the collection of expectations and risk aversion measures is on non-cognitive skills. Questionnaire 1 (Q1) corresponds to the reference category in this regression. In Q1, the most recent information prior to the collection of expectations and risk aversion is on past experiences with labour market and schooling outcomes. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A7. Heterogeneous Effects on Expectations and Risk Aversion**Non-Cognitive Skills****OLS regression coefficients**

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Expectations							Risk Aversion
	Education	Work	Family & Friends	Health	Crime	Wealth	All	
Questionnaire 2 (Q2)	0.00593 (0.130)	0.843** (0.343)	-0.119 (0.124)	-0.207 (0.130)	-0.104 (0.121)	0.172 (0.128)	0.121 (0.123)	-0.016 (0.065)
Male (=1)	0.0389 (0.0326)	-0.126 (0.085)	-0.158*** (0.033)	-0.001 (0.032)	0.088** (0.035)	-0.199*** (0.033)	-0.155*** (0.033)	-0.076*** (0.017)
Edu: Secondary incomplete	0.129* (0.0704)	-0.073 (0.184)	0.160** (0.069)	0.017 (0.075)	-0.055 (0.075)	0.194*** (0.072)	0.163** (0.072)	0.016 (0.033)
Edu: Secondary complete	0.0482 (0.0601)	-0.450*** (0.156)	0.093* (0.056)	-0.027 (0.061)	-0.084 (0.063)	0.141** (0.061)	0.036 (0.061)	0.002 (0.028)
Edu:Tertiary	0.180** (0.0723)	-0.418** (0.186)	0.036 (0.070)	-0.041 (0.073)	-0.115 (0.073)	0.148** (0.074)	0.042 (0.072)	-0.124*** (0.034)
Agreeableness (=1, somehow agree)	0.0799* (0.0471)	0.142 (0.121)	-0.038 (0.044)	-0.017 (0.046)	-0.037 (0.044)	0.076* (0.045)	0.048 (0.046)	-0.000 (0.022)
Conscientiousness (=1, somehow agree)	-0.0784 (0.0538)	-0.087 (0.140)	-0.045 (0.053)	-0.191*** (0.061)	-0.041 (0.052)	0.057 (0.054)	-0.039 (0.053)	-0.007 (0.026)
Extraversion (=1, somehow agree)	-0.0538 (0.0509)	0.110 (0.129)	-0.018 (0.047)	0.004 (0.045)	-0.081* (0.042)	0.036 (0.048)	0.005 (0.048)	-0.026 (0.024)
Neuroticism (=1, somehow agree)	-0.0706 (0.0744)	-0.051 (0.203)	0.038 (0.068)	0.131 (0.092)	0.159* (0.087)	0.013 (0.069)	0.022 (0.072)	0.006 (0.035)
Openness to experience (=1, somehow agree)	0.000881 (0.0772)	0.505** (0.212)	-0.024 (0.070)	0.019 (0.078)	0.001 (0.070)	0.036 (0.074)	0.079 (0.074)	-0.010 (0.036)
Rosenberg self-esteem: lowest tercile	-0.180*** (0.0612)	-0.651*** (0.159)	-0.196*** (0.056)	0.042 (0.051)	-0.013 (0.051)	-0.131** (0.058)	-0.267*** (0.059)	-0.050* (0.030)
Rosenberg self-esteem: middle tercile	-0.101* (0.0589)	-0.285* (0.147)	-0.090* (0.054)	0.072 (0.053)	0.067 (0.055)	-0.041 (0.055)	-0.107* (0.055)	-0.047* (0.029)
Q2*Agreeableness (=1, somehow agree)	-0.0457 (0.0632)	-0.329** (0.164)	-0.058 (0.063)	0.038 (0.063)	0.016 (0.062)	-0.111* (0.062)	-0.124** (0.062)	0.006 (0.032)
Q2*Conscientiousness (=1, somehow agree)	0.0370 (0.0716)	0.089 (0.187)	0.079 (0.073)	0.251*** (0.072)	0.013 (0.074)	-0.049 (0.073)	0.050 (0.072)	0.015 (0.036)
Q2*Extraversion (=1, somehow agree)	0.0628 (0.0676)	-0.061 (0.172)	0.032 (0.066)	-0.046 (0.064)	0.049 (0.061)	0.068 (0.065)	0.048 (0.065)	-0.010 (0.033)
Q2*Neuroticism (=1, somehow agree)	0.123 (0.0985)	0.024 (0.267)	-0.061 (0.093)	-0.138 (0.112)	-0.112 (0.110)	-0.058 (0.096)	-0.040 (0.096)	-0.039 (0.049)
Q2*Openness to experience (=1, somehow agree)	0.0526 (0.103)	-0.623** (0.272)	0.051 (0.095)	0.060 (0.090)	0.105 (0.091)	-0.013 (0.100)	-0.049 (0.097)	-0.014 (0.050)
Q2*Rosenberg self-esteem: lowest tercile	0.0187 (0.0819)	0.302 (0.209)	0.164** (0.079)	-0.029 (0.072)	0.026 (0.073)	0.044 (0.079)	0.137* (0.079)	0.001 (0.041)
Q2*Rosenberg self-esteem: middle tercile	0.0528 (0.0796)	-0.006 (0.200)	0.090 (0.077)	-0.032 (0.077)	-0.008 (0.078)	0.010 (0.078)	0.050 (0.076)	-0.000 (0.040)
Constant	-0.00456 (0.112)	12.671*** (0.303)	0.125 (0.105)	0.094 (0.129)	0.066 (0.110)	-0.179 (0.111)	0.004 (0.109)	0.638*** (0.053)
Observations	4,123	4,123	4,123	4,123	4,123	4,123	4,123	4,068
R-squared	0.014	0.017	0.013	0.007	0.007	0.021	0.021	0.016

Note: Questionnaire 2 (Q2) corresponds to the questionnaire where the most recent information prior to the collection of expectations and risk aversion measures is on non-cognitive skills. Questionnaire 1 (Q1) corresponds to the reference category in this regression. In Q1, the most recent information prior to the collection of expectations and risk aversion is on past experiences with labour market and schooling outcomes. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1