UNIVERSITY OF WARSAW

Faculty of Economic Sciences

Working Papers

No. 24/2022 (400)

Farmers' risk preferences in eleven European farming systems: A multi-country replication of Bocquého et al. (2014)

Jens Rommel, Julian Sagebiel, Marieke Cornelia Baaken, Jesús Barreiro-Hurlé, Douadia Bougherara, Luigi Cembalo, Marija Cerjak, Tajana Čop, Mikołaj Czajkowski, María Espinosa-Goded, Julia Höhler, Laure Kuhfuss, Carl-Johan Lagerkvist, Margaux Lapierre, Marianne Lefebvre, Bettina Matzdorf, Edward Ott, Antonio Paparella, Erika Quendler, Macario Rodriguez-Entrena, Christoph Schulze, Tanja Šumrada, Annika Tensi, Sophie Thoyer, Marina Tomić Maksan, Riccardo Vecchio, Marc Willinger, Katarzyna Zagórska

Warsaw 2022

UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
WORKING PAPERS

Farmers' risk preferences in eleven European farming systems: A multi-country replication of Bocquého et al. (2014)

Jens Rommel[a], Julian Sagebiel[b], Marieke Cornelia Baaken[c], Jesús Barreiro-Hurlé[d], Douadia Bougherara[e], Luigi Cembalo[f], Marija Cerjak[g], Tajana Čop[g], Mikołaj Czajkowski[h], María Espinosa-Goded[i], Julia Höhler[j], Laure Kuhfuss[k], Carl-Johan Lagerkvist[a], Margaux Lapierre[e], Marianne Lefebvre[l], Bettina Matzdorf[r], Edward Ott[m], Antonio Paparella[f], Erika Quendler[n], Macario Rodriguez-Entrena[o], Christoph Schulze[m], Tanja Šumrada[p], Annika Tensi[j], Sophie Thoyer[e], Marina Tomić Maksan[g], Riccardo Vecchio[f], Marc Willinger[e], Katarzyna Zagórska[h]

[a] *Department of Economics, Swedish University of Agricultural Sciences*
[b] *German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig*
[c] *Department of Environmental Politics, Helmholtz Centre for Environmental Research – UFZ*
[d] *European Commission, Joint Research Centre (JRC)*
[e] *CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro*
[f] *Department of Agricultural Sciences, University of Naples Federico II*
[g] *Faculty of Agriculture, University of Zagreb*
[h] *Faculty of Economic Sciences, University of Warsaw*
[i] *Department of Economic Analysis and Political Economy, Faculty of Economics and Business Sciences, University of Sevilla*
[j] *Business Economics Group, Wageningen University & Research*
[k] *Social Economic and Geographical Sciences Department, the James Hutton Institute*
[l] *GRANEM n° 7456, Université d'Angers*
[m] *Leibniz Centre for Agricultural Landscape Research (ZALF)*
[n] *Federal Institute of Agricultural Economics, Rural and Mountain Research*
[o] *WEARE - Water, Environmental, and Agricultural Resources Economics Research Group, Universidad de Córdoba*
[p] *Biotechnical Faculty, University of Ljubljana*
[r] *Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg and Institute of Environmental Planning, Leibniz University of Hannover*

Corresponding author: jens.rommel@slu.se

Abstract: We replicate Bocquého et al. (2014), who used multiple price lists to investigate risk preferences of 107 French farmers. We collect new data from 1,430 participants in eleven European farming systems. In agreement with the original study, farmers' risk preferences are best described by Cumulative Prospect Theory. Structural model estimates show that farmers in the new samples are, on average, less loss averse and more susceptible to probability distortion than in the original study. Explorative analyses indicate differences between estimation approaches, as well as heterogeneity between and within samples. We discuss challenges in replications of economic experiments with farmers across farming contexts.

Keywords: Risk Attitudes, Agriculture, Cumulative Prospect Theory, Expected Utility Theory, Artefactual Field Experiment

JEL codes: D81, D90, Q12

Aknowledgements and disclaimer: We are grateful to the authors of Bocquého et al. (2014). In particular, we thank Geraldine Bocquého for sharing all her material and for always responding so fast to our numerous questions. We would like to thank Ciaran Blanchflower, Olivia Stopek, Nino Cavallaro, and the REECAP network for their continued support of the project. Comments from three anonymous reviewers helped us to improve the paper.

Jens Rommel and Julian Sagebiel are grateful for funding received from the Swedish Research Council for Sustainable Development (FORMAS) for the project "Building resilience for Swedish food production under increasing risk of extreme weather" under registration number 2019-00993. Douadia Bougherara, Sophie Thoyer, and Marc Willinger acknowledge financial support from the University of Montpellier (SRUM) and from the French National Research Agency under the Investments for the Future Program, referred to as ANR-16-CONV-0004 (#DigitAg), and are grateful to F. Brun and N. Bernard Le Gall (ACTA), M. Beurey (APCA), E. Gourdain and C. Hannon (Arvalis), and L. Lemeur (UNPT) for their contribution to survey dissemination. Julia Höhler and Annika Tensi acknowledge financial support from Wageningen University & Research and the Business Economics Group. Katarzyna Zagórska would like to thank the Faculty of Economic Sciences, University of Warsaw, which funded the data collection in Poland with a grant for young scholars (Badania Młodych 2020, resources from the Ministry of Science and Higher Education). Laure Kuhfuss acknowledges funding received from the Scottish Government RESAS division in the framework of the 2016-2021 Strategic Research Programme. Marianne Lefebvre acknowledges funding received from the Pays de Loire Region through RFI Alliance Europa (2020-2022). She is also grateful to Alexandre Perais and Ivan Dufeu for their contribution to survey dissemination. Tanja Šumrada acknowledges financial support from the Slovenian Research Agency (research programme P4-0022 (B)) and is grateful to the Association of Slovenian Rural Youth for their help with data collection. The Croatian and Swedish teams acknowledge funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952303 (AgriFoodBoost).

The Italian team received a grant from the Project PRIN DRASTIC "Driving The Italian Agri-Food System Into A Circular Economy Model," PRIN-MIUR – Call 2017, funded by the Italian Ministry of Education, University and Research (MIUR). Macario Rodriguez-Entrena, María Espinosa-Goded, and Jesús Barreiro-Hurlé acknowledge the funding via a specific contract between the European Commission (EC) and the University of Cordoba (JRC/SVQ/2021/VLVP/0333 - Programming and data-gathering for replication of the Bocquého et al. (2013) experiment).

The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

Risk and uncertainty are at the core of many questions in agricultural economics, and researchers have long been interested in estimating farmers' risk preferences (e.g., Binswanger, 1980; Collins et al., 1991). Risk preferences are an important modeling input of broad interest to policy makers and insurance companies. There is a large diversity of approaches to study farmers' risk preferences, but the literature is scattered (Iyer et al., 2020). Few studies systematically investigate differences in risk preferences across countries and farming systems to capture heterogeneity and assess distributional effects of risk-related policies. Large multi-country datasets collected under consistent protocols are mostly unavailable, and some geographical regions and farming systems in Europe are underrepresented in the literature on risk preferences (Iyer et al., 2020).

To reveal risk preferences in controlled environments, researchers have often used experiments with incentivized gambles (e.g., Holt & Laury, 2002). Such gambles, as part of multiple price lists, have been widely applied to investigate farmers' risk preferences in the gain domain (Iyer et al., 2020). A widely cited study in the field of agricultural economics is Bocquého et al. (2014) who used the multiple price lists developed by Tanaka et al. (2010) to investigate risk preferences of a sample of French farmers. The authors compared two major theories of decision-making under risk: (1) Expected Utility Theory (EUT, von Neumann and Morgenstern, 1947), which captures risk tolerance through the curvature of the utility function, and (2) Cumulative Prospect Theory (CPT, Tversky & Kahneman, 1992), which is based on reference dependence, leading to gain-loss asymmetry in behavior towards risk, and probability weighting. Bocquého et al. (2014) used structural models to estimate EUT and CPT parameters, including models that adjust for socio-demographic heterogeneity. The authors found that farmers are risk averse in the gain domain under EUT and exhibit loss aversion and probability distortion under CPT. CPT provided a better overall fit to the data.

The main objective of this paper is to explore the robustness of the results by Bocquého et al. (2014) based on data gathered under a similar protocol on new populations from eleven European farming systems. We cover a wide range of farming systems and national contexts, namely arable farmers in Austria, Germany, the Netherlands, and Sweden; wine growers in Croatia; potato farmers in Northern France; organic farmers in the North-West of France; olive growers in Apulia (Southern Italy); young farmers in Slovenia; members of two olive oil cooperatives in Andalusia (Spain); and farmers of different specializations in Poland. Our study is a conceptual replication of an experiment (i.e., the same protocol is broadly followed, but applied to a different context or sample), but we also engage in verification and reanalysis of the original data (see Christensen et al., 2019 for a discussion of different forms or replication and Rahwan et al. 2019 for an example of a conceptual replication of an artefactual field experiment). We focus on the same questions as the original study: Are farmers' risk preferences better described by EUT or CPT? And, what are the parameters for standard specifications of structural models under these theories? Hence, our study also contributes to a better understanding of farmers' decision-making under behavioral biases (Dessart et al., 2019) and the broader debate on the replicability of prospect theory across different study populations (Ruggeri et al., 2020).

Theory, experimental design, data, and estimation strategy

Utility functions and weighting function

Bocquého et al. (2014) estimated three structural models: (1) an EUT power function model with a reflected utility function at zero with parameter *r*, (2) an EUT expo-power function model with a reflected utility function at zero with parameters *α* and *β*, and (3) a CPT model with parameters *σ*, *λ*, and *γ*. The authors also estimated the impact of socio-economic covariates on the size of these parameters. In the following, we shortly introduce the utility and weighting functions underlying the structural models.

Under the EUT power model, 1 – *r* is the Constant Relative Risk Aversion parameter (i.e., *r* is an anti-index of risk aversion for positive payoffs). The utility over payoffs *y* from a risky prospect *u(y)* is defined as follows:

$$u(y) = \begin{cases} y^r & if\ y \geq 0 \\ -(-y)^r & if\ y < 0 \end{cases}, \qquad (1)$$

For gains (*y* ≥ 0), *r* < 1 implies concavity and risk aversion. Because the function is reflected for losses, concavity and convexity reverse for *y* < 0. Note that this specification also implies reference dependence (but no loss aversion) due to the reflection of the function.

The EUT power model implies constant relative risk aversion. In contrast, the EUT expo-power function model allows for varying degrees of absolute and relative risk aversion (Saha, 1993). In expression (2), parameters *α* and *β* capture risk aversion for gains (y ≥ 0), assuming the following utility function:

$$u(y) = \begin{cases} [1 - \exp(-\beta y^{\alpha})]/\beta & if\ y \geq 0 \\ [1 - \exp(\beta(-y)^{\alpha})]/\beta & if\ y < 0 \end{cases}, \qquad (2)$$

CPT has three parameters: (1) *σ* is an anti-index of concavity for gains where values of *σ* < 1 indicate risk aversion in the gain domain; (2) *λ* is an index of loss aversion, where values of *λ* > 1 imply that the utility function is steeper in the loss domain (loss aversion); and (3) *γ* is an anti-index of probability distortion towards overweighting of small probabilities, where, for binary prospects, values of *γ* < 1 imply probability distortion (overweighting of small probabilities and underweighting of large probabilities).[i] Here, we assume the status quo as the only reference point, although there has been a long debate on the identification of different (endogenous) reference points in prospect theory (Kőszegi & Rabin, 2007; Barberis, 2013) or the possibility of more than one reference point (Koop & Johnson, 2012). The utility function of CPT is defined as follows:

$$u(y) = \begin{cases} y^{\sigma} & if\ y > 0 \\ 0 & if\ y = 0\ , \\ -\lambda(-y)^{\sigma} & if\ y < 0 \end{cases} \qquad (3)$$

Under CPT, probabilities of all gambles are weighted, and, here, we use the one-parameter weighting function of Prelec (1998) which is strictly increasing from the unit interval into itself, i.e., for any two probabilities $p_a > p_b$, it maintains the order for the assigned weights such that $\omega(p_a) > \omega(p_b)$. By defining $\omega(0) = p_0 = 0$ and $\omega(1) = p_1 = 1$, the start and end points of the weighted and unweighted probabilities are the same. For any $p > 0$ and $p \leq 1$ the probability weights ω are defined as follows:

$$\omega(p) = \ exp[-(-\ln p)^{\gamma}]. \qquad (4)$$

Bocquého et al. (2014) used maximum likelihood for the estimation of all three structural models, which we also follow in our analysis.[ii] This estimation strategy uses a latent choice index, which is the difference between the expected and prospect utilities for the gambles. Note that for $\lambda = \gamma = 1$, $\omega(p) = p$ and $\sigma = r$, i.e., the CPT model becomes the EUT power model. This feature allows for a direct comparison of model fit between the EUT and CPT specifications. For all models that reject the null of $\lambda = \gamma = 1$, CPT provides a better fit than EUT. Because CPT has more parameters, we will follow Bocquého et al. (2014) and also use the Bayesian information criterion (BIC) to decide on the better model fit.

Experimental design and sample of Bocquého et al. (2014)

The experiment of Bocquého et al. (2014) was based on a modification of the three multiple price lists designed by Tanaka et al. (2010) to approximate parameters for CPT in a three-parameter specification in rural Vietnam. The modification of Bocquého et al. (2014) consisted of multiplying the lottery stakes by ten and deleting two rows from the first price list. The Tanaka et al. (2010) task has been used in other studies with farmers (e.g., Liu, 2013; Bougherara et al., 2017; Sagemüller & Mußhoff, 2020; Villacis et al., 2021), but Bocquého et al. (2014) is one of the few studies dealing with CPT in European agriculture (see Bonjean, 2019 or Kreft et al., 2021 for other European examples). In each of the three multiple price lists (cf. table 1), participants had to choose the row at which they preferred to switch from the *safer lottery* (Option A) to the *riskier lottery* (Option B). The task of choosing the switching row, rather than picking a lottery row by row as in Holt and Laury (2002), prevents multiple switches. Therefore, without having to discard inconsistent responses, the task enforces monotonicity, which allows to approximate CPT parameters per respondent based on their switching points by the so-called mid-point approach (Tanaka et al., 2010).

Table 1. Multiple price lists used in this study

Row	**Option A**		**Option B**		**Expected payoff difference (Option A – Option B)**
Series 1	*Probability 30%*	*Probability 70%*	*Probability 10%*	*Probability 90%*	
1	400	100	680	50	77
2	400	100	750	50	70
3	400	100	830	50	60
4	400	100	930	50	52
5	400	100	1060	50	39
6	400	100	1250	50	20
7	400	100	1500	50	– 5
8	400	100	1850	50	– 40
9	400	100	2200	50	– 75
10	400	100	3000	50	– 155
11	400	100	4000	50	– 255
12	400	100	6000	50	– 455
Series 2	*Probability 90%*	*Probability 10%*	*Probability 70%*	*Probability 30%*	
1	400	300	540	50	– 3
2	400	300	560	50	– 17
3	400	300	580	50	– 31
4	400	300	600	50	– 45
5	400	300	620	50	– 59
6	400	300	650	50	– 80
7	400	300	680	50	– 101
8	400	300	720	50	– 129

9	400	300	770	50	– 164
10	400	300	830	50	– 206
11	400	300	900	50	– 255
12	400	300	1000	50	– 325
13	400	300	1100	50	– 395
14	400	300	1300	50	– 535
Series 3	*Probability 50%*	*Probability 50%*	*Probability 50%*	*Probability 50%*	
1	250	– 40	300	– 210	60
2	40	– 40	300	– 210	– 45
3	10	– 40	300	– 210	– 60
4	10	– 40	300	– 160	– 85
5	10	– 80	300	– 160	– 105
6	10	– 80	300	– 140	– 115
7	10	– 80	300	– 110	– 130

Note: Adapted from Tanaka et al., 2010; Displayed units are points. Note that in accordance with the original study, the expected payoff difference was not shown to participants.

Bocquého et al. (2014) used a stratified random sampling strategy to build a representative sample of 107 farmers of the Burgundy Region in France. Their experiment was conducted face to face from February to June 2010 as part of a 2.5-hour long survey. Participating farmers were told that they all would receive a fraction of Euro amounts displayed in lotteries and that this fraction was predetermined and hidden in a closed envelope. Payments were calculated with an exchange rate of two percent of the amounts displayed as points (Euro) in table 1. The exchange rate of two percent was revealed *after* farmers completed their decisions.

Protocol and adjustments of the replication

The replication idea emerged from discussions in the "Research Network on Economic Experiments for the Common Agricultural Policy," a group of researchers using and promoting experiments for the evaluation of agricultural policy.[iii] The idea was shared with researchers who had experience in data collection with farmers and could reasonably offer a sample of participants. These researchers were invited to a series of joint meetings in which the experimental protocols were developed and discussed. Eventually, eleven teams joined the efforts and were involved in designing the experiment and data collection in different farming populations.

A few adjustments to Bocquého et al. (2014)'s protocol were agreed upon. First, we only included the survey parts that were relevant to the study of risk preferences to reduce response time.

Second, we modified incentives in the experiment. We used a common point system to display rewards consistently across all samples and currencies. This allowed us to share materials and videos more easily across multiple samples, also enhancing experimental control. We adjusted exchange rates between points and monetary rewards to account for variation in opportunity costs of participation time in the respective samples. The goal was to achieve 150 to 200 percent of a typical participant's opportunity costs for 20 minutes (see supplementary material for more details per sample). We also allowed paying out higher amounts to only a fraction of participants to limit administrative costs without losing the incentive effect (as for instance in Rommel et al., 2019; see Charness et al., 2016 for a general discussion).

Third, in contrast to Bocquého et al. (2014), in all instances, we revealed the exchange rate from points to monetary rewards *before* the experiment started. We believe that this is a more transparent procedure, beneficial in terms of experimental control and unobserved heterogeneity, because respondents are less likely to form heterogeneous beliefs on their lottery payments.[iv] Revealing the exchange rate only ex-post can be perceived as even more intransparent in online studies than in face-to-face studies, and we wanted to remove this additional confound.

Fourth, due to the pandemic and to lower the cost of data collection, we chose to allow both face-to-face (as in the original study) and web-based experiments, as the option of a lab-in-the-field experiment was unavailable in many instances. Note that in many instances, farm population data for probabilistic sampling and the application of survey weights was not available or accessible to the researchers. For instance, in the Spanish case, the research team's inquiry to the regional government about making available an anonymized list of the Andalusian olive growers and the corresponding emails accounts was rejected due to concerns about personal data protection. Hence, we decided to accept convenience samples to allow for a broader coverage of more diverse farming systems and a larger total sample size.

Finally, we agreed on a set of key covariates, in common with Bocquého et al. (2014) as explanatory of risk preferences. These covariates are (1) the age of the respondent in years, (2) the number of children, (3) education beyond secondary school (dummy), (4) (self-stated) general trust towards other people (dummy), (5) the total arable area of the farm, (6) the proportion of land owned, and (7) whether the farm is a sole proprietorship or a society with only one associate (dummy). Whenever we refer to covariates in the models, we mean these seven variables that were also part of the original study. All teams followed the same procedure when gathering the data.[v]

A questionnaire was developed in English and then translated to national languages.[vi] A jointly produced, approximately four-minute long, instruction video explaining the task and payment procedures was offered to all online respondents. The video used one of the examples, and the instructions used the same examples as the original study. The video was also publicly screened in the Spanish and Austrian face-to-face data collections. Note that in some cases (e.g., France I), additional data were gathered *after* the experiment, serving other research purposes. We also asked about respondents' comprehension.[vii] All instructions and other material are available online (Anonymous, 2021a).

According to national regulations, no ethical approval was required for the study, nevertheless expedite ethical approval was obtained from the German Association of Experimental Economics as a joint commitment of the group to ethical research practices (link to certificate contains names, we will include it in the final version after review). The study design was pre-registered as a replication under the open science framework (Anonymous, 2021a).

We obtained informed consent from all respondents. The consent form, which had to be actively accepted by participants, explained the broad purpose, data treatment, and payment procedures, as well as an indication of the range of the variable component of the payment. It also contained contact information. No deception was used, and no personal data were recorded without consent. A debrief in the form of a summary of the results was offered to interested participants by email.

Samples and recruitment across countries and farming systems

There are few attempts to elicit risk preferences across a large number of countries in incentivized experiments with students (Vieider et al., 2015) or surveys of the general population (Falk et al., 2018; Meissner et al., 2022). Overall, these studies find large within and between-sample heterogeneity, highlighting the need to investigate risk preferences in many different contexts and generally rejecting the idea of homogeneous preferences.

The original Bocquého et al. (2014) study aimed at a probabilistic sample of a small area in Burgundy. In the replication, we had three types of sampling strategies: (1) targeting specific farming systems with homogeneous production (e.g., potato growers in Northern France or olive growers in Apulia, Italy), (2) randomly or non-randomly sampling the overall population of farmers at national (Sweden) or regional levels (Slovenia), or (3) targeting a specific type of farming practice within a smaller region (e.g., organic Farmers in North-West France, members of two olive grower cooperatives in Spain). Note that these strategies limit the comparability with the original study. All data were collected in the summer and fall of 2021. We provide a short narrative with more details for each of our samples in section 3 of the appendix. An overview of all samples is provided in table 2.

Table 2. Overview on study samples

Sample name	Target population	Sampling and survey mode	Payments	Sample size for analysis
Original study, BJR2014	Farms from 64 towns in Burgundy, France	Probabilistic sample from registry; face-to-face lab-in-the-field experiment	Cash payments to all participants (19 Euro on average, p.145)	107
Austria	Arable farmers in Austria (region: Lower Austria)	Convenience sample; face-to-face experiment in group meetings which were held in cooperation with the Chamber of Agriculture	Payments as vouchers for local farm shop to all participants (average of 11.87 Euro, ranging from 2.90 Euro to 65 Euro)	128
Croatia	Winegrowers and wine producers from Croatia	Convenience sample from web-scraped contact information; online survey	One in ten participants received a voucher (average of 11 Euro, ranging from 7 Euro to 19 Euro)	104
France I	Potato farmers mainly located in Northern France (Hauts de France, Grand Est and Normandie)	Convenience sample of farmers contacted through various networks, newsletters, and emails; online survey	Payments with vouchers to all participants (average of 26.36 Euro, ranging from 8.70 Euro to 195 Euro)	96
France II	Organic farmers (vegetable growers, livestock and crops) of North-West of France	Convenience sample of farmers contacted through agricultural chambers and networks of organic farmers via newsletters and a mailing list; online survey	Payments with vouchers to all participants (average of 20 Euro, ranging from 13 Euro to 27 Euro)	28
Germany	Arable farmers in Germany	Randomly selected farmers from database of a market research company; online survey	Bank transfer to all participants (average of 8.83 Euro, ranging from 2.90 Euro to 65 Euro)	153

Italy	Olive growers of Apulia region (Southern Italy)	Convenience sample; individual face-to-face interviews	Cash payments to all participants (average of 10.02 Euro, ranging from 2.90 Euro to 65 Euro)	130
Netherlands	Arable farmers in the Netherlands	Randomly selected farmers from database of a market research company; online survey	Bank transfer to all participants (average of 16.09 Euro, ranging from 4.35 Euro to 97.50 Euro)	160
Poland	Various farmers in Poland	Mixed convenience sample of online participants from market research company and face-to-face interviews (recruited by farm advisors)	Bank transfer to 94 eligible participants (average of 9.34 Euro, ranging from 4.24 Euro to 19.56 Euro)	169
Slovenia	Young farmers, members of the Slovenian rural youth association	Mixed convenience sample of online participants and face-to-face interviews during farmer events	Payments with vouchers to all participants (average of 9.15 Euro, ranging from 2.90 Euro to 45 Euro)	114
Spain	Members of two olive oil cooperatives in Andalusia	Self-selected sample of members invited to the meetings at the premises; survey filled online on individual mobile devices	Voucher payment to all to be used to purchase olive oil in the cooperative's shop (average of 15.70 Euro, ranging from 5.80 Euro to 36 Euro)	130
Sweden	All registered farming businesses with an email address	Simple random sample; personalized link to online survey	Bank transfer to one in ten participants (average of 132 Euro, ranging from 66 Euro to 202 Euro)	218

Note: Where applicable, local currencies were converted to Euro (1 Euro were approximately 4.60 Polish Złoty and approximately 10.30 Swedish Crowns at the time of the study). All newly collected samples can be classified as lab-in-the-field experiments.

Table 3 gives an overview of respondent farm characteristics pooled across all newly collected samples. Note that we include only covariates that were common across all samples and the original study. Disaggregated and additional summary statistics for each sample are presented in section 1 of the appendix. We also discuss how representative for the underlying populations the samples are in section 2 of the appendix.

Table 3. Summary statistics of pooled data

Variable	Description	N	Missings (%)	Mean	SD
Age	Age in years	1371	4.13	45.96	13.90
NbChildren	Number of children	1298	9.23	0.92	1.15
EducSup	= 1 if more than secondary education	1402	1.96	0.42	
Trust	= 1 if self-reported as trusting other people	1368	4.34	0.37	
FarmSize	Arable land area in 100 ha	1317	7.90	0.81	2.33
LandOwned	Proportion of land owned	1357	5.10	0.66	0.35
IndivOwner	= 1 if sole legal owner of the farm	1383	3.29	0.73	

Analysis for replication

The analysis combined the estimation of three pre-registered structural models with further explorative analysis. We obtained the data and code from the authors of the original study (for Stata), and we successfully verified all analyses (in Stata and R). The original study weighted responses by population level statistics (survey weights in Stata, see footnote 15, page 147 in Bocquého et al., 2014). Because we did not use probabilistic sampling, we also did not apply survey weights in our analysis. The original data and code for additional verification is available online (Anonymous, 2021b).

Results

Structural models

Tables 4, 5, and 6 present coefficient estimates and 95 percent confidence intervals in brackets for all structural models without covariates.[viii] We include the original study's estimates (denoted as BJR2014) with and without survey weights for better comparison. We also report statistics of model fit: the log-likelihood of a model without parameters (LL null), the log-likelihood of the reported conversion (LL converge), and the Bayesian information criterion (BIC).

Table 4. Structural estimates of EUT model

	New samples pooled	BJR2014	BJR2014 (weighted)	Austria	Croatia	France_I	France_II	Germany	Italy	Netherlands	Poland	Slovenia	Spain	Sweden
r	0.214	0.227	0.212	0.232	0.229	0.183	0.187	0.229	0.193	0.240	0.212	0.206	0.140	0.232
	[0.206; 0.223]	[0.201; 0.254]	[0.173; 0.251]	[0.202; 0.261]	[0.202; 0.257]	[0.137; 0.229]	[0.119; 0.256]	[0.208; 0.251]	[0.164; 0.223]	[0.218; 0.261]	[0.187; 0.237]	[0.173; 0.239]	[0.089; 0.192]	[0.214; 0.249]
LL null	-32492.814	-2397.471	-2397.471	-2831.234	-2377.753	-2165.515	-638.889	-3483.439	-2909.458	-3594.941	-3865.606	-2604.863	-2949.745	-4968.972
LL converge	-31304.487	-2294.777	-7462.748	-2782.017	-2242.197	-2146.947	-623.350	-3299.214	-2891.252	-3420.294	-3694.140	-2513.158	-2945.381	-4674.464
N choices	47190	3531	3531	4224	3432	3168	924	5049	4290	5280	5577	3762	4290	7194
N respondents	1430	107	107	128	104	96	28	153	130	160	169	114	130	218
BIC	62619.736	4597.723	14933.665	5572.383	4492.535	4301.954	1253.529	6606.956	5790.868	6849.160	7396.907	5034.548	5899.125	9357.810

Table 5. Structural estimates of EUT expo-power function model

	New samples pooled	BJR2014	BJR2014 (weighted)	Austria	Croatia	France_I	France_II	Germany	Italy	Netherlands	Poland	Slovenia	Spain	Sweden
α	0.217	0.293	0.288	0.213	0.254	0.193	0.206	0.230	0.180	0.199	0.240	0.232	0.195	0.213
	[0.207; 0.227]	[0.265; 0.321]	[0.252; 0.324]	[0.184; 0.241]	[0.224; 0.283]	[0.151; 0.235]	[0.133; 0.280]	[0.205; 0.256]	[0.138; 0.222]	[0.176; 0.223]	[0.216; 0.265]	[0.200; 0.263]	[0.153; 0.237]	[0.186; 0.240]
β	0.010	0.107	0.119	-0.081	0.063	0.041	0.068	0.003	-0.064	-0.164	0.076	0.080	0.206	-0.062
	[-0.016; 0.037]	[0.078; 0.135]	[0.076; 0.161]	[-0.187; 0.026]	[0.014; 0.111]	[-0.097; 0.180]	[-0.130; 0.265]	[-0.067; 0.072]	[-0.219; 0.090]	[-0.255; -0.072]	[0.027; 0.125]	[0.010; 0.150]	[0.094; 0.319]	[-0.147; 0.022]
LL null	-32492.814	-2397.471	-2397.471	-2831.234	-2377.753	-2165.515	-638.889	-3483.439	-2909.458	-3594.941	-3865.606	-2604.863	-2949.745	-4968.972
LL converge	-31303.356	-2226.730	-7218.557	-2776.620	-2234.563	-2146.308	-622.629	-3299.204	-2889.648	-3395.308	-3680.582	-2505.216	-2920.949	-4668.515
N choices	94380	7062	7062	8448	6864	6336	1848	10098	8580	10560	11154	7524	8580	14388
N respondents	1430	107	107	128	104	96	28	153	130	160	169	114	130	218
BIC	62629.623	4471.184	14454.839	5571.323	4486.795	4310.123	1260.302	6616.849	5797.410	6809.147	7379.803	5028.283	5860.012	9356.178

Table 6. Structural estimates of CPT model

	New samples pooled	BJR2014	BJR2014 (weighted)	Austria	Croatia	France_I	France_II	Germany	Italy	Netherlands	Poland	Slovenia	Spain	Sweden
σ	0.314	0.297	0.280	0.322	0.333	0.289	0.284	0.334	0.297	0.314	0.304	0.322	0.284	0.329
	[0.307; 0.320]	[0.276; 0.318]	[0.255; 0.306]	[0.297; 0.348]	[0.313; 0.354]	[0.254; 0.325]	[0.232; 0.337]	[0.318; 0.350]	[0.269; 0.324]	[0.294; 0.333]	[0.286; 0.322]	[0.298; 0.345]	[0.253; 0.315]	[0.315; 0.342]
λ	1.601	2.174	2.274	1.531	1.817	1.701	1.751	1.574	1.457	1.187	1.807	1.848	2.162	1.352
	[1.529; 1.674]	[1.852; 2.497]	[1.804; 2.744]	[1.316; 1.747]	[1.575; 2.059]	[1.358; 2.044]	[1.074; 2.428]	[1.386; 1.763]	[1.181; 1.733]	[0.979; 1.396]	[1.563; 2.051]	[1.577; 2.120]	[1.843; 2.480]	[1.185; 1.520]
γ	0.574	0.681	0.657	0.643	0.595	0.563	0.562	0.571	0.546	0.627	0.591	0.562	0.487	0.552
	[0.555; 0.594]	[0.580; 0.781]	[0.507; 0.806]	[0.579; 0.707]	[0.535; 0.655]	[0.464; 0.661]	[0.401; 0.723]	[0.516; 0.625]	[0.485; 0.607]	[0.566; 0.689]	[0.527; 0.656]	[0.498; 0.625]	[0.404; 0.570]	[0.506; 0.597]
LL null	-32492.814	-2397.471	-2397.471	-2831.234	-2377.753	-2165.515	-638.889	-3483.439	-2909.458	-3594.941	-3865.606	-2604.863	-2949.745	-4968.972
LL converge	-29400.747	-2148.738	-7027.904	-2664.898	-2037.316	-2064.864	-597.967	-3004.128	-2762.964	-3276.823	-3474.950	-2306.617	-2768.928	-4254.937
N choices	141570	10593	10593	12672	10296	9504	2772	15147	12870	15840	16731	11286	12870	21582
N respondents	1430	107	107	128	104	96	28	153	130	160	169	114	130	218
BIC	58837.075	4325.280	14083.611	5358.137	4102.351	4157.207	1219.715	6037.133	5554.315	6582.658	6979.076	4641.227	5566.243	8539.814

Recall that in the EUT power specification with a reflected utility function at zero, parameter *r* is an anti-index of risk aversion (in the gain domain). In the gain domain, for $r < 1$, the utility function is concave, i.e. participants are risk averse on average ($r = 1$ indicates risk neutrality, $r > 1$ indicates risk seeking behavior).

Under the EUT power model, responding farmers, in all samples, are risk averse in the gain domain, on average. All estimates of *r* are in a rather narrow range. The point estimates of *r* in five samples (Austria, Croatia, Germany, Netherlands, Sweden) are slightly higher than in the original study (BJR2014 unweighted), whereas six samples (France I, France II, Italy, Poland, Slovenia, and Spain) show lower *r* estimates, i.e., a higher degree of risk aversion in the gain domain and risk seeking in losses.

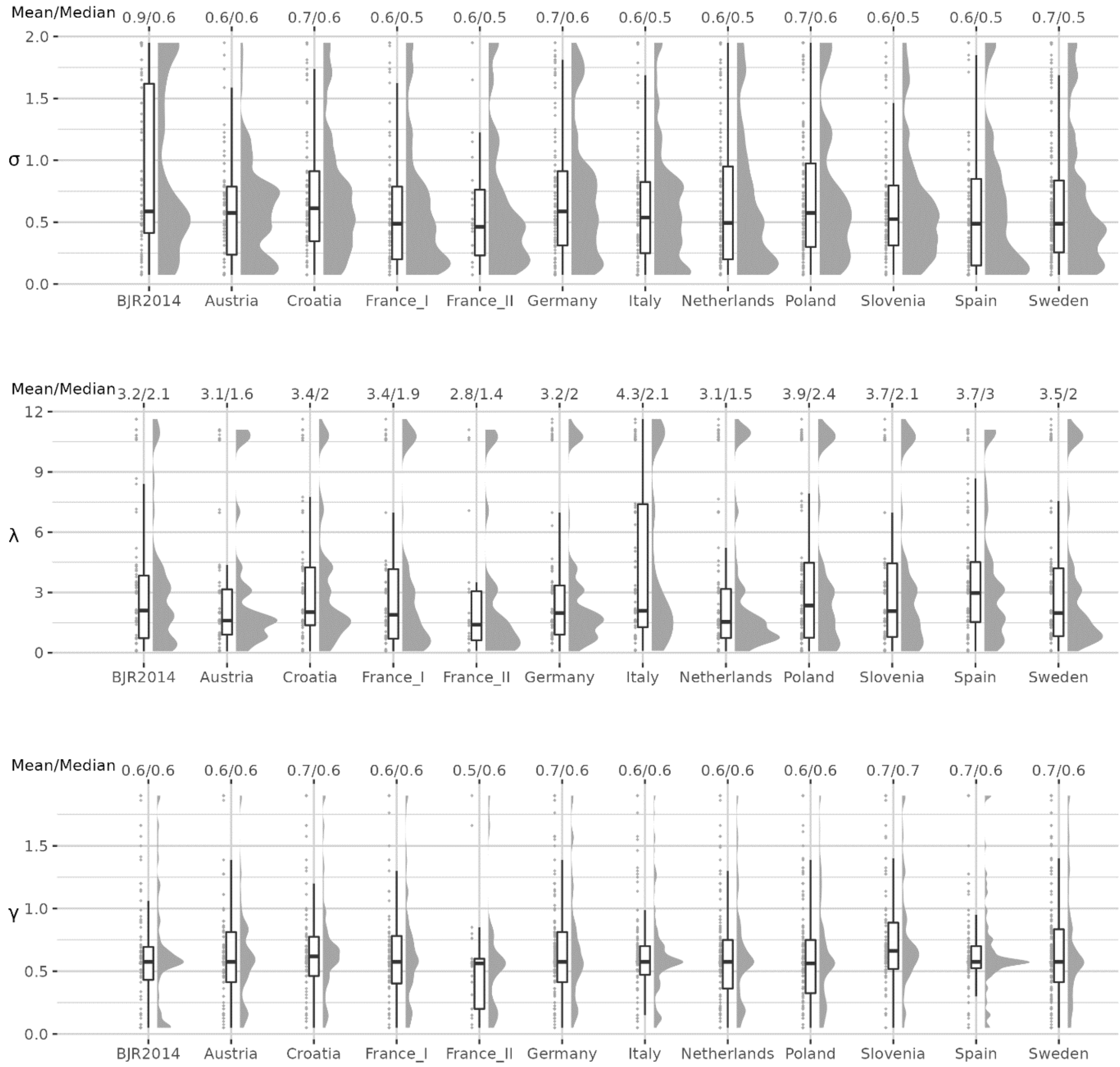
The EUT expo–power model estimates two parameters to allow for varying degrees of absolute and relative risk aversion. Compared with the original study, point estimates of α are lower for all samples, whereas point estimates of β are lower in all but one sample (Spain), in which they are higher. Note that the EUT expo–power model must satisfy $\alpha \times \beta > 0$. Although this is not the case for all point estimates (Austria, Italy, Sweden), it holds true in most instances for combinations of values in the 95 percent confidence intervals. We refrain from additional constraints on the model specification to avoid poor convergence of the demanding computation of the maximum likelihood models.

Under CPT, most samples (Austria, Croatia, Germany, Netherlands, Poland, Slovenia, Sweden) show values above .3 for point estimates of σ. The two French, the Spanish, and the Italian samples are below this value and slightly below the estimates based on data from the original study. For the newly added samples, all point estimates for λ are below two and show smaller values than the data from the original study, with only the exception of the Spanish sample, which is slightly above two. The Dutch and the Swedish samples have particularly low point estimates for λ; and in the Dutch case the null hypothesis of no loss aversion ($\lambda = 1$) is included in the confidence interval. Point estimates for γ are below the estimates for the original data in all samples, implying a greater distortion of probabilities across all samples compared to the original data. For all samples, we reject the null hypothesis of $\lambda = \gamma = 1$, i.e., risk preferences are better explained by CPT than by EUT. All BIC values for the three-parameter specification of the CPT model are lower than for the nested one-parameter EUT model, indicating a better fit of CPT even if model complexity is taken into account.

CPT mid-point approach

Figure 1 presents kernel density distribution estimates, mean and median values (on top of each figure) for the three CPT parameters σ, λ, and γ for all samples and the original study denoted as BJR2014 obtained by the mid-point approach. Additional summary statistics and high-resolution figures per country for the mid-point approach approximations are presented in sections 5.1 and 5.2 in the appendix.

Figure 1 Caption: Kernel density estimates of the distribution of CPT parameters, using the mid-point approach



For all samples, and in line with the original study, the approximated parameters differ from the structural models. For example, mean and median values of *σ* and *λ* are substantially higher than the estimated values in the structural models. Median values are closer to the structural estimates, because the distribution is highly skewed (with a few large values of *λ* affecting the means disproportionately). The range of parameter values is high. Although the order of the medians across the samples is not strongly affected, a few samples change their relative position when we compare medians from the mid-point approach with the structural estimates.

Additional robustness tests

In the appendix, we report and discuss additional analysis and robustness tests. First, we re-estimated all structural models with covariates whenever the sample size allowed us to do so (appendix section 4). Similar to Tanaka et al. (2010) and Bocquého et al. (2014), we include covariates in a regression with each of the parameters as the dependent variable to explore heterogeneity in the CPT parameters for the mid-point approach (appendix section 5.3). Both approaches indicate that observed heterogeneity in the selected common farm and individual characteristics do not have a high predictive power. Risk preferences were not strongly related to any of the covariates. Second, we re-estimated the structural models using only respondents who took at least six minutes to respond to the survey in online samples (appendix section 6.2). Third, we re-estimated the structural models removing respondents who indicated poor understanding or random choices (appendix section 6.3). Fourth, as in Bocquého et al. (2014), we re-estimated the structural models, using a reduced set of observations per respondent by including only the one or two rows per multiple price list at which the switch occurs (appendix section 6.1). Note that the results for all models are generally robust to the re-estimation.

Discussion

Summary and general discussion

In the upper panel of table 7, we compare the models and results in Bocquého et al. (2014) with the models and results in the newly collected samples. The first column presents an overview of the estimated models in Bocquého et al. (2014), the second column presents how we have dealt with this analysis, and the third column summarizes our main conclusions from the comparison. In the lower panel, we also present additional analysis (beyond the scope of Bocquého et al., 2014).

Table 7. Comparison of original study and replication results

Bocquého et al. (2014)	Approach in this study	Comparison
Structural modelling of EUT (power and expo-power) and CPT utility functions with Maximum Likelihood without covariates	Verified using original code with survey weights; replicated with and without survey weights on original and newly collected data (pre-registered analysis) (tables 4, 5, 6)	Similar results in all new samples; point estimates of λ and γ are slightly lower for CPT than in Bocquého et al. (2014)
Deriving individual parameters for CPT with mid-point approach	Verified using original code (figure 1 and appendix section 5.1); models conceptually replicated	Overall patterns and order of structural models are mostly maintained between samples, but similar to the original study, mean and medians differ rather substantially from structural models (especially for λ)
Structural modelling of EUT (power and expo-power) and CPT utility functions with Maximum Likelihood with covariates	Verified using original code; slightly adapted models conceptually replicated for consistency due to non-applicable covariates (appendix section 4)	Adapted model parameter estimates following from small differences in the samples due to list-wise missing covariates are not qualitatively different from models without covariates, similar to the original study, low explanatory power of covariates within samples and pooled data
Estimating the impact of covariates on CPT parameters derived from mid-point technique	Verified using original code; slightly adapted models conceptually replicated for consistency due to non-applicable covariates (appendix section 5.3)	Similar to the original study, low explanatory power of covariates within sample and pooled data
Robustness checks with reduced observations per respondent	Verified analysis in original code and applied to new samples (appendix section 6.1)	Results are robust
Robustness checks with different reference points in CPT	Verified analysis in original code, but not applied due to difficulty of defining alternative reference points coherently across samples	Not applicable

Robustness checks with varying exchange rates	Verified analysis in original code, but not applicable (because no exchange rate ambiguity in new samples as described above)	Not applicable
Further robustness checks, not applied in the original study		
	Removal of respondents taking less than six minutes in online samples (appendix section 6.2)	Results are robust
	Removal of respondents who self-reported to have poor understanding of or random decisions in the task (appendix section 6.3)	Results are robust

We verified all analysis, using the original code from the authors. We also replicated the original study results in eleven additional samples. Although our results are not substantially different, we noted some deviations of our estimates from the original study, as described in the table. However, we can also confirm the original study's main conclusion that farmers' risk preferences, on average, are substantially better described by CPT than by EUT, as under all approaches and for all samples, λ was consistently estimated as greater than one and γ was consistently estimated as smaller than one. In addition, the overall model statistics (and BIC in particular) also indicated a better fit of the data to CPT.

There was considerable heterogeneity both between and within samples. For example, in the structural models without covariates, the parameter for loss aversion λ ranged from a point estimate of more than two in the Spanish sample to less than 1.2 in the Dutch case. Using the mid-point technique, the same parameter λ had minimum, median, and maximum values of .08, 1.54, and 11.62 in the Netherlands; whereas the respective values were .08, 2.97, 11.1 in Spain. The mid-point technique can yield high values of λ for some respondents, which resulted in average values for λ of 3.09 and 3.66 in the Netherlands and Spain, respectively (see section 5.1 in the appendix). In other words, not only are the estimates heterogeneous within and across samples, but substantially different conclusions may also arise from using different estimation strategies on the same data. While the structural models offer a point estimate for the whole sample using an error term that accounts for individual choice errors, the mid-point technique can give direct insights into the distribution of parameters for CPT based on raw choices.

Because different estimation approaches can yield different results, we suggest the estimation of a large number of plausible models. This enables readers to assess the robustness and uncertainty associated with an estimate. More importantly, an open science approach is pivotal: the sharing of data and code allows the community to run further robustness tests and to integrate results in meta-analysis.

Our findings have important implications for policy. As stated by Colen et al. (2016), “behavioral findings (such as evidence of loss aversion), replicated over time and across domains, can safely be assumed to be valid everywhere and at any time and can therefore help understand reactions to policy of a large share of the EU farming population.” Here, our estimates provide plausible ranges of parameters, which can be included in agricultural policy models.

Many agricultural measures are based on farmers’ voluntary enrolment. Ex-ante evaluations are set-up to predict the expected uptake of such voluntary measures. This requires that behavioral drivers of far-reaching economic decisions or technology adoption, including risk and loss aversion, are better anticipated (Dessart et al., 2019; Spiegel et al., 2021). This type of ex-ante information can help to fine-tune policies so as to obtain the desired level of participation, or to optimize the outcome for a given budget. For instance, under the assumption that new measures involve greater risks, the high prevalence of risk and loss aversion signals the need to increase the compensation for agri-environmental measures or other green farm practices beyond cost forgone for risk-averse and loss-averse decision-makers which can pose a main barrier for transformational shifts in farming (Koetse and Bouma, 2022). Likewise, risk aversion and loss aversion have welfare and policy implications for insurance design. For instance, Dalhaus et al. (2020) show how taking into account loss aversion in insurance design may increase farmers’ uptake of well-designed insurance.

Challenges in multi-country replications of experiments with farmers

Conventional laboratory experiments with students are typically replicated under the exact same conditions, with only the timing and subjects being different. Uniform recruitment software and sampling, underlying population, localities, as well as experimental protocols and payment procedures can be used (e.g., Camerer et al., 2016). In contrast, artefactual field experiments, i.e., experiments with non-standard subjects (Harrison and List, 2004), such as farmers and other professionals, can create challenges for replication. For instance, in the replication attempt on dishonesty in the banking sector, Rahwan et al. (2019) had to work with a distinctive sample, which hampered a direct comparison of results with the original study, not least because of selection effects.

University laboratories typically work with long-term and experienced staff, whereas field experiments (outside of the laboratory) often build upon diverse teams, involving newly trained assistants. As a result, there are likely more confounds in such replications, such as small changes in wording, gestures and other cues from field staff, or even changes in the sampling frame and payment procedures. These many changes will almost inevitably differ across multiple samples, affecting experimental control and, hence, causal interpretations of differences in outcomes. One could be very strict in enforcing the exact same protocol across multiple countries (e.g., as in Vieider et al., 2016 or Dessart et al., 2021). However, in this study, we have chosen a more flexible approach of building a strong network of collaborators with a good mutual understanding of the case at hand, but also open to small adjustments in the experimental procedures. Thanks to this flexibility, we could include several research teams with different constraints for data collection and obtained a large dataset. Although crucial materials to replicate are not always accessible (Palm-Forster et al., 2019; Palm-Forster & Messer, 2021), luckily, we could build on the well-documented instructions, codes, datasets, and other material provided in Bocquého et al. (2014) and in later communication with the authors.

Harmonizing and improving infrastructures for social science research with farmers is an important task also for obtaining higher quality samples. For instance, some collaborators of this project could collect data through third-party farmer panels (Netherlands, Germany) or a general registry of all farming businesses (Sweden). Others (e.g., Spain or Italy), in contrast, had to work with convenience and snowball samples. A stronger grouping of cases and further harmonization of samples within these sub-groups (e.g., by farming system, region, or sampling procedure) could lead to better benchmarks for comparison by the removal of additional confounds. Although additional challenges such as different legal treatments and taxation of cash rewards will likely remain, a coordinated effort to build social science research infrastructures and networks for primary data collection with farmers could facilitate cross-national research (see Lefebvre et al., 2021 for more discussion on this). As a first step in this direction, farmers who are part of the European Union's Farm Accountancy Data Network could be invited to voluntarily commit to participate in experiments under high cash rewards on a regular basis.

Recruitment of and access to farmers representing a target population is a hard task (Weigel et al., 2021), and the sampling procedures likely affects options of statistical inference. While a snowballing approach and the use of convenience samples recruited by advertising the link to the online experiment in farmers' networks was successful for some subsamples (e.g., France), we would like to note that such open links must be used with caution for incentivized online experiments. Indeed, in the first attempt at data collection in Scotland, the survey link was hijacked and bots generated multiple successive answers until fully filling the maximum number of respondents set for the survey, probably to scam payments. The problem was early identified, and consequently unique links to the online experiment were shared with verified farmers, which we recommend for future online experiments. However, one should also be aware that this form of recruitment takes a lot of time and resources. Eventually, the final sample size was not large enough to be included in this study.

Online panel providers (Germany, Netherlands) or official farm registry data (Sweden) have the advantage that access is more restricted and individual invitation links to the survey instrument can be used. One can also more plausibly apply probabilistic sampling. However, with panel providers or registry data, response rates were well below ten percent, raising concerns about selection bias. Based on quotes obtained by the Spanish team, the price per response for a 20-minute online survey increased from than ten Euros for the general population to 75 Euros for a farmer sample. On top of that, many panel providers could not guarantee minimum sample sizes due to the limited availability of farmers in (3.9% in Spain).

The Swedish team received several inquiries from invited farmers asking about the seriousness of the study, because the described payments seemed dubious to many. However, our recruitment efforts through email have been generally more successful than in the study of Weigel et al. (2021) who sent email invitations to more than 4,700 respondents in two experiments and – in spite of substantial monetary compensation for taking part – did not receive a single response. One can only speculate that the high levels of trust towards research institutions and the familiarity with being contacted and dealing with errands online in Sweden may have led to this greater success rate. For future recruitments, one may also consider sending paper mail invitations which in in the United Kingdom has led to a response rate of more than seven percent (Howley & Ocean, 2021).

The costs of collecting data for one respondent differed substantially between countries. Research teams who used market research companies, paid approximately 50 Euro per farmer response, whereas those recruiting through their own networks often paid less than 20 Euros per response (including incentives). However, in the latter case, the additional transaction costs (meeting and convincing partners to invite farmer participants) can be substantial. Overall, our experiences confirm that consistent data collection for social science research with farmers across Europe remains a challenge (Lefebvre et al., 2021).

Farmers complained in some instances about the experimental task being tedious, abstract, or difficult to understand. However, according to a self-assessment, problems with comprehension and poorly motivated responses were not severe (appendix section 7). Notably, two samples that used face-to-face data collection (Austria and Spain) were at the opposite ends of the assessments for most of the questions (appendix section 7). Self-assessed comprehension and response quality were very high in Austria, whereas participants in Spain faced more difficulties with the task. Finally, Italy (who also used face-to-face data collection) was close to the overall mean for these questions. Hence, we cannot draw very strong conclusions regarding the use of online vs. face-to-face data collection, but we note less between-sample heterogeneity on these aspects for online data collection. It is an open question as to how far adjusting tasks for improved simplicity and comprehension, as well as realism and engagement (e.g., Charness & Viceisza, 2016; Meraner et al., 2018; Menapace et al., 2016; Villacis et al., 2021) could have improved these scores.

Model extensions and future research

Our analysis can be extended with the available data in many ways. Unlike the original study, we have not applied survey weights in the analysis. This could be included in further robustness tests, based on known or assumed information of key covariates in the underlying farmer populations, although this might not be possible in all cases. The use of survey weights is also hampered by the fact that we have not used probabilistic sampling.

One could also test for other reference points in the CPT models. Here, we have focused on the standard assumption of the utility function being kinked at the status quo, but other or more reference points could apply. We have also assumed that σ and γ do not differ between the loss and gain domains. One could use different lotteries to estimate a five-parameter CPT specification (with parameters $\sigma+$, $\sigma-$, λ, $\gamma+$, *and* $\gamma-$) and test for differences of the σ and γ parameters in relation to the reference point (as for instance in Bocquého et al., 2022).

We have followed the original and other studies in the application of the mid-point approach (Tanaka et al., 2010; Bocquého et al., 2014; Villacis et al., 2021), but the mid-point approach can only provide an approximation of the CPT parameters because it elicits intervals. Cameron and Huppert (1989) have used interval regression to correct biases that may arise from using mid-points rather than interval limits for payment card data in contingent valuation. In the same fashion, an interval regression could be applied to better account for covariates. The data also offer additional potential to explore how predictive elicited parameters are of real-world behavior under risk, such as the purchase of insurance or the use of irrigation (Charness et al., 2020). More can also be done to further explore observed and unobserved heterogeneity. By using a finite mixture model, for example, one could estimate propensities of respondents to either belonging to a EUT or a CPT group (Harrison & Rutström, 2009).

Our data and results are also useful for the integration with farm-level models. Although such models rarely consider risk and uncertainty (Huber et al., 2018), there is a growing trend towards a more realistic representation of economic agents in these models, including an increasing openness towards the behavioral economics and prospect theory literature in farm-level modeling (e.g., Appel & Balmann, 2019; Huber et al., 2022). Our study provides a rich data source for modelers to parametrize such models, including an overview of the distribution of these parameters and how key farm characteristics may correlate with them.

Conclusions

The objective of this study was to verify the analysis of Bocquého et al. (2014) and to test the robustness of their results in a replication in eleven European farming systems. Provided with the original code and data from the authors, we succeeded in verifying all parameters drawn from the original study. In line with the original study and the broader social science literature (Ruggeri et al., 2020), we confirmed that CPT provides a better fit to describe farmers' risk attitudes than EUT. This conclusion holds in all additional samples, albeit we also found considerable heterogeneity within and across samples. Similar to the original study, we faced the challenge of different methods yielding substantially different results for the CPT parameters. We conclude that pre-registration of a preferred specification, a wide range of additional robustness tests, and open methods and data are the best way to deal with these challenges.

References

Andersen, Steffen, Seda Ertaç, Uri Gneezy, Moshe Hoffman, and John A. List. 2011. “Stakes matter in ultimatum games.” *American Economic Review* 101(7): 3427–39.

Anonymous, 2021a. *Pre-registration of the replication study (anonymous version)*. https://osf.io/g7q92/?view_only=3ba60b86729d4befa4dad8650d282745

Anonymous, 2021b. *Data and code of the study on Github*. (We will provide a public link upon publication; for the review process, we share data and code as files)

Appel, Franziska, and Alfons Balmann. 2019. “Human behaviour versus optimising agents and the resilience of farms–Insights from agent-based participatory experiments with FarmAgriPoliS.” *Ecological Complexity* 40: 100731.

Barberis, Nicholas C. 2013. “Thirty years of prospect theory in economics: A review and assessment.” *Journal of Economic Perspectives* 27(1): 173–96.

Binswanger, Hans P. 1980. “Attitudes toward risk: Experimental measurement in rural India.” *American Journal of Agricultural Economics* 62(3): 395–407.

Bocquého, Géraldine, Florence Jacquet, and Arnaud Reynaud. 2014. “Expected utility or prospect theory maximisers? Assessing farmers' risk behavior from field-experiment data.” *European Review of Agricultural Economics* 41(1): 135–72.

Bocquého, Geraldine, Julien Jacob, and Marielle Brunette. 2022. “Prospect theory in multiple price list experiments: further insights on behaviour in the loss domain.” *Theory and Decision*, 1–44.

Bonjean, Isabelle. 2019. “Who are the Loss-Averse Farmers? Experimental Evidence from Structurally Estimated Risk Preferences,” Working Papers 291526, Katholieke Universiteit Leuven, Centre for Agricultural and Food Economics, https://ideas.repec.org/p/ags/kucawp/291526.html.

Bougherara, Douadia, Xavier Gassmann, Laurent Piet, and Arnaud Reynaud. 2017. “Structural estimation of farmers’ risk and ambiguity preferences: a field experiment.” *European Review of Agricultural Economics* 44(5): 782–808.

Camerer, Colin F., Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan, Emma Heikensten, Felix Holzmeister, Taisuke Imai, Siri Isaksson, Gideon Nave, Thomas Pfeiffer, Michael Razen, and Hang Wu. 2016. “Evaluating replicability of laboratory experiments in economics.” *Science* 351(6280): 1433–6.

Camerer, Colin F., and Robin M. Hogarth. 1999. “The effects of financial incentives in experiments: A review and capital-labor-production framework.” *Journal of Risk and Uncertainty* 19(1): 7–42.

Cameron, Trudy A., and Daniel D. Huppert. 1989. “OLS versus ML estimation of non-market resource values with payment card interval data.” *Journal of Environmental Economics and Management* 17(3): 230–46.

Charness, Gary, Thomas Garcia, Theo Offerman, and Marie C. Villeval. 2020. “Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory?” *Journal of Risk and Uncertainty* 60(2): 99–123.

Charness, Gary, Uri Gneezy, and Brianna Halladay. 2016. “Experimental methods: Pay one or pay all.” *Journal of Economic Behavior & Organization* 131: 141–50.

Charness, Gary, and Angelino Viceisza. 2016. “Three risk-elicitation methods in the field: Evidence from rural Senegal.” *Review of Behavioral Economics* 3(2): 145–71.

Christensen, Garett, Jeremy Freese, and Edward Miguel. 2019. *Transparent and reproducible social science research*. Oakland: University of California Press.

Colen, Liesbeth, Sergio Gomez y Paloma, Uwe Latacz-Lohmann, Marianne Lefebvre, Raphaële Préget, and Sophie Thoyer. 2016. “Economic experiments as a tool for agricultural policy evaluation: insights from the European CAP.” *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 64(4): 667–694.

Collins, Alan, Wesley N. Musser, and Robert Mason. 1991. “Prospect theory and risk preferences of Oregon seed producers.” *American Journal of Agricultural Economics* 73(2): 429–35.

Dalhaus, Tobias, Barry J. Barnett, and Robert Finger. 2020. “Behavioral weather insurance: Applying cumulative prospect theory to agricultural insurance design under narrow framing.” *PloS one* 15(5): e0232267.

Dessart, François J., Jesús Barreiro-Hurlé, and René Van Bavel. 2019. “Behavioural factors affecting the adoption of sustainable farming practices: a policy-oriented review.” *European Review of Agricultural Economics* 46(3): 417–71.

Dessart, François J., Jens Rommel, Jesús Barreiro-Hurlé, Fabian Thomas, Macario Rodríguez-Entrena, María Espinosa-Goded, Katarzyna Zagórska, Mikołaj Czajkowski, and René Van Bavel. 2021. “Farmers and the new green architecture of the Common Agricultural Policy: a behavioural experiment.” *JRC Science for Policy Report*, EUR 30706 EN, doi: 10.2760/718383, JRC123832.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global evidence on economic preferences." *The Quarterly Journal of Economics* 133(4): 1645–92.

Harrison, Glenn W., and John A. List. 2004. "Field experiments." *Journal of Economic Literature* 42(4): 1009–1055.

Harrison, Glenn W., and Elisabet Rutström. 2009. "Expected utility theory and prospect theory: One wedding and a decent funeral." *Experimental Economics* 12(2): 133–58.

Henningsen, Arne, and Ott Toomet. 2011. "maxLik: A package for maximum likelihood estimation in R." *Computational Statistics* 26(3): 443–58.

Hirschauer, Norbert, Sven Grüner, Oliver Mußhoff, Claudia Becker, and Antje Jantsch. 2021. "Inference using non-random samples? Stop right there!" *Significance* 18(5): 20–24.

Holt, Charles A., and Susan K. Laury. 2002. "Risk aversion and incentive effects." *American Economic Review* 92(5): 1644–55.

Howley, Peter, and Neil Ocean. 2021. "Can nudging only get you so far? Testing for nudge combination effects." *European Review of Agricultural Economics*: in press. doi: 10.1093/erae/jbab041

Huber, Robert, Martha Bakker, Alfons Balmann, Thomas Berger, Mike Bithell, Calum Brown, Adrienne Grêt-Regamey, Hang Xiong, Quang Bao Le, Gabriele Mack. Patrick Meyfroidt. James Millington, Birgit Müller, J. Gareth Polhill, Zhanli Sun, Roman Seidl, Christian Troost, and Robert Finger. 2018. "Representation of decision-making in European agricultural agent-based models." *Agricultural Systems* 167: 143–60.

Huber, Robert, Hang Xiong, Kevin Keller, and Robert Finger. 2022. "Bridging behavioural factors and standard bio-economic modelling in an agent-based modelling framework." *Journal of Agricultural Economics* 73(1): 35–63.

Iyer, Poorvi, Martina Bozzola, Stefan Hirsch, Manuela Meraner, and Robert Finger. 2020. "Measuring farmer risk preferences in Europe: a systematic review." *Journal of Agricultural Economics* 71(1): 3–26.

Koetse, Mark J., and Jetske A. Bouma. 2022. "Incentivizing a regime change in Dutch agriculture." *Environmental Innovation and Societal Transitions* 44: 265–282.

Koop, Gregory J., and Joseph G. Johnson. 2012. "The use of multiple reference points in risky decision making." *Journal of Behavioral Decision Making* 25(1): 49–62.

Kőszegi, B., and Matthew Rabin. 2007. "Reference-dependent risk attitudes." *American Economic Review* 97(4): 1047–73.

Kreft, Cordelia, Robert Huber, David Wuepper, and Robert Finger. 2021. “The role of non-cognitive skills in farmers' adoption of climate change mitigation measures.” *Ecological Economics* 189: 107169.

Lefebvre, Marianne, Jesus Barreiro-Hurlé, Ciaran Blanchflower, Liesbeth Colen, Laure Kuhfuss, Jens Rommel, Tanja Šumrada, Fabian Thomas, and Sophie Thoyer. 2021. “Can Economic Experiments Contribute to a More Effective CAP?” *EuroChoices* 20(3): 42–49.

Liu, Elaine 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): 1386–403.

Mechera-Ostrovsky, Tehila, Steven Heinke, Sandra Andraszewicz, and Jörg Rieskamp. 2022. ”Cognitive abilities affect decision errors but not risk preferences: A meta-analysis.” *Psychonomic Bulletin & Review*: 1–32.

Meissner, Thomas, Xavier Gassmann, Corinne Faure, and Joachim Schleich. 2022. *Individual characteristics associated with risk and time preferences – A multi-country representative survey*, https://arxiv.org/abs/2204.13664.

Menapace, Luisa, Gregory Colson, and Roberta Raffaelli. 2016. “A comparison of hypothetical risk attitude elicitation instruments for explaining farmer crop insurance purchases.” *European Review of Agricultural Economics* 43(1): 113–35.

Meraner, Manuela, Oliver Mußhoff, and Robert Finger. 2018. “Using involvement to reduce inconsistencies in risk preference elicitation.” *Journal of Behavioral and Experimental Economics* 73: 22–33.

von Neumann, John, and Oskar Morgenstern. 1947. *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Palm-Forster, Leah H., Paul J. Ferraro, Nicholas Janusch, Christian Vossler, and Kent D. Messer. 2019. “Behavioral and experimental agri-environmental research: methodological challenges, literature gaps, and recommendations.” *Environmental and Resource Economics* 73(3): 719–42.

Palm-Forster, Leah H., and Kent D. Messer. 2021. “Chapter 80 – Experimental and behavioral economics to inform agri-environmental programs and policies.” In *Handbook of Agricultural Economics, Vol. 5*, edited by Christopher B. Barrett, David R. Just, 4331-406. Amsterdam: Elsevier.

Prelec, Drazen. 1998. “The probability weighting function.” *Econometrica* 66(3): 497–527.

Rahwan, Zoe, Erez Yoeli, and Barbara Fasolo. 2019. "Heterogeneity in banker culture and its influence on dishonesty." *Nature* 575(7782): 345–9.

Rommel, Jens, Daniel Hermann, Malte Müller, and Oliver Mußhoff. 2019. "Contextual framing and monetary incentives in field experiments on risk preferences: Evidence from German farmers." *Journal of Agricultural Economics* 70(2): 408–25.

Ruggeri, Kai, Sonia Alí, Mari Louise Berge, Giulia Bertoldo, Ludvig D. Bjørndal, Anna Cortijos-Bernabeu, Clair Davison, Emir Demić, Celia Esteban-Serna, Maja Friedemann, Shannon P. Gibson, Hannes Jarke, Ralitsa Karakasheva, Peggah R. Khorrami, Jakob Kveder, Thomas Lind Andersen, Ingvild S. Lofthus, Lucy McGill, Ana E. Nieto, Jacobo Pérez, Sahana K. Quail, Charlotte Rutherford, Felice L. Tavera, Nastja Tomat, Chiara Van Reyn, Bojana Većkalov, Keying Wang, Aleksandra Yosifova, Francesca Papa, Enrico Rubaltelli, Sander van der Linden, and Tomas Folke. 2020. "Replicating patterns of prospect theory for decision under risk." *Nature Human Behaviour* 4(6): 622–33.

Sagemüller, Frederik, and Oliver Mußhoff. 2020. "Effects of Household Shocks on Risk Preferences and Loss Aversion: Evidence from Upland Smallholders of South East Asia." *The Journal of Development Studies* 56(11): 2061–78.

Saha, Atanu. 1993. "Expo-power utility: A 'flexible' form for absolute and relative risk aversion." *American Journal of Agricultural Economics* 75(4): 905–13.

Spiegel, Alisa, Wolfgang Britz, and Robert Finger. 2021. "Risk, risk aversion, and agricultural technology adoption — A novel valuation method based on real options and inverse stochastic dominance." *Q Open* 1(2): qoab016.

Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen. 2010. "Risk and time preferences: Linking experimental and household survey data from Vietnam." *American Economic Review* 100(1): 557–71.

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty* 5(4): 297–323.

Vieider, Ferdinand M., Mathieu Lefebvre, Ranoua Bouchouicha, Thorsten Chmura, Rustamdjan Hakimov, Michal Krawczyk, and Peter Martinsson. 2015. "Common components of risk and uncertainty attitudes across contexts and domains: Evidence from 30 countries." *Journal of the European Economic Association* 13(3): 421–52.

Villacis, Alexis H., Jeffrey R. Alwang, and Victor Barrera. 2021. "Linking risk preferences and risk perceptions of climate change: A prospect theory approach." *Agricultural Economics* 52(5): 863–77.

Weigel, Collin, Laura A. Paul, Paul J. Ferraro, and Kent D. Messer. 2021. "Challenges in recruiting US farmers for policy-relevant economic field experiments." *Applied Economic Perspectives and Policy* 43(2): 556–572.

---

[i] Note that Kahneman and Tversky allowed for σ to differ in the gain and loss domains. The same applies to γ in Prelec's probability weighting function.

[ii] The original code was written for Stata. We estimated all models in R, using the maxLik package for maximizing the likelihood functions of the structural models (Henningsen & Toomet, 2011).

[iii] For more information on the network, the reader is referred to www.reecap.org. For a discussion on experiments for agricultural policy-making see Colen et al. (2016).

[iv] Suppose there are two types of respondents: optimists (believing in a high payment) and pessimists (believing in a low payment). If the assumed size of the payment affects choices or noise, unobserved heterogeneity increases (because we do not know respondents' beliefs). While there is some evidence of stake size effects in ultimatum games (e.g., Andersen et al., 2011), the main concern with choices in risky gambles is noise (Camerer & Hogarth, 1999; Mechera-Ostrovsky et al., 2022).

[v] In addition to these covariates, Bocquého et al. (2014) included the proportion of the household income coming from another profession than farming, a dummy for deferred payments, a dummy for livestock, the proportion of idle land out of the arable area in 2009, a dummy if the farm has no successor despite looking for one, a dummy for farms located in the Northern part of the study area, and the importance of risk faced on soft wheat (1–5 score). We did not include these covariates, because they did not fit the more diverse farming contexts we were dealing with.

[vi] Due to time constraints, we did not back-translate the instructions and videos. The involved researchers were all experienced in field work with farmers and familiar with the used terminology. In all instances, multiple team members reviewed the texts for clarity and to closely match the English master version. In several instances, pre-tests were run with colleagues or farmers (e.g., a small pilot was conducted with five farmers in the Netherlands).

[vii] Respondents could indicate their agreement with the three statements "It was difficult to understand the task.", "My choices were random.", and "There were too many different lotteries." on a five-point scale (see appendix for more details).

[viii] Estimated standard errors are clustered per respondent. Strictly speaking, standard errors could be inflated and misspecified in the structural models due to enforced monotonous switching (choices between options are not independent within lotteries). To address this issue, we use only the observations for the rows per multiple price list for which the switch occurs (resulting in three to six observations per respondent instead of 33). We provide these

estimates as a robustness test in section 6.1 of the appendix. Note that for the mid-point approach this issue does not exist. Another issue is that we applied random sampling from a population only in Germany, the Netherlands, and Sweden. Hence, throughout, estimated standard errors (and subsequent statistics such as confidence intervals) must be viewed as an approximation of true sampling error (Hirschauer et al., 2021). Alternatively, one can interpret the point estimates exclusively for the given group of study subjects rather than as estimates of a population parameter. Of course, in this case, uncertainty based on sampling error is irrelevant.

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl