How Well Can Experts Predict Farmers' Risk Preferences ?

Henning Schaak, Jens Rommel, Julian Sagebiel, Jesus Barreiro-Hurlé, Douadia Bougherara, Luigi Cemablo, Marija Cerjak, Tajana Čop, Mikolaj Czajkowski, María Espinosa-Goded, et al.

How Well Can Experts Predict Farmers' Risk Preferences?

Henning Schaak
Department of Economics, Swedish University of Agricultural Sciences
henning.schaak@slu.se

Jens Rommel
Department of Economics, Swedish University of Agricultural Sciences
jens.rommel@slu.se

Julian Sagebiel
Biodiversity Economics, German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig
julian.sagebiel@idiv.de

Jesus Barreiro-Hurlé
European Commission, Joint Research Centre (JRC)
jesus.barreiro-hurle@ec.europa.eu

Douadia Bougherara
CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier
douadia.bougherara@inrae.fr

Luigi Cemablo
Department of Agricultural Sciences, University of Naples Federico II
cembalo@unina.it

Marija Cerjak
Faculty of Agriculture, University of Zagreb
mcerjak@agr.hr

Tajana Čop
Faculty of Agriculture, University of Zagreb
tcop@agr.hr

Mikołaj Czajkowski
Faculty of Economic Sciences, University of Warsaw
mc@uw.edu.pl

María Espinosa-Goded
Faculty of Economic and Business Science, University of Sevilla
megoded@us.es

Julia Höhler
Business Economics Group, Wageningen University & Research
julia.hoehler@wur.nl

Carl-Johan Lagerkvist
Department of Economics, Swedish University of Agricultural Sciences
carl-johan.lagerkvist@slu.se

Macario Rodriguez-Entrena
WEARE - Water, Environmental, and Agricultural Resources Economics Research Group, Universidad de Córdoba
mrentrena@uco.es

Annika Tensi
Business Economics Group, Wageningen University & Research
annika.tensi@wur.nl

Sophie Thoyer
CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier
sophie.thoyer@inrae.fr

Marina Tomić Maksan
Faculty of Agriculture, University of Zagreb
matomic@agr.hr

Riccardo Vecchio
Department of Agricultural Sciences, University of Naples Federico II
riccardo.vecchio@unina.it

Katarzyna Zagórska
Faculty of Economic Sciences, University of Warsaw
kzagorska@uw.edu.pl

Selected Paper prepared for presentation at the 2022 Agricultural & Applied Economics Association Annual Meeting, Anaheim, CA; July 31-August 2

Abstract

Risk is ubiquitous in agriculture and a core interest among agricultural economists. While farmers' risk preferences are well studied, there is limited knowledge on the perspectives of stakeholders on farmers' risk preferences. We address this gap by eliciting predictions from 561 experts, which allows us to understand how well these experts understand farmers' risk preferences. First, we compare the accuracy of predictions by distinguishing different groups of experts. Second, we investigate whether the risk preferences of farmers from different production systems differ in terms of predictability for the experts. Third, we examine the effectiveness of expert predictions by randomly assigning experts to different incentives schemes. We find that an international group of researchers in experimental economics provide the most accurate predictions if compared to farm advisors and other experts from different countries as well as students of agriculture. Differences in predictions across the eight samples of farmers from different production systems are small. Incentivizing predictions by either a tournament scheme (the best prediction receives a reward) or high accuracy (randomly selected participants are paid depending on the quality of their prediction) do not strongly affect accuracy, but tournament schemes show somewhat smaller standard deviations.

Keywords

Risk attitudes; Expert predictions; Multiple prices lists; Meta-science; Experimental economics

1 Introduction

Predictions of research results by experts can improve the effectiveness of the research process in the social sciences for at least three reasons (DELLAVIGNA ET AL., 2019). First, predictions offer a systematic way to elicit a community's ex-ante beliefs on a study, alleviating hindsight bias. Establishing a clear benchmark for what is known on what and by whom in combination with a debriefing can help to update experts' views. Second, a benchmark of what experts predict ex-ante can facilitate the acceptance of null results in particular when the null deviates from experts' views. Third, systematic and regular predictions from an expert community can facilitate more accurate predictions. Said predictions can then inform future research designs, for instance when selecting treatments for designing effective behavioral interventions (DELLAVIGNA and POPE, 2018b; MILKMAN ET AL., 2022). Furthermore, it has been shown that prior beliefs of experts and policy-makers can differ (VIVALT and COVILLE, 2021). In these cases, expert predictions can provide new information to policy-makers, leading to an update of beliefs.

Predictions of research results in Economics have often focused on laboratory experiments (DELLAVIGNA and POPE, 2018a,b). In Agricultural Economics, predictions of experimental outcomes have focused on narrow topics such as the behavior of German farmers under different treatments of a public goods game (ROMMEL ET AL., 2021), using professional academics and graduate students as experts. To the best of our knowledge, there is no comprehensive study eliciting what experts know about an important topic such as risk preferences in European agriculture (IYVER ET AL., 2020).

It is the first objective of this study to present results from a cross-country prediction study of farmers' risk preferences. Simply put, we investigate who knows what about risk preferences of farmers. Multiple samples of more than 500 experts in total (Polish, French, Croatian, and Italian farm advisors; Swedish students of agriculture; a group of mixed experts from Spain; as well as experimental economists) predicted the outcomes of multiple samples of farmers' risk preferences.

It is the second objective of this study to understand whether the behavior of some farmers is easier to predict than of others. Farmers, whose behavior was asked to predict, took part in an incentivized multiple price list (TANAKA ET AL., 2010) based on economic gambles to elicit their risk preferences (including wine growers in Croatia, olive farmers in Italy and Spain, potato growers in France as well as arable farmers in the Netherlands, Sweden, and Germany). By varying topical expertise and local context knowledge in our samples of forecasters, our analysis can focus on whose predictions are most accurate for whom. In other words, we want to assess whether the predictive accuracy differs by production system, as this allows us to evaluate in which cases experts may "fill-in" for farmer data.

It is the third objective of this study to understand better how to incentivize accurate predictions. Previous research on expert predictions has focused on the impact of elicitation formats. Notably, DELLAVIGNA ET AL. (2020) tested the impact of (1) reference values, (2) raw units vs. standard deviations, (3) sliders vs. text entry, and (4) different slider bounds had on expert evaluations, finding that only slider bounds have had a small impact on predictions. We augment this line of research by focusing on another important question: the role of incentives. Specifically, by randomly assigning participants to one of five conditions in a between-subjects design (denoted by numbers in brackets), we test two different tournament incentive rates against two schemes which incentivize accuracy independent of what others do. For the tournament incentives, the most accurate predictions within a reference group receive a (low [1] or high [2]). For accuracy-based incentive, the expected reward solely depends on the deviation from the actual outcome (low [3] or high [4] punishment for deviating from the true value). We compare these schemes to a control condition of rewards unrelated to accuracy [5]. Treatments 3-5 can be described as between subjects random incentivized systems (CLOT ET AL., 2018; CHARNESS ET AL., 2016).

In the next section, we will introduce the experimental design. Then, we present and discuss the results. In the final section, we present an outlook on additional analyses and some conclusions.

2 Experimental design and data

2.1 Experimental design

The to-be-predicted data from farmers were gathered as part of a large-scale cross-country effort to replicate the study of BOCQUÉHO ET AL. (2014) in different European Union member states (see ROMMEL ET AL., 2022 for details). In this study, which took place in the second half of 2021, farmers had to make a choice between riskier and safer options in the risk preferences elicitation task of TANAKA ET AL. (2010). Monotonous switching was enforced in this study, i.e., farmers could only indicate a single switch from option A to option B. The data collection for the prediction study took place after the farmer data were collected but before the outcomes of farmers' choices were known (in late 2021 and early 2022). Authors of the replication study were not allowed to take part in the prediction.

We obtained informed consent from all participants and used no deception in the study. Participants were offered a debrief by allowing them to subscribe to a short summary of the research results. We pre-registered basic analysis before data collection (see https://aspredicted.org/Z8Z_FV7).

For eight farmer samples, we asked experts to predict the average number of times a farmer participant in a specific sample would choose the safer option A for one of the multiple price lists. Hence, higher numbers indicate higher predicted average risk aversion. Farmers choosing the safer option A seven times or more are risk averse. We elicited predictions on a scale from

0 (farmers on average never choose the safer option A) to 12 (farmers on average always choose the safer option A). This format was perceived as the most intuitive by the research team. Predictions of means had to be entered with a one decimal point accuracy on a slider for each of the eight samples. Table 1 displays the price list, including the expected payoff difference (which was shown neither to forecasters nor farmers participating in the original study).

Table 1: Multiple price list used in this study and difference in expected value

Row	**Option A**		**Option B**		**Expected payoff difference (A – 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

Note: Adapted from Tanaka et al., 2010; Displayed units are experimental currency units.

Our main outcome variable of interest is the accuracy of the predictions, defined as a predictor's absolute deviation from a sample's actual average. Note that this definition implies that *smaller* values (lower bound at zero) indicate predictions with *higher* accuracy. Recall that we obtained eight predictions per participant. We used a nonparametric multi-comparison Kruskal-Wallis test to investigate whether the accuracy of different samples of forecasters come from the same underlying distribution. That is, we asked whether some forecasters are better or worse than others. We also used a Kruskal-Wallis test to investigate whether the distribution of accuracy differs for the farmer samples. That is, we ask whether some farmers' behavior is easier or more difficult to predict. Further, accuracy was used as the dependent variable in a linear regression model to explain accuracy by experience and knowledge in the subject field of risk and uncertainty in agriculture.

Accurate predictions were incentivized in four out of five treatments, which were implemented between subjects (see Table 2 for an overview). In treatment ACCLOW, one randomly selected participant from a group of 50 participants was offered a payment calculated as 300 Euro minus the squared deviation of one randomly selected prediction out of the total of eight predictions per participant. In treatment ACCHIGH, the payment was calculated as 300 Euro minus twice the squared deviation in order to test for incentive effects, i.e., in ACCHIGH deviations were punished relatively more. In TOURHIGH and TOURLOW (the two tournament schemes), payments of 300 and 100 Euro were offered to the best prediction (from a randomly selected sample) among a group of 50 participants. In CONTROL, a payment of 300 Euro was offered to a randomly selected participant from a group of 50 participants. We received between 100 and 150 responses per treatment. Hence, we offered payments to three participants per treatment for a total of 15 payments. Groups to decide on the winner were divided into equal size (i.e., the actual group size was a bit smaller than 50 which is equivalent to rounding up payments). We successfully contacted and exchanged banking details and executed payments with 10 out these 15 respondents. One respondent explicitly declined the payment, and four others did not respond to our attempt to contact them.

Table 2: Overview of the experimental treatments

Row	Type	Selection criterion for Payment	Payable amount	Relevant prediction
ACCLOW	Accuracy	Randomly selected	€ 300 – the squared deviation of the prediction from true value	Randomly selected
ACCHIGH	Accuracy	Randomly selected	€ 300 – two times the squared deviation of the prediction from true value	Randomly selected
TOURHIGH	Tournament	Most accurate prediction	€ 300	Randomly selected
TOURLOW	Tournament	Most accurate prediction	€ 100	Randomly selected
CONTROL	Control	Randomly selected	€ 300	None

2.2 Data

Data were collected through an online survey between 15 December 2021 and 28 January 2022. The survey was available in multiple languages (Croatian, English, French, German, Italian, Polish, and Spanish) and distributed through multiple channels, including researchers, advisor associations, and students, as well as networks of the authors. After the participants were welcomed and introduced to the survey's objectives, informed consent was obtained. Predictions were explained and elicited at the beginning of the survey. Depending on the treatment, the incentive mechanism was introduced. In a later part of the survey, a manipulation check asked to select the assigned incentive mechanism, in order to understand whether it was salient and well-understood by the participant. Finally, socioeconomic information about the participants, as well as their assessment of the prediction task (e.g. perceived difficulty, confidence in the predictions) were collected. In total, 561 participants completed the survey. As each respondent predicted the outcomes of the eight samples, the final dataset contains 4,488 predictions. Summary statistics of the sociodemographic characteristics of the participants are presented in Table 3.

Table 3: Descriptive Statistics

	(N = 561)
Age	
Mean ± Standard Deviation	38.26 ± 11.92
Median	37
Min	20
Max	84
Female	
If respondent is female	240/555 (43.2%)
Professional background	
Economics or Business Studies	184 (32.8%)
Agricultural Sciences/Farming	238 (42.4%)
Other	139 (24.8%)
Sample	
Polish farm advisors	109 (19.4%)
Croatian farm advisors	56 (10.0%)
French farm advisors	72 (12.8%)
Italian farm advisors	51 (9.1%)
Spanish experts	59 (10.5%)
Swedish students	69 (12.3%)
International researches	76 (13.6%)
Other	69 (12.3%)

Source: Own calculations

3 Results

3.1 Accuracy of predictions by samples and difficulty of predicting a sample

Table 4 displays the true means by farmer sample and the absolute deviations from these true means for each of the samples of forecasters. Farmers are, on average, slightly risk-seeking, with the Polish farmers being the most and the Spanish farmers being the least risk-averse. Means range from 4.74 in Spain to 6.30 in Poland, i.e., with a range of 1.64 in the mean, there is a rather large heterogeneity in how farmers respond to the multiple price lists.

We do not report average predictions of the forecasters here, but focus on accuracy, defined as the deviation from the sample average. The last column (*Pooled*) indicates how much the forecaster samples deviate, on average, from the true means across all eight samples. The second row (*Pooled Predictions*) displays how much, on average, all pooled predictions deviate from the true mean for each of the eight samples. In other words, low values in the last column indicate high predictive accuracy of a group of forecasters; low values in the third row indicate that a sample is easier to predict. Note that the sample of researchers provided the most accurate predictions on average, whereas the sample of French farmers was the easiest to predict. The range is smaller when considering the diversity of predicted samples (0.28 – 2.13 for France to 2.41 for Spain) than when considering the diversity of forecasters samples (0.90 – 1.80 for the researchers to 2.70 for the Polish farm advisors). Formal testing reveals that the average deviations of the predictions per participant do not come from the same distribution across all samples of predictors (Kruskal-Wallis test; $X^2 = 41.01$; $p < 0.001$), indicating that at least two samples of predictors in our data come from a different distribution (i.e., most probably our data point towards a higher accuracy of the international researchers in comparison to the Polish farm advisors). In contrast, regarding the to-be-predicted samples, we could not reject the null of accuracies coming from the same distribution (Kruskal-Wallis test; $X^2 = 8.76$; $p = .28$).

Table 4: True means of farmer samples and absolute deviations of expert predictions

		Predicted farmer samples								
Expert samples	N	Sweden	Germany	Poland	Netherlands	Spain	Italy	Croatia	France	Pooled
True Mean		5.70	5.71	6.30	5.80	4.74	4.96	6.05	5.28	5.61
Pooled Predictions	561	2.21	2.17	2.26	2.20	**2.41**	2.34	2.25	**2.13**	2.25
Farm Advisors Poland	109	2.59	2.89	2.92	2.60	2.74	2.82	2.66	2.35	**2.70**
International Researchers	76	1.63	1.56	1.88	1.53	2.18	2.08	1.73	1.81	**1.80**
Farm Advisors Croatia	56	2.12	2.09	2.36	2.69	2.52	2.48	2.63	2.22	2.39
Farm Advisors France	72	1.97	1.91	2.11	1.93	1.99	1.90	1.98	1.95	1.97
Farm Advisors Italy	51	2.82	2.46	2.54	2.81	2.60	2.78	2.45	2.05	2.56
Experts Spain	59	2.14	2.19	2.16	2.17	2.41	2.02	2.14	2.08	2.17
Swedish students	69	2.31	2.06	1.98	2.00	2.32	2.28	2.30	2.20	2.18
Other	69	2.09	1.92	1.87	1.93	2.42	2.21	2.08	2.29	2.10

Source: Own calculations (True means based on ROMMEL ET AL., 2022), Note: Bold values for highest and lowest absolute deviation across predicted samples and category of experts

3.2 Role of incentives

Table 5 summarizes the distribution of all predictions for the incentive treatments. Overall, the mean accuracies are similar across treatments. A Kruskal-Wallis test ($X^2 = 4.28$; $p = 0.37$) does not reject the null of equal distributions. Differences in the standard deviations are relatively large, and pairwise F-tests reveal at least some incompatibility of the data with the null (e.g., testing the standard deviation of *TOURHIGH* against *CONTROL* yields an F-ratio of 0.64 with p = .018 for the two-sided test). This indicates that incentives may not necessarily bias the results, but could help in enhancing the efficiency of predictions (see CAMERER and HOGARTH, 1999 for a discussion on the effect of incentives on the variation of experimental outcomes depending on effort).

Table 5: Accuracy by incentive treatments

Treatment	N	Minimum	Q1	Q2/Median	Q3	Maximum	Mean	SD
TOURLOW	107	0.52	1.43	1.95	2.77	6.43	2.13	1.12
TOURHIGH	112	0.47	1.36	1.95	2.79	5.46	2.15	1.06
ACCLOW	116	0.64	1.43	1.94	2.67	6.33	2.16	1.10
ACCHIGH	118	0.42	1.52	2.23	3.11	6.43	2.41	1.28
CONTROL	108	0.42	1.39	2.10	3.15	6.37	2.37	1.33

Source: own calculations

After the respondents made their predictions, we implemented a manipulation check on the incentives treatments by asking them correctly identify the incentive scheme they were assigned to. As seen in Table 6, between 50 and 70 of the respondents could correctly identify their exact treatment. In addition, 15% could at least identify the correct incentive mechanism (tournament or accuracy). Since the correct answers were not incentivized, these numbers can be considered large. One may remove respondents who did not provide correct answers for a robustness check in a later stage of the analysis.

Table 6: Control question for treatment mechanism

		Answer					
Assigned treatment		ACCHIGH	ACCLOW	CONTROL	TOURHIGH	TOURNLOW	I don’t know
ACCHIGH	N	67	19	8	9	2	13
	%	56.8	16.1	6.8	7.6	1.7	11.0
ACCLOW	N	17	68	9	11	1	10
	%	14.7	58.6	7.8	9.5	0.9	8.6
CONTROL	N	4	12	71	16	0	5
	%	3.7	11.1	65.7	14.8	0.0	4.6
TOURHIGH	N	11	20	8	60	0	13
	%	9.8	17.9	7.1	53.6	0.0	11.6
TOURNLOW	N	5	17	10	14	48	13
	%	4.7	15.9	9.3	13.1	44.9	12.1
All	N	104	136	106	110	51	54
	%	18.5	24.2	18.9	19.6	9.1	9.6

Source: own calculations

3.3 Heterogeneity in predictions

To investigate what drives the accuracy of predictions further, Table 7 presents regression results. The basic specification (1) only includes an intercept and binary controls for the predicted sample (estimates omitted for brevity). Model 2 additionally includes four dummy variables for the five treatments (reference category = *CONTROL*). Model 3 adds the covariates to adjust for the samples of forecasters (reference category: Polish experts), and Model 4 adds socioeconomic characteristics. Standard errors are clustered at the individual level to account for correlated predictions within participants.

The regressions indicate only small and statistically insignificant effects of the treatments, supporting the results of the previous subsection. Further, we did not find that the data on the predicted samples deviate from the null. In contrast, we find that accuracy differs by the respondent groups. Model 4 suggests that some expert groups (International Researchers, Farm

Advisors from France, and the miscellaneous group "other") made more precise forecasts than the largest participant subgroup ("Polish Advisors"). Gender, age and professional background showed no statistically significant effects on the prediction accuracy.

Table 7: Linear regressions with accuracy as dependent variable

	Model 1	Model 2	Model 3	Model 4
(Intercept)	2.253***	2.375***	2.834***	2.936***
	(0.070)	(0.136)	(0.183)	(0.355)
ACCHIGH		0.046	0.067	0.045
		(0.174)	(0.170)	(0.172)
ACCLOW		-0.209	-0.248	-0.242
		(0.164)	(0.159)	(0.159)
TOURHIGH		-0.218	-0.226	-0.271+
		(0.162)	(0.159)	(0.159)
TOURLOW		-0.236	-0.221	-0.232
		(0.168)	(0.166)	(0.166)
Expert: International			-0.891***	-0.862***
			(0.185)	(0.229)
Expert: Farm_Advisors_Croatia			-0.301	-0.271
			(0.200)	(0.205)
Expert: Farm_Advisors_France			-0.741***	-0.690***
			(0.181)	(0.187)
Expert: Farm_Advisors_Italy			-0.113	-0.061
			(0.198)	(0.206)
Expert: Experts_Spain			-0.554**	-0.409*
			(0.198)	(0.203)
Expert: Swedish_students			-0.542**	-0.571*
			(0.197)	(0.227)
Expert: Other			-0.611***	-0.588**
			(0.182)	(0.218)
Female				0.147
				(0.103)
Age				-0.005
				(0.005)
Background Agricultural Sciences/Farming				0.003
				(0.142)
Background Other				-0.125
				(0.152)
Num. Obs.	4,488	4,488	4,488	4,408
R^2	0.003	0.008	0.042	0.043
AIC	17,351.3	17,335.5	17,194.5	16,871.4
Prediction sample FE	Yes	Yes	Yes	Yes

Source: own calculations; Notes: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001, Clustered standard errors in parentheses

4 Discussion

It has to be noted, that the current analysis leaves a lot of room for further research. It would be interesting to investigate whether the effects on the prediction accuracy are asymmetric, i.e., whether some samples are generally predicted as too high, whereas others are predicted as too low. Here, we have only focused on the absolute *deviations*. Investigating the average deviations of a sample for forecasters to assess as to how far there is a "wisdom of crowd effect" would also be a relevant extension. This could also include more exploration of the distribution of predictions in relation to the distribution of farmers' choices.

The results presented above focused on the average prediction accuracy. The results do not indicate significant effects of the monetary incentives on the accuracy, but suggest that they could lead to less noisy forecasts. The results give a small indication that tournament schemes might perform better, but this needs further analysis. In this context, one should also investigate whether tournament based incentives exhibit a gender-heterogeneous treatment effect if compared to accuracy-based incentives on either accuracy or variation (NIEDERLE and WESTERLUND, 2007). This could be done by extending the analysis to distributional regression techniques, in order also model the variance parameter of the accuracy (cf. HOHBERG ET AL. 2020). Future research could consider focusing on this mechanism, in order to improve use available sample pools in a way that increases the statistical power.

Correlations between the prediction accuracy and variation, as well as the forecasters' certainty about their predictions and the perceived ease of the predictions offer further room for exploring the rich dataset. The certainty about predictions or the display of a good understanding by correctly responding to the manipulation check could be starting points for robustness checks. We have offered all respondents a short summary of the research results. It would be instructive to see whether or not exerts update their beliefs after taking part in a prediction (VIVALT and COVILLE, 2021).

It has to be taken into account that predictions were only obtained for one experimental risk preference elicitation task. The original study included three multiple price lists to elicit parameters for cumulative prospect theory. Here, we have only used one of the list to understand risk preferences. Although this has arguably allowed us to substantially simplify the task for respondents and to obtain a larger sample, it comes at the cost of understanding more about other aspects of risk preferences, such as the degree of loss aversion or probability weighting. While one may carry out such investigations in the future, one should probably also keep in mind that this could limit the sample of forecasters. Further investigating how elicitation formats and the complexity of instructions drive response rates and accuracy is, hence, important.

5 Conclusions

There is no in-depth understanding of stakeholder perceptions of farmers risk preferences. By analyzing the predictions of 561 agricultural experts of farmers' behavior in a multiple-price-list experiment for the determination of risk preferences for different groups of farmers, this study provided first insights into this previously neglected issue. By making differences between prior beliefs and experimental results visible, the study can enable participants to update their prior beliefs. We further find that different financial incentives have no statistically significant effect on the overall prediction accuracy.

References

BOCQUÉHO, G., F. JACQUET, and A. REYNAUD (2014): Expected utility or prospect theory maximisers? Assessing farmers' risk behaviour from field-experiment data. In: European Review of Agricultural Economics 41 (1): 135-172.

CAMERER, C. F. and R. M. HOGARTH (1999): The effects of financial incentives in experiments: A review and capital-labor-production framework. In: Journal of Risk and Uncertainty 19 (1): 7-42.

CHARNESS, G., U. GNEEZY and B. HALLADAY (2016): Experimental methods: Pay one or pay all. In: Journal of Economic Behavior & Organization 131: 141-150.

CLOT, S., G. GROLLEAU, and L. IBANEZ (2018): Shall we pay all? An experimental test of Random Incentivized Systems. In: Journal of Behavioral and Experimental Economics 73: 93-98.

DELLAVIGNA, S., D. POPE, and E. VIVALT (2019): Predict science to improve science. In: Science 366 (6464): 428-429.

DELLAVIGNA, S., and D. POPE (2018a): Predicting experimental results: who knows what? In: Journal of Political Economy 126 (6): 2410-2456.

DELLAVIGNA, S., and D. POPE (2018b): What motivates effort? Evidence and expert forecasts. In: The Review of Economic Studies 85 (2): 1029-1069.

DELLAVIGNA, S., N. OTIS, and E. VIVALT (2020): Forecasting the results of experiments: Piloting an elicitation strategy. In: American Economic Review, Papers and Proceedings 110 (May 2020): 75-79.

IYER, P., M. BOZZOLA, S. HIRSCH, M. MERANER, and R. FINGER (2020): Measuring farmer risk preferences in Europe: a systematic review. In: Journal of Agricultural Economics 71 (1): 3-26.

HOHBERG, M., P. PÜTZ, and T. KNEIB (2020): Treatment effects beyond the mean using distributional regression: Methods and guidance. In: PLOS ONE 15, e0226514.

MILKMAN, K L., ET AL. (2022): A 680,000-person megastudy of nudges to encourage vaccination in pharmacies. In: Proceedings of the National Academy of Sciences 119(6), e2115126119.

NIEDERLE, M., and L. VESTERLUND (2007): Do women shy away from competition? Do men compete too much? In: The Quarterly Journal of Economics 122 (3): 1067-1101.

ROMMEL, J., ET AL. (2022): Farmers' risk preferences in eleven European farming systems: A multi-country replication of Bocquého et al. (2014). Mimeo. https://tinyurl.com/riskprefs22

ROMMEL, J., ET AL. (2021): Environmental Cooperation at Landscape Scales: First Insights from Co-Designing Public Goods Games with Farmers in Four EU Member States. https://pub.epsilon.slu.se/23419/

TANAKA, T., C. F. CAMERER, and Q. NGUYEN (2010): Risk and time preferences: Linking experimental and household survey data from Vietnam. In: American Economic Review 100 (1): 557-571.

VIVALT, E., and A. COVILLE (2021): How Do Policymakers Update? Mimeo. http://evavivalt.com/wp-content/uploads/How-Do-Policymakers-Update.pdf