UNIVERSITY of York

HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 22/26

An assessment of physicians' risk attitudes using laboratory and field data

Massimo F. Castro; Calogero Guccio and Domenica Romeo

September 2022

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

An assessment of physicians' risk attitudes using laboratory and field data

Massimo Finocchiaro Castro[a,b], Calogero Guccio[c,b,*] & Domenica Romeo [b,c a]

[a] *Department of Law and Economics, Mediterranean University of Reggio Calabria, Italy*
b Health [b] Econometrics and Data Group, University of York, UK
[c] *Department of Economics and Business, University of Catania, Italy*

Abstract

By employing a large sample of both laboratory and field data, we investigate whether attitudes towards risk significantly vary between physicians, medical students and non-medical students. Also, we look for differences in risk propensity between laboratory and artefactual field experimental sessions and control for individuals' characteristics that may affect risk attitude. Results show significant variation in risk attitude, regardless of the estimation technique employed (linear regression, interval regression and maximum likelihood estimation), suggesting constant relative risk aversion (CRRA) as a supported representation of risk preferences. Finally, data consistently show that physicians are more risk-seeking in the monetary domain than other subject groups.

Keywords: Risk aversion; Field experiments; Laboratory experiment; Physicians' behaviour.

JEL Classification: I1; C81; C93; D81

* **Corresponding author**: Calogero Guccio, Department of Economics and Business, University of Catania, Italy, guccio@unict.it.

1. Introduction

Attitude towards risk significantly matters in individual decision making. For instance, tourists' destination selection (George, 2010), voting (Nadeau et al., 1999), entrepreneurship (Caliendo et al., 2009), production decisions (Berg, 2003) are all circumstances varying on subjects' risk propensity. Also in the medical sector risk attitude plays a fundamental role, since physicians usually take decisions driven by their risk attitude, being exposed to uncertainty and time pressure (McKibbon, 2005; Méndez et al., 2021). Thus, physicians' different propensity to risk is generally responsible for practice variation (Miraldo et al., 2019). For instance, risk aversion can affect physicians in two different ways: first physicians may be reluctant to prescribe a high-risk treatment in the fear of harming a patient; second, they could avoid the high-risk treatment afraid of being sued for malpractice (Chandra et al., 2012). While evidence is often mixed, Franks et al. (2000) find that referral likelihood is positively associated with doctors' risk avoiding behaviour. Similarly, physicians less prone to risk are more likely to admit patients than their risk averse peers (Pearson et al., 1995). Also, risk aversion affects obstetricians' willingness to perform caesareans (O'leary et al., 2007), as well as physicians' attitude towards vaccination against seasonal and pandemic influenza regarding themselves and patients (Massin et al., 2015). Finally, limited consultation of resources (Allison et al., 1998) and early adoption of new drugs are driven by higher risks tolerance (Zhang et al., 2019).

Therefore, measuring subjects' risk preferences becomes important to predict behaviours. For this reason, eliciting risk preferences in experimental sessions whose core tasks involve decisions under uncertainty is already common practice in several domains but the medical one (see for instance Shiv et al., 2005; Wakolbinger and Haigner, 2009; Maggian and Montinari, 2017; Brosig-Koch et al., 2019; Fallucchi et al., 2020), although Galizzi et al. (2016a) and Arrieta et al. (2017) represent two of the limited exceptions. To the best of our knowledge, we are the first one assessing physicians' risk attitude pooling data from laboratory (Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021) and artefactual field experiments (Finocchiaro and Romeo, 2021a; Finocchiaro and Romeo, 2021b).

In each of the four experiments, participants have been asked to complete the well-known Holt and Laury (2002) questionnaire (HL, hereafter) with hypothetical payoffs, prior to start the main experimental tasks. According to individual's choices in the HL questionnaire, it is possible to classify her as risk averse, neutral, or loving. Also, we control for the effects of several individual's characteristics and, as a robustness check, we perform the analysis excluding what can be classified as inconsistent choices form the sample (Chuang and Schecter, 2015; Filippin and Crosetto, 2016).

Non-parametric analysis shows that participants in the lab are more risk averse than subjects joining the artefactual field experiments. Also, the results show significant differences in risk attitude among types of participants: physicians tend to be less risk averse than any other subject groups. Moreover, medical students are more inclined to risk-averse decisions than doctors. By excluding inconsistent subjects, differences still hold.

Given the data at hand and applying different empirical techniques such as linear regression, interval regression, and maximum likelihood estimation, we suggest the constant relative risk aversion (CRRA) as the supported representation of preferences. Finally, we also investigate the patterns of inconsistent choices, using a structural model estimated with maximum likelihood. We show that different types of subjects follow different sequences of inconsistent choices.

The rest of the paper is organized as follows. Section 2 reviews the related literature. In Section 3, we briefly describe data and non-parametric analysis. Section 4 presents the parametric analysis. Finally, Section 5 concludes the study.

2. Literature Review

Although different measurement procedures to infer subjects' attitude towards risk exist (Harrison and Rutström, 2008; Charness et al., 2013), we focus our review on three of the most widely adopted methodologies. The first one is the Holt and Laury (2002) approach that suggests a multiple price list paired lottery method. HL questionnaire consists of 10 hypothetical choices between a safer lottery called A and a riskier lottery called B. Payoff and probabilities are distributed such that the number of times a subject chose lottery A could be used to estimate his attitude towards risk. According to HL, as the probability associated to the high payoff outcome increases, subjects should shift from option A to option B (see Table 1).

<<Table 1 around here>>

The second approach, introduced by Binswanger (1981) and then proposed in an alternative version by Eckel and Grossman (2002), is the ordered lottery selection task. Subjects are asked to select one among a set of five gambles, each containing two possible outcomes, with a linearly increasing expected value but a higher standard deviation. While the probability of each outcome is set at 50%, the outcome varies from one lottery to another. Finally, in the third approach, the risk elicitation task is framed as an investment decision (Gneezy and Potters, 1997). In a series of twelve identical but independent rounds of lotteries, subjects endowed with four euro must choose how much of the given amount to bet in a lottery with a 2/3 probability of losing that amount and a 1/3 probability of winning two and a half times the amount invested. Although the three methods somehow differ, all of them allow to estimate the individuals' coefficient of risk aversion (Crosetto and Filippin, 2013), once specific assumptions on the utility function are made[1].

Besides the approach to be taken, several factors appear to affect individual's attitude toward risk. For instance, the propensity towards risk may change according to whether participants face hypothetical or real payoff, although evidence is mixed. Whereas some authors show that risk taking is modulated differently when real consequences are at stake (Ettenson and Coughlin, 1982; Xu et al., 2016), others find no significant differences between hypothetical and real rewards (Beattie and Loomes, 1997; Camerer and Hogarth, 1999)[2]. Some authors, instead, investigate how risk attitude varies according to the choice domain (Weber et al., 2002; Wang et al., 2016). Finally, Galizzi et al. (2018) review the literature comparing risk preferences across domains and conclude that attitudes towards risk are domain specific.

[1] See, for instance, Andersen et al. (2008).
[2] See Kühberger et al. (2002) for a comprehensive literature review.

Although, in the literature, the focus is often on the comparison of risk preferences across different domains (Schoemaker, 1990; Prosser and Wittenberg, 2007; Riddel, 2012), a limited number of studies assess risk preferences in the health sector. For example, Galizzi et al. (2016b) run a field experiment in which 300 patients of a Greek hospital are asked to complete the HL questionnaire adapted to the health and the financial context, with hypothetical payoffs. Results show that subjects are more risk averse in the health domain than in financial one. The same experimental design has been applied by Galizzi, et al. (2016a) to compare risk and time preferences between doctors and patients. Although the risk preferences of two groups are almost the same in the health domain, differences become significant in the financial domain. Specifically, whereas doctors are found to be risk averse, patients show risk neutrality. Similarly, Zhu et al. (2019) use the HL task adapted to the health domain to assess patients' risk and time attitude. Differently from previous studies, Goldzahl (2017) adopts a procedure close to Eckel and Grossman (2002) to elicit 178 women's risk preferences and finds that risk aversion is responsible for 30% of the variance in breast cancer screening regularity. Relationship between health conditions and behaviour towards risk is examined by Martín-Fernández et al. (2018), using both a self-assessment scale and a lottery game. Results show that subjects suffering from poor mental conditions are more risk averse, whereas the same does not apply to participants with poor general health status. Finally, Arrieta et al. (2017) design a lab experiment where both medical and nonmedical students, playing the role of a physician, provide treatments to patients. The authors show that, although participants generally show risk aversion, risk attitudes are health context dependent. Additionally, medical students tend to be more risk averse than their peers, and such tendency is mitigated when real rewards for third parties are introduced.

Although the interest in experimental health economics is growing fast, there are still few papers investigating the risk attitudes of physicians and of subjects playing the role of physicians. Hence, we are the first to analyze such a huge experimental data set and exploit its variation in terms of types of participants and their characteristics to offer a novel and solid empirical contribution to the analysis of risk preferences.

3. Data and Non-parametric Tests

The dataset has been built on four experimental papers[3], all adopting HL task with hypothetical payoffs to assess participant attitude towards risk. Before starting the main task of each experiment, participants have been asked to fill the HL questionnaire in. Each participant was aware that payoffs were hypothetical and that her profit would derive from the choices made in the following stages of each experiment. Given that those tasks have been communicated to participants only after they completed the HL questionnaire, they did not affect the assessment of subjects' risk attitude in any way. Hence, the discussion of each experimental design is out of the scope of the present paper. The differences among the four sources of data in terms of type of experiment and participants are depicted in Table 2. The total subjects pool counts 433 participants, distinguished as follows: 232 students, 42 medical students and 159 physicians (of which 12 in the laboratory and 147 in the field). Subjects are almost equally divided by gender: around 57% of the sample are men, the remaining are women. Regarding the type of experiment, we use data from two framed field experiments and two laboratory experiments[4].

<<Table 2 around here>>

Table 3 reports the descriptive statistics of the variables used in both non-parametric and parametric analyses. The last 5 rows of Table 3 report the widely used classification of participants according to their choices in the HL questionnaire (Filippin and Crosetto, 2016).

<<Table 3 around here>>

[3] The papers we refer to are Finocchiaro Castro et al. (2019), Finocchiaro Castro et al. (2021), Finocchiaro Castro and Romeo (2021a) and, Finocchiaro Castro and Romeo (2021b).

[4] The two framed field experiments, done with medical doctors only, have been run at the main Hospital of Reggio Calabria. Differently, the two laboratory experiments have been conducted at the MEBEL lab of the Mediterranean University of Reggio Calabria with students, medical students, and medical doctors.

3.1 Number of Safe Choice

Like in Filippin and Crosetto (2016), the variable at the core of our empirical analysis is the number of safe choices, *N_safe*, (choosing option A in the HL questionnaire) made by participants. Thus, we report in Table 4 the breakdown of *N_safe* by type of experiment, subject pool, age, and gender.

<<Table 4 around here>>

Non-parametric tests, whose results are reported in Table 5, show that most of the pairwise comparisons turn out to be significant. First, participants in the lab are more risk averse than the ones joining the field experiments (*p-value*<0.01). Given that only physicians joined the field experiments, the reported difference can be ascribed to the higher risk-seeking attitude among physicians with respect to other subjects, resulting in a lower average number of safe choices (3.817 vs 4.734).

<<Table 5 around here>>

As far as subjects' type is concerned, evidence is mixed. Medical students are more risk lovers than non-medical students (*p-value*<0.05), confirming Arrieta et al. (2017)[5], but still more conservative than physicians (*p-value*<0.01). The latter can be explained by the lower level of experience in dealing with risky situations of medical students with respect to physicians. As reported by Lawton et al. (2019), less experienced physicians, which in our context may correspond to medical students, are more inclined to risk-averse decisions. Finally, gender differences in risk attitudes can be observed among students (*p-value*<0.01), where women turn out to be less risk seeking than men. Hence, we provide support to the results of Dave et al (2010), and Eckel and Grossman (2008). Moreover, when we compare

[5] Notice that Arrieta et al. (2017) find controversial evidence depending on the subjects' choice domain. They show that medical students are more risk averse than nonmedical students in the health domain, but the opposite is true for the monetary domain (HL).

physicians' choices according to the gender, the differences appear negligible (3.829 vs 3.805) and not significant.

3.2 Inconsistent choices

Everyday life shows that people usually commit errors while making decisions. In particular, experimental economists who observe individual behavior both in the lab and in the field very often record inconsistent choices, especially when participants deal with risk or uncertainty (Chuang and Schechter, 2015). While some authors believe that inconsistent patterns represent only an exception (Holt and Laury, 2002; Abdellaoui et al., 2011), others show that they can represent a significant portion of the whole sample (Jacobson and Petrie, 2009; Charness and Viceisza, 2011). Unfortunately, how to deal with inconsistent choices from HL results is still under question (Hirschauer et al. 2014; Engel et al., 2019). In fact, in HL task we should observe that as the probability associated to the high payoff outcome increases, subjects should shift from option A to option B at a certain point. From a standard microeconomic perspective, a utility maximizing individual should shift only once from A to B, without going back. However, different choice patterns can be usually observed in experimental sessions, such as multiple switch sequences. For instance, subjects always selecting A, even when the more valuable outcome B becomes a certainty, and players always choosing option B, which violates the axioms of the expected utility theory, can be classified as irrational. Hence, the switching point is commonly used to classify subjects according to their risk aversion coefficients.

Table 6 provides the summary statistics of inconsistent choices divided according to the type of experiment, of participant, age, and gender. To investigate the patterns of inconsistent choices, we have built the variable *inconsistent*, which is a dummy taking the value 1 if the pattern of choices can be classified as inconsistent, and 0 otherwise. Interestingly, the average level of inconsistent choices observed in artefactual field experiments is higher than the one reached in the lab experiments (0.347 vs. 0.297), supporting the findings of Jacobson and Petrie (2009). Table 7 lists the Wilcoxon Mann-Whitney tests. Differences are significant for most of the pairwise comparisons. As shown by Filippin and Crosetto (2016), we also report significant gender differences in both students and physicians' samples, where men appear to be more consistent than women.

<<Table 6 around here>>

<<Table 7 around here>>

The detailed pattern of inconsistent choices is reported in Table 8. Inconsistent choices in the HL questionnaire commonly take three different forms: switching from lottery B to A, always choosing lottery A, always choosing lottery B.

Looking at data, switching from option B to option A is the most frequent inconsistent behaviour and it has been significantly more often chosen by students than physicians (*p-value*<0.001). The opposite can be observed in the case of dominated choices (Always Option A or always Option B), where physicians inconsistency levels are significantly higher than the ones of students (in both cases *p-value*<0.001)[6]. Although the empirical analysis of inconsistent choice represents a deviation from the main aim of our paper, it represents a robustness check of the results described in the following Section. Hence, we report the econometric investigation on inconsistent choices in the Appendix B.

<<Table 8 around here>>

4. Structural estimates using interval regression

To assess the robustness of previous results by means of a parametric empirical analysis, we assume a theoretical representation of individual risk preferences. For, in this section, we assume that risk preferences can be represented by a utility function with a constant relative risk aversion (CRRA) coefficient r on monetary outcomes M, which would make a subject indifferent between lottery A and B in the HL task,

[6] For a detailed discussion about the role of inconsistent choices in our sample, see Appendix A.

$$U(M,r) = \frac{M^{1-r}}{1-r} \qquad (1)$$

Hence, a value of 0 denotes risk neutrality, negative values indicate risk-loving, and positive values indicate risk aversion. Thus, we can infer on r based on the 10 decisions made by participants into the HL lottery. To assess the effects of experimental designs, and at the same time, controlling for individual characteristics, we adopt an interval regression model (Coller and Williams, 1999). The dependent variable is the CRRA interval, where the risk parameter r lies, that each participant implicitly chose when she switches from the safe option (treatment A) to the risky option (treatment B) (Harrison and Rutström, 2008; Arrieta et al., 2017). More specifically, the lower r_{lb} and upper bounds r_{ub} of the interval of r are associated to the switching point from the safe option (lottery A) to the risky option (lottery B). For instance, a subject who switches from A to B between the fifth and the sixth row would result in an r located in an interval between 0.15 and 0.41 (Harrison and Rutström, 2008). This implies that the higher r, the higher subject's degree of risk aversion.

Using panel interval regression model, we control for all the individual characteristics collected in the four experiments. In addition, we can account for the different features of the experimental designs. Thus, we can assume that the coefficient of relative risk aversion r_i of a subject I follows a linear function of the individual characteristics X_i, and a stochastic term ε_i,

$$r_i = X_i\beta + \varepsilon_i \qquad (2)$$

In our context, individuals' characteristics refer to the subjects' type (either physician, student, or medical student), age and gender. Moreover, the variation of age across the sample can be seen as a proxy for the differences in wealth between physicians and students. The error term ε_i is assumed to follow a normal distribution with censoring, which considers the choices at the extremes of the interval of r when

r_{lb} or r_{ub} goes to infinity. Hence, we estimate Model (2) on the entire sample and on the subsample of consistent choices only[7].

Tables 9 and 10 report the results of model estimation (2) for the full sample and the sub-sample of consistent choices, respectively. The results largely confirm the previous findings. Regardless of the inconsistent choices and of any individual characteristic, physicians are constantly more willing to take risks than any other type of participants.

<<Table 9 around here>>

<<Table 10 around here>>

5. Concluding remarks

The relevant literature has clearly shown that attitude towards risk significantly matters in individual decision making across different domains. For instance, in the health sector, the role of physician's risk propensity plays, indeed, a fundamental role, since they usually take decisions driven by their risk attitude, being exposed to uncertainty and time pressure (McKibbon, 2005; Méndez et al., 2021).

Surprisingly, few papers have tackled this issue analysing experimental data. Hence, our work attempts at filling this gap in the literature providing a robust empirical analysis of risk attitude employing a massive sample of physicians, medical, and non-medical students. For, we have pooled data from four experimental papers where all participants have been asked to complete the Holt and Laury (2005) questionnaire before starting the main task of each experimental design.

[7] Specifically, we estimate the full sample first and then we estimate the subsample of only consistent choices, by excluding also multiple switches. Contrarily, Arrieta et al. (2017) and Andersen et al. (2006) consider also multiple switches inferring the lower bound of the interval from the first switch and the upper bound from the last switch. We apply the same approach of Arrieta et al. (2017) obtaining estimates which intermediate to those reported in Tables 9 and 10. They largely confirm our results. The estimates are available from the authors on request. As a robustness test, we also performed estimates on model (2) using the approach proposed by Harrison and Ruström (2008) and applying the script provided by Harrison (2008). Results are provided in Appendix B.

Results suggest that physicians are more risk-loving than both types of students and that on average physicians account for most of inconsistent choices observed in the experimental sessions. Additionally, evidence on medical students appears mixed. Their behavior does not differ from the one of non-medical students when we consider the whole set of choices. Differently, non-medical students tend to behave like physicians when we exclude the inconsistent choices from the analysis. Also, we confirm the findings of Brosig-Koch et al. (2016) showing that the level of answer to incentives changes according to subjects' pool: physicians' behavior is less affected than students. Hence, experimenters need to be careful when selecting the subject pools to test health economics predictions, not taking for granted that medical students can proxy physicians in experimental settings. In this case, policy implications based merely on estimates should be cautiously assessed. As robustness check, we have also estimated the risk aversion parameters, using two different approaches: the interval regression and the Maximum Likelihood (see Appendix B). Estimations confirm our main result, showing that physicians are more risk seekers than other subjects.

Although our results are robust also to robustness checks, some limitations should be acknowledged. First, our sample is not only composed by physicians, so, although we control for subject type in the regressions, caution is required when drawing insights also from students acting as physicians. Second, all the four designs report the HL questionnaire with hypothetical payoff. Thus, we cannot exclude that some participants have not been fully motivated in answering the HL lottery list. Finally, and maybe more relevant, differently from Galizzi et al. (2016a), we only focus the HL questionnaire on the monetary domain, instead of introducing the health domain. Although the evidence on the role of the health domain is somewhat mixed, we cannot exclude that individual answers may have been different if the experiments would have been run also in the health domain. Therefore, further research activities need to be devoted to the role of health domain compared to the monetary domain in health economics experiments.

References

Abdellaoui, M., Driouchi, A. and L'Haridon, O. (2011) Risk aversion elicitation: reconciling tractability and bias mini- mization, *Theory and Decision*, 71, 63–80.

Allison, J. J., Kiefe, C. I., Cook, E. F., Gerrity, M. S., Orav, E. J., & Centor, R. (1998). The association of physician attitudes about uncertainty and risk taking with resource use in a Medicare HMO. *Medical Decision Making*, *18*(3), 320-329.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, *9*(4), 383-405.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3), 583-618.

Arrieta, A., García-Prado, A., González, P., & Pinto-Prades, J. L. (2017). Risk attitudes in medical decisions for others: an experimental approach. *Health Economics*, 26, 97-113.

Beattie, J., & Loomes, G. (1997). The impact of incentives upon risky choice experiments. *Journal of Risk and Uncertainty*, 14(2), 155-168.

Berg, E. (2003, August). Modeling the impacts of uncertainty and attitudes towards risk on production decisions in arable farming. In *Poster Paper submitted for the 2^{5t}h International Conference of IAAE, Durban, South Africa* (pp. 16-22).

Binswanger, H. P. (1981). Attitudes toward risk: Theoretical implications of an experiment in rural India. *The Economic Journal*, 91(364), 867-890.

Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., Kokot, J., & Wiesen, D. (2019). Physician performance pay: Experimental evidence. *SSRN Electron. J, 3467583*.

Caliendo, M., Fossen, F. M., & Kritikos, A. S. (2009). Risk attitudes of nascent entrepreneurs–new evidence from an experimentally validated survey. *Small business economics*, *32*(2), 153-167.

Camerer, C.F., Hogarth, R.M. (1999) The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty*, 19, 7–42.

Chandra, A., Cutler, D., & Song, Z. (2011). Who ordered that? The economics of treatment choices in medical care. *Handbook of health economics*, *2*, 397-432.

Charness, G. and Viceisza, A. (2011) Comprehension and risk elicitation in the field: evidence from rural Senegal. IFPRI Discussion Papers 1135, International Food Policy Research Institute (IFPRI), Washington DC

Charness, G., Gneezy, U., & Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization*, *87*, 43-51.

Chuang, Y., & Schechter, L. (2015). Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. *Journal of development economics*, *117*, 151-170.

Crosetto, P., & Filippin, A. (2013). The “bomb” risk elicitation task. *Journal of Risk and Uncertainty*, 47(1), 31-65.

Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better?. *Journal of Risk and Uncertainty*, *41*(3), 219-243.

Eckel, C. C., & Grossman, P. J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and human behavior*, 23(4), 281-295.

Eckel, C. C., & Grossman, P. J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization*, *68*(1), 1-17.

Engel, C., & Kirchkamp, O. (2019). How to deal with inconsistent choices on multiple price lists. *Journal of Economic Behavior & Organization*, *160*, 138-157.

Ettenson, R. T., & Coughlin, R. C. (1982). Effects of type of payoff and instructions on individual risk-taking behavior. *Psychological Reports*, 51(3), 855-860.

Fallucchi, F., Nosenzo, D., & Reuben, E. (2020). Measuring preferences for competition with experimentally-validated survey questions. *Journal of Economic Behavior & Organization*, *178*, 402-423.

Filippin, A., & Crosetto, P. (2016). A reconsideration of gender differences in risk attitudes. *Management Science*, *62*(11), 3138-3160.

Finocchiaro Castro, M., & Romeo, D. (2021a). *An experimental analysis of patient dumping under different payment systems*. Società italiana di Economia pubblica, Working papers 106.

Finocchiaro Castro, M., & Romeo, D. (2021b). Do National Health Guidelines increase coordination level among physicians? An experimental investigation (No. 21/09). HEDG, c/o Department of Economics, University of York.

Finocchiaro Castro, M., Ferrara, P. L., Guccio, C., & Lisi, D. (2021). Optimal mixed payment system and medical liability. A laboratory study. MPRA Paper No. 110276, https://mpra.ub.uni-muenchen.de/110276/.

Finocchiaro Castro, M., Ferrara, P. L., Guccio, C., & Lisi, D. (2019). Medical malpractice liability and physicians' behavior: Experimental evidence. *Journal of Economic Behavior & Organization*, 166, 646-666.

Franks, P., Williams, G. C., Zwanziger, J., Mooney, C., & Sorbero, M. (2000). Why do physicians vary so widely in their referral rates?. *Journal of general internal medicine*, *15*(3), 163-168.

Galizzi, M. M., Harrison, G. W., & Miraldo, M. (2018). Experimental methods and behavioral insights in health economics: Estimating risk and time preferences in health. In Health Econometrics. Emerald Publishing Limited.

Galizzi, M. M., Miraldo, M., Stavropoulou, C., & Van Der Pol, M. (2016a). Doctor–patient differences in risk and time preferences: A field experiment. *Journal of Health Economics*, 50, 171-182.

Galizzi, M. M., Miraldo, M., & Stavropoulou, C. (2016b). In sickness but not in wealth: field evidence on patients' risk preferences in financial and health domains. *Medical Decision Making*, 36(4), 503-517.

George, R. (2010). Visitor perceptions of crime-safety and attitudes towards risk: The case of Table Mountain National Park, Cape Town. *Tourism management*, *31*(6), 806-815.

Goldzahl, L. (2017). Contributions of risk preference, time orientation and perceptions to breast cancer screening regularity. *Social Science & Medicine*, 185, 147-157.

Harrison, G. W., & Rutström, E. E. (2008). Risk aversion in the laboratory. In *Risk aversion in experiments*. Emerald Group Publishing Limited.

Harrison, G. W., Lau, M. I., Rutström, E. E., & Sullivan, M. B. (2005). Eliciting risk and time preferences using field experiments: Some methodological issues. In *Field experiments in economics*. Emerald Group Publishing Limited.

Hirschauer, N., Musshoff, O., Maart-Noelck, S. C., & Gruener, S. (2014). Eliciting risk attitudes–how to avoid mean and variance bias in Holt-and-Laury lotteries. *Applied Economics Letters*, *21*(1), 35-38.

Jacobson, S., & Petrie, R. (2009). Learning from mistakes: What do inconsistent choices over risk tell us?. *Journal of risk and uncertainty*, *38*(2), 143-158.

Kühberger, A., Schulte-Mecklenbeck, M., & Perner, J. (2002). Framing decisions: Hypothetical and real. *Organizational Behavior and Human Decision Processes*, 89(2), 1162-1175.

Lawton, R., Robinson, O., Harrison, R., Mason, S., Conner, M., & Wilson, B. (2019). Are more experienced clinicians better able to tolerate uncertainty and manage risks? A vignette study of doctors in three NHS emergency departments in England. *BMJ quality & safety*, *28*(5), 382-388.

Maggian, V., & Montinari, N. (2017). The spillover effects of gender quotas on dishonesty. *Economics Letters, 159*, 33-36.

Martín-Fernández, J., Medina-Palomino, H. J., Ariza-Cardiel, G., Polentinos-Castro, E., & Rutkowski, A. (2018). Health condition and risk attitude in the Dutch population: an exploratory approach. *Health, Risk & Society*, 20(3-4), 126-146.

Massin, S., Ventelou, B., Nebout, A., Verger, P., & Pulcini, C. (2015). Cross-sectional survey: risk-averse French general practitioners are more favorable toward influenza vaccination. *Vaccine*, *33*(5), 610-614.

McKibbon, K. A. (2005). *The effect of risk attitude and uncertainty comfort on primary care physicians' use of electronic information resources*. University of Pittsburgh.

Méndez, S. J., Scott, A., & Zhang, Y. (2021). Gender differences in physician decisions to adopt new prescription drugs. *Social Science & Medicine*, *277*, 113886.

Miraldo, M., Hauck, K., Vernet, A., & Wheelock, A. (2019). Variations in the adoption of healthcare innovation: a literature review. *Oxford Research Encyclopedia of Economics and Finance*.

Nadeau, R., Martin, P., & Blais, A. (1999). Attitude towards risk-taking and individual choice in the Quebec referendum on sovereignty. *British Journal of Political Science*, *29*(3), 523-539.

O'leary, C. M., De Klerk, N., Keogh, J., Pennell, C., De Groot, J., York, L., ... & Stanley, F. J. (2007). Trends in mode of delivery during 1984–2003: can they be explained by pregnancy and delivery complications?. *BJOG: An International Journal of Obstetrics & Gynaecology*, *114*(7), 855-864.

Pearson, S. D., Goldman, L., Orav, E. J., Guadagnoli, E., Garcia, T. B., Johnson, P. A., & Lee, T. H. (1995). Triage decisions for emergency department patients with chest pain. *Journal of General Internal Medicine*, *10*(10), 557-564.

Prosser, L.A., & Wittenberg, E. (2007). Do risk attitude differ across domains and respondent types? *Medical Decision Making*, 27(3), 281-287.

Riddel, M. (2012). Comparing risk preferences over financial and environmental lotteries. *Journal of Risk and Uncertainty*, *45*(2), 135-157.

Schoemaker, P. J. (1990). Are risk-attitudes related across domains and response modes?. *Management science*, 36(12), 1451-1463.

Shiv, B., Loewenstein, G., Bechara, A., Damasio, H., & Damasio, A. R. (2005). Investment behavior and the negative side of emotion. *Psychological science*, *16*(6), 435-439.

Wakolbinger, F., & Haigner, S. D. (2009). Peer advice in a tax-evasion experiment. *Economics Bulletin*, *29*(3), 1653-1669.

Wang, X. T., Zheng, R., Xuan, Y. H., Chen, J., & Li, S. (2016). Not all risks are created equal: A twin study and meta-analyses of risk taking across seven domains. *Journal of experimental psychology: general*, 145(11), 1548.

Weber, E. U., Blais, A. R., & Betz, N. E. (2002). A domain‐specific risk‐attitude scale: Measuring risk perceptions and risk behaviors. *Journal of behavioral decision making*, 15(4), 263-290.

Xu, S., Pan, Y., Wang, Y., Spaeth, A. M., Qu, Z., & Rao, H. (2016). Real and hypothetical monetary rewards modulate risk taking in the brain. *Scientific Reports*, 6(1), 1-7.

Zhang, Y., Méndez, S. J., & Scott, A. (2019). Factors affecting general practitioners' decisions to adopt new prescription drugs–cohort analyses using Australian longitudinal physician survey data. *BMC health services research*, *19*(1), 1-12.

Zhu, J., Shi, Y., Li, J., & Zhang, Z. (2020). Role of risk attitude and time preference in preventive aspirin use adherence. *Journal of evaluation in clinical practice*, 26(3), 819-825.

TABLES AND FIGURES

Table 1. Adaptated of Holt and Laury (2002) questionnaire

Lottery A	Lottery B	Your choice
2€ with probability 1/10 1,60€ with probability 9/10	3,85€ with probability 1/10 0,10€ with probability 9/10	
2€ with probability 2/10 1,60€with probability 8/10	3,85€with probability 2/10 0,10€ with probability 8/10	
2€ with probability 3/10 1,60€ with probability 7/10	3,85€ with probability 3/10 0,10€ with probability 7/10	
2€ with probability 4/10 1,60€ with probability 6/10	3,85€ with probability 4/10 0,10€ with probability 6/10	
2€ with probability 5/10 1,60€ with probability 5/10	3,85€ with probability 5/10 0,10€ with probability 5/10	
2€ with probability 6/10 1,60€ with probability 4/10	3,85€ with probability 6/10 0,10€ with probability 4/10	
2€ with probability 7/10 1,60€ with probability 3/10	3,85€ with probability 7/10 0,10€ with probability 3/10	
2€ with probability 8/10 1,60€ with probability 2/10	3,85€ with probability 8/10 0,10€ with probability 2/10	
2€ with probability 9/10 1,60€ with probability 1/10	3,85€ with probability 9/10 0,10€ with probability 1/10	
2€ with probability 10/10 1,60€ with probability 0/10	3,85€ with probability 10/10 0,10€ with probability 0/10	

Source: our adaptation from Holt and Laury (2002)

Table 2. Subjects pool

Type of data	Subjects	Male	Female	Average age
		Laboratory data		
Students	232	142	90	24.65
Medical students	42	22	20	24.71
Doctors in the lab	12	2	10	33.25
Total in the lab	**286**	**166**	**120**	**25.02**
		Field data		
Doctors in the field	147	80	67	48.64
Total	**433**	**246**	**187**	**33.00**

Source: our elaboration using data from Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021; Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b.

Table 3. Descriptive statistics of the sample

Variables	Description	N	mean	sd	min	max
Age	Age in years	433	33.00	13.78	16	69
Female	Dummy for gender	433	0.432	0.496	0	1
Student	Dummy for student	433	0.536	0.499	0	1
Medical student	Dummy for medical student	433	0.097	0.296	0	1
Physician	Dummy for physician	433	0.367	0.482	0	1
Experiment	Categorial variable for the experiment	433	2.178	0.897	1	4
Field	Dummy for field experiment	433	0.339	0.474	0	1
Safe	Dummy for choosing option A in the HL lottery	433	0.443	0.497	0	1
N_safe	Number of safe choices	433	4.397	1.892	0	10
Only_A	=1 if subject always selects A	433	0.021	0.143	0	1
Only_B	=1 if subject always selects B	433	0.065	0.246	0	1
Multi_Switch	=1 if subject selects A after B	433	0.229	0.420	0	1
Inconsistent	Dummy for inconsistent choices	433	0.314	0.465	0	1

Source: our elaboration using data from Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021; Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b.

Table 4. Number of safe choices

Subject	Subjects	Mean	St. dev.
Type of data			
Laboratory subjects	286	**4.689**	**1.469**
Field subjects	147	**3.830**	**2.421**
Subject type			
Students	232	4.763	1.442
Medical students	42	4.571	1.516
Doctors	159	3.818	2.361
Age			
Between 18 and 24	166	4.548	1.220
Between 25 and 34	128	4.679	1.705
Between 35 and 49	59	3.814	2.105
50 or more	80	4.062	2.801
Gender			
Male	246	4.382	1.831
Female	187	4.417	1.970
Male students	142	4.697	1.478
Female students	90	4.866	1.376
Male medical students	22	4.409	1.530
Female medical students	20	4.750	1.482
Male doctors	82	3.829	2.274
Female doctors	77	3.805	2.451
Total	**433**	**4.397**	**1.892**

Source: our elaboration using data from Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021; Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b.

Table 5. Mann-Whitney tests

Test	N_1	N_2	N_safe_1	N_safe_2	P-value
Laboratory vs. Field data	286	147	4.689	3.830	0.0001
All subjects vs. Doctors	274	159	4.734	3.818	0.0001
Medical Students vs. Doctors	42	159	4.571	3.818	0.0001
Students vs. Medical students	232	42	4.762	4.571	0.0289
Male vs. Female students	142	90	4.697	4.866	0.0061
Male vs Female medical students	22	20	4.409	4.750	0.7704
Male vs Female doctors	82	77	3.829	3.805	0.6710

Source: our elaboration using data from Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021; Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b.

Table 6. Inconsistent choices

Subject	Observation	Mean	St. dev.
	Type of data		
Laboratory	286	0.297	0.457
Field data	147	0.347	0.476
	Subject type		
Students	232	0.293	0.455
Medical students	42	0.333	0.472
Doctors	159	0.340	0.474
	Age		
Between 18 and 24	166	0.265	0.442
Between 25 and 34	128	0.313	0.464
Between 35 and 49	59	0.237	0.426
50 or more	80	0.476	0.499
	Gender		
Male	246	0.272	0.445
Female	187	0.369	0.483
Male students	142	0.239	0.427
Female students	90	0.378	0.485
Male medical students	22	0.364	0.482
Female medical students	20	0.300	0.459
Male doctors	82	0.305	0.461
Female doctors	77	0.377	0.485
Total	**433**	**0.3140**	**0.464**

Source: our elaboration using data from Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021; Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b.

Table 7. Mann-Whitney tests - inconsistent choices

Test	$Inconsistent_1$	$Inconsistent_2$	P-value
Laboratory vs. Field data	0.297	0.347	0.0008
All subjects vs. Doctors	0.458	0.340	0.0058
Medical students vs. Doctors	0.333	0.340	0.8086
Students vs. Medical students	0.293	0.333	0.0976
Male vs. Female students	0.239	0.378	0.0000
Male vs Female medical students	0.364	0.300	0.1676
Male vs Female doctors	0.305	0.377	0.0025

Source: our elaboration using data from Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021; Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b.

Table 8. Summary statistics of inconsistent choices by subject pool

Type	Inconsistent choices		*% of inconsistent choices*		
	Number	**Out of**	***Physicians***	***Non-medical and medical students***	***Total***
Switching from B to A	99	433	*16.35*	*26.64*	*22.86*
Always Option A	9	433	*3.14*	*1.46*	*2.08*
Always Option B	28	433	*14.47*	*1.82*	*6.47*
Total inconsistent choices	**136**	**433**	***33.96***	***29.93***	***31.41***

Source: our elaboration using data from Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021; Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b.

Table 9. Interval regression model estimation of coefficient of relative risk aversion – full sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**
Constant	0.1812***	0.2336***	0.1663***	0.2050***	0.0310***	0.2118***	0.0416	0.1033**	-0.0072	0.0525	-0.3556***	-0.2568***
	(0.0125)	(0.0204)	(0.0123)	(0.0200)	(0.0109)	(0.0207)	(0.0459)	(0.0502)	(0.0460)	(0.0494)	(0.0621)	(0.0770)
Doctor	-0.3851***	-0.3821***					-0.3972***	-0.3600***				
	(0.0207)	(0.0612)					(0.0452)	(0.0689)				
Field			-0.3726***	-0.6044***					-0.3118***	-0.5195***		
			(0.0212)	(0.0402)					(0.0428)	(0.0523)		
Medical students					0.0980***	-0.0445					0.3972***	0.3600***
					(0.0349)	(0.0345)					(0.0452)	(0.0689)
Age							0.0028**	0.0021*	0.0021*	0.0011	0.0028**	0.0021*
							(0.0012)	(0.0013)	(0.0013)	(0.0013)	(0.0012)	(0.0013)
Female							0.0366*	0.0425**	0.0222	0.0296	0.0366*	0.0425**
							(0.0204)	(0.0204)	(0.0204)	(0.0203)	(0.0204)	(0.0204)
Student							0.0651*	0.0677**	0.1371***	0.1423***	0.4622***	0.4277***
							(0.0347)	(0.0345)	(0.0316)	(0.0315)	(0.0352)	(0.0630)
Experiment fixed effects	No	yes	no	yes	no	yes	no	yes	no	yes	no	yes
Observations	433	433	433	433	433	433	433	433	433	433	433	433
Log-likelihood	-9313.525	-9287.889	-9330.4472	-9307.2921	-9476.5876	-9306.4595	-9308.1756	-9282.9523	-9320.0946	-9296.5642	-9308.1756	-9282.9523

Source: our elaboration using data from Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021; Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b.
Notes: Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 10. Interval regression model estimation of coefficient of relative risk aversion – subsample of consistent choice

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**	**Relative_risk aversion**
Constant	0.1785***	0.1929***	0.1699***	0.1754***	0.1515***	0.1926***	0.0313	0.0641*	0.0326	0.0469	0.0071	-0.0497
	(0.0086)	(0.0147)	(0.0084)	(0.0143)	(0.0073)	(0.0148)	(0.0330)	(0.0353)	(0.0331)	(0.0347)	(0.0442)	(0.0523)
Doctor	-0.0960***	-0.1993***					-0.0542*	-0.1138**				
	(0.0144)	(0.0407)					(0.0315)	(0.0460)				
Field			-0.0784***	-0.2454***					0.0523**	-0.1390***		
			(0.0148)	(0.0291)					(0.0296)	(0.0374)		
Medical students					-0.0728***	-0.1065***					0.0242	0.1138**
					(0.0237)	(0.0240)					(0.0315)	(0.0460)
Age							0.0014	0.0005	0.0003	0.0001	0.0014	0.0005
							(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Female							0.0284**	0.0352**	0.0268*	0.0312**	0.0284**	0.0352**
							(0.0144)	(0.0143)	(0.0144)	(0.0142)	(0.0144)	(0.0143)
Student							0.1216***	0.1245***	0.1459***	0.1497***	0.1458***	0.2383***
							(0.0243)	(0.0241)	(0.0219)	(0.0219)	(0.0247)	(0.0420)
Experiment fixed effects	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
Observations	297	297	297	297	297	297	297	297	297	297	297	297
Log-likelihood	-4968.2641	-4937.9424	-4976.2347	-4949.9227	-4985.541	-4954.0804	-4922.7071	-4953.7795	-4953.7795	-4925.7587	-4954.0804	-4922.7071

Source: our elaboration using data from Finocchiaro Castro et al., 2019; Finocchiaro Castro et al., 2021; Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b.
Notes: Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

APPENDICES

APPENDIX A

In this Appendix we explore the role of consistent and inconsistent choices in our sample. Table A.1 replicates Table 5, considering consistent choices of a smaller sample of subjects. By excluding inconsistent subjects, differences are still significant for some pairwise comparisons: laboratory versus field experiment, all subjects versus doctors, and students versus medical students. Focusing on the latter, the result is even strengthened with respect to Table 5, confirming that, also when looking at consistent choices only, students are more risk averse than their peers enrolled in medical degrees. In this subsample, the differences in risk profiles between medical students and doctors are no longer significant. On the contrary, gender differences become significant for the physicians' sample, confirming the mixed results in the literature extensively reviewed by Filippin and Crosetto (2016).

<<Table A.1 around here>>

The next step is to identify the determinants of the number of safe choices for consistent and inconsistent subjects. In Table A.2, we report the OLS estimates of the determinants of number of safe choices for the full sample, whereas Table A.3 reports the estimates for the restricted dataset of consistent choices. We employ a parsimonious approach starting from the simplest estimation model and moving to the full model, which also includes several variables to control for characteristics of the experimental designs. Our analysis focuses on three groups: doctors (159), medical students (42) and nonmedical students (232).

The results in Table A.2 confirm that physicians are more risk seeking. Than other groups. However, we confirm the results from the non-parametric tests where medical students did not behave significantly differently from other students and were significantly more risk-averse than doctors.

<<Table A.2 around here>>

Considering our restricted dataset, Table A.3 reports the regression analysis for the same models in Table A.2, starting again from unconditional model. The results shown in Table A.3 largely confirm those for the full sample. As we would expect looking at the results of the non-parametric tests shown in Table A.1, medical students now generally show a greater propensity to take risks than the rest of the sample. However, the results become again positive and significant when we consider the model with full controls.

<<Table A.3 around here>>

Overall, the results reported here show that inconsistent choices could play a considerable role. Nevertheless, even when considering only individuals who made consistent choices, the results show that physicians have a higher propensity to take risks than other groups, regardless of their age, gender and whether the experiment is conducted in the field or in the laboratory.

Furthermore, the choices we defined as inconsistent may not be such. In fact, they could be due either to random errors or to a violation of the expected utility model underlined in Holt and Laury (2002).

Table A.1. Mann-Whitney test by type of data, subject, and gender - subsample of consistent choices

Test	N_1	N_2	N_safe_1	N_safe_2	P-value
Laboratory vs. Field data	201	96	4.637	4.427	0.0000
All subjects vs. Doctors	192	105	4.667	4.390	0.0000
Medical students vs. Doctors	28	105	4.321	4.390	0.2376
Students vs. Medical students	164	28	4.726	4.321	0.0000
Male vs. Female students	122	70	4.731	4.714	0.0442
Male vs Female medical students	14	14	4.357	4.286	0.5921
Male vs Female doctors	57	48	4.298	4.500	0.0021

Source: our elaboration using data from Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b; Finocchiaro Castro et al., 2021 and Finocchiaro Castro et al., 2019

Table A.2. Estimates of the determinants of safe choices – full sample

VARIABLES	(1) N_safe	(2) N_safe	(3) N_safe	(4) N_safe	(5) N_safe	(6) N_safe	(7) N_safes	(8) N_safe	(9) N_safes	(10) N_safe	(11) N_safes	(12) N_safe
Constant	4.7336***	4.8684***	4.6888***	4.7830***	4.3785***	4.8041***	4.0674***	4.2199***	3.9290***	4.0562***	2.9193***	3.0725***
	(0.0352)	(0.0578)	(0.0346)	(0.0566)	(0.0302)	(0.0584)	(0.1284)	(0.1411)	(0.1287)	(0.1389)	(0.1735)	(0.2161)
Doctor	-0.9160***	-1.1318***					-1.1482***	-1.1475***				
	(0.0580)	(0.1724)					(0.1267)	(0.1937)				
Field			-0.8589***	-1.3941***					-0.8394***	-1.2991***		
			(0.0593)	(0.1123)					(0.1200)	(0.1463)		
Medical students					0.1929**	-0.1394					1.1482***	1.1475***
					(0.0971)	(0.0973)					(0.1267)	(0.1937)
Age							0.0173***	0.0153***	0.0142***	0.0119***	0.0173***	0.0153***
							(0.0034)	(0.0036)	(0.0035)	(0.0035)	(0.0034)	(0.0036)
Female							0.1585***	0.1746***	0.1161**	0.1335**	0.1585***	0.1746***
							(0.0572)	(0.0574)	(0.0571)	(0.0572)	(0.0572)	(0.0574)
Student							0.2065**	0.2126**	0.4397***	0.4509***	1.3547***	1.3601***
							(0.0974)	(0.0970)	(0.0888)	(0.0887)	(0.0984)	(0.1771)
Experiment fixed effects	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
Observations	433	433	433	433	433	433	433	433	433	433	433	433
R-squared	0.0545	0.0625	0.0462	0.0532	0.0009	0.0536	0.0619	0.0688	0.0547	0.0612	0.0619	0.0688

Source: our elaboration using data from Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b; Finocchiaro Castro et al., 2021 and Finocchiaro Castro et al., 2019

Notes: Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A.3. Estimates of the determinants of safe choices - subsample of consistent choices

VARIABLES	(1) N_safe	(2) N_safe	(3) N_safe	(4) N_safe	(5) N_safe	(6) N_safe	(7) N_safes	(8) N_safe	(9) N_safes	(10) N_safe	(11) N_safes	(12) N_safe
Constant	4.6667***	4.7077***	4.6368***	4.6471***	4.5948***	4.7066***	4.0719***	4.1749***	4.0735***	4.1120***	3.9641***	3.7566***
	(0.0311)	(0.0534)	(0.0304)	(0.0519)	(0.0263)	(0.0537)	(0.1195)	(0.1282)	(0.1196)	(0.1258)	(0.1599)	(0.1896)
Doctor	-0.2762***	-0.6875***					-0.1378**	-0.4184**				
	(0.0522)	(0.1473)					(0.0640)	(0.1670)				
Field			-0.2097***	-0.7380***					-0.1255*	-0.4424***		
			(0.0535)	(0.1051)					(0.0731)	(0.1352)		
Medical students					-0.2734***	-0.3682***					0.1078	0.4184**
					(0.0857)	(0.0871)					(0.1140)	(0.1670)
Age							0.0085***	0.0056*	0.0047	0.0042	0.0085***	0.0056*
							(0.0032)	(0.0034)	(0.0033)	(0.0033)	(0.0032)	(0.0034)
Female							0.0857	0.1088**	0.0795	0.0939*	0.0857	0.1088**
							(0.0522)	(0.0520)	(0.0520)	(0.0517)	(0.0522)	(0.0520)
Student							0.4205***	0.4293***	0.5114***	0.5220***	0.5283***	0.8477***
							(0.0880)	(0.0875)	(0.0794)	(0.0794)	(0.0893)	(0.1523)
Experiment fixed effects	No	yes	no	yes	no	yes	no	yes	no	yes	no	yes
Observations	297	297	297	297	297	297	297	297	297	297	297	297
R-squared	0.0093	0.0261	0.0052	0.0190	0.0034	0.0248	0.0191	0.0354	0.0191	0.0333	0.0191	0.0354

Source: our elaboration using data from Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b; Finocchiaro Castro et al., 2021 and Finocchiaro Castro et al., 2019

Notes: Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

APPENDIX B

In this Appendix, we assume that the sequence of choices made by the subject in the HL multiple price list lottery can be considered as an independent observation. This approach requires some highly restrictive assumptions and is only reported here as a robustness check of the results provided in Section 4.

As pointed out by Filippin and Crosetto (2016), facing patterns close to that described in Appendix A does not necessarily result in a violation of the expected utility theory. To be more specific, an individual can still behave in accordance with expected utility theory but simultaneously make mistakes. Under such assumption, we will next report a model which includes a stochastic component for each of the pair choices. In fact, one of the limitations of the previous models suggested by Harrison and Rutström (2008) is that by estimating only one single parameter *r* we infer the bounds which make the subject indifferent between lottery A and B. Thus, we exclude subjects making multiple switches, which instead would require multiple parameters to be estimated. Here, instead, we will derive 10 parameters for each of the subjects, as binary choice (lottery A or lottery B), thus allowing for multiple switches. In fact, including also inconsistent subjects is the only way to assess whether a pool's behaviour differs from another one and eventually which variable impacts on such difference.

More specifically, to assess the lottery choices while controlling for the level of noise in decision making, we build a random utility structural model and estimate it through maximum likelihood. To perform our estimations, we use the script provided by Harrison (2008) and the error specification of Holt and Laury (2002).

For each of the binary choice, each subject will select the option depicted on the right choice (B) (the riskier one) whenever the expected utility (EU) from that option is larger than the expected utility from the left option (A) (the safer one) plus the random component μ. More specifically, we assume the simple version of CRRA utility function assuming that subjects are expected utility maximisers characterized by $U(x) = x^r$, and that they can make an evaluation error μ when comparing the utility between A and B choices. Under these assumptions the probability of choosing the safe lottery is

$$Prob(S) = \frac{EU_A^{\frac{1}{\mu}}}{EU_B^{\frac{1}{\mu}} - EU_A^{\frac{1}{\mu}}} \text{ and for subject } i \;\; EU_i = \sum_j p_j (x_j)^r \qquad \text{(B.1.)}$$

where A is the safe lottery, B the risky lottery and μ is the noise parameter.

Given the above assumptions we can write the loglikelihood function such that:

$$LogLik = \begin{cases} ln\ 1 - Pr(S) & if\ subject\ selects\ lottery\ A \\ ln\ \Pr(S) & if\ subject\ selects\ lottery\ B \end{cases} \qquad (B.2)$$

and then estimate a structural model of choice using maximum likelihood and clustering standard errors by subject. The model is estimated using the amounts in euros of the Table 1. The estimates resulting from the structural model through the maximum likelihood are provided in Table B.1.

<<Table B.1 around here>>

The results in Table B.1 largely confirm the previous findings on risk attitude of physicians. The estimated risk parameter *r* is around 0.45 on average and, considering that we have only one treatment in our sample, is substantially in line with the estimates of Holt and Laury (2002). As far as *r* is concerned, physicians show a significantly lower risk aversion compared with other subjects. The same result is also obtained when considering the data collected in the field. Both results are also robust when we include controls for age, gender, and experiment. In Table B.1 the average noise level μ is about 3. This value is rather high but not surprising considering the analysis on the number of inconsistent choices in our sample reported in Section 3. Physicians have a significantly lower μ than students suggesting that doctors make fewer errors in their choices than others. Once again, these results are comparable both for the group of doctors and when considering the results in the field and controlling for age, gender, and experiment.

Summing up, the robustness check reported in this Appendix confirm that doctors are less risk-averse in the financial domain than other subject.

Table B.1 – Estimated risk aversion parameter under CRRA - Maximum-Likelihood estimates

CRRA specification $u(x) = x^r$				
	(1)	(2)	(3)	(4)
r	0.3555*** (0.0962)	0.4730*** (0.1081)	0.3893*** (0.1043)	0.4895*** (0,1104)
r_{doctor}	-0.3978*** (0.1016)	-0.6284*** (0.2264)		
r_{field}			-0.3680*** (0.1071)	-0.7326*** (0.2105)
μ	3.5035*** (0.6139)	3.7224*** (0.7424)	3.4298*** (0.5983)	2.9465*** (0.5242)
μ_{doctor}	-4.1664*** (1.1155)	-2.9456*** (1.0221)		
μ_{field}			-4.5600*** (1.3788)	-2.1880*** (0.6145)
Full controls	no	yes	no	yes
Log likelihood	-1860.7827	-1856.3575	-1862.6079	-1849.1810

Source: our elaboration using data from Finocchiaro Castro and Romeo, 2021a; Finocchiaro Castro and Romeo, 2021b; Finocchiaro Castro et al., 2021 and Finocchiaro Castro et al., 2019

Notes: number of decisions 4,330; number of subjects 433. Robust standard error clustered by subjects reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively