2016

Bringing a Natural Experiment into the Laboratory: The Measurement of Individual Risk Attitudes

Zuzana Brokesova
University of Economics in Bratislava

Cary Deck
Chapman University, deck@chapman.edu

Jana Peliova
University of Economics in Bratislava

Bringing a Natural Experiment into the Laboratory: the Measurement of Individual Risk Attitudes*

Zuzana BROKESOVA[a], Cary DECK[b], Jana PELIOVA[c]

[a] Corresponding author, Faculty of National Economy, University of Economics in Bratislava, Dolnozemska cesta 1, 852 35 Bratislava, Slovak Republic, zuzana.brokesova@euba.sk, +421 2 672 91 562

[b] College of Business and Public Policy University of Alaska Anchorage, 3211 Providence Drive, Anchorage, AK 99508, United States; Sam M. Walton College of Business, University of Arkansas, Fayetteville, AR 72701, United States & Economic Science Institute, Chapman University, One University Drive, Orange, CA 92866, United States, cdeck@walton.uark.edu

[c] Faculty of National Economy, University of Economics in Bratislava, Dolnozemska cesta 1, 852 35 Bratislava, Slovak Republic, jana.peliova@euba.sk

* This paper is part of a research project No. 2015 - 3-02/6 entitled 'Support for the internationalization of scientific research in economic sciences' supported by the Foundation VUB (Nadácia VÚB) and research grants VEGA No. 1/0849/15 and No. 1/0964/15 supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic. We would like to acknowledge Mr. Marek Marcinka for help in design of experimental software.

Bringing a Natural Experiment into the Laboratory: the Measurement of Individual Risk Attitudes

Abstract

Controlled laboratory experiments have become a generally accepted method for studying economic behavior, but there are two issues that regularly arise with such work. The first pertains to the ability to generalize experimental results outside the laboratory. While laboratory experiments are typically designed to mimic naturally occurring situations, ceteris paribus comparisons are rare. Using data from a promotional campaign by a bank and a matching laboratory experiment, we find similar patterns of risk taking behavior controlling for gender and age. The second issue pertains to the impact that the payment procedure in an experiment has on observed risk taking behavior. Specifically, we compare behavior on a risk taking task where that is the only task and payment is assured, where it is one of several similar tasks of which one will be randomly selected for payment, and where it is the only task but there is only a small probability of receiving payment. We find similar behavior across these three payment procedures.

Keywords: Natural Experiment, Laboratory Experiment, Experimental Methodology, Risk Attitudes

JEL codes: C91, C99, D81

1. Introduction

Almost every economic decision involves risk. Therefore, it is unsurprising that numerous researchers have attempted to measure risk attitudes of individuals. Unfortunately, empirical work is often plagued by important parameters, such as the probability of an outcome or its associated payoff to the decision maker, being unobservable. For this reason, much of the risk measurement work has been conducted in the lab. This raises the question of to what degree do these choices made in the lab reflect risky choices that would be made outside of the lab.

In a well-known paper, List (2006) compares a gift exchange game with private information across the lab, field, and intermediate settings. The conclusions suggest that strategic behavior differs between the field and the lab.[1] Ideally, one would like to compare risk taking in the lab with the same choice faced outside of the lab, but there is relatively little direct basis for such a comparison because of the normally unobservable characteristics of risky choices in life. One exception, is television game shows such as *Deal or No Deal* where the rules are well defined, which explains why a series of papers came out when the show was introduced (see Blavatskyy and Pogrebna 2008, Deck et al. 2008, Post et al. 2008, De Roos and Sarafidis 2010, and Andersen et al. 2006, 2014). Both Baltussen et al. (2008) and Deck et al. (2013) develop versions of this game show for use in the laboratory, but with smaller stakes, no audience, and no host egging on the player thus thwarting a direct comparison of risk attitudes between the lab and the field.

In this paper, we exploit a promotional campaign conducted by a bank that offered potential customers a chance to receive up to 1,000 € to measure risk attitudes in the field. We also implement a parallel decision including the stakes and presentation in the lab. The observed behavior is similar between the two data sets controlling for the age and gender of the decision maker suggesting that lab experiments provide a reasonable measure of risk aversion.

[1] However, as pointed out by Falk and Heckman (2009), a specific field experiment is no more generalizable than a specific lab experiment. Indeed, when studying a treatment effect, either the effect is separable from other characteristics in which case the lab versus field distinction is moot or it is not separable in which case neither a field study nor a lab experiment can claim greater insight for some third set of characteristics without further auxiliary assumptions. In the case of risk taking, the maintained assumption is that individuals have a risk tolerance (for financial risks) that is invariant to the setting.

One drawback to the analyzed data is that the level of precision in the measured risk attitude is weak. This is due to the structure of the choice that decision makers faced in the field: a binary choice between a certain payment and a risky payment. Hence, one can only draw inferences that a person is more or less risk averse than some threshold. This contrasts with the approach of most laboratory experiments. Perhaps the most popular approach for measuring risk tolerance in the lab is that of Holt and Laury (2002) in which subjects are asked to make a series of binary choices over lottery pairs.[2] By presenting a subject with several choices, a finer partitioning of a respondent's risk attitude is possible. However, having a respondent make multiple decisions necessitates a design choice by the researcher. If all (or even just two) of the choices are used to determine payment then potential wealth effects are introduced. For this reason, researchers instead frequently select one task at random for payment. While some have provided evidence to suggest that this random payment incentive does not alter behavior (e.g. Cubitt et al. 1998 and Starmer and Sugden 1991), recent work by Cox et al. (2015) argues that this technique is not incentive compatible if the independence axiom does not hold and behaviorally can bias behavior.[3]

To provide further evidence on the degree to which presenting subjects multiple tasks, one of which is selected at random for payment, impacts risk taking behavior, we introduce a second laboratory treatment. Here a subject faces five risky choices, including the one faced by the bank customers, one of which is randomly selected for payment. This treatment also allows us to partition risk attitudes into similar bins as those used in Holt and Laury (2002). We find that subjects are generally consistent with their choices across tasks. Further, most of the subjects can be classified as being modestly risk averse. But most importantly, we do not observe any differences in behavior between this treatment and either the field data or the single task experiment.

Finally, a third treatment is included that examines how another common payment technique impacts risk taking behavior. Rather than paying each participant, some researchers randomly

[2] For a review of experimental methodologies to elicit and assess individual risk attitudes see Charness, Gneezy and Imas (2013).
[3] Approaches such as that of Eckel and Grossman (2002), avoid this issue by presenting a single choice among several lotteries.

select one (or possibly more) of the *n* respondents to actually receive payment (e.g. Tversky and Kahneman 1981, Langer and Weber 2008). Such procedures are typically accompanied by a statement reminding the participant that her choices might determine her payment and therefore it is in her best interest to respond as if they will. This approach allows a researcher to collect *n* times more observations or increases the nominal stakes *n*-fold for the same expected cost. While the attraction to this approach is clear, it is important to identify to what degree it influences behavior.[4] Ultimately, we find that subjects in this treatment make similar decisions to those in the other two treatments and in the field.

2. Risk Data from the Field

In a natural experiment, a private bank in the Slovak Republic conducted a marketing promotion to attract new clients. Individuals, who had a minimum of 1,000 € could open a savings account to which the bank would add either a fixed amount of 20 € or a randomly determined amount.[5] The critical feature of this campaign is that the distribution for the risky payment was available to the decision makers. This distribution, shown in Table 1, has an expected value of 27.5 €.

Table 1. Distribution of Random Payment by Bank

Payoff	**Likelihood**
10 €	50%
20 €	39%
30 €	5%
50 €	3%
100 €	2%
1,000 €	1%

[4] In the previous research, there were identified differences in individuals choices based on the different applied random incentive system (Baltussen et al. 2012, Harrison and Swarthout 2014, March et al. 2014). In the area of risk elicitation, Baltussen et al. (2012) find out that between-subjects randomization (i.e. a randomly selected subsample of all subjects are payed - Pay Random Task) reduces risk aversion, while within-subjects randomization (i.e. a payment method where only one of the series of tasks/decisions is randomly selected for payment - Pay Random Subject) delivers unbiased measurements of risk aversion. Similarly Harrison and Swarthout (2014) find the effect of the payment protocols on preferences for risk.

[5] The award amount was instantly reflected in the account balance, but to receive the money a person had to keep a minimum balance of 1,000 € excluding the award in the new account for 3 months.

To make the decision to take the safe or risky payment, a person went to the bank's website where the official rules were available. Then, one would click the button associated with their choice. If the person chose to "Roll the Dice," an image of die rolled across their screen with each face displaying one of the six possible prize amounts. Under the assumption of constant relative risk aversion (CRRA), an individual with a risk parameter of $r = 0.38$ would be indifferent between the two options.

A total of 3,917 people participated in the promotion, 75.64% of whom opted for the random payment. Of the participants, 69.9% were male and the average age was 36.92 (more summary statistics are provided in Table 4). Several laboratory experiments have reported that women are more risk averse then men (Powell and Ansic 1997, Eckel and Grossman 2002, 2008, Fehr-Duda, Gennaro, and Schubert 2006, Agnew 2008, Borghans et al. 2009, Charness and Gneezy 2010, Dohmen et al. 2011).[6] However, relatively few studies have examined the effect of age on risk taking (e.g. Harbaugh, Krause and Vesterlund 2002, Harrison, Lau and Rutström 2007, Tymula et al. 2013).[7] Figure 1 plots the percentage of males and females who opted for the risky payment by age. The size of the marker indicates the proportion of the data accounted for by a particular age and gender combination. Based on Figure 1, it appears that there is no gender difference in the behavior of the bank customers, but behavior does appear to vary with age as older people are less willing to take the risky option. These conclusions are supported econometrically in Table 2, which reports the results of estimating a probit model allowing for gender, age, and an interaction between the two. The dependent variable takes a value of 1 if the person opted to "Roll the Dice" and accept the risky payment. Otherwise, it is 0. The coefficient on age is negative and significant while the coefficient on gender is not statistically significant.

[6] The pattern is not universal. For example, Charness and Gneezy (2004) and Schubert et al. (1999) find that men are more risk averse in the domain of loss. Schubert (2000), Filippin and Crosetto (2014) and Harbaugh, Krause and Vesterlund (2002) find no difference between men and women. Brinig (1995) finds also no evidence of a sex differences, however in case of the interaction of sex and age it becomes a significant factor of prediction risk-taking behavior.

[7] Results are not uniform, e.g. Tymula et al. (2013) conclude that in the gain domain, older individuals do take less risk than younger individuals. On the other hand, Harbaugh, Krause and Vesterlund (2002) find no age difference in risk taking and Harrison, Lau and Rutström (2007) reveals age effect on risk attitudes in controlled experiments. In meta-analysis of 29 papers, Mata et al. (2011) find that age difference is a function of task characteristics.

Figure 1. Risk Taking in Natural Experiment

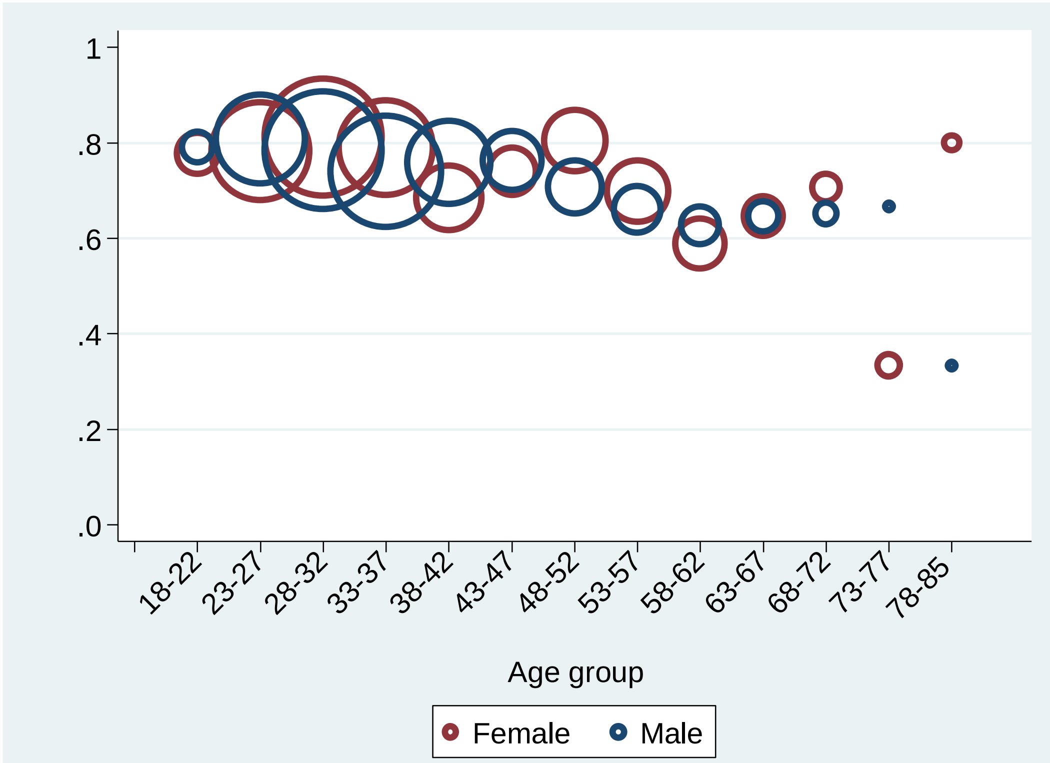

Table 2. Probit Analysis of Natural Experiment

Variable	**Coefficient**
Gender	.0359 (.1555)
Age	-.0124** (.0025)
Gender x Age	-.0002 (.0039)
Constant	1.1504** (.0952)
Pseudo R^2	0.0099
Observations	3,917

Dependent variable is binary and equals to 1 if the person opted to "Roll the Dice" and accept the risky payment. Otherwise, it is 0. Standard errors are in parentheses. *, ** denote significance at the 5% and 1% level, respectively.

3. Risk Data from the Lab

Lab data were collected from laboratory subjects in three between-subject treatments. The first, which is referred to as the Baseline, was designed to closely match the field data from the bank promotion. The other two treatments are designed to evaluate the effect of two common payment procedures used in laboratory experiments. Referred to as “Pay Random Task” and “Pay Random Subject”, these treatments are explained in detail below.

The experiments were conducted in the Economic Laboratory at the University of Economics in Bratislava. Participants were primarily students, but some were staff and others were from the general public. Because participants in the bank promotion had to open an account with a 1,000 € balance, potential laboratory subjects completed a short questionnaire by email that included a question about their current wealth and their ability to open such an account. Only those who would have been able to open an account were invited to participate in the experiment, but the questionnaire also contained other distraction questions so that respondents were not aware of this fact.

In the Baseline treatment, subjects read “rules” that closely mirrored those of the bank promotion (English translations of the bank promotion rules and the experimental instructions are included in the appendix while copies in the original Slovak are available upon request). The choice problem was the same in terms of stakes and probabilities. The visual presentation of the task and the manner of providing information regarding the distribution for the risky payment were virtually identical. The one difference is that subjects in the laboratory were not required to actually open an account and wait three months for the payment. Instead, they were paid in cash at the conclusion of the experiment. Hence, the real stakes were slightly higher in the experiment than in the field.

The Pay Random Task was identical to the Baseline except that subjects made five choices, in a randomly determined order, and only one of these choices was used to determine the subjects' payment. For each choice task, the safe lottery was a certain payment of 20 €. The distribution for the risky payments for each task is shown in Table 3. No new payment amounts were introduced, but the probabilities were changed so that the level of CRRA for which a subject would be indifferent varied from task to task. The specific CRRA parameter thresholds were chosen to be similar to those in Holt and Laury (2002). For the first task, subjects observed the full set of rules, but on subsequent tasks, only the changes (i.e. the relevant distribution) were presented. This was done out of concern if the full rules were used each time then subjects might begin reading the "rules" and conclude that it was an identical decision because the first part of the text would have been repetitive. One should recall that the rules closely follow those implemented by the bank.

Table 3. Distributions of Risky Payment for Pay Random Task

Payoff	**Likelihood**				
	Game 1	**Game 2**	**Game 3 (Baseline)**	**Game 4**	**Game 5**
10 €	24%	33%	50%	85%	99%
20 €	55%	49%	39%	5%	0%
30 €	15%	12%	5%	4%	0%
50 €	3%	3%	3%	3%	0%
100 €	2%	2%	2%	2%	0%
1000 €	1%	1%	1%	1%	1%
CRRA parameter for which one is indifferent between lottery and €20	0.96	0.68	0.38	0.15	-0.01

The Pay Random Subject treatment was identical to the Baseline except that each participant was informed that only one person in the session would actually be paid based upon her choice. A session, meaning a group of people in the laboratory at the same time, consisted of 16-18 people and this was the case for all three treatments.

Experimental sessions lasted approximately 30 minutes and the average salient earnings were 19.03 €. For comparison, the average hourly wage in the Slovak Republic is 5.02 € and therefore the top possible prize of 1,000 € represents a substantial payment. Participants in the Baseline and Pay Random Task did not receive any payment besides the salient earnings, but those subjects who were not randomly selected in the Pay Random Subject earned a flat 3 € for their time. After completing the experiment, subjects in each treatment completed a brief survey which included demographic information as well as the three questions Cognitive Reflection Task (CRT) based on Frederick (2005) general risk taking question from the German Socio-Economic Panel (SOEP)[8].

Table 4 provides summary statistics for each of the three treatments as well as the field data. Clearly, the laboratory subjects tended to be younger than the respondents in the field and a slightly larger percentage were female. Two-sample Hotelling's T-Square test statistic verifies that there were no differences in the composition of subjects in each of the three treatments.[9] Therefore, behavioral comparisons between experimental treatments are based on two sample proportions tests whereas as comparisons between the field data and laboratory data are based upon probit regression analysis.

Table 4. Summary Statistics

	Field Data	**Laboratory Experiments**			
		Combined	**Baseline**	**Pay Random Task[A]**	**Pay Random Subject**
Observations	3,917	162	56	55	51
% Safe	24.36	14.81	12.50	16.36	15.69
% Male	66.90	55.56	57.14	54.55	54.90
Age					
Average	36.92	26.07	25.45	26.09	26.61
Range	18-85	18-70	19-65	18-66	19-70
Average CRT Score	-	1.62	1.80	1.58	1.59
Average Risk Taking[B]	-	.40	.39	.39	.43

[A] Behavior reported here is based on game 3 which is the same choice as in the other treatments.
[B] Average investment measured by general risk taking question from the SOEP.

First, we compare behavior across the three experimental treatments. The two sample proportion test p-values for Baseline vs Pay Random Task, Baseline vs Pay Random Subject, and Pay Random Task vs Pay Random Subject are 0.5626, 0.6350, and 0.9252, respectively. Therefore, we

[8] General question measuring risk attitudes were added to the SOEP in the 2004 wave.
[9] The null hypothesis is that the vector of characteristic means is the same in the two treatments being compared. The p-values for comparing Baseline vs Pay Random Task, Baseline vs Pay Random Subject, and Pay Random Task vs Pay Random Subject are 0.8658, 0.8194, and 0.9291, respectively.

conclude that in this case behavior was the same in the one shot baseline as in the Pay Random Task treatment and the Pay One Subject treatments.[10] This is formalized in Finding 1.

Finding 1. *Subjects make similar choices regarding risk taking under both a Pay Random Task and a Pay Random Subject protocol to the choices made in a salient standalone choice.*

If we repeat this exercise separately for males and for females, we again find no evidence of a treatment effect for either gender (the smallest of the 6 p-values is 0.2646). We also compare males to females in each of the three treatments and find no evidence of a gender difference (the smallest p-value is 0.7620, which arises in the Pay Random Subject treatment).

Having established that the three laboratory treatments yield similar behavior, we combine these data to compare behavior in the lab with behavior in the field. Table 5 reports a probit regression comparing the decision to "Roll the Dice" across the two data sources. The explanatory variable Lab takes the value 1 if the decision was made in the lab and is 0 otherwise. The statistical analysis reveals that, controlling for gender and age, behavior is similar in the lab and field.[11] This is formalized in Finding 2.

Finding 2. *Subjects in the laboratory and participants in a natural experiment make similar risk taking choices.*

Table 5. Probit Regression Comparing Risk Taking in the Field and in the Lab.

Variable	**Coefficient**
Lab	.2085
	(.1247)
Age	-.0120**
	(.0019)
Gender	.0363

[10] This conclusion is also supported through probit regression.
[11] A similar result is found if one allows for separate indicator variables for each of the three laboratory treatments.

	(.0473)
Constant	1.1329**
	(.0744)
Pseudo R^2	0.0110
Observations	4,079

Dependent variable is binary and equals to 1 if the person opted to "Roll the Dice" and accept the risky payment. Otherwise, it is 0. Standard errors are in parentheses. * and ** denote significance at the 5% and 1% level, respectively.

While the main focus of this paper is on comparing behavior for the choice that is common to all four conditions, the Pay Random Task treatment affords a finer analysis of individual behavior (which is why it is commonly employed). A subject's choices in Pay Random Task are deemed to be consistent if when a subject opts to "Roll the Dice" in Game *i* then the subject also opts to "Roll the Dice" in game *j* for *j* > *i* as this behavior can be rationalized by a CRRA function. Overall, 74.54% of the subjects were consistent. By comparison, in their baseline condition Holt and Laury (2002) found that 13.2% of subjects behaved consistently. Table 6 compares the distribution of consistent subjects in our study with those from Holt and Laury (2002) for the 20x real treatment, which has the most similar payoff level to ours. It is clear from Table 6, that our subjects were far more risk tolerant than theirs. This is conclusion is supported by a Wilcoxon rank sum test (p-value = 0.000 for testing the two distributions are the same).

Table 6. Distribution of Risk Attitudes

CRRA parameter	**Proportion of choices**	
	Holt and Laury (2002) (20x Real Treatment)	**Our Lab Data**
r ≤ 0.15 (risk neutral or risk loving)	0.19	0.76
0.15 ≤ r ≤ 0.40(slightly risk averse)	0.19	0.17
0.40 ≤ r ≤ 0.68 (risk averse)	0.23	0.00
0.68 ≤ r ≤ 0.96 (very risk averse)	0.22	0.00
0.96 ≤ r (highly risk averse)	0.17	0.07

The boundaries are approximate as there were slight differences in the interval between the two studies. The Holt and Laury (2002) data can be further separated into those that are risk neutral (-0.15 ≤ r ≤ 0.15) and those that are strictly risk loving (r ≤ -0.15). The groups account for 13% and 6 % of their respondents, respectively. Our data can further separated into those that are mildly risk averse (-0.01 ≤ r ≤ 0.15) and those that are risk loving (r ≤ -0.01). These groups account for 15% and 61% of our subjects, respectively.

Finally, we consider how survey responses, specifically to the CRT and general risk taking question, correlate with observed behavior. With respect to CRT, previous experimental research has found that increasing the decision maker's cognitive load can lead to more risk averse behavior (e.g. Whitney et al. 2008, Gerhardt 2013, Deck and Jahedi 2015).[12] Deck and Jahedi (2015) argue that increased cognitive load increases reliance on the brain's more intuitive System 1 rather than the cognitive System 2 (see Kahneman 2003 for a discussion of the dual system). As the CRT can be viewed as a measure of the degree to which a person relies upon System 2 in decision making, one would expect those who score higher on the CRT to be more willing to "Roll the Dice." To test this hypothesis, we conducted additional analysis as reported in Table 7. For this analysis, data from all three laboratory treatments are combined. The first specification (Model 1) only controls for CRT while the second specification (Model 2) includes demographic controls. Consistent with past work that has found that males perform better on the CRT (Frederick 2005, Oechssler et al. 2009, Hoppe and Kusterer 2011), we observe a statistical difference between males and females (t-test for average score has p-value < 0.0001). Therefore, model 2 includes an interaction between Female and CRT but it does not affect subjects' risky decisions. Model 3 of Table 7 controls for the subject's risk attitude as measured in the post experiment survey. Based on the results presented in Table 7, a subject's response to the risk taking question does not predict actual risk taking behavior as the coefficient on the survey measure of risk taking is not significant.

Table 7. Probit analysis of the role of cognitive ability in decision to roll a dice

Variable	**Model 1**	**Model 2**	**Model 3**
CRT	.0325 (.1447)	.1466 (.1856)	
Survey Risk			.2516 (.5591)
Female		.5257 (.5649)	.1659 (.2581)
Age		.0086 (.0145)	.0101 (.0145)

[12] While cognitive ability and reflection are distinct concepts, there is also evidence that those with lower cognitive ability tend to be more risk averse (see Burks et al. 2009, Dohmen et al. 2010, Benjamin et al. 2013). Once can also view increasing cognitive load as effectively reducing cognitive ability.

CRT x Female		-.2244	
		(.3201)	
Constant	.9908**	.4924	.6163
	(.2667)	(.4958)	(.4644)
Pseudo R^2	0.0004	0.0128	0.009
Observations	162	162	162

4. Discussion

This paper provides insight on two recurring concerns in economic experiments measuring risk attitudes. The first pertains to the similarity in risk taking in and out of the lab. Exploiting a bank promotion that presented people with a risky choice, we develop a laboratory experiment designed to match the promotion as closely as possible. Almost four thousand decisions were obtained via the natural experiment, revealing clear age effect on risk attitude, but no evidence of a gender effect. The laboratory data for the same decision also reveal no gender difference in risk attitudes, consistent with Schubert (2000), Harbaugh, Krause and Vesterlund (2002) and Filippin and Crosetto (2014). More importantly, controlling for the difference in the age composition of the two groups, behavior in the lab and the field are similar suggesting that lab experiments do provide a usable representation of risk taking outside of the lab.

The second contribution regards the behavioral consequences of common procedures in risk experiments. Having subjects complete several tasks and randomly selecting one for payment did not impact behavior. Neither did having only some randomly selected people actually receive payment for their choices. This suggests that the effect of such design choices may be limited, at least in some situations. Of course, this does not show that these factors never affect risk taking. Further research is needed to identify the task characteristics that determine when such design choices are reasonable and when they may be problematic.

References

Agnew, J. R., Anderson, L. R., Gerlach, J. R., and Szykman, L. R. 2008. Who chooses annuities? an experimental investigation of the role of gender, framing, and defaults. The American Economic Review, 98, 418-422.

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. 2006. Dynamic choice behavior in a natural experiment. University of Central Florida Department of Economics: Working Paper, 06-10.

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. 2014. Dual criteria decisions. Journal of Economic Psychology, 41, 101-113.

Baltussen, G., Post, G. T., Van Den Assem, M. J., and Wakker, P. P. 2012. Random incentive systems in a dynamic choice experiment. Experimental Economics, 15, 418-443.

Baltussen, G., Post, T., and Van den Assem, M. J. 2008. Risky choice and the relative size of stakes. Available at SSRN 989242.

Benjamin, D. J., Brown, S. A., and Shapiro, J. M. 2013. Who is 'behavioral'? Cognitive ability and anomalous preferences. Journal of the European Economic Association, 11, 1231-1255.

Blavatskyy, P., and Pogrebna, G. 2008. Risk aversion when gains are likely and unlikely: Evidence from a natural experiment with large stakes. Theory and Decision, 64, 395-420.

Borghans, L., Heckman, J. J., Golsteyn, B. H., and Meijers, H. 2009. Gender differences in risk aversion and ambiguity aversion. Journal of the European Economic Association, 7, 649-658.

Brinig, M. F. 1995. Does mediation systematically disadvantage women. William and Mary Journal of Women and the Law, 2, 1-34.

Burks, S. V., Carpenter, J. P., Goette, L., and Rustichini, A. 2009. Cognitive skills affect economic preferences, strategic behavior, and job attachment. Proceedings of the National Academy of Sciences, 106, 7745-7750.

Charness, G., and Gneezy, U. 2004. Gender, Framing, and Investment. mimeo.

Charness, G., and Gneezy, U. 2010. Portfolio choice and risk attitudes: An experiment. Economic Inquiry, 48, 133-146.

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

Cox, J. C., Sadiraj, V., and Schmidt, U. 2015. Paradoxes and mechanisms for choice under risk. Experimental Economics, 18, 215-250.

Cubitt, R. P., Starmer, C., and Sugden, R. 1998. On the validity of the random lottery incentive system. Experimental Economics, 1, 115-131.

De Roos, N., and Sarafidis, Y. 2010. Decision making under risk in Deal or No Deal. Journal of Applied Econometrics, 25, 987-1027.

Deck, C., and Jahedi, S. 2015. The effect of cognitive load on economic decision making: A survey and new experiments. European Economic Review, 78, 97-119.

Deck, C., Lee, J. and Reyes, J. 2008. Risk attitudes in large stake gambles: evidence from a game show. Applied Economics, 40, 41-52.

Deck, C., Lee, J., Reyes, J. A., and Rosen, C. C. 2013. A failed attempt to explain within subject variation in risk taking behavior using domain specific risk attitudes. Journal of Economic Behavior and Organization, 87, 1-24.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. 2010. Are Risk Aversion and Impatience Related to Cognitive Ability?. American Economic Review, 100, 1238-60.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. Journal of the European Economic Association, 9, 522-550.

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

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

Falk, A., and Heckman, J. 2009. Lab Experiments Are a Major Source of Knowledge in the Social Sciences. Science, 326, 535-538.

Fehr-Duda, H., De Gennaro, M., and Schubert, R. 2006. Gender, financial risk, and probability weights. Theory and decision, 60, 283-313.

Filippin, A. and Crosetto, P. 2014. A reconsideration of gender differences in risk attitudes. Grenoble Applied Economics Laboratory: Working Papers.

Frederick, S. 2005. Cognitive reflection and decision making. Journal of Economic perspectives, 19, 25-42.

Gerhardt, H. 2013. Essays in experimental and neuroeconomics. Humboldt-Universität zu Berlin: Doctoral dissertation.

Harbaugh, W. T., Krause, K., and Vesterlund, L. 2002. Risk attitudes of children and adults: Choices over small and large probability gains and losses. Experimental Economics, 5, 53-84.

Harrison, G. W. and Swarthout, J. T. 2014. Experimental payment protocols and the bipolar behaviorist. Theory and Decision, 77, 423-438.

Harrison, G. W., Lau, M. I., and Rutström, E. E. 2007. Estimating risk attitudes in Denmark: A field experiment. The Scandinavian Journal of Economics, 109, 341-368.

Holt, C. and Laury, S. 2002. Risk Aversion and Incentive Effects, American Economic Review, 92, 1644-1655.

Hoppe, E. I. and Kusterer, D. J. 2011. Behavioral biases and cognitive reflection. Economics Letters, 110, 97-100.

Kahneman, D. 2003. Maps of bounded rationality: Psychology for behavioral economics. American economic review 93(5): 1449-1475.

Langer, T., and Weber, M. 2008. Does commitment or feedback influence myopic loss aversion? An experimental analysis. Journal of Economic Behavior and Organization, 67, 810–819.

List, J.A. 2006. The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transaction, Journal of Political Economy, 114, 1-37.

March, C., Ziegelmeyer, A., Greiner, B., and Cyranek, R. 2014. Monetary Incentives in Large-Scale Experiments: A Case Study of Risk Aversion. Unpublished.

Mata, R., Josef, A. K., Samanez-Larkin, G. R., and Hertwig, R. 2011. Age differences in risky choice: A meta-analysis. Annals of the New York Academy of Sciences, 1235, 18-29.

Oechssler, J., Roider, A. and Schmitz, P. W. 2009. Cognitive abilities and behavioral biases. Journal of Economic Behavior and Organization, 72, 147–152.

Post, T., van den Assem, M.J., Baltussen, G. and Thaler, R.H. 2008. Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show, The American Economic Review, 98, 38-71.

Powell, M., and Ansic, D. 1997. Gender differences in risk behaviour in financial decision-making: An experimental analysis. Journal of economic psychology, 18, 605-628.

Schubert, R., Brown, M., Gysler, M., and Brachinger, H. W. 1999. Financial decision-making: are women really more risk-averse?. American Economic Review, 82, 381-385.

Schubert, R., Brown, M., Gysler, M., and Brachinger, H. W. 2000. Gender specific attitudes towards risk and ambiguity: an experimental investigation. Institut für Wirtschaftsforschung, Eidgenössische Technische Hochschule: Working paper.

Starmer, C., and Sugden, R. 1991. Does the random-lottery incentive system elicit true preferences? An experimental investigation. The American Economic Review, 81, 971-978.

Tversky, A., and Kahneman, D. 1981. The framing of decisions and the psychology of choice. Science, 211, 453–458.

Tymula, A., Belmaker, L. A. R., Ruderman, L., Glimcher, P. W., and Levy, I. 2013. Like cognitive function, decision making across the life span shows profound age-related changes. Proceedings of the National Academy of Sciences, 110, 17143-17148.

Whitney, P., Rinehart, C. A., and Hinson, J. M. 2008. Framing effects under cognitive load: The role of working memory in risky decisions. Psychonomic Bulletin and Review, 15, 1179-1184.