Journal of Economic Behavior & Organization 87 (2013) 43–51

Contents lists available at SciVerse ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Experimental methods: Eliciting risk preferences

Gary Charness[a,*], Uri Gneezy[b,c], Alex Imas[b]

[a] University of California, Santa Barbara, United States
[b] University of California, San Diego, United States
[c] CREED, University of Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 23 May 2012
Received in revised form 7 December 2012
Accepted 13 December 2012
Available online 31 December 2012

JEL classification:
C91
D03
D81

Keywords:
Risk preferences
Elicitation methods
Experimental methodology

ABSTRACT

Economists and psychologists have developed a variety of experimental methodologies to elicit and assess individual risk attitudes. Choosing which to utilize, however, is largely dependent on the question one wants to answer, as well as the characteristics of the sample population. The goal of this paper is to present a series of prevailing methods for eliciting risk preferences and outline the advantages and disadvantages of each. We do not attempt to give a comprehensive account of all the methods or nuances of measuring risk, but rather to outline some advantages and disadvantages of different methods.

Risk is ubiquitous in decision-making. The extent to which people are willing to take on risk constitutes their *risk preferences.* Assessing and measuring the risk preferences of individuals is critical for economic analysis and policy prescriptions. Hallmark theories of risk and uncertainty such as subjective expected utility (Savage, 1954) and prospect theory (Kahneman and Tversky, 1979) leave risk attitudes as a free parameter, and individuals may have differing attitudes toward risk.

For example, when given a chance to purchase a lottery ticket with equal chances of winning either $10 or $0, a risk-neutral individual will be willing to pay up to $5—the expected value of the lottery. Individuals who are only willing to pay less than $5 are considered to be risk-averse, while those willing to pay more are considered to be risk-seeking. Any of these risk attitudes is consistent with all theories, given some risk-attitude parameters.

Economists and psychologists have developed a variety of experimental methodologies to elicit and assess individual risk attitudes. Choosing which to utilize, however, is largely dependent on the question one wants to answer, as well as the characteristics of the sample population. The goal of this paper is to present a series of prevailing methods for eliciting risk preferences and outline the advantages and disadvantages of each. We do not attempt to give a comprehensive account of all the methods or nuances of measuring risk, but rather to outline some advantages and disadvantages of different methods. For a more comprehensive discussion (with different conclusions), see for example Cox and Harrison (2008).

We characterize elicitation methods according to their complexity. More complex methods are typically used for estimating parameters corresponding to risk preferences of a model that makes particular functional form assumptions. These parameters can then be used as evidence for or against particular theories of decision-making. Complex methods also demand

* Corresponding author.
E-mail address: charness@econ.ucsb.edu (G. Charness).

0167-2681/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jebo.2012.12.023

more understanding and mathematical sophistication from the subjects, or else comprehension suffers and the results may be less meaningful.

Simple elicitation methods tend to be substantially easier for participants to understand. For example, the Balloon Analogue Risk Task (Lejuez et al., 2002) presents subjects with a sequence of choices of whether or not to gain additional money by pumping more air into a balloon, each pump coming with the risk of losing the accumulated gains if the balloon pops. Other simple elicitation methods include a single choice of how much to allocate between a safe and risky asset (Gneezy and Potters, 1997; Charness and Gneezy, 2012), a single choice between gambles (Eckel and Grossman, 2002; Dave et al., 2010), and non-incentivized questionnaires (Weber et al., 2002; Dohmen et al., 2011). Simple methods are most useful when trying to capture treatment effects and differences in individual risk preferences.

We wish to emphasize that we are not suggesting that one method is better than another. Rather, we simply highlight the costs and benefits of different methods, urging researchers to take these into account when choosing the one that best fits their needs. Elicitation approaches should be tailored to the underlying question being asked, with attention paid to the abilities of the participants to understand the questions.

We now proceed to examine specific elicitation techniques that have been used in the literature, discussing the relative advantages and disadvantages of each.

1. Elicitation methods

1.1. The Balloon Analogue Risk Task (BART)

The Balloon Analogue Risk Task (BART) measures risk preferences by presenting individuals with a computer simulation of pumping air into a series of balloons (Lejuez et al., 2002). Balloons of three different colors (blue, yellow and orange) are presented one at a time. For each successive pump, the balloon grows in size and the individual earns money that is deposited into a temporary reserve. The value of the reserve is never revealed to the participant. As the balloon becomes bigger, the chances that it would pop after another pump grows as well; the probability of popping is negligible before the first pump and grows to certainty after the balloon reaches a particular size. If the balloon pops, all earnings in the temporary reserve disappear and a new balloon appears. At any given time, the participant can either pump the balloon or collect what she has earned so far. If the participant chooses to collect her earnings, that money is deposited into her permanent account and a new balloon appears. She then faces the same scenario with the next balloon.

The BART is a good example of a class of methods that aims to elicit risk preferences in a context that is familiar and easy to grasp (see also Crosetto and Filippin, 2012). It has also been used to study risk attitudes across a variety of subfields such as neuroscience (Fecteau et al., 2007), drug addiction (Bornovalova et al., 2005) and psychopathology (Hunt et al., 2005).

The probability of popping increases monotonically with each successive pump and evolves according to a function specific to the color of the balloon. The blue balloon, for example, has a 1/128 probability of popping after the first pump, 1/127 after the second, 1/126 after the third, and so on. After 128 pumps, the balloon is guaranteed to pop. It takes 8 and 32 pumps respectively to pop the orange and yellow balloons with certainty. Participants are told that the balloons would explode after enough pumps, but the actual probability function is never revealed. This procedure was designed to model situations where excessive risk-taking leads to diminishing returns and greater hazards. Particularly, with each successive pump, the amount that could potentially be lost increases and the overall expected gain decreases. Individuals are presented with 90 balloons in total, with the colors randomized accordingly.

Since each successive pump carried an increased risk of causing the balloon to pop, the authors took the average number of pumps, excluding balloons that exploded, to be the adjusted value corresponding to the individual's risk preference. This value correlated significantly with reported real-world risky behavior such as gambling, drug use and unprotected sex, as well as self-reported risk constructs that measured impulsiveness and sensation seeking. Lejuez et al. (2003) found that risk preferences elicited using the BART correlated significantly with self-reported risk-taking in daily life in a population of high school students, as measured by the CDC Youth Risk Behavior Surveillance System.

However, it is not clear if risk preferences elicited through these methods extend to other domains, in particular financial decision-making (which is of particular interest to economists), or if they are associated with risk preferences elicited through other methods (Buelow and Suhr, 2009). In addition, the BART requires a computer and multiple trials to implement. Hence, it may not be ideal for being paired with other decision-making tasks, when time is a factor, or for field work where access to computers is limited.

1.2. Questionnaires

Questionnaires are a commonly used method of eliciting risk preferences that rely on the individual's self-reported propensity for risk. A typical general risk question comes in the form of: "*Rate your willingness to take risks in general*" on a 10-point scale, with 1-*completely unwilling* and 10-*completely willing*.

Such general risk questions implicitly assume that they are measuring a single, stable risk preference that influences behavior across various domains. In turn, risk preferences derived through this method are commonly used as indicators for the propensity to engage in behavior ranging from portfolio selection to smoking. However, a substantial amount of evidence suggests that the measured risk preferences are highly dependent on the domains in which they are elicited. The risk attitudes of company managers, for example, appear to differ substantially depending on whether risk was in the recreational or financial domain (Maccrimmon and Wehrung, 1990). Dreber et al. (2011) show that the propensity for risk amongst female tournament bridge players differed substantially between the domains of bridge and financial decision-making.

To capture this variation, Weber et al. (2002) developed the *domain-specific risk-taking* (DOSPERT) scale. The full DOSPERT scale contains 40 items: eight items in the domains of recreational, health, social, and ethical risks, and four items in the domains of gambling and investment. Each item asks the subject to rate on a 5-point scale how likely she would be to engage in a particular behavior such as “Drinking heavily at a social function” or “Gambling a week’s income at a casino.” Using homogenous, non-standard subsamples such as sky-divers, smokers, and gamblers, Hanoch et al. (2006) used the DOSPERT to demonstrate the domain-specific nature of risk preferences. For example, individuals who engaged in recreational risks had high risk-taking scores in that domain but did not necessarily score high in other domains such as finance.

In a large study of the German population combining survey data and field experiments, Dohmen et al. (2011) examined the association between risk preferences elicited through different methods and how well they predicted individual behavior. The authors found that the general risk question did fairly well in predicting risk-taking behavior in a field experiment with real monetary stakes. The association between the elicited risk preferences and self-reported risk-taking behavior in domains such as job choices and portfolio selection was also examined. Although the general risk question had some predictive power across domains—performing better than more complex methods—the best predictor of behavior in a particular domain was the corresponding domain-specific measure elicited through a method akin to the DOSPERT scale.

Charness and Viceisza (2011) used the general risk question to examine risk preferences of individuals in rural Senegal. They found that participants, particularly women, reported significantly greater tolerance for risk than typical experimental subjects in the western world (including the Dohmen et al. (2011) study).

While simple to understand, questionnaires are typically not directly incentivized. Hence, it is up for debate whether the elicited risk preferences reflect an individual’s true attitudes toward risk, particularly in the domain of financial decision-making. It can be argued that if measures of risk preferences are to be associated with actual risk-taking behavior, their elicitation should be incentivized in order to ensure that choices reflect true underlying attitudes toward risk. There is some strong evidence for the importance of incentive-compatibility in risk elicitation (e.g., Holt and Laury, 2002). In turn, when choosing to use questionnaires, researchers should be aware of the tradeoff between the simplicity of the method and the possibility of gratuitously-expressed preferences for risk.

1.3. The Gneezy and Potters method

The elicitation method of Gneezy and Potters (1997) provides a measure of risk preferences in the context of financial decision-making with real monetary payoffs. Here, the decision maker receives \$$X$ and is asked to choose how much of it, \$$x$, she wishes to invest in a risky option and how much to keep. The amount invested yields a dividend of \$$kx$ ($k > 1$) with probability p and is lost with probability $1 - p$. The money not invested \$$(X - x)$ is kept by the investor. The payoffs are then \$$(X - x + kx)$ with probability p, and \$$(X - x)$ with $1 - p$. In all cases, p and k are chosen so that $p \times k > 1$, making the expected value of investing higher than the expected value of not investing; thus, a risk-neutral (or risk-seeking) person should invest \$$X$, while a risk-averse person may invest less. The choice of x is the only decision the participants make in the experiment (see Charness and Gneezy, 2012 for a survey).

For example, consider the case in which the participant receives an endowment of 100 cents. She is then asked to choose what part of this endowment (x) she would like to invest in a risky asset and how much to keep. The risky asset returns 2.5 times the amount invested with a probability of one-half and nothing with a probability of one-half. The participant keeps the money that she does not invest ($100 - x$). The amount invested is then used as the measure of risk preferences.

Note that for these parameters, risk-neutral (and, in turn, risk-seeking) individuals should invest their entire endowment. Hence, a disadvantage of this method is that it cannot distinguish between risk-seeking and risk-neutral preferences. However, since risk-seeking preferences appear to be relatively uncommon, and a fairly small fraction of participants choose to invest the entire amount of points, the amount invested x provides a good metric for capturing treatment effects and differences in attitude toward risk between individuals.

This elicitation method has been used to provide support for myopic loss aversion in the financial decisions of students (Gneezy and Potters, 1997), as well as professional traders (Haigh and List, 2005). The method has also been used to show a positive correlation between risk taking, testosterone levels, and facial masculinity (Apicella et al., 2008), and to compare gender differences in risk attitudes (Charness and Gneezy, 2012).

The relative simplicity of the method, combined with the fact that it can be implemented with one trial and basic experimental tools, makes it a useful instrument for assessing risk preferences in the field. Gneezy et al. (2009) used the method to examine gender differences in risk preferences between a patriarchal society in Tanzania and a matriarchal society in India. Dreber et al. (2011) used it to look at the risk preferences of bridge players at a professional championship. In eliciting the risk preferences of people in rural Senegal, the method produced sensible data in line with previous findings, while a more complex elicitation method produced largely inconsistent results (Charness and Viceisza, 2011).

1.4. The Eckel and Grossman method

The method developed by Eckel and Grossman (2002), was explicitly designed to be a simple way of eliciting risk preferences that produced enough heterogeneity in choices to allow for the estimation of utility parameters. The method asks

Table 1
The Eckel and Grossman measure.

Choice (50/50 Gamble)	Low payoff	High payoff	Expected return	Standard deviation	Implied CRRA range
Gamble 1	28	28	28	0	$3.46 < r$
Gamble 2	24	36	30	6	$1.16 < r < 3.46$
Gamble 3	20	44	32	12	$0.71 < r < 1.16$
Gamble 4	16	52	34	18	$0.50 < r < 0.71$
Gamble 5	12	60	36	24	$0 < r < 0.50$
Gamble 6	2	70	36	34	$R < 0$

subjects to make only one choice; participants are presented with a number of gambles and are asked to choose one that they would like to play. The number of presented gambles can be varied. For example, Dave et al. (2010) presented participants with six gambles. Each of the gambles, listed in Table 1, involves a 50% chance of receiving the low payoff and a 50% chance of the high payoff. One of the gambles is a sure thing: in this case, Gamble 1 with a certain payoff of $28. For Gambles 1–5, the expected payoff increases linearly with risk, as represented by the standard deviation. Note that Gamble 6 has the same expected payoff as Gamble 5 but with a higher standard deviation. The gambles are designed so that risk-averse subjects should choose those with a lower standard deviation (Gambles 1–4), risk-neutral subjects should choose the gamble with the higher expected return (Gamble 5), and risk-seeking subjects should choose Gamble 6.

This method allows for parameter estimation: the chosen gamble implies an interval for the risk coefficient under the assumption of constant relative risk aversion (CRRA). Under this assumption, utility can be represented by the function $u(x) = x^{1-r}$, with r corresponding to the coefficient of relative risk aversion and x corresponding to wealth. Individuals with $r > 0$ can be classified as risk averse, $r < 0$ as risk loving and $r = 0$ as risk neutral. Table 1 contains intervals for the risk coefficient corresponding to each chosen gamble. The intervals are determined by calculating the value of r that would make the individual indifferent between the gamble she chose and the two adjacent gambles. For example, a choice of Gamble 3 implies a risk coefficient in the interval of (0.71, 1.16): indifference between Gambles 3 and 4 corresponds to $r = 0.71$, and indifference between Gambles 2 and 3 to $r = 1.16$.

This measure has been used in Eckel and Grossman (2008) to demonstrate that women are significantly more risk averse than men. The authors also examined the stereotyping of risk attitudes by asking subjects to guess the gamble choice of others and found that both men and women predicted greater risk aversion for women. In a field experiment with French farmers, Reynaud and Couture (2012) compared several risk elicitation methods and found the measure elicited using the Eckel and Grossman method correlated significantly with those elicited through the other methods. In addition, Dave et al. (2010) demonstrated that the method produced significantly less noisy estimates of risk preferences than more complex elicitation methods, particularly when participants had lower math abilities.

The measure is relatively easy for individuals to understand. However, it cannot differentiate between different degrees of risk-seeking behavior.

2. The multiple price list method

The methods discussed so far share the advantage of being relatively easy to understand and implement, and are good in capturing differences in individual risk preferences and identifying treatment effects. Risk preferences elicited through these methods have also been shown to correlate with other individual characteristics and real world risk-taking behavior.

More complex methods often elicit risk preferences by presenting subjects with a series of choices between gambles. Given assumptions of a particular functional form and fine enough choice data over gambles, it is possible to estimate increasingly precise intervals for the curvature parameters of a utility function or to use econometric methods to obtain point values of those estimates.

One of the most commonly used elicitation methods in economics represents choices between gambles as multiple price lists (MPL). An early incentivized use of the MPL method can be found in Binswanger (1981), who elicited risk preferences of farmers in rural India. Other researchers have gone on to use the MPL method to price commodities (Kahneman et al., 1990) and elicit discount rates (Coller and Williams, 1999); see Andersen et al. (2006) for a more complete review of the MPL literature.

In an influential paper Holt and Laury (2002) popularized the MPL, using the method to estimate risk parameters of a utility function; in fact, many researchers refer to the MPL method as the Holt–Laury measure of risk aversion. The prevalent use of the Holt–Laury measure has allowed researchers to compare risk attitudes across a wide array of contexts and environments. In turn, this has facilitated a less fragmented approach to the study of risk preferences that minimizes methodological differences and aims to characterize a more general phenomenon.

When using this method, a participant is typically presented with a list of 10 decisions between paired gambles, as in Table 2. The two gambles for each decision are stacked in rows, with gambles in the left and right columns labeled Option A and Option B, respectively. The participant then chooses which gamble she prefers to play from each pair by picking either Option A or B, making this choice for every decision row. The payoffs of gambles in Option A and Option B remain constant; the only thing that changes between decision rows is the probability associated with each payoff. For the first decision row, there is only a 1/10 chance of getting the high payoff for either option, and the expected payoff of Option A is $1.17 greater

Table 2
MPL method.

Option A	Option B	Option A	Option B
1/10 of $2, 9/10 of $1.60	1/10 of $3.85, 9/10 of $0.10	☐	☐
2/10 of $2, 8/10 of $1.60	2/10 of $3.85, 8/10 of $0.10	☐	☐
3/10 of $2, 7/10 of $1.60	3/10 of $3.85, 7/10 of $0.10	☐	☐
4/10 of $2, 6/10 of $1.60	4/10 of $3.85, 6/10 of $0.10	☐	☐
5/10 of $2, 5/10 of $1.60	5/10 of $3.85, 5/10 of $0.10	☐	☐
6/10 of $2, 4/10 of $1.60	6/10 of $3.85, 4/10 of $0.10	☐	☐
7/10 of $2, 3/10 of $1.60	7/10 of $3.85, 3/10 of $0.10	☐	☐
8/10 of $2, 2/10 of $1.60	8/10 of $3.85, 2/10 of $0.10	☐	☐
9/10 of $2, 1/10 of $1.60	9/10 of $3.85, 1/10 of $0.10	☐	☐
10/10 of $2, 0/10 of $1.60	10/10 of $3.85, 0/10 of $0.10	☐	☐

From Holt and Laury (2002).

than of Option B. Hence, only a very risk-seeking individual would choose Option B in the first decision row. Moving down the rows, the probability of the high payoff increases, and by the last decision the choice is between $2.00 and $3.85 with certainty. If the individual understands the instructions and prefers more money to less, she should choose Option B for Decision 10. For all but the most risk-seeking, this implies a pattern where individuals start by choosing Option A for the first decision and switch over to Option B at some point before the last decision row. This switch point is then used as the measure of the individual's risk preference.

This method can be used to estimate coefficients of risk preferences provided assumptions on the functional form of utility. Under an assumption of CRRA, payoffs are constructed such that the gamble pair at which the individual switched over from Option A to Option B, termed the "switch point," can be used to provide an estimated interval for r. For example, choosing Option A for the first four decisions and Option B for the rest would be consistent with risk neutrality—an r in the interval (−0.15, 0.15). Choosing the safe Option A for the first six decisions and Option B for the rest would indicate risk aversion with the coefficient r in the interval (0.41, 0.68).

When implementing the MPL method, participants are typically informed that after all decisions are made, one decision will be selected at random and the chosen gamble will be played for real. Subjects are then paid according to that outcome. If she believes that her choices have no effect on what decision row is chosen for payment, then the MPL represents a series of binary choices and incentivizes the subject to reveal her preferences truthfully. Azrieli et al. (2012) demonstrate theoretically that selecting one decision for payment at random is the only incentive-compatible way to utilize the MPL method, and Laury (2006) found that there is no significant difference between choice behavior when all decision rows are implemented or one is chosen at random.

The MPL method has been used to evaluate risk preferences in a variety of contexts. Dohmen et al. (2010) used the MPL method to demonstrate that lower cognitive ability is associated with greater risk aversion, controlling for demographic factors. Multiple MPLs have also been used to jointly estimate time discount factors and risk coefficients for particular models of choice. Andersen et al. (2008) presented subjects with two sets of MPLs: one set was similar to the one in Table 2 and was used to elicit risk preferences, the other set was used to elicit time preferences. This technique, known as the double multiple price list (DMPL), was used to jointly estimate parameters corresponding to the risk coefficient and the discount rate under the assumptions of a CRRA utility function that is separable and stationary over time.

In a study of Vietnamese villagers, Tanaka et al. (2010) constructed modified multiple price lists to elicit risk and time preferences, using the data to estimate a total of six parameters. Making strict functional form assumptions, the authors estimated three parameters of a prospect theory model (corresponding to concavity, loss aversion, and weighting function parameters) and three parameters of a general time discounting model (corresponding to time discounting, present-bias, and hyperbolicity parameters). The results indicated that mean village income was correlated with risk preferences, and that individuals in poorer villages were more averse to loss but not necessarily to income variation. Household income, however, was not correlated with risk preferences.

One of the main disadvantages of complex methods of eliciting risk preferences is that, depending on the population, a significant number of subjects will fail to understand the procedure. This reduces the reliability of the risk preference measure and can potentially bias the results. With the standard MPL method, individuals are typically allowed to switch freely between Options A and B as they progress down the decision rows. As such, participants may make inconsistent decisions either by switching more than once or making "backwards" choices—starting with Option A and switching to B (Dave et al., 2010; Holt and Laury, 2002). Depending on the participant pool, this problem could be significant. For example, Jacobson and Petrie (2009) found that in a sample of Rwandan adults, 55% made inconsistent choices; Charness and Viceisza (2011) found that 75% of farmers from rural Senegal made inconsistent choices (51% switched more than once and 24% always chose Option A). This poses an obvious problem since the inference of risk preferences, and in turn, parameter estimation requires a unique switch point and such inconsistent behavior is difficult to rationalize under standard assumptions on preferences.

Andersen et al. (2006) argue that multiple switch points may reflect indifference between the choices since multiple price lists typically do not include an explicit option to signify indifference. Other researchers view such behavior as indicative of

Table 3
Methods of eliciting risk preferences.

Method	Simple or complex	Used in	Paid	Summary	Population	Examined gender difference	Estimated parameters
The Balloon Analogue Risk Task (BART)	Simple	Lejuez et al. (2002)	Yes	Correlated with self-reported risky behavior such as gambling and drug use, as well as measures of impulsivity and sensation seeking	Students	M > F	No
		Fecteau et al. (2007)	Yes	Stimulation to the dorsolateral prefrontal cortex (DLPFC) decreases risk-taking behavior	Students	M = F	No
		Bornovalova et al. (2005)	Yes	Crack-cocaine users less risk averse than heroin users	Drug users	No	No
		Hunt et al. (2005)	Yes	Higher self-reported psychopathy correlated with greater risk taking. Men were less risk-averse than women	Students	M > F	No
Questionnaires	Simple	Maccrimmon and Wehrung (1990)	No	Measured risk attitudes across domains. Success of business executives negatively correlated with risk aversion in domain of financial risk-taking. Little association between domains	Business executives	No	No
		Weber et al. (2002)	No	Developed domain specific scale (DOSPERT) to measure risk attitudes. Found risk attitudes were domain specific	Students	M > F	No
		Hanoch et al. (2006)	No	Used DOSPERT to show little association in risk attitudes across domains using non-standard population	Targeted subsamples	M = F	No
The Gneezy and Potters method	Simple	Gneezy and Potters (1997)	Yes	Provided support for myopic loss aversion	Students	No	No
		Charness and Gneezy (2012)	Yes	Subjects preferred to receive information on investment outcomes more frequently	Students	M > F	No
		Haigh and List (2005)	Yes	Demonstrated professional traders exhibited greater myopic loss aversion than students	Professional traders and students	No	No
		Apicella et al. (2008)	Yes	Risk taking correlates positively with testosterone and facial masculinity	Students	No	No
		Gneezy et al. (2009)	Yes	No gender difference in risk attitudes in either subpopulation	Members of the Maasai and the Khasi tribes	M = F	No
		Dreber et al. (2011)	Yes	Showed a difference in risk attitudes in domain of bridge and financial decision-making	Professional bridge players	M > F	No
The Eckel and Grossman method	Simple	Eckel and Grossman (2002); Eckel and Grossman (2008)	Yes	Subjects predicted gender difference in risk aversion	Students	M > F	Yes
		Eckel et al. (2011)	Yes	Peer and school quality effects associated with risk preferences	9th and 11th graders	M > F	Yes
The multiple price list method	Complex	Binswanger (1981)	Yes	Elicited risk preferences did not correlate with most demographic variables; did not exhibit predictive power for actual risk-taking behavior (adoption of novel farming techniques)	Farmers in India	M = F	Yes
		Holt and Laury (2002)	Yes	Found that subjects exhibited substantially greater risk aversion at higher stakes	Students	M > F at low stakes; M = F at high stakes	Yes
		Laury (2006)	Yes	No significant difference found when risk preferences are elicited using the random-choice payment method or subjects are paid for every decision	Students	M = F	Yes

		Dohmen et al. (2010)	Yes	Higher cognitive ability associated with less risk aversion	Representative sample of German population	M = F	Yes
		Andersen et al. (2008)	Yes	Used double multiple price list (DMPL) method to estimate risk and time preferences jointly. Estimated discount rates more in line with market rates	Representative sample of Danish population	M = F	Yes
		Tanaka et al. (2010)	Yes	Village income but not household income correlated with elicited risk preferences	Vietnamese villagers	M = F	Yes
		Jacobson and Petrie (2009)	Yes	More than half of subjects made at least one inconsistent choice. Mistakes associated with sub-optimal decision making in other domains	Representative sample of Rwandan adults	M = F	Yes
		Andersen et al. (2006)	Yes	Parameters estimated using the MPL method found to be sensitive to procedures, subject pools and format. New forms of the MPL method are proposed and tested	Students	M = F	Yes
		von Gaudecker et al. (2011)	Yes	Elicited risk preferences correlated with age and education	Representative sample of Dutch population	M > F	Yes
		Harrison et al. (2007)	Yes	Average subject was risk averse. Elicited risk preferences associated with age and education	Representative sample of Danish population	M = F	Yes
Multiple methods	Simple and Complex	Dohmen et al. (2011)	Real and hypo	Domain-specific questions best predictors of self-reported risky behavior in respective domain; general risk question predicted behavior across all domains. MPL measure had little predictive power	Representative sample of German population	M > F	No
		Charness and Viceisza (2011)	Real and hypo	Results from MPL and general questionnaire suggested significant confusion. Results from Gneezy and Potters method largely consistent with previous findings	Villagers in rural Senegal	M = F	No
		Anderson and Mellor (2009)	Real and hypo	No correlation between risk preferences elicited using questionnaires and MPL method	Students	M > F	Yes
		Lonnqvist et al. (2011)	Real and hypo	No correlation between risk preferences elicited using questionnaires and MPL method. Questionnaire data predicted actual risk-taking behavior; MPL data did not	Students	M = F	No
		Reynaud and Couture (2012)	Yes	Find significant relationship between risk preferences elicited using MPL and Eckel and Grossman methods	French farmers	No	Yes
		Dave et al. (2010)	Yes	Compare risk preferences elicited using MPL and Eckel and Grossman methods. MPL method produced noisier data, particularly in those with low math ability	Subgroups of Canadian residents	M > F	Yes

confusion on the part of the participants and typically remove this data from the analysis. Depending on what proportion of individuals make inconsistent choices, this may significantly hinder the usability of this method.

Several techniques have been proposed to address the issue of multiple switch points. Rather than asking individuals which of the two gambles they prefer for each decision row, Andersen et al. (2006) ask at which decision row they would like to switch from Option A to B—guaranteeing a unique switch point by design. Tanaka et al. (2010) use a similar procedure of enforcing unique switch points to examine the preferences of Vietnamese villagers. This method imposes strict monotonicity on revealed preferences and enforces transitivity.

While enforcing consistent choices, this technique may significantly bias the results. If inconsistent choice data is treated as noise and dropped, it can be said with some confidence that the individuals who are left understood the instructions and are revealing their true preferences. Enforcing a single switch point not only imposes added assumptions on preferences (which may or may not hold), but confused individuals who would have made inconsistent choices if they were allowed to switch freely are now included in the sample. This forces noise into the data, which may bias the results and, in turn, the estimated preferences.

In addition, while parameter estimation is one of the major advantages of complex elicitation methods, it is not clear whether this estimated level of risk aversion could be used to predict behavior in other domains. Parameter estimation requires substantial functional form assumptions, and when those assumptions are relaxed or changed, the estimated parameters change as well. For example, using the DMPL method, Andersen et al. (2008) estimated a risk coefficient corresponding to significant concavity of the utility function and demonstrated that using this parameter under CRRA and time-separability assumptions leads to estimates of discount rates that are significantly lower than in previous studies that assumed linear utility. Using a different protocol, Andreoni and Sprenger (2012) found substantially less concavity, while Tanaka et al. (2010) found more (albeit in a different population) when making different functional form assumptions.

Friedman and Sunder (2011) selectively survey 60 years of empirical research on the estimation of risk preferences under the assumption that individuals maximize some specific functional form of utility. The authors find that thus far, this approach has produced limited applicable results for out-of-sample predictions since the estimates are often unstable and highly context-dependent. Given the sensitivity of estimated parameters to the protocol used and the assumptions made, researchers should be cautious in using these parameters to forecast behavior in other domains and extrapolate policy implications.

3. Discussion

Table 3 provides a summary, albeit an incomplete one, of some of the more prevalent methods used to elicit risk preferences. These methods have provided researchers with the methodological tools to study risk attitudes in a variety of environments, using a wide array of sample populations. As discussed in the previous sections, choosing which method to use largely depends on the context as well as the question being asked. For example, the BART method may work well in the lab with a student population, but would be less feasible to implement in the field where access to computers is limited.

More broadly, the methods explored above highlight an important issue regarding the advantages and limitations of using experiments to elicit risk preferences. Given the level of control over the design and incentive structure, experiments are best at identifying treatment effects and differences between individuals—saying that one individual is more risk averse than another—as well as testing economic theory. The Gneezy and Potters method, for example, was effective in providing support for myopic loss aversion.

Given their stylized nature, however, experiments may not be ideal for identifying levels of risk aversion, to the extent that the estimated risk preferences and parameters may not have predictive power for behavior across domains. For example, Dohmen et al. (2011) found that a standard lottery measure that allows for the estimation of risk coefficients had little predictive power for real world, risky behavior in such domains as employment, car driving and health. Similarly, in a large-scale field study of truck drivers, Anderson et al. (2011) found that risk preferences elicited through MPL had little predictive power for economic variables like credit scores and job persistence, or non-economic behavior such as driving accidents and smoking.

In view of the limited predictive power of experimentally-elicited risk preferences across domains, sometimes the advantages of complex methods may well be outweighed by the disadvantage of decreased comprehension, and consequently, producing substantially noisier data. Simpler methods have the advantage of being more straightforward and capable of eliciting sensible risk preferences from a broader set of individuals, but similar care should be taken in extrapolating from these measures to other domains. Given the advantages and disadvantages of the methods discussed above, researchers should choose the one that is best suited for the particular question being asked, as well as the sample population being used.

References

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.

Anderson, J., Burks, S., DeYoung, C., Rustichini, A., 2011. Toward the integration of personality theory and decision theory in the explanation of economic behavior. mimeo.

Anderson, L.R., Mellor, J.M., 2009. Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure. Journal of Risk and Uncertainty 39 (2), 137–160.
Andreoni, J., Sprenger, C., 2012. Estimating time preferences from convext budgets. American Economic Review 102 (7), 3333–3356.
Apicella, C., Dreber, a., Campbell, B., Gray, P., Hoffman, M., Little, A., 2008. Testosterone and financial risk preferences. Evolution and Human Behavior 29 (6), 384–390.
Azrieli, Y., Chambers, C.P., Healy, P.J., 2012. Incentives in experiments: a theoretical analysis. mimeo.
Binswanger, H.P., 1981. Attitudes toward risk: theoretical implications of an experiment in rural India. Economic Journal 91 (364), 867–890.
Bornovalova, M.A., Daughters, S.B., Hernandez, G.D., Richards, J.B., Lejuez, C.W., 2005. Differences in impulsivity and risk-taking propensity between primary users of crack cocaine and primary users of heroin in a residential substance-use program. Experimental and Clinical Psychopharmacology 13 (4), 311–318.
Buelow, M.T., Suhr, J.A., 2009. Construct validity of the Iowa gambling task. Neuropsychology Review 19 (1), 102–114.
Crosetto, P., Filippin, A., 2012. The Bomb risk elicitation task. mimeo.
Cox, J.C., Harrison, G.W., 2008. Risk Aversion in Experiments. Emerald Group Publishing Limited, United Kingdom.
Charness, G., Gneezy, U., 2012. Strong evidence for gender differences in risk-taking. Journal of Economic Behavior and Organization 83, 50–58.
Charness, G., Viceisza, A., 2011. Comprehension and risk elicitation in the field: evidence from rural Senegal. mimeo.
Coller, M., Williams, M.B., 1999. Eliciting individual discount rates. Experimental Economics 2 (2), 107–127.
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.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? American Economic Review 100 (3), 1238–1260.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2011. Individual risk attitudes: measurement, determinants and behavioral consequences. Journal of the European Economic Association 9 (3), 522–550.
Dreber, A., Rand, D.G., Wernerfelt, N., Garcia, J.R., Vilar, M.G., Lum, J.K., Zeckhauser, R., 2011. Dopamine and risk choices in different domains: findings among serious tournament bridge players. Journal of Risk and Uncertainty 43 (1), 19–38.
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 and Organization 68 (1), 1–17.
Eckel, C., Grossman, P.J., Johnson, C.A., de Oliveira, A.C.M., Rojas, C., Wilson, R., 2011. On the development of risk preferences: experimental evidence. mimeo.
Fecteau, S., Pascual-Leone, A., Zald, D.H., Liguori, P., Théoret, H., Boggio, P.S., Fregni, F., 2007. Activation of prefrontal cortex by transcranial direct current stimulation reduces appetite for risk during ambiguous decision making. The Journal of Neuroscience 27 (23), 6212–6218.
Friedman, D., Sunder, S., 2011. Risky curves: from unobservable utility to observable opportunity sets. mimeo.
Gneezy, U., Leonard, K.L., List, J.A., 2009. Gender differences in competition: evidence from a matrilineal and a patriarchal society. Econometrica 77 (5), 1637–1664.
Gneezy, U., Potters, J., 1997. An experiment on risk taking and evaluation periods. Quarterly Journal of Economics 112 (2), 631–645.
Haigh, M.S., List, J.A., 2005. Do professional traders exhibit myopic loss aversion? An experimental analysis. The Journal of Finance 60 (1), 523–534.
Hanoch, Y., Johnson, J.G., Wilke, A., 2006. Domain specificity in experimental measures and participant recruitment. Psychological Science 17 (4), 300–304.
Harrison, G.W., Lau, M.I., Rutström, E.E., 2007. Estimating risk attitudes in Denmark: a field experiment. Scandinavian Journal of Economics 109 (2), 341–368.
Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects. American Economic Review 92 (5), 1644–1655.
Hunt, M.K., Hopko, D.R., Bare, R., Lejuez, C.W., Robinson, E.V., 2005. Construct validity of the Balloon Analog Risk Task (BART): associations with psychopathy and impulsivity. Assessment 12 (4), 416–428.
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.
Kahneman, D., Knetsch, J.L., Thaler, R.H., 1990. Experimental tests of the endowment effect and the Coase theorem. Journal of Political Economy 98 (6), 1325–1348.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–292.
Laury, S.K., 2006. Pay one or pay all: random selection of one choice for payment. mimeo.
Lejuez, C.W., Aklin, W.M., Zvolensky, M.J., Pedulla, C.M., 2003. Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. Journal of Adolescence 26 (4), 475–479.
Lejuez, C.W., Read, J.P., Kahler, C.W., Richards, J.B., Ramsey, S.E., Stuart, G.L., Strong, D.R., et al., 2002. Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). Journal of Experimental Psychology: Applied 8 (2), 75–84.
Lonnqvist, J.E., Verkasalo, M., Walkowitz, G., Wichardt, P.C., 2011. Measuring individual risk attitudes in the lab: task or ask? An empirical comparison. mimeo.
Maccrimmon, K.R., Wehrung, D.A., 1990. Characteristics of risk taking executives. Management Science 36 (4), 422–435.
Reynaud, A., Couture, S., 2012. Stability of risk preference measures: results from a field experiment on French farmers. Theory and Decision 73 (2), 203–221.
Savage, L.J., 1954. The Foundations of Statistics. John Wiley, New York.
Tanaka, T., Camerer, C.F., Nguyen, Q., 2010. Risk and time preferences: linking experimental and household survey data from Vietnam. American Economic Review 100 (1), 557–571.
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.
von Gaudecker, H.M., van Soest, A., Wengstrom, E., 2011. Heterogeneity in risky choice behavior in a broad population. American Economic Review 101 (2), 664–694.