NETSPAR INDUSTRY SERIES

A survey of risk preference measures and their relation to field behavior

Paul Bokern
Jona Linde
Arno Riedl
Hans Schmeets
Peter Werner

SURVEY PAPER 58

Survey Papers are part of the **refereed Industry Paper Series**, which are refereed by the Netspar Editorial Board. Survey papers provide a concise summary of the ever-growing body of scientific literature on the effects of an aging society and, in addition, provide support for a better theoretical underpinning of policy advice. Netspar Survey papers attempt to present an overview of the latest, most relevant research, explain it in non-technical terms and offer a summary of the policy implications. Survey Papers are presented for discussion at Netspar events. Representatives of academic and private sector partners, are invited to these events. Survey Papers are published at the Netspar website.

CONTENTS

Affiliations

Paul Bokern – Maastricht University, Statistics Netherlands (Centraal Bureau voor de Statistiek)

Jona Linde – Maastricht University

Arno Riedl – Maastricht University

Hans Schmeets – Maastricht University, Statistics Netherlands (Centraal Bureau voor de Statistiek)

Peter Werner – Maastricht University

Abstract

Many economic and financial decisions involve risk. A crucial question in this context is how much risk people are willing to take. The importance of measuring the willingness of people to take risk is acknowledged by Dutch law, which requires financial institutions to take into account their customers' risk preferences when offering products and services. Additionally, knowledge of participants' risk preferences play an important role in the recent pension agreement in the Netherlands, since it requires pension providers to invest in line with the age composition and risk attitudes of participants.

Here we discuss the concept of risk preferences and common ways to measure them as proposed in the economics literature. We summarize traditional and behavioral economics perspectives on decisions under risk. Next, we present several incentivized experimental measures and non-incentivized survey measures of risk preferences and describe how they correlate with field behavior, and we discuss advantages and disadvantages of the different methods. We also discuss what is known, according to the economics literature, about the relationship between risk preferences and age.

Samenvatting

Risico is inherent aan veel economische en financiële keuzes. Een cruciale vraag hierbij is hoeveel risico mensen bereid zijn te nemen. Het belang van het vaststellen van risicobereidheid en het in kaart brengen van risicovoorkeuren wordt erkend door de Nederlandse wet, die financiële instellingen verplicht rekening te houden met de risicovoorkeuren van klanten bij het aanbieden van producten en diensten. Risicovoorkeuren spelen bijvoorbeeld ook een rol in het nieuwe pensioenakkoord, waarbij van fondsen verwacht wordt dat ze de risicohouding per leeftijdsgroep bepalen en beleggingen daarop afstemmen.

Hier bespreken we het concept risicovoorkeuren en hoe deze gebruikelijk gemeten worden volgens de economische literatuur. We bespreken traditionele en gedragseconomische perspectieven op keuzes onder risico. Vervolgens gaan we in op experimentele methodes met financiële prikkels alsmede op vragenlijsten die worden gebruikt om risicovoorkeuren vast te stellen, bespreken de correlatie tussen de gemeten risicovoorkeuren en gedrag, en presenteren de voor- en nadelen van de verschillende methodes. Ten slotte wordt vanuit de economische literatuur besproken wat er bekend is over de relatie tussen risicovoorkeuren en leeftijd.

Introduction

Many economic and financial decisions involve risk and uncertainty.[1] These decisions range from relatively low stakes and minor consequences to relatively high stakes and substantial long-term implications. For example, people decide to buy a bicycle with or without theft insurance, but they may also have to decide whether to invest a substantial amount in financial markets, to buy a house, and to take out insurance. An important determinant of how individuals decide in these and other contexts involving risk is their *risk preferences*, which determine how much risk people are willing to take. Given the importance of such decisions, it comes as no surprise that there is abundant academic literature in economics that studies risk preferences and how to measure them.

Risk preferences also play an important role in the recent pension agreement in the Netherlands. In particular, the Social Economic Council (SER) acknowledges the importance of investment policies that are in line with the age composition and risk attitude of participants (Koolmees, 2019). The Ministry of Social Affairs and Employment furthermore distinguishes between two aspects of risk attitude[2], namely the willingness to take risk – risk preference ("*risicobereidheid*") – and the extent to which participants can bear risk – risk capacity ("*risicodraagvlak*"). Pension funds therefore have the responsibility to measure risk attitudes accurately, to ensure that the investment policies are in the best interest of the pension participants. In this paper, we concentrate on how the willingness of individuals to take risk is measured in the field of economics, thus on what is called "*risicobereidheid*".[3]

The importance of measuring people's willingness to take risk is acknowledged by Dutch law, which requires financial institutions to consider their customers' risk

---

1   It is important to note that risk and uncertainty are related but distinct concepts. Risk refers to a situation where the final outcome is unknown whereas the probability distribution over potential outcomes is known. Uncertainty refers to a situation where both the outcome and the probability distribution are unknown (Knight, 1921). Both are relevant in economic and financial decision-making. In this survey we focus exclusively on risk.

2   Note that in the academic literature the term 'risk preference' is often used interchangeably with 'risk attitude', 'risk tolerance', or 'sensitivity to risk' (Hertwig, Wulff, & Mata, 2019).

3   Willingness to take risk and risk capacity may both depend on wealth and human capital, as well as on other economically relevant factors which may lead to a correlation between these two concepts. Importantly, for the elicitation of risk preferences (our main interest here), it should not matter whether they are dependent or not. Rather, it is important to measure risk preferences accurately. It is an empirical question to determine whether elicited risk preferences are correlated with the capacity to bear risk.

preferences when offering products and services.[4] However, the regulation does not stipulate how these preferences are to be measured. Consequently, financial institutions have been using a wide range of measures that are often not validated scientifically (Landman, 2013, 2019; Loonen & van Raaij, 2008). It is therefore unclear to what extent these measures accurately reflect the preferences of customers. Loonen and van Raaij (2008), for instance, find that elicited risk preferences differ substantially across banks using different measures.

Several authors have proposed methods from the academic literature which financial institutions could apply (Alserda, Dellaert, Swinkels, & van der Lecq, 2016, 2019; Dellaert, Donkers, Turlings, Steenkamp, & Vermeulen, 2016; Potters, Riedl, & Smeets, 2016; van der Meeren, de Cloe-Cos, & van Geen, 2019). Contrary to hypothetical questions used currently in practice, these papers propose incentivized methods, in which participants make real decisions with real financial consequences, from which preferences can be inferred. Hypothetical questions commonly elicit so-called 'stated preferences', whereas incentivized methods are considered to elicit so-called 'revealed preferences'.

Economists are involved in a lively debate regarding the methods that should be used to elicit risk preferences properly (for extensive reviews see Charness, Gneezy, & Imas, 2013; Harrison & Rutström, 2008). Charness et al. (2013) discuss methods based on their complexity. Methods range from very simple non-incentivized survey questions that ask about the self-reported propensity of individuals to take risk (Dohmen et al., 2011) to more complex experimental tasks such as a Multiple Price List (MPL), where participants make real decisions involving real risk (Holt & Laury, 2002). It is important to understand the advantages and disadvantages of the different methods before deciding on the best elicitation method for the intended purposes . We discuss these methods in Section 3 in more detail.

The main purpose of this paper is to discuss risk preferences and how they are most commonly measured according to the economic literature. As a first step, we explain two major theoretical frameworks to analyze individual decision-making under risk. This is followed by a presentation of several often-used incentivized experimental tasks and non-incentivized survey questions to elicit individual risk preferences and by a discussion of their advantages and disadvantages. We also report on results how the elicited risk preferences correlate with (self-reported) field behavior. Finally, we document the evidence regarding the relationship between risk

4   Art. 4:23 Wet op het financieel toezicht.
    See https://wetten.overheid.nl/BWBR0020368/2020-12-29

preferences and age. This relationship is particularly important in view of the new pension agreement in the Netherlands, as pension funds are expected to determine the risk attitude of different age groups and to invest accordingly (Koolmees, 2020).

The standard framework in economics to study risk preferences is *Expected Utility Theory* (hereafter EUT; Von Neumann & Morgenstern, 1944). EUT builds on a number of reasonable axioms of rational decision– making, especially transitivity, and it is therefore considered a *normative* theory of decisions under risk. An important aspect of EUT is that risk preferences can be represented by an *expected utility function*. Risk–averse or risk–seeking preferences are captured by the shape of this utility function. EUT is considered to be tractable and to apply in a wide range of decision domains. Its normative appeal is largely based on the notion that an individual can only make consistent, situationally independent, and non–exploitable decisions under risk when deciding according to EUT. Therefore, economists largely interpret EUT as an adequate theory of rational decision–making.

Despite its normative appeal, EUT has been criticized for not being a *descriptively valid* theory of choice (Starmer, 2000). In fact, numerous systematic deviations from EUT have been identified both in the lab and in the field. Consequently, risk prefer–ences elicited from the answers people give to survey questions or choices they make in decision situations involving risk may not necessarily obey the axioms of EUT and thus rational choice. This presents a challenge for financial institutions as they seek to make rational investment decisions on behalf of their clients.

One response to this limitation of EUT as a descriptive theory is to use measures that try to minimize the effect of known deviations. This is the approach taken by, for example, Alserda et al. (2019). An alternative response is to explicitly take into account the way people deviate from EUT when interpreting elicited preferences. To do so, a descriptively accurate theory of decision–making under risk is required. In their semi–nal study, Kahneman and Tversky (1979) propose *Prospect Theory* (hereafter PT) as such an alternative. Since its inception, PT has proven to be very influential.

Three central concepts of PT that have had a large impact on economics and finance are *reference dependence*, *loss aversion*, and *probability weighting*. These concepts entail that people (1) evaluate potential outcomes based on utility gains or losses relative to some reference point rather than evaluating final outcomes, (2) are more sensitive to losses than gains of equal magnitude, and (3) weigh probabilities non–linearly. Many authors consider PT the most accurate descriptive theory of how people make choices under risk (Barberis, 2013).

The paper proceeds as follows. Section 2 presents the concept of risk preferences in economics. Section 3 provides an overview of elicitation methods, how they are used,

and their advantages and disadvantages. Section 4 reviews studies that link decisions from risk preference elicitation tasks to decisions in the field involving risk. Section 5 discusses studies that link measures for risk preferences with the age of decision makers. Lastly, section 6 discusses the implications of our findings.

2. Risk preferences in economic theory

2.1 Expected Utility Theory

In economic science, risk commonly refers to variance in outcomes. Suppose, for instance, that a gamble is presented where a decision maker is asked to choose between €100 for sure and receiving either nothing or €200 with equal probability. The expected value is €100 in both options, while the variance of the outcome is larger in the latter option. Risk preferences determine the willingness (or lack thereof) to choose in favor of options with higher variance. In this example, a risk-averse individual would prefer the certain option, while a risk-seeking individual would prefer the risky option. An individual who is indifferent between the two options is classified as risk-neutral.

A formalization of the idea that people evaluate options not only on the basis of the expected value but also consider variance was first proposed by Bernoulli (1738). He argued that people do not take into account the objective value of outcomes, but rather assign "moral value" (i.e. utility) when evaluating a prospect. In particular, individuals assign utility to each of the outcomes and weigh these against the probability of occurrence of the respective outcome.[5] A utility-maximizing individual chooses the prospect that yields the highest expected utility. This concept is known as Expected Utility Theory (EUT) and is widely used in economic theory and applications (Von Neumann & Morgenstern, 1944).[6]

Under EUT, there is a one-to-one relationship between the shape of the utility function and the class of risk preferences it represents. A concave utility function represents risk aversion, while a convex utility function represents risk seeking. Figure 1 shows an example of a concave utility function. The concavity of the utility function implies that the individual receives a higher utility from a certain amount equal to the expected value of a gamble compared to the risky gamble itself. Hence, the individual represented by this example is risk-averse.

Individual risk preferences – and thus the shape of the utility function – are personal to the decision maker and therefore unknown to the researcher. To infer individual risk preferences, economists have developed numerous experimental methods

---

5 Expected Utility (EU) is defined as follows: $EU = \sum_i p_i * u(x_i)$ where $u(\cdot)$ is a Von Neumann–Morgenstern utility function (Von Neumann & Morgenstern, 1944) defined on a set of outcomes $x_i \in X$. Note that the argument of the utility function is not part of the axiomatic foundation of EUT. Previous studies have used terminal wealth, income, and consumption (see Harrison, Lau, Ross, and Swarthout (2017) for a discussion).

6 See Starmer (2000) for a review on EUT, its limitations, and alternatives.

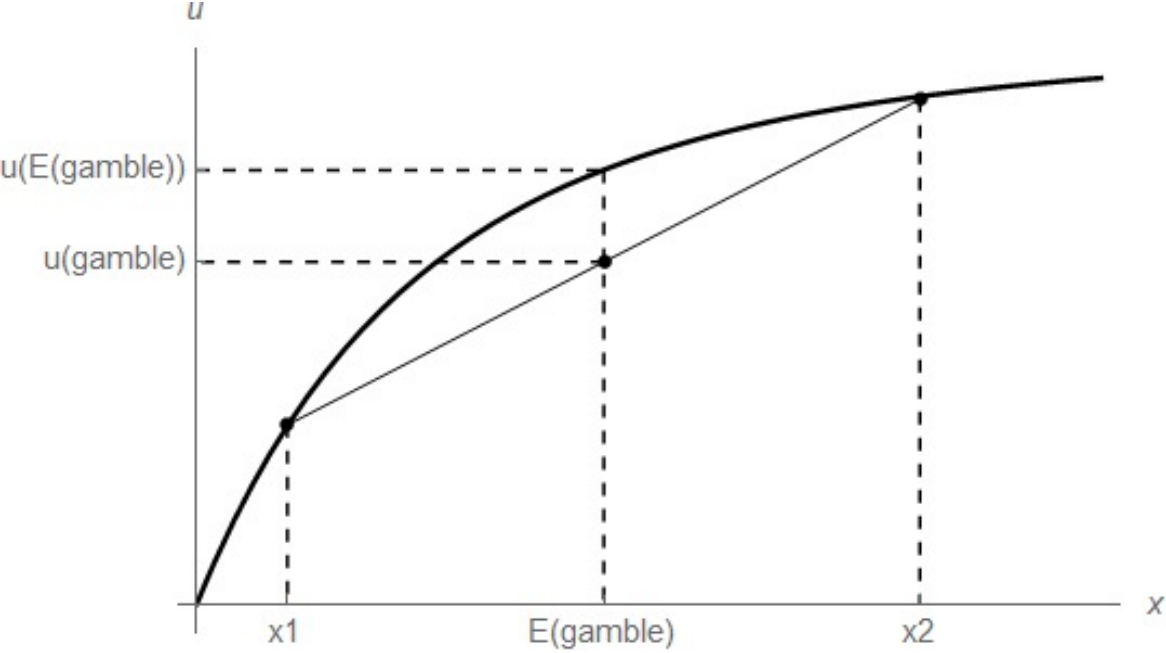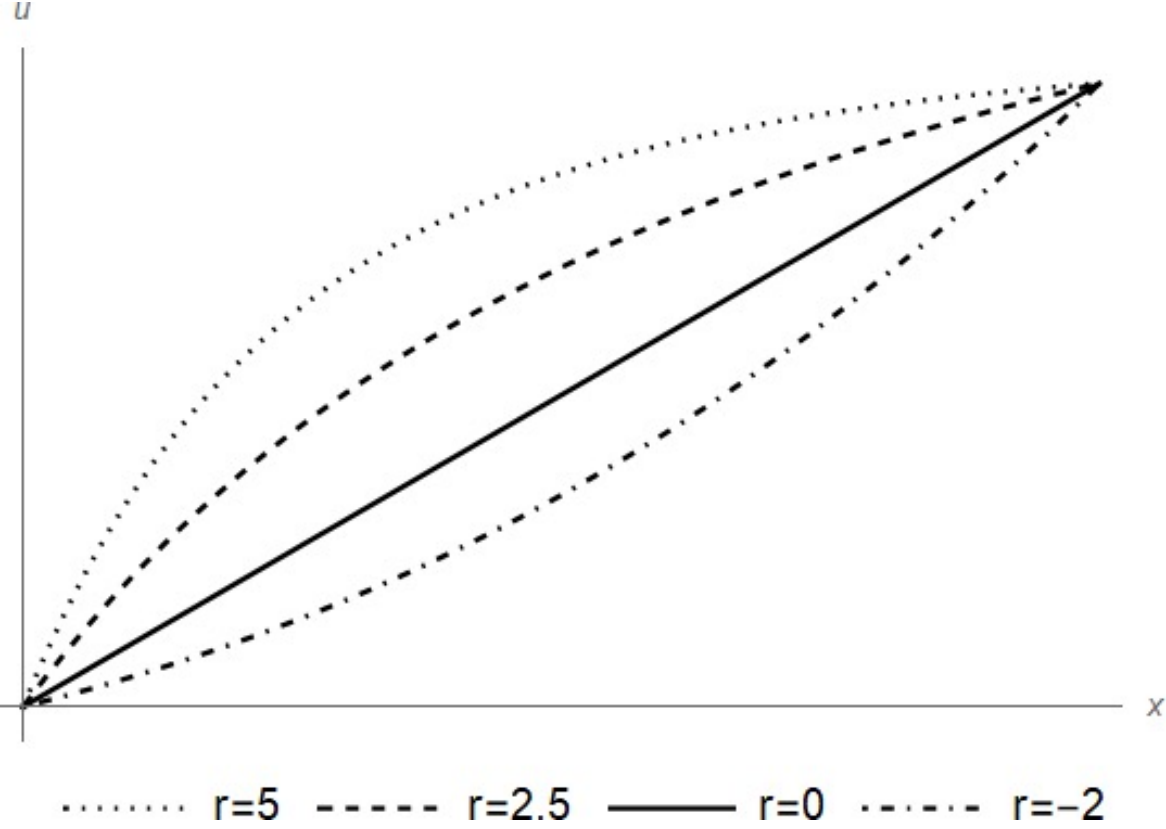*Figure 1. Example of a concave utility function.*

Figure 2. Example of CRRA utility functions.

Box 1. Higher Order Risk Preference: Prudence and Temperance.
In an expected utility framework, risk aversion and risk seeking are represented by the concavity and convexity, respectively, of the utility function, which can formally be assessed by looking at its second derivate ($u'' < 0$). More recently, theoretical work has shown how higher order risk preferences may also be relevant for economic decisions. In a life-cycle savings model, for instance, the sign of the third derivative affects the relationship between savings and the uncertainty of future income, the so-called background risk. A positive third derivative ($u''' > 0$) is defined by Kimball (1990) as prudence. A prudent individual prefers to increase savings in response to higher background risk and thus prefers more precautionary savings. Similarly, the sign of the fourth derivative affects the relationship between background risk and the risk level of an individual's portfolio. Kimball (1993) defined a negative fourth derivative ($u'''' < 0$) as temperance. A temperate individual prefers to make less risky investments as background risk increases.

In their seminal work, Eeckhoudt and Schlesinger (2006) provide a behavioral foundation of prudence and temperance, which can be used to elicit them in an experimental setting. Following this approach, Noussair, Trautmann, and Van de Kuilen (2013) elicit higher order risk attitudes in a representative sample of the Dutch population and link it to self-reported field behavior. Their results indicate that higher prudence correlates positively with savings and negatively with credit card debt. Stronger temperance was found to correlate negatively with having risky investments and the share of risky investments in portfolios. Using a different method, Schneider and Sutter (2020) also find that prudence is positively related to savings and that temperance is negatively related to risky investment. Additionally, their measure for temperance is positively related to savings. For an extensive review, see Trautmann and Van de Kuilen (2018).

(see Section 3). Typically, the researcher assumes some specific functional form that defines the utility function. One functional form that is widely used is constant relative risk aversion (CRRA), which captures risk aversion with only one parameter (Harrison & Rutström, 2008).[7] Figure 2 displays CRRA utility functions for different values of this parameter, here called *r*. As *r* increases, the utility function becomes more concave and hence implies a higher level of risk aversion.

The literature in economics has also gone beyond the simple version of CRRA, exploring other aspects that may affect people's economic decisions under risk. Within the EUT framework, for instance, higher order risk preferences have been shown to be theoretically and empirically relevant (Noussair et al., 2013) (see Box 1 for more information on higher order risk preferences). Outside of the EUT framework, PT (Kahneman & Tversky, 1979) has had considerable impact. We discuss the key elements of PT next.

---

7   CRRA is commonly defined as follows: $u(x) = x^{1-r}/(1-r)$ for $r \neq 1$ and $u(x) = \ln(x)$ for $r = 1$. This formulation implies risk seeking for $r < 0$, risk neutrality for $r = 0$ and risk aversion for $r > 0$ (see, e.g., Wakker (2008) for more details).

Figure 3. Example of a value function.
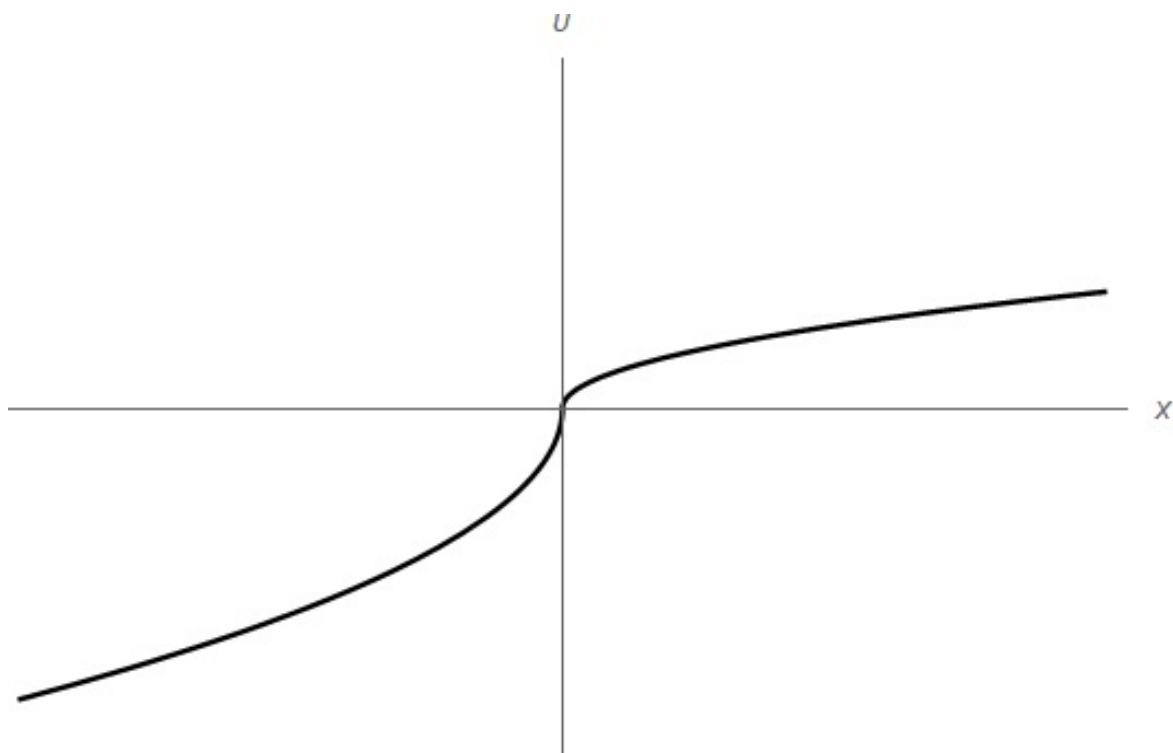
2.2 Prospect Theory

In their seminal paper, Kahneman and Tversky (1979) show in a series of experiments that people systematically deviate from EUT and use this insight to propose PT as an alternative theory.[8] PT applies insights from psychology to provide an empirically more accurate theory of choice under risk. In particular, the utility function over outcomes used in EUT is replaced by a value function, which is defined over gains and losses relative to a reference point. Additionally, instead of assuming that people weigh utility over outcomes with their respective probabilities, as in EUT, PT assumes that people weigh probabilities non- linearly, assigning decision weights to each potential outcome that are transformations of the objective probabilities.[9] Figure 3

---

8   Tversky and Kahneman published the original version of PT in 1979. To overcome some limitations of this version, they published a modified version called Cumulative Prospect Theory in 1992. See Barberis (2013) for a recent review.

9   The PT value is defined as follows: $V = \sum_i \pi_i * v(x_i)$ where π_i represents decision weights and $v(\cdot)$ the value function, which is an increasing function with $v(0) = 0$ (Kahneman & Tversky, 1979). Tversky and Kahneman (1992) propose the following functional form for the value function: x^a for gains and $-\lambda * (-x)^a$ for losses and the following functional form for probability weighting: $\pi(p) = p^\gamma / (p^\gamma + (1-p)^\gamma)^{1-\gamma}$.

shows an example of a value function.[10] In the example, the reference point is at 0 and the value on the horizontal axis (*x*) represents the gain or loss relative to the reference point.

The value function illustrates two key principles of PT. The first key principle is reference dependence, meaning that people evaluate gains and losses relative to a reference point. As an illustration, consider a decision situation where the decision maker is endowed with €20 and is in addition offered a gamble involving a possibility of losing €10 or gaining €15 with equal probability. Under EUT, the individual would base the choice on comparing the certain outcome of €20 with the gamble of receiving €10 or €35 with equal probability.[11] Under PT, assuming that the reference point is the endowment given, the individual would compare the value of €0 with certainty and the value of a prospect of losing €10 or gaining €15 with equal weighted probability.

The second key principle is that losses loom larger than gains. In other words, people dislike the loss of a given amount more than the amount of pleasure that they derive from gaining the same amount. This effect is represented by the value function in Figure 3, where the function is steeper in the domain of losses than in the domain of gains. Taking the previous example, an individual who is sufficiently loss-averse may decline the gamble, because the perceived negative value from losing €10 may outweigh the perceived value from gaining €15.

The third key principle is probability weighting. In the EUT framework, it holds that people weigh utilities with their respective objective or subjective probabilities. In other words, utility is weighted linearly. Contrary to this, Kahneman and Tversky (1979) posit that people distort probabilities systematically when they evaluate prospects, assigning decision weights to risky outcomes that typically deviate from the attached probabilities. Figure 4 displays the weighting function proposed by Tversky and Kahneman (1992). The horizontal axis refers to objective probabilities, while the vertical axis shows the decision weights that decision makers attach to these objective probabilities. The dashed 45-degree line represents the situation where the decision weights match the probabilities. Comparing the two lines, it is clear that the decision

---

10 The example in Figure 3 shows a value function that is convex in the loss domain, as found experimentally by Tversky and Kahneman (1992). Note, however, that this finding is contested in the literature (see Bilsen, Laeven, & Nijman, 2020 and the references therein). Here we primarily focus on risk attitudes in the domain of gains.

11 Note that we assume here that decision makers only consider the outcomes in the decision situation and ignore their lifetime wealth. For a discussion on this, see also Footnote 5 above.

Figure 4. Example of a probability weighting function.
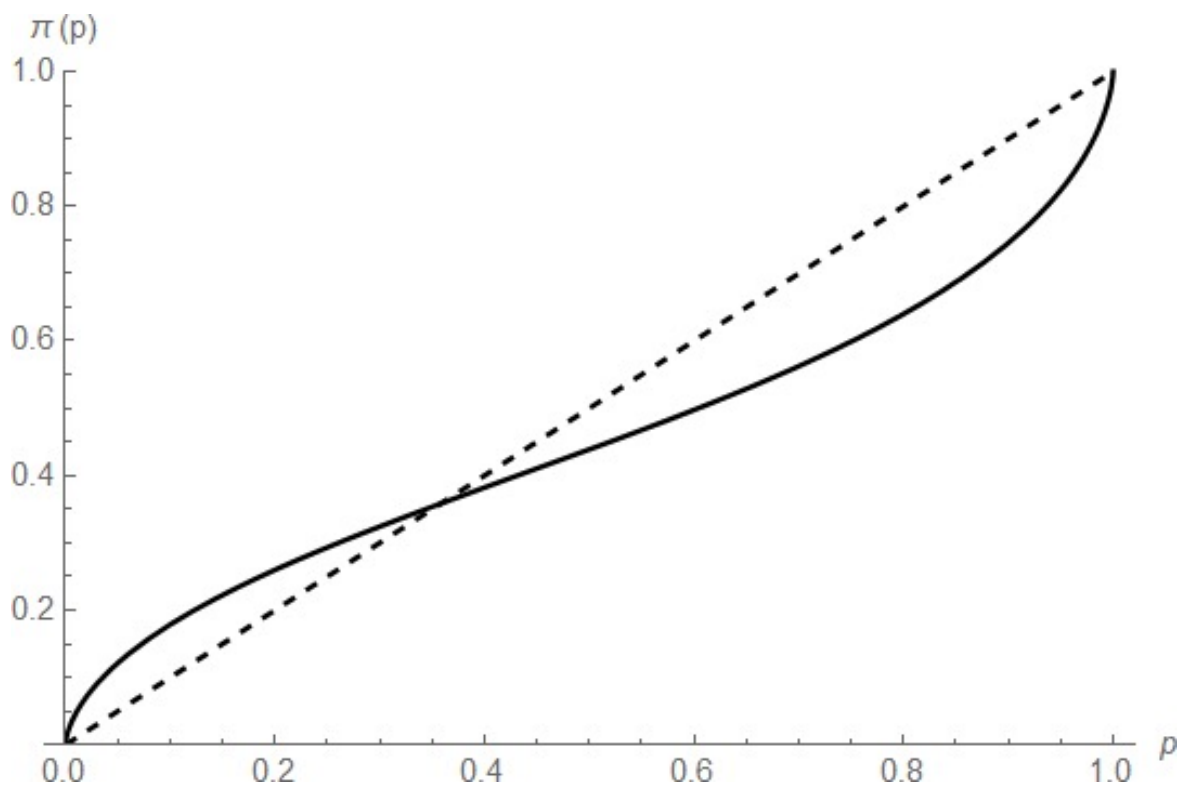
weights differ from the probabilities in most cases: a typical weighting function over-weights low probabilities and underweights relatively high probabilities.

The shape of the value function and the probability weighting function are specific to the individual and thus unknown to the researcher. It is important to note that some methods used to elicit risk preferences may be confounded with principles from PT. The interested reader is referred to, for instance, Tanaka, Camerer, and Nguyen (2010) for simultaneous measurement of the shape of the value and probability weighting functions.[12] Some of the methods in Section 3 also allow for the measurement of one or more elements of PT.

12 See also Abdellaoui, Bleichrodt, and Paraschiv (2007), Abdellaoui, Bleichrodt, L'Haridon, and Van Dolder (2016), and Chapman, Snowberg, Wang, and Camerer (2018) for loss aversion, and Abdellaoui (2000) for probability weighting.

3. Methodologies to elicit risk preferences

In economics, people's risk preferences are most commonly measured through the revealed preference approach. In this approach, preferences are inferred from choices with real, usually financial, consequences, either in the field or in the laboratory (Dohmen, Quercia, & Willrodt, 2019). Numerous methodologies have been proposed for this purpose (for extensive reviews, see Charness et al., 2013; Harrison & Rutström, 2008). It is important to note that these methods do not directly reveal a person's risk preferences. The elicited measures need to be interpreted and then translated into parameters that relate to the underlying assumed model of decision-making, which can be used to predict behavior, for instance related to investment decisions. In particular, to make sure that the measures are interpreted correctly, it is necessary to assess the extent to which a measure could be affected by deviations from EUT and potentially to explicitly estimate all relevant parameters in a PT framework.[13]

More recently, economists have also started to use stated preferences measures. In this approach, people are directly asked to assess their willingness to take risk. For these measures, usually no underlying utility model such as the EUT or PT is assumed. That makes it difficult to translate them into decision-making models based on EUT or PT that could be used to predict behavior such as investment decisions. We discuss examples of both approaches.

3.1 Multiple Price List

The multiple price list (MPL) risk elicitation method was popularized by Holt and Laury (hereafter HL, 2000) and has been applied in different versions numerous times thereafter.[14] Participants in this experimental task receive a list of pairs of lotteries. In each pair, the participant must choose between the two options; say, option A and option B. The list is designed such that the expected value of both or one of the options changes down the list. Assuming EUT, the point where the participant switches from choosing one option to the other allows the researcher to infer information about the participant's risk preference. If PT rather than EUT is assumed, additional lists are required to identify PT's additional parameters and still be able to identify curvature of the value function (see, for instance, Tanaka et al., 2010).

---

13 The econometric methods required to perform this estimation go beyond the scope of this paper.

14 In the Dutch pension domain, for instance, Alserda et al. (2016, 2019) propose an MPL that is tailored to the pension context to infer risk preferences.

Table 1. Original MPL in HL (2002).

Lottery	Option A				EV(A)	Option B				EV(B)
	$p(\$2)$		$p(\$1.60)$			$p(\$3.85)$		$p(\$0.10)$		
#1	0.1	$2.00	0.9	$1.60	$1.64	0.1	$3.85	0.9	$0.10	$0.48
#2	0.2	$2.00	0.8	$1.60	$1.68	0.2	$3.85	0.8	$0.10	$0.85
#3	0.3	$2.00	0.7	$1.60	$1.72	0.3	$3.85	0.7	$0.10	$1.23
#4	0.4	$2.00	0.6	$1.60	$1.76	0.4	$3.85	0.6	$0.10	$1.60
#5	0.5	$2.00	0.5	$1.60	$1.80	0.5	$3.85	0.5	$0.10	$1.98
#6	0.6	$2.00	0.4	$1.60	$1.84	0.6	$3.85	0.4	$0.10	$2.35
#7	0.7	$2.00	0.3	$1.60	$1.88	0.7	$3.85	0.3	$0.10	$2.73
#8	0.8	$2.00	0.2	$1.60	$1.92	0.8	$3.85	0.2	$0.10	$3.10
#9	0.9	$2.00	0.1	$1.60	$1.96	0.9	$3.85	0.1	$0.10	$3.48
#10	1	$2.00	0	$1.60	$2.00	1	$3.85	0	$0.10	$3.85

Note: the columns labeled EV(A) and EV(B) list the expected value of the related lottery.

To illustrate this, Table 1 displays the original MPL introduced by HL. In this MPL, the options differ in terms of the monetary amounts that the participant can receive. In particular, the monetary amounts in option B are more variable than in option A, making option B the more risky option. Moving down the list, the probability that each outcome occurs changes such that the expected value (EV) of both options increases, but at a higher rate in option B. A risk-neutral person would compare the EV of both options and switch to option B once the EV is larger than in the other option (in the example that is the case from #5 onwards). A risk-averse participant switches later and a risk-seeking participant earlier, depending on the strength of the participant's aversion to risk.

Assuming EUT and the CRRA functional form, one can formally calculate the bounds of the risk aversion parameter r that is implied by a particular switching point from option A to option B. In Table 2, we display the CRRA range and risk preference classification implied by the number of safe choices (option A) in the original MPL by HL.[15]

---

15 To calculate the CRRA range, Holt and Laury (2002) implicitly assume that expected utility is defined over income earned from the experiment. As a result, the implied CRRA parameter in a specific row can be calculated by equating the utilities of option A and option B. For instance, to calculate the implied parameter of #2, one has to solve the following equation to find $r \approx -0.95$:

$$EU(A) = EU(B) \leftrightarrow (0.2 * \frac{2.00^{1-r}}{1-r} + 0.8 * \frac{1.60^{1-r}}{1-r}) = (0.2 * \frac{3.85^{1-r}}{1-r} + 0.8 * \frac{0.10^{1-r}}{1-r})$$

The assumption that expected utility is defined over income earned from the experiment, rather than integrated with terminal wealth or monthly income, is referred to as narrow bracketing. Several authors have found empirical evidence in favor of this assumption (Andersen, Cox, et al., 2018; Harrison et al., 2017; Schechter, 2007).

Table 2. Risk classification of original MPL in HL.

Number of safe choices	CRRA range
0–1	r < −0.95
2	−0.95 < r < −0.49
3	−0.49 < r < −0.15
4	−0.15 < r < 0.15
5	0.15 < r < 0.41
6	0.41 < r < 0.68
7	0.68 < r < 0.97
8	0.97 < r < 1.37
9–10	1.37 < r

Using several MPLs, it is possible to get a more precise estimate of the risk aversion parameter, which can then be structurally estimated.[16]

The MPL proposed by HL keeps the outcomes within each option constant, but it varies probabilities in each row in such a way that the risky option becomes more attractive when moving down the list. An issue with this format is that it may confound risk preferences with the extent to which participants weight probabilities non-linearly, an issue already recognized by Wakker and Deneffe (1996). In line with this, Drichoutis and Lusk (2016) find evidence that the MPL used by HL is more likely to provide accurate information about the shape of the participant's probability weighting function than about the shape of the utility function. They propose an alternative format where, instead of probabilities, the outcomes of the lotteries are varied. Probabilities are kept constant in every row at a probability level of 50% that either outcome occurs.

In addition to the format proposed by Drichoutis and Lusk (2016), there exist many other types of MPLs. Csermely and Rabas (2016) identify nine different MPL formats used in the literature. These formats are grouped into two categories: paired gamble (PG) formats and standard gamble (SG) formats. PG formats have generic lotteries on both sides of the table, as in the original MPL by HL. SG formats have a certain payoff (i.e., a degenerate lottery) on one side of the table (the amount is paid out with certainty) and a generic lottery on the other side. The types of MPLs differ from each other in terms of what is changed when moving along the list: probabilities, highest payoff, lowest payoff, sure payoff, or a combination.

All formats of the MPL method aim to measure the true underlying risk preferences of the participant. However, Csermely and Rabas (2016) find in their experiments that different MPL formats may yield different estimated parameter values. They conclude:

16 For a practical explanation of this approach, see Appendix F in Harrison and Rutström (2008).

"This implies that an arbitrary selection of a particular risk assessment method can lead to differing results and misleading revealed preferences." (p.130). Comparing the different measures in terms of within-method and between-method consistency, they suggest using the MPL format proposed by Drichoutis and Lusk (2016) when one plans to measure risk preferences using the MPL method.

The main advantage of the MPL method is that the task is relatively transparent for participants and provides clear incentives to reveal preferences truthfully (Andersen, Harrison, Lau, & Rutström, 2006).

However, the results from Csermely and Rabas (2016) and Drichoutis and Lusk (2016) illustrate that one has to be careful in choosing the format that is used. Another concern is that the method is potentially susceptible to framing effects (Andersen et al., 2006; Andersen, Harrison, Lau, & Rutström, 2008). For instance, participants may be drawn to the middle of the table or the risk-neutral choice. This problem can be avoided by presenting each row separately and in random order. A downside of such an approach, however, could be that the data become noisier due to mistakes. Furthermore, the method may result in inconsistent choices if people switch multiple times between decision options. This is not necessarily a problem if one structurally estimates preference parameters and includes a parameter for making errors (Von Gaudecker, Van Soest, & Wengstrom, 2011). However, it may be problematic if one aims to infer risk preferences directly from switching points in the MPLs. Combining these two possible drawbacks – framing and inconsistent choice – Andersson, Holm, Tyran, and Wengström (2016) show that the propensity to make errors also depends on where the neutral switching point is placed in the list. To account for this, they suggest using several MPLs with varying switching points for eliciting risk preferences.

A related approach that some researchers have taken is asking lotteries in a sequence, using a so-called staircase method instead of displaying an entire list of lotteries (see, for instance, Barsky, Juster, Kimball, & Shapiro, 1997; Falk, Becker, Dohmen, Huffman, & Sunde, 2016). Van der Meeren et al. (2019) propose this method in the context of pensions. In this method, participants make a decision only between two lotteries at a time. The decision that the participant makes determines the next pair of lotteries in the sequence that the participant receives. Compared to the standard MPLs, this addresses the concerns of framing and multiple switching. In addition, it potentially calls for a lower number of decisions per participant to obtain a relatively precise estimate of the risk preference parameter. A major disadvantage is that mistakes cannot be easily identified by the researcher, while such mistakes will lead participants into a sequence that may not reflect their preferences at all. In addition, participants may become aware of the procedure and could exploit its

sequential structure by choosing the option that will lead them to the most profitable sequence if the experiment is incentivized.

3.2 Convex Time Budget

The Convex Time Budget (CTB) is a more recent elicitation method, proposed by Andreoni and Sprenger (2012a) and applied to the pension domain by Potters et al. (2016). In this method, participants receive a monetary budget that they are free to allocate to either of two accounts. The first account pays out some amount of money at an early date with certainty. The second account pays out a larger amount of money at a later date but only with some probability. Figure 5 shows an example of a decision screen. In this example, the early amount is paid out exactly eight weeks after the decision is made. The late amount is paid out with a probability of 90% exactly 24 weeks after the decision is made. In this example, the participant can earn 6% interest by waiting longer.

Participants have to make decisions in several choice sets. The choice sets vary in a couple of parameters: the timing of the payment at the later date, the probability that the later amount will be paid out, and the interest rate that the participant can earn

Figure 5. CTB Decision Screen Example
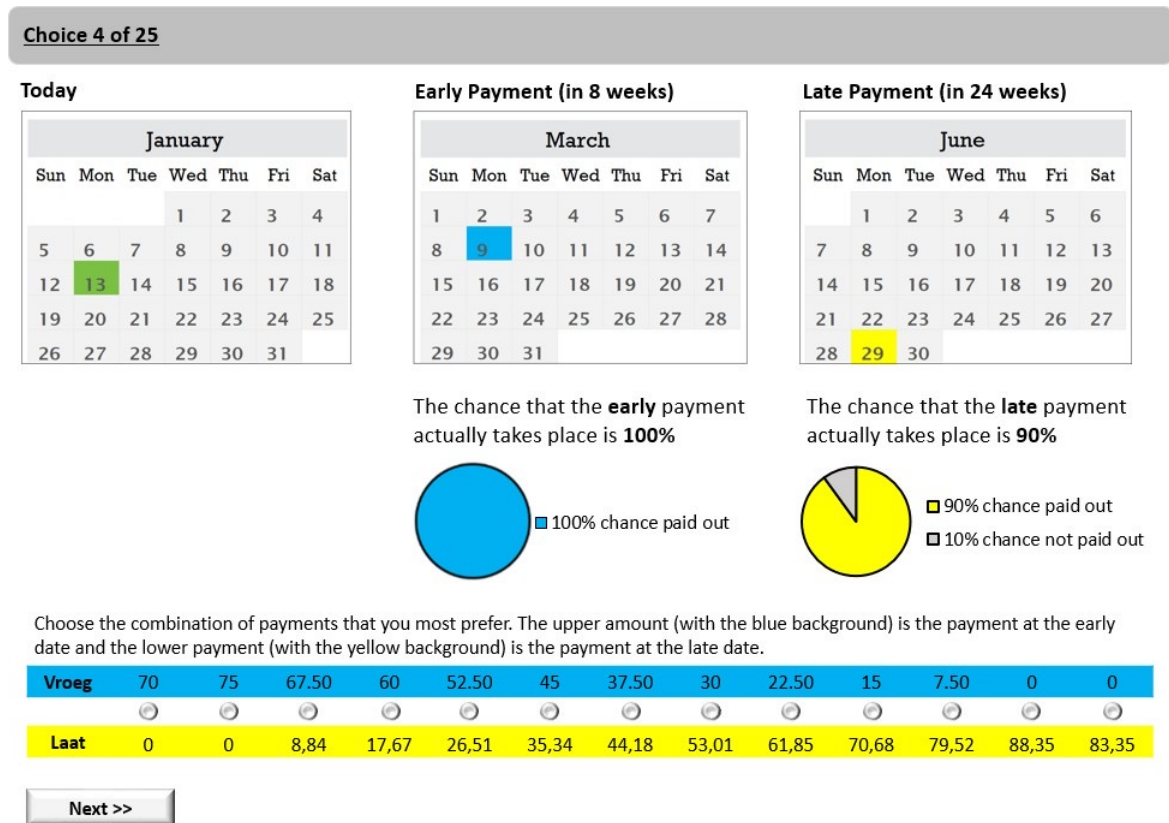
by waiting for the larger amount at the later date (see Potters et al. (2016, pp. 22–23) for an overview of possible parameters). An important advantage of this method is that it allows the researcher to infer risk preferences, probability weighting, and time preferences simultaneously.[17] It also allows for measuring present bias, if it is feasible to pay out the early payment immediately (Imai, Rutter, & Camerer, 2020). The interested reader is referred to Andreoni and Sprenger (2012a) and Andreoni and Sprenger (2012b) for more information on the estimation of preference parameters that apply this method.

The possibility of estimating individual parameters of risk preferences, time preferences, and probability weighting at the same time is especially relevant in the context of financial decisions in the field, where these factors likely play a role simultaneously. A disadvantage of the method is that it is somewhat more complex than other elicitation methods. To the best of our knowledge, there are no studies so far that correlate risk preferences estimated using this method with behavior outside the lab.

3.3 Investment Task

The Investment Task (IT) is an elicitation method proposed by Gneezy and Potters (1997). In this method, participants receive a monetary endowment of €X and are asked to decide how much of this they want to invest into a risky option and how much to keep liquid. Any money invested will yield a dividend of $k*$€x with a probability p and is otherwise lost. The money that is not invested will be kept by the participant. The parameters p and k are chosen such that investing more always raises the expected value and the variance. In this set-up, individuals who are risk-neutral or risk-seeking should always invest their entire endowment. Risk-averse individuals, on the other hand, may invest less depending on how risk-averse they are. The value €x, which is the only decision participants make in this experiment, is thus the measure of risk aversion.

The main advantage of this method is that it requires only one trial and is relatively simple for participants to understand. A disadvantage of this method is that it cannot distinguish between individuals who are risk-neutral and those who are risk-seeking. Without varying p, and therefore introducing additional tasks, it is also not possible to identify the influence of probability weighting.

---

17 Time preferences reflect the trade-off between consumption today and consumption in the future. Time preferences are also relevant for financial decisions, but beyond the scope of this article. See Cohen, Ericson, Laibson, and White (2020) for a review.

3.4 Distribution Builder

The Distribution Builder (DB) is an elicitation method introduced by Sharpe, Goldstein, and Blythe (2000) and that has been proposed for the Dutch pension context by Dellaert et al. (2016). The DB is developed to elicit information about the risk preferences of investors. Participants in this task choose their most preferred probability distribution of future wealth among a wide range of alternatives with equal cost. Goldstein, Johnson, and Sharpe (2008) apply the method to investment preferences for retirement income.

In this method, participants are asked to create a probability distribution over a range of final outcomes. In the pension context, the final outcomes are different levels of yearly income during retirement. The probability distribution is represented by 100 movable units that are displayed in an interactive interface. The participant can freely move units up and down to reach the desired probability distribution, provided that the cost of the distribution does not exceed the available budget.

A key principle of the DB is that the cost of each possible probability distribution is determined using a method that is consistent with a model of equilibrium pricing in capital markets (Sharpe et al., 2000).[18] In particular, the cost function assigns a unique price to each of the 100 units. The cost of the distribution is determined by weighing each of the outcomes by the associated price. Participants therefore have to make a trade-off between the chance of receiving a higher outcome and the higher cost associated with it. Due to the budget constraint that participants face, they thus have to accept the probability of receiving a lower income. This reflects the risk-return trade-off principle of capital markets.

The preferred probability distributions indicated by the participants is used to infer their risk preferences. Sharpe et al. (2000) assume EUT and illustrate how to approach this for an investor with a CRRA utility function. Goldstein et al. (2008) take the same approach, but additionally use the DB to measure loss aversion parameters. Donkers, Lourenço, Dellaert, and Goldstein (2013) assume PT and estimate parameters for the value function and probability weighting.

There are several potential advantages associated with the DB (Dellaert et al., 2016). First, the DB displays probabilities in terms of frequencies (units) in an intuitive graphic interface, which may improve understanding compared to directly displaying the probabilities. Second, the participant experiences trade-offs that are made in real investment decisions in an interactive and playful manner (gamification). This makes

18 Details of the cost function can be found in Sharpe et al. (2000) and Sharpe (2001). See also Donkers, Lourenço, Goldstein, and Dellaert (2013) for a discussion of design considerations, in particular for financial investment and pension decision-making.

the complex task of making investment decisions more accessible but may at the same time have a framing effect and thereby affect risk-taking behavior.

Some remarks are warranted here. First, the DB requires participants to make decisions in a context that aims to resemble actual investment decisions, although in a simplified matter. If the DB is successful in resembling an actual decision, the inferred preferences should describe how people make decisions in such a context. This implies that for many participants these decisions do not merely reflect risk preferences but also probability weighting and loss aversion (especially with an explicit reference point as in Goldstein et al. (2008)). Consequently, assuming EUT and CRRA may not be justified, and risk preferences cannot be inferred directly from decisions. Second, to the best of our knowledge, little research has been conducted to compare the DB with other established measures and no research to test its external validity. Thus, it is important to conduct more research before the DB relative to other elicitation methods can be assessed.

3.5 General Risk Question

The General Risk Question (GRQ) is a widely known non-incentivized risk measure that simply asks participants to state their willingness to take risk (Dohmen et al., 2011). The question is stated as follows: "How do you see yourself? Are you generally a person who is fully prepared to take risks, or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means not at all willing to take risks and the value 10 means fully prepared to take risks". In addition to this general question, the authors propose domain-specific risk questions. These questions use the same wording as the general question, but then referring to a specific context such as financial matters, health, or leisure.

A clear advantage of this method is its simplicity. A drawback of the method is that it is not clear what it precisely measures. In addition, the measure is not incentive-compatible, which makes economists wary of inattention or responses that participants perceive as socially desirable (Arslan et al., 2020). To address the first issue, Arslan et al. (2020) investigate what a person reasons when responding to the GRQ. Their results indicate that participants are highly heterogeneous in the risk that they think of. However, major life decisions, such as taking risk in investments, careers, or relationships are much more often mentioned than smaller compound risks such as smoking or unprotected sex. The latter issue is addressed by Dohmen et al. (2011), who find that the GRQ is a reliable predictor for risk-taking behavior in an incentivized MPL.

4. External validity

The risk elicitation methods discussed in the previous section all aim to measure the true latent risk preferences of participants. An important consideration in this regard is to what extent risk preferences elicited with incentivized experimental measures or hypothetical survey questions correlate with behavioral patterns in the field. Specifically, with regard to individual investors or pension savers, an important question is to what extent the elicited risk preferences link to their actual financial decision-making. In this section, we discuss research that has elicited risk preferences, using one or more of the methods discussed in this paper and correlate them with field behavior. Table 3 provides an overview of the literature.

Table 3. Overview of studies that investigate the relation between field behavior and risk preferences, measured by the risk elicitation methods discussed in this paper: Multiple Price List (MPL), Convex Time Budget (CTB), Investment Task (IT), Distribution Builder (DB), and General Risk Question (GRQ).

Article	Sample	Measure	Field Behavior Financial	Field Behavior Other
Dustmann et al. (in press)	General population China (n=6332)	GRQ		Migrating for work***(>3 months away from home for work or business)
Charness et al. (2020)	Representative sample Netherlands (n=1122)	MPL (PGp)	Savings, % income in risky investments, owning real estate, financial insurance, voluntary deductible health insurance, Self-employment	
		IT	Savings, % income in risky investments, owning real estate, financial insurance, voluntary deductible health insurance, self-employment	
		GRQ	Savings, % income in risky investments, owning real estate, financial insurance, voluntary deductible health insurance, self-employment	
Dohmen et al. (2019)	Representative sample Germany (n=9325)	GRQ	Holding stocks***, self-employment***	
Schleich et al. (2019)	Representative sample eight EU countries	MPL (PGhigh)		Adoption of LEDs, energy-efficient appliances*, and retrofit measures
Sanou et al. (2018)	Rural households Niger (n=640)	IT		Agricultural technology adaption***
		MPL (SGlow)		Agricultural technology adaption**
		GRQ		Agricultural technology adaption

Castillo et al. (2018)	Rural households Peru (n=9682)	MPL (PGp and PGall)	Asked for informal credit, asked for formal credit*	Age of first pregnancy***, age of marriage*, participation in social organizations**, unhealthy habits***, diseases, purchased seeds***, purchased fertilizer***
Menkhoff and Sakha (2017)	Rural households Thailand (n=760)	MPL (SGsure)	Lottery expenditure, planned lottery expenditure, self-employment, % income in risky investments, planning to invest in business***, borrowing, engaging in risk hedging activities, number of insurance contracts, having health insurance	Farming***, BMI
		IT	Lottery expenditure, planned lottery expenditure**, self-employment**, % income in risky investments***, planning to invest in business, borrowing, engaging in risk hedging activities, number of insurance contracts, having health insurance	Farming, BMI
		GRQ	Lottery expenditure**, planned lottery expenditure, self-employment**, % income in risky investments, planning to invest in business, borrowing, engaging in risk hedging activities, number of insurance contracts, having health insurance	Farming, BMI
Beauchamp et al. (2017)	Twin registry Sweden (n=11000)	GRQ	$Portfolio risk for retirement wealth***, share of equity in stock of assets***, having/had own business***	Alcohol consumption***, smoking***
Galizzi et al. (2016)	Representative sample UK (n=661)	MPL (PGp)	Amount of savings, saving regularly, saving time horizon, having a private pension fund	Smoking, junk food consumption, Fruit and vegetables consumption*, alcohol consumption, BMI**
		GRQ	Amount of savings**, saving regularly, saving time horizon, having a private pension fund	Smoking, junk food consumption, fruit and vegetables consumption, alcohol consumption, BMI
Verschoor et al. (2016)	Rural households Uganda (n=1803)	IT		Purchasing fertilizer***, growing of cash crops
Qiu et al. (2014)	Representative sample USA (n=432)	MPL (PGp)		Conducted energy retrofits to home*, purchased energy efficient appliances**, purchased energy efficient air-conditioning
Hardeweg et al. (2013)	Rural Households Thailand (n=934)	MPL (SGsure)	Lottery expenditure*, self-employment	
		GRQ	Lottery expenditure***, self-employment***	

Kim and Lee (2012)	Representative sample South–Korea (n=7553)	MPL (PGpall)	Self–employment*	
Dohmen et al. (2011)	Representative sample Germany (n=22019)	GRQ	Self–employment***, investment in stocks***	Active sports***, smoking***
Jaeger et al. (2010)	General population Germany (n=10115)	GRQ		Migrating*** (movingfrom one region to another)
Ding et al. (2010)	Students China (n=121)	MPL (SGsure)	Buying lottery tickets, buying stocks	Drinking, smoking, taking vitamin pills, rock climbing***, exam preparation
		GRQ	Buying lottery tickets, buying stocks***	Drinking*, smoking**, taking vitamin pills, rock climbing***, exam preparation**
Anderson and Mellor (2008)	General population USA (n=1091)	MPL (PGp)		Smoking*, heavy drinking*, obesity**, seat belt non–use*, speeding
Lusk and Coble (2005)	Students USA (n=50)	MPL (PGp)		Willingness to eat*, purchase*, and accept* genetically modified food

Note: The table only includes measures discussed in this paper and does not report other risk elicitation methods that authors may have used. With the exception of Kim and Lee (2012) and Qiu et al. (2014), all papers use incentives for the MPL and IT elicitation. We also report the type of MPL that was employed. PGp, PGall, PGpall are Paired Gamble formats that vary probabilities, all prizes, or both respectively), while SGlow and SGsure are Standard Gamble formats that vary the low or sure payment respectively. Sample size is full sample reported by the authors; regression specifications reported in the papers may include a smaller number of observations. We indicate that a sample is representative if this is indicated as such by the authors. $ indicates observed field behavior, whereas all other variables are based on stated field behavior. Stars indicate whether the inferred risk preferences significantly correlate with a given field behavior variable in regressions reported in the respective paper: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

From surveying the literature, the following observations can be made. First, the studies are very fragmented in both the specific method applied and the types of field behavior focused on. The majority of papers listed in Table 3 elicit risk preferences using the MPL method and/or the GRQ. Additionally, several papers adopt IT. To our best knowledge, no papers have to date been published that correlate field behavior with risk preferences measured using the CTB or the DB. A potential explanation for not using the CTB is that the method is relatively complex, and researchers may be hesitant to use it for surveys with a general population sample. In addition, compared

to the MPL and IT, this measure has been developed relatively recently. A potential explanation for the lack of studies that correlate decisions from the DB method with field behavior is that its implementation is relatively cumbersome compared to other measures, as it requires the development of an interactive graphical interface and heavy engagement by participants.

Second, several different formats of the MPL method are used. Many papers use the format popularized by HL, where participants receive a list of paired gambles with varying probabilities (PGp). Other papers use a standard gamble format, in which participants must choose between a certain payoff and a risky gamble, with a varying sure payoff (SGsure), varying low payoff in the gamble (SGlow), a paired gamble format where all payoffs are varied (PGall), or a paired gamble where only the high payoff in the gamble is varied (PGhigh). As discussed above, it is important to carefully consider which format is used, because elicited risk preferences may be confounded with probability weighting. Several possibilities to minimize the effect of probability weighting are feasible. For instance, Csermely and Rabas (2016) suggest a format that is composed of paired gambles where the high payoff is varied (PGhigh). This has been implemented by one of the surveyed studies (Schleich et al., 2019). Alternatively, probabilities in the MPL could be varied in order to explicitly identify probability weighting. This approach has not been applied in any of the surveyed papers.

Third, in all but one study, the field behavior variables considered are self-reported. This constitutes a potential problem for testing external validity as it is unclear how self-reported field behavior correlates with actual field behavior. This makes reported (non-)correlations between risk preference measures and field behavior variables difficult to interpret. The only exception is Beauchamp et al. (2017), who take advantage of a reform in the Swedish pension system that introduced individualized pension saving accounts. The new system requires all Swedes to decide how to invest part of their pension savings and to construct an investment portfolio. This allowed the researchers to correlate risk preference measures with a variable that measures the average risk level of the funds chosen by the individual. Their results indicate that the GRQ is a statistically significant predictor for the risk that people take in their portfolio. The authors did not elicit any incentivized measures.

Finally, looking at the variety of studies and the variables that significantly correlate with risk preference measures as reported in the papers, no clear picture emerges with respect to correlations with financial decision-making in the field. In studies that elicit risk preferences with incentivized MPLs of various formats, no relationship is found for savings (Charness, Garcia, Offerman, & Villeval, 2020; Galizzi et al., 2016), borrowing and applying for informal credit (Castillo et al., 2018; Menkhoff

& Sakha, 2017), buying stocks (Charness et al., 2020; Ding et al., 2010), and having insurance (Charness et al., 2020; Menkhoff & Sakha, 2017). However, other papers find marginally significant correlations for buying lottery tickets (Hardeweg et al., 2013) and a statistically significant relationship with the plans to invest in business (Menkhoff & Sakha, 2017). A single paper reports a marginally significant correlation with self-employment (Kim & Lee, 2012), but several others report no relationship (Charness et al., 2020; Hardeweg et al., 2013; Menkhoff & Sakha, 2017).

Also in studies that elicit the GRQ a somewhat ambiguous picture emerges. The GRQ significantly correlates with whether people hold stocks (Ding et al., 2010; Dohmen et al., 2011; Dohmen et al., 2019) and with the share of equity in stocks and portfolio risk (Beauchamp et al., 2017), but not with the percentage of income in risky investments (Charness et al., 2020; Menkhoff & Sakha, 2017). It also correlates significantly with being self-employed (Beauchamp et al., 2017; Dohmen et al., 2011; Dohmen et al., 2019; Hardeweg et al., 2013; Menkhoff & Sakha, 2017), except in Charness et al. (2020). The results are less clear for savings, where Galizzi et al. (2016) find a significant relationship while Charness et al. (2020) do not. This is also the case for lottery expenditures where a significant relationship is found in rural Thailand (Hardeweg et al., 2013; Menkhoff & Sakha, 2017) but not for Chinese students (Ding et al., 2010). No relationship is found for borrowing (Menkhoff & Sakha, 2017) and having insurance (Charness et al., 2020; Menkhoff & Sakha, 2017). Overall, however, it seems that the GRQ correlates more often with self-reported field behavior than MPLs do.

The IT method is applied less often, which makes it difficult to draw conclusions. Menkhoff and Sakha (2017) find that it significantly correlates with self-employment, percentage of income invested in risky investments, and planned lottery expenditure, but not with current lottery expenditure, borrowing, and having insurance. On the other hand, Charness et al. (2020) do not find any relationship with self- employment, percentage of income in risky investments, saving, and having insurance.

Overall, the economics literature is so far inconclusive regarding the link between risk preference measures and field behavior. The measure that correlates most often is the GRQ. Importantly, the GRQ is always asked in the same way, whereas researchers have not yet agreed on the optimal form of MPLs to be used. As argued above, the variety of forms used may negatively affect the precision of measurement of individual risk preferences. Moreover, it should be noted that in almost all studies that use the self-reported GRQ, also field behavior is self-reported. It is, therefore, impossible to rule out that the observed correlations are spurious and in fact reflect a consistency bias in self-reports. The other measures discussed in Section 3 have not been studied sufficiently to draw conclusions regarding their external validity.

5. Age effects

In this section, we discuss the relationship between risk preferences and age. Given that the new pension agreement in the Netherlands requires pension funds to consider the risk attitudes of different age groups in their investment strategies (Koolmees, 2020), it is important to explore whether a link exists between age and risk preferences. In the following, we will focus on studies with representative population samples, except for a large study on Dutch pension participants.[19] For completeness' sake, we also include some papers that use measures that have not been discussed in Section 3. Table 4 provides an overview of the surveyed literature. We report the presence and direction of a relation between age and risk aversion in column 4 and discuss results for specific age groups in the comments if reported by the authors.

The table provides a very clear picture with respect to stated preference questions. All papers that include survey questions that measure the general willingness to take risk report a positive relation between age and risk aversion. Thus, older people report less willingness to take risks. Three additional observations are relevant. First, there is some evidence that the effect is not linear over the entire life course. Dohmen, Falk, Golsteyn, Huffman, and Sunde (2017) take advantage of a panel dataset and report that the positive relationship between age and risk aversion gets weaker for people over 65. Using cross-sectional data, Alserda et al. (2019) similarly find that the positive relation between age and risk aversion decreases at higher age. Second, there is some evidence of substantial heterogeneity in the relation between age and risk aversion with regard to socio-economic status. Schurer (2015) reports a positive relation between age and risk aversion up to age 40 for all socio-economic groups. After age 40, however, this relationship only remains for those at the bottom of the income and education ladder, while risk aversion does not change or even decreases for those at the top of the income and education ladder. Third, the age effect remains after controlling for birth cohort effects (Dohmen et al., 2017; Schurer, 2015).

The picture is less clear-cut for experimental measures. Looking at papers that included incentives, one can see that the majority of studies do not report any relationship between age and risk aversion. More specifically, eight papers report no relation between age and risk preferences, while only three papers report a significantly positive correlation with risk aversion. In addition, two papers report a significant

19 See also König (2020), who reviews the literature specifically for self-reported risk attitudes including smaller studies, and Zilker, Hertwig, and Pachur (2020), who review the literature also for lotteries in the loss and gain/loss domain.

quadratic relation, and one paper finds significant differences depending on some parameters in the task. Table 4 covers a wide range of experimental measures that may not always be comparable. However, when comparing papers, no clear pattern is observed when controlling for the format and type of measurement.

Table 4. Overview of studies investigating the heterogeneity of risk preferences with respect to age.

Paper	Sample	Measure	Incentives	Age Effect on Risk Aversion	Comments
Frey et al. (2020)	Germany (n=916)	8 lottery questions	X	None	
		GRQ		Positive	
Dohmen et al. (2019)	Germany (n=9325)	GRQ		Positive	
Alserda et al. (2019)	Pension plan participants Netherlands (n=7894)	MPL (PGp)		Positive (quadratic)	Quadratic relationship: the positive effect declines with age.
		2 survey questions on pension risk		Positive (quadratic)	Quadratic relationship: the positive effect declines with age.
Boschini et al. (2019)	Sweden (n=997)	MPL (SGsure)	X	None	
Chapman et al. (2018)	USA (n=2000)	DOSE	X	None	DOSE is a staircase method with Bayesian updating, proposed by the authors.
Chapman et al. (2018)	USA (n=1000)	Composite measure 1	X	None	Based on PCA, the authors combine seven risk measures into two composite measures: 1. MPL gains (SGsure), MPL losses (SGsure), MPL gains/losses (SGsure), willingness to accept. 2. Common ratio certain, common ratio lottery, willingness to pay.
		Composite measure 2	X	Positive	
Falk et al. (2018)	76 countries (n=80000)	Risk pref-erence sur-vey module (Falk et al., 2016)		Positive	Risk preference survey module combines a MPL staircase method with the GRQ in a composite measure.
Andersen et al. (2018)	Denmark (n=413)	MPL (PGp)	X	Positive	Young (<40) are less risk-averse compared to old (>=40).
Dohmen et al. (2017)	Netherlands (n=35173)	6 survey questions		Positive	Controlling for cohort effects, the authors find in both samples that after age 65 the willingness to take risk keeps decreasing but with a flatter slope.
	Germany (n=120837)	GRQ		Positive	
Lee and Kang (2016)	Korea (n=1086)	MPL (SGp)		Positive	

Mamerow et al. (2016)	Switzerland (n=973)	MPL (SGall)	X	Positive / Negative	Age effect depends on framing of task.
		Balloon analogue risk task (Lejuez, Aklin, Zvolensky, & Pedulla, 2003)	X	None / Positive	Age effect depends on framing of task.
		GRQ	Positive		
Josef et al. (2016)	Germany (n=433)	MPL (SGsure)	X	Quadratic	Quadratic relationship U-shaped: people first become less risk averse, then more risk averse, with evident increases from age 30.
	(n=44076)	GRQ		Positive	
Galizzi et al. (2016)	UK (n=661)	MPL (PGp)	X	None	
		GRQ		Positive	
		Gamble choice task (Eckel & Grossman, 2002, 2008)	X	None	
Andersson et al. (2016)	Denmark (n=2289)	MPL (PGhigh)	X	None	Age group >65 is significantly less risk-averse in some specifications.
Schurer (2015)	Germany (n=36105)	GRQ		Positive	Clear age group effects. Compared to 36–40, the age groups 31–35, 26–30, 20–25, <20 are less risk-averse with increasing effect sizes, the age groups 41–45, 46–50, 51–55, 56–60, 61–65, 66–70, 71–75, >76 are more risk-averse with increasing effect sizes. However, clear heterogeneity is observed among different socio-economic groups. After mid-age (40) only disadvantaged groups (in terms of socio-economic status) become more risk-averse. Willingness to take risk roughly stabilizes for more advantaged groups.
Qiu et al. (2014)	USA (n=432)	MPL (PGp)		None	
Kim and Lee (2012)	South Korea (n=7553)	MPL (PGpall)		Positive	
Von Gaudecker et al. (2011)	Netherlands (n=1422)	MPL (PGp)	X	Positive	No significant differences between age groups 18–34, 35–44, 55–64. Age groups 45–54 and 65+ significantly more risk-averse, with the largest effect for the latter group.

Dohmen et al. (2011)	Germany (n=22019)	GRQ		Positive	
Jaeger et al. (2010)	Working population Germany (n=10115)	GRQ		Positive	
Dohmen et al. (2010)	Germany (n=452)	MPL (SGsure)	X	Negative (quadratic)	Quadratic relationship: the negative effect declines with age.
	(n=1012)	GRQ		Positive	
Harrison et al. (2007)	Denmark (n=253)	MPL (PGp)	X	None	Age group 40–50 less risk–averse than those older and younger. No significant differences for other age groups (<30, 30–39, >50)
Donkers et al. (2001)	Netherlands (n=3949)	8 lottery questions		Positive	

6. Discussion and Conclusion

The purpose of this paper was to provide an overview of the concept of risk preferences and how they are measured in the academic literature. To that extent, we discussed the two most prominent economic theories of decision-making under risk – *Expected Utility Theory* (EUT) and *Prospect Theory* (PT). Next, we discussed a number of experimental tasks that have been proposed in the economic literature for measuring risk preferences. Three of these tasks have already been proposed explicitly for applications in the financial sector (Alserda et al., 2016, 2019; Dellaert et al., 2016; Potters et al., 2016; Van der Meeren et al., 2019). First, there is the Multiple Price List (MPL), a method where people receive a list of pairs of lotteries and need to decide for each of them between a safer and a riskier option. Second, we have the Convex Time Budget (CTB), a method where people receive a monetary budget and need to decide how to divide this over a safe early account and a potentially risky late account. The third method we discussed is the Distribution Builder (DB), which asks participants to create a preferred probability distribution over final wealth outcomes, for instance income during retirement. Additionally, we discussed the Investment Task (IT), a method explicitly designed for the investment context. Finally, we reviewed studies that use the General Risk Question, a stated preference method where people are asked to self-report their willingness to take risk.

In order to transfer these risk preference measures to actual practice, such as in the pension field, and to assess their applicability for eliciting the risk preferences of individual persons, an important criterion is their correlation with field behavior. Surveying the literature, we observe the following. First, there are no papers so far that investigate the external validity of risk preferences measured by the DB and CTB. Second, correlations between the two other preference measures, MPL and IT, and field behavior vary in different domains, and many studies do not find any significant link between MPL measures and financial behaviors. Finally, the GRQ correlates overall most often with self-reported field behavior.

An open question is why correlations between revealed preference measures and field behavior are observed infrequently. One potential explanation is that most papers consider *stated* field behavior that is inferred from survey questions rather than from observation of actual behavior. This may also partly explain the differences found between behavioral measures and the GRQ, which is also a survey question. More generally, given the variety of implemented MPL formats and the specific field behavior variables analyzed in the existing literature, it is difficult to draw strong conclusions. More systematic evidence is needed that links different forms of MPLs both

with survey data and objective (e.g. administrative) data in order to establish robust links between risk preference measures and decisions in the field.

Another possible explanation is that most experimental studies involve low stake decisions. The inferred risk preference parameter from small stake experiments may not translate one-to-one to economic and financial decisions that involve much higher stakes. Indeed, some authors deal with this issue by eliciting risk preferences using much higher hypothetical stakes (for instance, Alserda et al., 2019). The disadvantage is that experiments with high stakes are too costly to incentivize, and people may behave differently when their choices are hypothetical than when they are real. In our view, both incentivized experiments with small stakes and hypothetical experiments with high stakes have merit, and ideally both are used, allowing the researcher to compare behavior in both approaches.

Recent advances in the literature point to other possible explanations for low correlations between preferences measures and field behavior. Gillen, Snowberg, and Yariv (2019) stress the importance of measurement error in experimental elicitation. Measurement errors may arise because of inattention and rounding in finite choice menus. A related concern is that, if people make random choices, then the inferred preference may depend on the format of the task. Andersson et al. (2016) show this for two MPLs (PGhigh), where they vary the row of the risk-neutral switching point. They find that cognitive ability is positively correlated with risk aversion in one task but negatively in the other, thus indicating a spurious relationship. Following this, Andersson, Holm, Tyran, and Wengström (2020) stress the importance of a balanced design (multiple MPLs with different risk-neutral switching points) and of controlling for decision errors in structural estimations.

The concern about the format of the task used is more general, as already briefly discussed above. Some elicitation methods may confound risk preferences with probability weighting (Drichoutis & Lusk, 2016). Risk preferences from other elicitation methods may be confounded with loss aversion, especially when a method includes a specific reference point (Goldstein et al., 2008). In those cases, it is important that parameters of the underlying probability weighting and utility or value function are estimated simultaneously, such that the estimate accounts for the potentially confounding effect. Specifically, one should avoid using these methods to directly infer the CRRA risk aversion parameter, for instance from the number of safe choices made in an MPL. Moreover, note that a measure that confounds risk preferences with probability weighting and/or loss aversion may be a better predictor of field behavior than a measure which does not, because field behavior is driven by both risk aversion, loss aversion, and probability weighting. A stronger correlation with field behavior

therefore does not directly identify a better measure of risk preferences for use in pension and other practice.

Another important criterion for transferring risk preference measures to practice and to assess their applicability for eliciting risk preferences of individual persons is the cross-validity of measures. Most of the experimental methodologies are designed explicitly for measuring risk preferences under EUT. If the elicitation were done in a context that minimizes measurement biases, then one would expect that estimations in different methods correlate with each other. However, there is evidence of low correlation between various types of behavioral measures (Charness et al., 2020; Csermely & Rabas, 2016; Deck, Lee, Reyes, & Rosen, 2013; Ding et al., 2010; Frey, Pedroni, Mata, Rieskamp, & Hertwig, 2017; Galizzi et al., 2016; He, Veronesi, & Engel, 2018; Loomes & Pogrebna, 2014; Menkhoff & Sakha, 2017; Pedroni et al., 2017; Reynaud & Couture, 2012).

We have also provided an overview of the literature on the relationship between age and risk preference. Given the importance that the pension system reform attaches to this, it is important to note that the literature shows by and large that higher age is associated with a lower willingness to take risks. In addition, many studies have linked risk preference measures to a variety of demographic characteristics (such as gender and education level) and found significant effects (see, for instance, Dohmen et al., 2011, and, Alserda et al. 2019, for a study that focuses on the pension context). Whereas some observable background characteristics, such as gender, seem to exhibit a relatively robust correlation with risk preference measures, demographics only explain part of the individual heterogeneity in the measures for risk preferences. Hence, it might indeed be worthwhile for policymakers and the finance industry to take the heterogeneity of individuals into account, which requires adequate measures for risk preferences.

The current state of the academic literature does not provide a clear answer as to which measurement method is optimal for the elicitation of risk preferences for the financial sector. As such, more structural research is needed into (1) the empirical relation between different elicitation methods, (2) their relation to alternative theories of decision-making under risk, (3) their relation to not only stated but also actual field behavior in large representative samples, (4) the relation between elicitation methods that use low real stakes, high hypothetical stakes and high real stakes, and (5) the relation between preferences when probabilities are known (risk) and when they are unknown (uncertainty), among others. Only in this way can further steps be made towards a clear understanding of which measures are most suited to identify the underlying risk preferences of decision makers.

Nevertheless, following recent advances in the literature, an important recommendation for applied risk preferences elicitation can for now be made. Elicitation of risk preferences requires the use of stated preferences questions, together with several incentivized experimental measures, and needs to control for decision and measurement errors in the structural estimation of parameters.

References

Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. *Management Science, 46*(11), 1497–1512. doi:10.1287/mnsc.46.11.1497.12080

Abdellaoui, M., Bleichrodt, H., L'Haridon, O., & Van Dolder, D. (2016). Measuring Loss Aversion under Ambiguity: A Method to Make Prospect Theory Completely Observable. *Journal of Risk and Uncertainty, 52*(1), 1–20. doi:10.1007/s11166-016-9234-y

Abdellaoui, M., Bleichrodt, H., & Paraschiv, C. (2007). Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. *Management Science, 53*(10), 1659–1674. doi:10.1287/mnsc.1070.0711

Alserda, G. A. G., Dellaert, B. G. C., Swinkels, L., & Van der Lecq, F. S. G. (2016). Pension risk preferences: A personalized elicitation method and its impact on asset allocation. *Netspar Industry Paper Series (Design paper 62).*

Alserda, G. A. G., Dellaert, B. G. C., Swinkels, L., & Van der Lecq, F. S. G. (2019). Individual pension risk preference elicitation and collective asset allocation with heterogeneity. *Journal of Banking & Finance, 101*, 206–225. doi:https://doi.org/10.1016/j.jbankfin.2019.02.014

Andersen, S., Cox, J. C., Harrison, G. W., Lau, M. I., Rutström, E. E., & Sadiraj, V. (2018). Asset Integration and Attitudes toward Risk: Theory and Evidence. *The Review of Economics and Statistics, 100*(5), 816–830. doi:10.1162/rest_a_00719

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics, 9*(4), 383–405. doi:10.1007/s10683-006-7055-6

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting Risk and Time Preferences. *Econometrica, 76*(3), 583–618. doi:10.1111/j.1468-0262.2008.00848.x

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2018). Multiattribute utility theory, Intertemporal utility, and correlation aversion. *International Economic Review, 59*(2), 537–555. doi:https://doi.org/10.1111/iere.12279

Anderson, L. R. & Mellor, J. M. (2008). Predicting health behaviors with an experimental measure of risk preference. *Journal of Health Economics, 27*(5), 1260–1274. doi:https://doi.org/10.1016/j.jhealeco.2008.05.011

Andersson, O., Holm, H. J., Tyran, J.-R., & Wengström, E. (2016). Risk Aversion Relates to Cognitive Ability: Preferences or Noise? *Journal of the European Economic Association, 14*(5), 1129–1154. doi:10.1111/jeea.12179

Andersson, O., Holm, H. J., Tyran, J.-R., & Wengström, E. (2020). Robust inference in risk elicitation tasks. *Journal of Risk and Uncertainty, 61*(3), 195–209. doi:10.1007/s11166-020-09341-6

Andreoni, J. & Sprenger, C. (2012a). Estimating Time Preferences from Convex Budgets. *American Economic Review, 102*(7), 3333–3356. doi:10.1257/aer.102.7.3333

Andreoni, J. & Sprenger, C. (2012b). Risk Preferences Are Not Time Preferences. *American Economic Review, 102*(7), 3357–3376. doi:10.1257/aer.102.7.3357

Arslan, R. C., Brümmer, M., Dohmen, T., Drewelies, J., Hertwig, R., & Wagner, G. G. (2020). How people know their risk preference. *Scientific Reports, 10*(1), 15365. doi:10.1038/s41598-020-72077-5

Barberis, N. C. (2013). Thirty Years of Prospect Theory in Economics: A Review and Assessment. *Journal of Economic Perspectives, 27*(1), 173–196. doi:10.1257/jep.27.1.173

Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *The Quarterly Journal of Economics, 112*(2), 537–579. doi:10.1162/003355397555280

Beauchamp, J. P., Cesarini, D., & Johannesson, M. (2017). The psychometric and empirical properties of measures of risk preferences. *Journal of Risk and Uncertainty, 54*(3), 203–237. doi:10.1007/s11166-017-9261-3

Bernoulli, D. (1738). Exposition of a new theory on the measurement of risk. Translated by L. Sommer. *Econometrica, 22*(1), 23–36.

Bilsen, S. v., Laeven, R. J. A., & Nijman, T. E. (2020). Consumption and Portfolio Choice Under Loss Aversion and Endogenous Updating of the Reference Level. *Management Science, 66*(9), 3927–3955. doi:10.1287/mnsc.2019.3393

Boschini, A., Dreber, A., von Essen, E., Muren, A., & Ranehill, E. (2019). Gender, risk preferences and willingness to compete in a random sample of the Swedish population. *Journal of Behavioral and Experimental Economics, 83*, 101467. doi:https://doi.org/10.1016/j.socec.2019.101467

Castillo, M., Petrie, R., Cotla, C. R., & Torero, M. (2018). *Risk preferences and decision quality of the poor* (Working Paper). Retrieved from http://www.marcocastillo.org/research.html

Chapman, J., Dean, M., Ortoleva, P., Snowberg, E., & Camerer, C. (2018). *Econographics* (NBER Working Paper No. 24931). Retrieved from https://www.nber.org/papers/w24931

Chapman, J., Snowberg, E., Wang, S., & Camerer, C. (2018). *Loss Attitudes in the U.S. Population: Evidence from Dynamically Optimized Sequential Experimentation (DOSE)* (NBER Working Paper No. 25072). Retrieved from https://www.nber.org/papers/w25072

Charness, G., Garcia, T., Offerman, T., & Villeval, M. C. (2020). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *Journal of Risk and Uncertainty, 60*(2), 99–123. doi:10.1007/s11166-020-09325-6

Charness, G., Gneezy, U., & Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization, 87*, 43–51. doi:https://doi.org/10.1016/j.jebo.2012.12.023

Charness, G., & Viceisza, A. (2012). *Comprehension and risk elicitation in the field: Evidence from rural Senegal* (IFPRI discussion paper No. 1135). Retrieved from https://EconPapers.repec.org/RePEc:fpr:ifprid:1135

Cohen, J., Ericson, K. M., Laibson, D., & White, J. M. (2020). Measuring Time Preferences. *Journal of Economic Literature, 58*(2), 299–347. doi:10.1257/jel.20191074

Csermely, T., & Rabas, A. (2016). How to reveal people's preferences: Comparing time consistency and predictive power of multiple price list risk elicitation methods. *Journal of Risk and Uncertainty, 53*(2), 107–136. doi:10.1007/s11166-016-9247-6

Deck, C., Lee, J., Reyes, J. A., & 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 & Organization, 87*, 1–24. doi:https://doi.org/10.1016/j.jebo.2012.11.010

Dellaert, B., Donkers, B., Turlings, M., Steenkamp, T., & Vermeulen, E. (2016). Naar een nieuwe aanpak voor risicoprofielmeting voor deelnemers in pensioenregelingen. *Netspar Industry Paper Series (Design paper 49).*

Ding, X., Hartog, J., & Sun, Y. (2010). *Can we measure individual risk attitudes in a survey?* (IZA Discussion Paper No. 4807). Retrieved from https://www.iza.org/publications/dp/4807/can-we-measure- individual-risk-attitudes-in-a-survey

Dohmen, T., Falk, A., Golsteyn, B. H. H., Huffman, D., & Sunde, U. (2017). Risk Attitudes Across the Life Course. *The Economic Journal, 127*(605), F95–F116. doi:https://doi.org/10.1111/ecoj.12322

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. doi:10.1257/aer.100.3.1238

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association, 9*(3), 522–550. doi:10.1111/j.1542-4774.2011.01015.x

Dohmen, T., Quercia, S., & Willrodt, J. (2019). *Willingness to take risk: The role of risk conception and optimism* (SOEPpapers on Multidisciplinary Panel Data Research 1026). Retrieved from https://ideas.repec.org/p/diw/diwsop/diw_sp1026.html

Donkers, B., Lourenço, C., Dellaert, B., & Goldstein, D. G. (2013). Using Preferred Outcome Distributions to Estimate Value and Probability Weighting Functions in Decisions under Risk. *Netspar Discussion Paper (05/2013-014)*.

Donkers, B., Lourenço, C., Goldstein, D. G., & Dellaert, B. (2013). Building a distribution builder: Design considerations for financial investment and pension decisions. *Netspar Design Paper (20)*.

Donkers, B., Melenberg, B., & Van Soest, A. (2001). Estimating Risk Attitudes using Lotteries: A Large Sample Approach. *Journal of Risk and Uncertainty, 22*(2), 165–195. doi:10.1023/A:1011109625844

Drichoutis, A. C. & Lusk, J. L. (2016). What can multiple price lists really tell us about risk preferences? *Journal of Risk and Uncertainty, 53*(2/3), 89–106. doi:10.2307/45158452

Dustmann, C., Fasani, F., Meng, X., & Minale, L. (in press). Risk Attitudes and Household Migration Decisions. *The Journal of Human Resources.* doi:10.3368/jhr.58.3.1019-10513R1

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. doi:https://doi.org/10.1016/S1090-5138(02)00097-1

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. doi:https://doi.org/10.1016/j.jebo.2008.04.006

Eeckhoudt, L. & Schlesinger, H. (2006). Putting Risk in Its Proper Place. *American Economic Review, 96*(1), 280–289. doi:10.1257/000282806776157777

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global Evidence on Economic Preferences. *The Quarterly Journal of Economics, 133*(4), 1645–1692. doi:10.1093/qje/qjy013

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2016). *The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences* (IZA DP No. 9674). Retrieved from https://www.iza.org/publications/dp/9674/the-preference-survey-module-a-validated-instrument-for-measuring-risk-time-and-social-preferences

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances, 3*(10), e1701381. doi:10.1126/sciadv.1701381

Frey, R., Richter, D., Schupp, J., Hertwig, R., & Mata, R. (2020). Identifying robust correlates of risk preference: A systematic approach using specification curve analysis. *Journal of Personality and Social Psychology.* doi:10.1037/pspp0000287

Galizzi, M. M., Machado, S. R., & Miniaci, R. (2016). *Temporal stability, cross-validity, and external validity of risk preferences measures: Experimental evidence from a UK representative sample* (Working Paper). Retrieved from The London School of Economics and Political Science, Department of Social Policy, London, UK: http://eprints.lse.ac.uk/67554/

Gillen, B., Snowberg, E., & Yariv, L. (2019). Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study. *Journal of Political Economy, 127*(4), 1826–1863. doi:10.1086/701681

Gneezy, U. & Potters, J. (1997). An Experiment on Risk Taking and Evaluation Periods. *The Quarterly Journal of Economics, 112*(2), 631–645. doi:10.1162/003355397555217

Goldstein, D. G., Johnson, E. J., & Sharpe, W. F. (2008). Choosing Outcomes versus Choosing Products: Consumer-Focused Retirement Investment Advice. *Journal of Consumer Research, 35*(3), 440- 456. doi:10.1086/589562

Hardeweg, B., Menkhoff, L., & Waibel, H. (2013). Experimentally Validated Survey Evidence on Individual Risk Attitudes in Rural Thailand. *Economic Development and Cultural Change, 61*(4), 859-888. doi:10.1086/670378

Harrison, G. W., Lau, M. I., Ross, D., & Swarthout, J. T. (2017). Small stakes risk aversion in the laboratory: A reconsideration. *Economics Letters, 160*, 24–28. doi:https://doi.org/10.1016/j.econlet.2017.08.003

Harrison, G. W., Lau, M. I., & Rutström, E. E. (2007). Estimating Risk Attitudes in Denmark: A Field Experiment. *The Scandinavian Journal of Economics, 109*(2), 341–368. doi:10.1111/j.1467-9442.2007.00496.x

Harrison, G. W. & Rutström, E. E. (2008). Risk Aversion in the Laboratory. In J. C. Cox & G. W. Harrison (Eds.), *Risk Aversion in Experiments* (Vol. 12, pp. 41–196): Emerald Group Publishing Limited.

He, P., Veronesi, M., & Engel, S. (2018). Consistency of Risk Preference Measures: An Artefactual Field Experiment from Rural China. *The Journal of Development Studies, 54*(11), 1955–1973. doi:10.1080/00220388.2017.1336542

Holt, C. A. & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *The American Economic Review, 92*(5), 1644–1655. Retrieved from http://www.jstor.org/stable/3083270

Imai, T., Rutter, T. A., & Camerer, C. F. (2020). Meta-Analysis of Present-Bias Estimation Using Convex Time Budgets. *The Economic Journal.* doi:10.1093/ej/ueaa115

Jaeger, D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., & Bonin, H. (2010). Direct Evidence on Risk Attitudes and Migration. *The Review of Economics and Statistics, 92*(3), 684–689. doi:10.1162/REST_a_00020

Josef, A. K., Richter, D., Samanez-Larkin, G. R., Wagner, G. G., Hertwig, R., & Mata, R. (2016). Stability and change in risk-taking propensity across the adult life span. *Journal of Personality and Social Psychology, 111*(3), 430-450. doi:10.1037/pspp0000090

Kahneman, D. & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263. doi:10.2307/1914185

Kim, Y.-I. & Lee, J. (2012). Estimating Risk Aversion Using Individual-Level Survey Data. *The Korean Economic Review, 28*(2), 221-239.

Kimball, M. S. (1990). Precautionary Saving in the Small and in the Large. *Econometrica, 58*(1), 53–73. doi:10.2307/2938334

Kimball, M. S. (1993). Standard Risk Aversion. *Econometrica, 61*(3), 589-611. doi:10.2307/2951719
Knight, F. H. (1921). *Risk, uncertainty and profit* (Vol. 31): Houghton Mifflin.

König, A. N. (2020). Domain-specific risk attitudes and aging—A systematic review. *Journal of Behavioral Decision Making*, 1–20. doi:https://doi.org/10.1002/bdm.2215

Koolmees, W. (2019). *Principeakkoord vernieuwing pensioenstelsel [Kamerbrief].* Retrieved from https://www.rijksoverheid.nl/documenten/kamerstukken/2019/06/05/kamerbrief-principeakkoord-vernieuwing-pensioenstelsel

Koolmees, W. (2020). *Beantwoording vragen over de uitwerking van het pensioenakkoord [Kamerbrief].* Retrieved from https://www.rijksoverheid.nl/documenten/kamerstukken/2020/07/10/antwoorden-op- kamervragen-over-pensioenakkoord

Landman, W. (2013). Uniformiteit risicoprofielen van particuliere beleggers. *Economisch Statistische Berichten*.

Landman, W. (2019). Meten van risicovoorkeuren beleggers is volstrekt zinloos. *Economisch Statistische Berichten*.

Lee, S. H. & Kang, H. G. (2016). Integrated framework for the external cost assessment of nuclear power plant accident considering risk aversion: The Korean case. *Energy Policy, 92*, 111–123. doi:https://doi.org/10.1016/j.enpol.2016.01.035

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. doi:https://doi.org/10.1016/S0140-1971(03)00036-8

Loomes, G., & Pogrebna, G. (2014). Measuring Individual Risk Attitudes when Preferences are Imprecise. *The Economic Journal, 124*(576), 569–593. doi:10.1111/ecoj.12143

Loonen, T. & Van Raaij, F. (2008). Risicoprofielen onderling weinig consistent. *Economisch Statistische Berichten*.

Lusk, J. L. & Coble, K. H. (2005). Risk Perceptions, Risk Preference, and Acceptance of Risky Food. *American Journal of Agricultural Economics, 87*(2), 393–405. doi:https://doi.org/10.1111/j.1467-8276.2005.00730.x

Mamerow, L., Frey, R., & Mata, R. (2016). Risk taking across the life span: A comparison of self-report and behavioral measures of risk taking. *Psychology and Aging, 31*(7), 711–723. doi:10.1037/pag0000124

Menkhoff, L. & Sakha, S. (2017). Estimating risky behavior with multiple-item risk measures. *Journal of Economic Psychology, 59*, 59–86. doi:https://doi.org/10.1016/j.joep.2017.02.005

Noussair, C. N., Trautmann, S. T., & Van de Kuilen, G. (2013). Higher Order Risk Attitudes, Demographics, and Financial Decisions. *The Review of Economic Studies, 81*(1), 325–355. doi:10.1093/restud/rdt032

Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. *Nature Human Behaviour, 1*(11), 803–809. doi:10.1038/s41562-017-0219-x

Potters, J., Riedl, A., & Smeets, P. (2016). Towards a practical and scientifically sound tool for measuring time and risk preferences in pension savings decisions. *Netspar Industry Paper Series (Design paper 59)*.

Qiu, Y., Colson, G., & Grebitus, C. (2014). Risk preferences and purchase of energy-efficient technologies in the residential sector. *Ecological Economics, 107*, 216–229. doi:https://doi.org/10.1016/j.ecolecon.2014.09.002

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. doi:10.1007/s11238-012-9296-5

Sanou, A., Liverpool-Tasie, L. S. O., & Shupp, R. (2018). Eliciting Risk Attitudes in the Field: Surveys or Experimental Methods? An Empirical Comparison in Rural Niger. *The Journal of Development Studies, 54*(8), 1450–1470. doi:10.1080/00220388.2017.1404034

Schechter, L. (2007). Risk aversion and expected-utility theory: A calibration exercise. *Journal of Risk and Uncertainty, 35*(1), 67–76. doi:10.1007/s11166-007-9017-6

Schleich, J., Gassmann, X., Meissner, T., & Faure, C. (2019). A large-scale test of the effects of time discounting, risk aversion, loss aversion, and present bias on household adoption of energy-efficient technologies. *Energy Economics, 80*, 377–393. doi:https://doi.org/10.1016/j.eneco.2018.12.018

Schneider, S. O. & Sutter, M. (2020). *Higher Order Risk Preferences: New Experimental Measures, Determinants and Field Behavior* (MPI Collective Goods Discussion Paper No. 2020/22). Retrieved from https://ssrn.com/abstract=3683782

Schurer, S. (2015). Lifecycle patterns in the socioeconomic gradient of risk preferences. *Journal of Economic Behavior & Organization, 119*, 482–495. doi:https://doi.org/10.1016/j.jebo.2015.09.024

Sharpe, W. F. (2001). *Individual Risk and Return Preferences: A Preliminary Survey* (Working Paper). Retrieved from https://web.stanford.edu/~wfsharpe/art/rrsurvey/vienna2001.htm Sharpe, W. F., Goldstein, D. G., & Blythe, P. W. (2000). *The Distribution Builder: A Tool for Inferring Investor Preferences* (Working Paper). Retrieved from https://web.stanford.edu/~wfsharpe/art/qpaper/qpaper.html

Starmer, C. (2000). Developments in Non-expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk. *Journal of Economic Literature, 38*(2), 332–382. doi:10.1257/jel.38.2.332

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. doi:10.1257/aer.100.1.557

Trautmann, S. T., & Van de Kuilen, G. (2018). Higher order risk attitudes: A review of experimental evidence. *European Economic Review, 103*, 108–124. doi:https://doi.org/10.1016/j.euroecorev.2018.01.007

Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297–323. doi:10.1007/bf00122574

Van der Meeren, G., de Cloe-Cos, H., & Van Geen, A. (2019). Meet risicobereidheid met een kwantitatieve methode. *Economisch Statistische Berichten.*

Verschoor, A., D'Exelle, B., & Perez-Viana, B. (2016). Lab and life: Does risky choice behaviour observed in experiments reflect that in the real world? *Journal of Economic Behavior & Organization, 128*, 134–148. doi:https://doi.org/10.1016/j.jebo.2016.05.009

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. doi:10.1257/aer.101.2.664

Von Neumann, J. & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Wakker, P. & Deneffe, D. (1996). Eliciting von Neumann-Morgenstern Utilities When Probabilities Are Distorted or Unknown. *Management Science, 42*(8), 1131–1150. doi:10.1287/mnsc.42.8.1131

Zilker, V., Hertwig, R., & Pachur, T. (2020). Age differences in risk attitude are shaped by option complexity. *Journal of Experimental Psychology: General, 149*(9), 1644–1683. doi:10.1037/xge0000741

OVERZICHT UITGAVEN
IN DE SURVEY PAPER SERIE

1. Saving and investing over the life cycle and the role of collective pension funds
Lans bovenberg , Ralph Koijen, Theo Nijman and Coen Teulings

2. What does behavioural economics mean for policy? Challenges to savings and health policies in the Netherlands
Peter Kooreman and Henriëtte Prast

3. Housing wealth and household portfolios in an aging society
Jan Rouwendal

4. Birth is the sessenger of death – but policy may help to postpone the bad news
Gerard van den Berg and Maarten Lindeboom

5. Phased and partial retirement: preferences and limitations
Arthur van Soest and Tunga Kantarci

6. Retirement Patterns in Europe and the U.S. (2008)
Arie Kapteyn and Tatiana Andreyeva

7. Compression of morbidity: A promising approach to alleviate the societal consequences of population aging? (2008)
Johan Mackenbach, Wilma Nusselder, Suzanne Polinder and Anton Kunst

8. Strategic asset allocation (2008)
Frank de Jong, Peter Schotman and Bas Werker

9. Pension Systems, Aging and the Stability and Growth Pact (2008) Revised version
Roel Beetsma and Heikki Oksanen

10. Life course changes in income: An exploration of age- and stage effects in a 15-year panel in the Netherlands (2008)
Matthijs Kalmijn and Rob Alessie

11. Market-Consistent Valuation of Pension Liabilities (2009)
Antoon Pelsser and Peter Vlaar

12. Socioeconomic Differences in Health over the Life Cycle: Evidence and Explanations (2009)
Eddy van Doorslaer, Hans van Kippersluis, Owen O'Donnell and Tom Van Ourti

13. Computable Stochastic Equilibrium Models and their Use in Pension- and Ageing Research (2009)
Hans Fehr

14. Longevity risk (2009)
Anja De Waegenaere, Bertrand Melenberg and Ralph Stevens

15. Population ageing and the international capital market (2009)
Yvonne Adema, Bas van Groezen and Lex Meijdam

16. Financial Literacy: Evidence and Implications for Consumer Education (2009)
Annamaria Lusardi and Maarten van Rooij

17. Health, Disability and Work: Patterns for the Working-age Population (2009)
Pilar García-Gómez, Hans-Martin von Gaudecker and Maarten Lindeboom

18. Retirement Expectations, Preferences, and Decisions (2010)
Luc Bissonnette, Arthur van Soest

19. Interactive Online Decision Aids for Complex Consumer Decisions: Opportunities and Challenges for Pension Decision Support (2010)
Benedict Dellaert

20. Preferences for Redistribution and Pensions. What Can We Learn from Experiments? (2010)
Jan Potters, Arno Riedl and Franziska Tausch

21. Risk Factors in Pension Returns (2010)
Peter Broer, Thijs Knaap and Ed Westerhout

22. Determinants of Health Care Expenditure in an Aging Society (2010)
Marc Koopmanschap, Claudine de Meijer, Bram Wouterse and Johan Polder