Control and Cybernetics

vol. **38** (2009) No. 3

Comparison of linear interpolation and arctan approximation of one-dimensional monotonic utility functions based on experimental data*

by

Natalia Nikolova[1], Kiril Tenekedjiev[1], Fangyan Dong[2] and Kaoru Hirota[2]

[1]Technical University – Varna
1 Studentska Str., 9010 Varna, Bulgaria

[2]Hirota Lab, Dept. Computational Intelligence and Systems Science
Interdisciplinary Graduate School of Science and Engineering
Tokyo Institute of Technology,
G3-49, 4259 Nagatsuta, Midori-ku, Yokohama-city 226-8502, Japan

e-mail: {natalia,kiril}@dilogos.com, {tou,hirota}@hrt.dis.titech.ac.jp

Abstract: Elicitation of utilities is among the most time consuming tasks in decision analysis. We search for ways to shorten this phase without compromising the quality of results. We use the results from an empirical experiment with 104 participants. They elicited 9 inner nodes from their one-dimensional utility function over monetary gains and losses using three elicitation techniques. A specific feature of the results is their interval character, as the elicitators are fuzzy rational individuals. The data is used to construct arctan-approximated and linearly interpolated utilities and to compare the results. We form partial samples with 3, 4 and 5 nodes for each participant and each elicitation method, and again interpolate/approximate the utilities. We introduce goodness-of-fit and deterioration measures to analyze the decrease in quality of the utility function due to reduced data nodes. The analysis, using paired-sample tests, leads to the following conclusions: 1) arctan-approximation is more adequate than linear interpolation over the whole samples; 2) 5 inner nodes are sufficient to construct a satisfactory arctan-approximation; 3) arctan-approximation and linear interpolation are almost equal in quality over the partial samples, but the local risk aversion of the linearly interpolated utility function is of poor quality unlike that of the arctan-approximated utility function.

Keywords: utility function, interpolation, approximation, weighted least square, goodness-of-fit, statistical tests.

*Submitted: February 2008; Accepted: May 2009.

1. Introduction

Utility theory (von Neumann, Morgenstern, 1947) has been applied to problems in industry, policy making, concurrent engineering and economics (French, 1993). It models risky alternatives as lotteries, i.e. sets of disjoint events, associated with a consequence (prize), and its probability. The preferences of a decision maker (DM) are modeled by the utility function $u(.)$, increasing in prize. Expected utility is the quantitative criterion that rational decision analysis is based on.

In most decision problems one has to construct a one-dimensional (1-D) $u(.)$ over continuous prizes. The common procedure is to elicit only several nodes, and then approximate/interpolate the entire function. The first task may be approached by the methods of probability equivalence (PE), certainty equivalence (CE) (Farquhar, 1984), lottery equivalence (LE) (McCord, De Neufville, 1986), uncertain equivalence (UE) (Tenekedjiev et al., 2006), trade-off (TO) (Wakker, Deneffe, 1996). They solve preferential equations via dichotomy, Press et al. (1992). An ideal DM has infinite discriminating abilities and obeys the axioms of rationality, therefore she/he elicits unique estimates (French, Insua, 2000). Real DMs identify uncertainty intervals using triple dichotomy (Tenekedjiev et al., 2004). Their preferences disobey some of the rationality axioms and are partially non-transitive, and such DMs are called fuzzy rational (FRDM), Nikolova et al. (2005). An analytical approximation of $u(.)$ should adequately interpret the typical risk attitude of the DM, represented by her/his local risk aversion function $r(.)$, Pratt (1964). Most analytical forms apply only for specific risk attitude or prize range (Clemen, 1996; Keeney, Raiffa, 1993). Nikolova (2007) presented an arctan form of 1-D $u(.)$ of a FRDM. It applies over gains and losses and its corresponding $r(.)$ reflects the risk attitude of most DMs. Other publications (e.g., Tenekedjiev et al., 2008) compare the arctan-approximation with another analytical form, power approximation.

Here we report on a further investigation of the properties of arctan-approximation. We use the weighted least square method to estimate the unknown parameters of the arctan-approximated utility function, as in Tenekedjiev et al. (2008). It weighs the deviation of the model from the subjective estimate in a given node by the width of the uncertainty intervals. Each elicitation method solves a specific preference equation, but the solutions take the form of uncertainty intervals for the FRDM. Some methods, like CE, UE and TO, generate uncertainty intervals on the abscissa, x, i.e. on prizes. Other methods, like PE and LE, generate uncertain intervals on the ordinate, u, i.e. utility. That is why it is required that analytical forms of the utility function have analytical inverses.

This paper considers two forms of utility function – arctan-approximated and linearly interpolated. We focus on the possibility of constructing $u(.)$ for monotonic preferences using a reduced number of elicited nodes, which can substantially facilitate utility analysis. This is expected to be better than linear

interpolation. We hypothesize that: 1) arctan-approximation represents better the typical risk attitude of the DM and so is more adequate than linear interpolation; 2) five inner nodes suffice to construct satisfactory arctan-approximation; 3) both the arctan-approximated and the linear interpolated utility functions have roughly the same goodness-of-fit measures when constructed over the reduced samples, but the local risk aversion of the former is more precise. Also, its quality is practically the same when the full number of nodes is used.

An empirical experiment was used to test these hypotheses, involving 104 volunteers. With three elicitation methods, CE, LE and UE, they elicited 9 inner nodes of their utility functions over monetary prizes in the interval from -10,000 Bulgarian leva (BGN) to 30,000 BGN. We use weighted least squares method in two modifications (depending on the type of elicited uncertainty intervals), to estimate the parameters of $3 \times 104 = 312$ utility functions of arctan type. All utility and local risk aversion functions were constructed. The counterpart set of utility functions were linearly interpolated on the midpoints of the uncertainty intervals of inner nodes. Similar procedures were performed for samples containing 3, 4 or 5 inner nodes. The goodness-of-fit measures of the model to the data were calculated. We introduced measures of deterioration of the resulting function due to reduction in data points, and employ four tests for paired samples (Bootstrap mean test, Bootstrap median test, sign test, and sign rank test) to prove statistical significance of the differences in precision of the arctan-approximation and linear interpolation.

In what follows, Section 2 gives an overview of the form in which FRDMs elicit utility nodes, major concern being how to interpret the typical risk attitude of the DM. Section 3 presents the content and structure of the empirical study that provides evidence for the three hypotheses regarding the properties of the arctan-approximation. Three appendices contain formal descriptions and algorithms employed in the main text.

2. Constructing monotonic 1-D utility functions

Assume that X is a continuous 1-D set of prizes and an FRDM has monotonically increasing preferences over it ($\succ$ stands for strict preference):

$$x_i \succ x_j \Leftrightarrow x_i > x_j, \text{ for } x_i \in X,\ x_j \in X. \tag{1}$$

Then the most and the least preferred prizes x_{best} and x_{worst} are, respectively, the supremum and the infimum of X. The problem is to construct a 1-D utility function $u(.)$ over the interval $[x_{worst}; x_{best}]$. It is impossible to elicit the utilities of all prizes in the domain, only a couple of nodes can be assessed for subsequent interpolation/approximation. The utility elicitation techniques require from the DMs to solve preference equations between prizes and/or lotteries. They need to change one parameter in this equation until the indifference of options is reached. A general scheme of the equation and the characteristics of the main elicitation techniques were outlined in Tenekedjiev et al. (2006).

Assume that an FRDM elicited $z-2$ inner nodes of $u(.)$ with coordinates (x_{u_l}, u_l), for $l = 2, 3, \ldots, z-1$, where x_{u_l} and u_l are, respectively, a utility quantile and a utility quantile index. The end nodes are known: $(x_{u_1}, u_1) = (x_{worst}, 0)$ and $(x_{u_z}, u_z) = (x_{best}, 1)$.

Methods like PE and LE select several quantiles x_{u_l} $(l = 2, 3, \ldots, z-1)$ and elicit their corresponding quantile indices $\hat{u}_l$. The degree of belief of DMs regarding the quantile index takes the form of (uncertainty) intervals $\hat{u}_l \in [\hat{u}_l^d; \hat{u}_l^u]$ $(l = 2, 3, \ldots, z-1)$. Then the utility function should be constructed on $z > 1$ number of nodes with the following characteristics:

$$\begin{aligned}
&\{(x_l; \hat{u}_l^d; \hat{u}_l^u) \mid l = 1, 2, \ldots, z\},\\
&x_1 < x_2 < ... < x_z,\\
&0 = \hat{u}_1^d \leq \hat{u}_2^d \leq ... \leq \hat{u}_z^d = 1,\\
&0 = \hat{u}_1^u \leq \hat{u}_2^u \leq ... \leq \hat{u}_z^u = 1,\\
&\hat{u}_l^d < \hat{u}_l^u, \text{ for } l = 2, 3, \ldots, z-1.
\end{aligned} \tag{2}$$

The quantile index $\hat{u}_l$ is a random variable, dependent on x_{u_l}, belonging to the interval $[\hat{u}_l^d; \hat{u}_l^u]$. So, in a strictly increasing utility function, the bounds need not only increase, but may also coincide.

Methods like CE and UE select utility quantile indices u_l $(l = 2, 3, \ldots, z-1)$ and elicit their corresponding quantiles $\hat{x}_{u_l}$. Again, the FRDM elicit uncertainty intervals, here of the form $\hat{x}_{u_l} \in [\hat{x}_{u_l}^d; \hat{x}_{u_l}^u]$ $(l = 2, 3, \ldots, z-1)$. Then the utility function is constructed on $z > 1$ nodes with the following characteristics:

$$\begin{aligned}
&\{(\hat{x}_{u_l}^d; \hat{x}_{u_l}^u; u_l) \mid l = 1, 2, \ldots, z\},\\
&\hat{x}_{u_1}^d \leq \hat{x}_{u_2}^d \leq ... \leq \hat{x}_{u_z}^d,\\
&\hat{x}_{u_1}^d = \hat{x}_{u_1}^u \leq \hat{x}_{u_2}^u \leq ... \leq \hat{x}_{u_z}^u = \hat{x}_{u_z}^d,\\
&\hat{x}_{u_l}^d < \hat{x}_{u_l}^u, \text{ for } l = 2, 3, \ldots, z-1,\\
&0 = u_1 < u_2 < ... < u_z = 1.
\end{aligned} \tag{3}$$

The quantile $\hat{x}_{u_l}$ is a random variable, dependent on u_l, belonging to the interval $[\hat{x}_{u_l}^d; \hat{x}_{u_l}^u]$. So, in a strictly increasing utility function, the bounds need not only increase, but may also coincide.

Once a set of elicited nodes is available, the utility function may be interpolated or approximated. The selected method should precisely interpret the true utility function and preserve the risk attitude of each DM. The typical risk attitude is best described by the local risk aversion function $r(x) = -u''(x)/u'(x)$, Pratt (1964). Empirical studies reveal that DMs are risk averse for gains and small losses and their risk aversion decreases with the growth of gains. In the same time they are risk prone for losses and small gains and their risk proneness decreases with the growth of losses, French (1993).

Tenekedjiev et al. (2008) discuss in detail analytical approximation of the utility function in the case of strictly increasing preferences. The procedures

are summarized in Section A1.1 of Appendix 1. Analytical approximation is appropriate if only few elicited nodes are available or if the uncertainty intervals are too large. If the selected mathematical form describes correctly the risk attitude of the FRDM, the analytical construction of the utility function filters the elicitation errors.

Keeney and Raiffa (1993), a rich source of analytical forms, suggest that for risk neutrality $u(x) = x$, for constant risk aversion $u(x) \underset{p}{\sim} 1 - e^{-x/R}$ $(R > 0)$ and $u(x) \underset{p}{\sim} -e^{-cx}$ $(c > 0)$ ($\underset{p}{\sim}$ denoting proportionality). For decreasing risk aversion $u(x) \underset{p}{\sim} \ln(x)$ and $u(x) \underset{p}{\sim} (x+b)^a$ $(b > -x,\ a < 1)$. For constant risk proneness $u(x) \underset{p}{\sim} e^{-cx}$ $(c < 0)$, whereas for decreasing risk proneness $u(x) \underset{p}{\sim} x^2$ and $u(x) \underset{p}{\sim} (x+b)^a$ $(b > -x,\ a > 1)$. Trautmann and Weihs (2006) discuss Harrington's desirability function $u(x) \underset{p}{\sim} e^{-e^{-b-ax}}$. This dependence generates a function with a typical local risk aversion and is applicable to problems, where consequences are gains and losses. It can also be applied in decision problems under strict certainty.

Nikolova (2007) proposed a modified arctan-approximation of the utility function (A1.9). It suggests that if the optimal approximated curve passes through the uncertainty interval of the nodes then it is very likely that the arctan-approximation would reduce the elicitation error, since it used correct prior information for risk attitude. If the optimal approximated curve substantially deviates from the uncertainty intervals, then the arctan-approximation should be replaced by another approximation, as the risk attitude of that DM is not typical.

Tenekedjiev et al. (2008) compare arctan-approximation with power approximation (4), whose local risk aversion is as in (5):

$$u(x) = \frac{(x - x_d + x_0)^a - x_0^a}{(x_u - x_d + x_0)^a - x_0^a}\,, \tag{4}$$

$$r(x) = -\frac{u''(x)}{u'(x)} = \frac{1-a}{x - x_d + x_0}\,. \tag{5}$$

3. Empirical study

3.1. Experimental setup

The empirical experiment was reported in Tenekedjiev et al. (2008): 104 volunteers (university students, taking a quantitative decision analysis course) constructed their utility functions in the interval [-10000 BGN; 30000 BGN] using CE, UE, and LE. The volunteers elicited nine utility quantiles with indices $0.1, 0.2, \ldots, 0.9$ using CE and UE. They also elicited the utilities of nine inner prize values, -6000 BGN, -2000 BGN, 2000 BGN, 6000 BGN, 10000 BGN, 14000 BGN, 18000 BGN, 22000 BGN and 26000 BGN, using LE. Each participant was

interviewed in three 2-hour sessions, one session per method, with 48h between each session. The order of methods and elicited nodes was random for each participant in order to avoid the anchoring effect (Tversky, Kahneman, 1974). An example of data collected from a participant is shown in Table 1.

The data from the j-th DM can be grouped in 12 samples. Samples $\mathrm{CE}_{9,j}$, $\mathrm{UE}_{9,j}$ and $\mathrm{LE}_{9,j}$ containing all the nine inner nodes, elicited respectively using CE, UE and LE, are called *whole samples*. Samples $\mathrm{CE}_{3,j}$ and $\mathrm{UE}_{3,j}$, consisting of nodes with quantile indices 0.2, 0.5 and 0.8, and samples $\mathrm{LE}_{3,j}$ that include 3 nodes with quantiles -2000 BGN, 10000 BGN and 22000 BGN, are called *3-partial samples*. Samples $\mathrm{CE}_{4,j}$ and $\mathrm{UE}_{4,j}$, consisting of nodes with quantile indices 0.2, 0.4, 0.6 and 0.8, and samples $\mathrm{LE}_{4,j}$ of nodes with quantiles -2000 BGN, 6000 BGN, 14000 BGN and 22000 BGN, are called *4-partial samples*. Samples $\mathrm{CE}_{5,j}$ and $\mathrm{UE}_{5,j}$ of nodes with quantile indices 0.1, 0.3, 0.5, 0.7, 0.9, and samples $\mathrm{LE}_{5,j}$ of nodes with quantiles -6000 BGN, 2000 BGN, 10000 BGN, 18000 BGN and 26000 BGN, are called *5-partial samples*. The 3, 4 and 5-partial samples together shall be referred to as *partial samples*.

Table 1. Quantile and quantile index uncertainty intervals of inner nodes of the utility function in the interval [−10000 BGN; 30000BGN] for DM No. 84, elicited using CE, UE and LE. The shaded boxes indicate the values that the DM did not elicit

	l	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
	u_l	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
CE	$\hat{x}^d_{u_l}$	−10000	−7700	−6100	−5800	−4400	−2400	3000	3800	6300	8000	30000
CE	$\hat{x}^u_{u_l}$	−10000	−6100	−4500	−4000	−2000	4000	5400	6000	8300	10200	30000
UE	$\hat{x}^d_{u_l}$	−10000	−6400	−2000	−1600	−400	1000	2700	4400	6500	9900	30000
UE	$\hat{x}^u_{u_l}$	−10000	−4000	4000	6000	2400	3800	5100	7200	8900	12100	30000

	x_l	**−10000**	**−6000**	**−2000**	**2000**	**6000**	**10000**	**14000**	**18000**	**22000**	**26000**	**30000**
LE	$\hat{u}^d_l$	0	0.03	0.12	0.16	0.40	0.63	0.72	0.85	0.89	0.92	1
LE	$\hat{u}^u_l$	0	0.09	0.20	0.28	0.56	0.81	0.88	0.95	0.97	0.98	1

Assume that the goodness-of-fit measure (A1.4) is

$$\chi^2_u(I^{test}, I^{par}, form). \tag{6}$$

It is calculated using sample I^{test} of type (2) and analytical dependence (A1.1) of type $form$ with parameters $\vec{p}^{I^{par}}_{opt,form} = (p^{I^{par}}_{1,opt,form}, p^{I^{par}}_{2,opt,form}, \ldots, p^{I^{par}}_{n,opt,form})$, identified by optimization on the sample I^{par} of type (2). Similarly, the goodness-of-fit measure (A1.8) is:

$$\chi^2_x(I^{test}, I^{par}, form). \tag{7}$$

It is calculated using sample I^{test} of type (3) and analytical dependence (A1.1) of type $form$ with parameters $\vec{p}^{I^{par}}_{opt,form} = (p^{I^{par}}_{1,opt,form}, p^{I^{par}}_{2,opt,form}, \ldots, p^{I^{par}}_{n,opt,form})$, identified by optimization using the sample I^{par} of type (3).

For the arctan-approximated utility, *form* is *arc*, $n = 2$, $p^{I^{par}}_{1,opt,arc} = a^{I^{par}}_{opt}$, $p^{I^{par}}_{2,opt,arc} = x^{I^{par}}_{0,opt}$. For the utility linearly interpolated on the midpoints of the uncertainty intervals in I^{par}, *form* is *lin*, $n = z - 2$, where z is the number of nodes in I^{par}. For (6), $p^{I^{par}}_{l-1,opt,lin} = (\hat{u}^{d,I^{par}}_l + \hat{u}^{u,I^{par}}_l)/2$, for $l = 2, 3, \ldots, z-1$, whereas for (7), $p^{I^{par}}_{l-1,opt,lin} = (\hat{x}^{d,I^{par}}_l + \hat{x}^{u,I^{par}}_l)/2$, for $l = 2, 3, \ldots, z-1$.

3.2. Linear interpolation versus arctan-approximation over the whole samples

This section provides a qualitative proof of the first hypothesis, formulated in the Introduction.

For each DM we can construct three utility functions using linear interpolation on the midpoints of the uncertainty intervals in $CE_{9,j}$, $UE_{9,j}$ and $LE_{9,j}$. The resulting graphics for $j = 84$ are shown in Fig. 1a, b, c (upper section). The local risk aversion function of the linearly interpolated utilities may be constructed using the procedures from section A1.2 of Appendix 1. The lower sections of Fig. 1a, b, c depict the results. The rough shape of the local risk aversion is a typical result from the differentiation of unsmoothed functions. The reduction of the number of points makes the interpolation rather rough, and its local risk aversion becomes practically useless (see Fig. 2a, b, c).

The data in $CE_{9,84}$, $UE_{9,84}$ and $LE_{9,84}$ are used to arctan-approximate $u(.)$ using analytical form (A1.9) with parameters minimizing (A1.4) and (A1.8). The optimization is executed using MATLAB functions *optparam_u* and *optparam_x*. The resulting utility functions and their local risk aversion functions for DM No. 84 are presented in Fig. 3. A visual comparison of Figs. 1 and 3 reveals that for the analyzed samples, the arctan-approximation adequately encapsulates the elicited nodes, and the corresponding local risk aversions are of much higher quality than those of the linearly interpolated function. Similar results were obtained for all the experiment participants.

We can conclude the following: 1) linear interpolation of $u(.)$ does not account for the width of uncertainty intervals; 2) linearly interpolated local risk aversion function is of no practical use; 3) arctan-approximation fits to the elicited nodes and the average goodness-of-fit measures (6) and (7) are satisfactory; 4) local risk aversion of the arctan-approximated $u(.)$ is of much higher quality than of the linearly interpolated $u(.)$.

Thus, if the arctan-approximation fits well to the data, it should be preferred over the linear interpolation. Otherwise, the risk attitude of the FRDM is not typical and perhaps the best option is linear interpolation.

3.3. Arctan-approximation over the partial samples

This section provides quantitative proof of the second hypothesis from the Introduction. If there is an appropriate analytical form of $u(.)$, then it can be

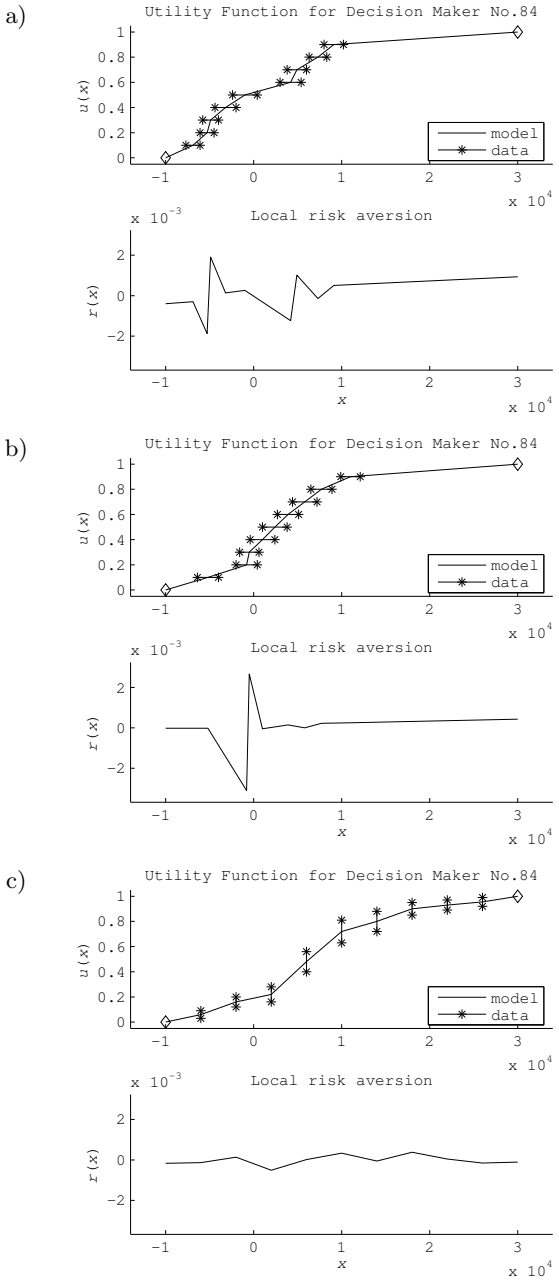

Figure 1. Linearly interpolated utility function (up) and local risk aversion function (down) for DM No. 84 using: a) CE; b) UE; c) LE, all on nine inner nodes. The uncertainty intervals of the elicited nodes are depicted by horizontal/vertical lines.

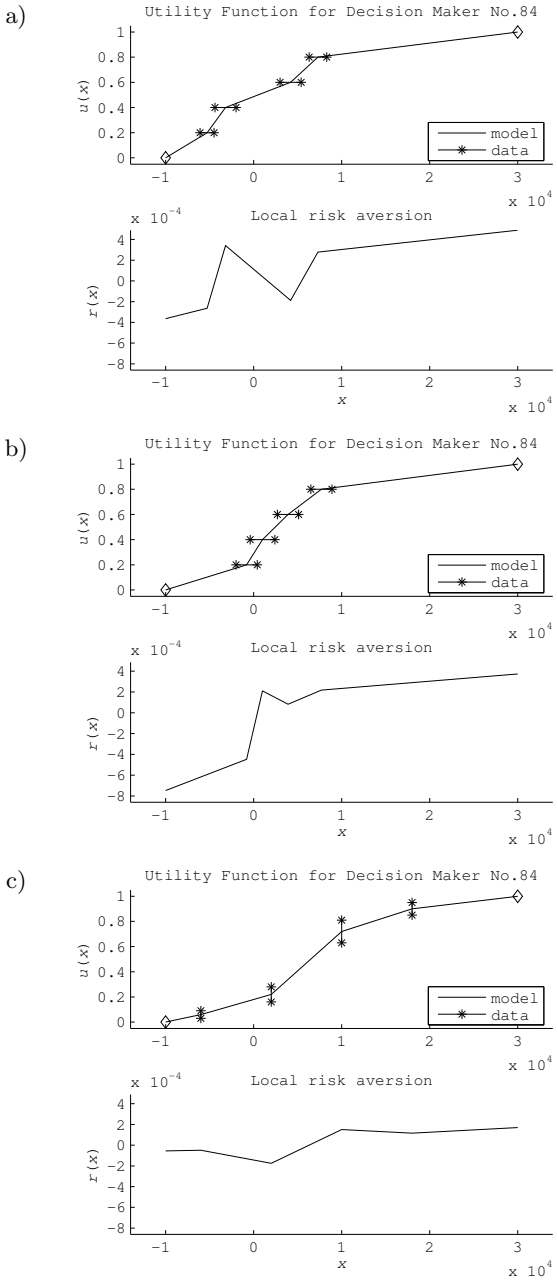

Figure 2. Linearly interpolated utility function (up) and local risk aversion function (down) for DM No. 84 using: a) CE; b) UE; c) LE, all on four inner nodes. The uncertainty intervals of the elicited nodes are depicted by horizontal/vertical lines.

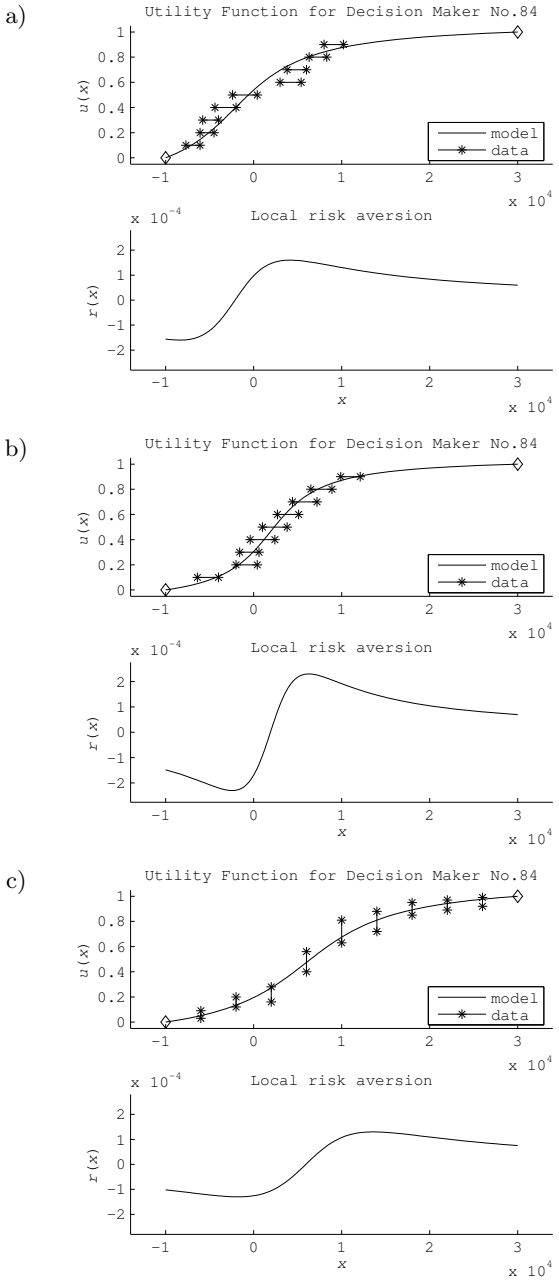

Figure 3. Arctan-approximated utility function (up) and local risk aversion function (down) for DM No. 84 using: a) CE; b) UE; c) LE, all on nine inner nodes. The uncertainty intervals of the elicited nodes are depicted by horizontal/vertical lines

adequately constructed using a small number of data points. To test this proposition, we estimated the parameters of the utility function over the partial samples. For example, the following optimal parameters and their corresponding minimal values of the measures (A1.4) and (A1.8) were estimated for the data in the partial samples:

a) for the 3-partial samples:

$$a_{opt}^{CE_{3,84}} = 1.3\text{e-}4, \quad x_{0,opt}^{CE_{3,84}} = -2673, \quad \chi_x^2(CE_{3,84}, CE_{3,84}, arc) = 1.3,$$
$$a_{opt}^{UE_{3,84}} = 2.5\text{e-}4, \quad x_{0,opt}^{UE_{3,84}} = 2555, \quad \chi_x^2(UE_{3,84}, UE_{3,84}, arc) = 0.043,$$
$$a_{opt}^{LE_{3,84}} = 1.2\text{e-}4, \quad x_{0,opt}^{LE_{3,84}} = 4749, \quad \chi_u^2(LE_{3,84}, LE_{3,84}, arc) = 0.039;$$

b) for the 4-partial samples:

$$a_{opt}^{CE_{4,84}} = 1.2\text{e-}4, \quad x_{0,opt}^{CE_{4,84}} = -2353, \quad \chi_x^2(CE_{4,84}, CE_{4,84}, arc) = 2.5,$$
$$a_{opt}^{UE_{4,84}} = 2.4\text{e-}4, \quad x_{0,opt}^{UE_{4,84}} = 2412, \quad \chi_x^2(UE_{4,84}, UE_{4,84}, arc) = \mathit{0.087},$$
$$a_{opt}^{LE_{4,84}} = 1.1\text{e-}4, \quad x_{0,opt}^{LE_{4,84}} = 5450, \quad \chi_u^2(LE_{4,84}, LE_{4,84}, arc) = \mathit{0.0058};$$

c) for the 5-partial samples:

$$a_{opt}^{CE_{5,84}} = 1.9\text{e-}4, \quad x_{0,opt}^{CE_{5,84}} = -2229, \quad \chi_x^2(CE_{5,84}, CE_{5,84}, arc) = 4.4,$$
$$a_{opt}^{UE_{5,84}} = 2.3\text{e-}4, \quad x_{0,opt}^{UE_{5,84}} = 1528, \quad \chi_x^2(UE_{5,84}, UE_{5,84}, arc) = 0.29,$$
$$a_{opt}^{LE_{5,84}} = 1.5\text{e-}4, \quad x_{0,opt}^{LE_{5,84}} = 6304, \quad \chi_u^2(LE_{5,84}, LE_{5,84}, arc) = 0.28.$$

For all three methods, the resulting parameters slightly deviate from those estimated using whole samples. The resulting utilities for the 3-partial and 5-partial samples are depicted in Figs. 4 and 5. A dotted line presents the utility function and the local risk aversion function. The latter are constructed using $CE_{9,j}$, $UE_{9,j}$, and $LE_{9,j}$. The figures show negligible deterioration of quality due to the reduction of the number of nodes.

The measures (A1.4) and (A1.8) are calculated over the whole samples, with parameters acquired from the partial samples in order to get a quantitative estimate of the deterioration.

A possible absolute measure of deterioration due to the reduced number of nodes is the difference between the calculated values of (A1.4) or (A1.8), and their minimal values (i.e. the optimal parameters calculated using the whole samples for $r = 3, 4, 5$; $j = 1, 2, \ldots, 104$:

$$\begin{aligned}
\Delta_{abs}^{CE_{r,j}} &= \chi_x^2(CE_{9,j}, CE_{r,j}, arc) - \chi_x^2(CE_{9,j}, CE_{9,j}, arc),\\
\Delta_{abs}^{UE_{r,j}} &= \chi_x^2(UE_{9,j}, UE_{r,j}, arc) - \chi_x^2(UE_{9,j}, UE_{9,j}, arc),\\
\Delta_{abs}^{LE_{r,j}} &= \chi_u^2(LE_{9,j}, LE_{r,j}, arc) - \chi_u^2(LE_{9,j}, LE_{9,j}, arc).
\end{aligned} \tag{8}$$

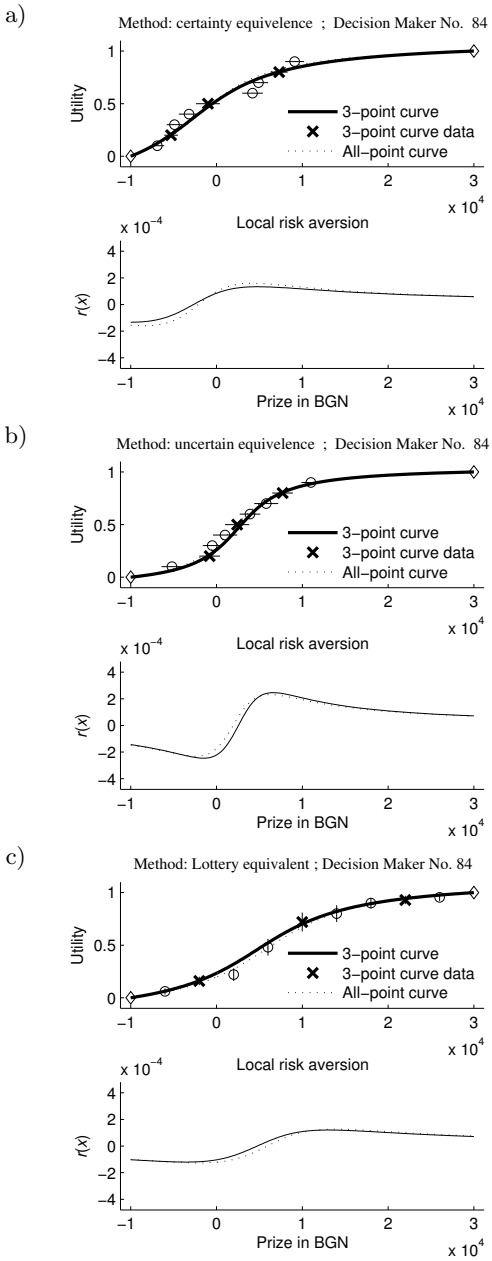

Figure 4. Arctan-approximated utility function (up) and local risk aversion function (down) for DM No. 84 using: a) CE; b) UE; c) LE, all on three and nine inner nodes. The uncertainty intervals of the elicited nodes are depicted by horizontal/vertical lines.

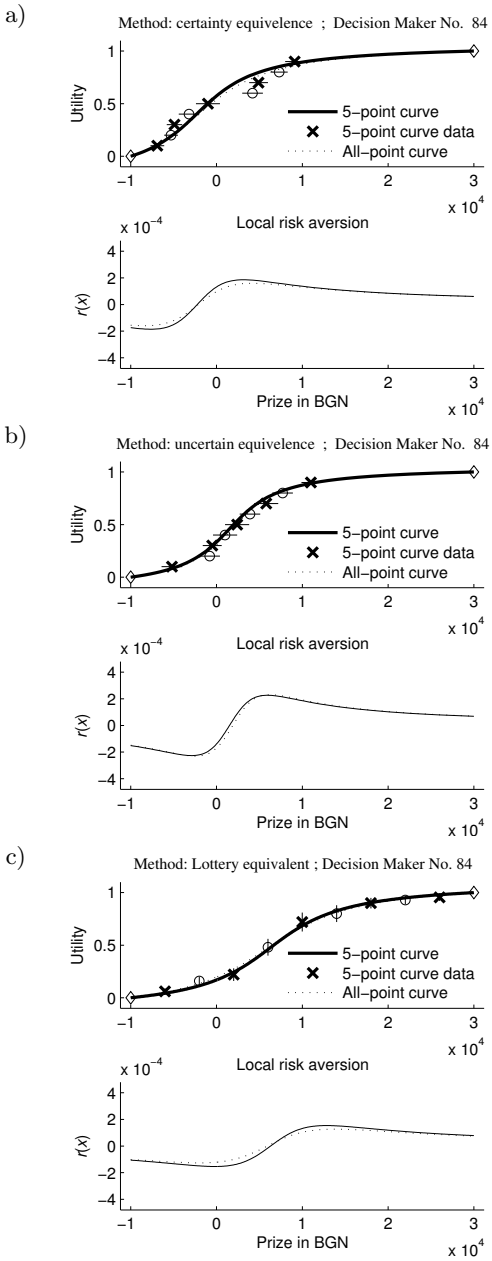

Figure 5. Arctan-approximated utility function (up) and local risk aversion function (down) for DM No. 84 using: a) CE; b) UE; c) LE, all on five and nine inner nodes. The uncertainty intervals of the elicited nodes are depicted by horizontal/vertical lines.

A possible relative measure of deterioration is the absolute measure (8) divided by the minimum value of (A1.4) or (A1.8), multiplied by 100 (in %) for $r = 3, 4, 5$; $j = 1, 2, \ldots, 104$:

$$\begin{aligned}
\Delta_{rel}^{CE_{r,j}} &= 100\Delta_{abs}^{CE_{r,j}}/\chi_x^2(CE_{9,j}, CE_{9,j}, arc),\\
\Delta_{rel}^{UE_{r,j}} &= 100\Delta_{abs}^{UE_{r,j}}/\chi_x^2(UE_{9,j}, UE_{9,j}, arc),\\
\Delta_{rel}^{LE_{r,j}} &= 100\Delta_{abs}^{LE_{r,j}}/\chi_u^2(LE_{9,j}, LE_{9,j}, arc).
\end{aligned} \tag{9}$$

The measures of deterioration (8) and (9) are calculated for all DMs. Tables 2, 3 and 4 present their mean (m), standard deviation (σ), median ($x_{0.5}$), interquartile range ($x_{0.75} - x_{0.25}$), and the goodness-of-fit measures with parameters, calculated over the partial samples, a nd tested over the whole sample.

If there is sufficient amount of data, a statistical test would reject a null hypothesis that the deterioration due to reduction of the number of data points is zero, therefore the test is useless for our purposes. Although deterioration of the quality of approximation is statistically significant, it is negligible. Tables 2, 3 and 4 show that, for all methods, the increase of the number of nodes from 3 to 5 decreases the measures of deterioration. In the case of 5 nodes the highest average relative difference (9) is 14% and the highest average absolute difference (8) is 0.15, a bit more than 1/7 of an uncertainty interval. Thus, 5 nodes are sufficient to construct feasible approximation of $u(.)$ and $r(.)$, although more inner points would (insignificantly) increase the precision of the utility function. When the number of the inner nodes is further reduced, the approximation would not indicate whether it was successful or not. This is important, because in the latter case the analytical approximation should be replaced by linear interpolation.

Table 2. Comparison of arctan-approximation on the whole and partial samples of CE. Rows contain mean, standard deviation, median and interquartile range for 10 samples of 104 goodness-of-fit and deterioration measures, corresponding to the following columns, i.e. **1:** $\chi_x^2(CE_{9,j}, CE_{9,j}, arc)$ (7); **2:** $\chi_x^2(CE_{9,j}, CE_{3,j}, arc)$ (7); **3:** $\Delta_{abs}^{CE_{3,j}}$ (8); **4:** $\Delta_{rel}^{CE_{3,j}}$ (9); **5:** $\chi_x^2(CE_{9,j}, CE_{4,j}, arc)$ (7); **6:** $\Delta_{abs}^{CE_{4,j}}$ (8); **7:** $\Delta_{rel}^{CE_{4,j}}$ (9); **8:** $\chi_x^2(CE_{9,j}, CE_{5,j}, arc)$ (7); **9:** $\Delta_{abs}^{CE_{5,j}}$ (8); **10:** $\Delta_{rel}^{CE_{5,j}}$ (9).

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
m	1.9	2.3	0.36	24	2.3	0.36	22	2.1	0.16	9.5
σ	1.81	2.15	0.517	31.2	2.21	0.671	28.6	1.96	0.294	11.2
$x_{0.5}$	1.4	1.7	0.15	13	1.8	0.13	10	1.6	0.066	5.1
$x_{0.75}$-$x_{0.25}$	1.86	1.87	0.371	27.0	2.06	0.390	29.0	1.90	0.171	12.3

Table 3. Comparison of arctan-approximation on the whole and partial samples of UE. Rows contain mean, standard deviation, median and interquartile range for 10 samples of goodness-of-fit and deterioration measures, corresponding to the following columns, i.e. **1:** $\chi^2_x(UE_{9,j}, UE_{9,j}, arc)$ (7); **2:** $\chi^2_x(UE_{9,j}, UE_{3,j}, arc)$ (7); **3:** $\Delta_{abs}^{UE_{3,j}}$ (8); **4:** $\Delta_{rel}^{UE_{3,j}}$ (9); **5:** $\chi^2_x(UE_{9,j}, UE_{4,j}, arc)$ (7); **6:** $\Delta_{abs}^{UE_{4,j}}$ (8); **7:** $\Delta_{rel}^{UE_{4,j}}$ (9); **8:** $\chi^2_x(UE_{9,j}, UE_{5,j}, arc)$ (7); **9:** $\Delta_{abs}^{UE_{5,j}}$ (8); **10:** $\Delta_{rel}^{UE_{5,j}}$ (9)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
m	2.4	2.8	0.48	25	2.8	0.41	18	2.5	0.15	8.0
σ	1.92	2.53	1.04	32.4	2.53	0.992	23.8	2.05	0.293	9.57
$x_{0.5}$	1.9	2.4	0.15	9.9	2.4	0.088	6.2	2.1	0.054	3.8
$x_{0.75}$-$x_{0.25}$	2.69	2.61	0.348	34.9	2.76	0.355	25.4	2.68	0.147	11.0

Table 4. Comparison of arctan-approximation on the whole and the partial samples of LE. Rows contain mean, standard deviation, median and interquartile range for 10 samples of goodness-of-fit and deterioration measures, corresponding to the following columns, i.e. **1:** $\chi^2_u(LE_{9,j}, LE_{9,j}, arc)$ (6); **2:** $\chi^2_u(LE_{9,j}, LE_{3,j}, arc)$ (6); **3:** $\Delta_{abs}^{LE_{3,j}}$ (8); **4:** $\Delta_{rel}^{LE_{3,j}}$ (9); **5:** $\chi^2_u(LE_{9,j}, LE_{4,j}, arc)$ (6); **6:** $\Delta_{abs}^{LE_{4,j}}$ (8); **7:** $\Delta_{rel}^{LE_{4,j}}$ (9); **8:** $\chi^2_u(LE_{9,j}, LE_{5,j}, arc)$ (6); **9:** $\Delta_{abs}^{LE_{5,j}}$ (8); **10:** $\Delta_{rel}^{LE_{5,j}}$ (9)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
m	0.78	1.1	0.23	34	0.89	0.11	18	0.86	0.08	14
σ	0.705	0.873	0.336	37.3	0.774	0.164	22.6	0.747	0.100	14.0
$x_{0.5}$	0.54	0.75	0.093	19	0.62	0.052	9.9	0.66	0.051	8.9
$x_{0.75}$–$x_{0.25}$	0.803	1.11	0.271	43.3	0.999	0.115	22.1	0.920	0.0944	18.3

Similar results are obtained through additional goodness-of-fit measures of the number of times when the approximated values of (A1.9) or (A1.12) are outside their initial uncertainty intervals. These measures are hardly suitable for optimization, since they are discrete variables, but they could be natural deterioration measures between the whole and the partial samples. The numerical results do not provide additional insight on the problem of reducing the sample size and are omitted.

Eliciting inner nodes is one of the most time-consuming tasks in decision analysis, and so identification of the minimum required number of inner nodes is of high practical importance.

3.4. Arctan-approximation versus linear interpolation over the partial samples

This section provides quantitative proof for the third hypothesis from the Introduction.

We compared linear interpolation and arctan-approximation in Section 3.2, but only qualitatively. Although linear interpolation does not generate "errors" in the elicited nodes, it does generate error if constructed over the partial samples with goodness-of-fit measures, estimated using the whole samples. The utility functions for DM No. 84, linearly interpolated on the midpoints of the uncertainty intervals of 3-partial and 5-partial CE, UE and LE samples and their corresponding local risk aversion functions are shown in Figs. 6 and 7 with a solid line. A dotted line represents the interpolated functions over the whole sample. Visual comparison shows that even at five nodes, the local risk aversion approximation is of poor quality.

An absolute measure of deterioration of linear interpolation compared to arctan-approximation is the difference between the goodness-of-fit measures for both forms for $r = 3, 4, 5$; $j = 1, 2, \ldots, 104$:

$$\begin{aligned}
\Delta_{abs,lin}^{CE_{r,j}} &= \chi_x^2(CE_{9,j}, CE_{r,j}, lin) - \chi_x^2(CE_{9,j}, CE_{r,j}, arc),\\
\Delta_{abs,lin}^{UE_{r,j}} &= \chi_x^2(UE_{9,j}, UE_{r,j}, lin) - \chi_x^2(UE_{9,j}, UE_{r,j}, arc),\\
\Delta_{abs,lin}^{LE_{r,j}} &= \chi_u^2(LE_{9,j}, LE_{r,j}, lin) - \chi_u^2(LE_{9,j}, LE_{r,j}, arc),
\end{aligned} \tag{10}$$

The values of (10), like the values of its components, are calculated for all DMs. The mean (m), standard deviation (σ), median ($x_{0.5}$), as well as the interquartile range ($x_{0.75} - x_{0.25}$) of the absolute measures of deterioration on the partial samples for all DMs are given in columns 3, 6 and 9 of Tables 5, 6 and 7.

We use the following paired sample statistical tests to prove the statistical significance of the differences in the goodness-of-fit measures of the arctan-approximation and linear interpolation:

1) Bootstrap mean test to analyze whether the mean value of the difference Δ of the two samples is zero. The null hypothesis is H_0: $E_\Delta = 0$, and the alternative hypothesis is H_1: $E_\Delta > 0$.

Table 5. Comparison of linear interpolation and arctan-approximation over the partial samples of CE. Rows contain mean, standard deviation, median and interquartile range for nine samples of goodness-of-fit and deterioration measures, corresponding to the following columns, i.e. **1:** $\chi_x^2(CE_{9,j}, CE_{3,j}, arc)$ (7); **2:** $\chi_x^2(CE_{9,j}, CE_{3,j}, lin)$ (7); **3:** $\Delta_{abs,lin}^{CE_{3,j}}$ (10); **4:** $\chi_x^2(CE_{9,j}, CE_{4,j}, arc)$ (7); **5:** $\chi_x^2(CE_{9,j}, CE_{4,j}, lin)$ (7); **6:** $\Delta_{abs,lin}^{CE_{4,j}}$ (10); **7:** $\chi_x^2(CE_{9,j}, CE_{5,j}, arc)$ (7); **8:** $\chi_x^2(CE_{9,j}, CE_{5,j}, lin)$ (7); **9:** $\Delta_{abs,lin}^{CE_{5,j}}$ (10)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
m	2.3	5.3	3.0	2.3	5.0	2.7	2.1	0.53	-1.5
σ	2.15	6.92	6.43	2.21	6.87	6.39	1.96	0.492	1.99
$x_{0.5}$	1.7	3.1	0.75	1.8	2.9	0.66	1.6	0.32	-1.4
$x_{0.75}$–$x_{0.25}$	1.87	5.23	4.67	2.06	5.12	4.59	1.90	0.594	1.89

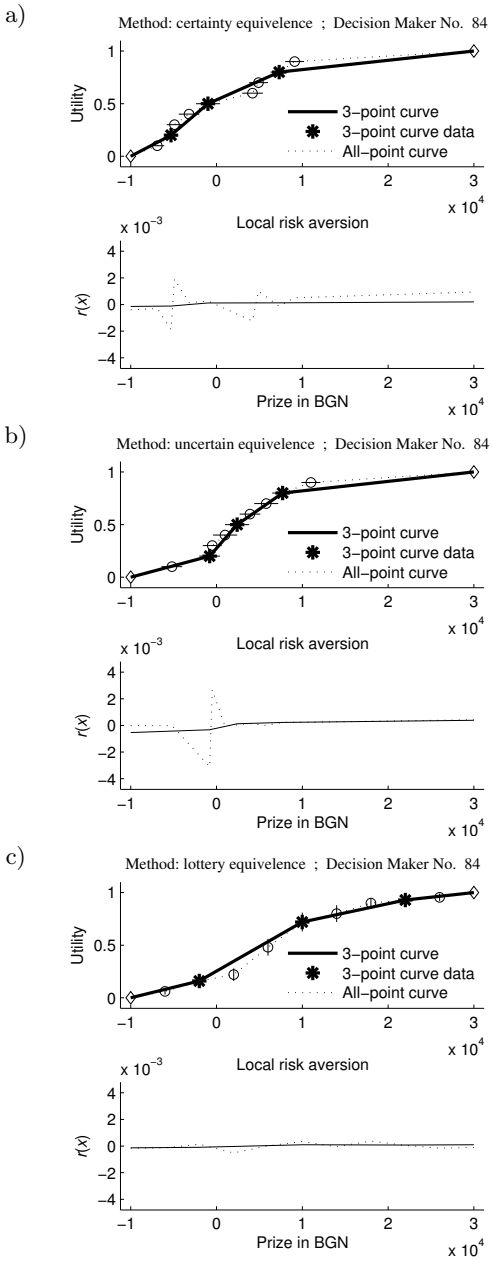

Figure 6. Linearly interpolated utility function (up) and local risk aversion function (down) for DM No. 84 using: a) CE; b) UE; c) LE, all on 3 and 9 inner nodes. The uncertainty intervals of the elicited nodes are depicted by horizontal/vertical lines.

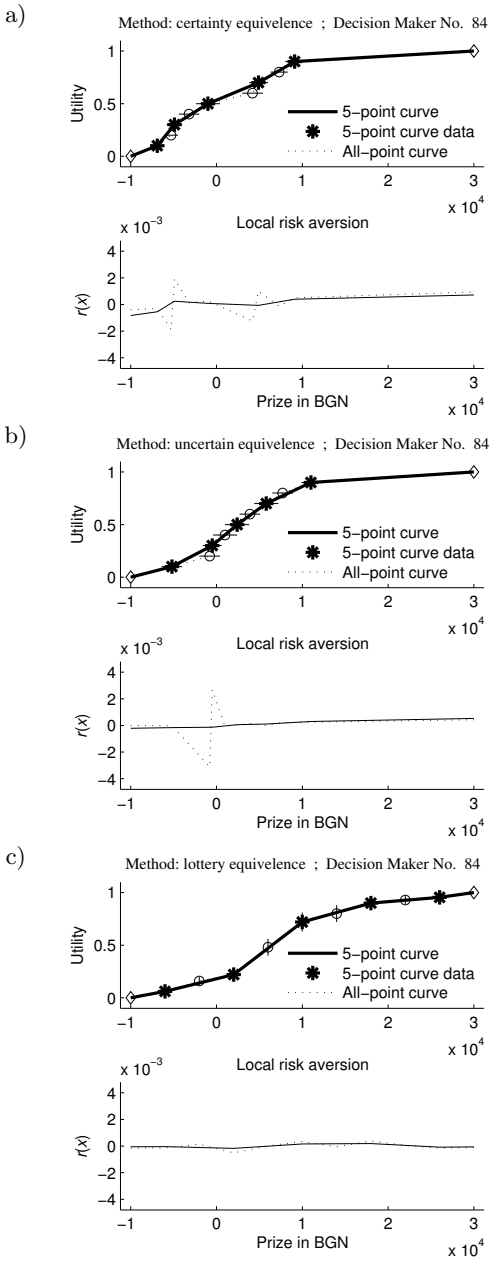

Figure 7. Linearly interpolated utility function (up) and local risk aversion function (down) for DM No. 84 using: a) CE; b) UE; c) LE,all on 5 and 9 inner nodes. The uncertainty intervals of the elicited nodes are depicted by horizontal/vertical lines.

Table 6. Comparison of linear interpolation and arctan-approximation over the partial samples of UE. Rows contain mean, standard deviation, median and interquartile range for nine samples of goodness-of-fit and deterioration measures, corresponding to the following columns, i.e. **1:** $\chi^2_x(UE_{9,j}, UE_{3,j}, arc)$ (7); **2:** $\chi^2_x(UE_{9,j}, UE_{3,j}, lin)$ (7); **3:** $\Delta^{UE_{3,j}}_{abs,lin}$ (10); **4:** $\chi^2_x(UE_{9,j}, UE_{4,j}, arc)$ (7); **5:** $\chi^2_x(UE_{9,j}, UE_{4,j}, lin)$ (7); **6:** $\Delta^{UE_{4,j}}_{abs,lin}$ (10); **7:** $\chi^2_x(UE_{9,j}, UE_{5,j}, arc)$ (7); **8:** $\chi^2_x(UE_{9,j}, UE_{5,j}, lin)$ (7); **9:** $\Delta^{UE_{5,j}}_{abs,lin}$ (10)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
m	2.8	3.6	0.81	2.8	3.4	0.59	2.5	0.62	-1.9
σ	2.53	5.80	4.94	2.53	5.86	5.08	2.05	1.09	2.01
$x_{0.5}$	2.4	2.1	-0.090	2.4	1.8	-0.16	2.1	0.39	-1.3
$x_{0.75}$–$x_{0.25}$	2.61	2.43	2.40	2.76	2.07	2.37	2.69	0.564	2.75

Table 7. Comparison of linear interpolation and arctan-approximation over the partial samples of LE. Rows contain mean, standard deviation, median and interquartile range for nine samples of goodness-of-fit and deterioration measures, corresponding to the following columns, i.e. **1:** $\chi^2_u(LE_{9,j}, LE_{3,j}, arc)$ (6); **2:** $\chi^2_u(LE_{9,j}, LE_{3,j}, lin)$ (6); **3:** $\Delta^{LE_{3,j}}_{abs,lin}$ (10); **4:** $\chi^2_u(LE_{9,j}, LE_{4,j}, arc)$ (6); **5:** $\chi^2_u(LE_{9,j}, LE_{4,j}, lin)$ (6); **6:** $\Delta^{LE_{4,j}}_{abs,lin}$ (10); **7:** $\chi^2_u(LE_{9,j}, LE_{5,j}, arc)$ (6); **8:** $\chi^2_u(LE_{9,j}, LE_{5,j}, lin)$ (6); **9:** $\Delta^{LE_{5,j}}_{abs,lin}$ (10);

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
m	1.0	0.78	-0.22	0.89	0.51	-0.38	0.86	0.34	-0.52
σ	0.873	0.700	0.656	0.774	0.433	0.571	0.747	0.416	0.676
$x_{0.5}$	0.75	0.53	-0.11	0.62	0.37	-0.20	0.66	0.21	-0.29
$x_{0.75}$–$x_{0.25}$	1.11	0.746	0.454	0.999	0.406	0.475	0.920	0.337	0.579

2) Bootstrap median test to analyze whether the median of the difference Δ of the two samples is zero. The null hypothesis is H_0: $\Delta_{0.5} = 0$, and the alternative hypothesis is H_1: $\Delta_{0.5} > 0$ (detailed description of both Bootstrap tests is given in Appendix 3).

3) sign test to analyze whether the median of the difference of the two samples Δ is zero. The null hypothesis is H_0: $\Delta_{0.5}=0$, and the alternative hypothesis is H_1: $\Delta_{0.5} > 0$. The value of $p_{value,3}$ may be calculated using a modification of the *signtest* function of the MATLAB Statistical Toolbox (The MathWorks, 2006).

4) sign rank test to analyze whether the median of the difference of the two samples Δ is zero. The null hypothesis is H_0: $\Delta_{0.5} = 0$, whereas the alternative hypothesis is H_1: $\Delta_{0.5} > 0$. The value of $p_{value,4}$ may be calculated using a modification of the *signrank* function of the MATLAB Statistical Toolbox (The MathWorks, 2006).

The p_{value} was calculated for each test and each elicitation method with N=100,000 simulation cycles for the Bootstrap tests. The results are summarized in Table 8. The p_{value}, calculated by tests that analyzed whether the

arctan-approximation was better than linear interpolation are in shaded boxes. The tests that analyzed whether the linear interpolation is better than arctan-approximation are given in white boxes. The tests that reject the null hypothesis at a significance level α=0.001 are in normal font, whereas the tests that failed to reject the null hypothesis are in bold. The results of Table 8 lead to the following conclusions:

1) the advantage of the arctan-approximation is statistically significant according to all tests for samples $CE_{3,j}$ and $CE_{4,j}$;

2) the advantage of the linear interpolation is statistically significant according to all tests for the partial samples of LE (except for the sign test on $LE_{3,j}$); such result could have been expected since LE generates a flatter utility function;

3) the results obtained on the basis of samples $UE_{3,j}$ and $UE_{4,j}$ show that the null hypothesis cannot be rejected, because, for all eight tests, $p_{value} > \alpha = 0.001$; moreover, the medians and the mean values point at different directions;

4) the analysis of the 5-partial samples of CE, UE and LE shows statistical significance in favor of linear interpolation; that might have been expected because here, linear interpolation uses five parameters compared to only two of the arctan-approximation.

These conclusions show that the third hypothesis does not contradict the data.

Table 8. Values of p_{value} of the four paired-sample tests over the partial samples. Shaded p_{value} indicate the tests of whether the arctan-approximation is better than linear interpolation, whereas non-shaded ones - whether linear interpolation is better than arctan-approximation. Normal font p_{value} are for tests that rejected H_0 at α=0.001, whereas bold p_{value} are for tests that fail to reject H_0.

	CE	**UE**	**LE**	**CE**	**UE**	**LE**	**CE**	**UE**	**LE**
	On 3 nodes			**On 4 nodes**			**On 5 nodes**		
$p_{value,1}$	0	**1.05e–2**	5e–6	5e–6	**4.65e–2**	0	0	0	0
$p_{value,2}$	0	**4.51e–2**	0	0	**1.19e–2**	0	0	0	0
$p_{value,3}$	5.11e–6	**1.41e–1**	**2.20e–3**	**1.20e–3**	**3.12e–2**	6.62e–12	2.17e–14	2.26e–19	2.17e–14
$p_{value,4}$	1.97e–7	**3.01e–1**	1.78e–5	5.91e–6	**7.40e–2**	1.05e–13	6.89e–15	3.50e–17	5.04e–16

4. Conclusions

This paper investigated the advantages of arctan-approximation over linear interpolation of a 1-D utility function on the basis of empirical data. Initially, the utility function was linearly interpolated on the midpoints of the uncertainty intervals. This approach did not take into account the width of the uncertainty intervals. Graphical results demonstrated that: 1) the constructed local risk

aversion was too imprecise; 2) the reduction of the number of inner nodes from 9 to 4 additionally compromised the results.

The analytical approximation of the utility function using the arctan(.) form (A1.9) qualitatively demonstrated that: 1) arctan-approximation was more adequate than linear interpolation as it correctly represented the risk attitude of the FRDM; 2) even if 4 inner nodes were used, arctan-approximation still correctly encapsulated the local risk aversion of the FRDM.

Another step of the analysis was to arctan-approximate $u(.)$ over partial samples. Goodness-of-fit measures, as well as measures of deterioration were identified, based on χ^2. Even though the deterioration of quality was statistically significant, it was actually negligible. Thus, 5 nodes were assumed to be sufficient to acquire feasible approximation of $u(.)$ and $r(.)$. Further reduction of the number of elicited nodes is not recommended, because it is not possible to check whether the approximation was successful or not. A final step of the analysis was to compare arctan-approximation and linear interpolation over the partial samples. Four paired-sample tests were defined to prove the third hypothesis.

As a result of the study, we propose to elicit five inner nodes using a preliminarily selected elicitation method as a first step in constructing the utility function. Then $u(.)$ may be arctan-approximated on these nodes. If the later does not fit well to the data, then it should be replaced by linear interpolation on the midpoints of the node uncertainty intervals.

All calculation and visualization procedures in this study were performed using original MATLAB program functions available free of charge upon request from the authors. Results reported in the paper shall facilitate further utility analysis, as it can be based on a short elicitation phase, followed by calculation procedures performed by program functions. If the FRDM has typical risk attitude, then such a utility analysis would apply, it will be less time consuming, and the results will be less influenced by elicitation errors. As the reported results differ depending on the elicitation method, it is necessary to continue research by collecting elicitation results via other methods, such as TO or the modification of UE – the chaining UE.

References

Clemen, R. (1996) *Making Hard Decisions: an Introduction to Decision Analysis*. Second Edition. Duxbury Press, Wadsworth Publishing Company.

Efron, B. and Tibshirani, R.J. (1993) *An Introduction to the Bootstrap*. Chapman & Hall.

Farquhar, P.H. (1984) Utility Assessment Methods. *Management Science* **30** (11), 1283–1300.

French, S. (1993) *Decision Theory: an Introduction to the Mathematics of Rationality*. Ellis Horwood.

French, S. and Insua, D.R. (2000) *Statistical Decision Theory*. Arnold.

KEENEY, R.L. and RAIFFA, H. (1993) *Decisions with Multiple Objectives: Preference and Value Tradeoffs*. Cambridge University Press.

MCCORD, M. and DE NEUFVILLE, R. (1986) Lottery Equivalents: Reduction of the Certainty Effect Problem in Utility Assessment. *Management Science* **32**, 56–60.

NIKOLOVA, N.D. (2007) Elicitation Errors in One-Dimensional Non-Monotonic Utility Functions. *Computer Science and Technology* **2**, 48–59.

NIKOLOVA, N.D., SHULUS, A., TONEVA, D. and TENEKEDJIEV, K. (2005) Fuzzy Rationality in Quantitative Decision Analysis. *Journal of Advanced Computational Intelligence and Intelligent Informatics* **9** (1), 65–69.

PRATT, J.W. (1964) Risk aversion in the small and in the large. *Econometrica* **32**, 122–136.

PRESS, W.H., TEUKOLSKY, S.A., VETTERLING, W.T. and FLANNERY, B.P. (1992) *Numerical Recipes – the Art of Scientific Computing*. Cambridge University Press.

TENEKEDJIEV, K., NIKOLOVA, N.D. and DIMITRAKIEV, D. (2004) Application of the Triple Bisection Method for Extraction of Subjective Utility Information. *Proc. Second International Conference "Management and Engineering '2004"*, **2** (7), 115–117, Sofia, Bulgaria.

TENEKEDJIEV, K., NIKOLOVA, N.D. and DIMITRAKIEV, D. (2008) Analytical One-Dimensional Utility – Comparison of Power and Arctg-Approximation. *Engineering Sciences* **4**, 20–34 (in Bulgarian).

TENEKEDJIEV, K., NIKOLOVA, N.D. and PFLIEGL, R. (2006) Utility Elicitation with the Uncertain Equivalence Method. *Comptes rendus de l'academie Bulgare des sciences* **59** (3), 283–288.

THE MATHWORKS (2006) *MATLAB Statistical Toolbox 3 – User's Guide*. The MathWorks Inc.

TRAUTMANN, H. and WEIHS, C. (2006) On the Distribution of the Desirability Index using Harrington's Desirability Function. *Metrika* **63** (2), 207–213.

TVERSKY, A. and KAHNEMAN, D. (1974) Judgment under Uncertainty: Heuristics and Biases. *Science* **185**, 1124–1131.

VON NEUMANN, J. and MORGENSTERN, O. (1947) *Theory of Games and Economic Behaviour*, Second Edition. Princeton University Press.

WAKKER, P. and DENEFFE, D. (1996) Eliciting Von Neumann-Morgenstern Utilities when Probabilities are Distorted or Unknown. *Management Science* **42**, 1131–1150.

Appendix 1. Analytical approximation of a utility function

Equation (A1.1) represents analytical dependence approximating utility of an FRDM:

$$u = u(x, \vec{\mathbf{p}}), \tag{A1.1}$$

where $\vec{\mathbf{p}} = (p_1, p_2, \ldots, p_n)$ is an n-dimensional vector of unknown parameters (from an n-dimensional set Π), defining the form of $u(.)$. (A1.1) must be strictly increasing, which implies

$$u(x_i, \vec{\mathbf{p}}) > u(x_j, \vec{\mathbf{p}}) \Leftrightarrow x_i > x_j, \text{for } x_i \in [x_{worst}; x_{best}], \\ x_j \in [x_{worst}; x_{best}] \text{ and } \vec{\mathbf{p}} \in \Pi. \quad \text{(A1.2)}$$

The end nodes should be error-free, thus

$$u(x_{worst}, \vec{\mathbf{p}}) = 0,\ u(x_{best}, \vec{\mathbf{p}}) = 1,\ \text{for } \vec{\mathbf{p}} \in \Pi. \quad \text{(A1.3)}$$

If the estimated nodes are of the form (2), then the unknown parameters of (A1.1) may be identified using a weighted least square method (Press et al., 1992). Deviation of the model from the best subjective point estimate in a given node is weighted by the width of the uncertainty interval.

Each $\hat{u}_l$ is a random variable, dependent on x_{u_l}. The difference from the classical regression analysis is that we measure the confidence interval $[\hat{u}_l^d; \hat{u}_l^u]$ with confidence level approaching 1, rather than several realizations of the dependent variable. The distribution of $\hat{u}_l$ is unknown, but it is assumed to be of the same type for all $l = 2, 3, ..., z-1$. Then, for each symmetrical distribution, the result of the measurement of the dependent variable may be treated as if the average measured value $(\hat{u}_l^d + \hat{u}_l^u)/2$ with standard deviation proportional to $(\hat{u}_l^u - \hat{u}_l^d)$ were obtained. This is a good approximation even when the distribution of $\hat{u}_l$ is asymmetric, but unimodal.

The goodness-of-fit measure of (A1.1) if the data are of type (2) is:

$$\chi_u^2 = \sum_{l=2}^{z-1} \left(\frac{u(x_l, \vec{\mathbf{p}}) - \left(\hat{u}_l^d + \hat{u}_l^u\right)/2}{\hat{u}_l^u - \hat{u}_l^d} \right)^2 = \sum_{l=2}^{z-1} \left(\frac{2u(x_l, \vec{\mathbf{p}}) - \hat{u}_l^d - \hat{u}_l^u}{2\hat{u}_l^u - 2\hat{u}_l^d} \right)^2. \quad \text{(A1.4)}$$

The optimal parameters $\vec{\mathbf{p}}_{opt}$ may be identified by n-dimensional minimization of χ_u^2 over $\vec{\mathbf{p}}$.

Since (2) contains random data, $\vec{\mathbf{p}}_{opt}$ is also random. So, $u = u(x, \vec{\mathbf{p}}_{opt})$ is a random function.

From (A1.2) it follows that there exists an inverse function of (A1.1):

$$x = x(u, \vec{\mathbf{p}}) = u^{-1}(x, \vec{\mathbf{p}}). \quad \text{(A1.5)}$$

The function (A1.5), similarly to (A1.1), would be strictly increasing and fixed at its ends:

$$x(u_i, \vec{\mathbf{p}}) > x(u_j, \vec{\mathbf{p}}) \Leftrightarrow u_i > u_j,\ \text{for } u_i \in [0;1], u_j \in [0;1] \text{ and } \vec{\mathbf{p}} \in \Pi, \quad \text{(A1.6)}$$

$$x(0, \vec{\mathbf{p}}) = x_{worst}, x(1, \vec{\mathbf{p}}) = x_{best},\ \text{for } \vec{\mathbf{p}} \in \Pi. \quad \text{(A1.7)}$$

If the elicited nodes are of the form (3), then the unknown parameters of (A1.1) may be identified using a weighted least squares method.

Each $\hat{x}_{u_l}$ is a random variable, dependent on u_l. The difference from the classical regression analysis is that we measure a confidence interval $[\hat{x}^d_{u_l}; \hat{x}^u_{u_l}]$ with confidence level approaching 1, rather than several realizations of the dependent variable. The distribution of $\hat{x}_{u_l}$ is unknown, but it is assumed to be of the same type for all $l = 2, 3, ..., z-1$. Then, for each symmetrical distribution, the result of the measurement of the dependent variable may be treated as if the average measured value $(\hat{x}^d_{u_l} + \hat{x}^u_{u_l})/2$ were obtained, with standard deviation proportional to $(\hat{x}^u_{u_l} - \hat{x}^d_{u_l})$. This is a good approximation even when the distribution of $\hat{x}_{u_l}$ is asymmetric, but unimodal.

A possible goodness-of-fit measure of (A1.1) if the data is of type (3) is

$$\chi^2_x = \sum_{l=2}^{z-1} \left(\frac{x(u_l, \vec{\mathbf{p}}) - \left(\hat{x}^d_{u_l} + \hat{x}^u_{u_l}\right)/2}{\hat{x}^u_{u_l} - \hat{x}^d_{u_l}} \right)^2 = \sum_{l=2}^{z-1} \left(\frac{2x(u_l, \vec{\mathbf{p}}) - \hat{x}^d_{u_l} - \hat{x}^u_{u_l}}{2\hat{x}^u_{u_l} - 2\hat{x}^d_{u_l}} \right)^2 \tag{A1.8}$$

The optimal parameters $\vec{\mathbf{p}}_{opt}$ may be found using n-dimensional minimization of χ^2_x on $\vec{\mathbf{p}}$. This task can be much easier if the inverse function (A1.5) of (A1.1) is analytical. Otherwise, each calculation of (A1.8) would require numerical solving of $z-2$ nonlinear algebraic equations with a single unknown.

Since (3) contains random data, $\vec{\mathbf{p}}_{opt}$ is also random, and $x = u^{-1}(x, \vec{\mathbf{p}}_{opt})$ is a random function.

Nikolova (2007) proposed an analytical dependence (A1.9) that obeys (A1.2) and (A1.3):

$$u(x) = \frac{\arctan[a(x - x_0)] - \arctan[a(x_1 - x_0)]}{\arctan[a(x_z - x_0)] - \arctan[a(x_1 - x_0)]}. \tag{A1.9}$$

In (A1.9), the vector of unknown parameters is a two-dimensional $\vec{\mathbf{p}} = (a, x_0)$, and the two-dimensional set Π is defined as:

$$\Pi = \{(a, x_0 / a \in (0, \infty) \wedge x_0 \in (-\infty, \infty)\}. \tag{A1.10}$$

The local risk aversion function that corresponds to the utility (A1.9) is

$$r(x) = -\frac{u''(x)}{u'(x)} = \frac{2a^2(x - x_0)}{1 + a^2(x - x_0)^2}. \tag{A1.11}$$

After substituting (A1.9) in (A1.4), χ^2_u transforms into a function of a and x_0.

The utility function (A1.9) has an analytical inverse:

$$x(u) = \frac{tg\{u \times \arctan[a(x_z - x_o)] + (1 - u)\arctan[a(x_1 - x_o)]\}}{a} + x_0. \tag{A1.12}$$

After substituting (A1.12) in (A1.8), χ^2_x transforms into a function of a and x_0.

The analytical approximation (A1.9) shall be referred to as arctan-approximation.

Appendix 2. Local risk aversion of the linearly interpolated utility function

If several elicited nodes of the form (2) or (3) are available, the utility function may be linearly interpolated on the midpoints of the uncertainty intervals. It is necessary to construct the local risk aversion function of the linearly interpolated utility. The left and right bound of $r(x)$ may be calculated for each inner node (x_l, u_l), for $l = 2, 3, \ldots, z-1$, using approximation of the derivatives with finite differences:

$$r(x_l^-) = -2\left(\frac{u_{l+1}-u_l}{x_{l+1}-x_l} - \frac{u_l - u_{l-1}}{x_l - x_{l-1}}\right)\frac{x_l - x_{l-1}}{(x_{l+1}-x_{l-1})\,(u_l - u_{l-1})}, \quad \text{for } l = 2, 3, \ldots, z-1, \tag{A2.1}$$

$$r(x_l^+) = -2\left(\frac{u_{l+1}-u_l}{x_{l+1}-x_l} - \frac{u_l - u_{l-1}}{x_l - x_{l-1}}\right)\frac{x_{l+1} - x_l}{(x_{l+1}-x_{l-1})\,(u_{l+1} - u_l)}, \quad \text{for } l = 2, 3, \ldots, z-1. \tag{A2.2}$$

Since x_l^- and x_l^+ are equal, then direct application of (A2.1) and (A2.2) leads to vertical sections in $r(x)$. The vertical sections should be replaced by their midpoints in order to represent risk aversion as a function:

$$r(x_l) = \frac{r(x_l^-) + r(x_l^+)}{2}, \text{ for } l = 2, 3, \ldots, z-1. \tag{A2.3}$$

The derivatives of the first and the last nodes may be estimated respectively as left and right derivatives:

$$r(x_1) = r(x_2^-), \tag{A2.4}$$

$$r(x_z) = r(x_{z-1}^+). \tag{A2.5}$$

The function $r(x)$ may be linearly interpolated on the nodes calculated via (A2.3), (A2.4) and (A2.5).

Appendix 3. Bootstrap mean and median tests

The Bootstrap mean test analyzes whether the mean value of the difference Δ of two paired samples is zero. The null hypothesis is H_0: the mean value of Δ is zero, and the alternative hypothesis is H_1: the mean value of Δ is positive. If the null hypothesis is true, then the distribution of Δ is symmetrical around its mean. Then each synthetic sample of n realizations of Δ may be doubled to a synthetic sample of $2n$ realizations by adding all values symmetrical to the initial ones with respect to the mean (i.e. zero). After generating N synthetic samples, N synthetic estimates of the mean value of Δ may be calculated and doubled to a $2N$ number of synthetic estimates by adding estimates symmetrical to the initial ones with respect to the mean (i.e. with respect to zero). The p_{value} is the

number of synthetic estimates of the mean value of Δ, exceeding the observed mean divided by $2N$ (Efron, Tibshirani, 1993).

The Bootstrap median test analyzes whether the median of the difference Δ of the two samples is zero. It is the same as the Bootstrap mean test, but it uses the median as a test statistic. Assume that $\{a_i|i = 1, 2, \ldots, n\}$ and $\{b_i|i = 1, 2, \ldots, n\}$ are two paired samples, such that the mean of the first sample is lower than that of the second. Also assume that N is the number of Bootstrap replicas. Let r be a discrete random variable that takes the integer values $1, 2, 3, \ldots, 2n$, with equal probabilities $(2n)^{-1}$. The following algorithms are elaborated to calculate p_{value} of the Bootstrap mean and the median tests.

Algorithm A3.1. Calculation of p_{value} of a Bootstrap mean test

1. Form a sample of differences between the pairs of values a_i and b_i: $\{\Delta_i|i = 1, 2, \ldots, 2n\}$, where $\Delta_i = b_i - a_i$ (for $i = 1, 2, \ldots, n$), and $\Delta_{i+n} = a_i - b_i$ (for $i = 1, 2, \ldots, n$);
2. Calculate the mean $m(\Delta) = \frac{1}{2n}\sum_{i=1}^{2n}\Delta_i$ of the sample of differences $\{\Delta_i|i = 1, 2, \ldots, 2n\}$;
3. Put $j = 0$;
4. Put $j = j + 1$;
5. Form a synthetic sample of differences $\{\Delta_i^{s,j}|i = 1, 2, \ldots, 2n\}$:
   (a) Put $i = 0$;
   (b) Put $i = i + 1$;
   (c) Generate a random realization r_i of r;
   (d) Put $\Delta_i^{s,j} = \Delta_{r_i}$;
   (e) If $i < 2n$, then go to step 5.b.
6. Calculate means of the j-th synthetic sample $m_j^s(\Delta^s) = \frac{1}{2n}\sum_{i=1}^{2n}\Delta_i^{s,j}$ and put $m_{j+N}^s(\Delta^s) = -m_j^s(\Delta^s)$;
7. If $j < N$, then go to step 4.
8. Calculate $p_{value,1}$ of the Bootstrap mean test as follows:

$$p_{value,1} = \sum_{\substack{j=1 \\ m_j^s(\Delta^s) > m(\Delta)}}^{2N} \frac{1}{2N}.$$

Algorithm A3.2. Calculation of p_{value} of a Bootstrap median test

1. Form a sample of differences between the pairs of values a_i and b_i: $\{\Delta_i|i = 1, 2, \ldots, 2n\}$, where $\Delta_i = b_i - a_i$ (for $i = 1, 2, \ldots, n$), and $\Delta_{i+n} = a_i - b_i$ (for $i = 1, 2, \ldots, n$);
2. Sort the elements $\{\Delta_i|i = 1, 2, \ldots, 2n\}$ in ascending order such that $\Delta_1 \leq \Delta_2 \leq \ldots \leq \Delta_{2n}$;
3. Calculate the median $\Delta_{0.5}(\Delta) = \frac{\Delta_n + \Delta_{n+1}}{2}$ of the sorted sample of differences $\{\Delta_i|i = 1, 2, \ldots, 2n\}$

4. Put $j = 0$;
5. Put $j = j + 1$;
6. Form a sorted synthetic sample of differences $\{\Delta_i^{s,j} | i = 1, 2, \ldots, 2n\}$:
   (a) Put $i = 0$;
   (b) Put $i = i + 1$;
   (c) Generate a random realization r_i of r;
   (d) Put $\Delta_i^{s,j} = \Delta_{r_i}$;
   (e) If $i < 2n$, then go to step 6.b;
   (f) Sort $\{\Delta_i^{s,j} | i = 1, 2, \ldots, 2n\}$ in ascending order such that $\Delta_1^{s,j} \leq \Delta_2^{s,j} \leq \ldots \leq \Delta_{2n}^{s,j}$.
7. Calculate the median of the j-th synthetic sample $\Delta_{0.5,j}^s(\Delta^s) = \frac{\Delta_n^{s,j} + \Delta_{n+1}^{s,j}}{2}$ and put $\Delta_{0.5,j+N}^s(\Delta^s) = -\Delta_{0.5,j}^s(\Delta^s)$;
8. If $j < N$, then go to step 5;
9. Calculate $p_{value,2}$ of the Bootstrap median test as follows:

$$p_{value,2} = \sum_{\substack{j=1 \\ \Delta_{0.5,j}^s(\Delta^s) > \Delta_{0.5}(\Delta)}}^{2N} \frac{1}{2N}.$$