What can multiple price lists really tell us about risk preferences?

Drichoutis, Andreas and Lusk, Jayson

University of Ioannina, Department of Economics, Oklahoma State University, Department of Agricultural Economics

27 March 2012

What Can Multiple Price Lists Really Tell Us about Risk Preferences?[*]

Andreas C. Drichoutis[†]   Jayson L. Lusk[‡]

University of Ioannina   Oklahoma State University

October 22, 2012[§]

Abstract

Multiple price lists have emerged as a simple and popular method for eliciting risk preferences. Despite their popularity, a key downside of multiple price lists has not been widely recognized - namely that the approach is unlikely to generate sufficient information to accurately identify different dimensions of risk preferences. The most popular theories of decision making under risk posit that preference for risk are driven by a combination of two factors: the curvature of the utility function and the extent to which probabilities are weighted non-linearly. In this paper, we show that the widely used multiple price list introduced by Holt and Laury (2002) is likely more accurate at eliciting the latter, and we introduce a different multiple price list that is likely more accurate at eliciting the former. We show that by combining information from different multiple price lists, greater predictive performance can be achieved.

Keywords: expected utility theory; experiment; multiple price list; probability weighting; rank dependent utility; risk.

JEL Classification Numbers: C91, D81.

1

1 Introduction

The abundance of uncertainty in life has prompted a great many investigations into humans' response to risk. The interest in understanding risk preferences has created a latent demand for effective, easy-to-use risk preference elicitation devises. Following a long line of previous research by Becker et al. (1964), Binswanger (1980, 1981), and many others, in 2002 Holt and Laury (H&L) introduced a risk preference elicitation method that has subsequently become a mainstay. In a testament to the general interest in risk preference elicitation and to the specific appeal of the approach introduced by H&L, their work has been cited more than 1,800 times according to Google Scholar and is the second most highly cited paper published by the American Economic Review since 2002 according to ISIs Web of Knowledge. The approach used by H&L has subsequently come to be referred to as a type of multiple price list (MPL) (Andersen et al., 2006; Harrison and Rutstrom, 2008), an approach thought to have been first used by Miller et al. (1969).[1] The key advantage of the MPL is its ease of use. Respondents make a series of consecutive choices between two outcomes, where the expected value of one outcome increases at a higher rate than the other. The point at which an individual switches from choosing one outcome over the other is often used as a measure of risk aversion.

Despite the fact that MPLs are easy to use and easy for participants to understand, the approach has some weaknesses. (Harrison et al., 2005) pointed out that inferences from MPLs can be influenced by order effects (see also Holt and Laury, 2005), and Andersen et al. (2006) discussed the potential for choices in MPLs to be influenced by the ranges of values used. Here, we point to a more fundamental problem with MPLs that seems to have been overlooked by practitioners. In particular, the H&L approach is subject to Wakker and Deneffe's (1996) critique that many risk preference elicitation methods confound estimates of the curvature of the utility function (i.e., the traditional notion of risk preference) with an

---

[1]The word "multiple" in multiple price list is redundant since the word "list" already implies repetitive choices. Nevertheless, we adopt the phrasing MPL in this paper as it is more commonly used in the literature than other variants such as "choice list."

estimate of the extent to which an individual weights probabilities non-linearly. These are two conceptually different constructs that have different implications for individuals' behavior under risk, and without controlling for one, biased estimates of the other are obtained.

This observation about MPLs is well known to experts in the field of risk preference elicitation, and yet in our experience, it is not well known to newcomers or those outside the field. To illustrate our point, consider the responses by Peter Wakker (an accomplished and well known risk researcher) to a query posted on the "ESA Experimental Methods Discussion" Google Group about the advisability of using H&Ls elicitation technique:

> Probably you, and many other readers of this list, just want a quick and convenient index of risk aversion, to be used in regressions and comparisons. Then my advice is not to use choice lists as in [H&L]. Choice lists use many choice questions that are very similar, with either one outcome or one probability gradually changing and all else kept fixed. They do not cover the empirical domain of interest well or efficiently. In general, I recommend using choices between some lottery pairs with a variation in outcomes and probabilities to have a more representative sample of the domain.

And

> Many young experimental economists think that [H&L] invented the measurement of risk attitudes, with nothing existing before 2002, and that [expected utility] is default. My emails have argued that much of interest existed before. I hope that experimental economists will use the refined modern techniques to advance, rather than forget or (inefficiently losing time) reinvent the preceding literature.

The purpose of this paper is to further elucidate some of the issues raised by Wakker and more widely disseminate this knowledge among the (apparently large) audience of individuals interested in risk preference elicitation. Moreover, while we agree that the use of a single

"choice list" or MPL, may not perform well in fully capturing the multidimensional aspects of risk preferences, it must be acknowledged that their popularity results from ease of use. Accordingly, in this paper, we show that different types of MPLs are better able to capture some risk dimensions than others and that by using two (or more) easy to use MPLs, a researcher might achieve a more balanced picture of risk preferences, and thus might attain improved predictive validity.[2]

In what follows, we show that H&L's original MPL is, perhaps ironically, not particularly well suited to measuring the traditional notion of risk preferences - the curvature of the utility function. Rather, it is likely to provide a better approximation to the curvature of the probability weighting function. We then introduce an alternative MPL that has exactly the opposite property. By combining the information gained from both types of MPLs, we show that greater prediction performance can be attained.

2   Effect of Probability Weighting in MPLs

In the base-line MPL used by H&L, individuals were asked to make a series of 10 decisions between two options (see Table 1). In option A, the high payoff amount is fixed at $2 and the low payoff amount is fixed at $1.60 across all 10 decision tasks. In option B, the high payoff amount is fixed at $3.85 and the low payoff amount is fixed at $0.10. The only thing changing across the 10 decisions are the probabilities assigned to the high and low payoffs. Initially the probability of receiving the high payoff is 0.10 but by the tenth decision task, the probability is 1.0.

The expected value of lottery A exceeds the expected value of lottery B for the first four decision tasks. Thus, someone who prefers lottery A for the first four decision tasks and then switches and prefers lottery B for the remainder is often said to have near-risk neutral preferences. Analysts often use the number of "safe choices" (the number of times option A

---

[2]If interest rests solely in creating a single index of risk preference without committing to a single theory, there are some relatively simple methods available such as the one shown in exercise 3.6.3 in Wakker (2010).

Table 1: The H&L Multiple Price List

Lottery A				Lottery B				EVA (€)	EVB (€)	Difference (€)	Open CRRA interval if subject switches to Lottery B (assumes EUT)	
p	€	p	€	p	€	p	€					
0.1	2.00	0.9	1.60	0.1	3.85	0.9	0.10	1.640	0.475	1.17	$-\infty$	-1.71
0.2	2.00	0.8	1.60	0.2	3.85	0.8	0.10	1.680	0.850	0.83	-1.71	-0.95
0.3	2.00	0.7	1.60	0.3	3.85	0.7	0.10	1.720	1.225	0.50	-0.95	-0.49
0.4	2.00	0.6	1.60	0.4	3.85	0.6	0.10	1.760	1.600	0.16	-0.49	-0.15
0.5	2.00	0.5	1.60	0.5	3.85	0.5	0.10	1.800	1.975	-0.18	-0.15	0.14
0.6	2.00	0.4	1.60	0.6	3.85	0.4	0.10	1.840	2.350	-0.51	0.14	0.41
0.7	2.00	0.3	1.60	0.7	3.85	0.3	0.10	1.880	2.725	-0.85	0.41	0.68
0.8	2.00	0.2	1.60	0.8	3.85	0.2	0.10	1.920	3.100	-1.18	0.68	0.97
0.9	2.00	0.1	1.60	0.9	3.85	0.1	0.10	1.960	3.475	-1.52	0.97	1.37
1	2.00	0	1.60	1	3.85	0	0.10	2.000	3.850	-1.85	1.37	$+\infty$

Note: Last four columns showing expected values and implied CRRA intervals were not shown to subjects.

was chosen) or the A-B switching point to describe risk preferences and to infer the shape of an assumed utility function (Bellemare and Shearer, 2010; Bruner et al., 2008; Eckel and Wilson, 2004; Glockner and Hochman, 2011; Lusk and Coble, 2005).

Perhaps the first thing that should be noted about the original H&L MPL is that it entails choices made over only four dollar amounts ($0.10, $1.60, $2.00 and $3.85). Because a utility function is unique only up to an affine transformation, one must fix two of these points and can only identify the relative difference implied by the other two. Stated differently, the original H&L MPL reveals little information about the curvature of the utility function.[3] By contrast, the H&L MPLs entails choices over 11 different probability amounts (from 0 to 1 in increments of 0.1). Thus, the approach contains much more information about the potential shape of the probability weighting function over the entire probability domain.

To more formally address these issues, assume peoples preferences are represented by rank-dependent utility theory introduced by Quiggin (1982) and incorporated into cumulative prospect theory by Tversky and Kahneman (1992). Applying the theory to the H&L

---

[3]One can of course utilize several MPLs and scale up the payoffs as H&L did to allow for a wider range of dollar amounts (thus providing more information on the shape of the utility function). However, those researchers interested in adding a quick and simple risk preference elicitation devise to their studies are unlikely to want to add numerous MPLs simply to get an informed shape of the utility function.

MPL, the rank-dependent utility of option A is $RDU_A = w(p)U(2) + (1 - w(p))U(1.6)$ and the rank-dependent utility of option B is $RDU_B = w(p)U(3.85) + (1 - w(p))U(0.1)$, where p is the probability of receiving the higher payoff amount in each option. A person chooses option A over B when $RDU_A > RDU_B$ or when $w(p)U(2) + (1 - w(p))U(1.6) > w(p)U(3.85) + (1 - w(p))U(0.1)$. Re-arranging, one can see that option A is chosen when:

$$\frac{w(p)}{1 - w(p)} < \frac{U(1.6) - U(0.1)}{U(3.85) - U(2)} \tag{1}$$

Equation (1) reveals two important facts. First, the choice between option A and B in the H&L task is driven both by the shape of $w(p)$ and the shape of $U(x)$ - i.e., it does not separately identify only the curvature of the utility function or the coefficient of relative risk aversion as is often presumed. Second, equation (1) shows that, at most, one can identify only two utility differences $U(1.6) - U(0.1)$ and $U(3.85) - U(2)$, which is clearly a small amount of information to be gleaned about the shape of $U(x)$.

To illustrate the first point, note that many experimental studies have estimated the shape of $w(p)$ using functional forms such as $w(p) = p^{\gamma}/[p^{\gamma} + (1 - p)^{\gamma}]^{1/\gamma}$. Estimates of γ typically fall in the range of 0.56 to 0.71 (e.g., see Camerer and Ho, 1994; Tversky and Kahneman, 1992; Wu and Gonzalez, 1996), which implies an S-shaped probability weighting function that over-weights low probability events and under-weights high probability events.

Now, consider a simple example where individuals have a linear utility function (i.e., they are risk neutral in the traditional sense), $U(x) = x$. With the traditional H&L task, a risk neutral person with $U(x) = x$ and $\gamma = 1$ would switch from option A to B at the fifth decision task. However, if the person weights probabilities non-linearly, say with a value of $\gamma = 0.6$, then they would instead switch from option A to B at the sixth decision task. Thus, in the original H&L decision task, an individual with $\gamma = 0.6$, will *appear* to have a concave utility function (if one ignores probability weighting) *even* though they have a linear utility function, $U(x) = x$. The problem is further exasperated as γ diverges from one. Of course in

reality, people may weight probabilities non-linearly and exhibit diminishing marginal utility of earnings, but the point remains: simply observing the A-B switching point in the H&L decision task is insufficient to identify the shape of $U(x)$ and the shape of $w(p)$. The two are confounded. While it is possible to use data from the H&L technique to estimate these two constructs, $U(x)$ and $w(p)$, *ex post*, we argue that more information is contained about $w(p)$ than $U(x)$ in the original H&L MPL.

In addition to the above arguments that choices in the H&L MPL are likely to provide more information on the shape of $w(p)$ than $U(x)$ relates to the moderate level of payoffs used in many experiments using MPLs. Several authors have argued that the utility function should be linear over relatively low payoff amounts (Selten et al., 1999; Wakker, 2010). If true, this would suggest that the risk averse behavior previously observed in H&L tasks may well relate to $w(p)$ than to $U(x)$. A final piece of evidence suggesting that the original H&L task is more likely to elicit probability weights than utility curvature are the findings that in repeated choice tasks people are more likely to pay attention to the factors changing across the tasks (which in the case of H&L are the probabilities). Because probabilities are changing in the original H&L task, people are more likely to pay attention to this dimension of choice (Bleichrodt, 2002).

2.1 A New MPL

Given the preceding discussion, one might ask if there is a simple way to use a MPL that yields more information about $U(x)$ and, at least in some special cases, avoids the confound between $w(p)$ and $U(x)$. One can indeed achieve such an outcome by following an approach like the one used by Wakker and Deneffe (1996) in which probabilities are held constant. Using this insight, we modify the H&L task such that probabilities remain constant across the ten decision tasks and instead change the dollar payoffs down the ten tasks. Our approach is similar to that used in prior research such as that by Cohen et al. (1987) where certainty equivalents are elicited from subjects by using repeated choices with varying payoff amounts.

Table 2 shows a new MPL. In this MPL, the probabilities of all payouts are held constant at 0.5. We constructed the new MPL shown in table 2 so that it matched the original H&L MPL in terms of the coefficient of relative risk aversion (CRRA) implied by a switch between choosing option A and option B under the assumption of expected utility (EU) preferences. For example, if an individual (with EU preferences) switched from choosing option A to option B on the sixth row of the original H&L task, it would imply a CRRA between 0.14 and 0.41. Likewise, in the new MPL with constant probabilities, a switch from choosing option A to option B on the sixth row would also imply (assuming EU preferences) a CRRA between 0.14 and 0.41.

Table 2: New MPL with Constant Probabilities

Lottery A				Lottery B				EVA (€)	EVB (€)	Difference (€)	Open CRRA interval if subject switches to Lottery B (assumes EUT)	
p	€	p	€	p	€	p	€					
0.5	1.68	0.5	1.60	0.5	2.01	0.5	1.00	1.640	1.506	0.13	$-\infty$	-1.71
0.5	1.76	0.5	1.60	0.5	2.17	0.5	1.00	1.680	1.583	0.10	-1.71	-0.95
0.5	1.84	0.5	1.60	0.5	2.32	0.5	1.00	1.720	1.658	0.06	-0.95	-0.49
0.5	1.92	0.5	1.60	0.5	2.48	0.5	1.00	1.760	1.738	0.02	-0.49	-0.15
0.5	2.00	0.5	1.60	0.5	2.65	0.5	1.00	1.800	1.827	-0.03	-0.15	0.14
0.5	2.08	0.5	1.60	0.5	2.86	0.5	1.00	1.840	1.932	-0.09	0.14	0.41
0.5	2.16	0.5	1.60	0.5	3.14	0.5	1.00	1.880	2.068	-0.19	0.41	0.68
0.5	2.24	0.5	1.60	0.5	3.54	0.5	1.00	1.920	2.272	-0.35	0.68	0.97
0.5	2.32	0.5	1.60	0.5	4.50	0.5	1.00	1.960	2.748	-0.79	0.97	1.37
0.5	2.40	0.5	1.60	0.5	4.70	0.5	1.00	2.000	2.852	-0.85	1.37	$+\infty$

Note: Last four columns showing expected values and implied CRRA intervals were not shown to subjects.

What are the advantages and disadvantages of the new MPL compared to the H&L MPL? At the onset, one can see that because the new MPL only utilizes one probability level, 0.5, it cannot reveal much about the shape of the probability weighting function. However, the new MPL entails choices over 22 different dollar payouts. Thus, the new MPL has the potential to yield much more information about the shape of the utility function than does the traditional H&L MPL.

To consider these ideas more formally, again assume individuals have rank-dependent

preferences and note that option A will be chosen over option B if $w(0.5)U(A^H) + (1 - w(0.5))U(1.6) > w(0.5)U(B^H) + (1 - w(0.5))U(1)$, where A^H and B^H are the higher payoffs for options A and B, respectively (values which changes over the 10 decision tasks), and where $A^H > 1.6$, $B^H > 1$, and $B^H > A^H$. Re-arranging terms, one can see that option A is chosen if:

$$\frac{w(0.5)}{1 - w(0.5)} < \frac{U(1.6) - U(1)}{U(B^H) - U(A^H)} \tag{2}$$

Comparing equation (2) with equation (1), one can see that the original H&L task can utilize 10 points to estimate the function for $w(p)$ but by contrast, the new task can only estimate a single point, $w(0.5)$. In contrast, whereas the original H&L task can only estimate two utility differences, the new task can estimate 11. Thus, the new MPL reveals more information about the shape of $U(x)$ than the original H&L MPL, but the original H&L MPL reveals more information about the shape of $w(p)$ than does the new task.

2.1.1 The New MPL under Original Prospect Theory Preferences

There is one additional feature of the new MPL shown in table (2) that bears mention. Although it does not totally do away with the aforementioned confound between $w(p)$ and $U(x)$ assuming rank-dependent preferences, the confound completely disappears if people weight probabilities as in original prospect theory.

If people weight probabilities according to original prospect theory, in the new MPL they will choose option A when $w(0.5)U(A^H) + w(0.5)U(1.6) > w(0.5)U(B^H) + w(0.5)U(1)$. One can divide both sides of this inequality by $w(0.5)$ to see that option A will be chosen when $U(A^H) + U(1.6) > U(B^H) + U(1)$, or rewriting:

$$1 < \frac{U(1.6) - U(1)}{U(B^H) - U(A^H)} \tag{3}$$

Because equation (3) does not contain the term w(0.5), the choice of option A over B cannot be explained by probability weighting. Stated differently, *even* if an individual weights

probabilities non-linearly in the fashion given by original prospect theory, only the shape of $U(x)$ will dictate their choices in the new MPL show in table 2.

The condition in (3) could only be obtained because of our choice of the probability value 0.5. For any other probability value the weighting function does not drop out and the confound remains.[4] Thus, in the original H&L task (which uses probabilities from 0 to 1), the confound between $w(p)$ and $U(x)$ remains even if preferences are given by original prospect theory.

3   Experiment

To investigate some of the issues discussed above, a laboratory experiment was conducted to compare behavior in the original H&L MPL and our new MPL. Moreover, the experiment was designed to see which MPL (or whether a combination of the two) could better predict a hold-out sample of choices. The next sub-section describes the subjects, recruiting, and experimental environment. Then, we describe the different treatments used in the study.

3.1   Description of the experiment set-up

A lab experiment was conducted using the z-Tree software (Fischbacher, 2007). Subjects consisted of undergraduate students at the University of Ioannina, Greece and were recruited using the ORSEE recruiting system (Greiner, 2004). During the recruitment, subjects were told that they would be given the chance to make more money during the experiment.[5] Stochastic fees have been shown to be able to generate samples that are less risk averse than would otherwise have been observed (Harrison et al., 2009).

Subjects participated in sessions of group sizes that varied from 9 to 11 subjects per session (all but two sessions involved groups of 10 subjects). In total, 100 subjects participated

---

[4]Most empirical estimates suggest that $w(0.3) \approx 0.3$. It is possible to also use this empirical relation to create a MPL that avoids probability weighting.

[5]Subjects were told that "In addition to a fixed fee of 10, you will have a chance of receiving additional money up to 25. This will depend on the decisions you make during the experiment."

in 10 sessions that were conducted between December 2011 and January 2012. Each session lasted about 45 minutes and subjects were paid a 10€participation fee. Subjects were given a power point presentation explaining the risk preferences tasks as well as printed copies of instructions. They were also initially given a five-choice training task to familiarize them with the choice screens that would appear in the real task. Subjects were told that choices in the training phase would not count toward their earnings and that this phase was purely hypothetical.

Full anonymity was ensured by asking subjects to choose a unique three-digit code from a jar. The code was then entered at an input stage once the computerized experiment started. The experimenter only knew correspondence between digit codes and profits. Profits and participation fees were put in sealed envelopes (the digit code was written on the outside) and were exchanged with printed digit codes at the end of the experiment. No names were asked at any point of the experiment. Subjects were told that their decisions were independent from other subjects, and that they could finish the experiment at their own convenience. Average total payouts including lottery earnings were 15.2€ (S.D.=4.56).

3.2 Risk preference elicitation

Our experiment entailed a 2x2 within-subject design, where each subject completed two different multiple price lists (MPL) at two payout (low vs. high) amounts. As shown in Table 3, the baseline (or control) involved the original H&L task at their low payoff amounts (a task we refer to as H&L1).

Table 3: Treatments in experiment

Multiple Price List	Payout	
	low(x1)	high (x5)
H&L	H&L1	H&L5
New MPL with constant probabilities	nMPL1	nMPL5
Hold-out task	H1	H5

The baseline H&L MPL presented subjects with a choice between two lotteries, A or B.

For each lottery choice show in Table 1, a subject chose A or B, and one choice was randomly selected as binding for the payout. The last choice shown in Table 1 is a simple test of whether subjects understood the instructions correctly.[6] The second treatment (H&L5) is identical to the first (H&L1) except that all payouts are scaled up by a magnitude of five.

In addition to the choices in treatments H&L1 and H&L5, subjects also completed the new MPL show in Table 2 (nMPL1) and another set of choices identical to the ones shown in Table 2 except that all payoffs were scaled up by a magnitude of five (nMPL5).

Instead of providing a table of choices arrayed in an ordered manner all appearing on the same page as in H&L, each choice was presented separately showing probabilities and prizes (as in Andersen et al., 2011). Subjects could move back and forth between screens in a given table but not between tables. Once all ten choices in a table were made, the table was effectively inaccessible. The order of appearance of the treatments for each subject was completely randomized to avoid order effects (Harrison et al., 2005). An example of one of the decision tasks is shown in Figure 1.

One of the implicit arguments made thus far is that the original H&L task can better estimate the probability weighting function and the new MPL can better estimate the utility function. As such, a combination of the insights attained by the two approaches might result in a better overall model. To determine whether this combination is indeed "better" than either used alone, we used out-of-sample prediction as our measure of performance. Thus, as shown in Table 3, the study also included two hold-out tasks which we use as the basis of measuring prediction performance. We constructed these hold-out tasks by creating yet another MPL that modified the original H&L design such that the probability of receiving the higher payout option increased nonlinearly down the list (see Table 4). The MPL is constructed so that it matched the original H&L task in terms of the coefficient of relative risk aversion (CRRA) implied by a switch between choosing option A and option B under the assumption that subjects have prospect-theory preferences where they weigh probabilities

---

[6]16 out of 100 subjects failed to pass this test concerning comprehension of lotteries and were omitted from our sample.

Figure 1: Example Decision Task

nonlinearly with $w(p) = p^{0.6}/[p^{0.6} + (1-p)^{0.6}]^{1/0.6}$.

Because each subject completed three MPLs (with 10 choices each) at two payouts, they each made 60 binary choices. For each subject, one of the 60 choices was randomly chosen and paid out.

4   Data analysis and results

4.1   Descriptive analysis

Figure 2 illustrates the proportion of subjects choosing option A for the original H&L task and the new task for small and large payoff amounts. Note that all four tasks were designed to elicit the same switching point for a given risk aversion coefficient under the assumption of expected utility preferences (an assumption we will show later to be descriptively invalid).

The two H&L tasks and the nMPL5 tasks, imply significant risk averse behavior as subjects switch, on average, far after task four. However, low-payoff new task, nMPL1,

Table 4: Holdout Task

Lottery A				Lottery B				EVA (€)	EVB (€)	Difference (€)	Open CRRA interval if subject switches to Lottery B (assumes EUT)	
p	€	p	€	p	€	p	€					
0.03	2.00	0.97	1.60	0.03	3.85	0.97	0.10	1.610	0.194	1.42	$-\infty$	-1.71
0.09	2.00	0.91	1.60	0.09	3.85	0.91	0.10	1.636	0.439	1.20	-1.71	-0.95
0.20	2.00	0.80	1.60	0.20	3.85	0.80	0.10	1.678	0.835	0.84	-0.95	-0.49
0.34	2.00	0.66	1.60	0.34	3.85	0.66	0.10	1.735	1.365	0.37	-0.49	-0.15
0.50	2.00	0.50	1.60	0.50	3.85	0.50	0.10	1.800	1.975	-0.17	-0.15	0.14
0.66	2.00	0.34	1.60	0.66	3.85	0.34	0.10	1.865	2.585	-0.72	0.14	0.41
0.80	2.00	0.20	1.60	0.80	3.85	0.20	0.10	1.922	3.116	-1.19	0.41	0.68
0.91	2.00	0.09	1.60	0.91	3.85	0.09	0.10	1.964	3.512	-1.55	0.68	0.97
0.97	2.00	0.03	1.60	0.97	3.85	0.03	0.10	1.990	3.756	-1.77	0.97	1.37
1	2.00	0	1.60	1	3.85	0	0.10	2.000	3.850	-1.85	1.37	$+\infty$

Note: Last four columns showing expected values and implied CRRA intervals were not shown to subjects.

where probabilities are held constant, a different picture emerges. Subjects appear less risk averse in the constant-probability task than in the conventional H&L task.

One striking difference in nMPL1 task is the fact that the percent choosing option A remains at about 50% for the first five decision task, and, in fact, slightly increases over this range. One explanation for this trend is that the new task generated more multiple switching points than the standard H&L task.[7] If we calculate the number of choices that violate monotonicity, we find that the average subject made 0.21 and 0.11 such violations in the original H&L task at low and high payouts, respectively. By contrast, in our new MPL tasks with constant probabilities, the average subject made 0.85 and 0.69 such violations in the low and high payout tasks, respectively. Over the first few choices in the new decision task at low payoffs (nMPL1), the difference in the expected values between lottery options A and B were relatively small, and this might partially explain why the task generated more switching behavior. However, it should be noted that such small differences in expected values were *required* to generate the same implied CRRA intervals as the original H&L task

---

[7]In our experiment, we did not impose monotonicity on choices or provide warnings when monotonicity was violated. Although such a procedure could be implemented, it is unclear if it is superior to simply observing how people behave when unconstrained.
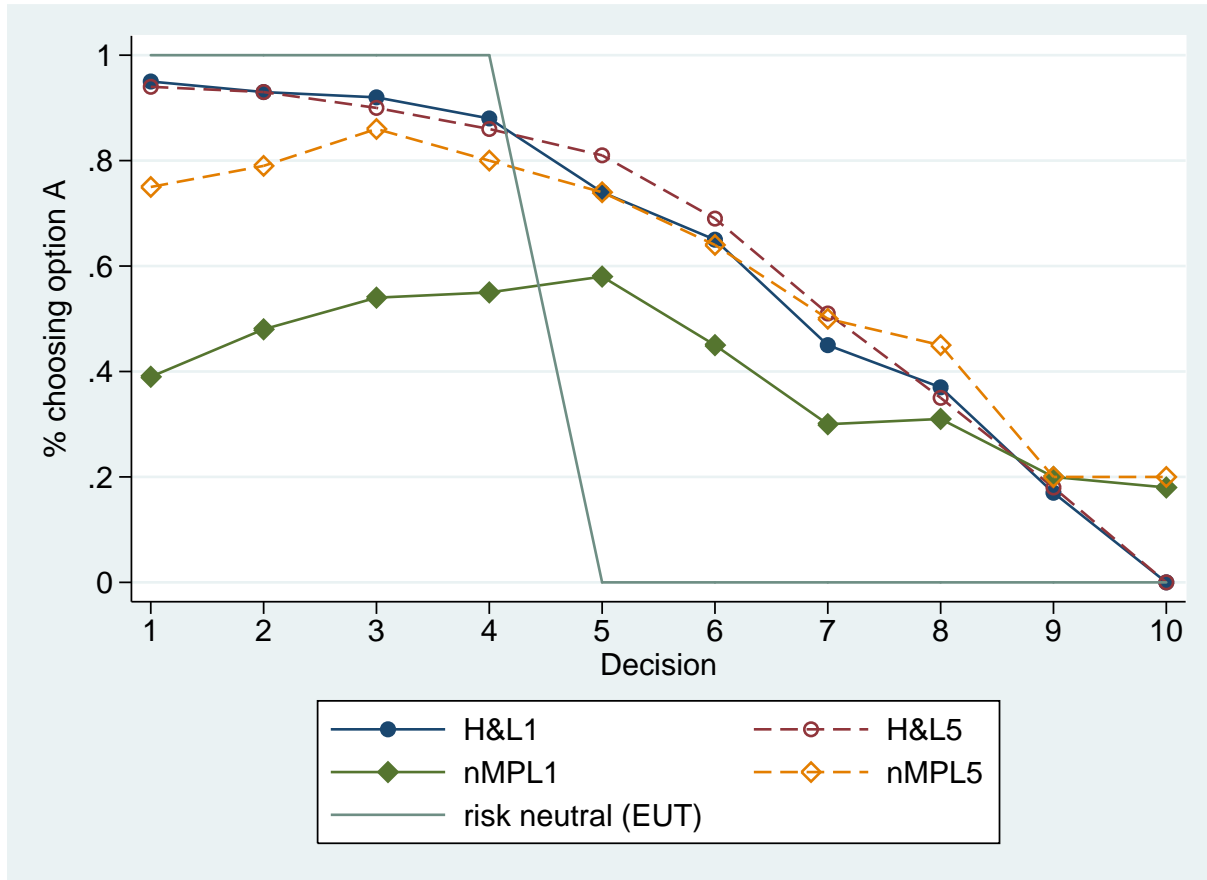
Figure 2: Percentage of respondents choosing option A for each decision task

given the overall payout magnitudes. Thus, this is not a feature of the new task *per se* but rather a feature of constant relative risk aversion and expected utility theory applied to lotteries with payouts of the magnitude considered in the original H&L task but with constant probabilities. Importantly, we have analyzed our data removing individuals that significantly violated monotonicity (i.e., made three or more inconsistent choices), and our econometric estimates (discussed momentarily) are virtually unchanged.

Figure 2 also illustrates the effects of scaling off payoffs. For the traditional H&L task, increasing payoffs had very little effect on the percentage of times option A was chosen. However, increasing payoffs had a much larger effect on our new MPL. The issue of monotonicity does not appear as problematic in the new MPL when payoffs are scaled up by a factor of five. This might be because the expected value differences between options A and

B (shown in table 2) are also scaled up by a factor of five in this task.

4.2 Econometric modeling approach

To explore the results in terms of the curvature of the utility and probability weighting functions, we utilize the random utility approach also used by Andersen et al. (2008) and use the rank-dependent utility model as the base-line model of analysis. Letting the random rank-dependent utility of option A experienced by individual i in choice j be:

$$V_{ij}^A = Z_{ij}^A + \varepsilon_{ij}^A \tag{4}$$

and similarly for option B:

$$V_{ij}^B = Z_{ij}^B + \varepsilon_{ij}^B \tag{5}$$

where ε_{ij} is a stochastic error term assumed to be known to the individual but unobservable to the analyst, and where Z_{ij}^A and Z_{ij}^B are the systematic portions of the utility functions assumed to follow rank-dependent preferences, i.e., $Z_{ij}^A = w(p_j)U(A_j^H) + (1 - w(p_j))U(A_j^L)$, where A_j^H is the high payoff and A_j^L is the low payoff for option A in choice j.

The probability of option A being chosen over option B is the probability that $V_{ij}^A > V_{ij}^B$. Assuming the difference in the error terms is distributed i.i.d. normal with standard deviation equal to σ, the probability A is chosen is:

$$P_{ij}^A = \Phi((Z_{ij}^A - Z_{ij}^B)/\sigma) \tag{6}$$

Thus, a log likelihood function can be defined for estimation:

$$LLF = \sum_{i=1}^{N} \sum_{j=1}^{J} [y_{ij} ln(P_{ij}^A) + (1 - y_{ij}) ln(1 - P_{ij}^A)] \tag{7}$$

where $y_{ij} = 1$ if option A is chosen and $y_{ij} = 0$ if option B is chosen. If an individual

indicates indifference to A and B, $y_{ij} = 0.5$.

In the analyses that follow, we consider several specifications for $w(p)$ and $U(x)$. The base-line specifications for the utility function is the constant relative risk aversion specification: $U(x) = \frac{x^{1-r}}{1-r}$, where r is the coefficient of relative risk aversion.[8] In the new MPLs we have many more points on the utility function and can also estimate an expo-power utility function (Saha, 1993): $U(x) = (1 - exp(-\alpha x^{1-r}))/\alpha$.

For the probability weighting function in the H&L MPLs, we consider the function used by Tversky and Kahneman (1992) and others: $w(p) = p^{\gamma}/[p^{\gamma} + (1-p)^{\gamma}]^{1/\gamma}$. We also estimated the probability weighting function proposed by Prelec (1998): $w(p) = exp(-\beta(-ln(p)^{\tau}))$. For the new MPL, there is only a single probability point and thus we need only estimate a single parameter, θ, representing the weight placed on the 0.5 probability, i.e., $w(0.5) = \theta$.

We estimated each of these competing specifications separately for high and low payoffs (note: likelihood ratio tests reject the hypothesis of equality of parameters across high and low payoffs for each specification). Moreover, we used the AIC and BIC model selection criteria to determine the best fitting model for each dataset.

4.3   Econometric Results

For the traditional H&L MPLs, each of the aforementioned model variations was estimated (see the appendix Table A.1). For both high and low payoffs, the AIC and BIC model selection criteria indicate a preference for the rank-dependent models over the prospect-theory models. Within the rank-dependent models, the Tversky and Kahneman (1992) probability weighting function is preferred to the Prelec (1998) weighting function according to the AIC and BIC.

The preferred model for both the H&L1 and H&L5 treatments is the CRRA model with

---

[8]In the original H&L task, we can also estimate a "non-parametric" utility function and instead estimate the two utility differences shown in equation (1): $[U(1.6) - U(0.1)]$ and $[U(3.85) - U(2)]$. In this latter case, however, the standard deviation, σ, is no longer separately identified and must be normalized to one. In the H&L MPL, this formulation is actually observationally equivalent to the CRRA specification with σ freely estimated; both utility specifications give identical maximum likelihood function values and probability weighting estimates. These results are shown in the appendix Table A.1.

the Tversky and Kahneman (1992) probability weighting function. For low-payoffs, the estimate of the coefficient of relative risk aversion ($r = 0.004$) was not statistically different from zero, but the estimated parameter on the probability weighting function, $\gamma = 0.501$, was statistically different from one, indicating a rejection of the expected utility model in favor of the rank-dependent model. Similarly, for high-payoffs, the coefficient of relative risk aversion ($r = -0.135$) was not statistically different from zero, but the estimated parameter on the probability weighting function, $\gamma = 0.45$, was statistically different from one according to a likelihood ratio test. Thus, at least for our subjects, the apparent risk averse behavior shown in Figure 2 is solely a result of probability weighting rather than utility function curvature for the conventional H&L tasks. The implication is that practitioners using the H&L task to infer curvature of the utility function would have arrived at erroneous conclusions had they not also jointly estimated the extent to which people weight probabilities non-linearly.

Figure 3 plots the estimated probability weighting functions for the low-payoff H&L1 MPL. In addition to the two aforementioned functional forms, we also show the results of a "non-parametric" estimation in which a single parameter is estimated for each of the 11 probability points available in the H&L task (with the lowest normalized to zero and the highest normalized to one). Although the "non parametric" form is not preferred to the parametric forms according to AIC and BIC, the results reveal the level of information about probability weights obtainable from the H&L task. In all specifications, the results reveal significant over-weighting of low probability events and under-weighting of high probability events.

Turning to the new MPLs, the AIC and BIC indicate that the most preferred models (see full results in appendix Table A.2) are the models assuming constant relative risk aversion with the error variance normalized to one for both the low and high payoff tasks. The results reveal that at low payoffs, the coefficient of relative risk aversion ($r = 0.116$) was not statistically different from zero and the estimated weight applied to probability 0.50 was $w(0.5) = 0.581$, a difference ($0.581 0.50 = 0.081$) which is not statistically different from zero,
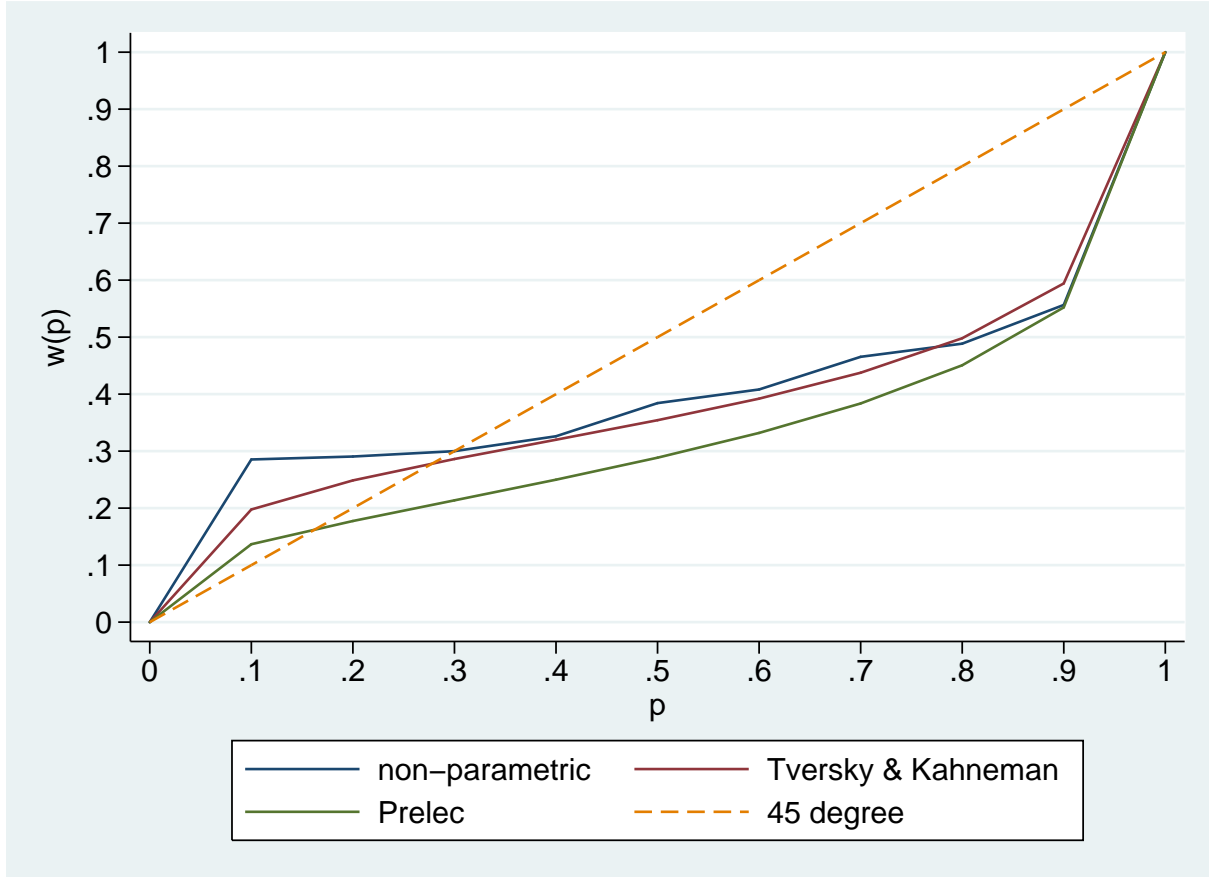
18

Figure 3: Percentage of respondents choosing option A for each decision task

implying linear probability weighting in the vicinity of $p = 0.05$. Taken together, for low payoffs, the estimates imply near risk-neutral behavior for the new MPL.

At high payoffs, however, a different picture emerges. For our new MPL which utilizes much more variation over payoff amounts than the H&L task, we find that for high payoffs, a statistically significant estimate for the coefficient of relative risk aversion ($r = 0.233$) emerges. Moreover, we find that the estimated weight applied to probability 0.50 was $w(0.5) = 0.366$. This estimate of probability weighing is very similar to that implied by the high-payoff H&L task (with $\gamma = 0.45$, the H&L5 task implies $w(0.5) = 0.313$).

4.4 Prediction performance

Because of the larger variation in probabilities in the H&L task, we have argued that this task should yield better estimates of the curvature of the probability weighting function. By contrast, because of the larger variation in the dollar amounts in our new MPL, we have argued that this task should yield better estimates of the curvature of the utility function. To put these conjectures to the test, we now see how well the aforementioned estimates are able to predict the holdout tasks at low (H1) and high (H2) payoff amounts.

In particular, we compare the predictive performance of three models: i) a model based on the estimate of r and γ from the H&L MPL, ii) a model based on the estimate of r (and for lack of a better choice assuming $\gamma = 1$) from the nMPL, and finally iii) a composite model in which we use the estimate of r from our new MPL and the estimate of γ from the H&L MPL.[9] To judge predictive fit, we use two criteria: 1) the percent of correct predictions and 2) the value of the likelihood function observed at out-of-sample values — the out-of-sample log-likelihood function (OSLLF). The out-of-sample log-likelihood function approach has long been used in the marketing literature for model selection (Erdem, 1996; Roy et al., 1996) and further elucidated in the economics literature by Norwood, Roberts, and Lusk (2004); Norwood, Lusk, and Brorsen (2004). The OSLLF has desirable properties in judging the predictive fit of discrete choice models and it is our preferred selection criteria.

Table 5 shows the performance of the three models in predicting the out-of-sample hold-out choices. For low payoffs, the composite model generates the same % correct predictions but has a lower OSLLF than the H&L MPL. Although a paired-test indicates no significant difference in the composite-model and H&L1 OSLLF values, the non-parametric sign-rank test indicates the two are significantly different (p-value < 0.01). The composite model

---

[9]Our composite model takes the estimate of the curvature of the probability weighting from the H&L task and the estimate of the curvature of the utility function from the new MPL. An alternative approach is to pool the two data sets and estimate a combined model. When we do this for the low payoff task, we find an estimate of $r = 0.3249$ and $\gamma = 0.6904$, and $\sigma = 0.5883$, all of which are significantly different from zero. However, this model exhibits significantly poorer out of sample predictions with the OSLLF$= -0.5222$ and %correct predictions$=76.42\%$ than the composite model discussed in the main text. A similar result holds for the high payoff task.

outperforms the nMPL both in terms of percent of correct predictions and in terms of OSLLF (both the t-test and signed-rank test indicate OSLLFs are significantly different at $p < 0.01$ level).

Table 5: Out-of-sample prediction performance of three competing models

	H&L predictions	nMPL predictions	Composite model predictions
low payoffs			
OSLLF	-0.4068	-0.4677	-0.4025
% correct	83.57%	76.43%	83.57%
high payoffs			
OSLLF	-0.4696	-1.0565	-0.3876
% correct	83.09%	83.09%	82.60%

A similar result is obtained for the high-payoff values. Although all three models generate similar performance in terms of the percentage of correct predictions, the composite model far outperforms the H&L5 and nMPL5 tasks in isolation according to the OSLLF values (the composite model yields significantly different OSLLF values as compared to the H&L5 and nMPL5 tasks according to t-tests and signed-rank tests at the $p < 0.01$ level).

Taken together, the results in Table 5 largely confirm our intuition that better predictions can be made by using the H&L task to infer the curvature of the probability weighting function and the new MPL to infer the curvature of the utility function.

5 Conclusion

Although H&L introduced a useful tool for characterizing risk taking behavior, their approach is limited in being able to identify why a particular behavior under risk was observed. Risk averse behavior could result from curvature of the utility function, curvature of the probability weighting function, or both. The obvious implication is that caution should be taken in directly using a single number like "number of safe choices" from H&Ls risk preference elicitation method to infer curvature of the utility function, the theoretical concept

that is often of interest, because risk averse behavior may be driven by probability weighting. In fact, we show that, if anything, the H&L task is probably best suited to measuring the curvature of the probability weighting function.

We introduced a modified version of the H&L task which held probabilities constant at 0.50 and provided much more variation in the payoff amounts. By providing more variation in payoff amounts, we hoped to obtain better estimates of the curvature of the utility function. By and large, that's what our experimental results imply. At both low and high payoff amounts, econometric estimates suggest that behavior is almost totally driven by the curvature of the probability weighting function (the estimated CRRA is not significantly different from zero in either case). Only with our new MPL under high payoffs did we observe significant curvature in the utility function.

To test our intuition about the relative merits of the two elicitation approaches, we sought to determine whether a composite model that combined the estimate of the curvature of the utility function from our new MPL with the estimate of the curvature of the probability weighting function from the H&L task would exhibit better out-of-sample prediction performance with a hold-out task than either model used in isolation. Our results implied that the composite model did indeed generate lower OSLLF values than the estimates from the conventional H&L task or the MPL used alone.

References

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

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

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutstrom (2011). Discounting behavior: A reconsideration. *Center for the Economic Analysis of Risk, Working Paper 2011-03*.

Becker, G. M., M. H. Degroot, and J. Marschak (1964). Measuring utility by a single-response sequential method. *Behavioral Science 9*(3), 226–232.

Bellemare, C. and B. Shearer (2010). Sorting, incentives and risk preferences: Evidence from a field experiment. *Economics Letters 108*(3), 345–348.

Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural india. *American Journal of Agricultural Economics 62*(3), 395–407.

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

Bleichrodt, H. (2002). A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health Economics 11*(5), 447–456.

Bruner, D., M. McKee, and R. Santore (2008). Hand in the cookie jar: An experimental investigation of equity-based compensation and managerial fraud. *Southern Economic Journal 75*(1), 261–278.

Camerer, C. F. and T.-H. Ho (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty 8*(2), 167–196.

Cohen, M., J.-Y. Jaffray, and T. Said (1987). Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. *Organizational Behavior and Human Decision Processes 39*(1), 1–22.

Eckel, C. C. and R. K. Wilson (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization 55*(4), 447–465.

Erdem, T. (1996). A dynamic analysis of market structure based on panel data. *Marketing Science 15*(4), 359–378.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Glockner, A. and G. Hochman (2011). The interplay of experience-based affective and probabilistic cues in decision making. *Experimental Psychology 58*(2), 132–141.

Greiner, B. (2004). An online recruitment system for economic experiments. In K. Kremer and V. Macho (Eds.), *Forschung Und Wissenschaftliches Rechnen. Gwdg Bericht 63. Ges. Fr Wiss*, pp. 79–93. Datenverarbeitung, Gttingen.

Harrison, G. W., E. Johnson, M. M. McInnes, and E. E. Rutstrom (2005). Risk aversion and incentive effects: Comment. *American Economic Review 95*(3), 897–901.

Harrison, G. W., M. I. Lau, and E. E. Rutstrom (2009). Risk attitudes, randomization to treatment, and self-selection into experiments. *Journal of Economic Behavior & Organization 70*(3), 498–507.

Harrison, G. W. and E. E. Rutstrom (2008). Risk aversion in the laboratory. In J. C. Cox and G. W. Harrison (Eds.), *Research in Experimental Economics Vol 12: Risk Aversion in Experiments*, Volume 12, pp. 41–196. Bingley, UK: Emerald Group Publishing Limited.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review 92*(5), 1644–1655.

Holt, C. A. and S. K. Laury (2005). Risk aversion and incentive effects: New data without order effects. *American Economic Review 95*(3), 902–904.

Lusk, J. L. and K. H. Coble (2005). Risk perceptions, risk preference, and acceptance of risky food. *American Journal of Agricultural Economics 87*(2), 393–405.

Miller, L., D. E. Meyer, and J. T. Lanzetta (1969). Choice among equal expected value alternatives: Sequential effects of winning probability level on risk preferences. *Journal of Experimental Psychology 79*(3), 419–423.

Norwood, B. F., J. L. Lusk, and B. W. Brorsen (2004). Model selection for discrete dependent variables: Better statistics for better steaks. *Journal of Agricultural and Resource Economics 29*(3), 404–419.

Norwood, B. F., M. C. Roberts, and J. L. Lusk (2004). Ranking crop yield models using out-of-sample likelihood functions. *American Journal of Agricultural Economics 86*(4), 1032–1043.

Prelec, D. (1998). The probability weighting function. *Econometrica 66*(3), 497–528.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization 3*(4), 323–343.

Roy, R., P. K. Chintagunta, and S. Haldar (1996). A framework for investigating habits, "the hand of the past," and heterogeneity in dynamic brand choice. *Marketing Science 15*(3), 280–299.

Saha, A. (1993). Expo-power utility: A flexible form for absolute and relative risk aversion. *American Journal of Agricultural Economics 75*(4), 905–913.

Selten, R., A. Sadrieh, and K. Abbink (1999). Money does not induce risk neutral behavior, but binary lotteries do even worse. *Theory and Decision 46*(3), 213–252.

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

Wakker, P. and D. Deneffe (1996). Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management Science 42*(8), 1131–1150.

Wakker, P. P. (2010). *Prospect theory for risk and ambiguity.* Cambridge, UK: Cambridge University Press.

Wu, G. and R. Gonzalez (1996). Curvature of the probability weighting function. *Management Science 42*(12), 1676–1690.

A   Appendix

Table A.1: Competing estimates for the H&L1 and H&L5 MPLs

| | Rank Dependent Theory | | | | Prospect Theory | |
| | T&K weight function | | Prelec weight function | | T&K weight function | Prelec weight function |
Parameters	Non-Parametric	CRRA	Non-Parametric	CRRA	Non-Parametric	Non-Parametric
low payoff						
r	—	0.004 (0.183)	—	-0.162 (0.320)	—	—
$U(1.6)-U(1)$	3.312* (0.27)	—	2.696* (1.995)	—	5.159* (1.729)	4.388* (2.705)
$U(3.85)-U(2)$	4.063* (1.286)	—	4.150* (3.444)	—	3.419* (1.359)	3.681* (2.764)
γ	0.501* (0.066)	0.501* (0.066)	—	—	0.479* (0.099)	—
σ	1	0.454* (0.06)	1	0.529 (0.391)	1	—
α	—	—	0.392 (0.337)	0.392 (0.337)	—	1.31* (0.391)
τ	—	—	1.435* (0.583)	1.435* (0.583)	—	0.164* (0.11)
LLF	-349.9	-349.9	-349.15	-349.15	-354	-351.65
AIC	705.8	705.8	706.3	706.3	714	711.3
BIC	720	720	725.2	725.2	728.2	730.2
high payoff						
r	—	-0.135 (0.218)	—	-0.314 (0.324)	—	—
$U(8)-U(5)$	3.447* (0.300)	—	2.077* (0.826)	—	5.792* (2.131)	3.864* (1.582)
$U(19.25)-U(10)$	5.118* (1.846)	—	3.901 (2.507)	—	3.809* (1.653)	3.119 (1.622)
γ	0.450* (0.062)	0.450* (0.062)	—	—	0.445* (0.093)	—
σ	1	2.592* (0.608)	1	5.483 (2.869)	1	—
α	—	—	0.461 (0.279)	0.461 (0.279)	—	1.548* (0.385)
τ	—	—	1.818* (0.487)	1.818* (0.487)	—	0.157* (0.073)
LLF	-347.85	-347.85	-347.1	-347.1	-354.1	-350.85
AIC	701.7	701.7	702.2	702.2	714.2	709.7
BIC	715.9	715.9	721.2	721.2	728.4	728.6

Note: **(*) Statistically significant at the 5%(10%) level. Numbers in parenthesis denote standard errors.

Table A.2: Competing estimates for the nMPL1 and nMPL5 tasks

	CRRA	normalized CRRA	expo power
low payoff			
r	-0.610 (1.518)	0.116 (0.180)	0.368 (1.341)
θ	—	—	-0.340 (2.275)
$w(0.5)$	0.517* (0.153)	0.581* (0.072)	0.581* (0.074)
σ	2.376 (4.339)	1	1
LLF	-539.6	-539.7	-539.65
AIC	1085.2	1083.4	1085.3
BIC	1099.4	1092.9	1099.5
high payoff			
r	1.022 (0.949)	0.233* (0.030)	-0.083 (0.171)
θ	—	—	0.048* (0.009)
$w(0.5)$	0.526* (0.192)	0.366* (0.011)	0.434* (0.062)
σ	0.134 (0.337)	1	1
LLF	-477.1	-477.45	-476.9
AIC	960.2	958.9	959.8
BIC	974.4	968.4	974

Note: **(*) Statistically significant at the 5%(10%) level.
Numbers in parenthesis denote standard errors.