An Experimental Study of Decision-Making under Uncertainty -- Individual, Group and Panel Data

Chinn-Ping Fan, Mei-Hua Tsai and Bih-Shiow Chen

Abstract

With two- and three-year panel data of lottery pricing experiment, we estimate four variants of model on decision making under uncertainty with individual and pooled group data. We first study the stability of individual behavior. Although there exists some very stable subjects, but for most subjects, not all the parameters of two consecutive years are equal, although these significant differences are often very small in values. Our subjects are quite stable in terms of the nonlinearity of $w(p_i)$, even though the exact functional form of $w(p_i)$ may change. As for the comparison of individual and group estimations, we find a high degree of correspondence in terms of parameter values. But in terms of model selection, it is possible to reach different conclusion from individual and group estimations. Finally, for the comparison of EUT versus PT, our data clearly show that. PT is the better theory for most subjects, and especially so if requiring for two year consistency.

Keywrods: Choice under uncertainty, Expected utility theory, Prospect theory, Panel experiment, Individual estimation

JEL Classifications: C81; C91; D81

C. Fan
Corresponding author, Department of Economics, Soochow University. e-*mail*: cpfan@scu.edu.tw.

M. Tsai
Department of Business Management, Chin Min Institute of Technology, Miaoli, Taiwan,
e-*mail*:mhtsai@ms.chinmin.edu.tw

B. Chen
Department of Economics, Soochow University.
e-*mail*: bschen@scu.edu.tw.

1. Introduction

Within the vast realm of experimental literatures studying decision-making under uncertainty, two research strategies could be identified; one treats a single individual as the unit of analysis and the other analyzes the pooled data of a group of subjects. In this paper, these two approaches will be referred to as individual estimation and group estimation. To give just two examples: Tversky and Kahneman (1992) analyzed data of individual subject behavior[1], while Loomes, Moffatt and Sugden (2002) pooled the observations from all subjects to make statistical inferences.

These two research strategies tackle the same issue from different perspective, each has it's own significance. It may seem that individual estimation is closer to the theoretical context, but the question of stability of individual behavior becomes an important issue. And quotation[2] like following makes one worries about this stability.

> *…in a number of tightly-controlled experiments in which subjects have confronted exactly the same pairwise choice problem on two occasions, separated only by a short time interval, the proportion who choose differently in the two cases has often been found to be of the order of 20 to 30 per cent.*

It is generally acknowledged that subject behavior contains stochastic component. But, could it be possible that besides stochastic variations, there are also structure changes? If this is the case, then we may observe the same subject in different times and reach different conclusions. Our first research question stems from this concern. We collected panel data for

[1] For example, Table 4 in p. 308 listed the percentage of risk-seeking choices of 25 individual subjects.
[2] Loomes, Moffatt and Sugden (2002), in the last paragraph of p. 103.

the same subjects and tested for structure changes with individual estimations.

Our second research question focuses on the comparison of individual estimations and group estimations. For the same group of subjects, logic requires that the conclusions from the two be more or less supportive of each other. If this is not the case, if the results from individual estimations and group estimations are contradictory, then researchers would not know which one to believe. Also, we need to have this consistency established so that we may directly compare literature results using different estimation strategies.

And finally, our third research question concerns the testing of theories of individual decision making under uncertainty. There are many good survey papers, for example, Camerer (1995) and Stamer (2000), discussing various competing theories. But in order to keep a proper focus on the first two research questions, so within the many candidates, we estimate and compare only expected utility theory (abbreviated EUT hereafter) and prospect theory (abbreviated PT), others are left for future research. With the many individual and group estimations of this research, we want to see if our data provide justification for the probability weights of EUT or PT.

Section 2 describes the background in more details; Section 3 explains the experiment designs and procedures. The experimental findings are reported in Section 4 and Section 5 concludes this paper.

2. Background

In this research, an uncertainty is represented as a three-outcome lottery $L = (\boldsymbol{x}, \boldsymbol{p})$ where $\boldsymbol{x} = (x_1, x_2, x_3)$ is the ranked outcome vector with $x_1 > x_2 > x_3$ and $\boldsymbol{p} = (p_1, p_2, p_3)$ is the probability vector. The utility function over lottery could be written in the general format: $U(L) = \sum_{i=1}^{3} \pi(p_i) u(x_i)$, where $\pi(p_i)$ is the decision weighting function and $u(x_i)$ is the value function[3].

According to EUT, $\pi(p_i)$ is the identity function, i.e., probabilities enter into $U(L)$ as linear weights. On the other hand, PT, among other things, suggests a nonlinear probability weighting function since people often overestimate small probabilities and underestimate large probabilities. In order to accommodate multi-outcome lotteries, Tversky and Kahneman (1992) proposed that the decision weights should be constructed by probability weights so that $\pi(p_i) = w\left(\sum_{j=1}^{i} p_j\right) - w\left(\sum_{j=1}^{i-1} p_j\right)$. And since people may behave differently for gains and losses, so $\pi(p_i)$ should be estimated for these two outcome regions separately.

Based on previous literature results, we estimate the following functional forms in this paper. For value function, we assume the power function $u(x_i) = x_i^{\alpha}$, this is probably the most commonly used form[4]. As for the probability weighting function, EUT suggested the identity function of Equation (1),

---

[3] Or in EUT, $u(x_i)$ would be referred to as the monetary utility function.

[4] We have also tried the quadratic functional form, but consistent with literature results, power function is generally better.

$$w(p_i) = p_i . \qquad (1)$$

And for PT, scholars tested for various functional forms. The original paper of Tversky and Kahneman (1992) introduced Equation (2),

$$w(p_i) = \frac{p_i^r}{\left(p_i^r + (1-p_i)^r\right)^{\frac{1}{r}}} . \qquad (2)$$

Prelec (1998) suggested two functional forms, Equation (3) and (4),

$$w(p_i) = \exp\left(-\left(-\ln p_i\right)^r\right), \qquad (3)$$

$$w(p_i) = \exp\left(-s\left(-\ln p_i\right)^r\right). \qquad (4)$$

Many researches[5] find the two-parameter $w(p_i)$ of Equation (4) to be a good choice. Scott (2006) considered a total of 256 model variants with various forms of value function, probability weighting functions and stochastic error specifications. He concluded that, together with the power value function and a Logit stochastic error, the one-parameter $w(p_i)$ of Prelec (1998), Equation (3), is the best functional form for probability weights.

In this research, with the power value function and four probability weighting functions listed above in Equations (1) to (4), we construct and estimate four $U(L)$ models. The three PT models, with respectively, Equations (2), (3) and (4) as $w(p_i)$, are abbreviated as the T&K, Pr1, and Pr2 models.

Also note that, with the same purpose of eliciting subjects' true preference, different researches require subjects to perform different tasks. In Tversky and Kahneman (1992),

[5] For example, Gonzalez and Wu (1999), and Bleichrodt and Pinto (2000). For more discussion, see also Scott (2006), p.110.

subjects compared lotteries with fixed sums of money and performed a series of ranking. Hey and Orme (1994) and many others asked subjects to make pairwise choices. James (2007) employed the BDM procedure of Becker, DeGroot and Marschak (1964) in which subjects determined the valuations of the lotteries. The issue of preference elicitation is in itself an important research topic and the famous preference reversal example reminds us that the behavior results may not be procedure invariant. In this paper, we adopt the BDM procedure similar to James (2007) and leave the comparison of other elicitation methods for future work.

In the BDM procedure, the valuations determined by the subjects could be either the willingness to pay (WTP) or the willingness to accept (WTA), depending on the specific setup. These valuations will all be generally abbreviated as *CE* (the certainty equivalent values) of the lottery in this paper. Our research consisted of a total of ten sessions of experiment, one session involved *K* lotteries, L_k, $k = 1,...,K$. In each session, the data we collected from each subject were $CE(L_k)$, $k = 1,...,K$. Assuming that $U(L_k)$ and $CE(L_k)$ had the same ranking over the *K* lotteries, so $CE(L_k) \approx U(L_k) = \sum_{i=1}^{3} \pi(p_{ki}) u(x_{ki})$.

Finally, we need to specify the stochastic error term. For the binary experimental data collected from pairwise choices, Loomes, Moffatt and Sugden (2002) and Scott (2006) contains comprehensive discussions on the specifications of the stochastic error. But the *CE* data collected from the BDM procedure are continuous variables, so we use a simpler specification similar to that of James (2007). We assume additive error terms ε_k with independent identical distributions with zero means and finite variances. Hence, with the *K* three-outcome lotteries in a session, we estimate four variants of model Equation (5), with the four $w(p_i)$ specified above.

$$CE(L_k) = \sum_{i=1}^{3} \pi(p_{ki})u(x_{ki}) + \varepsilon_k \tag{5}$$

A key element of our research design is to recruit the same group of subjects repeatedly and collect panel data. There are researches with time factor incorporated into the experiment design. For example, in Tversky and Kahneman (1992) and Carbone and Hey (2000), subjects participated in severa; sessions of experiment separated by a period of a few days. And in James (2007), the 52 rounds of BDM evaluation consisted of four 13-round periods with different setups. The time spans of these literatures are all quite short. To our knowledge, we are the first to collect two- and three-year panel data and test for long-term stability.

When pooling the data of all subjects together to run group estimation, a common interpretation is that we study the behavior of some fictitious representative individual. Harrison and Ruström (2006) introduced a different concept. They suggested the possibility that both EUT and PT are valid; each used by some proportions of the subjects. And hence, they estimated the respective model parameters and also the proportions with which the two theories are mixed. Our research design is different, and we will compare our results with theirs in Section 4.

3. Experiment

Two groups of subjects were recruited. Group I is a three-year panel consisted of 26 subjects of the freshmen class of 2005, there were 16 econ majors and 10 law majors, 8 male and 18 female. Group I participated the experiment for three consecutive years, from 2005 to 2007. Group II is a two-year panel of 2006 and 2007, with 19 subjects, 12 econ majors and 7 law majors, 9 male and 10 female, they were freshmen class of 2006. With this panel structure, data sets will be named according to the following convention. Roman numbers I and II refer to Group I and Group II (with subjects I1 to I26 and II1 to II19). Letters G and L refers to gain and loss lotteries; and the number that follows indicates the year of the experiment. We have run a total of ten sessions of experiment and collected the following ten data sets: IG5, IG6, IG7, IL5, IL6, IL7 for Group I and IIG6, IIG7, IIL6, IIL7 for Group II.

The experiment was programmed with the software z-Tree (Fischbacher 2007), and conducted in the Computer Lab of Department of Economics, Soochow University, Taipei, Taiwan. We use the BDM procedure to elicit subjects' valuations of the lotteries. Figure 1 is an example of a gain lottery shown on the computer screen. Subjects were asked to determine the WTA for this lottery. Afterwards, a market price would later be randomly selected. If the WTA was less or equal to the market price, then the lottery was successfully sold and subjects received the market price as reward. And if WTA was higher than the market price, the lottery was not sold but subjects may still receive earning, pending on the operating of a random mechanism corresponding to the probability structure of the lottery.

I would like to sell this lottery for __________NT dollars.

Reward	$200	$50	$0
Probability	0.08	0.02	0.90

Figure 1 A sample gain lottery

The loss lotteries were described as possible financial damages from natural disasters. Subjects determined how much they were willing to pay (*WTP*) to purchase an insurance policy to cover this natural disaster. The market price of the insurance premium was also determined randomly. If *WTP* was smaller than the market premium, the subject failed to purchase the insurance policy, so he might later incur various losses. And if *WTP* was greater than or equal to the market premium, a subject would successfully purchase an insurance policy. This subject would pay the insurance premium and any future losses would be fully covered by the insurance.

I would like to pay __________NT dollars to purchase insurance policy.

Losses	\$0	\$−50	\$−150
Probability	0.55	0.25	0.20

Figure 2 A sample loss lottery

In 2005, there were 96 gain and 24 loss lotteries constructed with 5 outcome vectors and 24 probability vectors. The lottery sets of 2006 and 2007 were identical, there were 54 gains and 54 losses constructed with 18 outcome vectors and 6 probability vector[6]. The differences of these lottery sets were quite large. The 05 lottery sets involved seven monetary values ranging from NT\$-150 to NT\$200[7], and twenty three probability values ranging from 0.00 to 0.95. The 06 and 07 lottery sets involved thirteen monetary values ranging from NT\$-420 to NT\$420, and thirteen probability values ranging from 0.05 to 0.90. Since we estimate both the value function and probability function, so the lottery set of 2005 seems a bit imbalance in the sense that there were many more probability values than monetary values. This is the reason why we changed the lottery set in 2006. With this design, we can also compare the estimates from different years to observe if subject behavior change when facing different lotteries.

Each year subjects received a show-up fee of NT\$200. In 2005, 3 gain and 1 loss

[6] SM1 in the Electronic Supplementary Materials contains the complete list of the lotteries.

lotteries were randomly selected *ex post* and played to determined subject monetary payoff; and in 2006 and 2007, 2 gains and 1 loss lotteries. To reduce boredom and fatigue, there were half-time breaks in all years. An average session took about 90~120 minutes and subjects' monetary rewards varied, ranging from NT$120 to NT$800.

4. Results

The estimations were done with the nonlinear least square procedure of Limdep. In this section, we analyze the results of individual and group estimations, for both single-year and panel data.

4.1 Individual single-year estimation

With the four $U(L)$ variants, EUT, T&K, Pr1, and Pr2, considered in this paper, we estimate a total of 928 individual single-year models[8]. Table 1 reports the median coefficients of all models.

[7] NT$100 is about 2.13 euro.

[8] 624 models (26 subjects $\times$ 3 years $\times$ 4 models $\times$ 2 payoff regions, gains and losses) for Group I, and 304 (19 subjects $\times$ 2 years $\times$ 4 models $\times$ 2 payoff regions, gains and losses) for Group II. SM2 of the Supplementary material contains complete list of individual single-year regression coefficients.

Table 1 Median coefficients, individual single-year estimation

	EUT	T&K		Pr1		Pr2		
	α	α	r	α	r	α	r	s
Gains								
IG5	0.919	0.936	0.732	0.936	0.627	0.915	0.918	0.697
IG6	0.998	1.010	0.802	1.006	0.771	0.998	0.978	0.876
IG7	0.987	1.002	0.723	0.997	0.703	0.990	0.803	0.912
IIG6	1.008	1.016	0.733	1.017	0.656	0.992	0.798	0.667
IIG7	1.011	1.027	0.793	1.019	0.625	1.002	0.968	0.767
Losses								
IL5	0.959	0.950	0.611	0.967	0.519	0.975	0.668	1.085
IL6	0.998	1.000	0.877	0.999	0.891	0.994	1.057	1.218
IL7	0.997	0.989	0.776	0.997	0.746	0.989	0.819	1.307
IIL6	0.993	0.989	0.779	0.995	0.724	0.992	0.752	1.035
IIL7	0.987	0.974	0.753	0.985	0.839	0.989	0.725	0.947

The first thing we notice in Table 1 is that the median coefficients of 06 and 07 are somewhat similar, but that of 05 are quite different from the other two years. It appears that the median estimations change when subjects faced different lottery sets. For 06 and 07, in all four models and for both gains and losses, the values of α are all close to 1, so the value functions are almost linear. In EUT, a linear value function means risk neutral and expected value maximization. Many scholars claimed that, for small payoffs, subjects should be approximately risk neutral[9]. The α values in Table 1 are consistent with these literature findings. As for the coefficients of probability weighting functions, literature shows a wide range of values. It would be difficult to make comparison because past researches contain different setups and various results. But in general, it seems that our results are reasonable comparing with that of Scott (2006) and James (2007).

We may also want to know the following: are there systematic differences between

[9] Appendix V in Conlisk (1989) p. 407 contains a proof under EUT. Also see Holt and Laury (2002), p.1644.

Group I and Group II, between gains and losses, between different years for the same groups? The answers to the first two questions seem negative, and we will answer the third question with panel estimation in the next subsection.

After reviewing the median coefficients, we then want to pick one best model for each subject's single-year estimation. To do this , we must first make sure that models have significant parameters. Note that the four models we estimated are nested, so if the r parameters in T&K and Pr1 are not significantly different from zero[10], then these two models would be reduced to EUT; and if s is not significant, then the Pr2 model would be reduced to Pr1. Models that have non-significant parameters or are non-convergent will be rejected first, and the rest are arranged according to their AIC values(Akaike information criterion). The model with the smallest AIC value would be denoted as the best model for the subject. Appendix I reports the model selection results for single-year individual estimations.

With the best model selected for each subject, we could calculate the relative frequency. For example, for data set IG5, 5 (19.23%) subjects had EUT as the best model, and the numbers for T&K, Pr1, and Pr2 models are respectively, 4 (15.38%), 4 (15.38%), and 13 (50.00%). These summary percentages are reported in Table 2.

[10] We use a 5% significance level through out this paper.

Table 2 Model selection for individual single-year estimation, in percent

	(1) EUT	(2) T&K	(3) Pr1	(4) Pr2	(5)=(2)+(3)+(4) Total PT
Gains					
IG5	19.23	15.38	15.38	50.00	80.77
IG6	19.23	11.54	19.23	50.00	80.77
IG7	11.54	30.77	15.38	42.31	88.46
IIG6	21.05	10.53	26.32	42.11	78.95
IIG7	15.79	26.32	10.53	47.37	84.21
Losses					
IL5	19.23	23.08	53.85	3.85	80.77
IL6	19.23	11.54	23.08	46.15	80.77
IL7	15.38	26.92	3.85	53.85	84.62
IIL6	10.53	21.05	26.32	42.11	89.47
IIL7	5.26	15.79	21.05	57.89	94.74

Although EUT is still the best model for a minor part of the subjects, but the dominance of PT over EUT is clear in Table 2. It seems that people are different, some (less than 20%) behave just like EUT claims, treating probability linearly with no subjectively adjustments. But most subjects do have some form of nonlinear weighting adjustments for the probabilities, and hence, PT is the proper theory for them. We also find that within PT, the best functional form for probability weighting is Pr2, this is similar to previously mentioned literature results

IL5 is the only data set where Pr2 does not have a clear superiority. But recall the peculiar structure of lottery set IL5, consisting of only 24 loss lotteries involving two negative payoff values and twenty three probability values. It also shows in Table 1 that, for all four models, the median coefficients of IL5 are different from others. Therefore, we would rather think of IL5 as an exception.

It is also interesting to compare IG5 with other gain lottery sets. The differences in lottery structure also leads to differences in the median parameters, but the best model proportions of IG5 are quite similar to the other gain data sets. Further studies are needed before we can make general conclusions on the relationship between lottery sets and estimation results. .

An important research question of this paper is to compare, for each individual subject, the behavior patterns of two years and test for structural changes. In the next subsection, we report the result of panel estimation and structural change tests.

4.2 Individual panel estimation

In the above analysis, we choose one best model among the four candidates in order to justify the nonlinearity of probability weights. But when taking a closer look at the estimation results, we find that for some subjects, the values of the AIC statistics of the four models are quite close. Therefore, to be cautious, we first take a comprehensive approach and run panel estimations for all four models. Table 3 shows the single-year and panel estimations for Subject I26, Model Pr2, data sets IG5, IG6, IG7. We use this case to explain our research strategy.

Table3 Estimated coefficients of Subject I26, Model Pr2[11]

	α	r	s	AIC	LR
IG5	0.952*	0.545*	0.599*	9.202	
IG6	0.988*	0.230*	0.628*	10.563	
IG7	0.975*	0.550*	0.717*	10.494	
IG5&6	0.968*	0.381*	0.560*	10.102	24.862*
IG6&7	0.980*	0.388*	0.662*	10.577	11.192*

[11] The * sign means significant at 5% confidence level. We use the 5% level throughout this paper

After estimating single-year models, we pool data sets IG5 and IG6 (IG6 and IG7) together to run panel estimation IG5&6 (IG6&7)[12]. For the hull hypothesis of no structural change, H_0: $\alpha_t = \alpha_{t+1}$, $r_t = r_{t+1}$ and $s_t = s_{t+1}$, the last column of Table 3 reports the likelihood ratio test statistics[13]. The *'s in this column tell us that, for both panel estimations, the null hypothesis of no structural change are rejected at 5% significance level. Therefore we conclude that for Subject I26, Model Pr2, there are structural changes between 05 and 06, and also between 06 and 07.

A prerequisite for the above procedure is that both single-year estimations must have significant parameters. For example, if the *s* parameter of IG5 in Table 3 is not significant, then model Pr2 would be rejected and reduced to Pr1. And hence, it would be meaningless to test for structural change of model Pr2 for this panel.

Table 4 summarizes the structural change test results for individual estimations. Let us again use Group I, Model Pr2, Panel IG5&6 as an example to explain. We first exclude those subjects for whom the Pr2 models are not significant in both years. Of the 26 subjects of Group I, 4 (15.38%) subjects belong to this category, labeled as "Not 2Y-Sig." in Table 4. Within the remaining subjects who had significant Pr2 models for both years, the null hypothesis of no structural change could not be rejected for only 1 (3.85%) subject. And we find that there are indeed structural changes for a great majority of subjects (21 subjects, 80.77%).

[12] SM3 of the supplementary material contains complete list of individual panel regression coefficients.
[13] See Judge *et al* (1980) for the likelihood ratio test of structural change for nonlinear statistical models.

Table 4 Proportions of structural change test results, individual panel estimation, in percent

	EUT			T&K			Pr1			Pr2		
	Not 2Y-Sig.	No Structural Change	Structural Change	Not 2Y-Sig.	No Structural Change	Structural Change	Not 2Y-Sig.	No Structural Change	Structural Change	Not 2Y-Sig.	No Structural Change	Structural Change
Gains												
IG5&6	0.00	19.23	80.77	3.85	15.38	80.77	7.69	7.69	84.62	15.38	3.85	80.77
IG6&7	0.00	50.00	50.00	0.00	46.15	53.85	15.38	26.92	57.69	7.69	19.23	73.08
IIG6&7	0.00	68.42	31.58	0.00	52.63	47.37	21.05	36.84	42.11	42.11	21.05	36.84
Losses												
IL5&6	0.00	53.85	46.15	0.00	30.77	69.23	3.85	42.31	53.85	69.23	19.23	11.54
IL6&7	0.00	30.77	69.23	0.00	23.08	76.92	3.85	30.77	65.38	7.69	23.08	69.23
IIL6&7	0.00	36.84	63.16	0.00	26.32	73.68	10.53	21.05	68.42	15.79	10.53	73.68

Table 5 Model selection consistency, in percent

	(1) 2Y-EUT	(2) 2Y-T&K	(3) 2Y-Pr1	(4) 2Y-Pr2	(5) 2Y-PT-diff. $w(p_i)$	(6)=(2)+(3)+(4)+(5) Total 2Y-PT	(7)=1−((1)+(6)) Change
Gains							
IG5&6	7.69	3.85	3.85	34.62	26.92	69.23	23.08
IG6&7	7.69	3.85	3.85	23.08	46.15	76.92	15.38
IIG6&7	10.53	10.53	0.00	21.05	42.11	73.68	15.79
Losses							
IL5&6	0.00	7.69	15.38	0.00	38.46	61.54	38.46
IL6&7	7.69	7.69	0.00	23.08	42.31	73.08	19.23
IIL6&7	0.00	5.26	5.26	26.32	47.37	84.21	15.79

The null hypothesis tests for the equality of all parameters, and hence, it becomes more demanding for models with more parameters. The EUT models estimate only one parameter, and accordingly, Table 4 shows that the no structural change proportions are the highest for EUT. The Pr2 models estimate three parameters, and the proportions of no structural change are much lower. T&K and Pr1 both estimate two parameters, but Table 4 shows that T&K models have higher no structural change proportions for gains. Also, excluding panel IG5&6, the T&K models are two-year significant for all subjects.

Recall that the same lottery sets was used for 06 and 07, and they were quite different from that of 05. We tend to think, and it also seems reasonable, that the proportions of structural change should be higher for the 5&6 panels with drastically different lottery sets. However, for some IL5&6 models, the no structural change proportions are quite high. This seems strange. It could be just a coincidence, but for now, we could only say that we need to do more study before making general conclusion.

Although test results vary with different models, but the general conclusion seems to be that for most subjects, there are structural changes, some of the parameters are different between two years. But still, the null hypothesis can not be rejected for some subjects. There exists a minor proportion of very stable subjects, we can estimate their behavior between two years and find no significant differences in the parameters.

Requiring the equality of all parameters may be too stringent a condition, so we now try a different perspective and observe model selection consistency. Again let us take Subject I26 as an example. Appendix I shows that Pr2 is the best model selected for IG5, IG6 and IG7. Hence, for panels IG5&6 and IG6&7, Subject I26 is denoted as 2Y-Pr2, shown in column (4) of Table 5. For panel IG5&6, there is a total of 9 (34.62%) subjects in the 2Y-Pr2 cell.

Considering the relevancy of PT, besides those 2Y-T&K (1 subject), 2Y-Pr1 (1 subject), and 2Y-Pr2 (9 subject) subjects, there are also 7 (23.08%) subjects whose selected models are PT but with different $w(p_i)$ functional forms. Altogether, almost 70% of the subject could be classified as 2Y-PT (Column 6 of Table 5) for panel IG5&6.

The superiority of PT over EUT is again confirmed in Table 5. For single-year estimation, Table 2 shows that around 20% of the subjects have EUT selected as the best model. But according to Table 5, the proportions of subjects who had EUT as the best models consecutively for two years (Column (1)) are very low. And the proportions of subjects who change between the two theories (Column ((7)), i.e., one year EUT and another year PT, are also quite low[1]. Most subjects, between 70% to 80%, are in Column (6), whose best models contain nonlinear probability weights. And again, Pr2 comes up as the best functional form for $w(p_i)$.

Our first research question concerns the stability of individual behavior, and it can be answered from different perspectives. If we take the strict standard of equality of parameters, then the general conclusion seems to be that for most subjects, there are structural changes; the estimated parameters of two consecutive years are different. But the degree of instability should not be exaggerated since many of those differences are significantly different from zero but very small in values. We also take a different perspective and consider model selection consistency. Table 5 finds that the two year consistencies of PT models are quite strong and it is much weaker for EUT, proportions vary between 0.00% to 10.53%.

[1] less than 20% if we ignore panel period 5&6.

4.3 Group single-year estimation

The data of all subjects in a session are pooled together to make group estimations. The models are similar to Equation (5) but with the following changes. The assumption that the error terms are identically distributed is now released, so the estimated standard errors of the parameters are corrected for heteroscedasticity. Also, we incorporate demographic factors into the value function and estimate[2]:

$$u(x_i) = x_i^{\alpha + g \times \text{Gender} + m \times \text{Major}} \qquad (6)$$

with Gender=1 if female and Major=1 if economics.

Our second research question concerns the consistency of individual and group estimations. This question is answered by examining the following dimensions. Firstly, when comparing group estimation coefficients (Table 6) with the median of individual coefficients (Table 1), we find close correspondence for all four models. This is formal evidence to support the statement that, for the same group of subjects, whether estimated individually or with pooled data, the results are very similar. With this consistency formally established, we are assured that literatures with different estimation strategies are directly comparable.

[2] This is, of course, not the only way to handle demographic factors, but rather a first attempt.

Table 6 Group single-year estimation

	Gains							Losses					
	α	g	m	r	s	AIC		α	g	m	r	s	AIC
EUT							EUT						
IG5	0.900*	0.027*	-0.004			10.124	IL5	0.972*	0.012	-0.031			9.894
IG6	0.996*	-0.003	0.009*			11.025	IL6	0.984*	0.009*	0.005			10.926
IG7	0.987*	-0.002	0.004			11.252	IL7	0.978*	0.004	0.021*			11.273
IIG6	0.998*	-0.011	0.018*			11.418	IIL6	0.997*	0.037*	-0.040			11.522
IIG7	1.007*	0.003	0.002			11.656	IIL7	0.991*	0.033*	-0.040			11.309
T&K							T&K						
IG5	0.913*	0.027*	-0.003	0.669*		10.091	IL5	1.030*	0.015*	-0.034	0.498*		9.698
IG6	0.997*	-0.003	0.009*	0.785*		10.986	IL6	0.982*	0.009*	0.004	0.888*		**10.909**
IG7	0.992*	-0.003	0.003	0.723*		11.206	IL7	0.977*	0.004	0.021*	0.767*		11.216
IIG6	1.006*	-0.012	0.018*	0.665*		11.347	IIL6	0.997*	0.038*	-0.042	0.710*		11.455
IIG7	1.011*	0.002	0.002	0.724*		11.62	IIL7	0.992*	0.031*	-0.040	0.724*		**11.234**
Pr1							Pr1						
IG5	0.913*	0.027*	-0.003	0.636*		10.100	IL5	0.979*	0.016*	-0.034	0.461*		**9.689**
IG6	1.007*	-0.003	0.009*	0.701*		10.971	IL6	0.983*	0.009*	0.004	0.894*		10.916
IG7	1.001*	-0.003	0.003	0.656*		11.199	IL7	0.977*	0.004	0.021*	0.743*		11.215
IIG6	1.014*	-0.011	0.018*	0.594*		11.332	IIL6	0.995*	0.039*	-0.042	0.678*		**11.447**
IIG7	1.022*	0.002	0.002	0.633*		11.602	IIL7	0.990*	0.031*	-0.040	0.700*		11.236
Pr2							Pr2						
IG5	0.887*	0.027*	-0.003	0.860*	0.697*	**10.085**	IL5	0.976*	0.016*	-0.034	0.466*	1.027*	9.693
IG6	0.992*	-0.003	0.009*	0.791*	0.821*	**10.952**	IL6	0.974*	0.009*	0.005	1.008*	1.233*	10.911
IG7	0.990*	-0.003	0.003	0.716*	0.870*	**11.192**	IL7	0.961*	0.004	0.021*	0.897*	1.383*	**11.206**
IIG6	0.990*	-0.012	0.018*	0.726*	0.725*	**11.293**	IIL6	0.988*	0.039*	-0.042	0.725*	1.135*	**11.447**
IIG7	0.999*	0.002	0.001	0.771*	0.729*	**11.572**	IIL7	0.986*	0.031*	-0.040	0.730*	1.079*	11.238

We then follow the same procedure as in individual estimation and choose, among the four candidates, one best model with the smallest AIC values, these are shown in bold figures in the last column of Table 6. For gain lotteries, the best models are Pr2 for all five data sets. Recall that in Table 2, the proportions of subjects who have Pr2 as the best model are also the highest. The conclusions are exactly the same, so for gains lotteries, Pr2 is the best model for individual and group estimations.

But for losses models, the correspondence is not so perfect. Considering data sets IL6 and IIL7, for individual estimations, Table 2 shows that Pr2 is the best model. But according to Table 6, T&K is the best model for the group estimation for these two data sets. Also, for IIL6, Table 6 shows that models Pr1 and Pr2 are equally good, but Table 2 tells us that 42% of the subjects have Pr2 as the best model and only 26% for Pr1. Hence, if we use AIC as the model selection criteria, it is possible to come up with different conclusions from individual estimations and group estimations. Again we think this observation should be considered as a warning only. The divergence may not be that serious since there are only small differences in the AIC values of group estimations.

Also note that for group estimations, EUT is never selected as the best model, but Tables 2 and 5 show that for a minor proportion of individual subjects, EUT is the best model, even two years consecutively. It seems that with individual estimations, the multiple facets of the subject group are captured more completely. Harrison and Ruström (2006) captured this multiple facets character with the mixture model. They found that the proportions the EUT and PT being the latent model are about equal[1]. Our research strategy is different and hence the results may not be directly comparable. But our individual estimations findings definitely show lower proportions for EUT.

As for the two demographic dummies; recall that they enter into the value function in the same way for all four models. Therefore, it is only reasonable that in Table 6, the coefficients of the dummy variables are also almost the same across all four models. It appears that gender is not a significant factor for gains but it is mostly significant for losses, so maybe women and men are different only when facing possible losses, but not for potential gains. Most of the dummies of subject's major are not significant, and even when they are, the values are very small.

4.4 Group panel estimation

Finally, the coefficients of group panel regressions are reported in Table 7. For most group panels, the null hypotheses of no structural change are again rejected, similar to the conclusion of individual estimation. And again, we should be careful in interpreting the group estimation structural change test results. Similar to individual estimations, there are significant changes in some of the parameters, but the differences are often very small in values. The last column also shows that the LR test statistics are much larger for panel 5&6; this is due to the fact that the lottery sets of 05 were very different from the other two years

[1] Or, put it exactly, EUT 55% and PT 45%, see page 21, Table 1 of Harrison and Ruström (2006).

Table 7 Group panel estimation

	α	g	m	r	s	AIC	LR
Gains							
EUT							
IG5&6	0.978*	0.002	0.007*			10.790	953.969*
IG6&7	0.992*	-0.003	0.007*			11.156	37.711*
IIG6&7	1.003*	-0.004	0.010*			11.549	15.605*
T&K							
IG5&6	0.985*	0.002	0.007*	0.683*		**10.731**	870.581*
IG6&7	0.994*	-0.003	0.006*	0.756*		11.114	41.653*
IIG6&7	1.009*	-0.005	0.010*	0.694*		11.498	17.451*
Pr1							
IG5&6	0.990*	0.002	0.007*	0.660*		10.758	994.502*
IG6&7	1.004*	-0.003	0.006*	0.679*		11.103	40.950*
IIG6&7	1.018*	-0.005	0.010*	0.614*		11.481	17.131*
Pr2							
IG5&6	0.968*	0.002	0.008*	0.875*	0.754*	10.742	1003.220*
IG6&7	0.991*	-0.003	0.006*	0.755*	0.845*	**11.091**	42.815*
IIG6&7	0.995*	-0.005	0.010*	0.748*	0.728*	**11.447**	18.473*
Losses							
EUT							
IL5&6	0.984*	0.009*	0.003			10.720	34.623*
IL6&7	0.981*	0.007*	0.013*			11.118	17.050*
IIL6&7	0.994*	0.035*	-0.040*			11.422	7.814
T&K							
IL5&6	0.981*	0.009*	0.003	0.851*		**10.692**	59.666*
IL6&7	0.979*	0.007*	0.013*	0.828*		11.082	31.858*
IIL6&7	0.995*	0.034*	-0.041*	0.718*		11.352	9.595*
Pr1							
IL5&6	0.983*	0.009*	0.003	0.855*		10.702	67.540*
IL6&7	0.980*	0.007*	0.013*	0.816*		11.087	37.097*
IIL6&7	0.993*	0.035*	-0.041*	0.690*		**11.348**	10.187*
Pr2							
IL5&6	0.977*	0.009*	0.003	0.909*	1.116*	10.701	75.966*
IL6&7	0.968*	0.007*	0.013*	0.947*	1.299*	**11.080**	38.783*
IIL6&7	0.987*	0.035*	-0.041*	0.731*	1.112*	**11.348**	10.350

IIL6&7 is the only panel where the no structural change hypothesis can not be rejected for models EUT and Pr2, and rejected only by a close margin for the other two models. But Table 2 provides a very different answer for individual estimations. For all four models, the no structural change hypotheses are rejected for over sixty percents of the subject. This case serves to remind us that it is possible for individual divergence to be smoothed out in group estimation with pooled data. However, there are still traces to be found in individual estimation that could lead to this result. Let us again observe the median coefficients of individual estimations in Table 1. Comparing the relationship of IIL6 and IIL7 with other panel rows, we can see that indeed the differences in median coefficients are smaller between IIL6 and IIL7. Therefore, for panel IIL6&7, the general pictures shown by the median coefficients of individual estimations and group estimation tell us that the models of the two years are very similar. But still, the degree of divergence at individual level is shown more clearly in the individual structural change tests of Table 4.

The influences of the dummy variables become much clearer in the panel estimation. Similar to single-year estimations, gender has no significant influence for gain models, but it does for losses. All the coefficients of gender dummy are positive, so the value functions of females are closer to linear form then male. In Table 7, excluding panel IL5&6, all the coefficients of subjects' major are significant, the signs are all positive for gains model.

5. Conclusion

Collecting two- and three-year panel data of lottery pricing experiment, we estimate four variants of the model on decision-making under uncertainty. Our first research question concerns the stability of individual behavior. Estimating each year's data with exactly the same econometric procedures, we find that, for most subjects, not all the parameters of two consecutive years are equal. Although there exists some stable subjects, but for more subjects, there are structural changes. However, we should be careful in interpreting this result since many of the significant differences are very small in value.

Among the four variants, one best model (with the smallest AIC) is chosen for each individual single-year estimation. Most subjects had PT models two years consecutively, even though the exact functional form of $w(p_i)$ may be different. Therefore, our subjects are quite stable in terms of the nonlinearity of $w(p_i)$.

We also pool the data together to run group estimations and compare these results with individual estimations. It is assuring to find that there is a high degree of correspondence in terms of parameter values. However, in terms of model selection and comparison, it is possible to reach different conclusion from individual and group estimations.

Finally, for the comparison of EUT versus PT, our results clearly show that PT is the better theory for most subjects, and especially for two year consistency. We find that within the not so many EUT subjects, even less were consecutively EUT for two years.This maybe a new blow to EUT.

For future works, we are now considering the following. To begin with, only EUT and PT are estimated in this paper, more competing theories could be considered. Also, with the vast amount of data, we could also try nonparametric estimation.

Even with only four variant models, we observe some intriguing results. For example, if the model selection is done according to AIC criteria, that T&K would not be considered a very good model. However, we also find that almost all of the individual single-year T&K models are significant, and the proportions of no structural change subjects are also very high. We plan to do more study on the issue model selection criteria.

We also observe that with the different lottery sets of 05 and 06, estimated parameters change; while 06 and 07 had the same lottery sets, so the estimated parameters are quite close in values. Does this means that parameters of the decision making model will change with the lottery sets? If this is the case, it partially explains why literature results show such a wide range of parameter values. Ant it also cautious us that we should be ready to accept a new set of parameter values when we change the lottery sets. This amazing idea will be the target of our next project.

Acknowledgement

The first author gratefully appreciates the financial support from T*he National Science Foundation, Taiwan, ROC*, under Grant No. 93-2415-H-031-001.

References

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

Bleichrodt, H. & Pinto, J. L. (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science*, 46, 1485-1496.

Conlisk, J. (1989). Three variants on the Allais example. *The American Economic Review*, 17, 31-57.

Carbone, E. & Hey, J. D. (2000). Which error story is best? *Journal of Risk and Uncertainty*, 20, 161-176.

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

Gonzalez, R. & Wu, G. (1999). On the shape of the probability weighting function, *Cognitive Psychology*, 38, 129-166.

Harrison, E. W., & Ruström, E. E. (2006) Expected utility theory and prospect theory: One wedding and a decent funeral, presented at the FUR XII conference, LUISS, Italy.

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data, *Econometrica*, 62, 1251-1289.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects, *The American Economic*

Review, 92, 1644-1655.

James, D. (2007). Stability of risk preference parameter estimates within the Becker-Degroot-Marschak procedure, *Experimental Economics*, 10,123-141.

Judge. G. G., W.E. Griffiths, W. E., Hill, R. C., Lutkepohl, H., & Tsoung-Chao Lee T-S. (1980). The Theory and Practice of Econometrics, 2nd edition, New York : John Wiley.

Loomes, G., Moffatt, P. G., & Sugden, R. (2002). A microeconometric test of alternative stochastic theories of risky choice, *Journal of Risk and Uncertainty*, 24, 103-130.

Prelec, D. (1998). The probability weighting function, *Econometrica*, 66, 497-527.

Scott, H. S. (2006). Cumulative prospect theory's functional menagerie, *Journal of Risk and Uncertainty*, 32, 101-130.

Tversky, A . & Kaheman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, 5, 297-323.

Appendix I. Model selection, single-year individual estimation

Gains							Losses						
Group I				Group II			Group I				Group II		
Sub.	05	06	07	Sub.	06	07	Sub.	05	06	07	Sub.	06	07
I1	PR2	T&K	PR2	II1	PR2	PR2	I1	PR1	PR1	PR2	II1	PR1	PR2
I2	PR1	PR2	T&K	II2	PR2	T&K	I2	T&K	PR2	PR2	II2	PR2	PR1
I3	PR2	PR2	PR2	II3	PR2	PR2	I3	PR1	PR2	EUT	II3	PR1	PR2
I4	PR1	PR1	EUT	II4	PR2	PR2	I4	EUT	PR2	T&K	II4	PR1	PR2
I5	T&K	T&K	PR2	II5	PR2	T&K	I5	PR1	EUT	EUT	II5	T&K	PR2
I6	PR1	T&K	T&K	II6	PR2	PR2	I6	PR1	EUT	PR2	II6	PR2	PR2
I7	PR2	PR1	PR1	II7	PR1	PR2	I7	T&K	PR1	PR2	II7	T&K	T&K
I8	PR2	PR2	T&K	II8	EUT	PR1	I8	T&K	PR2	PR2	II8	PR2	PR2
I9	PR2	PR2	T&K	II9	EUT	EUT	I9	PR1	PR2	PR2	II9	EUT	T&K
I10	PR1	PR2	T&K	II10	PR1	PR2	I10	PR1	PR2	PR2	II10	PR1	PR2
I11	EUT	PR1	PR2	II11	PR1	PR2	I11	EUT	PR2	T&K	II11	T&K	PR1
I12	T&K	EUT	PR1	II12	PR2	EUT	I12	T&K	T&K	T&K	II12	PR1	PR1
I13	PR2	PR2	T&K	II13	EUT	EUT	I13	PR1	EUT	PR2	II13	T&K	PR1
I14	T&K	EUT	EUT	II14	PR1	PR2	I14	PR1	EUT	EUT	II14	PR2	EUT
I15	T&K	PR1	PR2	II15	PR2	PR1	I15	EUT	PR2	EUT	II15	PR2	PR2
I16	PR2	EUT	T&K	II16	T&K	T&K	I16	T&K	T&K	T&K	II16	PR2	PR2
I17	PR2	PR2	PR2	II17	PR1	PR2	I17	PR1	PR2	PR2	II17	PR2	PR2
I18	PR2	PR1	PR2	II18	EUT	T&K	I18	T&K	PR1	PR2	II18	PR2	T&K
I19	EUT	PR2	PR2	II19	T&K	T&K	I19	EUT	PR2	PR1	II19	EUT	PR2
I20	PR2	PR2	PR2				I20	PR1	PR1	PR2			
I21	EUT	EUT	T&K				I21	PR1	T&K	PR2			
I22	PR2	PR2	PR1				I22	PR1	PR1	T&K			
I23	EUT	EUT	EUT				I23	PR2	EUT	T&K			
I24	EUT	PR2	PR2				I24	EUT	PR2	T&K			
I25	PR2	PR2	PR1				I25	PR1	PR1	PR2			
I26	PR2	PR2	PR2				I26	PR1	PR2	PR2			