To referees:

Many thanks for your comments and suggestions on my paper. I have incorporated with the comments and suggestions in the following revised version. For the convenient reading, all revised parts have been blue-colored in the text, and their page numbers and revised contents are specified below:

1. The works of Luce, Edwards, Galanter, Pliner, Breaut, Parker, Schneider, Siegler, Opfer, Dehaene, Izard, Spelke, and Pica are cited in the paper. They appear on

1) p.4, Galanter1962, 1990 are cited;
2) p.5, Galanter 1962, 1990, Galanter and Pliner 1974, Galanter et al. 1977, Breaut 1983, Parker and Schneider 1988 are cited;
3) p.6, Luce and Edwards 1958 is cited;
4) p.12, The scaling technique used in Galanter's works (1962, 1990) is compared with the double estimate;
5) p.15, Luce, 2010 is cited;
6) p.16, Galanter 1962, 1990, Galanter and Pliner 1974, Galanter et al. 1977, Breaut 1983, Parker and Schneider 1988 are cited as supporting evidences;
7) p.19, the psychophysical numeric scaling studies of Siegler and Opfer 2003, Dehaene, Izard, Spelke, and Pica, 2008 are compared with the utility scaling results in this paper;
8) p.20, Galanter 1962, 1990, Galanter and Pliner 1974, Galanter et al. 1977, Breaut 1983, Parker and Schneider 1988 are cited as supporting evidences.

2. Professor Luce commented that "the logarithmic law should be generally suitable for utility scaling" is an overly large claim. I agree this comment. To correct my inadequate wording in the original paper, a note for the limitation of the logarithmic law for utility measures is added on p.10: The logarithmic law should be suitable for the utility scaling in the case with a judgment pattern similar to bargaining behaviors in an ultimatum game. In Part 1.3.2, this case will be clearly defined as "single estimate".

3. p.10 and p.11, a note of b_i being a constant ratio in Klein-Rubin utility function.

4. p.12 and P.13, a note for the difference of categories between utility measures and psychophysical measures.

5. Some questions addressed in Professor Luce's report are replied here, they do not appear in the text of the revised paper:

1) This paper is an economics study, it only involves psychophysical research paradigm as a methodological tool, and is unnecessary to discuss all details of the contents in psychophysics, especially of those only relevant to physical stimuli but not economic behaviors. It is the ultimatum game experiment, but not the psychophysics during 20th century, to provide the evidence sporting the Weber fraction in economic judgments. In its methodological path for deriving the utility logarithmic law from Bernoulli's hypothesis, the emphasized point is that JND is unnecessary (see the first paragraph just below the title of Part 1.1 on p.6). This is just also the methodological gist presented in Luce and Edwards, 1958. In my original paper, such a methodological path was cited from Masin,

Zudini, and Antonelli, 2009, which contains a direct analysis on Bernoulli's hypothesis, and now in the revised version also from Luce and Edwards, 1958 (see p.6).

2) This paper is not a Fechnerian psychophysical interpretation for utility measures, and involves analyses on Fechner's logarithmic law, Stevens' power law, and the more foundational category of judgment types in utility measures. One of the conclusions in this paper is that the logarithmic law is suitable for the description of commodity choice and the power law for risk choice. The first page of my paper mentions "Kahneman and Tversky's power function for money-risk utility estimation", and from Part 1.3.2 (on p.12) to the last section, either the theoretical analysis or the experimental analysis always concentrates on the difference of single- and double-estimate utilities, which deliver logarithmic and power laws for utility judgments respectively. All those are not Fechnerian. Another conclusion in this paper is that the category framework for single- and double-estimates, which has never been discussed in psychophysics, offers a new category way to discriminate the different utilities for commodity choice and risk choice. This is one of the new results in the paper, neither Fechnerian nor Stevensian but a new framework. The traditional categories in psychophysics are inappropriate for this paper.

3) The English of this revised version had been edited by a professional language editing service.

Psychophysical Interpretation for Utility Measures

Yuqing He

Abstract The paper explores utility measures by combining experiments with mathematical derivations in the psychophysics paradigm. The analysis on the ultimatum game experiment reveals the evidence for the utility threshold and thus supports Bernoulli's utility logarithmic law. Both experimental results and theoretical derivations show that the logarithmic law is suitable for the description of commodity choice and the power law for risk choice. The further mathematical demonstration indicates the logarithmic law for utility scaling to be a Klein-Rubin utility function, a utility function well defined in microeconomics. Based on this, the experimental utility measure is connected with the econometric model Linear Expenditure System, and presents an experimental procedure for testing the utility maximization hypothesis, which will resolve a long unsettled perplexity in a fundamental stone of economics since Gossen proposed it in 1854.

JEL A12, D01
Keywords psychophysics, ultimatum game, utility function, logarithmic law, power law

Department of Public Administration, Jinan University, Guangzhou 510632, Guangdong Province, China; E-mail: hyq5.27 @163. com

Introduction

The paper seeks to achieve four goals: (i) reveal correspondences between sensation threshold in psychophysics and responder's acceptance threshold in the ultimatum game; (ii) experimentally and mathematically verify that utility is an observable psychological magnitude that agrees with the form of sensation scales in psychophysics; (iii) derive the quantitative attributes of commodity choice and risk choice; and (iv) present an experimental procedure to test the utility maximization hypothesis.

The first highlight of this paper is the shape of utility function. It is based on other researchers' earlier studies from three aspects: Bernoulli's derivation for the logarithmic function of utility (Bernoulli, 1738), Kahneman and Tversky's power function for money-risk utility estimation proposed in their experimental analysis on the cumulative prospect theory (Kahneman and Tversky, 1992) and Galanter's power law revealed in his utility measures (Galanter, 1962, 1990), and Ekmen's mathematical inference of deriving the power law from the logarithmic law in psychophysics (Ekmen, 1964).

To deal with the St. Petersburg paradox in a gamble (e.g. see Stigler, 1950), Swiss mathematician D. Bernoulli assumed in 1738 that the increase of one's subjective economic utility dU is directly proportional to the increment of wealth dx and inversely proportional to one's whole wealth x, or $dU=kdx/x$. By integrating on both sides, he derived a logarithmic law $U=k\ln x+C$ for the utility (Bernoulli, 1738). More than one hundred years later, this result agreed with Fechner's logarithmic law for the sensory response of physical stimulus. This was once regarded as an evidence of utility judgment to be a kind of psychophysical response by many researchers including Marshall and Fechner (Stigler, 1950). Evidently, Bernoulli's above derivation foundation is incomplete. To support his hypothesis, an empirical Weber fraction $\Delta x/x$ in utility judgment is required. It can only be resolved by empirical or experimental studies. Bernoulli did not continue and get to finish the empirical analysis for his hypothesis.

Differing from the suggestion of Bernoulli's logarithmic function, Kahneman and Tversky in their experimental analysis for the cumulative prospect theory found that it is appropriate if a power function is used to describe subjects' money-risk utility estimating data (Kahneman and Tversky, 1992). Also, in the framework of Stevens' magnitude estimation approach, Galanter proposed the power law in his utility measurements (e.g, Galanter, 1962, 1990). However, in systematic economics, those experimental results did not reveal a logical path paving to the core zone of utility theory, such as the utility maximization and its applied models in econometrics, the operational definition of utility concept, and etc., therefore attracted not so much attentions in existing economic literature. It is an insufficiently explored domain but had appeared captivating clues long ago.

Responding to the debate on the alternatives of Fechner's logarithmic law and Stevens' power law in psychophysics, Ekmen supposed the logarithmic law commonly applicable in psychophysical judgments, then theoretically introduced the logarithmic descriptions simultaneously into both the numeric judgment and physical stimulus judgment in Stevens' magnitude estimate approach (e.g., Stevens, 1959; Stevens and Guirao, 1963), and finally obtained a mathematical relation in which the power law could be derived from the logarithmic law (Ekmen, 1964). Nevertheless, Ekmen only discussed his formulations theoretically without any experimental test. It left an attractive inspiration to us rather than a final resolution to the debate.

Although the above three analyses are incomplete as conclusive discussions either in theory or in experiment, they at least give us such an impression that there may exist two alternatives for the shape of utility functions, one logarithmic and another power, and following Ekmen's analytical path, the two alternatives may be logically consistent but not conflicting.

This paper will remedy the defects among the above studies and integrate them into a logically consistent description, namely, reveal an experimental basis for Bernoulli's hypothesis, clear theoretical elements for Kahneman and Tversky's power function in money-risk utility estimation, and introduce experimental decomposition to test Ekmen's formulations.

Combining the psychophysics approach with the evidence from the ultimatum game experiment, Section 1 will reveal an experimental basis for Bernoulli's hypothesis, and will discuss the commodity choice and risk choice in a new framework of utility scaling. In this analysis, the ultimatum game experiment is given a new sense of utility threshold measurement, a typical psychophysics interpretation. In this section, a mathematical demonstration will identify the logarithmic law with the Klein-Rubin utility function (Klein and Rubin, 1947) in a new experimental scheme, and finish the key preparation for setting up an experimental procedure to test the utility maximization hypothesis.

In Section 2, by discriminating two kinds of estimates, single estimate and double estimate, in experimental utility measures, commodity choice and risk choice will be logically related in a utility scaling framework, and an experimental test to Ekmen's formulations will be naturally introduced into the discussions.

In the other hand, there are two traditional paths to treat the issue of utility measures. One is the individual utility measure, and another the average utility measure. The former is usually followed by researchers who probe the features of utility concept itself, often seen in most experimental studies (e.g., Fisher, 1892; Mosteller and Nogee, 1951; MacCrimmon and Toda, 1969; Kahneman and Tversky, 1979; Starmer, 2004), and the latter is usually followed by those who use econometric models to describe realistic markets or psychophysical methods to scaling subjects' responses, often seen in applied studies of market empirical data (e.g., Klein and Rubin, 1947; Stone, 1954; Liuch, 1973; Houthakker, 1960; Theil, 1965; Deaton and Muellbauer, 1980) or psychophysicists' experimental studies (e.g., Galanter, 1962, 1990; Galanter and Pliner, 1974; Galanter et al., 1977; Breaut, 1983; Parker and Schneider, 1988). The two paths separately stretch in three relatively independent fields of experimental economic measurements, econometric studies, and psychophysical researches, and cause isolation between them.

There were some debates on the average utility measures. The pioneer of experimental utility measurement, psychologist Thurstone firstly used average utility measure to determine the experimental indifference curve (Thurstone, 1931), but was criticized by other researchers (e.g., Wallis and Friedman, 1942). The criticism seems to continue today. Pooling all individuals' data and then fitting the model to the pooled data may be regarded as misrepresenting the true utility functions for individuals. For example, one can approximate a power function by combining exponential functions of different rates, furthermore, even if the form of the utility function is the same for every subject, the subjects may differ in parameters. Such a worry may be reasonable if one uses an average result to describe individual judgment. However, if the average result is limited to apply to describing group behaviors, for example, market empirical data, the above worry will be unnecessary. Even if the individuals' utilities follow exponential functions of different rates but the best asymptote for their average result is power function, we should use the power function but not the exponential function to depict the overall effect of all individuals in a market or an experiment, in which both exponential and power functions may be true but respectively for either individual or group. In this case, we should also forbid using an individual utility function to describe group behavior. Whether or not a utility report has economics sense depends on whether it presents a regular and stable result, as a regular and stable result will contribute a meaningful description in economics.

This paper is an experimental study that follows the path of average utility measure, agreeing with the psychophysics and econometrics paradigms. Namely, it will determine parameters in a utility function by using averages of subjects' performances in experiments. Another highlight in this paper is Klein-Rubin utility function (Klein and Rubin, 1947) based on average measures, which has been demonstrated in Section 1 as a psychophysical function for the utility and is given a behavioral economics sense. This will provide a realizable framework to estimate both the econometric model Linear Expenditure System (Stone, 1954) and the utility function in psychophysical paradigm simultaneously in a set of experimental data, and further, to form a new way to experimentally test the utility maximization hypothesis, which is so fundamental and important but has never been tested in experiments since Gossen proposed it in 1854.

Section 3 summarizes and discusses the findings in this paper, including an experimental procedure for testing the utility maximization hypothesis.

1 Experimental analysis to the utility threshold in ultimatum game

1.1 Weber fraction in the ultimatum game

Bernoulli's discussion in 1738 means Fechner's assumption of just noticeable difference (JND) as sensation unit (Fechner, 1860) unnecessary, and implies another mathematical path of deriving the logarithmic law directly from empirical Weber fraction (Masin, Zudini, and Antonelli, 2009). From a deeper analysis, Luce and Edwards also concluded that the logarithmic law can be derived only from Weber fraction without JND (Luce and Edwards, 1958). That is, as long as observations of utility scaling behavior support a Weber fraction, using Bernoulli's mathematical derivation, the logarithmic law will be possibly derived as an empirical law.

In psychophysics, Weber's law states that in the moderate stimulus intensity, the amount of change needed for the difference threshold ΔI is a constant fraction K of the initially presented physical stimulus intensity I, that is, $\Delta I / I = K$. Following Weber's discovery, Fechner related a least sensation variation ΔS to K, and obtained a relationship between the subjective sensation scaling S and the objective stimulus intensity I. Fechner's result $S = c\ln I + C$ can also be derived similarly by the derivations used by Bernoulli (1738), where c and C are two constants.

In a typical ultimatum game experiment (e.g., see Güth, 1995), there are two players, one plays proposer and the other responder, to distribute a "stake" (usually certain amount of money) between them. Both sides know the money stake q. At first, the proposer controls q. And then, the proposer proposes to the responder an ultimatum division of the stake, in which the proposer offers to the responder a money amount q_o cutting out from q, and retains $q - q_o$ to himself. Then the responder determines to accept or reject it. If the responder accepts, he gains q_o, and the proposer gains $q - q_o$. If the responder rejects, both sides obtain nothing. Usually for the proportion q_o/q there is a proportional rejection threshold q_{or}/q. Once the proportional offer is below such a threshold, the responder would rather probably reject the proposer's proposal. After investigating the results from sixteen research groups, Camerer found that half of the proposers' offers were rejected by responders below q_o/q=0.2 (Camerer, 2003). This proportional rejection threshold is similar to what the Weber fraction reveals in psychophysics.

As far as their corresponding relations, components of the subjective measure in the ultimatum game are completely identical to those leading us to derive Fechner's logarithmic law from sensation response of physical stimulus. Those components include: (i) standard stimulus, compared stimulus, absolute difference threshold, and they give the evidence to construct $\Delta S = c\Delta I/I$; and (ii) a constant Weber fraction, which supports an integral on both sides of $\Delta S = c\Delta I/I$ and finally derives the logarithmic law by Bernoulli's mathematical path.

In one of the difference threshold measurements for physical stimulus, a standard stimulus I is given, then a compared stimulus I_c is presented. The intensity of I is kept constant, while the intensity of I_c is adjusted. The difference between I_c and I is gradually diminished and approaches zero. And the subject is asked to compare the difference between I_c and I. In the adjusting course of I_c, there statistically is a critical intensity level $I_{cc} \neq I$, at which the subject reports the difference beginning to vanish between I_c and I; $\Delta I = |I - I_{cc}|$ is the absolute difference threshold, and Weber fraction $\Delta I/I$ is determined too (e.g. see Luce and Krumhansl, 1988).

In an ultimatum game, denote the proposer's offer q_{or} as the responder's rejection threshold, below which the rate of responders' rejections will evidently increase, while, above which will evidently decrease. In the ultimatum game experiment, a responder always faces to three "stimuli": (i) money stake q; (ii) offer q_o; and (iii) the part a proposer retains for himself $q - q_o$. Here, in q_o and $q - q_o$ only one is independent. The responder's utility estimates in an ultimatum game are interpreted as follows：
A responder's utility estimate to the offer is a comprehensive result involving various judgments, including the money amount, unfairness, and so on. In the responder's utility estimate, if $q_o < q_{or}$, the proposer's retained part $q - q_o$ will be regarded unfair, the responder's loss from unfairness will exceed his gain from q_o, the utility is null, and the offer will be rejected; if $q_o > q_{or}$, the proposer's retained part $q - q_o$ will be regarded acceptable, the responder's gain from q_o will exceed his loss from unfairness, the utility is positive, and the offer will be accepted. Comparing with the measurement to difference threshold of physical stimulus, q is a standard stimulus, $q - q_o$ a compared stimulus, and q_{or} the absolute difference threshold. If $q_o < q_{or}$, the responder feels no utility difference between q and $q - q_o$ (for the responder, it is equivalent to proposer occupying all money stake). It corresponds to the difference vanishing between I_c and I in the psychophysical measurement. Reversely, if $q_o > q_{or}$, then the responder feels the difference. It corresponds to a difference between I_c and I being detected in the psychophysical measurement. Here, a one-to-one correspondence exists between the sensation threshold in psychophysics and in the responder's acceptance threshold in an ultimatum game. The above explanation is called "utility scaling interpretation" for the ultimatum game experiment.

In fact, Fechner's logarithmic law in psychophysics depends on the primary relationship

$$\Delta S = c\frac{\Delta I}{I}.$$

Let S indicate a responder's subjective utility estimate in the ultimatum game experiment, then a similar relationship

$$\Delta S = c\frac{q_{or}}{q}$$

holds in the utility scaling interpretation, where q_{or} identifies with ΔI. Naturally, the ratio rejection threshold q_{or}/q corresponds to Weber fraction $\Delta I/I$.

The above analogue has revealed the structural correspondence between q_{or}/q and Weber fraction. The further problem is whether or not such a ratio q_{or}/q is stable in a specific ultimatum game. In 1994, Cameron had worked out an experimental study on the simple one-period ultimatum game with raising stakes (Cameron, 1999), from which we are able to derive a test to the stability of rejection threshold at different monetary levels.

1.2 Examination to Bernoulli's hypothesis

Cameron conducted experiments in Indonesia using real Indonesia money, the Rupiah, as stakes at three levels of 5,000, 40,000, and 200,000, about 0.075, 0.6, and 3.0 times the average monthly expenditure of a participant. This is a maximal span between low and high stakes that had been used in ultimatum game experiments. Three groups of subjects respectively participated in 5,000-, 40,000-, and 200,000-experiments. Cameron tested their homogeneity by simple t-tests and no significant difference was detected among them. Their results can be compared with each other. Only taking into account unrepeated games, there are three 5,000-experiments, one 40,000-experiment, and one 200,000-experiment provided in Cameron's study. As the statistical test in Cameron's paper showed no significant difference among three 5,000-experiments, their data will be pooled in this paper. We will concentrate on the responders' switching points from acceptance to rejection in the ultimatum games with different stakes. It is sufficient for us to discuss the results of q_o/q below 50%.

Table 1. Data of rejection and offer in Cameron's experiment

n = 101 pairs in three 5,000-experiments; n = 35 pairs in 40,000-experiment; n = 37 pairs in 200,000-experiment

Stake		Proposer's proportional offer interval (p_o/p) 0-0.09	0.1-0.19	0.2-0.29	0.3-0.39	0.4-0.49	0.5-0.59
5,000	Rejection rate	100%	50%	55%	20%	19%	3%
	Number of proposals	4	2	11	5	31	40
40,000	Rejection rate			25%	29%	0%	0%
	Number of proposals			4	7	3	20
200,000	Rejection rate	100%	100%	50%	0%	0%	0%
	Number of proposals	1	1	2	3	11	16

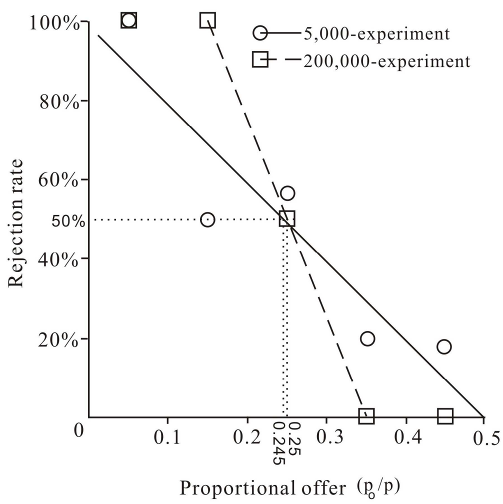

Fig. 1. Proportional rejection thresholds of 5,000-experiment and 200,000-experimrnt. There is little difference between them.

In Cameron's 5,000- and 200,000-experiments, proposers offered q_o/q from 0% through 50%

stakes; the switching points will be shown in these experiments directly. In 40,000-experiment, however, no offers were given by proposers below 20% stake, and no switching point can be observed in this case. Hence, only the results of 5,000- and 200,000-experiments can be compared. Cameron's experimental data are rearranged in Table 1, which is estimated (after discarding invalid data) from "FIGUREs 1, 2, and 3" in Cameron's paper (1999). In Table 1, the percentage of responder's rejections to proposer's proposals is described in six proportional offer intervals "0~0.09", "0.1~0.19", "0.2~0.29", "0.3~0.39", "0.4~0.49", and "0.5~0.59"; "Number of proposals" indicates the number of proposers who propose to responders in a specific proportional offer interval; and "Rejection rate" indicates the percentage of proposer's offers rejected by responders in a specific proportional offer interval. For instance, in the bottom-right in Table 1, "16" indicates sixteen proposers making proportional offers within the interval 0.5~0.59, and above "16", "0%" indicates none of the sixteen offers being rejected by responders. Results of 5,000- and 200,000-experiments in Table 1 are plotted in Fig. 1. The rejection threshold is defined to be the value of proportional offer that is rejected exactly 50 percent of the time, similar to that in a Yes-No detection procedure in psychophysics (e.g. Gescheider, 1976). In Fig. 1, the solid line describes 5,000-experiment, and the dashed line 200,000-experiment. They are the results of linear regression on data between proportional offers of 0.5 and 1.0. Fig. 1 shows that the rejection threshold is 0.245 for 5,000-experiment and 0.25 for 200,000-experiment. They are so close that they can be viewed as identical. That is, Cameron's experiment supports Bernoulli's hypothesis, and the logarithmic law could exist in the utility scaling.

Although the above discussion involves only a small data set in its key point and cannot be regarded as the ultimate evidence, doubtlessly, it first delivers a direct positive proof for Bernoulli's hypothesis since 1738. At least, with this proof, Bernoulli's utility logarithmic law represents not only purely a theoretical imagination but also an attractive possible reality. The further conclusion, of course, will be determined by a combination with experimental tests in Section 3.

List and Cherry (2000), Slonim and Roth (1998), and Roth et al. (1991) also compared frequencies of responders' rejection behavior for low and high stakes in ultimatum games. List and Cherry used stakes $20 and $400, Slonim and Roth used Slovak Crown as stakes at three levels Sk60, Sk300, and Sk1500, and Roth et al. used $10 and $30 in their first-round experiment for Pittsburgh subjects. Their first-round tests can be viewed as a simple one-period ultimatum game, equivalent to Cameron's experiment. Also investigating affections from low and high stakes in simple one-period ultimatum games, Straube and Murnighan (1995) used stakes $5 and $100, and Hoffman et al. (1996) $10 and $100. Their outcomes show only little different effects on subjects' rejection behavior for low and high stakes in the simple one-period ultimatum game. The stability of ratio rejection threshold in the simple one-period ultimatum game is a general phenomenon.

In addition, in repeated game studies, List and Cherry (2000) and Slonim and Roth (1998) found significant difference between low and high stakes in the rejection threshold of experienced participants (e.g. in the tenth round). The learning effect accounts for the behavioral change from the simple one-period ultimatum game to repeated ultimatum game: an experienced subject understands how to strategically take more benefit or diminish loss from his opponent (Roth, et al., 1991). It therefore does not deny Bernoulli's hypothesis. Repeated games describe strategic manipulations rather than intuitive performances. To reveal intuitive features, studying subjects' performance in the simple one-period cases is more appropriate.

1.3 Discussions on the logarithmic law of utility

To overcome the failure of Weber's law in very low stimulus intensity, a revised form of Weber's fraction $\Delta I/(I+a)=K$ was proposed in psychophysics, where a is a constant parameter (see Gescheider, 1997). Accordingly, the revised Fechner's law should be as follows:

$$S=c\ln(I+a)+C. \qquad (1)$$

To tally with the conventions in economics, (1) is rewritten as

$$S=c\ln(q+a)+C, \qquad (2)$$

where q is the economic quantity, and c, a, and C are constants determined by experiment. Evidently, compared with the logarithmic law in psychophysics, the subjective scaling S in (2) has a meaning of intuition scaling of economic quantity q. It is similar to the sensation scaling of physical stimulus. As there is no particular modality for utility judgment, the word "intuition" is used here to replace "sensation". However, whether it is sensation or intuition, stimulus–response is their common attribute.

The sensation measurement in psychophysics and the utility scaling in the ultimatum game are identified in their components and quantitative stabilities, and present a similar stimulus-response pattern that reveals an experimental foundation for Bernoulli's hypothesis. The logarithmic law should be suitable for the utility scaling in the case with a judgment pattern similar to bargaining behaviors in an ultimatum game. In Part 1.3.2, this case will be clearly defined as "single estimate".

1.3.1 Utility implication of the logarithmic law

If a logarithmic law exists in the above mentioned utility scaling, does it identify with a utility function well defined in microeconomics?

In econometrics, Linear Expenditure System (LES) (Stone, 1954) and Extended Linear Expenditure System (ELES) (Liuch, 1973) are two successful econometric models in empirical studies. They were mathematically derived by using the Klein-Rubin utility function in the consumer theory (e.g., see Liuch, 1973). The Klein-Rubin utility function is

$$U=\sum b_i \ln(q_i-r_i)\,,$$

where, b_i is a ratio constant, which determines the distribution of consumption proportions for a consumer's consumed commodities, and r_i is a constant too. Counting q_i and r_i by the monetary value, $b_i=\dfrac{q_i-r_i}{\sum(q_i-r_i)}$, $0<b_i\le 1$, $\sum b_i=1$, q_i is the consuming quantity of commodity i, and r_i the consumer's necessary consuming quantity of commodity i, a constant.

Measuring the commodity quantity by its monetary value, any commodity quantity q_i, i=1,2,3,…n, can always be added up to a sum Q, or $Q=\sum q_i$. It can be demonstrated mathematically that the logarithmic law for Q represents a Klein-Rubin utility function. Namely,

$$c\ln(Q+a)+C=\sum b_i \ln(q_i-r_i)\,.$$

It is demonstrated as follows:

Taking differentials on the both sides of $Q=\sum q_i$,

$$dQ=\sum dq_i\,,$$

$$d(Q+a)=\sum d(q_i-r_i)$$

where a is a constant and $0\le r_i\le q_i$ is a constant for q_i. Identically transform the equation

$$\frac{Q+a}{Q+a}d(Q+a)=\sum\frac{q_i-r_i}{q_i-r_i}d(q_i-r_i)\,,$$

$$(Q+a)d\ln(Q+a)=\sum(q_i-r_i)d\ln(q_i-r_i)\,. \tag{3}$$

Set $\frac{\sum(q_i-r_i)}{Q+a}=\lambda$, then $\sum(q_i-r_i)=\lambda(Q+a)$. Denote $b_i=\frac{q_i-r_i}{\sum(q_i-r_i)}=\frac{q_i-r_i}{\lambda(Q+a)}$, which naturally satisfies $0<b_i\leq 1$ and $\sum b_i=1$. (3) is rewritten

$$d\ln(Q+a)=\lambda\sum\frac{q_i-r_i}{\lambda(Q+a)}d\ln(q_i-r_i)=\lambda\sum b_i d\ln(q_i-r_i)\,. \tag{4}$$

When LES and ELES are derived from Lagrange's method, b_i is assumed as a constant ratio, independent from q_i (e.g. see Intriligator, 1980), thus, independent from $q_i–r_i$, and further from $\ln(q_i–r_i)$. Integrating on both sides in (4) under the condition suitable for the derivations of LES and ELES, finally get

$$\frac{1}{\lambda}\ln(Q+a)+C=\sum b_i\ln(q_i-r_i)\,,$$

$$c\ln(Q+a)+C=\sum b_i\ln(q_i-r_i)\,, \tag{5}$$

where $c=1/\lambda$. Given b_i is a constant ratio as mentioned above, λ is thereby a constant, and further, c a constant too.

In (5), the left-hand side is the logarithmic law for Q, and the right-hand side is a Klein-Rubin utility function. $c\ln(Q+a)+C$, the logarithmic law for $Q=\sum q_i$, is a Klein-Rubin utility function. Furthermore, measuring an economic quantity by its monetary value, an economic quantity can always theoretically be decomposed into n parts q_i, i = 1,2,3,$\cdots$,n. We are always able to construct $Q=\sum q_i$ for any economic quantity theoretically, and thus derive (5) for them. This means that the intuitive scaling value for an economic quantity, if it holds as a logarithmic law, always represents a Klein-Rubin utility function. Experiments indeed approve the utility scaling value to hold as the logarithmic law (see Section 2, later). It is therefore concluded that such a logarithmic law obtained from psychophysical paradigm is of a utility description in economics. The Klein-Rubin utility function implies the relationship between subjective and objective commodity quantities. In other words, the essence of utility maximization in LES and ELES is the maximization of subjective commodity quantities. Revealing the relationship between subjective and objective quantities is just the basic task of psychophysics. Psychophysical paradigm will play an important role in economic analyses.

In the above demonstration, no special restriction is made to the type of q_i. The demonstration holds for all economic quantities provided they are additive. Therefore, it is generally suitable for money amount, commodity quantity, risk degree, and so on.

We now interpret the Klein-Rubin utility function clearly from the viewpoints of behavioral economics and the utility scaling interpretation: in this utility function, $(q_i–r_i)$ indicates the measures for economic quantities q_i relative to the reference point r_i (in LES and ELES, r_i is interpreted as the base consumption share to maintain one's life), and $\sum b_i\ln(q_i-r_i)$ is a linear combination of utility scaling values (logarithmic laws) for economic quantity q_i compared to r_i.

A subjective scaling value of economic quantity is generally a utility value, and evidently, it is the cardinal utility. An ultimatum game experiment is essentially to measure the intuition threshold of utility.

There are two parallel processes of a responder's intuitions in an ultimatum game experiment. One is the objective observation of the number that tells the responder the objective amounts of money, and another, the subjective utility scaling that tells the responder the subjective utility obtained from proposer's offers. When q_o is below q_{or}, the former is still bigger than null but the latter diminishes to nothing. The utility scaling interpretation suits the latter. The responder's rejection behavior in the ultimatum game is a behavioral presentation of the utility scaling. Two independent processes are not mutually exclusive. The noticeable difference between q and $q-q_o$ in objective observation of the number does not affect the vanishing of difference between q and $q-q_o$ in utility scaling below the absolute difference threshold q_{or}. As the two processes are always mixed together, the utility implication of a responder's intuition scaling of money amounts in an ultimatum game might be ignored for too long a time.

1.3.2 Single estimate and double estimate

The above demonstration relies on $Q = \sum q_i$. This relation as a perceptible object means that subjects clearly know the sizes of Q and q_i and only the utility is required to estimate. It will be called "single estimate" below. Therefore, the demonstration contains that the logarithmic law as a utility function applies to the single estimate. Commodity choice is usually a single estimate. And typically, the bargaining judgment in an ultimatum game is a single estimate too.

If the meaningful sizes of Q and q_i are not informed to subjects, both quantity and utility need to be estimated. It will be called "double estimate". In the cases involving risk decision, one only possibly performs through the double estimate. For example, a kind of Allais' paradox used in Kahneman and Tversky's experiment is presented as the following (Kahneman and Tversky, 1979):

PROBLEM 1: Choose between

A: 4000 with probability 0.80, or B: 3000 with certainty.

PROBLEM 2: Choose between

C: 4000 with probability 0.20, or D: 3000 with probability 0.25.

Similar choices were also used in Kahneman and Tversky's experiment conducted for the cumulative prospect theory (Kahneman and Tversky, 1992). It is a money-risk estimate. Money-risk estimate is always a double estimate for the following reasons: 1) Clearly telling subjects the probability of a risk is unlikely to change their intuitive assessment into the precise quantitative calculation for an issue such as assessing worth of a risk, that is, it is still objectively un-meaningful for estimating a quantity; and 2) there usually is no objective standard of market price for a risk in a trifling item such as a risk of losing a chance to get 3000. In this case, one has to estimate subjective utilities of two factors simultaneously.

The utility scaling technique used by Galanter is also a double estimate. In an experiment, Galanter (1962) asked student subjects to report a money amount matching to a double happiness that a gift of \$10, \$100, or \$1,000 would bring them, and derived a power function from his experimental data. Evidently, asking subjects to judge "a double happiness" could not change their intuitive assessment into the precise quantitative calculation, and there certainly was no objective standard of market price for the students' happiness of a gift, and thus, Galanter's measure is a typical double estimate too. Other double estimates used by Galanter include scaling the pleasure (displeasure) for "you are given a brand-new bicycle" ("your brand-new bicycle is stolen"), "how much money would you have to lose to make you exactly twice as upset as losing \$5?"and so on (Galanter, 1990).

Kahneman and Tversky proposed the power law for money-risk estimate (Kahneman and Tversky, 1992), and Galanter also proposed the power law for the above mentioned monetary and nonmonetary utility measures (Galanter, 1962, 1990). In the light of the above analysis, the double estimate should follow the power law. The next section will answer this question from both theoretical and experimental analyses.

In psychophysics, the logarithmic law and power law are derived respectively from the partition estimate and ratio estimate. In utility measures here, the category is different. They are single estimate and double estimate. And evidently, the category of single and double estimates is discriminated by constructable measurement factors, and contains a potential experimental decomposition for directly testing Ekmen relations (Ekmen, 1964). By the way, the categories of single and double estimates are also applicable to the psychophysical scaling on physical stimuli (see p.19 of Supplemental Files).

2 Experimental tests of single- and double-estimate utilities

After a pilot study (see Part 1 in Supplemental Files), 192 graduate student volunteers were recruited from Jinan University as subjects participating in experiments. Five simple experiments and one Latin square experiment (see Bailey, 1996) were carried out among those subjects. The Latin square experiment is usually used to reduce order effects revealed in simple experimental measures so that high-quality regression curves could be obtained.

Five simple experiments (with 80 subjects) tested the differences between subjects with and without economics-study backgrounds. They found no significant difference between the two types of subjects in their experimental performances, namely, subjects' behaviors observed in valid data have no reliance on the specialized knowledge in economics. These results and other interested outcomes of the five simple experiments are presented in Parts 2~4 of Supplemental Files. To save the pages, the coming text mainly discusses the Latin square experiment, briefly L-experiment, in detail. Detailed Latin square designs are presented in Part 5 of Supplemental Files.

2.1 Methods

2.1.1 Participants

The L-experiment requires 60 subjects who deliver complete valid data in all measures. The subjects were divided into three groups, labeled Groups I~III, and each group contains five subgroups, labeled Subgroups i~v. Some subjects were bored by experimental measures and their measures had to be stopped midway and were therefore incomplete. There were totally 112 subjects (ages 22~25, majored in management or economics) randomly selected to enter the L-experiment, and 52 subjects failed to deliver complete valid data. In the final experimental results, each of Groups I~III contained 20 valid subjects. Every subgroup had 4 subjects (male 2, female 2). Experimental tests to single- and double-estimate utilities were performed by curve fits to the averages of the 60 valid subjects' data. To attract subjects, all participants were provided a free banquet or paid.

2.1.2 Measures

The test for every group consisted of three measurements, which investigated subjects' utility estimates respectively under three laboratory conditions of "quantity-price double estimate", "quantity single estimate", and "price single estimate".

Under the condition of quantity-price double estimate, labeled Meas. 1, the subjects experienced a sequence of consumptions, in which consuming quantities were unknown to them, and the prices were without existing recognized standards in practical markets. Subjects were asked to report money amounts they were willing to pay by estimating the consuming quantities through subjective assessments. This condition makes the utility scaling similar to one of Stevens' magnitude estimates in psychophysics, called cross-modality matching, a double estimate between the sensation scalings, for example, of handgrip and other physical stimuli such as loudness, brightness, visual length, vibration, electric shock, and so on (e.g., Stevens, 1959; Stevens and Guirao, 1963).

Under the condition of quantity single estimate, labeled Meas. 2, money amounts were clearly assigned to subjects for their consumption, and they were asked to estimate consuming quantities they think "should be" to match the money amounts.

Under the condition of price single estimate, labeled Meas. 3, subjects were clearly told about quantities consumed and asked to estimate the prices they were willing to pay.

Meas. 2 and 3 are decompositions from Meas. 1.

Electrical-power massage was used as consumption in the experiment. It was conducted through a portable electrical-power massage machine, which was handy for the experimenter to control. The duration of the electrical-power massage is the consuming quantity and the money a subject is willing to pay for such duration is the price chosen. Such an exact length of time is easily treated to be unknown to the subject. And, very importantly, till date in China, it is quite seldom to see commercial services of electrical-power massage, and the stimulus of electrical-power massage has no recognized standard market price. These satisfy the conditions required by the measurement of double-estimate utility. In the experiments, the subjects were not required to pay their estimates. It was a hypothetical test.

Taking Subgroup i of Group I as an example, Meas. 1, 2, and 3 were specified as follows:

Meas. 1: A subject was given five stimuli of the massage on the waist of durations 8, 16, 24, 32, and 40 seconds, sequentially; the subject was not informed about durations of the stimuli. After one stimulus was completely presented, the subject was immediately asked to report orally a money mount he was willing to pay for the stimulus just experienced, and so on for next stimulus. When all five stimuli were given and five money amounts were reported, Meas. 1 concluded. With the same subject, the experiment proceeded to Meas. 2.

Meas. 2: The experimenter posed the following question to the subject: If you are asked to pay ¥1.0, what is duration of the massage you will demand? The experimenter began the massage on the subject's waist. The subject estimated the time for the money amount ¥1.0 and said an "OK" immediately when he felt that an appropriate length of time had passed. The experimenter recorded the time with a stopwatch. And so on for ¥2.0, ¥3.0, ¥4.0, and ¥5.0 sequentially. When all five durations were reported, Meas. 2 concluded. With the same subject, the experiment proceeded to Meas. 3.

Meas. 3: The experimenter posed the following question to the subject: If you are given an electrical-power massage for 8 seconds, how much will you be willing to pay? The experimenter began the massage on the subject's waist for 8 seconds to acquaint him with such a time. After the subject reported on an answer sheet and the experimenter collected the answer sheet, the experimenter posed the second question to the subject: If you are given an electrical-power massage for 16 seconds, how much will you be willing to pay? And the subject reported on another answer sheet. If the subject asked, 16 seconds of the massage could be experienced by him (in fact, no subject did so in the experiments). And so on until 24, 32, and 40 seconds were presented to the subject. Finally, the subject reported five money amounts respectively for 8, 16, 24, 32, and 40 seconds on five separate answer sheets. Meas. 1, 2, and 3 all concluded.

Massage durations and money amounts assigned in L-experiment are shown in Table 2.

Table 2. Stimulus levels of Meas. 1~3

Meas. 1	8"	16"	24"	32"	40"
Meas. 2	¥1.0	¥2.0	¥3.0	¥4.0	¥5.0
Meas. 3	8"	16"	24"	32"	40"

There are two kinds of order effects taken into account in Latin square designs. The first is the order effect between Meas. 1~3, and the second between five stimuli assigned in each of Meas. 1~3.

Meas. 1~3 were carried on sequentially one by one in an independent room and spent about 10~15 minutes in total for a subject's complete program. During every measurement, only one subject was allowed into the room.

2.2 Ekman relations between single and double estimates

Meas. 2 and 3 are single estimate. They follow the logarithmic law $S = c\ln(q+a) + C$. Thus, for Meas. 2,

$$S_2 = c_2\ln(m + a_2) + C_2; \quad (6)$$

and for Meas. 3,

$$S_3 = c_3\ln(q + a_3) + C_3; \quad (7)$$

where, S_2 denotes the utility scaling value of money measured by consuming quantity (the massage time), m the money amount assigned in Meas. 2; S_3 denotes the utility scaling value of consuming quantity (the massage time) measured by money amount, q the duration of the massage assigned in Meas. 3. In the coordinate with horizontal scale of massage duration and vertical scale of money amount, (6) presents a convex curve for Meas. 2 and (7) a concave curve for Meas. 3. They will be tested by the experiments (see Part 2.3.1). These expected concave and convex changes are resulted from replacements between informed factors and estimated factors.

It has been extensively accepted that the shape of a utility curve can be changed by different judgment factors. In the prospect theory, based on one's facing "gain" or "loss", utility curves are distinguished by the concave or convex function (Kahneman and Tversky, 1979). In Luce's study, based on people's personality differences, theoretically, utility curves have three possible function forms: proportional, exponential, and negative exponential (Luce, 2010). Here, based on the mutual replacements between judgment factors, there are also two curve forms of concave and convex for utility representations.

Meas. 1 is a double estimate. Subjects have to estimate the utility of money in terms of quantity consumed and, at the same time, estimate the utility of consuming quantity in terms of money, all through their intuitions. In this estimation a subject makes effort to find a subjective scaling value S_2 to match with another subjective scaling value S_3, that is,

$$S_2 = bS_3. \quad (8)$$

Where b, named "preference coefficient", is a proportion reflecting subjects' preference to money or consuming in Meas. 1; here b increases as the money preference increases.

Substitute (6) and (7) in (8)

$$c_2\ln(m + a_2) + C_2 = bc_3\ln(q + a_3) + bC_3,$$

$$\ln(m+a_2)=\frac{bc_3}{c_2}\ln(q+a_3)+\frac{bC_3-C_2}{c_2}=\ln(q+a_3)^{\frac{bc_3}{c_2}}+\ln\exp\left(\frac{bC_3-C_2}{c_2}\right)$$

$$=\ln\left[(q+a_3)^{\frac{bc_3}{c_2}}\exp\left(\frac{bC_3-C_2}{c_2}\right)\right]=\ln[\beta(q+a_3)^{\alpha}],$$

$$m=\beta(q+a_3)^{\alpha}-a_2\,. \tag{9}$$

In (9), α and β follow

$$\alpha=\frac{bc_3}{c_2}\,; \tag{10}$$

$$\beta=\exp\left(\frac{bC_3-C_2}{c_2}\right). \tag{11}$$

Equation (9) shows that the relation between money amount (m) and consuming quantity (q) follows Stevens' power law in the double estimate of Meas. 1. The above inferences from (6) and (7) to (9) are similar to those of Ekman deriving Stevens' power law from Fechner's logarithmic law (Ekman, 1964). Below, (10) and (11) are together called Ekman relations.

The power law revealed in double-estimate utilities has long been discovered by many experimental studies (e.g., Kahneman and Tversky, 1992; Galanter, 1962, 1990; Galanter and Pliner, 1974; Galanter et al., 1977; Breaut, 1983; Parker and Schneider, 1988). In other words, as a result (9) has been supported by experimental observations. The following reported experiments will concentrate on two aspects: 1) compare the distinction between single- and double-estimate utilities that respectively correspond to the logarithmic law and power law, and are respectively followed by the commodity choice and risk choice; 2) test Ekman relations in the experiments as an analysis on the constructable relations implied in the logarithmic law and power law of utilities, and as assistant evidences to examine the power law.

2.3 Results

Original data and their treatments are presented in Part 5 of Supplemental Files. Before the original data were used in the analysis, a structural normalization had been made on them to equalize the structural contributions from every subject's data. This structural normalization completely reserves the overall quantitative levels and characters of all subjects' data (see Part 3 of Supplemental Files).

2.3.1 Patterns of utility scales

Fitting the power law (9) in means of structural-normalized data of Meas. 1 and the logarithmic laws (6) and (7) in means of structural-normalized data of Meas. 2 and 3, experimental values of a_2, c_2, C_2, a_3, c_3, C_3, α, and β are acquired as shown in Table 3.

Table 3. Observed parameters in Latin square experiment

a_2	c_2	C_2	a_3	c_3	C_3	α	β
10	74.6	-169.8	20	10.4	-32.2	0.596	1.88

Using the results of Table 3, the power law and logarithmic laws for Meas. 1~3 in the L-experiment are derived as the following

m=1.88(q+20)$^{0.596}$−10;

S_2=74.6ln(m+10) −169.8;

S_3=10.4ln(q+20) −32.2.

Fig. 2 plots the above curve estimations of Meas. 1~3. The data of Meas. 1 well approve the power law with R^2=0.99, and the data of Meas. 2 and 3 well approve the logarithmic laws with R^2=1.0 for Meas. 2 and R^2=0.99 for Meas. 3, respectively. Just as the above theoretical expectations, the logarithmic law for Meas. 2 is convex, whereas, the logarithmic law for Meas. 3 is concave. Similar results are also observed in five simple experiments (see Part 4 in Supplemental Files). The L-experiment and five simple experiments present similar utility curve groups, revealing the robustness of experimental results.

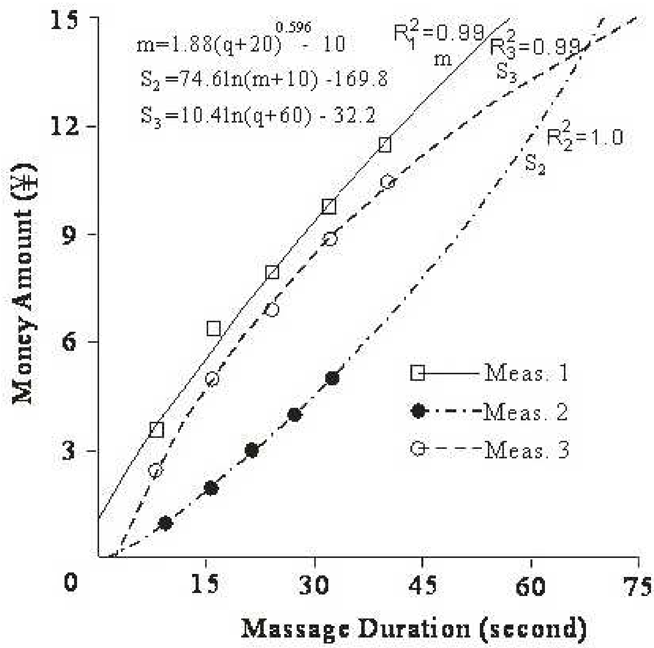

Fig. 2. Curve regressions for Meas. 1~3 in the Latin square experiment. The data of Meas. 1 well approve the power law with R^2=0.99, and the data of Meas. 2 and 3 well approve the logarithmic laws with R^2=1.0 for Meas. 2 and R^2=0.99 for Meas. 3, respectively. Though they are with identical stimulus intensities and presentation orders, the double estimate Meas. 1 and the single estimate Meas. 3 deliver two significantly different utility curves. And just as theoretical expectations, the logarithmic law for Meas. 2 is convex, whereas, the logarithmic law for Meas. 3 concave.

2.3.2 Relations between single- and double-estimate utilities

In the L-experiment, massage durations and presentation orders of Meas. 1 and 3 are the same, and the only distinction between Meas. 1 and 3 is that the former is a double estimate but the latter a single estimate. To examine the distinction between single and double estimates, and taking into account that fitted data for each measurement is a small sample (n=5), Wilcoxon test is selected as an examination tool. In the L-experiment, the Wilcoxon test for Meas. 1 and 3 presents Z=2.023, p=0.043. Their difference is significant. For five simple experiments, labeled Exps. I~V, Wilcoxon tests are also performed for Meas. 1 and 3 in every experiment. Table 4 collects all those testing results.

Table 4. Wilcoxon tests on Meas. 1 and 3 in L-experiment and Exps. I~V (n=5 in each)

L-exp.	Exp. I	Exp. II	Exp. III	Exp. IV	Exp. V
Z=2.023	Z=2.041	Z=1.095	Z=2.023	Z=2.023	Z=2.023
p=0.043	p=0.041	p=***0.273***	p=0.043	p=0.043	p=0.043

Table 4 shows the difference between Meas. 1 and 3 significant ($p<0.05$) in five of six experiments, and only Exp. II ($p=0.273$) is insignificant. That is, single and double estimates are generally different, they deliver two distinguished subjective measures.

Substituting the experimental values of c_2, C_2, c_3, and C_3 and taking b=4.0 in (10) and (11), obtain the theoretical values of α and β in the L-experiment

$$\alpha = \frac{bc_3}{c_2} = \frac{4.0\times 10.4}{74.6} \approx 0.558\,;$$

$$\beta = \exp\left(\frac{bC_3 - C_2}{c_2}\right) = \exp\left(\frac{-4.0\times 32.2 + 169.8}{74.6}\right) \approx e^{0.550} \approx 1.73\,.$$

Similarly, theoretical values of α and β in five simple experiments are also obtained. Table 5 presents all those results. The theoretical values of α and β agree with their experimental values with an average relative error [(Experimental−Theoretical) /Experimental] of 6.5% for L-experiment and five simple experiments. Ekman relations (10) and (11) hold at this error level. They reveal relationships between single- and double-estimate utility scales.

Table 5. Experimental and theoretical values of α and β

	b	α		β	
		Theory	Experiment	Theory	Experiment
L-exp.	4.0	0.558	0.596	1.73	1.88
Exp. I	2.6	0.641	0.704	0.72	0.77
Exp. II	2.5	0.730	0.786	0.50	0.51
Exp. III	3.1	0.903	0.950	0.25	0.26
Exp. IV	2.4	0.716	0.785	0.52	0.54
Exp. V	2.7	0.602	0.651	0.84	0.97

2.3.3 Constant errors

Table 5 shows a systematic deviation between theoretical and experimental values of α and β, where the theoretical values are identically less than the experimental ones. If the experimental values multiply a revision coefficient 0.92, their agreements would be greatly improved as shown in Table 6.

Table 6. Revised experimental values of α and β

	α		β	
	Theoretical	Revised experimental	Theoretical	Revised experimental
L-exp.	0.558	0.596×0.92≈0.548	1.73	1.88×0.92≈1.73
Exp. I	0.641	0.704×0.92≈0.648	0.72	0.77×0.92≈0.71

Exp. II	0.730	0.786×0.92≈0.723	0.50	0.51×0.92≈0.47
Exp. III	0.903	0.950×0.92≈0.874	0.25	0.26×0.92≈0.24
Exp. IV	0.716	0.785×0.92≈0.722	0.52	0.54×0.92≈0.50
Exp. IV	0.602	0.651×0.92≈0.599	0.84	0.97×0.92≈0.89

The deviation may be resulted from constant errors occurring in the measurements. For example, in the estimates of massage durations in Meas. 2, whatever subjects' immediate oral reports by saying "OK" or experimenters' manual records by pressing a stop watch tend to enlarge the massage durations but never to shorten them. It might evidently be one of the causes for constant errors. To prevent those errors, using real payoff and non-temporal test may be helpful.

Furthermore, different stimulus and response levels in Meas. 1~3 may cause errors in the estimations of α and β in each experiment. From "Fig. 4-1" in Part 4 of Supplemental Files, the fitting curves of Meas. 1 and 3 in Exps. I~III shows systematic variations as the stimulus levels wholly decrease from 15″~75″ to 8″~40″ for Exp. I to III: their up-right ends shift to the left from Exp. I to III. It means the differences between quantitative levels to yield constant shifts in subjects' estimations. The experimental estimations of α and β mainly rely on the data observed in Meas. 1, but the theoretical estimations of α and β on the relationships of fitting curves of Meas. 2 and 3. To ensure the basis identical to the experimental and theoretical estimations of α and β in an experiment, the best estimating condition in an experiment requests the data of Meas. 1~3 coming from identical quantitative levels. However, the assigned money amount levels in Meas. 2 are always lower than the response money amount levels in Meas. 1 and 3 in all Exps. I~V (see "Fig. 4-1" in Supplemental Files), and also in the L-experiment. When the data of Meas. 1~3 in an experiment are used respectively to estimate experimental and theoretical values of α and β, a constant error may be caused by these differences coming from the experimental setting distinctions.

Table 6 indicates that the revision coefficient for those constant errors is about 0.92.

2.3.4 Comparisons between utility and psychophysical numeric scalings

Some subjects delivered proportional data in their reports of Meas. 3. Proportions of proportional data in Meas. 3 are 35%, 42%, and 35% for Exps. I, II, and III, but merely 6% respectively for Exps. IV and V (see Part 4 of Supplemental Files), and almost disappears in L-experiment. Meas. 3 proportionally orders the stimuli to present to subjects in Exps. I, II, and III, but crossly disturbs the stimuli to present to subjects in Exps. IV and V and L-experiment. The proportional data in Meas. 3 almost disappears in the measures with the crossly disturbed stimulus presentations in Exps. IV and V and L-experiment. Obviously, the proportional data is resulted from the subjects' proportionally increasing expectation in the stimuli, a typical order effect.

In psychophysical studies of the close-ended numerical scaling, there are two response patterns, the logarithmic pattern for American immature subjects (e.g. second-grade children) and the linear pattern for American mature subjects (e.g. sixth-grade children and adults) (Siegler & Opfer, 2003). To the Mundurucu subjects coming from Amazonian indigene group with a reduced numerical lexicon and little or no formal education, an experimental study reveals at all ages subjects presenting a logarithmic scale in number-space mappings (Dehaene, et al., 2008). Combining the results from Siegler and Opfer with the results from Dehaene, et al., it has been concluded that the logarithmic pattern is followed by initial intuition scaling and the linear pattern is a cultural invention in formal education (Dehaene, et al., 2008). But the linear pattern in utility scaling is only resulted from the order effect. Though the utility scaling seems a kind of the numeric scaling, they are different substantially.

3. Conclusions and discussions

3.1 Syntheses of evidences

Utility scaling can be divided into two types of single and double estimates by their different scaling processes. Double-estimate Meas. 1 and single-estimate Meas. 3 in an experiment are completely the same as consumptions, but present significantly different utility scaling curves. The category of single and double estimates is necessary in the utility scaling.

The evidence from the ultimatum game experiment justifies an empirical basis of Bernoulli's hypothesis for deriving a utility logarithmic law. The mathematical demonstration confirms that the logarithmic law, if it exists in economic single estimate, has a sense of the Klein-Rubin utility function. In other words, it does agree with a usual utility concept in economics. Finally, experimental outcomes of single estimates Meas. 2 and 3 under distinguished circumstances approve the logarithmic law. To meet with the evidences from the above three aspects, that is, Bernoulli's hypothesis and its supporting evidence from the ultimatum game, the mathematical demonstration for utility implication of the logarithmic law, and the experimental observation of utility scaling, in a logical framework, the best conclusion is that single-estimate utility judgment follows the logarithmic law. In fact, a number of important utility functions broadly used in empirical studies which describe commodity choice behaviors, the typical single estimates, consist of the logarithmic components (e.g., Klein and Rubin, 1947; Stone, 1954; Liuch, 1973; Houthakker, 1960; Theil, 1965; Deaton and Muellbauer, 1980). The logarithmic law of single estimate had broad empirical evidences long ago.

Money-risk estimate is usually a double estimate. The subjects' estimate described in Kahneman and Tversky's experiment conducted for the cumulative prospect theory is a typical example of money-risk estimate. They obtained the experimental curves for those money-risk estimates and found that it is appropriate if a power function is used to describe them (Kahneman and Tversky, 1992). It supports the power law of double estimate in the utility scaling. On the other hands, in the framework of Stevens' magnitude estimation approach, a number of earlier studies also supported the power law for the double-estimate utilities (e.g., Galanter, 1962, 1990; Galanter and Pliner, 1974; Galanter et al., 1977; Breaut, 1983; Parker and Schneider, 1988). There are broad experimental evidences to stand up for the power law in the double estimate. In addition, experimental tests to Ekman relations deliver a positive result supporting the power law in Meas. 1, which derived from logarithmic laws in Meas. 2 and 3. Finally, utility-scaling experimental outcomes approve the power law in double-estimate Meas. 1.

Synthesizing the evidences from the three aspects, that is, experiments of cumulative prospect theory and Galanter et al.'s earlier experimental studies, tests of Ekman relations, and outcomes from Meas. 1, in a logical framework, the best conclusion is that double-estimate utility judgment follows the power law. Furthermore, experimental Ekman relations mean that in utility scaling, the logarithmic law is a fundamental law and the power law is only a corollary derived from such a fundamental law.

The above conclusions mean that at least at the experimental levels of money amounts and consumption quantities, commodity choice follows the logarithmic law, whereas, risk choice follows the power law. They were derived from synthesized evidences rather than from a single compelling one, that is, not from a single analysis on the ultimatum game, the experiments for cumulative prospect theory and in the framework of psychophysics, the mathematical demonstration, or the experimental curve regression in this paper, but from all of them.

3.2 Psychological attributes of utility judgment

Traditional concept of utility is vaguely addressed as an intuitive interpretation, for example, “the satisfaction derived from consuming commodities” (Henderson and Quant, 1971). However, such a utility concept has never been seriously applied in theoretical or empirical studies in economics. To give this concept an operational sense, a market investigation for consumers’ satisfaction seems feasible. Nonetheless, the existing recognized utility functions used in empirical studies, such as the Klein-Rubin utility function, are not obtained from satisfaction investigations. In fact, useful utility functions in empirical studies are proposed for commodity quantity measures but not for satisfaction measures. The traditional utility concept is an oral decorate artifact rather than a meaningful definition, without any direct contribution to current economic descriptions such as utility maximization, empirical analysis on consumers’ commodity choices, and so on. It may mix economics with marketing. This concept should be clarified from a more detailed analysis on its psychological attributes and economics requirement.

As commodity consumptions, subjects’ judgments in Meas. 1, 2, and 3 have no difference, and the satisfactions derived from them should be the same. However, they deliver significantly different utility descriptions. The results of Meas. 1 and Meas. 3 are two different utility curves subject to different functions; Meas. 2 follows a convex utility curve, but Meas. 3 a concave one. They are distinguished from each other only by their different estimate patterns. Thus, the judgment of utility relies on the psychological presentation patterns rather than on pure satisfaction derived from consuming commodities. Utility is a typical psychological magnitude that is differently yielded from various perceptual processes. Perceptual pattern is one of important determinants in it.

The utility curves of Meas. 2 and 3 were changed from convex to concave in the experiments. It is certainly not caused by the variation of subjects’ preference but by the variation of stimulus presentation patterns. In Meas. 2, the money amounts were informed to subjects who were asked to estimate utilities of those money amounts in terms of massage durations. In reverse, in Meas. 3, the massage durations were informed to subjects who were asked to estimate utilities of those massage durations in terms of money amounts. Such a change of perceptual process caused the variations from convex to concave. In Meas. 1 and 3, the two different perceptual processes, double estimate and single estimate, also caused two different types of utility curves. All these changes are attributed to the variations of perceptual patterns.

The utility function, either in current theoretical or empirical applications, e.g., in the theoretical analysis on utility maximization or econometric applications in empirical data, is always used to describe how consumers distribute their quantity choices among varieties of commodities but not to indicate the degrees of consumers’ satisfactions. The psychophysical interpretation for utility measure seeks to reveal the relationship between subjective and objective commodity quantities. It agrees with economics requirements of the utility analysis. The utility theory should emphasize more on the analysis of perceptual quantity.

In the light of discussions in Part 1.3.1, the Klein-Rubin utility function implies the relationship between subjective and objective commodity quantities. In other words, the essence of utility maximization, for example in LES and ELES, is the maximization of subjective commodity quantities. Therefore, “utility” should be operationally defined as “the subjective quantity of commodity or evaluation”—a measurable object and an applicable theoretical concept. It will enable us to relate the utility measure to the utility maximization, and further, to test the utility maximization hypothesis in a framework of theoretical cardinal utility maximization and empirical econometric model. It is one of evident differences of this paper from other experimental studies on utility measurements. That is, the experimental analyses finished in this paper can logically incorporate with the analytical system of utility theory. Other experimental utility studies have never exhibited such a character.

3.3 An experimental procedure for testing utility maximization

The cardinal utility maximization has been proposed for more than one hundred and fifty years (Gossen, 1854). Such an important hypothesis had been regarded for a long time as un-testable but replaceable by the ordinal utility maximization (e.g. Hicks, 1939), which seems measurable at a first glance. However, any experimental analysis on ordinal utility maximization must rely on the observable indifference curve. As seen in early experimental studies, no ideal indifference curves with precise shape had ever been obtained. In fact, those experimental studies fell into predicaments in which no ideal theoretical properties were found in objective observations unless the properties were introduced as a prior knowledge in the data arrangements (Thurstone, 1931; Roussears and Hart, 1951) or as a prior reminder to instill into subjects by a training session before they participated in the experiment (MacCrimmon and Toda, 1969). It has been found impossible to test the utility maximization experimentally within the ordinal utility framework. Ordinal utility may be another idle dream for classical economics. In this paper, cardinal utility had been revealed testable. Combining with the experimental outcome in the single estimate, the interpretation of the Klein-Rubin utility function for the logarithmic law implies a possibility of experimentally examining the cardinal utility maximization hypothesis.

The Klein-Rubin utility function $U = \sum b_i \ln(q_i - r_i)$, $b_i = \dfrac{p_i(q_i - r_i)}{\sum p_i(q_i - r_i)}$, is measurable provided b_i, q_i, and r_i can be determined in an experiment. Traditional application of LES uses empirical data of q_i to determine b_i and r_i. It can be transplanted to the experimental approach. The change in this transplantation is only that empirical q_i is replaced by experimental q_i. Nevertheless, to avoid the logic circulation, LES is impossible to independently test the Klein-Rubin utility maximization hypothesis because LES is just derived from such a utility maximization. However, now there is another way to determine the Klein-Rubin utility function experimentally using the utility scaling approach: assigning a fixed budget constraint and prices p_1, p_2, …, p_n in an n-commodity bundle for all subjects in an experiment, we can determine quantities q_1, q_2, …, q_n by observing subjects how distribute their choices in such an n-commodity bundle, derive constants r_1, r_2, …, r_n by measuring the logarithmic laws of the utilities of separate commodities (similar to Meas. 3), and combining data of q_i and r_i further estimate b_1, b_2, … b_n. In this way, we are able to construct the experimental Klein-Rubin utility function using the linear combination of the logarithmic laws experimentally in an n-commodity bundle. Meanwhile, as what has been previously mentioned, solving LES for b_i and r_i in data of the experimental q_i, we also obtain a Klein-Rubin utility function implied in LES for this n-commodity bundle. Denoting U_{Est} as the Klein-Rubin utility function experimentally estimated from the linear combination of logarithmic laws, and U_{LES} the Klein-Rubin utility function implied in LES derived from experimental data of q_i, the utility maximization will be experimentally tested by comparing U_{Est} with U_{LES} in a commodity bundle.

LES is the result of maximizing the Klein-Rubin utility function, thus, U_{LES} can be viewed as the utility function theoretically required by utility maximization for this commodity bundle. U_{Est} is the intuitive utility estimate for such a commodity bundle, directly indicating an outcome of behavioral observations. If the comparison reveals an agreement between U_{LES} and U_{Est}, the theoretic utility maximization implied in LES agrees with the behavioral utility evaluation in subjects' intuitive judgment and, therefore, the experiment supports the utility maximization hypothesis; otherwise, not.

This will be an important and inspiring progress that resolves a long unsettled perplexity in a fundamental stone of economics.

References

Bailey R. A. (1996). Orthogonal partitions for designed experiments, *Designs, Codes and Cryptography*, *8*, 45–77.

Bernoulli, D. (1738). Exposition of a new theory on the measurement of risk. Sommer, L. translated, *Econometrica*, *22*, (1954), 23–36.

Breaut, K. D.(1983). Psychological measurement and the validity of the modern economic approach: A presentation of methods and preliminary experiments. *Social Science Research*, *12*, 187–203.

Camerer, C. F. (2003). Behavioral Game Theory: Experiments in Strategic Interaction. Princeton: Princeton University Press. 43–59.

Cameron, L. A. (1999). Raising the stakes in the ultimatum game: Experimental evidence from Indonesia. *Economic Inquiry*, *37*, 47–59.

Deaton, A. and Muellbauer, J. (1980). An almost ideal demand system. *American Economic Review*, *70*, 312–326.

Dahaene, S., (1997). The number sense: How the mind creates mathematics. New York: Oxford University Press.

Dehaene, S., Izard, V., Spelke, E., & Pica, P., (2008). "Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures". *Science*, *320*, 1217–20.

Ekman, G. (1964). Is power law a special case of Fechner's law? *Perceptual Motor Skills*, *19*, 730.

Fechner, G. T. (1860). Elemente der Psychophysik. Reprinted (1999), Bristol: Thoemmes Press.

Fisher, I. (1892). Mathematical Investigations in the Theory of Value and Prices. New York: Augustus M. Kelley Publishers (1965).

Galanter, E. (1962). The direct measurement of utility and subjective probability. *American Journal of Psychology*, *75*, 208–220.

Galanter, E. (1990). Utility functions for nonmonetary events. *The American Journal of Psychology*, *103*, 449–470.

Galanter, E. and Pliner, P. (1974). Cross-modality matching of money against other continua. In: Moskowitz, H., Sharf, B., and Stevens, J. C. (Eds), Sensation and Measurement: Papers in honor of S. Stevens. Netherland: Reidel. 65–76.

Galanter, E., Popper, R., and Perera, T. (1977). Annoyance scales for simulated VTOL and CTOL overflights. *Journal of the Acoustical Society of America*, *8A* (*Suppl.*), 1.

Gescheider, G. A. (1976). Psychophysics: Methods and Theory. Hillsdale: Lawrence Erlbaum Associate. 61.

Gescheider, G. A. (1997). Psychophysics: The Fundamentals (3rd ed). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Gossen, H. H. (1854). The Laws of Human Relations and the Rules of Human Action Derived Therefrom. Blitz, R. C. translated, Cambridge: MIT Press (1983).

Güth, W. (1995). On ultimatum bargaining experiment—A personal review. *Journal of Economic Behavior and Organization, 27*, 329–344.

Henderson, J. M. and Quant, R. E., (1971), Microeconomic Theory: A Mathematical Approach. New York: McGraw-Hill. 6-8.

Hoffman, E., MacCabe, K. A., & Smith, V. L., (1996). On expectation and the monetary stakes in ultimatum game. *International Journal of Game Theory*, *25*, 289–301.

Houthakker, H. S. (1960). Additive Preferences. *Econometrica*, *28*, 244–257.

Hicks, J. R. (1939). Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory. Oxford: Clarendon Press.

Intriligator, M. D. (1980). Econometric Models, Techniques, and Applications. New Delhi: Prentice-Hall. 226–230.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kahneman, D. and Tversky, A. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–324.

Klein, L. R. and Rubin, H. (1947). A Constant-Utility Index of Cost of Living. *Review of Economic Studies*, *15***,** 78-83.

List, J. A. & Cherry, T. L., (2000). Learning to accept in ultimatum games: Evidence from an experimental design that generates low offers. *Experimental Economics*, *3*, 11–29.

Liuch, C. (1973). The expended linear expenditure system. *European Economic Review*, *4*, 21–32.

Luce, R. D. (2010). Interpersonal comparisons of utility for 2 or 3 types of people. *Theory and Decission*, *68*, 5–24.

Luce, R. D. and Krumhansl, C. L. (1988). Measurement, Scaling, and Psychophysics. In: Atkinson, R. C., Herrstein, R. J., Lindgey, G., & Luce, R. D. (Eds.), Stevens' Handbook of Experimental Psychology. New York: John Wiley & Sons, Inc. 3–73.

Luce, R. D. and Edwards, W. (1958). The derivation of subjective scales from just noticeable differences. *Psychological Review*, *65*, 222–237.

MacCrimmon, K. R. and Toda, M. (1969). The experimental determination of indifference curves. *Review of Economic Studies*, *36*, 433–451.

Masin S. C, Zudini V., and Antonelli M. (2009) Early alternative derivations of Fechner's law. *Journal of the History of Behavioral Sciences*, 2009, *45*, 56-65.

Mosteller, F. and Nogee, P. (1951). An experimental measurement of utility. *The Journal of Political Economy*, *59*, (5), 371- 404.

Parker, S. and Schneider, B. (1988). Conjoint scaling of the utility of money using paired comparisons. *Social Science Research*, *17*, 277–286.

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S., (1991). Bargaining and market behavior in Jerusalem, Ljubjana, Pittsbourgh, and Tokyo: An experimental study. *The American Economic Review*, *81*, 1068–1095.

Roussears, S. W. and Hart, A. G. (1951). Experimental verification of a composite indifference map. *Journal of Political Economy*, *59*, 288–318.

Siegler, R. S. & Opfer, J. E., (2003). "The development of numerical estimation: Evidence for multiple representations of numerical quantity". *Psychological Science*, *14*, (3), 237–243.

Starmer, C. (2004). Developments in nonexpected-utility theory: The hunt for a descriptive theory of choice under risk. Camerer, C. F., Loewenstein,G., & Rabin, M. (Eds), Advances in Behavioral

Economics. Princeton: Princeton University Press. 104–148.

Slonim, R. and Roth, A. E., (1998). Learning in high stakes ultimatum games: An experiment in the Slovak Republic. *Economica*, *66*, 569–596.

Stevens, S. S. (1959). Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *Journal of Experimental Psychology*, *57***,** 201–209.

Stevens, S. S. & Guirao, M. (1963). Subjective scaling of length and area and the matching of length to loudness and brightness. *Journal of Experimental Psychology*, *66***,** 177–186.

Stigler, G. J. (1950). The development of utility theory, II. *The Journal of Political Economy*, *58*, 373–396.

Stone, R. (1954). Linear expenditure system: An application to the pattern of British demand. *Economics Journal*, *64*, 511–527.

Straube, P. G. & Murnighan, J. K., (1995). An experimental investigation of ultimatum games: Information, fairness, expectations and lowest acceptable offers. *Journal of Economic Behavior and Organization*, 27, 345–364.

Theil, H. (1965). The information approach to demand analysis. *Econometrica*, *33*, 67–87.

Thurstone, L. L. (1931). The indifference function. *Journal of Social Psychology*, *2*, 139–167.

Wallis, W. A., and Friedman, M., (1942). The empirical derivation of indifference functions. In: Lange, O., McIntire, F., and Yntema, T., (ed), Studies in Mathematical Economics and Econometrics: In Memory of Henry Schultz. (University of Chicago Press, Chicago), 177-183.