UNIVERSITY OF NOTTINGHAM

Discussion Papers in Economics

Discussion Paper
No. 00/20

EXPECTED UTILITY AND THE ENDOWMENT EFFECT: SOME EXPERIMENTAL RESULTS

By Gwendolyn C. Morrison

October 2000

DP 00/20
ISSN 1360-2438

UNIVERSITY OF NOTTINGHAM

Discussion Papers in Economics

Discussion Paper
No. 00/20

EXPECTED UTILITY AND THE ENDOWMENT EFFECT: SOME EXPERIMENTAL RESULTS

by Gwendolyn C. Morrison

Gwendolyn Morrison is Lecturer, School of Economics, University of Nottingham

October 2000

Expected Utility and the Endowment Effect: Some Experimental Results

Gwendolyn C. Morrison[*]
University of Nottingham

Abstract:
Internal inconsistencies are so commonplace in studies using the contingent valuation method (CVM) that it has been argued that that method should be abandoned as a means of preference elicitation in favour of other methods such as standard gambles (SG). The experiment described in this paper finds that an anomaly that is well documented in the CVM literature, the endowment effect, is also present in SG. Therefore it is argued that it would be premature to use observed inconsistencies to reject one method in favour of the other, but that the way forward is to identify apparent inconsistencies that these methods have in common to help determine whether they are simply a product of one method or part of underlying preferences.

Key words: Expected Utility; Endowment Effect; Standard Gambles; Reference Points
JEL Classification: D1; D8; C9

[*] Corresponding author: Gwendolyn C. Morrison, School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, UK. Tel: +44 (115) 951-4768; FAX: +44 (115) 951-4159; email: wendy.morrison@nottingham.ac.uk

1. Introduction

The endowment effect first arose in the contingent valuation literature as a possible explanation for the unexpected, but frequently observed, disparity between willingness to pay (WTP) and willingness to accept (WTA) measures of value (Thaler, 1980; Kahneman et al., 1990). The endowment effect is a reference point effect whereby preferences are in part determined by the current endowment. In particular, people appear to place a higher value on goods when they already own them than they did before they acquired them. With respect to the contingent valuation method, the starting point of the elicitation question effectively endows respondents with a bundle, generally money and another good, and asks them to trade-off some of one to acquire some (more) of the other. The endowment effect embodies the observation that respondents appear to exhibit a sort of loss aversion (Kahneman and Tversky, 1979)—they are reluctant to give up any part of their current endowment. That is, in a WTP question respondents are resistant to parting with their money in exchange for a good, and in a WTA question they are reluctant to give up the good in exchange for money. Thus, an individuals' WTP to acquire a good is greatly exceeded by the amount of compensation they would require in order to willingly give up that same unit of the same good—the price at which they would buy is less than the price at which they would sell.

Some researchers have used such internal inconsistencies as an argument against the application of CVM in a policy setting context (Diamond & Hausman, 1994), and even advocated using other methods like expected utility (EU) estimation techniques—standard gambles (SG)—instead (Jones-Lee et al., 1995). Before using such anomalies as a basis for rejecting one method in favour of the other, we must first check that the problems in CVM are not also present in SG. Morrison (2000) describes how an endowment effect would manifest itself if present in EU estimation techniques and then demonstrates that internal inconsistencies that have been observed in SG studies[1] indicate that such an effect is also at work in SG. But those studies were not intended to test for this anomaly, and so a stronger test must be applied before any conclusions can

be drawn. This paper presents the results of an experiment that was designed to test for the presence of an endowment effect in SG.

2. Experimental Design

As presented by Morrison (1997a), the endowment effect can be illustrated as a pivot of an individual's indifference map from the point of endowment thereby effectively increasing the utility associated with the endowment relative to other bundles. Thus, an individual may be indifferent between two bundles, A and B, when simply asked to choose between them, but once endowed with bundle A, A becomes better than B and so they will not trade if given the opportunity. But this is a reference point effect and so if they were instead endowed with B and asked to trade it for A, they again would refuse to trade. In the context of EU estimation techniques, as illustrated in Morrison (2000), such an effect would make an individual appear more risk averse than they really are when responding to a PE question and less risk averse than they really are when responding to a CE question.

The rationale is as follows. The starting point of each PE question effectively endows each subject with a certain state. An endowment effect, if present, would lead the individual to state a probability of failure that is smaller than they would in the absence of such an effect. That is, in order to be persuaded to part with the certain state with which they were endowed, the respondent will select a gamble with a greater expected value than they would in the absence of an endowment effect. Therefore, when responding to a PE question, a risk neutral individual will appear risk averse. Conversely, CE questions endow individuals with a gamble and ask them to identify a state that would make them indifferent between having that selected state for certain and keeping the gamble with which they were endowed. An endowment effect effectively increases the value of that gamble to the individual and, thus, leads them to nominate a certain state of a higher value than they would do otherwise. Thus, a risk neutral individual will appear risk seeking when responding to a CE question.

However, it is not necessary to know individuals' risk attitudes in order to test their SG responses for the presence of an endowment effect. Given that an endowment effect will make individuals value goods more highly when they own them than when they do not, within subject comparison of valuations of a good when it is owned and when it is not should reveal an endowment effect if it is present. A two part experiment is used to test this: in the first part of the experiment respondents complete PE and CE questions; the second part asks those respondents to choose between the starting point of each PE and CE question and their own answer to that question. Since the choice questions do not issue respondents with an endowment, it follows that in the absence of an endowment there will be no endowment effect. Examining patterns of responses to the second part of the experiment provides a test for the endowment effect in SG questions. If there is no endowment effect operating in PE responses, then when given the choice between the starting point (endowment) of a PE question (say, 'X') and their own response to that PE question ('Y'), they should on average be indifferent between the starting point of the question and their own response. However, if an endowment effect is at work, then subjects should on average prefer their own response over the starting point of the question, because their own response incorporates an endowment effect.

CE questions were tailored to each respondent using their own answers to each PE question ('Y') as the starting point for the corresponding CE questions. As for the PE questions, if there is no endowment effect, then, when given the choice, the respondent should be indifferent between the starting point of the CE question and their own response to that question ('Z'). But, if on average the subjects prefer their own responses to the starting points of the questions, then this is taken as evidence of an endowment effect.

That the end point of the PE question is the starting point for the CE question provides a third test for an endowment effect. When given the choice between the starting point of a PE question and their own response to the corresponding CE question, transitivity dictates that respondents should consider the two to be of equal value (ie, since X=Y and Y=Z, it must follow

that X=Z). However, if there is an endowment effect, then respondents will prefer their own responses to the CE question to the starting point of the PE question (ie, if XAY and YAZ, then it must be that XAZ).

3. Questionnaire and Sample

Standard gambles are considered by many health economists to be the "gold standard" of health status measurement.[2] Since SG questions are often used to quantify health states (Torrance, 1986; Stiggelbout et al, 1994; Jones-Lee et al, 1995; Jansen et al, 1998), these experiments tested for an endowment effect in responses to SG questions set in the context of health. PE and CE questions were used to elicit EU estimates for four different injuries relative to normal health (which was assigned a utility of 100) and immediate death (assigned a utility of 0).

First respondents were asked to read the injury descriptions (these are listed in appendix A) and rank them from best to worst. Then they were asked to imagine that they had been injured in a road accident and, if the normal course of treatment were followed, then they would experience the injury described on the answer sheet (A, B, C, or D) for certain. However, they could instead accept a new risky treatment which, if successful would give them a full recovery, but if a failure would result in immediate death. Individuals were asked to indicate three kinds of answers on the response sheet (an example of the response sheets for a PE, CE, and choice question are in appendix B). First they were asked to put a check mark next to the chances of failure (or chances of success) for which they were sure that they would accept the risky treatment. Then, they were asked to put a cross (✗) next to any chances of failure (success) for which they were sure they would not take the risky treatment. Finally, they were asked to find the chances failure (success)—between the lowest percentage with a check and the highest percentage with a cross—that would make it most difficult for them to choose between the certain injury and the risky treatment.

A CE question followed each PE question, and each subject was asked to transcribe their response to the PE question, in terms of chances of failure, onto the starting point of the CE

question. So the starting point for the CE question was transparently their own response from the PE question. They would then be asked to indicate the state that would make it most difficult for them to choose between taking that state for certain and keeping the gamble, as in the example CE response sheet in appendix B.[3]

The final example question shown in appendix B is a Choice question. These were constructed from the starting points of and responses to the PE and CE questions. Each respondent was told that if they were roughly indifferent between two options then they should check the box indicating that they, "consider the two to be equal." Alternatively, if they preferred one option over the other, then they should check the box indicating they preferred that option. They were asked to indicate their preference in this way for each pair in their question sheet. There were a total of twelve pair-wise choices in the questionnaire: the starting point and their own response to each of the four PE questions, the starting point and their response to each CE question, and the starting point to each PE question and the response to the corresponding CE question.

In October 1998, students taking first year microeconomics for non-economics majors at the University of Nottingham took part in a two-part pilot experiment. The first part, which entailed responding to PE and CE questions, was conducted as part of a lecture. Responses to those questions were then used to construct individualised questions for the second part of the experiment—the choice experiment. Students were told that they would be paid £5 (≈US$7.50) to arrive half an hour early for the next lecture, which was one week later, in order to answer some more questions. Twenty-eight students completed both parts of the experiment.

4. Results

Table 1 summarises the results of the experiment. For each of the four injuries valued, subjects could respond to the choice questions in one of three ways: stating that they prefer the starting point of the question, they prefer their own response, or that they are roughly indifferent between the two (see appendix B for an example). Table 1 notes the number of people responding

in each way and presents the results of a one-tailed paired samples t-test for the null hypothesis that there is no endowment effect, against the alternative that an endowment effect is present.

[Table 1 about here]

First, consider the choice experiment responses with respect to the PE questions. If an endowment effect is at work, then individuals should choose the option that is in fact their own response. For injuries B, C, and D this is how most people responded and this is statistically significant at α=.01 for injuries B and D. Although the direction of preference goes against the endowment effect hypothesis for injury A, that is statistically insignificant. So, considering only the PE part of the experiment, there is some evidence in support of an endowment effect. The results, however, are stronger with respect to the CE questions, and for the comparison of the starting point of the PE questions with responses to the corresponding CE question. For all injuries being evaluated, the null of equality had to be rejected in favour of the alternative that individuals preferred their own responses with respect to the CE question and, again, in all four cases there is a significant preference for responses to the CE questions over the starting point of the corresponding PE question (α=.01). So, in ten of the twelve within subject pair-wise comparisons a significant pattern emerged from responses to the choice questions and, in all of those cases, that observed asymmetry is consistent with the existence of an endowment effect.

5. Conclusion

The accumulation of experimental evidence showing persistent internal inconsistencies has had different reactions from the profession with respect to the contingent valuation method and EU estimation techniques (ie, standard gambles). Experimental investigations into CVM could be described as a concerted effort to identify the source of systematic inconsistencies. Meanwhile, observed inconsistencies in standard gambles have, rather, led to a plethora of non-expected utility models (eg, Loomes and Sugden, 1982; Quiggin, 1982) with no apparent consensus as to the direction in which research should go.

Internal inconsistencies are now so commonplace in studies using the contingent valuation method that it has been argued that that method should be abandoned as a means of shadow pricing, and even that it be abandoned in favour of other methods of preference elicitation like standard gambles. Before adopting standard gambles in place of CVM, it would be prudent to first check that SG is free of inconsistencies—at the very least, we should ensure that the same inconsistencies that led to the rejection of CVM are not also present in SG. However, internal inconsistencies are also well documented in SG, and none of the non-EU models that have emerged to date have been found to be empirically superior to EU (Hey & Orme, 1994; Harless & Camerer, 1994). So, given the current state of the art, it is difficult to see how SG could be regarded as a better preference elicitation technique than CVM.

Despite the fact that experimental investigations into the two methods have been going on concurrently, little attention has been devoted to relating the inconsistencies found in EU estimates to those found in CVM studies. The experiment described this paper finds that an anomaly that is well documented in the CVM literature—the endowment effect—is also present in SG. So, calls to reject CVM, particularly in favour of SG, seem misguided or at least premature.

Rather than using apparent inconsistencies as a reason to reject one method, or trying to modify the EU model to accommodate inconsistencies observed in individual experiments, perhaps the way forward is to identify systematic inconsistencies that a number of elicitation methods have in common. This information might modify our understanding of basic preferences and, so, help to determine what does and does not constitute an "inconsistency." In turn, this can help direct us to likely sources of bias in survey questions and, therefore, to what should be defined as "best practice."

Acknowledgements
My thanks to the University of Nottingham Research Committee for providing the funds to conduct this experiment, and to the University of Nottingham students that took part.

References

Bleichrodt, H. (1996), "Applications of utility theory in the economic evaluation of health care", Ph.D. thesis, Erasmus University Rotterdam.

Diamond, P.A. and Hausman, J.A., (1994). Contingent Valuation: Is Some Number Better than No Number? Journal of Economic Perspectives 8(4), 45-64.

Harless, D.W. and Camerer, C.F., 1994, The predictive utility of generalized expected utility theories, Econometrica, 62(6), 1251-1289.

Hershey, J.C., and Schoemaker, P.J.H., 1982, Sources of bias in assessment procedures for utility functions, Management Science 28(8), 936-954.

Hershey, J.C., and Schoemaker, P.J.H., 1985, Probability versus Certainty Equivalence methods in utility measurement: Are they equivalent? Management Science 31(10), 1213-1231.

Hey, J.D. and Orme, C., 1994, Investigating generalizations of Expected Utility theory using experimental data, Econometrica, 62(6), 1291-1326.

Jansen, S.J.T.; Stiggelbout, A.M.; Wakker, P.P.; Vliet vlieland, T.P.M.; Leer, J.H.; and Nooy, M.A.; Kievit, J., 1998, Patient utilities for cancer treatments: A study of the chained procedure for the standard gamble and time trade-off, Medical Decision Making, 18(4), 391-399.

Jones-Lee, M.W.; Loomes, G.; and Philips, P.R., 1995, Valuing the prevention of non-fatal road injuries: contingent valuation vs standard gambles, Oxford Economic Papers, 47, 676-695.

Kahneman, D.; J.L. Knetsch; and R.H. Thaler, 1990, Experimental tests of the endowment effect and the coase theorem, Journal of Political Economy 98(6), 1325-1348.

Kahneman, D. and A. Tversky, 1979, Prospect theory: an analysis of decision under risk, Econometrica 47(2), 263-291.

Llewellyn-Thomas, H.; Sutherland, H.J.; Tibshirani, R.; Ciampi, A.; Till, J.E.; & Boyd, N.F., 1982, The measurement of patients' values in medicine, Medical Decision Making, 2, 449-462.

Loomes, G. and Sugden, R., 1982, Regret Theory: An alternative theory of rational choice under uncertainty, Economic Journal 92, 805-824.

Mehrez, A. and Gafni, A., 1991, The Healthy-years Equivalents: how to measure them using the standard gamble approach, Medical Decision Making, 11, 140-146.

Morrison, G.C. (1996), "An Assessment and Comparison of Selected Methods for Eliciting Quality of Life Valuations from the Public," D.Phil. Thesis, University of York, UK.

Morrison, G.C., 1997a, Resolving differences in willingness to pay and willingness to accept: comment, American Economic Review 87(1), 236-240.

Morrison, G.C., 1997b, HYE and TTO: what is the difference? Journal of Health Economics, 16(5), 563-578.

Morrison, G.C., 2000, The Endowment Effect and Expected Utility, Scottish Journal of Political Economy 47(2), 183-197.

Quiggin, J., 1982, A theory of anticipated utility, Journal of Economic Behaviour and Organization 3, 323-343.

Rutten-van Mölken, M.P.M.H.; Bakker, C.H.; van Doorslaer, E.K.A.; and van der Linden, S., 1995, "Methodological issues of patient utility measurement: experience from two clinical trials," Medical Care, 33(9), 922-937.

Stiggelbout, A.M.; Kiebert, G.M.; Kievit, J.; Leer, J.W.H.; Stoter, G.; & de Haes, J.C.J.M., 1994, Utility assessment in cancer patients: Adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores, Medical Decision Making, 14, 82-90.

Thaler, R., 1980, Toward a Positive Theory of Consumer Choice, Journal of Economic Behavior and Organization, 1, 39-60.

Torrance, G.W., 1986, Measurement of health state utilities for economic appraisal: A review, Journal of Health Economics, 5, 1-30.

Table 1
Choice Experiment

injury[d]	# prefer start	no pref	# prefer own response	t-stat[a,b,c]	endowment effect?
Probability Equivalent questions[a]					
A	15	1	12	-0.570	✗
B	3	7	18	4.091*	✓
C	8	9	11	0.862	✓
D	5	5	18	3.099*	✓
Certainty Equivalent questions[b]					
A	1	6	19	6.429*	✓
B	1	4	23	8.337*	✓
C	1	3	19	7.240*	✓
D	2	4	17	4.832*	✓
Starting point of PE vs Response to CE questions[c]					
A	5	3	18	3.138*	✓
B	0	3	25	15.000*	✓
C	1	2	20	8.068*	✓
D	0	1	24	24.000*	✓

[a], H_0: V(gamble)=V(certain state); H_1: V(gamble)>V(certain state); [ie, gamble preferred]
[b], H_0: V(certain state)=V(gamble); H_1: V(certain state)>V(gamble); [ie, certain state preferred]
[c], H_0: $V(\text{certain}_{\text{RESPONSE}})=V(\text{certain}_{\text{START}})$; H_1: $V(\text{certain}_{\text{RESPONSE}})>(\text{certain}_{\text{START}})$; [ie, own response preferred]
* significant at α=.01
[d], injury descriptions are in appendix A.

Appendix A.
Injury Descriptions

Injury A

In hospital
- 3 weeks
- in some pain

After hospital
- Loss of one leg (amp at knee)
- can walk with artificial limb, but can no longer take part in some sports
- live for 55 more years

Injury B

In hospital
- 3 weeks
- in some pain

After hospital
- Lost use of both arms and legs
- no change in mental abilities
- require a full-time carer for all basic needs (eating, toiletting, …)
- live for 55 more years

Injury C

In hospital
- 3 weeks
- in some pain

After hospital
- Complete loss of hearing (nb, a hearing aid cannot help)
- no other loss of abilities (neither physical nor mental)
- live for 55 more years

Injury D

In hospital
- 3 weeks
- in some pain

After hospital
- permanent brain damage
- no loss of physical abilities
- mental abilities significantly reduced (can no longer continue studies or current career aspirations)
- able to hold low skill job & to care for yourself & a family
- live for 55 more years

Injury F

- immediate death

Injury S

In hospital
- 3 weeks
- in some pain

After hospital
- full recovery
- live for 55 more years

Appendix B.

Example Probability Equivalent question # 1

Injury A
In hospital
• 3 weeks
• in some pain
After hospital
• Loss of one leg (amp at knee)
• can walk with artificial limb, but can no longer take part in some sports
• live for 55 more years

for certain

OR

Success outcome
In hospital
• 3 weeks
• in some pain
After hospital
• full recovery
• live for 55 more years

Failure Outcome
• immediate death

Chances of Success	Chances of Failure
	Won't accept treatment at any risk of failure
more than 99 in 100	less than 1 in 100
99 in 100	1 in 100
98 in 100	2 in 100
97 in 100	3 in 100
96 in 100	4 in 100
95 in 100	5 in 100
94 in 100	6 in 100
92 in 100	8 in 100
90 in 100	10 in 100
87 in 100	13 in 100
84 in 100	16 in 100
80 in 100	20 in 100
75 in 100	25 in 100
70 in 100	30 in 100
65 in 100	35 in 100
60 in 100	40 in 100
55 in 100	45 in 100
50 in 100	50 in 100
45 in 100	55 in 100
40 in 100	60 in 100
35 in 100	65 in 100
30 in 100	70 in 100
25 in 100	75 in 100
20 in 100	80 in 100
15 in 100	85 in 100
10 in 100	90 in 100
5 in 100	95 in 100
	Death is preferred

Please put a ✓ against the maximum risk of failure for which you are CONFIDENT that you would CHOOSE THE TREATMENT

Please put a ✘ against minimum risk of failure for which you are CONFIDENT that you would REJECT THE TREATMENT and accept the injury described (injury A)

Please circle the case where you consider the alternatives to be roughly equal

Example Certainty Equivalent question # 1

You now have a **choice between**,

1) the gamble that you just selected (circled)

Risky treatment with a risk of failure of _____ in 100

or,

2) ***experiencing the following health state for certain***

In hospital
- 3 weeks
- in some pain

After hospital
- full recovery

but only living for _____ more years (£ 55 more years)

State the minimum number of years that would make you indifferent between
- having the risky treatment, and
- having certain recovery with a shorter life expectancy
(be as precise as you wish—ie, state years/months/weeks/days if you want)

Example Choice question # 1

In this section please choose which of the alternatives you would prefer.

Certain Prognosis ~

In hospital
- 3 weeks
- in some pain

After hospital
- Loss of one leg (amputated at knee)
- can walk with artificial limb, but can no longer take part in some sports
- live for 55 more years

OR

Risky Treatment ~

Risk of Failure = _____
…….
Failure: Immediate Death

Success:
In hospital
- 3 weeks
- in some pain

After hospital
- full recovery
- live for 55 more years

Please check the box corresponding to your choice:

☐	☐	☐
Prefer the certain prognosis	Consider the two to be equal	Prefer to take the risky treatment

Footnotes

[1] Specifically the probability equivalent/certainty equivalent (PE/CE) disparity (Hershey and Shoemaker, 1982 & 1985) and the chained and unchained PE estimates (eg, Llewellyn-Thomas et al, 1982; Rutten-van Mölken et al, 1995; Bleichrodt, 1996; Morrison, 1996).

[2] For example, in order to estimate how much of a health improvement is brought about by a treatment, it is necessary to estimate the value of health before and after treatment. PE questions are often used in this context by assigning utilities of 100 and zero to perfect health and immediate death, respectively. One method of health status valuation, the Healthy-Years Equivalent, HYE (see Mehrez and Gafni, 1991), uses both PE and CE questions. Morrison (1997b) explains how the PE/CE disparity reported in the literature suggests that the HYE will yeild upward biased estimates.

[3] Note that the response sheet for the CE question held the quality of life constant and allowed the respondent to vary only the length of life. This is the format used when the Healthy-Years Equivalent (HYE) technique for valuing health is used (Mehrez & Gafni, 1991). Morrison (1997b) argued that the PE-CE disparity suggests that the HYE would result in systematically upward biased estimates of Time Trade-Off values. This is why the PE and CE questions used in this experiment were set in the context of health.