ELSEVIER

Journal of Public Economics xx (2004) xxx–xxx

JOURNAL OF PUBLIC ECONOMICS

www.elsevier.com/locate/econbase

Testing competing models of loss aversion: an adversarial collaboration

Ian Bateman[a], Daniel Kahneman[b], Alistair Munro[c], Chris Starmer[d], Robert Sugden[c,*]

[a]*School of Environmental Sciences, University of East Anglia, England*
[b]*Woodrow Wilson School for Public and International Affairs, Princeton University, United States*
[c]*School of Economics, University of East Anglia, Norwich NR4 7TJ, England*
[d]*School of Economics, University of Nottingham, United Kingdom*

Received 10 April 2003; received in revised form 28 November 2003; accepted 14 June 2004

Abstract

This paper reports an 'adversarial collaboration'—a project carried out by two individuals or research groups who, having proposed conflicting hypotheses, seek to resolve their dispute. It describes an experiment which investigates whether, when individuals consider giving up money in exchange for goods, they construe money outlays as losses or as foregone gains. This issue bears on the explanation of the widely observed disparity between willingness-to-pay (WTP) and willingness-to-accept (WTA) valuations of costs and benefits, which has proved problematic for contingent valuation studies. The results of the experiment are broadly consistent with the hypothesis that money outlays *are* perceived as losses.

JEL classification: D11 (consumer economics: theory); C91 (design of experiments: laboratory, individual behaviour); H43 (publicly provided goods: project evaluation)

Keywords: Adversarial collaboration; Loss aversion; Willingness to pay; Reference-dependent; Preferences; Status quo bias

* Corresponding author.
E-mail address: r.sugden@uea.ac.uk (R. Sugden).

doi:10.1016/j.jpubeco.2004.06.013

This paper has two goals. The first is to introduce a method—adversarial collaboration—which we believe is new to experimental economics.[1] The second is to use this method to discriminate between two opposing hypotheses concerning the boundary conditions for the occurrence of loss aversion, a phenomenon that has been implicated in some important economic debates.

In the context of public economics, loss aversion is particularly significant as a putative cause of disparities between willingness-to-pay (WTP) and willingness-to-accept (WTA) valuations of public goods, as revealed in contingent valuation studies. Contingent valuation is widely used to elicit individuals' preferences to guide public policy on health, safety and environmental issues (see, e.g., Bateman and Willis, 1999). However, its practitioners confront a recurrent problem: preferences, as elicited in contingent valuation surveys, often fail to satisfy properties of internal consistency and context-independence that are axiomatic in the theoretical framework of neoclassical economics; yet that is the framework that is used to organise, interpret and analyse the survey data. One of the most common and challenging of such problems is that WTA valuations often exceed WTP valuations to an extent that is inconsistent with received theory. Among practitioners, there is a widespread opinion that WTP data are more appropriate or reliable than WTA, but the justification for this position remains controversial.[2]

One explanation of the WTA/WTP disparity is offered by Tversky and Kahneman's (1991) *reference-dependent preference theory*, in which preferences are defined in relation to 'reference states'. Tversky and Kahneman propose that reference-dependent preferences between given options vary systematically according to the reference state from which they are assessed, as a result of *loss aversion*—the tendency for 'losses [to] loom larger than corresponding gains'. In this theory, the disparity arises because WTA and WTP valuations elicit preferences relative to different reference states. However, the mechanism by which loss aversion impacts on WTA and WTP is a subject of dispute.

As a subsidiary hypothesis, Tversky and Kahneman propose that, while the act of selling a good for money (WTA) is construed as a loss of that good, the act of giving up money to buy goods (WTP) is construed as a foregone gain of money, not as a loss. Thus, WTA reflects a comparison between *losses* of the good and *gains* of money, while WTP reflects a comparison between *gains* of the good and (foregone) *gains* of money. By virtue of this hypothesis, the WTA/WTP disparity is caused only by loss aversion in the good; loss aversion in money has no impact on the disparity. Tversky and Kahneman claim that this hypothesis is supported by the results of an experiment carried out by Kahneman et al. (1990), which we shall call the *Vancouver experiment*.

A different form of reference-dependent theory is assumed by Bateman et al. (1997) and formalised by Munro and Sugden (2003). In this version of the theory, there is symmetry between WTA and WTP: the acts of giving up goods when sold for money and of giving up money to buy goods are both construed as losses. On this account, the WTA/

[1] Adversarial collaboration has been used previously in psychology. Mellers et al. (2001) pioneer the method and propose a protocol for such collaboration.

[2] The influential NOAA Panel on Contingent Valuation recommended the WTP format (Arrow et al., 1993). Knetsch (2000) argues that WTP and WTA are conceptually symmetrical: WTP should be used to evaluate gains, WTA to evaluate losses.

WTP disparity is caused both by loss aversion in the good and by loss aversion in money. Bateman et al. appeal to this notion of symmetry to suggest that, if any valuation measure can be defended as 'neutral', *equivalent gain* (EG)—that is, the increment of money which an individual treats as equivalent to the gain of some given quantity of the relevant good—has a better claim to neutrality than either WTA or WTP. Bateman et al.'s hypothesis implies that EG is greater than WTP as a result of loss aversion in money, while Tversky and Kahneman's implies that the two measures are equal. Bateman et al. report an experiment—the *Norwich experiment*—which they interpret as confirming reference-dependent theory as formulated by them.

This paper reports an attempt to resolve the issue of whether money spent in buying goods is perceived as a loss. This investigation was carried out as an adversarial collaboration between Kahneman and the other authors (the 'Norwich group').[3] In Section 1, we explain the general principles of adversarial collaboration. In Sections 2 and 3, we explain the particular issue that our adversarial collaboration was intended to resolve, and derive specific predictions from our rival prior hypotheses. Section 4 describes the experiment we used to test these predictions. In Sections 5 and 6, we report our results and discuss their implications. In broad terms, the results of our experiment favour the hypothesis that money expenditures are treated as losses. However, taking account of conflicting evidence from some other related experiments, the question of which of the competing hypotheses is closer to the truth remains open.

1. Adversarial collaboration

In an adversarial collaboration, the two parties agree on the design of an experiment which they will conduct jointly. Before knowing what the experiment will find, they accept its validity as a test of their respective hypotheses.[4] Each party anticipates its interpretation of possible outcomes of the experiment, particularly those that it does *not* predict. The two parties agree that particular outcomes of the experiment would support one hypothesis, and particular other outcomes would support the other. Both parties commit to publishing the results, whatever they may be. We believe that this methodology has some advantages over the more conventional form of scientific debate, in which each research group designs and runs experiments independently, chooses which of its results to publish, and can challenge the validity of other groups' experimental designs after knowing the results those designs have produced. Adversarial collaboration encourages a more constructive approach to the resolution of disagreements.

Adversarial collaboration, as compared with conventional scientific debate, requires different attitudes on the part of researchers—in particular, more attention to understanding the other side's arguments, and less to rhetorical strategies for defeating them. But it also requires different expectations on the part of the scientific community as a whole. We are

[3] At the time of the collaboration, all the members of this group were based at the University of East Anglia, Norwich. Starmer has since moved to the University of Nottingham.

[4] Of course, for reasons widely recognised in the philosophy of science, no single test can ever be regarded as conclusive.

all used to reading journal articles which report apparently clear-cut experimental results and which draw strong conclusions from them. But as readers, we learn to apply some discount to such claims. We have to allow for confirmation bias in the design of experiments—the tendency for researchers to look for 'tests' which seem likely to confirm their prior hypotheses. We also have to keep in mind that the experiments that are reported in journals are not necessarily representative of the larger set of experiments that have been run: we have to allow for the possibility that research groups publish only their most 'successful' experiments, and use conformity with their prior hypotheses as one of their criteria of success. And we make allowance for the incentives that induce authors to talk up their conclusions, to draw wide-ranging implications from them and to play down doubts and ambiguities. Adversarial collaboration must be expected to lead to a different kind of publication.

Because the experimental designs used in adversarial collaboration have to be agreed by both parties, each party has to subject its hypothesis to a genuinely stringent test: one which that party regards as valid, but which the other party expects it to fail. Thus, one of the mechanisms which tends to generate *apparently* decisive experimental results in the existing literature, positive confirmation bias, is neutralised. The commitment to publication, backed up by the two parties' common knowledge of the outcomes of the experiment, neutralises another such mechanism: selection bias at the publication stage.

Adversarial collaboration will not always bring the parties into full agreement about the issue in dispute: they may have different interpretations of what their jointly conducted experiment has found. Scientific debate is better served if such differences are reported frankly than if they are concealed by bland generalities. From the reader's point of view, a report of this kind may be more useful than the superficially more definite conclusions that are customarily expected of non-adversarial research papers. Ultimately, however, the value of an adversarial collaboration is to be found in the validity and power of the experimental design it has adopted, and in the quality of the data generated. Whether, having seen the results, the parties to the collaboration agree on how they should be interpreted is a secondary matter: it is the reader who must draw the conclusions.

2. Theoretical background

The hypothesis that there is an asymmetry between individuals' attitudes to gains and to losses was first brought to the attention of economists by Kahneman and Tversky (1979) and by Thaler (1980). Since then, there has been an accumulation of evidence—from experiments, from survey data, and from the field—which suggests that individuals' choices are more responsive to anticipated losses than to equal and opposite anticipated gains.[5] These findings are inconsistent with the standard (or *Hicksian*) theory of consumer choice, in which preferences are defined over final consumption states.

[5] Bateman et al. (1997), Sugden (1999) and Starmer (2000) survey relevant experimental and survey research. See also Myagkov and Plott (1997), who find loss aversion in an experimental market setting, and Samuelson and Zeckhauser (1988), Benartzi et al. (1995), and List (2003), all of whom find evidence of loss aversion in 'real' markets.

In Tversky and Kahneman's (1991) theory of reference-dependent preferences, individuals have preference orderings over bundles of goods, as in Hicksian consumer theory, but these preferences are defined relative to *reference states*. A reference state is a point in goods space which the individual treats as the status quo or normal expectation; gains and losses are defined in terms of displacements from the reference state. In notation, reference states are represented by subscripting the preference relation; thus 'x is weakly preferred to y, viewed from the reference state r' is written as $x \succcurlyeq_r y$. Tversky and Kahneman propose a hypothesis of loss aversion which links changes in (reference-dependent) preferences with changes in the reference state. Let $x=(x_1, \ldots, x_n)$ and $y=(y_1, \ldots, y_n)$ be two bundles of the same n goods, such that for some good i, $y_i>x_i$. Let $r=(r_1, \ldots, r_n)$ and $s=(s_1, \ldots, s_n)$ be potential reference states, such that $r_i=x_i$, $s_i=y_i$, and $r_j=s_j$ for all $j \neq i$. For any given i, there is loss aversion in good i if, for all such x, y, r, s: $y \succcurlyeq_r x \Rightarrow y \succ_s x$. Tversky and Kahneman's hypothesis is that there is loss aversion, so defined, in all goods.

In order to apply this theory to a concrete choice problem, it is necessary to specify the chooser's reference state. The theory itself does not tell us how reference states are determined; in this respect, the formal model is left uninterpreted (Tversky and Kahneman, 1991, pp. 1046–1047). Thus, how the theory should be applied to specific decision-making environments can be a matter of judgement.

One natural interpretation, the *current endowment hypothesis* (CEH), is that an individual's reference state is the bundle of goods she currently owns: it is what she will hold or consume if she does not engage in any transactions beyond those she has already made. According to this hypothesis, when an individual deliberates about whether or not to buy a good, she does so in relation to a current endowment which includes her holdings of money. Thus, the money that would be spent in buying the good is perceived as a loss. This was the background hypothesis used to develop the tests in the Norwich experiment.[6] Notice that the definition of 'current endowment' presupposes a distinction that Hicksian choice theory does not recognise: that between transacting and not transacting. In Hicksian theory, a choice problem is represented simply as a set of mutually exclusive options, one of which must be chosen; there is no concept of a default option which will come about if the agent 'does nothing'. Thus, two choice problems that have identical representations in Hicksian theory may have different current endowments.

However, Tversky and Kahneman (1991) do not fully endorse CEH, even with respect to simple laboratory tasks. For example, consider the Vancouver experiment, in which subjects bought and sold coffee mugs for money, indicating their willingness to trade by reporting their valuations of mugs as 'buyers', as 'sellers' and as 'choosers'. Noting that choosers' valuations were found to be much closer to buyers' valuations than to sellers', Tversky and Kahneman suggest that buyers do not value the money they give up in a transaction as a loss. Their hypothesis is that a subject who is considering buying a coffee mug construes her reference state as including *neither* the mug *nor* the money she would

[6] The Norwich group does not claim that the current endowment hypothesis applies to all decision situations. For example, that hypothesis might not apply if decision-makers perceive current endowments as unfair or morally wrong.

have to spend to buy it. If she chooses to buy, she gains the mug; if not, she gains whatever she would buy with the money instead. We shall call the general form of this hypothesis *no loss in buying* (NLIB). Notice that NLIB is a hypothesis about reference states and not about loss aversion. It proposes that money spent in buying goods is not 'coded' as a loss: it does not propose that there is no loss aversion in money.

Kahneman has subsequently firmed up the NLIB hypothesis by proposing a theory of the conditions under which the gains and losses associated with a transaction are mentally *integrated* prior to evaluation, rather than being evaluated separately. When a loss and an equal and opposite gain are integrated in this way, painful perceptions of loss do not arise. Kahneman proposes that consumers normally have *budget reserves*, that is, reserves of money that are available for unanticipated spending. When an individual faces an unanticipated opportunity to buy a good, and is able to finance this spending from her budget reserve, gains and losses are integrated: the money that has to be spent to buy the good is already seen as a token for unspecified goods. In such circumstances, money outlays are not perceived as losses. In contrast, if the individual faces an unanticipated buying opportunity which she can finance only by forgoing some *specific* consumption plan, the act of buying involves a definite loss, separable from the gain; and so the money payment is perceived as a loss. In the converse case of selling, gains and losses are integrated if the proceeds of the sale are earmarked for the purchase of a replacement good, but they are treated separately if those proceeds will be added to the budget reserve (Kahneman and Novemsky, 2002).

The objective of our adversarial collaboration was to test CEH against NLIB. Our experimental design elicits subjects' willingness to engage in transactions involving money and low-value, non-staple goods. If the idea of budget reserves is accepted, it seems reasonable to assume that subjects' purchases of such goods can be financed from budget reserves, and that the proceeds from sales are not earmarked for replacements. Thus, we take it to be an implication of NLIB that, in such an experiment, gains and losses are integrated with respect to buying but not with respect to selling.

3. Implications of the competing hypotheses

Consider a model in which there are only two goods; quantities of these goods are represented by x_i, x_j. Usually, we interpret this model such that one good is some particular consumption good (for short, 'the good') and the other is an index of general purchasing power, or 'money'. When we use this interpretation, the good is denoted by G and money by M. This model may be interpreted either in relation to real transactions (such as those in our experiment) involving money and a private consumption good, or in relation to hypothetical transactions presented to participants in contingent valuation surveys. In the latter case, we assume that respondents report how they believe they would choose, were the relevant problem for real. Thus, reference states, endowments, gains and losses are defined in terms of the scenario presented in the relevant survey instrument.

For any given individual, consider how one unit of good i can be valued in units of good j. More specifically, for any given x'_i, x'_j, suppose that the individual is endowed with x'_j of good j, and consider how we might express in units of good j the value of consuming x'_i+1 units of good i rather than x'_i units. We begin by defining three measures of this value.

Willingness-to-pay (WTP). Suppose the individual's current endowment is (x'_i, x'_j). WTP_{ji} is the largest amount of good j that the individual would be willing to give up in return for a gain of one unit of good i.

Willingness-to-accept (WTA). Suppose the individual's current endowment is (x'_i+1, x'_j). WTA_{ji} is the smallest amount of good j that the individual would be willing to accept in return for accepting a loss of one unit of good i.

Equivalent gain (EG). Suppose the individual's current endowment is (x'_i, x'_j). EG_{ji} is the smallest amount of good j that the individual would be willing to accept in place of a gain of one unit of good i.[7]

WTP and WTA are commonly used in contingent valuation and cost–benefit studies. Although EG is used only rarely (Bateman et al., 2000, is an example), its mirror image, *equivalent loss* (EL) is a standard measure (see, for example, Rowe et al., 1980; Carson and Mitchell, 1993; Cummings and Taylor, 1999). EL is usually described as a measure of willingness to pay money to avoid reductions in the quality or quantity of a good. Analogously, EG can be thought of as a measure of willingness to accept money in place of an increase in the quality or quantity of a good.

By introducing risk into the elicitation process, we can define two further valuation measures, which play important roles in our experimental design:

Risky willingness to pay (RWTP). Suppose the individual's current endowment is (x'_i, x'_j). Consider a gamble with two mutually exclusive outcomes, each with probability 0.5. One outcome is that the individual gains one unit of good i, with no change in good j. The other is that she loses some amount of good j, with no change in good i. RWTP_{ji} is the largest such loss of good j consistent with her being willing to accept the gamble.

Risky willingness to accept (RWTA). Suppose the individual's current endowment is (x'_i+1, x'_j). Consider a gamble with two mutually exclusive outcomes, each with probability 0.5. One outcome is that the individual loses one unit of good i, with no change in good j. The other is that she gains some amount of good j, with no change in good i. RWTA_{ji} is the smallest such gain of good j consistent with her being willing to accept the gamble.

In most economic applications, valuation measures are used to express the value of given quantities of non-money goods in terms of money. That is, the *response mode*—the

[7] If choice problems are represented in terms of Hicksian theory, WTA_{ji} and EG_{ji} are identical. In Hicksian terms, there is some x_j* such that the individual reveals indifference between (x'_i+1, x'_j) and (x'_i, x'_j+x_j*); x_j* is measured both by WTA_{ji} and by EG_{ji}. However, WTA_{ji} and EG_{ji} are defined with respect to different current endowments. In the case of WTA_{ji}, the individual reveals indifference between sticking with her current endowment and making an exchange transaction in which she loses one unit of good i and gains x_j*. In the case of EG_{ji}, she reveals indifference between two unilateral transfers, in one of which she gains one unit of good i, in the other of which she gains x_j*.

unit in which the valuation is expressed—is money. But our definitions also allow these measures to be used to express the value of a given quantity of money in terms of a non-money good. Each of WTP_{MG}, WTA_{MG}, EG_{MG}, $RWTP_{MG}$ and $RWTA_{MG}$ can be interpreted as a measure of the money value of one unit of the good, elicited by allowing the respondent to select an amount of money which in some sense is equivalent to a fixed amount of the good. Symmetrically, each of WTP_{GM}, WTA_{GM}, EG_{GM}, $RWTP_{GM}$ and $RWTA_{GM}$ can be interpreted as a measure of the value, in units of the good, of one unit of money, elicited by using the good as the response mode.

Our analysis of these valuation measures uses a restricted model of reference-dependent preferences under risk, incorporating elements of prospect theory (Kahneman and Tversky, 1979) and the rank-dependent reformulation of that theory (Starmer and Sugden, 1989; Tversky and Kahneman, 1992). As in prospect theory, reference-dependent preferences over consumption bundles are assumed to depend only on the displacement of each bundle from the reference state; thus, there are no income effects. We assume that these preferences can be represented by an additively separable *value function*. Thus, in the case of two goods, the value function can be written as $v(\Delta x)=v_1(\Delta x_1)+v_2(\Delta x_2)$, where $\Delta x=(\Delta x_1, \Delta x_2)$ is a *displacement vector* of changes in consumption relative to the reference state. Set $\Delta x_1>0$ and $\Delta x_2<0$, and define the *gain component* of Δx as $\Delta x^+=(\Delta x_1, 0)$ and the loss component as $\Delta x^-=(0, \Delta x_2)$. Whether the individual prefers this lottery to the reference state depends on the sign of $\pi(0.5)v(\Delta x^+)+\pi(0.5)v(\Delta x^-)$, where $\pi(\cdot)$ is the *probability weighting function*. Equivalently (cancelling out the $\pi(0.5)$ terms and using additive separability), this preference depends on the sign of $v(\Delta x)$. Thus, if the individual is indifferent between the reference state and the displacement vector Δx, she is also indifferent between the reference state and a *balanced lottery* which gives her a 0.5 chance of the gain component of Δx and a 0.5 chance of its loss component. This property of *balanced-lottery risk neutrality* allows balanced lotteries to be used to elicit loss aversion.[8]

Following Tversky and Kahneman (1992), we assume the following functional form for the value function: for each good i, $v_i(\Delta x_i)=a_i(\Delta x_i)^\beta$ if $\Delta x_i\geq 0$ and $v_i(\Delta x_i)=-b_i(-\Delta x_i)^\beta$ if $\Delta x_i\leq 0$, where a_i, b_i and β are constants satisfying $b_i>a_i>0$ and $1\geq\beta>0$. Given this functional form, the value of b_i/a_i is a natural measure of loss aversion in good i; β is a parameter which represents the extent of 'diminishing sensitivity' for gains and losses, lower values of β corresponding with stronger effects of diminishing sensitivity. It is convenient to define $L_i\equiv(b_i/a_i)^{1/\beta}$. For any given value of β, L_G and L_M may be treated as indices of the extent of loss aversion in the good and in money respectively (with a value of unity representing the absence of loss aversion).

Given specific assumptions about the locations of reference states, each valuation measure can be expressed in terms of the parameters a_G, b_G, a_M, b_M and β. Table 1

[8] The prediction of balanced-lottery risk neutrality is not unique to prospect theory. In particular, expected utility theory implies balanced-lottery risk neutrality if the utility function is separable in the two goods. However, some theories predict balanced-lottery risk aversion. That is, if an individual is indifferent between the reference state and some displacement vector Δx, she prefers the reference state to a lottery which gives her a 0.5 chance of the gain component of Δx and a 0.5 chance of its loss component. For example, rank-dependent expected utility theory (Quiggin, 1993) predicts balanced-lottery risk-aversion if the probability weighting function has the standard property $\pi(0.5)<0.5$.

Table 1
Implications of rival hypotheses for valuation measures

Measure	Value of measure in terms of parameters of model, given hypothesis	
	CEH	NLIB
$RWTA_{MG}$, $1/RWTP_{GM}$	$(b_G/a_M)^{1/\beta}$	$(b_G/a_M)^{1/\beta}$
WTA_{MG}, $1/WTP_{GM}$	$(b_G/a_M)^{1/\beta}$	$(b_G/a_M)^{1/\beta}$
EG_{MG}, $1/EG_{GM}$	$(a_G/a_M)^{1/\beta}$	$(a_G/a_M)^{1/\beta}$
WTP_{MG}, $1/WTA_{GM}$	$(a_G/b_M)^{1/\beta}$	$(a_G/a_M)^{1/\beta}$
$RWTP_{MG}$, $1/RWTP_{GM}$	$(a_G/b_M)^{1/\beta}$	$(a_G/b_M)^{1/\beta}$

displays these implications in relation to CEH and to NLIB. (Thus, for example, the first row of the table reports that CEH implies $RWTA_{MG}=1/RWTP_{GM}=(b_G/a_M)^{1/\beta}$, while NLIB implies $RWTA_{MG}=1/RWTP_{GM}=(b_G/a_M)^{1/\beta}$.) Notice that the two hypotheses differ only in the cases of WTP_{MG} and WTA_{GM}. These are the two cases in which the individual is considering giving up money in exchange for the good—that is, in which she is considering a *buying* transaction. Notice also that, for each valuation measure which uses money as the response mode, there is a corresponding measure which uses the good as the response mode; in each such pair, one measure is the reciprocal of the other.

If we are to use the entries in Table 1 to design an experiment to discriminate between the two hypotheses, we must consider relationships between two or more valuation measures. One possible approach is to use the design of the Vancouver experiment. This design elicits WTA_{MG}, EG_{MG} and WTP_{MG}. CEH and NLIB both predict $WTA_{MG}>EG_{MG}$; this inequality is induced by loss aversion in the good (since, on either hypothesis, $WTA_{MG}/EG_{MG}=L_G$). CEH also predicts $EG_{MG}>WTP_{MG}$, as a result of loss aversion in money (i.e. $EG_{MG}/WTP_{MG}=L_M$). In contrast, NLIB predicts $EG_{MG}=WTP_{MG}$.

However, if used in isolation, this approach to discriminating between CEH and NLIB has a significant limitation: it cannot measure the extent of loss aversion in money, independently of the hypotheses being tested. Suppose, for example, that (as in the Vancouver experiment), median values of WTA_{MG} are considerably greater than EG_{MG}, while median values of EG_{MG} are only slightly greater than WTP_{MG}. Is this evidence in favour of NLIB, or evidence that loss aversion in money is less than loss aversion in the good? This problem can be avoided by using 'risky' valuations. Notice that, according to both hypotheses, $EG_{MG}/RWTP_{MG}=L_M$. Thus, if the value of $EG_{MG}/WTP_{MG}$is close to 1 while $WTP_{MG}/RWTP_{MG}$ is markedly greater than 1, that is evidence in favour of NLIB. Conversely, if EG_{MG}/WTP_{MG} is markedly greater than 1 while $WTP_{MG}/RWTP_{MG}$ is close to 1, that is evidence in favour of CEH. If both values are close to 1, we can infer that there is little or no loss aversion in money, in which case it is not possible to discriminate between the two hypotheses. This method of measuring loss aversion in money depends on the assumption of balanced-lottery risk neutrality. However, that assumption can itself be tested, since $RWTA_{MG}/WTA_{MG}=1$ is a direct implication of balanced-lottery risk neutrality.

The foregoing argument gives us a possible experimental design, in which $RWTA_{MG}$, WTA_{MG}, EG_{MG}, WTP_{MG} and $RWTP_{MG}$ are elicited. The core test is whether EG_{MG} is greater than WTP_{MG}, as predicted by CEH, or whether the two are equal, as predicted by

NLIB.[9] By using $RWTP_{MG}$ in conjunction with the assumption of balanced-lottery risk neutrality, we estimate the extent of loss aversion in money. By comparing $RWTA_{MG}$ and WTA_{MG}, we test that assumption. This is the *money response design*.

However, it would be possible to preserve the basic structure of the design described in the preceding paragraph while using a mix of response modes. In particular, consider the *incoming valuation design*, which elicits $RWTA_{MG}$, WTA_{MG}, EG_{GM}, WTA_{GM} and $RWTA_{GM}$. In this design, the core test is whether WTA_{GM} is greater than EG_{GM}, as predicted by CEH, or whether the two are equal, as predicted by NLIB. L_M is measured by $RWTA_{GM}/EG_{GM}$; balanced-lottery risk aversion implies $RWTA_{MG}/WTA_{MG}=1$.

One advantage of this latter design is that the core test is more tightly controlled. From the viewpoint of Hicksian theory, WTA_{GM} and EG_{GM} are *identical* to one another (this is explained in footnote 6). In contrast, EG_{MG} and WTP_{MG} measure different entities within Hicksian theory; differences between these measures can in principle result from income and substitution effects. If Hicksian assumptions hold and if EG_{MG} is small relative to the individual's total wealth, credible values of the rate of change of WTP_{MG} with respect to wealth imply that the value of EG_{MG}/WTP_{MG} is close to 1 (Sugden, 1999). Nevertheless, one might still prefer more control to less.

A second advantage of this design is that $RWTA_{MG}$, WTA_{MG}, EG_{GM}, WTA_{GM} and $RWTA_{GM}$ are all *incoming* valuations. An incoming valuation records the smallest amount of something (money or the good) that the individual is willing to accept as a transfer *from* someone else. In contrast, WTP_{MG} and $RWTP_{MG}$ are *outgoing* valuations: they record the largest amount of something (in these cases, money) that an individual is willing *to* transfer to someone else. In experiments which elicit valuations, a potentially confounding factor is that subjects may follow *tactical heuristics* that lead them to understate their true outgoing valuations and to overstate their true incoming valuations. Although such heuristics do not in fact serve a subject's interests in incentive-compatible experiments such as those we discuss in this paper, they may be well-adapted to many real-world situations in which terms of trade are determined through bargaining. The confounding effects of tactical heuristics can be reduced by eliciting only incoming valuations.

The Norwich group favoured the incoming valuation design, for the reasons explained in the preceding two paragraphs. Kahneman favoured the more conventional money response design. He had reservations about eliciting valuations of fixed amounts of money in units of a consumption good: he was concerned that subjects might have difficulty in understanding tasks of this kind, or not construe them as buying tasks. Recognising the advantages and disadvantages of both proposals, and in the spirit of adversarial collaboration, we opted for a composite design that would elicit all of the valuations EG_{MG}, WTP_{MG}, WTA_{MG}, $RWTP_{MG}$, $RWTA_{MG}$, EG_{GM}, WTA_{GM} and $RWTA_{GM}$, providing data for the tests proposed by both parties.

[9] These core predictions are implications of reference-dependent theory as presented by Tversky and Kahneman (1991). The restricted form of reference-dependent theory is needed only to justify the interpretations given to risky valuation measures.

4. The experiment

The experiment was carried out at the University of East Anglia in Norwich. Subjects were recruited from the undergraduate population by general e-mailed invitations; they were broadly representative of that population in terms of age, gender, and subject of study. Subjects were required to bring cash to the experiment, but were assured that any opportunities to spend money would be optional. Initially, we recruited 320 subjects. Each subject was allocated at random to one of eight treatments, each of which elicited one of the valuations EG_{MG}, WTP_{MG}, WTA_{MG}, $RWTP_{MG}$, $RWTA_{MG}$, EG_{GM}, WTA_{GM} and $RWTA_{GM}$. Thus, each subject confronted just one valuation task.[10]

The specific good took the form of luxury chocolates sold by a specialist shop located in the centre of Norwich, easily accessible from the university campus. These chocolates are sold by weight, at an average price of about £0.30 each. To allow exchanges in units of single chocolates to be carried out conveniently,[11] transactions within the experimental sessions were carried out in vouchers. A voucher entitled its holder to a specified number of chocolates, free of charge, when presented at the shop.

Depending on which task they had been assigned, subjects were given 'endowments' (which in some cases were 'nothing'). Subjects in the WTA_{GM} and $RWTA_{GM}$ groups were given £1.00. Those in the WTA_{MG} and $RWTA_{MG}$ groups were given 10 chocolates (in the form of vouchers). All other subjects were given nothing. Endowments (money or vouchers) were physically handed over to subjects. It was explained that subjects' endowments were theirs to keep if they so chose. The conditions for the use of the vouchers were explained, and samples of the chocolates were shown; no information was given about the price of the chocolates.

Valuations were elicited by using *multiple dichotomous choice* tasks.[12] For example, in the EG_{MG} treatment, each subject made 25 choices; in each choice problem, one option was described as 'We [i.e. the experimenters] give you 10 chocolates', while the other took the form 'We give you £*x*', where *x* took the values £0.30, £0.60, £0.90, ..., £7.50 in successive problems. In the WTA_{GM} treatment, one option was always 'You do not trade', while the other took the form 'You give us your £1.00 and take *y* chocolates in exchange', where *y* took the values 1, ..., 25. The other treatments were presented analogously. The order in which the values of *x* or *y* were presented (that is, either ascending order or descending order) was randomised. We required each subject's choices to be mutually consistent in the sense of respecting dominance.[13] When the response mode is money, this

[10] In this respect, our design matched that of the Vancouver experiment. In the Norwich experiment, each subject was presented with a series of separate tasks, only one of which (selected at random at the end of the experiment) was for real.

[11] Such exchanges are necessary for the EG_{GM}, WTA_{GM} and $RWTA_{GM}$ tasks, which elicit valuations of money in units of chocolate. The chocolates are not individually wrapped, but are boxed to order at the shop.

[12] We used dichotomous choices rather than open-ended tasks (e.g. 'What is the largest amount of money you would be willing to pay for ... ?') to minimise the salience of tactical considerations. In this respect, our design followed the Vancouver experiment and not the Norwich one.

[13] If a subject's responses were mutually inconsistent, the nature of the inconsistency was explained to him, and he was asked to revise those responses. In fact, all but 3 of the 427 subjects responded consistently at the first attempt. Since all the theories we consider imply this form of consistency, requiring responses to satisfy it does not bias our tests.

procedure elicits money valuations of 10 chocolates within £0.30 bands. When the response mode is chocolate, it elicits valuations of £1.00 in units of chocolate.

Before any tasks were undertaken, it was explained to subjects that one of their 25 choices—to be identified at the end of the experiment by a random process—would be for real. When all tasks had been completed, the real task was selected separately for each subject, who then carried out whatever transaction (if any) he had chosen in that problem. In the case of risky valuations, lotteries were described in terms of a bag of 100 numbered discs; if a subject's chosen transaction involved playing out a lottery, he drew a disc from the bag to determine the outcome.[14]

Two additional treatments were run in a follow-up experiment. The responses to the treatments described above turned out to indicate surprisingly low levels of loss aversion for chocolate. Kahneman conjectured that this was the result of our having used vouchers rather than actual chocolates in the experimental sessions. The use of vouchers, he suggested, might attenuate loss aversion by mentally distancing subjects from the consumption experiences associated with the chocolates. In addition, since vouchers have some of the properties of money, the psychological mechanisms which (on his account) give rise to NLIB might also affect tasks in which vouchers are given up in trade. To test this conjecture, we ran a follow-up experiment which repeated the EG_{MG} and WTA_{MG} treatments, exactly as before except for one detail: the '10 chocolates' took the form of a pre-packed box of 10 chocolates (the same kind as we had used before). We recruited an additional 107 subjects and divided them at random between the two additional treatments. These *immediate chocolate* treatments are denoted by EG_{MG}* and WTA_{MG}*.

5. Results

The responses to the 10 treatments are summarised in Tables 2 and 3.[15] In presenting the data, we use the following conventions. Recall that for any given subject in any given treatment, there are 25 choice problems. Since subjects' choices are required to respect dominance, there are 26 alternative permissible ways of answering any such set of problems.

We assign these responses the values 1, . . ., 26, in ascending order of the valuation of 10 chocolates (for tasks in which the response mode is money) or in ascending order of the valuation of £1.00 (for tasks in which the response mode is chocolate). Thus, for tasks in which the response mode is money, 1 corresponds with the range of money valuations of 10 chocolates from 0 to £0.30; 2 corresponds with valuations from £0.30 to £0.60, and so on up to 26 which corresponds with valuations from £7.50 upwards. For tasks in which the response mode is chocolate, 1 corresponds with the range of chocolate valuations of £1.00 from 0 to 1 chocolate; 2 corresponds with valuations from 1 to 2 chocolates, and so on up to 26 which corresponds with valuations from 25 chocolates upwards.

We also report subjects' *implicit preferences* between 10 chocolates and £1.00. When the response mode is money, a subject whose valuation is 1, 2 or 3 has chosen to have

[14] The instructions given to subjects are available at Supplementary data.

[15] The distribution of responses for each treatment are provided at Supplementary data.

Table 2
Responses to tasks with money as the response mode

	Task (in=incoming, out=outgoing)						
	$RWTP_{MG}$ (out)	WTP_{MG} (out)	EG_{MG} (in)	WTA_{MG} (in)	$RWTA_{MG}$ (in)	EG_{MG}* (in)	WTA_{MG}* (in)
Valuations of 10 chocolates (units of £0.30)							
Geometric mean	5.38	4.66	8.17	9.95	10.17	8.69	11.30
Arithmetic mean	6.75	5.55	10.00	10.80	12.70	10.24	12.46
Median	5	6	10	10	12	10	14
Standard deviation	4.99	2.82	5.05	4.58	7.35	4.94	5.19
Implicit preferences							
No. of subjects who:							
Prefer £1	9	9	5	0	3	3	3
Not clear[a]	7	4	1	1	1	8	1
Prefer 10 chocolates	24	27	34	39	36	44	48
(% who prefer chocolates)	(60.0)	(67.5)	(85.0)	(97.5)	(92.3)	(80.0)	(92.3)
Total	40	40	40	40	40	55	52

[a] Subjects whose responses indicated that the valuation of 10 chocolates was at least £0.90 but no more than £1.20.

£0.90 rather than 10 chocolates, and so can be presumed to prefer £1.00 to 10 chocolates. Conversely, a subject whose valuation is 5 or more has chosen to have 10 chocolates rather than £1.20, and so can be presumed to prefer 10 chocolates to £1.00. When the response mode is chocolate, the valuations 1, . . ., 10 reveal an implicit preference for £1.00 over 10 chocolates, while the valuations 11, . . . , 26 reveal the opposite preference. Implicit preferences are of interest because they are comparable across all treatments, irrespective of whether the response mode is money or chocolate. Notice that {10 chocolates, £1.00} is the only pair for which our design allows us to identify implicit preferences for all treatments.

Table 3
Responses to tasks with chocolate as the response mode

	Task (in=incoming, out=outgoing)		
	$RWTA_{GM}$ (in)	WTA_{GM} (in)	EG_{GM} (in)
Valuations of £1.00 (units of one chocolate)			
Geometric mean	9.42	9.62	7.52
Arithmetic mean	12.70	10.95	8.85
Median	12.5	10	8
Standard deviation	7.75	5.70	4.84
Implicit preferences			
No. of subjects who:			
Prefer £1	24	17	12
Prefer 10 chocolates	16	23	28
(% who prefer chocolates)	(40.0)	(57.5)	(70.0)
Total	40	40	40

Table 2 refers to the seven treatments for which the response mode is money. The upper part of the table reports, for each treatment, the geometric mean, arithmetic mean, median and standard deviation of the distribution of subjects' valuations.[16] The lower part of Table 2 reports, for each treatment, the distribution of responses classified by implicit preferences. Table 3 presents the corresponding data for the three treatments for which the response mode was chocolate.

Table 4 reports our tests. Each row compares two different valuation measures for a given good (immediate chocolates in row iii, chocolate vouchers in the others). The first entry in each row specifies a ratio of two valuations. The next two entries present the predicted value of this ratio for any given individual, according to each of CEH and NLIB. The following three entries indicate whether particular causal factors, if operating in the experiment as a whole, will impact on that ratio; if there is such an impact, its effect is to increase the value of that ratio above 1. These causal factors are, respectively, balanced-lottery risk aversion,[17] subjects' use of tactical heuristics, and Hicksian income and substitution effects. (In the case of tactical heuristics, 'yes' signifies that the numerator of the ratio is an incoming valuation and that the denominator is an outgoing valuation; 'no' signifies that both valuations are of the same type. In the case of Hicksian effects, firm predictions can be made only for choice under certainty. A 'no' entry in the 'Hicksian effects' column implies that Hicksian theory makes the firm prediction that the value of the relevant ratio is 1.) The sixth and seventh entries report, respectively, the ratio between the medians of the relevant valuations and the ratio between the arithmetic means. The final entry reports the t-statistic for a test of difference between arithmetic means of the valuations in the two treatments.

We use ratios between median valuations and ratios between mean valuations as summary statistics which can be compared with the corresponding predictions. For example, CEH and NLIB both predict $WTA_{MG}/EG_{MG}=L_G$ for any given individual (row ii). Thus, each of these hypotheses allows us to interpret a ratio between medians or means of WTA_{MG} and EG_{MG} as indicative of the value of L_G in the population. In this context, ratios between medians have a more precise interpretation: on the assumption that the ranking of individuals by their relative valuations of chocolate and money is constant across valuation measures and response modes, the ratio of median valuations from two treatments is an estimate of the ratio of the corresponding valuations for the median individual in the subject pool. However, because of the lumpiness of the valuation scales, one must be careful not to over-interpret small differences in ratios of medians.

To provide a benchmark for other comparisons, we begin by comparing WTA_{MG} and WTP_{MG}. Many experiments and surveys have found WTA to be greater than WTP in comparisons of this kind: we shall call this the classic WTA/WTP comparison. Row i of the table shows that this familiar result is replicated: the ratio of medians is 2.4 (1.95 for means) and the difference between means is overwhelmingly significant. This result is not

[16] Mean valuations (whether arithmetic or geometric) are sensitive to extreme values. However, subjects rarely used the extremes of the response scales. Of the 307 subjects using the money response mode, only 13 recorded the lowest valuation 1 and only 9 recorded the highest valuation 26. For the 120 subjects using the chocolate response mode, the corresponding numbers were 6 and 5.

[17] This concept is defined in footnote 8.

Table 4
Comparisons of valuations

Ratio	CEH predicts	NLIB predicts	Picks up balanced-lottery risk aversion	Picks up tactical heuristics	Picks up Hicksian effects	Ratio of medians	Ratio of means	t-test of difference between means
(i) WTA_{MG}/WTP_{MG}	$L_M L_G$	L_G	no	yes	yes	2.40	1.95	6.17**
(ii) WTA_{MG}/EG_{MG}	L_G	L_G	no	no	no	1.25	1.08	0.74
(iii) $WTA_{MG}*/EG_{MG}*$	L_G	L_G	no	no	no	1.40	1.22	2.28**
(iv) $RWTA_{MG}/WTA_{MG}$	1	1	yes	no	n.a.	1.25	1.18	1.39
(v) EG_{MG}/WTP_{MG}	L_M	1	no	yes	yes	1.67	1.80	4.86**
(vi) $WTP_{MG}/RWTP_{MG}$	1	L_M	yes	no	n.a.	1.20	0.82	–1.40
(vii) WTA_{GM}/EG_{GM}	L_M	1	no	no	no	1.25	1.24	1.76*
(viii) $RWTA_{GM}/WTA_{GM}$	1	L_M	yes	no	n.a.	1.25	1.16	1.15

$L_M \equiv (b_M/a_M)^{1/\beta} > 1$, $L_G \equiv (b_G/a_G)^{1/\beta} > 1$.
In the final column, * denotes that the numerator of the ratio is significantly greater than the denominator (5% significance level in a one-tail test); ** denotes significance at the 1% level.
n.a.=not applicable.

surprising, but it gives some assurance that our experiment is picking up whatever causal factors lie behind commonly observed differences between WTA and WTP.

To provide a further benchmark, we consider the extent of loss aversion in chocolate. Both CEH and NLIB imply $WTA_{MG}/EG_{MG}=L_G$. In the case of chocolate vouchers (row ii), the ratio of the medians of WTA_{MG} and EG_{MG} is 1.25 (1.08 for means), and there is no significant difference between the means. In the case of immediate chocolate (row iii), the ratio of medians is 1.40 (1.22 for means) and, given the larger sample size, the difference between means is significant at the 1% level. The latter result confirms that there is loss aversion in chocolate, as predicted by reference-dependent preference theory. That the value of this ratio is greater for immediate chocolates than for vouchers is consistent with Kahneman's conjecture that loss aversion for a good is weakened when that good is represented by vouchers.[18] Notice, however, that this conjecture has no implications for the validity of the tests which discriminate between CEH and NLIB, since these are concerned with loss aversion in money.

Recall that our analysis of risky valuations depends on the assumption of balanced-lottery risk neutrality, which can be tested by comparing $RWTA_{MG}$ and WTA_{MG} (row iv). Consistently with that assumption, we find no significant difference between these two valuations.

We now consider the money-response tests which discriminate between CEH and NLIB (rows v and vi). CEH predicts $EG_{MG}/WTP_{MG}=L_M>1$, while NLIB predicts $EG_{MG}/WTP_{MG}=1$. In fact, the ratio of medians is 1.67 (1.80 for means), and the difference in means is significant at the 1% level. CEH predicts $WTP_{MG}/RWTP_{MG}=1$, while NLIB predicts $WTP_{MG}/RWTP_{MG}=L_M>1$. The ratio of medians is 1.20 (0.82 for means); the

[18] One might expect immediate chocolates to be preferred to chocolate vouchers because of the transaction costs involved in using vouchers. This effect will tend to make valuations higher for immediate chocolates than for vouchers, but we see no reason to expect transaction costs to affect the relative values of WTA_{MG} and EG_{MG}.

difference in means is in the opposite direction to the NLIB prediction. Using ANOVA, we can make more powerful joint tests of the two hypotheses with respect to EG_{MG}, WTP_{MG} and $RWTP_{MG}$ taken together. The null hypothesis that all three means are equal can be rejected at the 1% level in an F-test. In a contrast test in which the alternative hypothesis is $EG_{MG}>WTP_{MG}=RWTP_{MG}$, as predicted by CEH, the null hypothesis of no differences between means can be rejected at the 1% level (t=4.19). If the alternative hypothesis is $EG_{MG}=WTP_{MG}>RWTP_{MG}$, as predicted by NLIB, the null hypothesis cannot be rejected (t=1.12). Clearly, the money-response tests favour CEH.

Finally, we consider the incoming-valuation tests (rows vii and viii). CEH predicts $WTA_{GM}/EG_{GM}=L_M>1$, while NLIB predicts $WTA_{GM}/EG_{GM}=1$. The ratio of medians is 1.25 (1.24 for means); the difference in means is significant at the 5% level. CEH predicts $RWTA_{GM}/WTA_{GM}=1$, while NLIB predicts $RWTA_{GM}/WTA_{GM}=L_M>1$. The ratio of medians is 1.25 (1.16 for means); the difference in means is not significant. The null hypothesis that all three means are equal can be rejected at the 5% level. In a contrast test in which the alternative hypothesis is $RWTA_{GM}=WTA_{GM}>EG_{GM}$, as predicted by CEH, the null hypothesis of no differences between means can be rejected at the 1% level (t=2.76). If the alternative hypothesis is $RWTA_{GM}>WTA_{GM}=EG_{GM}$, as predicted by NLIB, the null hypothesis can be rejected at the 5% level (t=2.06). On balance, the incoming-valuation tests favour CEH, but far from decisively.

6. Discussion

A surprising feature of our results is the relative weakness of loss aversion effects in all but one of the comparisons which, according to either or both of the parties to the collaboration, could be used to estimate the value of L_G or L_M in the subject pool. With the exception of the comparison between EG_{MG} and WTP_{MG}, all the tests legitimated by CEH imply that typical values of both L_G and L_M (whether inferred from medians or means) are in the range 1.08–1.40. Where NLIB and CEH differ, the tests legitimated by NLIB imply lower values of L_M. Kahneman's prior expectation was that L_G and L_M would take values close to 2.[19] Nevertheless, in the classic WTA/WTP comparison, we found a WTA_{MG}/WTP_{MG} ratio close to or greater than 2 (depending on whether the ratio is defined in terms of means or medians). This is in line both with Kahneman's expectation and with the results of other experiments. As explanations of these features of our data, we offer two conjectures, one of which is favoured by the Norwich group, the other by Kahneman.

The Norwich group interprets the results as consistent with CEH. Putting most weight on the tests that are not liable to be confounded by tactical heuristics, by income and substitution effects, or by the use of vouchers (that is, the tests reported in rows ii and vii of Table 4), the Norwich group concludes that loss aversion in chocolate and loss aversion

[19] Using data from an experiment in which subjects reported certainty equivalents for lotteries with money consequences, Tversky and Kahneman (1992) fit a model similar to that presented in Section 3, separately for each subject. The median value of b_M/a_M is 2.25; that of β is 0.88. From these data, one might expect typical values of L_M to be in the region of $2.25^{0.88}$=2.04. The results of the Vancouver experiment imply values of WTA_{MG}/EG_{MG}, i.e. of L_G, close to 2.

in money are real but relatively weak effects. Both of these effects contribute to the classic WTA/WTP disparity, but a third factor may be at work too: subjects may be using tactical heuristics. On this hypothesis, we should expect relatively chocolate-loving preferences to be revealed in those money-response tasks that elicit incoming valuations (i.e. EG_{MG}, WTA_{MG}, $RWTA_{MG}$, EG_{MG}* and WTA_{MG}*). Conversely, we should expect relatively money-loving preferences to be revealed both in the chocolate-response tasks (i.e. $RWTA_{GM}$, WTA_{GM} and EG_{GM}), since those tasks elicit incoming valuations in units of chocolate, and in those money-response tasks that elicit outgoing valuations (i.e. $RWTP_{MG}$ and WTP_{MG}). Implicit preferences do in fact show this general pattern (see Tables 2 and 3). This hypothesis implies that EG_{MG}/WTP_{MG} may overstate the true value of L_M. This might account for the marked difference between the money-response and incoming-valuation CEH estimates of L_M.

Kahneman, too, interprets the results of the joint experiment as failing to support NLIB. However, he interprets the observed difference between EG_{MG} and WTP_{MG} as evidence of strong loss aversion in money. For Kahenman, the puzzle is to explain why loss aversion in money shows up in this comparison, and not (as NLIB would predict) in the comparison between WTP_{MG} and $RWTP_{MG}$. His tentative interpretation is that, contrary to his prior expectation, subjects treated money given up in return for chocolates as a loss. This would be compatible with his general theory of the conditions under which gains and losses are integrated *if* the Norwich subjects were so financially constrained that they did not perceive themselves as having budget reserves. Kahneman conjectures that this may have been the case, and that this may amount to an unanticipated difference between the Norwich subject pool and the North American subject pools that he has used previously.

If we look beyond the results of this particular experiment, there is a more general problem: to find a unified explanation for *all* the data that have so far been generated within this experimental paradigm. Specifically, the experiment reported in this paper, the Vancouver and (original) Norwich experiments, and a set of additional experiments[20] reported by Kahneman and Novemsky (2002) together provide a very large body of data, which one might hope to be able to organise into a consistent pattern. But finding such a pattern is not easy. The results of our collaborative experiment are similar to those of the Norwich experiment, despite the different incentive mechanisms and elicitation procedures used in the two cases. Specifically: in each of these experiments, the results are consistent with CEH; loss aversion, both in money and in goods, exists but is relatively weak; and incoming valuations are markedly greater than outgoing valuations. However, the results of the Vancouver experiment and of the experiments reported by Kahneman and Novemsky show a different overall pattern. These results are consistent with NLIB; they indicate relatively strong loss aversion in both money and goods; and they show no

[20] Between 1986 and 1991, Kahneman ran a series of experiments which elicited various combinations of $RWTP_{MG}$, WTP_{MG}, EG_{MG}, WTA_{MG}, and $RWTA_{MG}$. At the time the joint experiment was designed, he believed that the results of these earlier experiments had been lost in a fire which destroyed his home in 1991. However, after the first stage of our joint experiment had been run, these data were rediscovered. These experiments are reported in the paper cited. Averaging across the experiments they report, Kahneman and Novemsky's computations of the ratios of median valuations are: $RWTA_{MG}/WTA_{MG}$=0.91, WTA_{MG}/EG_{MG}=1.80, EG_{MG}/WTP_{MG}=1.08, $WTP_{MG}/RWTP_{MG}$=2.23.

evidence of a systematic difference between incoming and outgoing valuations. Even within the North American data there is a considerable degree of variation between experiments—variation that cannot be explained merely as the result of random factors, given the hypothesis of a common North American subject pool.

One may also ask how the results of our experiment relate to other research on WTA/WTP disparities. In particular, several studies have found that the size of the classic WTA/WTP disparity tends to fall as individuals gain experience of trading the relevant goods, or of trading in the relevant environment. Typically, the effect of experience is to reduce WTA and to increase WTP; decreases in WTA tend to be greater in absolute terms than increases in WTP. For example, List (2003) elicits valuations from participants in organised sportscard markets; he finds that WTA/WTP disparities are smaller for those individuals who have had greater experience of trading in those markets; experience has a strong negative effect on WTA and a much weaker positive effect on WTP. While these results are not directly comparable with those reported in this paper, they suggest that the truth may lie somewhere between CEH and NLIB. If the effect of experience is to reduce the anticipated pain of moving away from a reference state, CEH implies that the effect will operate symmetrically on WTA and WTP, reducing the former and increasing the latter, while NLIB implies that it will operate only on WTA.

Summing up these substantive conclusions: the results of our experiment, considered in isolation, support CEH rather than NLIB. However, it seems that we can identify a range of putative effects—loss aversion in money as mediated by CEH, loss aversion in money as mediated by NLIB, loss aversion in goods, tactical heuristics—none of which is wholly robust in isolation, but each of which may sometimes contribute to the classic and highly reliable WTA/WTP disparity.

In terms of scientific method, we believe that our work has demonstrated the value of adversarial collaboration in experimental economics. We have been able to design an experiment which the proponents of two competing hypotheses could accept as a method of discriminating between those hypotheses, and we have been able to agree on the interpretation of the data generated by that experiment. Even though we have not reached a full consensus about the broader theoretical issues involved, we have gone a long way in narrowing down the areas of disagreement. In retrospect, the disciplines imposed by the collective activities of designing the experiment and of writing up the results were perhaps as important as the actual experimental results in generating this convergence and in clarifying the issues in dispute. For example, Kahneman's general hypothesis about the conditions under which the gains and losses associated with a transaction are 'integrated' was developed as a by-product of the collaboration: although the outline of this idea pre-existed our joint work, our attempts to design an agreed test of NLIB revealed the need for a more precise formulation. We recommend adversarial collaboration to other experimental researchers as a constructive way of resolving conflicts between rival hypotheses.

Acknowledgements

We are grateful to Colin Camerer, Jack Knetsch, the guest editors and three anonymous referees for comments on earlier versions of this paper. This research was carried out as part of the Programme in Environmental Decision Making, organised through the Centre for Social and Economic Research on the Global Environment, and supported by the Economic and Social Research Council of the UK (award no. M 535 25 5117). Robert Sugden's work was also supported by the Leverhulme Trust.

References

Arrow, Kenneth, Solow, Robert, Portney, Paul, Leamer, Edward, Radner, Roy, Schuman, Howard, 1993. Report of the NOAA Panel on Contingent Valuation. Report to the General Counsel of the US National Oceanic and Atmospheric Administration. Resources for the Future, Washington, DC.

Bateman, Ian, Willis, Ken (Eds.), 1999. Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries. Oxford University Press, Oxford.

Bateman, Ian, Munro, Alistair, Rhodes, Bruce, Starmer, Chris, Sugden, Robert, 1997. A test of the theory of reference-dependent preferences. Quarterly Journal of Economics 112, 479–505.

Bateman, Ian, Langford, Ian, Munro, Alistair, Starmer, Chris, Sudgen, Robert, 2000. Estimating four Hicksian welfare measures for a public good: a contingent valuation investigation. Land Economics 76, 355–373.

Benartzi, Shlomo, Thaler, Richard H., 1995. Myopic loss aversion and the equity premium puzzle. Quarterly Journal of Economics 110, 73–92.

Carson, Richard, Mitchell, Robert, 1993. The value of clean water: the public's willingness to pay for boatable, fishable, and swimmable water. Water Resources Research 29, 2445–2454.

Cummings, Ronald, Taylor, Laura, 1999. Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. American Economic Review 89, 649–665.

Kahneman, Daniel, Nathan, Novemsky, (2002). Loss aversion in risky and riskless transactions: choices, exchanges and gambles. Mimeo, Princeton University.

Kahneman, Daniel, Tversky, Amos, 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263–291.

Kahneman, Daniel, Knetsch, Jack L., Thaler, Richard H., 1990. Experimental tests of the endowment effect and the Coase theorem. Journal of Political Economy 98, 1325–1348.

Knetsch, Jack, 2000. Environmental valuations and standard theory: behavioural findings, context dependence and implications. In: Tietenberg, Tom, Folmer, Henk (Eds.), The International Yearbook of Environmental and Resource Economics 2000/2001: A Survey of Current Issues. Edward Elgar, Cheltenham.

List, John, 2003. Does market experience eliminate market anomalies? Quarterly Journal of Economics 118, 41–71.

Mellers, Barbara, Hertwig, Ralph, Kahneman, Daniel, 2001. Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. Psychological Science 12, 269–275.

Munro, Alistair, Sugden, Robert, 2003. On the theory of reference-dependent preferences. Journal of Economic Behavior and Organization 50, 407–428.

Myagkov, Mikhail, Plott, Charles R., 1997. Exchange economies and loss exposure: experiments exploring prospect theory and competitive equilibria in market environments. American Economic Review 87, 801–828.

Quiggin, John, 1993. Generalised Expected Utility Theory: the Rank-Dependent Model. Kluwer, Boston, MA.

Rowe, Robert, d'Arge, Ralph, Brookshire, David, 1980. An experiment on the economic value of visibility. Journal of Environmental Economics and Management 7, 1–19.

Samuelson, William, Zeckhauser, Richard, 1988. Status quo bias in decision making. Journal of Risk and Uncertainty 1, 7–59.

Starmer, Chris, 2000. Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. Journal of Economic Literature 38, 332–382.

Starmer, Chris, Sugden, Robert, 1989. Violations of the independence axiom in common ratio problems: an experimental test of some competing hypotheses. Annals of Operations Research 19, 79–102.

Sugden, Robert, 1999. Alternatives to the neo-classical theory of choice. In: Bateman and Willis (1999).

Thaler, Richard, 1980. Toward a positive theory of consumer choice. Journal of Economic Behavior and Organization 1, 39–60.

Tversky, Amos, Kahneman, Daniel, 1991. Loss aversion in riskless choice: a reference-dependent model. Quarterly Journal of Economics 106, 1039–1061.

Tversky, Amos, Kahneman, Daniel, 1992. Advances in prospect theory: cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 232–297.