23rd Australasian Transport Research Forum
Perth, Western Australia 29 September – 1 October 1999

The Willingness to Accept Value of Statistical Life Relative to the Willingness to Pay Value: Evidence and Policy Implications

Jagadish Guria, Joanne Leung,
Land Transport Safety Authority, New Zealand
Michael Jones-Lee, Graham Loomes
University of Newcastle upon Tyne, United Kingdom

Abstract

Large disparities between willingness to accept and willingness to pay based values of statistical life are commonly encountered in empirical studies. Theoretically, economists consider that if a public good is easily substitutable, then there should be no marked disparity between willingness to accept and pay values for the good, though the disparity increases with lack of substitutability. Psychologists, on the other hand, argue that people generally treat gains and losses asymmetrically and tend to require a substantially larger increase in wealth to compensate for a loss than the amount they would be willing to pay for an equivalent gain. Although most transport projects may aim to improve safety, situations arise when a relaxation of an existing regulation saves resources but increases the risk of death and injuries. A survey was recently carried out in New Zealand to determine people's willingness to pay to reduce road risks and their willingness to accept compensation for an increase in risk. This paper discusses the relationship observed between the two sets of responses

Contact Author

Jagadish Guria
Land Transport Safety Authority
PO Box 2840
Wellington
New Zealand

Phone: +64 4 494 8718 Fax: +64 4 494 8601
e-mail: jcg@ltsa.govt.nz

Introduction

When evaluating transport projects which entail expected reductions in the risks of death and injury, a number of countries now use a *value of statistical life* (VOSL)[1] to enable the safety benefits to be weighed against the costs of providing them. Ideally, this VOSL should reflect the collective *willingness to pay* (WTP) of members of the population for the relevant improvements in their own (and possibly others') safety.

Sometimes, however, situations arise when some existing regulation or piece of legislation is considered too restrictive and the possibility of relaxing it is explored. If such relaxation entails some increase in the risks of death and injury – in return, perhaps, for some saving of time and/or money – it is appropriate to value that change in risks on the basis of people's collective *willingness to accept* (WTA) compensation.

As discussed in the next section, under plausible assumptions standard economic theory suggests that the WTA figure should be, at most, only a few percentage points higher than the WTP figure. Yet many empirical studies in a variety of contexts have found that WTA is liable to greatly exceed WTP. Depending on the reason(s) for it, such a disparity may have important policy implications. To date, however, these implications appear to have received little attention, so that a major aim of this paper is to stimulate debate about the issues involved – especially since a recent value of safety (VOS) study conducted in New Zealand suggests a very substantial difference between WTA and WTP-based values, in the light of which, active consideration is being given to the possibility of explicitly using different values to evaluate projects, depending on whether those projects involve increases or decreases in the risks of death or injury on the roads.

The paper has five sections. The next section discusses various possible explanations of the disparity between WTA and WTP. The third section discusses the New Zealand VOS survey results. Policy implications are discussed in the fourth section and conclusions are drawn in the fifth.

WTA-WTP disparity: theoretical explanations

For the reader interested in a more detailed – but very accessible – discussion of the standard economic model and the main accounts of why the evidence may diverge from it, we would recommend Sugden (1999). This section takes the same general approach as that paper, focusing on the key points involved.

Standard economic models picture the typical individual as deriving welfare from consuming a bundle of goods and services in such a way that the loss of welfare from a small reduction in the amount of one good consumed can generally be offset by some upward adjustment in the quantity of one or more of the other goods consumed. Thus, if there is some small reduction in the good "physical safety" – as a result, say, of a

[1] The term value of preventing a (statistical) fatality (VPF) is increasingly being used instead of VOSL: the concept is the same, but the term VPF may convey it better

marginal increase in the risk of being injured on the roads – the standard assumption is that an individual can offset that loss by consuming more of other things that give utility; and the minimum amount of money necessary to buy the extra goods necessary to restore the individual to his/her level of welfare prior to the change in risk is that person's minimum willingness-to-accept (WTA) figure.

Equally, if there is some development which could reduce road risks and give the individual a little more utility from that good, we can imagine asking how much consumption of other goods she would be prepared to forego in order to obtain that benefit. The conventional assumption is that she would give up other consumption at the margin – and hence free up some money – up to the point where she returns to the same level of welfare she had before the safety improvement became available. This amount of money, then, is the person's maximum willingness-to-pay (WTP) response.

Under the standard assumption that all goods are to some extent substitutable for each other as sources of utility, and that the kinds of WTP and WTA amounts that would be appropriate for marginal changes in risk would constitute only a relatively small fraction of the typical individual's wealth, it would be surprising to find that the WTA and WTP figures diverged very much.

And yet many surveys and experiments have found the WTA responses to be substantially higher than the WTP responses, even for fairly familiar goods, such as chocolates, ball-point pens and coffee mugs (see Kahneman *et al.*, 1990 and Bateman *et al.*, 1997 for examples of the latter). As discussed below, not *all* of the evidence points in this same direction; but the disparity has been observed so often by many different researchers in such a variety of different settings that it clearly requires some explanation.

Types of explanation

Broadly speaking, four types of explanation have been offered, namely:

1) The conventional model is generally a good enough approximation, but for some goods the degree of substitutability is low, and this produces rather larger WTA-WTP disparities than would normally be expected.

2) The conventional model is not a very good approximation; rather, people's preferences are structured in such a way that they treat gains very differently from losses, and it is this feature which is reflected in the WTA-WTP disparities so often observed.

3) *At root*, people's preferences are more or less in line with the conventional model, but the kinds of questions presented to them in surveys are unfamiliar – perhaps even a little surreal – and they have little opportunity to reflect deeply enough to home in on their 'core' preferences; instead, they use inappropriate rules of thumb to answer the questions and/or fall prey to various cues and biases which tend to produce WTA and WTP responses which may not (both) reflect their 'true' underlying values

4) People simply don't have some core of comprehensive and highly coherent 'true' preferences – at least, not for most goods other than those which they consume frequently enough to enable them to evolve preferences by a process of repeated trial and adjustment. Thus when presented with unfamiliar goods in an unfamiliar question format, they have to construct their responses, perhaps combining some fairly general basic values with simple rules of thumb, with the choice of rule of thumb being liable to be influenced by cues provided by the subject matter of the questions and/or the way they are 'framed'.

Clearly, if either of explanations (1) or (2) is correct in the context of road safety, there is a *prima facie* case for reflecting people's preferences by using different values in road project appraisal (although this is not to say that such a strategy will not present some problems for public policy). If explanation (3) is correct, the implication would seem to be that further work needs to be done to develop survey instruments better able to elicit people's underlying preferences. Whereas if explanation (4) is correct, it is not immediately apparent what policy conclusions to draw – although we shall in due course discuss at least a couple of the possibilities.

Degree of substitutability

An influential paper in support of explanation (1) was produced by Hanemann (1991). He argued that the divergence between WTA and WTP values can be explained by the income elasticity of the inverse compensated demand function $\xi \equiv \frac{\eta}{\sigma_0}$ where η is the income elasticity of demand and σ_0 is the elasticity of substitution between the good being valued and all other goods. On the basis of this expression, Hanemann concluded that if the good in question is highly substitutable – i.e. σ_0 is relatively large – the divergence between WTP and WTA tends to disappear; whereas if the good has no close substitute – i.e. σ_0 tends towards 0 – then the divergence can be very substantial.

This hypothesis appears to be supported by an empirical study by Shogren *et al.* (1994). In this study, the divergence between WTP and WTA for goods with relatively close substitutes (in this case, a candy bar) disappeared with repeated exposure to market experience and feedback. "In contrast, for a private non-market good with no close substitute (reduction in human health risk) the divergence is robust and persistent, even given repeated market participation and full information on the characteristics of the good" (p256).

Against this, two points should be borne in mind. First, as mentioned earlier, there are other experiments where the disparities do not disappear, even for quite familiar goods (and it is possible that the Shogren *et al.* data could just as well be interpreted in terms of rules of thumb being used for valuing unfamiliar goods). Second, for Hanemann's conjecture to account for the kinds of disparities observed, it would require phenomenally implausible assumptions about the relevant elasticities (see Sugden, 1999, for examples), whereas casual observation suggests that even health and safety are not *such* special goods: in the context of transport safety, for example, most people

know that it is riskier to travel by car than by public transport, but seem willing to accept the higher risks in return for some combination of convenience, comfort, time-saving, etc.

Gains vs. losses

Part of the reason why Hanemann's conjecture seems so implausible is that it is trying to reconcile large disparities within a model where changes from the *status quo* are evaluated in the same way whether they are regarded as improvements/gains or as deteriorations/losses. However, an alternative approach is adopted by explanation (2), which models people's preferences differently; and in particular, allows that people may treat perceived gains and losses (relative to the *status quo* or some other salient *reference point*) in an asymmetrical fashion. Although this idea has a longer history, one fairly recent formulation is Tversky and Kahneman's (1991) *reference-dependent theory*.

Put simply, this model suggests that if some benefit – such as an x% reduction in the risk of being killed or injured on the roads – is offered to people, they will evaluate it as a gain, and express some willingness to pay for it. But if presented with some prospective *dis*benefit of the same magnitude – i.e. an *increase* of x% in their road risk – they will regard it as a loss, and the *loss aversion* entailed by the reference-dependent model will cause them to require a substantially larger amount of compensation than the amount they would be willing to pay for the corresponding gain. Moreover, once the *status quo* has changed, people may evaluate further increases or decreases relative to their new reference point, so that some WTA-WTP disparity is liable to occur at any and every point to which the respondent has become adjusted.

While this model appears to be much more compatible with the evidence (and with many people's intuitions about their own reactions to perceived gains and losses), it is not entirely problem-free. For example, this model would seem to require that even though people are assumed to have ready access to more-or-less complete and well-defined preferences, they somehow persistently treat money differently from other goods, in the sense that they exhibit less loss aversion towards money. After all, when responding to a WTP question, they are being asked to weigh some gain of a particular good like safety against the loss of the money it might cost (or, at a more fundamental level, the loss of whatever units of current consumption they would have to forego to release that money); whereas when responding to a WTA question, it is the other way round.

Strategic bias

To observe WTA substantially exceeding WTP across virtually all goods that have ever been the subject of such questions would require either that those goods happen to be ones involving greater loss aversion than the ones being foregone, or else that using the

medium of money somehow produces that effect. We shall return to this point below, when discussing possible implications for policy.

But first, let us consider explanations (3) and (4), which suggest that whether or not people have a 'core' of standard or non-standard preferences, their responses to survey questions may be susceptible to various sources of bias and distortion, and that it is this susceptibility which is largely responsible for the WTA-WTP disparity.

One possible source of such responses is what might be called *strategic bias*. Suppose an individual is rather hazy about his preferences, and considers that a particular change in risk is equivalent to some amount of money, but that this amount could be anything between $200 and $500. When asked how much he would be willing to pay for some given risk reduction, he may wish to err on the side of caution, or may perceive himself to be in something like the first round of a process of bargaining, and may therefore tend to pick a response from the lower end of the band of possible values; whereas when asked how much he would be willing to accept for an increase in risk of the same magnitude, erring on the side of caution, or behaving as if he were opening some negotiation, would cause him to select a value from the high end of the range.

There is good evidence that people *are* rather hazy about their values for such goods as road safety – see Dubourg *et al.* (1997) – and there is also some evidence that questions framed in terms of buying and selling may stimulate people to think in strategic terms (Burton *et al.*, 1999); so this kind of explanation may seem plausible. It would also fit with the notion that if people can be given more experience with buying and selling under conditions which provide incentives for truthful revelation and which give opportunities to refine their responses, the WTP/WTA disparity will tend to reduce or even disappear. And there is some evidence of this: for example, Coursey *et al.* (1987) organised an experimental auction where participants were asked about their willingness to accept compensation for consuming a foul-tasting (but harmless) substance, and their willingness to pay to avoid having to consume the substance, and found that although the initial WTA-WTP disparity was comparable to the findings in many surveys, the disparity greatly reduced (mostly as a result of falling WTA) with repeated rounds of the auction. On the other hand, Kahneman *et al.*'s (1990) auctions showed no such striking tendency for the disparity to shrink to an insignificant level, so the existing evidence on this issue is not conclusive.

Insufficient opportunity

Another possible source of any disparity may arise from people using their responses to convey a 'message'. For example, in the case of transport safety policy, some people may react against the idea of government agencies (who 'ought' to be looking after people's safety) exposing them and others to higher risks, and try to convey their disapproval by saying that "there is no amount of money" that would be acceptable. While on the other hand, they may feel that they already pay enough taxes, or existing tax revenues are wasted or misdirected, so that if they perceive a WTP question to be soliciting agreement to higher taxes, they may react against this by saying that they

would pay little or nothing. Moreover, extreme cases of such "protest votes" – i.e. infinitely high WTA and zero WTP – may only be the tip of the iceberg, in that some element of protest may influence other finite responses. However, although this could be an important issue in some surveys, it does not appear to be a strong candidate for explaining the disparity in cases where the question is about a private good, especially one with little 'moral' significance, such as a coffee mug or a ball-point pen.

For some more complex goods – and road safety might well count as one of these – it is possible to imagine other sources of disparate responses. If survey respondents have insufficient opportunity to give questions the consideration necessary to reach a balanced judgement, they may be liable to process information partially, and in ways which differ between WTA and WTP questions. For example, when asked what they would be willing to pay for a small reduction in an already small risk, respondents' attention may tend to be focused on the high likelihood that they won't have an accident in any case, and this may exert a downward influence on their responses. However, when asked what they would need to be paid in order to accept a change which increases their risk, attention may be switched to thinking about how they would feel if they did have such an accident and if this might have been attributable to accepting an increase in risk in return for some money – a line of thought which is liable to exert an upward influence on their response. Notice that such an explanation could account for the evidence produced by Shogren *et al.* (1994), where the WTA-WTP disparity persisted in the case of a complex good (involving a very small risk of food poisoning from consuming an unscreened sandwich) while it effectively disappeared in the case of a simple good (a risk-free candy bar).

All in all, then, the WTA-WTP disparity has been, and continues to be, the subject of both theoretical and empirical debate. In the next section, we shall describe a new study which was influenced by the debate and which adds some further evidence to the existing stock. We shall then return to the various explanations considered above, and examine the possible implications of the new data for transport policy in New Zealand – and perhaps elsewhere.

The New Zealand study

A WTP-based Value of Statistical Life (VOSL) has been used in New Zealand since 1991. This was based on a contingent valuation survey carried out during 1989/90 (Miller and Guria 1991). As some policy changes have the potential to increase risk of death or injury from its base level, the Land Transport Safety Authority of New Zealand, a body responsible for promoting safety in land transport at reasonable cost, considered it necessary to update the WTP-based VOSL and also to estimate the WTA-based VOSL and values for prevention of non-fatal injuries. Thus a new survey was designed and conducted in 1997/98 to estimate a fresh set of both WTP- and WTA-based values (Guria et al 1999).

The questionnaire

In addition to fatal injury, the survey questionnaire defined three types of non-fatal injuries: minor, temporary and permanent. The minor injury was defined as bruises, sprains or cuts not needing stitches or hospitalisation, but involving a day or two of reduced activities, with full recovery within 10 - 14 days in most cases. The temporary injury was defined as more serious but with full recovery within 2-4 months. The permanent injury was defined as having some long term impairments, with the possibility of very severe impairments in some cases.

During the early stages of the interview, respondents were given some information about average risks in their locality, and were asked to think through how their own and other household members' risks would compare with that average. They were also asked to think about how they would spend various amounts of money between $20 and $4,000.

After being invited to talk about how the various types of non-fatal injury would affect their own and other household members' quality of life, respondents were asked what sums of money might be required in order to make up for the loss of enjoyment entailed by suffering, respectively, a minor injury and a temporary injury. This was followed by a short set of questions eliciting their perceptions of the severity of all three types of non-fatal injury relative to each other and to death.

By this stage in the interview, then, respondents had been asked to reflect in some depth upon the kinds of injuries that might be sustained, their own and other household members' risks of sustaining such injuries, and the meaning of various sums of money as carriers of utility. It was only at this stage that they were presented with what were, in effect, the key WTP and WTA questions.

We shall describe the WTA question first, although which of the two was presented to respondents first was actually determined at random.

The basic scenario was that respondents were asked to suppose that they and other members of their household were having to relocate[2], and had to choose between Area X - where the cost of living and road safety record was the same as where they currently lived - and Area Y, where the risks to all household members of being killed or injured on the roads were 20% higher, but the cost of living was lower. Interviewers explained what a 20% increase would mean in terms of the subjective estimates of own and other household members' risks elicited earlier in the interview.

[2] This was important: if one option had been to stay put, this might have been desirable for numerous reasons other than safety – for example, avoiding the costs of moving, being certain about the desirability of the neighbourhood, not having to disrupt friendships, etc. - all–of which could have accounted for any WTA-WTP disparity.

Randomly chosen initial values

Because the survey was designed to use a combination of paper-based and laptop computer-based stimuli, it was possible to present the two Area options to respondents on the laptop screen and to insert the Area Y cost of living differential as one of three (randomly chosen) initial values: either $20, or $100, or $400 per year lower than Area X. There were three possible responses: definite preference for Area X; definite preference for Area Y; or a statement that it was hard to choose between the two. Whenever a preference for Area X was registered, the computer increased the size of the cost of living differential, and the respondent was invited to make a fresh choice. Whenever a preference for Area Y was expressed, the differential was reduced, and a new choice was requested. This process of iterative pairwise choice stopped when the differential became such that the respondent found it hard to choose between the two areas. After checking that this was indeed how the respondent felt (and if it was not, re-running the question until the point at which it was hard to choose was established and confirmed), that value was recorded as the household's 'best estimate' WTA for a 20% increase in household members' annual risks of being killed or injured.

The question was framed as a series of pairwise choices rather than as a question in the often-used form "What is the minimum amount you would accept ...." because, arguably, making a series of pairwise choices simplifies the task and is less likely to stimulate strategic thinking of the kind discussed in the previous section of the paper. Certainly, it allowed exactly the same format to be used for the WTP question, in which respondents were asked to choose between Area X and Area Z, where risks were either 20% or else 50% lower (this figure also being determined at random) while the cost of living was higher by the same amount as initially presented in the WTA question. The same iterative structure was then used to home in on a figure where the respondent found it hard to choose, and this figure (once confirmed) was then recorded as the 'best estimate' WTP response.

In both questions, if the iterative process took respondents to a point where even $5,000 was too small a differential to make it hard to choose, they were invited to state what differential *would* make it hard to choose. If they gave a finite amount, it was entered. If they stated that no finite amount would be enough, this too was recorded (and later coded as 999999).

Thus for each respondent we have both a WTA and a WTP figure, although for half of the sample, the WTP figure relates to a 50% lower risk, rather than the 20% presented to the other half of the sample; while all respondents' WTA figures are for a 20% higher risk.

The point of randomising respondents between two levels of risk reduction was to provide a check for sensitivity to the size of the benefit. There is now a great deal of evidence from earlier surveys and experiments that although the size of the risk change plays a crucial part in the computation of VOSLs, respondents may not be as sensitive

to this information as is required to produce a VOSL in which one might have confidence[3].

WTA - WTP disparity

Table 1 summarises the distribution of responses to both questions by the members of each subsample.

Comparison of the two columns of WTP responses shows that, in this study, respondents *did* display sufficient sensitivity to the size of the risk reduction to generate distributions of responses that were significantly different from each other in the direction that would be expected ($\chi^2 = 31.72$, 8 d.f.; $p < 0.001$). Considering the rather dismal history of previous surveys in this respect, such a result is very encouraging – and all the more so since it represents between-sample, rather than just within-subject, sensitivity. And although such a result, by itself, is not enough to establish the general validity of these responses, it can at least be regarded as providing some reassurance about the quality of the data.

Table 1: Distribution of WTA and WTP responses

	Subsample 1; n = 510		Subsample 2; n = 501	
Responses	WTA 20%	WTP 20%	WTA 20%	WTP 50%
NZ$				
0 – 10	10	89	14	56
11 –20	28	48	15	37
21 –100	41	78	33	50
101 – 400	63	86	71	86
401 – 1000	42	52	36	52
1001 – 1500	51	43	50	45
1501 – 4000	91	65	87	79
4001 – 5000	26	14	43	30
> 5000	148	35	152	66

At the same time, a comparison of WTA and WTP responses from the members of Subsample 1, who were asked about increases and decreases of the same magnitude, reveals a clear disparity between the two, of the kind so frequently reported in other studies ($\chi^2 = 162.7$, 8 d.f.; $p < 0.001$). The median WTA response in Subsample 1 was $2,250, compared with a median response of $275 to the WTP question. A comparison of means is a little more difficult, given more than 90 WTA responses refusing to state a

[3] Baron and Greene (1996) review some of this evidence and report a whole battery of further experiments showing insensitivity to the quantity of the good being valued; and Beattie *et al.* (1998) report a series of surveys where many respondents displayed little or no sensitivity to the size of the risk reduction, even when the relevant responses were elicited in consecutive questions and when respondents had their attention drawn to their answers and were given the opportunity to modify them if they wished.

finite amount; but even if we take a fairly radical (if not brutal) approach and set all such responses – and indeed, all responses involving finite figures greater than $20,000 – equal to $20,000, the mean WTA response would be $6,280, compared with a mean WTP response of $1,672[4].

To provide another perspective on the data, we can compute each individual's WTA/WTP ratio for all those who gave finite responses greater than zero to both questions. Table 2 reports the distributions for each subsample.

Table 2: Distribution of WTA/WTP ratio

WTA/WTP ratio	Subsample 1; n = 421 20% & 20%	Subsample 2; n = 397 20% & 50%
0 – 0.5	39	74
0.51 – 1.0	104	108
1.01 – 2.0	46	45
2.01 – 3.0	36	28
3.01 – 5.0	36	23
5.01 – 20.0	54	40
20.01 – 100.0	38	32
> 100	68	47

It can be seen that the higher WTP values for a risk reduction of 50% increase the denominator for Subsample 2 and thereby push the distribution of ratios significantly to the left ($\chi^2 = 20.54$, 7 d f; $p = 0.005$). However, the tendency for WTA to exceed WTP is clear for both subsamples, and in Subsample 1, where the magnitude of the risk change is 20% in both cases, the median ratio is 3.0, while the mean is very much higher[5]. If we allow for another 99 individuals in Subsample 1 who are not included in Table 2 because they did not give finite non-zero answers to both questions, the median ratio moves up to 4.0.

Overall, then, what emerges quite clearly from the responses of a large and representative sample of the New Zealand population is that even when every effort has been made (within the constraints imposed by conducting a national sample survey) to elicit WTA and WTP in as neutral a way as possible, after reasonable opportunity for reflection, and using a question format intended to avoid various known sources of bias, comparisons of relatively conservative measures of central tendency – medians and means based on responses severely capped at the upper end – show WTA to be at least three or four times greater than WTP. Given such a clear pattern in the data, the

[4] For Subsample 2, the median WTA was $3,000 as opposed to a median WTP (for a 50% risk reduction) of $600; setting maximum responses at $20,000, mean WTA was $6,704 as compared with a mean WTP of $2,869.

[5] The problem with computing a mean is to decide what to do with the cases where the ratio is extremely high. If we arbitrarily constrain the individual ratio not to exceed 20, the mean ratio is 9.1; if we set the maximum ratio at 10, the mean ratio falls to 5.3.

question then is: should the New Zealand public authorities take this as persuasive evidence that the people of the country want increases in road risks to be treated differently from reductions; and if so, should this be reflected by using a higher VOSL when considering increases than when evaluating reductions?

Policy implications

Before discussing the possible implications – including some potential problems – arising from using a WTA-based VOSL several times greater than the WTP-based figure, the first question to ask is: do we believe that the data are reflecting people's true preferences, or might they be predominantly the product of the kinds of biases discussed in the second section of this paper?

There is, of course, no obvious way of providing a conclusive answer on the basis of the data themselves. Inevitably, a degree of judgement must be used.

Large disparity

On the one hand, the responses were generated in a one-off interview consisting of unfamiliar hypothetical scenarios and lasting about an hour, with respondents having no opportunity to discuss with others or to reflect at their leisure about the answers they gave. When disparities occurred, they were not drawn to people's attention, and there was no chance for people to modify their answers should they have wished to do so. (It was simply not feasible to build such opportunities into the survey design.) Thus it could be argued that the disparities observed might in large part have been the result of partial and asymmetric information processing which would have disappeared – or at least, would have been greatly mitigated – if there had been time and real incentives to think the issues through carefully.

On the other hand, the format of the questions was designed to reduce the impact of known potential biases as far as possible. The questions were presented as a series of pairwise choices in order to simplify the decisions and to try to steer people away from taking buyer/seller strategic stances. The scenarios were framed in terms of households' private decisions about their own choice of location, rather than in terms of voluntary or tax-based contributions to a public good, and it was hoped that this would be less likely to provoke 'protest' responses. As much care as is possible for a survey of this kind was taken to enable respondents to familiarise themselves with the issues, and to put the risk changes in simple percentage terms based on their estimates of their own baseline risks. To the extent that responses were sensitive to the size of the risk reduction, it appears that the survey design was more successful than many of its predecessors.

Yet even so, the WTA-WTP disparity is observed extensively and is sufficiently strong that it registers as a ratio of 3:1 or 4:1 even when quite conservative measures are used. Thus even though there may still be some sources of bias at work, it would seem hard to believe that these could account for a difference so much greater than the few

percentage points entailed by standard theory under plausible assumptions. We are therefore inclined to take the view that the data reflect characteristics of people's preferences – possibly not highly articulated, but present at some basic level – which are consistent with explanation (2) and some form of reference-dependent model of choice.

But does it necessarily follow from such a proposition that the appropriate reaction of government agencies is to use a different VOSL for risk increases than for corresponding risk reductions?

Consistency over time

One issue relates to consistency over time and the possible non-reversibility of decisions. For example, one can imagine a situation in which a regulation is introduced in order to reduce risk at a time at which the WTP-based estimate of the benefits of reduced death and injury gives a Benefit/Cost (B/C) ratio that exceeds the threshold. But suppose that some time later it is discovered that circumstances have changed and that the B/C ratio has fallen below the threshold (or that the original B/C ratio estimate was overoptimistic). Under these circumstances, if the decision were to be taken afresh the regulation would not now be introduced – or to put things in a slightly different way, if the regulation had to be renewed every x years to stay in force, the B/C ratio would now not justify renewal and the regulation would lapse. But if, instead, the decision is framed in terms of removal of the *existing* regulation, thereby increasing risk in return for some other benefits in the form of time or cost savings, then the use of a higher WTA-based value of safety might lead to rejection of the proposal to remove the regulation, with the result that a regulation which would not now be introduced, in fact remains in force.

But is this really a problem? If the reference-dependent model really is capturing something that is essentially true of people's preferences – and which characterises their behaviour in the case of private goods (as Knetsch (1989) demonstrated and many people's introspection would suggest) – then perhaps the appropriate response for the policymaker is to appreciate that there may well be a degree of irreversibility in decisions concerning the introduction of new safety regulations and that particular care therefore needs to be taken to ensure the accuracy of risk reduction benefit estimates.

Gainers and losers

A second potential problem that may arise from having a WTA-based VOSL for risk increases that substantially exceeds its WTP-based counterpart for risk reductions relates to situations in which a safety programme or regulation reduces risk for some people but increases it for others. Thus, for example, suppose that a proposed safety-programme would involve the diversion of road traffic from a relatively high-density, high-risk route to a low-density, low-risk route. Suppose, in addition, that the result would be that 50% of the population in the area concerned would enjoy a risk reduction δr that would be expected to prevent 15 fatalities, while 30% of the population would

be subject to an increase in risk equal to 2/3 δr resulting in 6 additional fatalities, the remaining 20% of the population being unaffected. The overall impact of the proposed programme would therefore be a net reduction in the number of road fatalities in the area concerned of 15 – 6 = 9 and, very probably, a more equitable distribution of risk in the sense that risk on the high-risk route would fall somewhat while that on the low-risk route would rise a little. In terms of both efficiency and equity, therefore, the proposed programme would appear to be highly desirable, provided that it was not too costly to implement. But if the programme were to be assessed by applying a WTA-based VOSL to the 6 additional fatalities and a WTP-based VOSL to the 15 fatalities prevented, then with the WTA figure exceeding its WTP counterpart by a factor of, say, 3 then the programme would unequivocally be rejected even if it were costless to implement.

While this conclusion may, *prima facie,* seem somewhat uncomfortable, it should be stressed that if people's preferences really *are* reference-dependent to a degree that produces a WTA-based VOSL equal to three or more times its WTP-based counterpart then the gainers in the above example simply could not compensate the losers so that the programme would fail the hypothetical compensation test.

Against this, of course, it might be argued that a programme that would on balance save lives *and* result in a more equitable distribution of risk *ought* to be undertaken and that the results of a cost-benefit analysis using a mixture of WTA and WTP-based VOSLs should simply be overridden. With this in mind, it is tempting to suggest that one way forward with a decision such as this might be to ask those who would be potentially affected to cast their votes for or against the programme from behind a 'veil of ignorance', not knowing whether they would be gainers or losers but only that there would be a 0.5 probability of being a gainer with a risk reduction of δr, a 0.3 probability of ending up a loser, with an increase in risk equal to 2/3 δr and a 0.2 probability of being unaffected. But again, if preferences really are reference-dependent then even from behind the veil of ignorance a majority might well vote against the programme, preferring to 'stick with' the status quo. Indeed, a 'veil of ignorance' experiment *might* be the way forward as a means of testing the appropriateness of the reference-dependent hypothesis as an explanation for the WTA-WTP disparity.

Conclusions

While the WTA-WTP disparities reported in this paper *may* be the result of various biases, it is our belief that these disparities are in large part a genuine reflection of reference-dependent effects in which losses from a given reference (or starting) point loom considerably larger in people's preferences than corresponding gains. If this is so, then WTA-based VOSLs for risk increases can indeed be expected to exceed corresponding WTP-based figures for risk reductions by a considerable margin.

We also believe that the 'veil of ignorance' experiment outlined in the preceding section may offer some potentially fruitful insights.

Acknowledgement and disclaimer

The study was sponsored by the Land Transport Safety Authority, New Zealand. However, the views expressed in the paper are strictly those of the authors and are not necessarily shared by the Land Transport Safety Authority.

References

Baron, J. and Greene, J., (1996). "Determinants of Insensitivity to Quantity in Valuation of Public Goods: Contribution, Warm Glow, Budget Constraints, Availability and Prominence." *Journal of Experimental Psychology: Applied*, 2, 107-25.

Bateman, I.J., Munro, A., Rhodes, B., Starmer, C. and Sugden, R. (1997) "A Test of the Theory of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 112(2), 479-505.

Beattie, J., Covey, J., Dolan, P., Hopkins, L., Jones-Lee, M., Pidgeon, N., Robinson, A., and Spencer, A., (1998). "The Contingent Valuation of Safety and the Safety of Contingent Valuation, Part 1: Caveat Investigator." *Journal of Risk and Uncertainty*, 17, 5-25.

Burton, A., Chilton, S. and Jones, M., (1999). "The WTA/WTP Discrepancy – A Preliminary Qualitative Examination." Forthcoming in *Environmental Values*.

Coursey, D. L, Hovis, J. L and Schulze, W. D (1987). "The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value." *Quarterly Journal of Economics*, 102(3), 679-90.

Dubourg, W. R, Jones-Lee, M and Loomes, G (1997). "Imprecise Preferences and Survey Design in Contingent Valuation." *Economica*, 64, 681-702.

Guria, J. C, Jones, W, Jones-Lee, M. W, Leung, J, Loomes, G and Keall, M (1999). *The Values of Statistical Life and Prevention of Injuries*. Draft Report - not yet released for public use.

Hanemann, W. M. (1991), "Willingness To Pay and Willingness To Accept: How Much Can They Differ?" *American Economic Review*, 81(3), 635-646.

Kahneman, D, Knetsch, J. L and Thaler, R. H (1990). "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy*, 98, 1325-1348.

Knetsch, J., (1989). "The Endowment Effect and Evidence of Nonreversible Indifference Curves." *American Economic Review*, 79, 1277-84.

Miller, T. R and Guria, J. C (1991). *The Value of Statistical Life in New Zealand: Market Research on Road Safety*. Wellington: Land Transport Division, Ministry of Transport.

Shogren, J. F; Shin, S. Y; Hayes, D. J and Kliebenstein, J. B (1994). "Resolving Differencees in Willingness to Pay and Willingness to Accept." *American Economic Review*, 84(1), 255-270.

Sugden, R., (1999). "Public Goods and Contingent Valuation" and "Alternatives to the Neo-classical Theory of Choice", Chapters 5 and 6 in *Valuing Environmental Preferences*, eds. Bateman, I. and Willis, K., Oxford University Press

Tversky, A. and Kahneman, D., (1991). "Loss Aversion in Riskless Choice: A Reference-Dependent Model." *Quarterly Journal of Economics*, 106, 1039-61.