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

JAGADISH GURIA[1], JOANNE LEUNG[1], MICHAEL JONES-LEE[2] and GRAHAM LOOMES[3,*]

[1]Land Transport Safety Authority, Wellington, New Zealand; [2]Department of Economics, University of Newcastle upon Tyne, NE1 7RU, UK; [3]Center for the Economic and Behavioural Analysis of Risk and Decision, University of East Anglia, NR4 7TJ, Norwich, UK;
*Author for correspondence (e-mail: g.loomes@uea.ac.uk)

Abstract. Large disparities between willingness to accept (WTA) and willingness to pay (WTP) based values of statistical life are commonly encountered in empirical studies. Standard economic theory suggests that if a public good is easily substitutable there should be no marked disparity between WTA and WTP values for the good, though the disparity increases with reduced substitutability. However, psychologists have shown that people often 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 reports the disparity observed between the two measures and considers some of the problems posed for policymakers.

Key words: safety policy, willingness to accept disparity

1. Introduction

When evaluating transport projects which entail expected reductions in the risks of death and injury, a number of countries have for some time used *values of preventing a fatality* (VPF)[1] and *values of preventing an injury* (VPI) to enable the safety benefits to be weighed against the costs of providing them. Ideally, VPF and VPI should reflect the collective *willingness to pay* (WTP) of members of the population for the relevant improvements in their own (and possibly others') safety. More recently, there has been increasing interest in extending the use of such monetary values to other areas of health and safety: for example, to estimate the value of the health benefits of reducing air pollution (see, for example, Krupnick et al. 2002).

However, while most policy applications are aimed at reducing risks, there may be occasions when a change of policy could entail some increase in the risks of death and injury for at least some sections of the population. In such cases, it is appropriate to value that change in risks on the basis of people's *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 very 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 is organised as follows. The next section discusses various possible explanations of the disparity between WTA and WTP. The third section reports the disparities observed in the New Zealand VoS and discusses the status of those results. In the fourth section we consider possible problems for practical policy and discuss certain questions that should be explored before deciding whether the substantial difference between WTA and WTP should be incorporated into public decision making.

2. 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 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 measure which could reduce risks and give the individual a little more safety, the conventional assumption is that she would be prepared to forego some other consumption at the margin – and thus free up some money – to achieve that benefit, up to the point where she returns to the same level of welfare she had before the safety improvement became available. The money thus freed up 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.

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, where the choice of rule of thumb is 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, 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. We now consider each explanation in a little more detail.

2.1. 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" (p. 256).

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. 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.

2.2. GAINS VERSUS 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.

To observe WTA substantially exceeding WTP across a wide variety of goods 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 influence, and that it is this susceptibility which is largely responsible for the WTA–WTP disparity.

2.3. INFLUENCES ON THE CONSTRUCTION OF RESPONSES

One possible explanation of the disparity 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. 2000); 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.

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 for policy.

3. The New Zealand Study

A WTP-based VPF/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). Several years later, the Land Transport Safety Authority of New Zealand received funding to undertake a nationwide survey of travel behaviour and decided that this provided an opportunity to revisit and update the WTP-based value and extend the methodology to encompass non-fatal injuries of various severities. And since some policy changes have the potential to increase risk of death or injury from its base level, it was decided that the study should also elicit WTA values. The new survey was developed and conducted in 1997/98: a fuller description of the design and the results can be found in Guria et al. (1999). Below, we focus on the essential features of the study as they relate to the issue of the relationship between WTA and WTP.

Although the great majority of people in New Zealand and most other developed countries have considerable experience of being car drivers and/or passengers and are therefore conscious of road safety issues in general terms, the task of valuing changes in their risk levels is, for most people, an unfamiliar and demanding one. The first half of the interview protocol was therefore devoted to various "warm-up" tasks: asking respondents to think about several types of injury and the discomfort, inconvenience and/or distress entailed by them; giving them customised information about the average level of risk in their region and helping them to think through where their own and other household members' risks stood in relation to that average; encouraging them to think about the alternative consumption possibilities afforded by various sums of money from $20 to $5,000[2]; and so on. The intention was to give as much opportunity as is feasible in a one-to-one at-home interview lasting around 45–50 minutes for respondents to give some reasonable thought to several key considerations. They were also given some experience with the iterative value elicitation procedure (about which, more below). It was only at this stage that they were presented with the crucial 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,[3] 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 annual cost of living was lower. Interviewers drew attention to what a 20% increase would mean in terms of the subjective estimates of own and other household members' risks elicited earlier in the interview. Framing the

risk change in terms of a percentage of respondents' own baseline was a deliberate attempt to overcome (or at least, mitigate) the difficulty many respondents have with information about small changes in absolute risks (presented, for example, as 3 in 100,000).

Because the survey used a laptop computer display supplemented by paper/showcard material, 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 assigned) initial values: either $20, or $100, or $400 per year lower than Area X. There were three possible responses: a definite preference for Area X; a 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 frequently-used open-ended form "What is the minimum amount you would accept ..." because, arguably, making a series of dichotomous choices simplifies the task and may be less likely to stimulate "strategic bias" of the kind discussed in the previous section of the paper. Moreover, 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 such procedures, there is, of course, a danger that interviewees will "anchor" their responses on the first value presented to them and exhibit some degree of "starting point bias".[4] We tried to defuse this effect by using the same initial amount for all questions presented to a particular individual, including all the "warm-up" questions. Thus, if some anchoring bias had been in operation, using the same anchor for both WTA and WTP would arguably have had the effect of pulling both values towards the anchor, thereby tending to reduce any disparity. But, as it turned out, no significant anchoring effect could be detected in these questions: within any subsample

(i.e. after controlling for the order of presentation and the magnitude of the risk reduction in the WTP question) there was no systematic difference between the distributions of responses across the three starting values.[5] Nor did the order in which the questions were presented have any significant impact on the patterns of response.

In both questions, if the iterative process took respondents to a point where even $5,000 per year 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 VPFs, respondents may not be as sensitive to this information as is required to produce a VPF in which one might have confidence[6]. It is a noteworthy feature of this survey that respondents *did* appear to be sensitive to the difference in the size of the benefit, so that there was no significant difference between the WTP-based VPFs inferred from the two levels of risk reduction.

Overall, then, there are a number of reasons to believe that the responses to the crucial questions in this survey were less vulnerable to various well-known forms of bias and/or heuristic than many earlier studies. To summarise: within the constraints of this form of interview, strenuous efforts were made to enable respondents to consider the key elements in trade-offs between money and safety; an iterative pairwise choice procedure was used to make the WTP and WTA formats as similar as possible and to reduce the availability of strategic responses; common biases such as starting point effects and order effects were tested for and found to be insignificant; and WTP responses were found to be appropriately sensitive to the information about the magnitude of the benefit. In other words, if the responses to these questions exhibit a clear disparity between WTA and WTP, there are grounds for thinking that explanations (3) and/or (4) from the previous section may be relatively minor contributors, and that it is arguable that the responses reflect some rather more robust characteristic of people's preferences.

3.1. WTA–WTP RELATIVITIES

Given the absence of any evidence of starting point bias or of order effects, we divide the data into just two subsamples, differentiated by the size of the

Table I.

Responses	Subsample 1; n = 510		Subsample 2; n = 501	
	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

risk reduction in the WTP question. Table I summarises the distribution of responses to both questions by the members of each subsample.

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 finite amount; but even if we take a fairly radical 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.[7]

So 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. So although it would be rash to suppose that there are *no* elements 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 for public policymakers is to use values for risk increases substantially greater – perhaps three or four times higher – than the values for corresponding risk reductions? The next section explores some of the issues raised by that question.

4. Policy Implications

In this section, we shall focus on three scenarios where the use of very different WTA and WTP values raises questions/issues/potential problems.

The first scenario relates to consistency over time and possible irreversibility of policy decisions. Suppose that some new regulation or piece of legislation intended to improve safety is evaluated on the basis of a conventional *ex ante* social cost-benefit analysis using WTP-based values of safety. Suppose, furthermore, that the benefit-cost ratio exceeds the required level and that the proposal is therefore introduced. However, suppose that some time later the actual benefits of the regulation turn out to be somewhat lower than expected – either because of changed circumstances or over-optimism in the original *ex ante* predictions – and that if the actual benefits that have emerged had been incorporated in the original cost-benefit analysis then the regulation would not have been introduced. Nonetheless, it *has* been introduced and now constitutes part of the current *status quo* so that the issue is whether or not it should be rescinded. Rescinding it would deprive people of various safety benefits in return for a saving of resources, so that by conventional welfare economics standards the appropriate calculation would entail using the WTA-based value of safety. But suppose that the WTA value is, say, three times the WTP counterpart and that as a result the loss from abolishing the regulation would outweigh any cost savings. Thus, a regulation that would not now be introduced nonetheless remains in place.

Many people's initial reaction to this apparently somewhat paradoxical implication of a substantial WTA–WTP disparity is, to say the least, one of mild discomfort, and prompts us to ask whether that implication really is defensible.

A possible line of defence is that so long as the WTA values are based on an unbiased expression of well-informed preferences, then however much the configuration of those preferences may diverge from standard neo-classical assumptions, respect for those preferences does indeed call for rejection of the proposal to rescind the regulation. After all, it is possible that people's preferences *are* to some extent reference-dependent, so that once they have become accustomed to some new pattern of consumption their reference point changes and their preferences adapt accordingly. In other words, if

indeed members of the affected population *are* comfortable with the implications when they are confronted with them, then the case for using very different WTA and WTP values would seem unassailable (at least, from a standard liberal welfare economist's perspective).

If this is the case, it is important for policymakers 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 assessment.

The second scenario relates to possible asymmetries between the introduction and abolition of safety policies. For example, the practical reality of government spending is that the political process determines budgets for different areas of public expenditure and that cost-benefit analysis is then used as an aid to prioritising projects within those budget constraints. Typically, therefore, the lowest ranked project which can be afforded within a given budget will have a benefit:cost ratio significantly greater than unity. To illustrate, suppose that a project which is expected to prevent a total of 30 fatalities (for whom the WTP-based VPF is, say, $4 m) and which would cost $80 m – that is, a project which has a benefit:cost ratio of 1.5 – falls just outside the budget constraint. Suppose also that there is another proposal to rescind some existing regulation. It is estimated that this would save $80 m, but would involve some compromise of safety standards, expected to result in 10 fatalities that would not otherwise occur. If those fatalities were valued on the same WTP basis, the cost of the additional fatalities would be $40 m so that the benefit:cost ratio of rescinding the regulation would be 2.0, and if the money thereby released could be diverted to fund the project which would prevent 30 deaths,[8] there would be a net gain of 20 fatalities prevented. However, if the evaluation of the proposal to rescind the regulation entails using a significantly higher WTA figure – say, $10 m per fatality – the benefit:cost ratio is less than 1: the regulation remains in place, the marginal project is not implemented and the aggregate death toll is 20 higher than it might otherwise have been. This is a possible implication of incorporating a significant WTA–WTP disparity into policy.

Our third scenario relates to situations in which a project or regulation reduces risk for some people but increases it for others. For example, consider a scheme to divert traffic from one area to another. Suppose that Area A currently has relatively high traffic density and hence a high road accident rate, while Area B has relatively low traffic density and a low road accident rate. Now suppose that there is a proposal which will divert traffic from Area A to Area B, thereby reducing fatalities in Area A by 50 but increasing fatalities in Area B by 25 – a net benefit of 25 prevented fatalities which, if valued according to the WTP-based VPF used earlier, would be worth $100 m. If the costs of the programme were, say, $50 m, evaluating the proposal in this way would generate a benefit:cost ratio of 2.0. But if we take

the view that the proposal increases the risks to the population of Area B and that we should use a WTA-based value of $10 m per fatality to represent their preferences, then the benefits (the 50 deaths prevented in Area A, valued at $200 m) come to less than the costs ($50 m plus the $250 m value of the additional 25 deaths in Area B). Thus, under a conventional cost-benefit analysis, and appealing solely to considerations of efficiency, this proposal would be rejected, while an alternative proposal that would cost the same but prevent fewer deaths in total would be implemented.

To sum up: there may be cases – and our scenarios have tried to illustrate some of them – where the use of very different WTA and WTP values may raise implications with which we may not be altogether comfortable. So even though we may have strong grounds for believing that those different values are based on responses which reflect a genuine feature of people's preferences, we may wish to explore more thoroughly the issues raised by those cases before applying too generally a WTA value significantly greater than WTP.

Acknowledgements

Notes

1. This term is increasingly being used instead of the term *value of statistical life* (VOSL): the concept is the same, but the term VPF may convey it better.
2. All figures are in (1998) New Zealand dollars.
3. 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 distorted the WTA and WTP responses.
4. A review and discussion of evidence of "anchoring effects" and their possible impact on WTP responses can be found in Kahneman et al. (1999).
5. Interestingly, there *was* some evidence of starting point effects in the very first "practice" questions; but it is as if the opportunity to practice the procedure diminished the weight placed on the starting point in the later questions.
6. 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.

7. 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.
8. In practice, of course, this may not be so easy to achieve: *some* savings might accrue to government, but other savings might go to individuals. However, for the sake of our stylised example and the point it is aiming to illustrate, we suppose that the funds could be diverted.

References

Baron, J. and J. Greene (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**,107–125.

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

Beattie, J., J. Covey, P. Dolan, L. Hopkins, M. Jones-Lee, N. Pidgeon, a. Robinson and A. Spencer (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., S. Chilton and M. Jones (2000), 'The WTA/WTP Discrepancy – A Preliminary Qualitative Examination', *Environmental Values* **9**, 481–491.

Coursey, D. L., J. L. Hovis and W. D. Schulze (1987), 'The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value', *Quarterly Journal of Economics* **102**(3), 679–690.

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

Guria, J. C, W. Jones, M. W. Jones-Lee, J. Leung, G. Loomes and M. Keall (1999), The Values of Statistical Life and Prevention of Injuries. Draft Report.

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

Kahneman, D., I. Ritov and D. Schkade (1999), 'Economic Preferences or Attitude Expressions? An Analysis of Dollar Responses to Public Issues', *Journal of Risk and Uncertainty* **19**, 203–235.

Krupnick, A., A. Alberini, M. Cropper, N. Simon, B. O'Brien, R. Goeree and M. Heintzelman (2002), 'Age, Health and the Willingness To Pay for Mortality Risk Reductions: A Contingent Valuation Study of Ontario Residents', *Journal of Risk and Uncertainty* **24**, 161–186.

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

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

Sugden, R. (1999), 'Public Goods and Contingent Valuation', Chapter 5, pp. 131–152, and 'Alternatives to the Neo-classical Theory of Choice', Chapter 6, pp. 152–181 in I. Bateman I and K. Willis, eds., *Valuing Environmental Preferences*. Oxford: Oxford University Press.

Tversky, A. and D. Kahneman (1991), 'Loss Aversion in Riskless Choice: A Reference-Dependent Model', *Quarterly Journal of Economics* **106**, 1039–1061.