Towards a Willingness-to-Pay Based Value of Underground Safety

By Michael Jones-Lee and Graham Loomes*

1. Introduction

It is self-evident that London Underground Limited (LUL)[1] does not have access to unconstrained funding to finance safety improvement. This means that LUL faces hard choices in its appraisal of proposed safety projects. These choices will be virtually impossible to make in a systematic and consistent manner unless LUL has a monetary measure of the *value* of safety improvement so that value can be weighed against cost in reaching an informed decision as to whether or not any particular Underground safety project should be undertaken.

This then raises the question of *how* monetary values of safety should be defined and estimated for use in the Underground context. The clue to an answer lies in LUL's business objective which calls for it to *"maximise net social benefit within available funds and subject to a defined gross margin target"* (London Underground Limited, 1991). Clearly, any satisfactory definition of "social benefit" (or "social value") must take account of the interests and well-being of members of the public, and for present purposes this plainly means the travelling public (including, ideally, those who will benefit from the lower casualty rates, as well as the reduced road congestion and pollution, that will result from a substitution of Underground travel for road use). If one also allows that the interests and well-being of members of the travelling public are in large part determined by the nature of their *preferences*, then it follows that in LUL's case "social benefit" should be defined in such a way as to reflect those preferences.

---

* The University of Newcastle upon Tyne and the University of York, respectively. The research reported in this paper was carried out under contract to London Underground Limited. However, the opinions expressed in the paper are solely the responsibility of the authors and do not necessarily reflect the views of London Underground. The authors are grateful to two anonymous referees and the editor of this special issue for helpful comments and suggestions.

[1] London Underground Limited is the state-owned operator of London's underground railway system.

Given all of this, it is evident that LUL will require a measure of its customers' preferences — and more particularly their *strength* of preference for safety. Such a measure is naturally provided by customers' *willingness to pay* for safety, in that the maximum amount that a person is willing to pay for a good or service is a direct indication of what that good or service is worth to the person, relative to his or her other potential objects of expenditure. Furthermore, willingness to pay, being conditioned by ability to pay (that is, income), is an indication of strength of preferences, taking due account of resource constraints.

In summary, if LUL's safety expenditure decisions are to be systematic, consistent and pursuant of LUL's business objective, then:

- These decisions must involve a balancing of social benefit and cost, where the measure of social benefit reflects customers' strength of preference for safety.
- Since strength of preference is naturally measured by willingness to pay, LUL should ideally work with appropriately estimated *willingness-to-pay* (WTP) based values of safety.

Lest it be thought that this proposal is unduly radical, it should be appreciated that the UK Department of Transport and the Health and Safety Executive *both* inform their decisions by reference to WTP based values of safety (see Department of Transport, 1988, and Health and Safety Executive, 1992, Appendix 3), as do corresponding agencies in the USA, Canada, Sweden and New Zealand, with other countries showing clear signs of moving in the same direction. *More significantly, the Treasury has given unequivocal and emphatic advice that bodies such as LUL should adopt the WTP approach to the valuation of safety* (see H.M. Treasury, 1991, Annex B).

In the light of considerations such as these and because, hitherto, no estimates of WTP based values of Underground safety have been available, LUL instituted a programme of research, to be undertaken by the authors of this paper, aimed at obtaining such estimates. Following a desk study (designated "Phase 0") intended to identify the key issues and questions to be addressed, the first substantial phase of the research programme ("Phase 1") involved a pilot study, conducted during November 1992, designed to assess the feasibility of our proposed estimation procedures and provide some first indications of the possible order of magnitude of WTP based values of Underground safety. It is hoped that the main phase of the research programme ("Phase 2"), involving a large-scale sample survey, will be conducted in the near future. The purpose of this paper is to report the findings of the Phase 1 pilot study.

2. The Department of Transport's WTP Based Value of Preventing a Road Fatality

In principle, under the willingness-to-pay approach one seeks in the first instance to establish the amounts that affected individuals would be willing to pay for typically very small improvements in their own or others' safety. These (usually quite modest) amounts are then aggregated across all affected individuals to arrive at an overall monetary value for the safety improvement concerned. If the improvements in individual safety are

Table 1

Breakdown of DoT WTP-Based Value of Preventing a Road Fatality

Pure WTP component	£634,600
Component reflecting avoided medical and ambulance costs	£450
Component reflecting avoided output loss, net of consumption	£48,100
Total	£683,150

standardised so that — taken over the affected group — they can be expected, on average, to prevent one fatality, then the resultant aggregate willingness to pay is naturally referred to as the "value of preventing a *statistical* fatality" or more succinctly as the "value of a statistical life". Because it is cumbersome and somewhat forbidding, the term "statistical" is often dropped and one refers simply (but not entirely accurately) to the "value of preventing a fatality" or the "value of life".

Of course, in the process of aggregating willingness to pay across different individuals — in order to arrive at an overall value of safety — it is necessary to make a judgement about the *relative* social value of a £1 gain to different income groups. In conventional social cost-benefit analysis the aggregation is typically carried out on an *un*weighted basis, so that a £1 gain is effectively treated as having a constant social value, irrespective of the income level of the individual to whom it accrues. By contrast, in weighted cost-benefit analysis distributional weights are employed to reflect judgements about the way in which the social value of a £1 gain should vary with the income level of the recipient of the gain. All aggregate values referred to in this paper are unweighted.

To get a feel for the composition and magnitude of a WTP based value of statistical life it will be instructive to consider the figure currently used by the Department of Transport (DoT) in its analysis of proposed road projects.[2] This is shown in Table 1 in 1991 prices.

As can be seen, apart from the pure willingness-to-pay component (which is the DoT's estimate of aggregate willingness to pay for small safety improvements that, taken together, can be expected to prevent one road fatality) there are two further components. These components are included because there is strong evidence that individuals tend not to take account of them in determining their own individual willingness to pay for safety (see, for example, Jones-Lee, 1989, pp. 169-74). While the rationale for inclusion of the first of these additional components is self-evident, the reason for including the component reflecting avoided output loss, net of consumption, is somewhat less obvious. Essentially, this component is included because when an individual dies prematurely the rest of society (in the first instance, the individual's dependants, the Inland Revenue and DSS) loses the excess of what the individual would have produced during the remainder of his or her working life *over and above* what he or she would have consumed. The output loss figure is in fact an estimated *average* for this loss, net of consumption.

[2] The DoT decided to adopt the willingness-to-pay approach in place of its former gross-output based procedure in 1988 following a comprehensive review of the literature and careful consideration of the various issues involved. See Dalvi (1988) and Department of Transport (1988).

3. Towards a WTP-Based Value for Preventing an Underground Fatality[3]

How might a corresponding willingness-to-pay based value for the prevention of an Underground fatality be expected to break down? We would suggest that the following constitute the key components:

(a) Pure WTP component;
(b) Component reflecting the avoided medical and ambulance costs;
(c) Component reflecting the avoided cost of legal compensation (or, equivalently, the cost of insuring against claims for compensation).

As can be seen, the only component that has no obvious counterpart in the DoT breakdown for the road value is that for avoided costs of legal compensation. However, English courts award damages to the surviving dependants of the deceased victim of tortious negligence in an amount that is principally intended to compensate for their financial loss. In particular, the compensation awarded is intended to reflect the victim's "net income" — that is, the excess of his/her income over and above his/her consumption, had he/she survived. As such, one might therefore expect that the average compensation award for a fatality would be broadly similar to the net output loss component in the DoT road breakdown, though the former will of course be net of direct tax while the latter will be gross.

As far as the other components in the LUL breakdown are concerned, we would make the following observations:

(a) As argued in Annex B of the Treasury "Green Book", there is no *a priori* reason why the pure WTP component for a mode such as the Underground should have the same numerical magnitude as its DoT road counterpart (see H.M. Treasury, 1991, Annex B). In particular, it will be recalled that these components are intended to reflect the preferences and attitudes to risk of Underground passengers and road users respectively, and there are no grounds for supposing that these preferences and attitudes need necessarily be the same on the two modes.[4]

(b) By contrast, averaged over all Underground fatalities, medical and ambulance costs might not be expected to differ greatly from the corresponding costs averaged over all road fatalities.

4. Material Damage, Police and Administrative Costs and Lost Revenue

So far, we have said nothing about the damage to vehicles and property or the police and administrative costs that might be expected to accompany the typical road fatality; nor

[3] For a brief discussion of the value of preventing non-fatal Underground injuries, see Appendix A.

[4] It is also possible that differences in the perceived levels of the risks concerned may cause people's willingness to pay for risk reduction on the Underground to differ from that on the roads, though the direction of this effect is not unambiguous — see Jones-Lee (1989), p. 38.

have we alluded to the system damage and lost revenue that may result from an Underground accident. The reason for this is that while allowances for avoided material damage, police and administrative costs *are* included in the DoT's value for the prevention of a fatal road *accident*, no such allowances are included in the corresponding figure for a fatal casualty. The DoT's reason for proceeding in this way is that material damage and other things are "... part of the cost of an injury accident but ... are not specific to casualties" (Department of Transport, 1991).

While we are not entirely comfortable with the DoT's distinction on this point, we are prepared to concede that averaging damage, police and administrative costs over fatal injury accidents is slightly less arbitrary than doing so over fatal casualties. In any case, as far as road fatalities are concerned, the avoided material damage, police and administrative cost component of the value of preventing a fatal injury accident is only in the region of £3500, so that excluding this component from the corresponding casualty figure has a relatively minor effect. However, in the case of Underground fatalities matters are rather different. In particular, whereas even a large-scale road accident involving many vehicles will rarely, if ever, result in material damage costs in excess of £1m, LUL has estimated that the system-damage costs and lost revenue caused by the King's Cross fire amounted to some £30m. Furthermore, there are grounds for supposing that over an extended period, the total number of fatalities in large-scale Underground accidents (such as occurred at King's Cross in 1987 or Moorgate in 1975) might reasonably be expected to be similar to the overall number resulting from "small-scale" accidents, each involving only one or two fatalities (see Appendix B). It is therefore clear that averaged over *all* Underground fatalities, system-damage costs on the scale of King's Cross could be very substantial indeed.

Admittedly, given the very rare and quite unique nature of large-scale Underground accidents, it is impossible to predict with any measure of accuracy precisely what system-damage costs and lost revenue might be expected to be. However, in the light of the remarks in the previous paragraph it would not surprise us if, averaged over all Underground fatalities occurring during the next fifty or a hundred years, such costs were to amount to several hundred thousand pounds. Indeed, as a "back-of-the-envelope" calculation, suppose:

(a) that over a protracted period the total number of fatalities in large-scale and small-scale Underground accidents *are* approximately equal;
(b) that large-scale accidents will, on average, involve about 30 fatalities[5]; and
(c) that the system-damage and lost-revenue costs associated with the King's Cross fire can be taken as indicative of what might be expected in a large-scale accident.

On these assumptions, an estimated £30m of system-damage costs and lost revenue will be "spread" over 60 fatalities, giving an average figure of about £500,000 per fatality. Indeed, even if one assumes that large-scale accidents account for only about 20 per cent of all Underground fatalities, the average system-damage and lost-revenue costs would amount to some £200,000 per fatality, though admittedly this figure would fall somewhat

[5] This approximation is obtained from LUL's estimated F–N curve (see Appendix B) treating large-scale accidents as those involving between 10 and 100 fatalities.

if one also averaged the costs over those non-fatal injuries that might be expected to occur in a large-scale accident.

This having been said, we are not necessarily advocating that LUL should break with the DoT convention of averaging damage costs and lost revenue over accidents rather than casualties. What does concern us, however, is that given the possible magnitude of system-damage costs and lost revenue, these costs should not be overlooked in LUL's safety project appraisal. In short, while on a per accident (or per casualty) basis such costs are arguably small beer on the roads, they are almost certainly *not* in the Underground context.

5. Estimating the Pure WTP Component of the Underground Value

Regardless of how one elects to deal with the costs of system damage and lost revenue, the key question in determining a WTP based value of preventing an Underground fatality concerns the appropriate means by which to set about estimating the pure WTP component *per se*. Broadly speaking, two types of approach have been adopted in obtaining empirical estimates of pure WTP values in other contexts. Under the first of these approaches one seeks to observe situations in which people actually do trade off wealth or income for risk as in, for example, labour markets where riskier jobs can be expected to command identifiable wage premiums. Under the other approach one asks members of a representative sample of the affected population more or less directly about their individual willingness to pay for small improvements in their own and possibly others' safety.

It is our considered judgement that neither of these approaches is likely to be workable in the case of Underground risks.

As far as the first approach is concerned, its application would require that passengers should face a choice from a spectrum of *different* safety levels and a corresponding inversely-related fare structure on any particular Underground journey (for example, "very safe and expensive" *vs* "less safe but cheap"), and this is patently not the case, in that on any given journey on the Underground there is just one level of safety and one fare for each type of passenger (such as a one-day visitor to London or a season-ticket holder).

The difficulties with the second approach are more subtle but, we believe, nonetheless real. The first problem is that Underground risks are very small in relation to, say, road risks. This means that respondents in a sample survey would have to be asked about their willingness to pay for minuscule reductions in risk. While one might well get answers to these questions, the process by which such answers are then converted into an estimated value of statistical life would produce results with error bands so large as to render these results virtually worthless.[6]

[6] Thus, suppose that an individual indicates that she would be willing to pay £v for a reduction δp in the probability of being killed in an Underground accident during the coming year. Calculation of a value of statistical life then involves approximating the individual's "marginal rate of substitution" of wealth for risk of death by £v + δp. Suppose further that the person concerned is able to report her "true" willingness to pay only to an accuracy of ±£ε. The error band for the relevant marginal rate of substitution will then be ±£ε + δp which will clearly increase very rapidly as δp gets smaller.

The second difficulty with a process of direct questioning is that Underground passengers are *fare-paying*. Asked about their willingness to pay for safety improvement, passengers are therefore likely to believe that their responses could have a significant influence on future fare levels, so that one might reasonably expect that their answers would be subject to a substantial downward "strategic" bias.[7]

All of this suggests that if one is to form an empirical estimate of the pure WTP component of a value of statistical life for Underground risks it will be necessary to proceed *indirectly*, by trying to measure how members of the travelling public value reductions in the risk of death on the Underground *relative to* reductions in comparable risks on the roads. If such a measure can be obtained then it can be combined with the DoT figure for the roads to estimate the appropriate value to be used by LUL. Thus the main objective of Phase 1 of the study was to explore the feasibility of obtaining such a measure, and give some indication of its possible magnitude.

6. Summary of Phase 1 Fieldwork

Given the nature and complexity of the issues involved, and the importance of eliciting considered responses rather than instant, and therefore probably unreliable reactions, it was decided to organise 12 "focus" groups, each comprising 6 or 7 members of the public who would initially discuss the issues in an open-ended way (which was nevertheless guided by the group moderators) and who would then be presented with more formal quantitative questions at a later stage in the proceedings.

These quantitative questions addressed the two key characteristics of accidents with respect to which the Phase 0 desk study suggested that people's attitudes to safety on the Underground might differ from their attitudes to safety on the roads, namely: the *scale* of possible accidents; and the *context* in which death or injury might be sustained (see Jones-Lee and Loomes, 1992).

The scale issue arises from the fact that whereas only a very small proportion of all deaths on the roads are likely to occur in accidents in which 10 or more people are killed at the same time, such multiple-fatality accidents, although rare, are likely to account for a much higher proportion of all fatalities on the Underground over, say, a 20-30 year period. Some of the literature reviewed in the Phase 0 desk study suggested that reducing the risks of such multiple-fatality accidents might be given a disproportionately heavy weight in people's preferences relative to reductions in the risks of single-fatality accidents. Whether this is true, and if so, whether it is because of the sheer numbers involved, or because of the greater unpredictability associated with such events — or indeed, simply because of the greater media attention that such accidents typically receive — were questions we sought to explore during the discussion phase of each group meeting. Following the discussion phase, focus group members were asked to complete mini-questionnaires which contained at least one question intended to elicit some measure of the relative weights attached to reducing the risks of accidents according to their scale.

[7] For a discussion of strategic bias, see, for example, Cummings *et al.* (1986).

When asking about scale, the question held the context constant; that is to say, the "scale" question referred to different scales of accident *on the Underground.* The other main question, concerning contextual factors, held the scale of the accident constant and tried to focus purely on how people felt about a particular reduction in risk on the Underground *vis-à-vis* a comparable reduction in risk on the roads. The open-ended discussion on the context issue typically covered topics such as people's preferences for [a sense of] greater control over their safety, the degree of voluntariness involved in their decisions about the mode of transport used and the risks incurred, and their feelings about the extent to which responsibility for safety was (or ought to be) distributed differently when comparing public transport with private transport. In addition, it seems likely that many people will be particularly averse to the prospect of being involved in an accident in a remote and confined space deep below the surface of the earth. After the open-ended discussion, the corresponding more formal "context" question in the mini-questionnaire tried to elicit some measure of the relative weights involved.

During November 1992, twelve focus groups were convened. While the participants in these groups could not be regarded as a genuinely random sample of the London travelling public, care was taken to ensure a reasonable spread of ages, gender and social class. In addition, the locations used for these meetings were scattered around the area served by the Underground so as to obtain reasonable coverage of most of the lines.

The first four group meetings, held on 2 and 3 November 1992, each lasted approximately one and a half hours and were largely exploratory, with each group being presented with a somewhat different mini-questionnaire, so that different types of formats of questions could be tried out. The emphasis in these groups was less upon obtaining quantitative data and more upon allowing participants to express their feelings and concerns about safety (to ensure that there were no important elements or considerations that we had overlooked) and to make comments and suggestions about the question formats used and the likelihood that particular forms of question would elicit quantitative responses which were usable for our purposes, and in which we could have some confidence. More detailed discussion of these questionnaires and the responses to them is provided in the Phase 1 report (see Jones-Lee and Loomes, 1993), but particular points which surfaced repeatedly in the discussion in the first four group meetings — and were confirmed or expanded upon in the subsequent eight groups — were the following:

- When asked about the factors that influenced their decisions concerning choice of transport mode, safety was usually not one of the first two or three items to be volunteered by group participants; more frequently it was time, convenience and cost. Furthermore, in almost every group, the first thing that came to participants' minds when asked to think about safety on the Underground was the risk of *assault* and often the second thing was the risk of *bombs.*

- Initially, many participants expressed the view that "there is no limit" to the amount that should be spent to reduce risks and that putting a money value on safety is "callous". However, when encouraged to explore the matter further, many people

saw that they themselves place some limits on what they will do to reduce risk, that resources are not unconstrained, and that hard choices have to be made, though these are found to be more palatable when expressed in "opportunity cost" (for example, time) rather than "money cost" terms.

- The main reason given for using the Underground rather than the roads was the congestion on the latter and/or the difficulty or expense of finding a parking space. Indeed, for these reasons many participants believed that they had no real option but to use the Underground.

- Most of those who would have chosen to travel by car if travel time and cost were the same for both modes readily acknowledged that travelling by car was statistically more dangerous. To some extent this was offset against the greater sense of *control* that people felt when driving, together with the greater convenience and comfort. By contrast, people felt that the responsibility for safety on the Underground rested to a much greater extent with the management of LUL.

- Notions of responsibility and control also appear to have influenced a number of participants in their attitudes to large-scale as opposed to small-scale Underground accidents, though such considerations could work in one of two *opposite* directions. Thus, some people expressed the view that a number of single-fatality Underground accidents might have been avoidable if the victims themselves had taken greater care, whereas multiple-fatality Underground accidents merited extra weight because there was nothing that the victims could do to prevent this kind of accident. On the other hand, a feeling that there was relatively little that *anyone* could do to prevent the unpredictable coincidence of human and mechanical failure that often lay behind catastrophic accidents caused some people to prefer to target safety expenditure on the more "regular" — and in their view, preventable — small-scale types of accident. A related point concerned doubts about getting "value for money" from expenditure aimed at reducing the risk of large-scale accidents. Thus, suppose that large sums were to be spent in the hope of attenuating such risks and that, in the event, no large-scale Underground accidents were to occur. How would one know whether such an accident would *not have occurred anyway* and that the safety expenditure had therefore been effectively "wasted"?

In the final eight group sessions, each of which also lasted about one and a half hours, the preliminary discussion was also open-ended, but more structured than in the first four sessions, making greater use of certain standard (verbal) questions raised by the moderator, as well as exposing participants to tape-recorded "clips" of various (sometimes conflicting) expert opinions about road and rail safety policy. On the basis of the first four groups, the questionnaires had been refined and consisted of one "scale" question, one "context" question, and one "overview" question, as described below. The format used with the four groups held on 11 and 12 November was modified slightly for the remaining

four groups held on 24 and 25 November, but the modifications were small — the basic structure did not change, and in what follows, the responses from all eight groups are pooled.

The "scale" question asked participants to consider two alternative ways of spending extra money on safety on the Underground. The first option was to reduce the risk of multiple-fatality accidents to an extent expected to prevent one such accident occurring during the next 25 years, where such an accident, if it were to occur, would *on average* be likely to involve 25–30 fatalities. The alternative option was to spend the money on reducing the risks of single-fatality Underground accidents. Participants were asked to consider different possible levels of effectiveness of this alternative option, in terms of the total number of fatalities expected to be prevented over the same 25-year period, and to indicate at what levels they would definitely prefer the first option to the second, at what other levels they would definitely prefer the second option to the first, and at what (intermediate) levels — if any — they would be unable to state a clear preference for one or the other. This way of eliciting responses was intended to allow for the possibility that people might not have well-honed, precise preferences "immediately to hand" and might need to be allowed to feel their way towards a sort of "grey indifference" region in a more gradual manner.

For example, someone who placed a premium on reducing the risk of a multiple-fatality accident might continue to prefer the first (multiple-fatality accident prevention) option even if the second option were expected to prevent as many as 35-40 single-fatality accidents during the same period; might be undecided over the range 40-60, but might express a definite preference for switching the money to prevent single-fatality accidents if this were expected to prevent more than 60 fatalities. This sort of pattern of responses is probably most fruitfully thought of as reflecting a kind of "personal confidence interval" for the premium to be accorded to the value of preventing a fatality in a large-scale Underground accident. Accordingly, in summarising the results we have treated such a pattern of responses as entailing a "minimum" scale premium of $[(37.5 \div 27.5) - 1] \times 100 = 36\%$; a "maximum" scale premium of $[(60 \div 27.5) - 1] \times 100 = 118\%$; and a "best estimate" scale premium of $[(50 \div 27.5) - 1] \times 100 = 82\%$.[8]

Alternatively, someone else might prefer the first, multiple-fatality accident prevention option if the second would prevent only 10-15 single-fatality accidents, but might switch to the second option if it were expected to prevent more than 15 deaths, on the grounds that he or she preferred to have money spent on measures which could be seen over time to produce a tangible benefit, rather than invest it in an area where the benefits were uncertain and where one might never really know whether the expenditure had in fact been effective. In this case we would compute the "minimum" scale premium as $[(12.5 \div 27.5) - 1] \times 100 = -55\%$; the "maximum" scale premium as $[(15 \div 27.5) - 1] \times 100 = -45\%$; and the "best estimate" scale premium also as –45%.

Prior expectations, based on the literature reviewed in the Phase 0 report, were that the

[8] In these calculations, 35-40 fatalities were treated as 37.5; 25-30 as 27.5; and so on. "Best estimates" were derived from the mid-point of the range over which the participant was undecided or, if there was no indecision, from the point at which preferences switched from one option to the other.

Table 2

Mean and Median Estimates of the "Scale" Premium

	Mean	*Median*
"Minimum" Scale Premium	–13%	–18%
"Maximum" Scale Premium	35%	9%
"Best Estimate" Scale Premium	11%	–5%

Table 3

Mean and Median Estimates of the "Context" Premium

	Mean	*Median*
"Minimum" Context Premium	38%	0%
"Maximum" Context Premium	79%	20%
"Best Estimate" Context Premium	58%	10%

former attitudes would outweigh the latter, so that overall there would be a marked premium on preventing fatalities in large-scale accidents. And indeed there were a number of participants who responded in this way. But it seemed that there were others whose *initial* reactions were of this kind, but who, upon reflection tended towards the opposite view.

Altogether, 54 people took part in the last eight group meetings and of these, 46 gave usable responses to the "scale" question. And overall, the two opposing views outlined above largely tended to cancel one another out, so that *in aggregate* the premium associated with multiple-fatality accidents was substantially smaller than originally expected. The mean and median "minimum", "maximum" and "best estimate" of the scale premium are reported in Table 2.

The format of the "context" question was broadly similar to the one used in the "scale" question, in that it offered a "fixed" option — spending a sum of money to reduce the risk of single-fatality accidents on the Underground to an extent expected to prevent a total of 25 deaths over the next 25 years — and asked participants to weigh this against the alternative of spending the money to reduce the risk of such accidents on the *roads*, where the expected effectiveness of this alternative was varied.

As in the "scale" question, respondents were asked to indicate those levels of road risk reduction at which they would definitely prefer the Underground option, those levels at which they would definitely prefer the road option, and at what levels they would be unable to state a clear preference. Again, these responses were used to derive "minimum", "maximum" and "best estimates" of the context premium for preventing small-scale fatal Underground accidents relative to the value of preventing small-scale fatal road accidents.

For this question 52 out of 54 responses were usable and yielded the estimates summarised in Table 3.

Finally, the "overview" question simply told participants in round numbers the amount that the DoT was prepared to spend to prevent a road fatality and asked them to say how many times smaller or larger they felt the figure used on the Underground should be. This is, of course, a very crude form of question; but interestingly, most people (47) gave an answer. Only two thought the Underground figure should be lower, eighteen thought it should be the same as the one used for the roads, and the rest thought it should be higher, with the most frequent of the last group of responses suggesting it should be about twice as high. The mean response to this question would imply a premium of about 80 per cent, while the median would give a figure of 100 per cent, which is broadly consistent with the overall premium implied by the combination of the mean answers to the first two questions, namely 69 per cent.

Details of the computation by which the "scale" and "context" premiums should be combined to derive an overall premium are given in Jones-Lee and Loomes (1992). In particular, for a scale premium that is relatively small and a context premium that is relatively large, the implied overall premium is approximately equal to their sum. This result is based on the assumption that, over time, the number of fatalities in large-scale and small-scale Underground accidents are approximately equal (see Appendix B). However, because in the event "scale" effects turned out to be quite small, our results are relatively insensitive to this assumption. Thus, even if large-scale accidents account for only 20 per cent of all Underground fatalities, the overall premium implied by the combination of the mean responses to the first two questions would be about 61 per cent rather than 69 per cent.

7. Conclusion

From Tables 2 and 3 it is clear that for both the "scale" and "context" questions the mean responses substantially exceed the medians, reflecting the fact that the distributions of responses are highly right-skewed, as is commonly the case in quantitative questions related to physical risk (see, for example, Jones-Lee, 1989, chapter 4; Miller and Guria, 1991; and Persson, 1992, chapter 4). Furthermore, given that for both questions the median "best estimate" of the relevant premium does not differ greatly from zero, it is plain that those favouring a positive premium are roughly equal in number to those favouring a discount, but that the *magnitude* of the premium required by the former substantially exceeds the size of the discount sought by the latter.

This immediately raises the question of whether, in an exercise of this sort, one should be guided principally by means or by medians. Given that, as argued in the introductory remarks, the essential purpose of a WTP based value of Underground safety is to reflect customers' *strength* of preference for safety *vis-à-vis* other potential objects of expenditure, it seems clear that the mean, rather than a statistic such as the median, is the appropriate central-tendency measure. Indeed, insofar as a WTP based value of safety is

intended to reflect *aggregate* willingness to pay on the part of those affected by the safety improvement concerned, then the underlying methodology of the WTP approach *requires* that the mean should be employed. In fact, as one of the authors has argued elsewhere (see Jones-Lee, 1989, chapter 4), the only grounds on which one might sensibly contemplate the use of medians within the context of the WTP approach would be if there were serious doubts about the reliability of "extreme" responses, and we have no particular reason to harbour such doubts in the present study.

Focusing, then, on the mean responses to the three quantitative questions posed to participants in the Phase 1 study, it is clear that there is reasonably close correspondence between the overall premium implied by the combination of "best estimate" answers to the "scale" and "context" questions (69 per cent) and the premium implied by the responses to the "overview" question concerning relativity with the DoT value of preventing a road fatality (80 per cent). This suggests that an overall premium in the region of 75 per cent may well be warranted for the value of preventing an Underground fatality in relation to its road counterpart. Applied to the DoT's pure willingness-to-pay component of some £635,000, this would imply a WTP figure for the Underground in the region of £1.1m in 1991 prices. Adding a further £50,000 for avoided medical and ambulance costs and costs of legal compensation would then carry this figure nearer to £1.2m. Indeed, while we would not necessarily advocate that it should do so, were LUL to choose to take account of avoided costs of system damage and lost revenue by adding a further component to its WTP based value to reflect such factors, then the resultant overall figure could well be in the region of £1.7m.

All of this having been said, we must stress that these figures should be treated with some caution. While participants in the focus groups spanned a range of ages and social classes, they were not selected by the kind of procedure necessary to produce a truly random sample. Moreover, there are indications that some groups tended to respond in rather different ways from others, which may reflect the possibility that some members of those groups influenced others to an extent that might not persist if participants were allowed to give further independent thought to the issues and then answer the quantitative questions a little later in one-to-one interviews. In the light of this it is our intention to employ this sort of "two-stage" procedure in the Phase 2 main study.

There is one more point to be made concerning the use of the DoT value of preventing a road fatality as an "anchor" in determining a corresponding figure for the Underground. It may be worth bearing in mind that when the willingness-to-pay approach was first adopted by the DoT in 1988, the figure was deliberately set at the lower end of the range of values that were considered — in order to temper a radical change of methodology with an element of caution — and since then this figure has only been increased in line with inflation and economic growth. It is not inconceivable that the DoT or the Treasury may become amenable to the idea that the figure was originally set *too* low and should be raised for that reason, though such a revision in the near future seems somewhat unlikely. Even so, when considering the figure it wishes to use for its own project appraisal, LUL may prefer to anchor the premium discussed earlier to a higher figure than the arguably conservative £635,000. An anchor of £1m, for example, would produce an overall figure

for the Underground of something around £1.8m excluding avoided system-damage costs and lost revenue, or £2.3m if these were included, given the assumptions and estimates made above, and subject to the caveats already noted.

Finally, we feel bound to highlight the fact that, contrary to our prior expectations, it is "context" rather than "scale" effects that have done the lion's share of the work in producing a premium for the value of Underground safety in relation to the DoT figure for the roads. While we have always been sceptical of the argument — advanced in, for example, Wilson (1975) and Ferreira and Slesin (1976) — that scale, as such, provides a justification for treating the loss of n lives in a catastrophic accident as being significantly worse than the loss of n lives in separate small-scale accidents, we were prepared to entertain the possibility that people's aversion to the *ambiguity* associated with the probability of occurrence of large-scale accidents, together with the *uncertainty* concerning the extent of the consequent loss of life, might provide grounds for a significant "scale" premium (particularly as Kunreuther *et al.* (1992) report clear evidence of ambiguity and uncertainty effects in the context of financial risk). In the event, however, as already noted, such aversion appears to have been largely offset by countervailing doubts about getting demonstrable "value for money" from expenditure aimed at mitigating the risks of large-scale Underground accidents. This suggests that whatever justification there might be for setting the value of statistical life for catastrophe risks at a premium in relation to the value of preventing small-scale accident fatalities, anything other than a very modest premium may prove difficult to defend on the basis of the preferences and attitudes to risk of members of the public, *per se*.

Appendix A

While the Phase 1 study focused principally on Underground fatalities, some thought was also given to the value of preventing non-fatal Underground injuries. In recent work for the DoT the authors derived estimates of the WTP based value of preventing various severities of non-fatal road injury *relative* to the DoT's WTP based value for preventing road fatalities. These estimates were obtained from responses to a nationally representative sample survey — see Jones-Lee *et al.* (1993).

While it is argued in the main text that various factors can be expected to result in *absolute* values of safety that differ markedly as between the Underground and the roads, we have no strong grounds for supposing that the *ratio* of the value of preventing a particular severity of non-fatal injury relative to the corresponding value for a fatality will differ greatly between the two modes. If this is so, then the ratios estimated in the DoT road accident study can be used, at least as a first approximation, as a basis for deriving values for the prevention of non-fatal Underground injuries relative to the Underground figure for fatalities.

According to LUL statistics for 1992, non-fatal Underground injuries break down into "No first aid rendered" injuries (14.5 per cent); "Minor" injuries (82.2 per cent) and "Major" injuries (3.3 per cent). In terms of the DoT's classification, all of the "No first aid rendered" injuries would be treated as "slight", as would many of the "Minor" injuries.

Furthermore, with the exception of injuries incurred in large-scale accidents such as King's Cross, "Major" Underground injuries are apparently rarely worse than those classified as serious but involving no permanent disability in our non-fatal road injury study. Taken as a whole, therefore, it would appear that non-fatal Underground injuries are, in the main, substantially less severe than their road counterparts so that the weighted average ratio of the value of preventing these injuries — relative to the value of preventing an Underground fatality — will almost certainly be somewhat lower than the corresponding weighted average figure for the roads.

Appendix B

If one takes LUL's estimated *F–N* curve (which gives the predicted annual frequency, *F*, of accidents involving *N* or more fatalities) and treats large-scale accidents as those involving between 10 and 100 fatalities, then an approximation to the ratio of fatalities from large-scale and small-scale accidents follows from a straightforward exercise in integral calculus. For details, see Jones-Lee and Loomes (1992). However, if one considers historical accident data for the Underground the picture is somewhat different. Thus, data provided by the Editor for the period 1948-1990 indicate that there were only three accidents involving ten or more fatalities on the Underground during this period, namely at Stratford in 1953 (12 fatalities), Moorgate in 1975 (43 fatalities) and King's Cross in 1987 (31 fatalities). These constitute only about 20 per cent of the estimated total number of Underground fatalities for the period concerned if trespassers and suicides are excluded. The difference between the historical accident record and the implications of LUL's estimated *F-N* curve in part almost certainly reflects the fact that an *F-N* curve represents an attempt to predict and take account of all possible sources of system failure and hazard, including many that have, mercifully, not actually occurred and therefore do not show up in historical accident data.

References

Cummings, R. G., D. S. Brookshire and W. D. Schulze (1986): *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method.* Totowa, N.J., Rowman and Allanheld.

Dalvi, M. Q. (1988): *The Value of Life and Safety: A Search for a Consensus Estimate.* London, Department of Transport.

Department of Transport (1988): *Valuation of Road Accident Fatalities.* London, Department of Transport.

Department of Transport (1991): *Highways Economics Note No. 1.* London, Department of Transport.

Ferreira, J., and L. Slesin (1976): *Observations on the Social Impact of Large Accidents.* Technical Report No. 122, Massachusetts Institute of Technology, Operations Research Centre.

Health and Safety Executive (1992): *The Tolerability of Risk from Nuclear Power Stations.* London, HMSO.

H. M. Treasury (1991): *Economic Appraisal in Central Government. A Technical Guide for Government Departments.* ("The Green Book"); London, HMSO.

Jones-Lee, M. W. (1989): *The Economics of Safety and Physical Risk.* Oxford, Basil Blackwell.

Jones-Lee, M. W. and G. Loomes (1992): *The Monetary Value of Underground Safety Relative to Road Safety: An Exploratory Assessment (The "Phase 0 Report").* Report to London Underground Limited.

Jones-Lee, M. W. and G. Loomes (1993): *The Monetary Value of Underground Safety: Results of a Pilot Study (The "Phase 1 Report")*. Report to London Underground Limited.

Jones-Lee, M. W., G. Loomes, D. O'Reilly and P. Philips (1993): *The Value of Preventing Non-Fatal Road Injuries: Findings of a Willingness-to-Pay National Sample Survey.* Contractor Report 330, Crowthorne, Transport Research Laboratory.

Kunreuther, H., R. Hogarth and J. Meszaros (1992): "Insurer Ambiguity and Market Failure". Wharton Risk and Decision Processes Centre Working Paper 91-12-02.

London Underground Limited (1991): *London Underground Company Plan.* London, London Underground Limited.

Miller, T. R., and J. Guria (1991): *The Value of Statistical Life in New Zealand.* Wellington, N.Z., Land Transport Division, New Zealand Ministry of Transport.

Persson, U. (1992): *Three Economic Approaches to Valuing Benefits of Traffic Safety Measures.* Lund, Swedish Institute for Health Economics.

Wilson, R. (1975): "The Costs of Safety". *New Scientist*, 68, 274-75.

Date of receipt of final manuscript: September 1993