TI 2000-089/3
Tinbergen Institute Discussion Paper

The Value of Statistical Life in Road Safety: A Meta-analysis

A.T. de Blaeij
R.J.G.M. Florax
P. Rietveld
E. Verhoef

published in *Accident Analysis and Prevention* 35(6): 973-986

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Amsterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to wdriessen@few.eur.nl.
Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

The Value of Statistical Life in Road Safety: A Meta-Analysis

Arianne de Blaeij, Raymond J.G.M. Florax, Piet Rietveld, Erik Verhoef

Free University
Department of Spatial Economics
Master-point
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
Phone +31 20 44 46090
Fax +31 20 44 46004
E-mail: blaeij@tinbergen.nl
URL: http://www.econ.vu.nl/re/master-point

Abstract
Accident costs are an important component of external costs of traffic, a substantial part is related to fatal accidents. The evaluation of fatal accident costs crucially depends on the availability of an estimate for the economic value of a statistical life. The aim of the current paper is to present an overview of estimates contained in the empirical literature on the economic valuation of statistical life in road safety. Meta-analysis is used to determine which variables are appropriate to explain the variance of the value of statistical life (VOSL) estimates. The analysis shows, among other things, that the magnitude of estimates of the VOSL depends on the research method, as there is a significant difference between stated and revealed preference studies. It also shows that VOSL estimates cannot simply be averaged over studies, as the magnitude of a VOSL estimate is directly related to the initial level of risk to be caught up in a fatal traffic accident as well as the risk decline implied by the research set-up.

JEL Classification: D12, D81, I10
Keywords: transport safety, value of statistical life, meta-analysis

1. INTRODUCTION

Traffic accidents are a major issue in transport policies around the world. For example, in Europe approximately 40,000 fatalities occur in traffic accidents every year. In addition, the number of non-fatal accidents amounts to a multiple of this figure. Over the past few decades, the long run trend in road accidents has shown a decreasing number of casualties, even though transport volumes have increased substantially. It may well be, that the implementation of a broad range of safety enhancing measures in vehicles, infrastructure as well as in traffic behavior goes a long way in explaining this relative increase in road safety. Obviously, there is no guarantee that these improvements in accident rates will continue to occur, among other things because the introduction and adoption of additional safety enhancing measures imposes an increasing burden on both household and government budgets. Hence, a cost-benefit approach may help to better understand the economic efficiency of additional outlays on road safety improving policy measures.[1]

Accident costs, comprising both fatal and non-fatal damage costs, make up for an important part of the external costs of traffic. Damage costs include a variety of expenses related to e.g., medical treatment, material and immaterial damage, legal assistance, law enforcement, loss of time, etc. Some of these costs can be measured in monetary terms, as there is a market for these goods, others may not, due to lack of trading opportunities. Furthermore, some of these costs are borne by society as a whole (i.e., the taxpayer), whereas others merely constitute a financial burden to the traffic participant involved in an accident (Haight, 1994). The externality and the merit argument for government intervention are frequently used to justify allocative and regulatory decision-making by the public sector in relation to the provision of road safety — regardless whether monetary or non-monetary externalities are concerned, and whether strict economic or political, cultural or moral arguments

[1] The current paper is concerned with a strict economic approach to the issue of road safety. Obviously, there are also social, cultural and moral dimensions to this issue, but they are discarded in the current analysis.

prevail. It would be desirable if these decisions were based on robust and reliable measures of an appropriately defined, empirically assessable monetary value of road safety (Jones-Lee et al., 1995).

The European Union uses a value of one million EUROs per human life in safety cost-benefit analyses; this is generally referred to as the 'one-million EURO rule' (Despontin et al., 1998). The EURO figure is determined by means of on an output-based approach, or in other words, a gross output loss estimate. It implies that policy measures bringing about the saving of lives at a cost up to one million EURO per person can be justified on narrowly defined economic grounds alone. It should be noted that the use of this value also implies that any policy measure or project that leads to a reduction of one fatality will also result in a reduction of 8 serious injuries, 26 slight injuries, and 211 damage-only accidents. The one-million EURO test does, however, not take into account the willingness to pay to avoid pain and suffering (Despontin et al., 1998).

In Douglas Adam's "The Hitchhiker's Guide to the Galaxy", published in the late 1970s, a computer is confronted with the intriguing question: "What is the meaning of life?" After elaborate and tedious number crunching, the computer returns: "The answer to your question is 42." It is obvious, that the monetary valuation of traffic safety is not an easy task either. In fact, it implies answering an equally intriguing and difficult question, as an estimate of the economic value of a statistical life (VOSL) is needed. Although various methods and approaches have been put forward to estimate the VOSL, some scholars argue that valuing fatal injuries and hence human life is virtually impossible. They maintain that people do not nearly have sufficiently accurate preferences to make a sensible trade-off between road-safety and money. The (perceptions of) changes in risk levels are so small that making the trade-off is very difficult, if not impossible (Hauer, 1994).

Notwithstanding the above, there is an abundant empirical literature on the subject of VOSL in road safety. However, as shown in Table 1, the magnitude of VOSL estimates reported in the literature is vastly different, going all the way from less then 400,000 to 30 million in 1996 U.S. dollars (Cohen, 1980; Persson and Cedervall, 1991). Consequently, there does not seem to be agreement among economists as to an appropriate estimate of the VOSL that may be usefully applied in designing traffic and safety policies. This obviously enhances the difficulty of establishing consensus among policy makers, and complicates decision-making (Despontin et al., 1998). This paper therefore sets out to attain insight into the international literature about the economic valuation of statistical life in road safety. Only the value estimates regarding road safety are taken into account, because there is evidence that the VOSL depends on the context in which it is assessed (e.g., wage-risk, health, etc.; see Miller, 2000). In order to attain insight into the literature, the analytical tool of meta-analysis will be used. Meta-analysis is a methodology, comprising a vast array of statistical techniques, developed in order to systematically analyze differences between outcomes of studies, ultimately leading to a synthesis of the results (see also below).

[TABLE 1]

The remainder of the paper is organized as follows. A clarification of the concept of the statistical value of life is given in Section 2. This section also contains a discussion of various methodological aspects. In Section 3 the tool of meta-analysis to be used in the comparative study will be explained. Section 4 presents the empirical results of a meta-analysis. It explores whether the estimates are homogeneous within and between different groups of estimates. Furthermore, it contains the results of a meta-regression of 71 of the value of statistical life and the results of a meta-regression of a subset consisting of 51 estimates, wherefore the sample size is known. Section 5 contains the results of a meta-regression for a subset consisting of 31 VOSL-estimates obtained using the contingent valuation technique. Section 6 winds up this paper and presents conclusions and directions for future research.

2. THE VALUE OF A STATISTICAL LIFE AND ITS EMPIRICAL ASSESSMENT

For a cost-benefit analysis of traffic safety enhancing measures, estimates of the benefits of lives saved are needed. One might argue that the problem of estimating these benefits may be avoided by simply comparing the costs of policy measures that save a statistical life in transport with the costs in other realms of life where people are subject to risk. The optimal level of investment in traffic safety enhancing measures would then be determined by the marginal costs of saving a statistical life in other realms of life. This approach essentially boils down to a ranking of safety enhancing measures in transport according to the cost per statistical life saved, and the recommendation to continue investments in transport safety enhancement until the costs become higher than observed elsewhere. The valuable point in this approach is that it makes one aware of the issue of consistency of measures between various situations. However, the question of what the appropriate cut-off point of costs per statistical life saved is, remains unsolved.

It is essential to be aware of the fact that value of life studies refer to the value of a *statistical* life. Valuation of a statistical life is the valuation of a change of risk rather than the valuation of the life of a specific individual. The main reason is of course the impossibility to value the life of a specific individual as such. Therefore, the following operationalization in terms of the valuation of changes in risk is generally applied. The statement that the risk on a fatal accident is 1:100,000 implies that statistically there is one death per 100,000 people each year. Changes in the aforementioned risk level imply changes in the number of statistical lives saved, and can be given an economic value. The latter is essentially the marginal rate of substitution (MRS) of wealth for risk of death, due to any specific cause. The statistical value of life is then merely the average of a series of observations on this MRS, with the latter taken to be an appropriate estimator of the underlying (unobserved) population mean.

There is a widely shared opinion among economists that the monetary value of safety in public sector cost-benefit analyses should be determined in such a way that it reflects the preferences of those affected by a certain policy measure. This value should be expressed as the aggregate of the individuals' willingness to pay (WTP) for safety improvements, or alternatively the willingness to accept (WTA) compensation for increased risk levels. WTP (and WTA) values are individual trade-offs in terms of expenditures aimed at improving safety versus alternative types of consumption. Consequently, WTP (and WTA) values are explicitly intended to reflect the preferences, perceptions, and attitudes to risk of those members of the public that will be affected by the decisions in which the values are to be used. This implies that the WTP for a risk decrease can differ among hazardous situations.

The marginal WTP for a given reduction in probability of a fatal accident is an increasing function of the initial risk level. This is the standard hypothesis concerning the relationship between an individual's MRS, in most cases estimated by either the WTP or the WTA, and the size and nature of the risk reduction. More specifically, the total WTP is an increasing, strictly concave function of the reduction in probability of death during a forthcoming period (Jones-Lee, 1976, 1989). Consequently, the marginal WTP for reductions will be declining if the risk level declines.

The general form of the function representing the relationship between marginal WTP and risk level is depicted in Figure 1. This function can be interpreted as a demand function for safety. Risk levels can be measured as the percentage of safety, or equivalently as the probability on a fatal accident per 100,000 inhabitants. The WTP to avoid death with a probability of one equals the area under the marginal WTP curve. The VOSL is often calculated as the marginal willingness to pay for reductions in risk times the inverse of the risk reduction considered, which at the initial risk level P_0 yields the gray rectangle as VOSL.[2] Typically, the VOSL estimate depends directly on the initial risk level. Only under the assumption of a constant marginal WTP for reductions in risk levels along the full axis, this would not be the case. With a declining demand function, the willingness to pay to avoid

[2] Note that the VOSL is the population mean of the marginal rate of substitution.

a fatal accident will probably be larger than the VOSL, given that the marginal WTP will often be evaluated at the right hand side of Figure 1.

[FIGURE 1]

In the road safety valuation literature the initial probability of a fatal accident is typically low, between 5 and 50 fatal injuries per 100,000 inhabitants per year (McDaniels, 1992; Persson and Cedervall, 1991). At low risk levels, the function may be nearly horizontal. This implies that small differences in initial risk level among studies will not influence the VOSL estimates. This is implicitly a crucial assumption in the empirical VOSL literature.

The WTP (or WTA) can be empirically assessed by means of revealed or stated preference methods (Carthy et al., 1999; Jones-Lee and Loomis, 1995; Jones-Lee et al., 1995).[3] In the case of stated preference methods, the so-called contingent valuation method (CVM) is often used. The CV method boils down to more or less directly asking members of a representative sample of the population at risk to state their willingness to pay for a (typically small) hypothetical improvement in their own (and possibly other people's) safety. The respondents indicate their WTP for a risk reduction, e.g., a reduction in risk of a fatal accident from $4/10^5$ to $2/10^5$ in the coming year. The VOSL can then be assessed as the WTP times $1/(4/10^5 - 2/10^5)$. As mentioned above, the value of a statistical life is subsequently estimated as the arithmetic average of the individual marginal rates of substitution.

[3] Alternatively, output-based methods can be used. An example of these types of methods is the Human Capital Method that treats a human being as an element in the production process, and values death or incapacitation in terms of the estimated future loss to that process (occasionally minus the victim's estimated future consumption). As a result, it is possible that a fatal accident constitutes a benefit to society, because the victim costs more than he or she produces (Haight, 1994). It should be noted that the WTP (and WTA) approach does not rely directly on an individual's contribution to GDP, although a higher income may of course lead to a higher WTP.

A fundamental property of CVM is that respondents are asked directly to express their evaluation of a risk reduction in monetary terms. A related approach would be to confront respondents with two or more alternatives in terms of a monetary expenditure for safety and a corresponding risk level, and subsequently to ask the respondent to indicate the preferred alternative. The latter approach is known as the 'stated choice approach'.

An alternative to the approach based on stated preferences is the revealed preference approach. Consumers actually reveal their preferences when making decisions in which risk plays a role (e.g., buying a car with or without an airbag, using safety belts or not, etc.). If sufficient information is available regarding the choice alternatives actually considered by the consumers, the implicit trade-offs determining their behavior, will be revealed.

It is of essence to note that it is not merely individual consumers that reveal preferences through actual behavior; public sector agents implicitly do so as well. For example, governments implementing a safety enhancing measure that amounts to x monetary units per year, and that is intended to lead to a reduction in the expected number of fatal accidents of y, implicitly reveal the valuation of a statistical life to be $z \geq x/y$. As there is a persistent information problem facing governments, it may well be that preferences revealed by individual consumers markedly differ from the implicit individual marginal rates of substitution implied by governmental decisions on allocations. The difference between public and private valuation is one of the potential systematic causes for variation in statistical value of life estimates, and will be empirically investigated in the sequel.

The revealed and stated preference approaches both have their pros and cons; see Lanoie et al. (1995) for a treatment of this issue. The analysis of revealed preference data is often hampered by lack of data on the choice set considered by the actor and the perception of risks of the actor. Moreover, econometric difficulties (such as multicollinearity) may severely hamper the estimation of the trade-offs between money outlays and safety increases. These problems can be circumvented by the use of stated preference data, but then a major problem is that the responses of the interviewed may depend rather strongly on the way in which contextual information is being presented. A more general problem, relevant to both methods, is that traffic risks are very small and that many respondents will have difficulties in dealing with these rather abstract, small probabilities. In this respect, an advantage of stated preference approaches is that the information provided in the questionnaire can be used to guide respondents to a proper understanding of the 'good' to value (i.e., small risks; see Van Kooten, 1999).

Despite these conceptual difficulties, many VOSL estimates have been reported over the last couple of decades. Table 1 presents an annotated overview of the empirical studies of value of statistical life in road safety. Figure 2 shows the mean, and in the case of multiple estimates per study also the highest and the lowest VOSL estimate in the 25 studies that we consider. The mean value of all estimates equals 4.8 million (the median value is 2.8 million) . It is obvious from Table 1 and Figure 2 that there is considerable variation of VOSL estimates within and between studies. Moreover, the studies are mainly concerned with North America (particularly the US), and various countries in Europe, and the majority uses a stated preference approach for valuation. The empirical studies will be analyzed in more detail in the meta-analysis.

[FIGURE 2]

In an earlier exploratory meta-analysis of a subset of the studies that we consider in our sample, Elvik (1995) estimates the mean and median values of statistical life related to road safety and to occupational safety. The data he uses are VOSL estimates derived from studies using either stated and/or revealed preference methods. Taking into account this difference in valuation methods Elvik investigates the differences among different groups of studies and between high, intermediate and low validity studies by splitting up the data according to characteristics of the different groups.[4] He computes the mean VOSL within these groups and compares the means between groups without, however, discussing their respective significance levels. He also shows a graph of the inverse relationship between the level of risk and the value of life. His main conclusions are: the mean VOSL for occupational safety is higher than for transport safety; poorly designed stated preference studies result in higher estimates than more carefully designed studies; estimates of studies with high validity lead to lower variation; and lower risk levels result in higher VOSL estimates. The latter result is especially remarkable as it contradicts our expectation based on Figure 1, and we will come back to it in the context of the meta-analysis.

3. META-ANALYTICAL TOOLS FOR RESEARCH SYNTHESIS

In view of the substantial differences within and between studies reported in the foregoing section, we will now attend to a specific type of analysis in order to explain the variance among the estimates. Obviously, the situation of a rather scattered pattern of estimates is not at all uncommon in applied research. For instance, in agronomy, medicine, and psychology there is often a substantial number of studies available, mostly cast in the same or at least a similar experimental setting, that report effects of a specific treatment, be it a fertilizer, a drug, or a therapy. Over the years a distinct trend towards synthesizing available quantitative evidence can be witnessed in these disciplines. Instead of performing an additional study, attention increasingly focuses on the statistical analysis of research results attained earlier, as a way of either seeking common ground or investigating the causes for the

[4] The studies are grouped according to: seven types of activity (occupation versus transport), sample size (500 or more versus less than 500), median value reported, pretest of questionnaire, type of good (public versus private), inclusion of a rationality test, and WTP versus WTA.

differences observed. Glass (1976) coined the term 'meta-analysis' to refer to this synthesizing of research findings of different studies by means of statistical techniques.[5]

The main feature distinguishing meta-analysis from other types of summarizing techniques, such as state-of-the-art literature reviews, is its statistical nature. Meta-analysis is concerned with the statistical analysis of research results of studies performed previously, and should thus be distinguished from primary and secondary analysis (Glass, 1976).[6] Hunter and Schmidt (1990) succinctly explain the term 'meta-analysis' stating that meta-analysis is the 'analysis of analyses'. It should be recognized that research synthesis by means of narrative reviews, oftentimes using tabulations and graphical presentations, is severely hampered by selective sampling, lack of statistical rigor, and a misleading — although intuitively appealing — inference process.

A bias due to selective sampling may occur because the reviewer in most cases reports on a specific subset of the population of studies available, with the sample selection process in most cases being barely justified. But even if it would be, a bias may nevertheless occur, as editors of journals and books are more likely to accept and publish statistically significant research results, as opposed to insignificant ones (see Card and Krueger, 1995, for an example in economics). Olkin (1990, p. 5) denounces this practice as 'a form of statistical Star Wars'. The impact of this phenomenon may be partly circumvented by also including the fugitive literature (e.g., unpublished research memoranda, theses and dissertations) in a review. In the context of meta-analysis, various methodologies have been developed to assess and correct for publication bias (see Dalhuisen and Florax, 2000, for an overview).

---

[5] The following observation by Hunt (1997, p.13) may illustrate the popularity of meta-analysis: 20 years ago five major scientific databases (ERIC, PsycINFO, Scisearch, SOCIAL SCISEARCH, and MEDLINE) contained no listing of meta-analyses, whereas some five years ago their number already exceeded 3,000.

[6] Primary analysis refers to the (initial) analysis of a newly gathered database, whereas secondary analysis is concerned with a renewed analysis of an already existing database, using a different specification, adding an additional variable, etc.

The severest criticism of reviews is, however, the misleading inference process. In most reviews the author implicitly employs the technique of 'vote counting' (Light and Smith, 1971). In a stylized fashion the vote counting procedure is as follows. The available studies are sorted in three categories: those that yield a positive and significant result, those that yield a negative and significant result, and those that yield insignificant results. Now, the intuitively appealing train of thought is that whenever the proportion of study findings falling into one of the three categories exceeds a pre-fixed level C_0 (e.g., one-third), the characteristics of this group represent the true relationship or effect (Hedges and Olkin, 1985, pp. 48-49). However, Hedges and Olkin (1980) rigorously show that the Type-II errors of this decision procedure (i.e., the effect is nonzero, but the reviewer fails to count the effect as positive and significant) do not cancel. Consequently, the vote-counting method tends on average to lead to the wrong conclusion more often as the number of studies increases. Meta-analysis, may therefore be an attractive, statistically rigorous approach, although it also exhibits specific methodological deficiencies (see below). Most of the meta-analytical studies that have been performed in the area of economics are based on the so-called meta-regression technique.[7] Concisely, a meta-regression is based on some least square estimator of the following relation (Stanley and Jarrell, 1989):

$$y = f(p, x, r, t, l) + \varepsilon \tag{1}$$

where y is a specific effect measure observed in a series of studies, p the specific underlying cause, x moderator variables affecting the cause-effect relationship, and r, t, and l moderator variables representing differences among research designs, time-periods considered, and locations covered by the initial studies. In economics, the latter set of moderator variables has the distinct advantage of

[7] Van den Bergh et al. (1997) and Spierdijk et al. (2000) demonstrate the use of cluster techniques in meta-analysis, which is particularly useful if data on a nominal measurement scale are available.

discriminating between e.g., revealed preference and contingent valuation designs, and of the possibility to include temporal and spatial dynamics. In the context of the current analysis a series of estimates for the statistical value of life is used as the dependent variable. An overview of the set of explanatory variables will be presented in the next section.

Apart from the well-known criticism that meta-analysis is invalidated by trying to compare 'apples and oranges',[8] there are three evident methodological pitfalls (see also Glass et al., 1981). First, sample selection bias, for instance, due to selective sampling on the basis of theoretical framework, date of publication, publication as such, research design, etc. Second, dependence between the observations included in the sample, due to multiple sampling from the same study, dependencies over space and/or time, studies with the same author, etc. Third, heterogeneity among sample observations, which may show up in varying parameters (or heteroscedasticity in a regression context), due to different sample sizes of the initial studies, quality differences among studies, differences in research designs, etc.

The issue of sample selection bias obviously constitutes an important potential problem. It is straightforward to see that there is actually no need to analyze the complete population of studies if no systematic relationship between the sampling process and the effect size (in this case, the VOSL) exists. However, it is still an unsettled — and largely discarded — issue how misspecification testing should be carried out, and what solutions are available in the context of meta-analysis. Smith and Huang (1995) use a logit model to determine the likelihood of sample selection bias by means of the estimated probability of including a study in the meta-sample on the basis of year of publication, published or unpublished, etc. As the sample selection process in the current analysis has been geared towards obtaining the full population of studies, this issue will be ignored in the meta-analysis.

The methodological caveats of dependence and heterogeneity among the observations are not easy to treat with the relatively small sample that is available for the meta-analysis (i.e., 25 studies

[8] This is a fairly naïve critique. Hall et al. (1994, p. 20), for instance, present convincing counter arguments.

with 71 usable VOSL estimates). Although fixed and random effect specifications are a common solution, we have chosen to remedy potential dependence and heterogeneity, among other things, by including variables related to the location and time period that the underlying studies refer to. Heterogeneity in terms of heteroscedasticity is inherently present in meta-analytical samples, as the underlying studies typically differ in sample size. Its disturbing influence can be straightforwardly handled by means of standard econometric techniques, such as an appropriate estimator for a heteroscedastic error structure, or the use of White-adjusted variances. Heterogeneity due to quality differences, differences in research design, and the like are more difficult to handle and we have chosen to remedy the latter by including variables related to the type of valuation method used.

4. AN EMPIRICAL ASSESSMENT OF VOSL

The empirical assessment of the value of statistical life is investigated by means of the studies summarized in Table 1. The starting-point of the sample selection process was the reference list in Elvik (1995). Moreover, a literature search in 'Econbase' based on relevant keywords has been performed, recent volumes of the most relevant journals have been searched, and individual scholars have been approached via E-mail. This extensive search resulted in two additional studies. However, it was impossible to use all studies referenced in Elvik (1995). Two studies were intractable (Jones-Lee, 1977; Person and Cedervall, 1992); some studies did not contain a usable VOSL-estimate (Graham and Vaupel, 1981; Kamerud, 1983, 1988; Muller and Reutzel, 1984; Robertson, 1977; Winston and Mannering, 1984); and some studies used an inaccessible language (Hellquist et al., 1977; Person and Cedervall, 1992). Consequently, the meta-analysis is concerned with 25 studies, resulting in a total of 71 usable estimates. Figure 3 shows the distribution of VOSL estimates in 1996 U.S. dollars.

[FIGURE 3]

An interesting feature of the data to be explored is whether the VOSL estimates are homogeneous within and between different groups of estimates. As to the different groups to be explored we used a categorization into private vs public good, willingness to pay vs to accept, stated vs revealed preference, policy vs science-oriented, and Europe vs the rest of the world.

The dimension private vs public indicates whether the valuation of statistical life is based on the valuation of a risk reducing private (e.g., seat belts) or public good (e.g., road improvement). Economic theory suggests a difference, as due to the free-rider problem for public goods, the VOSL based on private good valuation is expected to be higher. Regarding the distinction between WTP and WTA, earlier studies (Lanoie et al., 1995; McDaniels, 1992) have shown that WTA estimates tend to be higher, as there is no income restriction. Likewise, it is to be expected that stated preference methods will exhibit higher VOSL estimates (see, e.g., Lanoie et al., 1995).

It is less clear whether one should expect a difference among studies that have been performed for policy purposes or for purely scholarly purposes. In the analysis the former type of studies has been distinguished on the basis of the government being the client or the performer of the research.

Finally, it would be ideal to investigate whether there are significant differences in VOSL estimates within and between urban and rural areas, as it is reasonable to assume that the type of accidents differ between urban and rural areas (Wadwha et al., 1998). However, for the underlying studies, no such information is available. Instead, we therefore distinguish between studies for European countries and for countries in the rest of the world.

An ANOVA-type analysis can be carried out (see Hedges, 1994) assuming that the data arise from a series of k independent studies, which can be divided into p disjoint groups on the basis of an a priori defined independent grouping variable. The number of estimates in the groups are labeled m_i, where $i = (1, 2, \ldots, p)$, the jth population value in the ith group is denoted by θ_{ij} and its estimate by T_{ij} with variance v_{ij}. The group mean effect estimate for the ith group $\overline{T}_{i\bullet}$ can then be defined by:

$$\overline{T}_{i\bullet} = \frac{\sum_{j=1}^{m_i} w_{ij} T_{ij}}{\sum_{j=1}^{m_i} w_{ij}} \qquad (2)$$

where the weight w_{ij} is the reciprocal of the variance of T_{ij},

$$w_{ij} = 1/v_{ij}. \qquad (3)$$

The grand mean $\overline{T}_{\bullet\bullet}$ is:

$$\overline{T}_{\bullet\bullet} = \frac{\sum_{i=1}^{p} \sum_{j=1}^{m_i} w_{ij} T_{ij}}{\sum_{i=1}^{p} \sum_{j=1}^{m_i} w_{ij}}. \qquad (4)$$

In most cases the variance is actually unknown and is conveniently replaced by an estimated variance, which is a function of the within-study sample size and the effect size estimate (Hedges, 1994, p. 286-287). In the meta-analysis sample the estimated variance of the VOSL estimates is, however, also unknown or not reported in a large number of the underlying studies. As the estimated variance is inversely proportional to the within-study sample size, we use the latter.[9] Figure 4 shows

[9] In fact we use the square root of the sample size for w_{ij}, and the logarithm of the VOSL as the latter is used in the regression analysis as well.

that this approach is accurate, as the variation in VOSL estimates for small sample sizes is much higher than for larger sample sizes.

[FIGURE 4]

Consequently, a test that there is no variation of the population VOSL estimates within the groups of studies Q_w is given by:

$$Q_w = \sum_{i=1}^{p} \sum_{j=1}^{m_i} w_{ij} \left(T_{ij} - \overline{T}_{i\bullet}\right)^2 . \qquad (5)$$

Under the null hypothesis $\theta_{i1} = \cdots = \theta_{im_i} = \theta_{i\bullet}$ the Q_w test is χ^2distributed with ($m_i - p$) degrees of freedom. Similarly, a between-groups variation test Q_b can be defined as:

$$Q_b = \sum_{i=1}^{p} w_{i\bullet} \left(\overline{T}_{i\bullet} - \overline{T}_{\bullet\bullet}\right)^2 \qquad (6)$$

which under the null hypothesis has a χ^2distribution with (p – 1) degrees of freedom. Finally, total variation Q_t can be computed as $Q_{w1}+Q_{w2}+Q_b$.

In the discussion above we assume that for each study the sample size is given. However, for certain types of studies the notion of sample size does not make sense. For example, revealed preference studies of behaviour of public authorities are not based on samples of consumers. Therefore we start our ANOVA approach by giving an equal weight to all studies. The results are presented in Table 2.

The Table shows that the means of stated and revealed preference studies, political and scientific studies, and studies referring to Europe and the rest of the world, are significantly different. Stated preference studies tend to yield higher estimates than revealed preference studies. A related result is that policy oriented studies tend to be lower than studies carried out in a scientific environment. The background is that studies on measures implemented by the public sector usually concern *lower limits* of the VOSL (if policy measure *x* with the associated VOSL *y* has been implemented the public sector reveals that the critical VOSL level is higher than or equal than *y*; see also Section 2). The bivariate analysis does not point to significant differences in the mean VOSL for valuation of private vs public goods, and willingness to pay vs willingness to accept studies. Table 2 also reveals in a statistical way the heterogeneity that we alluded to above. The *Q*-tests of both the within and between group variation are highly significant.

[TABLE 2]

The next step is an analysis based on a subset of 51 observations where weights can be used in a meaningful way. This is done by removing all revealed preference studies. In order to be able to interpret the effect of the introduction of weights we first give the unweighted results (see Table 3). Deleting the revealed preference studies has a clear impact on the statistical analysis of the means between studies. The impact of the country variable (Europe versus the rest of the world) disappears, and the political versus scientific variable now has a reversed sign: VOSL studies carried out in academia tend to yield lower values than studies that are closer to the political arena. Another difference is that a significant result is found now for the private versus public variable: within this subset the VOSL used in decisions on private measures (e.g., safer cars) is higher than in public policies (e.g., safer roads).

[TABLE 3]

Finally we present the results of the weighted analysis (see Table 4). In this case the only remaining significant difference between means of groups is the one between the studies in the private versus public domain.

[TABLE 4]

Although a bivariate analysis leads to useful insights, a multivariate approach is required to systematically explain the variation in the VOSL estimates. A so-called meta-regression is performed, with the variables partly taken from the underlying studies, and partly from additional data that have been gathered. The base specification is taken from the set-up specified in Equation (1). In addition to the variables used in the above analysis, two variables are added on the basis of theoretical considerations: GDP per capita (Jones Lee et al., 1983; Kidholm, 1995; Miller, 2000) and the actual risk level. The latter variable captures the impact of the actual risk level, which differs per country and per year, as theoretically it is expected that a higher initial risk level leads to a higher VOSL (see also Section 2). In many of the contingent valuation studies an initial risk level is given in the study. As this information is not available for all studies, we chose the statistical fatal accident rate as a proxy for the risk variable. Finally, in order to account for temporal differences a variable indicating the year to which the data pertain was added.

Table 5 presents the results of the meta-regression. In the first column with estimates the initial OLS results are reported as well as the results of various misspecification tests. The Jarque-Bera test indicates that the null hypothesis of a normally distributed error term has to be rejected. We therefore report the results of a logarithmic specification, with the dependent as well as the GDP per capita variable transformed (see also Miller, 2000). Although meta-samples are inherently heteroscedastic, due to the variation in sample size of the underlying studies, the results of the White test do not indicate that this is problematic statistically.

The results in the second column indicate that the variables STATED PREFERENCE and POLICY PURPOSE are significant, the former being positive and the latter negative. This indicates that a stated preference study tends to yield higher estimates, and that studies meant for policy purposes result in lower VOSL estimates. Note that similar results were found in the bivariate analysis of Table 2. It is interesting to observe that the coefficient for the variable ln(GDP PER CAPITA) is rather close to 1. Consequently, ln(VOSL/GDP PER CAPITA) can be used as the dependent variable, which circumvents the potentially problematic transformations to express the VOSL in a common real value. The last column of Table 5 shows that the regression results are comparable. Again stated preference studies result in a significantly higher VOSL estimate, which conforms to a priori expectations, and the variable POLICY PURPOSE is significantly different from zero and negative.

[TABLE 5]

Miller (2000) estimates the income elasticity of VOSL estimates across 68 studies from different fields, transport being one of them. The studies are concerned with wage-risk, contingent valuation surveys, and consumer behavior, and show an income elasticity ranging between 0.85 and 0.96. The wage-risk studies and contingent valuation surveys yield roughly comparable VOSLs, both of which are significantly higher than the VOSLs from consumer behavior studies. Table 5 shows that an income elasticity estimate based on road safety studies leads to a somewhat higher estimated income elasticity of 1.33. Subsequently, Miller (2000) uses the estimated income elasticity to come up with VOSLs for most developing and developed countries in the world, a procedure that is generally called ‘value’ or ‘benefit transfer’ (see Brouwer, 2000).

Finally we present the results of a meta-regression on the subset of stated preference studies, both unweighted and weighted (Table 6). In the weighted version we find significant results for the PRIVATE GOOD variable (similar to the bivariate analysis) and country type (VOSL estimates are higher in Europe as compared to other countries, ceteris paribus). It also appears that the actual risk level in a country has a positive impact on the VOSL: in high risk countries respondents appear to attach higher value to the VOSL. This is in line with the discussion in Section 2, where it was demonstrated that the current risk level may play a role in people's WTP for a certain reduction in risk. However, this calls for a deeper analysis since in stated preference studies it is not the *actual* risk but the *initial* risk that has to be considered. This will be the subject of the next section.

[TABLE 6]

5. THE VOSL IN RELATION TO THE INITAL RISK LEVEL

It was hypothesized in Section 2 that the WTP for a given reduction in probability is an increasing function of the initial risk level. This implies that a greater risk decline proposed in the valuation question leads to a declining WTP per statistical life. Close observation of the studies in the database reveals exactly this hypothesized pattern. For example, in Jones-Lee et al. (1983) the initial risk level is 10 per 100,000, and a risk decline of 2 per 100,000 results in a VOSL estimate of 2,210,000 British Pounds, whereas a risk reduction of 5 per 100,000 leads to a VOSL of 1,430,000 Pounds. Similar results are obtained for the data provided in Desaigues and Rabl (1995). Figure 5 shows the marginal WTP and risk levels computed on the basis of the data in the latter study.

[FIGURE 5]

The explanation for the declining VOSL with higher risk declines is that a different level of risk decline is not perceived as different when the valuation question is answered. Although this may be an explanation for the observed differences within studies, this is not the case between studies. Between studies differences are likely to result from differences in initial risk level. In order to investigate whether the pattern is common to all stated preference studies we investigate a subset of studies, with 33 usable CVM estimates, for which the initial risk level is known (risk decline is known for all CVM estimates).

[TABLE 7]

The independent variables in the regression are concerned with the initial risk level and the level of risk decline. The initial risk level refers to the level mentioned in the study,[10] and the risk decline is mentioned in the valuation questions. In addition, we add as independent variables the significant variables in the earlier regressions, reported in Table 1-6.

The results show that the variables INITIAL RISK and RISK DECLINE have the expected sign, and they are significantly different from zero at least at the 0.05-level. This indicates that the initial risk level matters, even if it is relatively low. Hence, the estimated VOSL depends on the initial risk level considered, implying that in terms of Figure 1 most studies are in the range for which the marginal WTP is still falling. These results also imply that the value transfer approach as proposed for example by Miller (2000), should not only take into account the level of GDP per capita, but also the initial risk level.

[10] The reader should note that this is different and in fact more precise than the variable ACTUAL RISK used in the earlier regressions.

6. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

From the 1970s onwards VOSL in road safety has been studied using stated and revealed preference methods. These studies were carried out in different countries and with respect to different years, resulting in a wide range of estimated values. We used meta-analysis to determine whether factors exist that systematically affect the value of life estimates.

Revealed preference studies lead to lower estimates than stated preference studies. An explanation of this outcome is that revealed preference studies tend to report lower limits of VOSL since they are based on measured implemented, not on measures that are discarded. Within the group of stated preference studies the WTP for risk reduction appears to be higher in the case of private goods (for example cars) compared with public goods (for example roads). Another conclusion is that the VOSL depends on the current risk level.

To assume that a single (average or mean) VOSL can be attained, as is frequently suggested in the literature as well as among policy-makers, is not sound from a theoretical perspective. As we have shown, the same conclusion can be drawn regarding the empirics of VOSL estimations. We have shown that the VOSL depends on the initial risk level regarding a (fatal) accident and on the risk decline that is being considered.

This obviously has implications for the validity of value transfer approaches. Short-cut methods to obtaining VOSL estimates over space and/or time are biased if they are solely based on, for instance, a measure such as GDP per capita. Furthermore, these results also create challenges for future VOSL studies. Care should be taken to very precisely transmit to respondents what initial risk level is being considered, and what the precise risk decline is.

Finally, our results raise the interesting question whether public investments in road safety — as well as policies affecting private investments — should be based on the ex ante benefits, as they result from a reduction in risk for the entire population affected by the policy measure, or alternatively on ex post values. With a falling marginal WTP to improve road safety, these two indicators may diverge, implying different optimal levels of investments in road safety.

REFERENCES

Atkinson, S.E., and Halvorson, R., 1990. The valuation of Risks to Life: Evidence from the Market for Automobiles. Review of Economics and Statistics 73, 133-136.

Baker, R.F., 1973. A critical analysis of highway safety policy. Accident Analysis and Prevention 5, 295-319.

Beattie, J., Covey, J., Dolan, P., Hopkins, L., Jones-Lee, M., Loomis, G., Pidgeon, N., Robertson A., and Spencer, A., 1998. On the Contingent Valuation of Safety and the Safety of Contingent Valuation: part 1- Caveat Investigator. Journal of Risk and Uncertainty 17, 5-25.

Bergh, J.C.J.M. van den, Button, K.J., Nijkamp, P. and Pepping, G.C., 1997. Meta-Analysis in Environmental Economics. Kluwer Academic Publishers, Dordrecht.

Blomquist, G.C., 1979. Value of Life Saving: Implication of Consumption Activity. Journal of Political Economy 87, 540-558.

Blomquist, G.C., and Miller, T.R., 1992. Values of life and time implied by motorist use of protection equipment. Unpublished manuscript. Department of Economics, University of Kentucky and the Urban Institute, Washington DC.

Button, K.J., 1995. Evaluation of transport externalities: what can we learn using meta-analysis. Regional Studies 29, 507-517.

Card, D., and Krueger, A.B., 1995. Time-Series Minimum Wage Studies: A Meta-Analysis. American Economic Review, Papers and Proceedings 85.

Carthy, T., Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N. and Spencer, A., 1999. On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 2- The CV/SG “Chained” Approach. Journal of Risk and Uncertainty 17(3), 187-213.

Cohen, B.L., 1980. Society’s valuation of life saving in radiation protection and other contexts. Health Physiscs 38, 33-51.

Dalhuisen, J.M., and Florax, R.J.G.M., 2000. Methodological Pitfalls in Environmental Economic Meta-Analyses: Publication Bias. In: Florax, R.J.G.M., Nijkamp, P. and Willis, K. (Eds.). Comparative Environmental Economic Assessment. Edward Elgar, Aldershot**,** forthcoming.

Desaigues, B., and Rabl, A., 1995. Reference Values for Human Life: an Econometric Analysis of a Contingent Valuation in France. In: Schwab Christe, N.G. and Soguel, N.C., (Eds.). Contingent Valuation, Transport Safety and the Value of Life. MASS, Kluwer Academic Publishers. Boston, pp. 85-112.

Despontin, M., Brucker, K. de, Coeck, C. and Verbeke, A., 1998. The economic evaluation of road safety in the European Union.

Dreyfus, M.K., and Viscusi, W.K., 1995. Rates of Time Preferences and Consumer Valuation in Automobile Safety and Fuel Efficiency. Journal of Law and Economics 38, 79-105.

Elvik, R., 1995. A meta-analysis of value of life estimates for occupational and transport safety. (Mimeo)

Florax, R.J.G.M., 1999. Research Synthesis in Economics: The Staggering Pile of Evidence. Department of Spatial Economics, Master-point, Free University, Working Paper, Amsterdam.

Ghosh, D., Lees, D. and Seal, W., 1975. Optimal motorway speed and some Valuations of Time and Life. The Manchester School 43, 134-143.

Glass, G.V., 1976. Primary, secondary and meta-analysis of research. Educational Research 5, 3-8.

Glass, G.V., McGaw, B. and Smith, M.L., 1981. Meta-Analysis in Social Research. Sage Publications, Beverly Hills.
Graham, J.D. and Vaupel, J.W., 1981. Value of Life: What difference does It Make? Risk Analysis, 1, 89-95 and Kamerud, D.B. (1983) The 55 MPH Speed Limit: Costs, Benefits and Implied Trade-Offs. Transportation Research, Series A 17A, 1, 51-64.
Haight, F.A., 1994. Problems in estimate comparative costs of safety and mobility. Journal of Transport Economics and Policy 28, 7-30.
Hall, J.A., Tickle-Degnen, L., Rosenthal, R. and Mosteller, F., 1994. Hypotheses and problems in research synthesis. In: Cooper, H., and Hedges, L.V. (Eds.). The handbook of research synthesis. Russell Sage Foundation, New York.
Hauer, E., 1994. Can one estimate the value of life or is it better dead than stuck in traffic? Tansportation Research 28(2), 109-118.
Hedges, L.G., and Olkin, I., 1980. Vote counting methods in research synthesis. Psychological Bulletin 88, 359-369.
Hedges, L.G., and Olkin, I., 1985. Statistical Methods for Meta-Analysis. Academic Press, San Diego.
Hellquist, B., Juas, B.;,Karlsson, C., Mattson, B. and Thompson, S., 1977. Värdering av risken för personskador. En jämförande studie av implicita och explicita värden. Karlsatd, Universitetet I Karlstad, Department of Economics, Memorandum 96.
Hunt, M., 1997. How Science Takes Stock: The Story of Meta-Analysis. Russell Sage Foundation, New York.
Hunter, J.E., and Schmidt, F.L., 1990. Methods of Meta-Analysis, Correcting Error and Bias in Research Findings. Sage Publication, London.
IMF, 1998. International financial statistical yearbook, LI.
Johannesson, M., Johansson, P.-O. and O'Conor, R., 1996. The Value of Private Safety Versus the Value of Public Safety. Journal of Risk and Uncertainty 13(3), 263-275.
Jondrow, J., Bowes, M. and Levy, R., 1983. The optimal Speed Limit. Economic Inquiry 21, 325-336.
Jones-Lee, M. W., 1976. The value of life, an economic analysis. Martin Robertson, London.
Jones-Lee, M. W., 1977. Some empirical rigor mortis: An empirical procedure for estimating the value of life from tyre replacement data. Paper presented at the SSRC Health Economists' Study Group, University of Newcastle-Upon-Tyne.
Jones-Lee, M. W., 1989. The economics of safety and physical risk. Basil Blackwell, Oxford.
Jones-Lee, M. W., and Loomis, G., 1995. Scale and Context Effects in the Valuation of Transport Safety. Journal of Risk and Uncertainty 11(3), 183-203.
Jones-Lee, M.W., Hammerton, M. and Abbott, V., 1983. The Value of Transport Safety: Results of a National Sample Survey. Report to the department of Transport. Newcastle-Upon-Tyne, University of Newcastle-Upon-Tyne, Department of Economics.
Jones-Lee, M.W., Hammerton, M. and Phillips , P.R., 1995. Valuing the prevention of Non-Fatal Road Injuries: Contingent Valuation vs Standard Gambles. Oxford Economic Papers 47, 676-695.
Kamerud, D.B., 1983. The 55 mph Speed Limit: Costs, Benefits, and Implied Trade-Offs, Transportation Research-A 17A(1), 51-64.
Kamerud, D.B., 1988. Benefits and Costs of 55mph Speed Limit: New Estimates and Their Implications, Journal of Policy Analysis and Management 7, 341-352.
Kidhom, K., 1995. Assessing the Value of Traffic Safety Using the Contingent Valuation Technique: The Danish Survey. In: Schwab Christe, N.G., and Soguel, N.C. (Eds.). Contingent Valuation, Transport Safety and the Value of Life. MASS, Kluwer Academic Publishers. Boston, pp. 45-62.
Kooten, G.C. van, Krcmar-Nozic, E. and Bulte, E.H., 1999. Fuzzy measures for a fuzzy concept: an alternative approach to nonmarket valuation. FEPA Working Paper, Vancouver: UBC.

Lanoie, P., Pedro, C. and Latour, R., 1995. The Value of a Statistical Life: A Comparison of two Approaches. Journal of Risk and Uncertainty 10, 235-257.
Light, R.J., and Smith, P.V., 1971. Accumulating evidence: procedures for resolving contradictions among different research studies. Harvard Educational Review 41, 429-471.
Maier, G., Gerking, S. and Weiss, P., 1989. The Economics of Traffic Accidents on Australian Roads: Risk Lovers or Policy Deficit? Empirica- Australian Economic Papers 16,177-192.
McDaniels, T.L., 1992. Reference Points, Loss aversion and Contingent Valuation for Auto Safety. Journal of Risk and Uncertainty 5, 187-200.
Melinek, S.J., 1974. A Method of Evaluating Human Life for Economic Purposes. Accident analysis and prevention 6, 103-114.
Miller, T.R., 2000. Variations between Countries in Values of Statistical Life. Journal of Transport Economics and Policy 34(2), 169-188.
Miller, T.R., and Guria, J., 1991. The Value of Statistical Life in New Zealand, Market Research on Road Safety. Land Transport Division, Ministry of Transport. Wellington.
Morrall, J.F. III, 1986. A Review of the Record. Regulation November/December, 25-34.
Muller, A. and Reutzel, T.J., 1984. Willingness to pay For Reduction in Fatality Risk: An Exploratory Survey. American Journal of Public Health, 74, 808-812.
Nijkamp, P., Rietveld, P. and Spierdijk, L., 1999. Classification techniques in quantitative comparitive research: A meta-comparison. Research memorandum 99-56. Department of Spatial Economics, Amsterdam.
OECD, 1999. OECD Statistics, main economic indicators, 4.
Olkin, I., 1990. History and Goals. In: Wachter, K.W., and Straf, M.L. (Eds.). The Future of Meta-Analysis. Russell Sage Foundation, New York.
Persson, U. and Cedervall, M., 1992. Trafiksäkerhet till lägsta möjliga kostnad. Ny samhällsekonomisk-kostnads-effekt-modell. TFB-rapport 1992:12. Stockholm, Transportforskningsberedningen.
Persson, U., and Cedervall, M., 1991. The Value of Risk Reduction: Results of a Swedish Sample Survey. IHE Working Paper 1991(6). The Swedish Institute for Health Economics, Lund.
Persson, U., Lugnér Norinder, A. and Svensson, M., 1995. Valuing the Benefits of Reducing the Risk of Non-Fatal Road Injuries: The Swedish Experience. In: Schwab Christe, N.G., and Soguel, N.C. (Eds.). Contingent Valuation, Transport Safety and the Value of Life. MASS, Kluwer Academic Publishers. Boston, pp. 63-84.
Raudenbush, S.W., 1994. Random Effects Models. In: Cooper, H., and Hedges, L.V. (Eds.). The Handbook of Research Synthesis. Russell Sage Foundation, New York.
Robertson, L.S., 1977. The Effects of Automobile Safety Regulation: Rejoinder, Journal of Economic Issues 11(3), 679-683.
Schwabb Christe, N.G., 1995. The Valuation of Human Costs by the Contingent Valuation Method: The Swiss Experience. In: Schwab Christe, N.G., and Soguel, N.C. (Eds.). Contingent Valuation, Transport Safety and the Value of Life. MASS, Kluwer Academic Publishers. Boston, pp. 19-44.
Smith, V.K., and Desvouges, W.H., 1987 An empirical Analysis of the Economic Value of Risk Changes. Journal of Political Economy 95, 89-114.
Smith, V.K., and Huang, J.C., 1995. Can Hedonic Models Value Air-quality? A Meta-analysis of Hedonic Property Values Models. Journal of Political Economy 103, 209-227.
Spierdijk et al. (2000)
Stanley, T.D., and Jarrell, S.B., 1989. Meta-regression Analysis: a Quantitative Method of Literature Surveys. Journal of Economic Surveys 3, 161-170.
Viscusi, W.K., Magat, W.A. and Huber, J., 1991. Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Trade-Offs for Chronic Bronchitis. Journal of Environmental economics and management 21, 32-51.

Wadhwa, L.C., Eng, D. and FIEAust, 1998. Highway Safety Issues in Rural and Urban Environments, Verification of Community Perception. Journal of International Association of Traffic and Safety Sciences 22(2), 79-85.
Winston, C. and F. Mannering, 1984. Consumer Demand for Automobile Safety. American Economic Review, AEA Papers and Proceedings 74, 316-319.
Winston, C., and Mannering, F., 1984. Consumer Demand for Automobile Safety. American Economic Review, AEA Papers and Proceedings 74, 316-319.

Figure 1: The relationship between the marginal willingness to pay and risk.

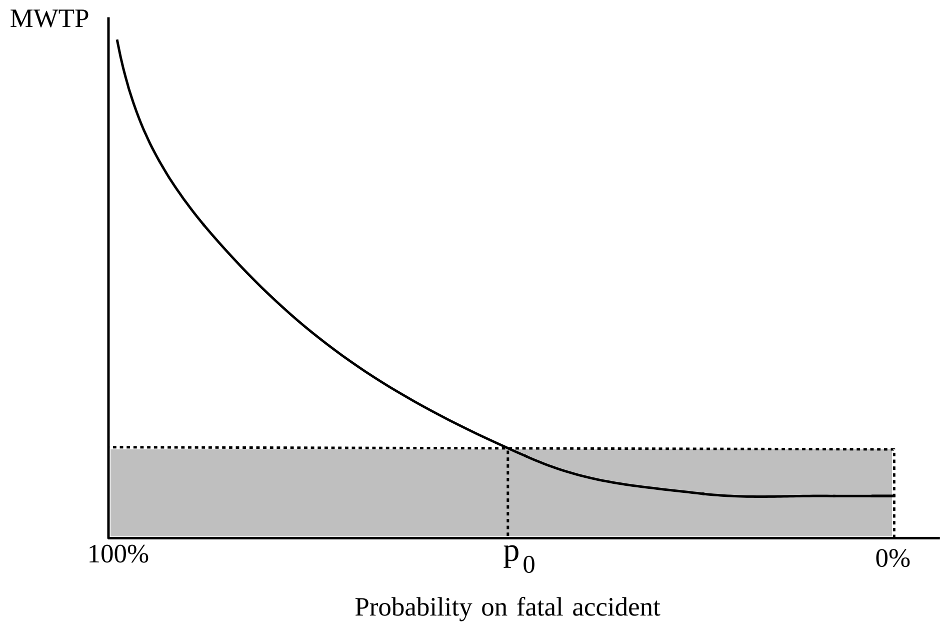

Figure 2: Mean, highest and lowest estimate of the VOSL (× 1,000,000), in 1996 U.S. dollars, in 25 road safety studies.

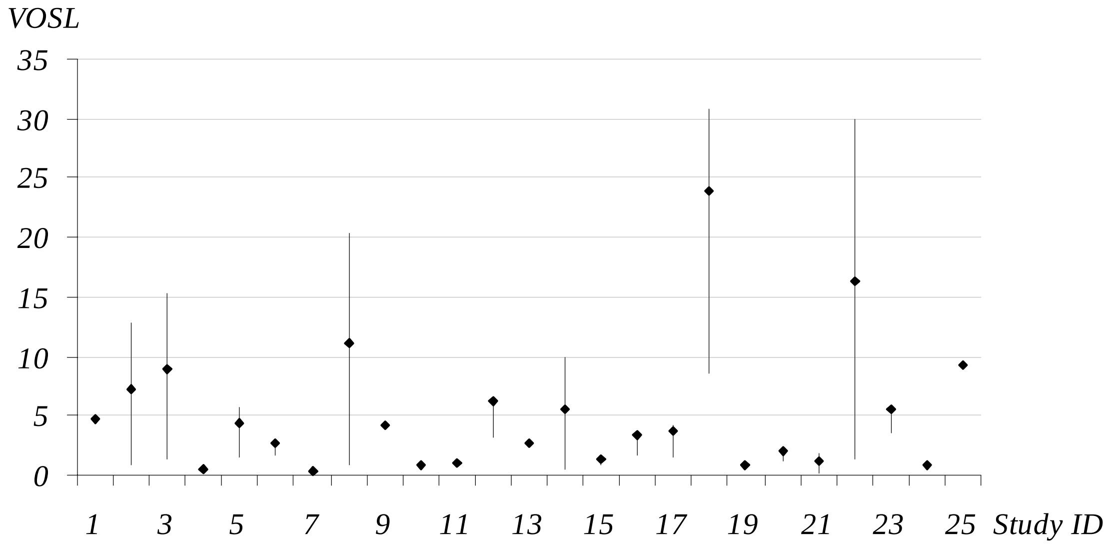

Figure 3: The distribution of VOSL (× 1,000,000) and ln(VOSL) estimates in 1996 U.S. dollars.

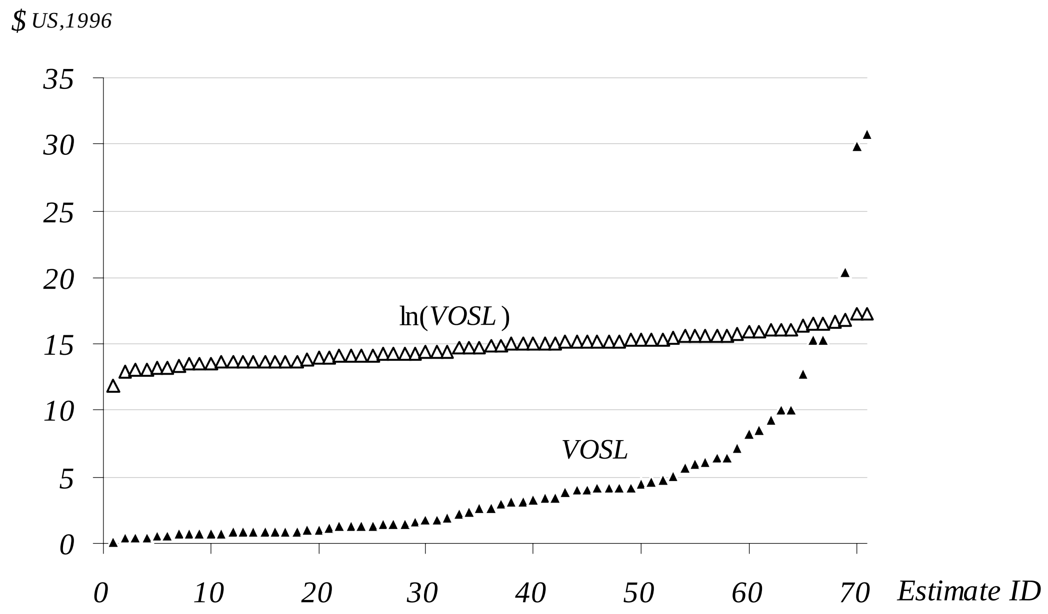

Figure 4: The distribution of VOSL estimates (× 1,000,000), in 1996 U.S. dollars, against sample size.

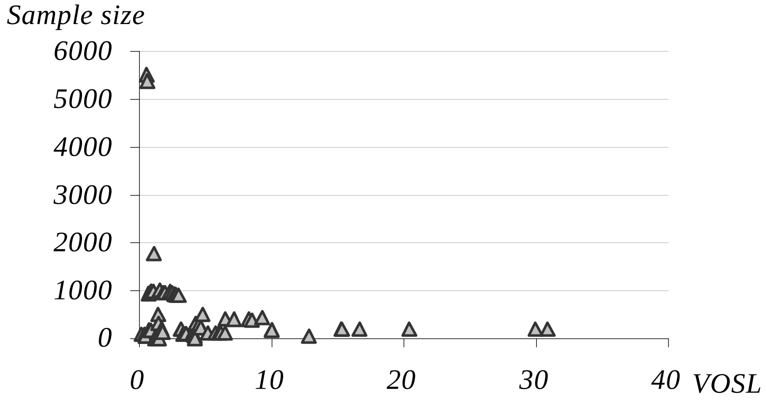

Figure 5: Marginal WTP (in 1996 U.S. dollars, logarithmic scale) versus risk levels.

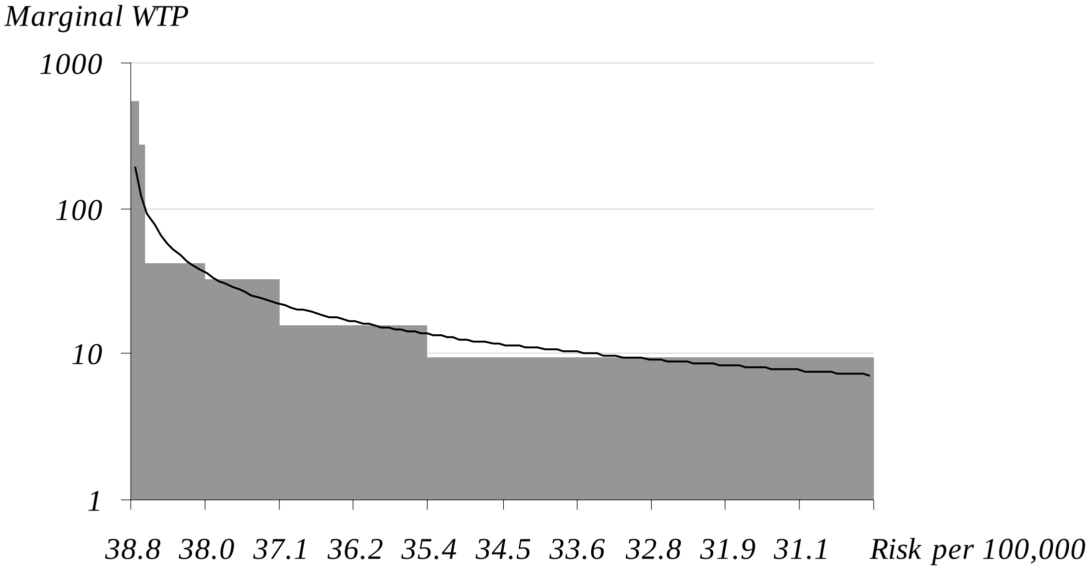

Table 1: An annotated overview of studies with empirical estimates of the value of statistical life (*1000) in road safety.[1]

Authors	Country	Year		Type of Study[3]	Number of estimates	Range of estimates in 1996 US dollars[4]		
		Publication	Data			Single estimate	Lowest estimate	Highest estimate
Atkinson and Halvorsen	U.S.	1990	1986	RP	1	4,636		
Baker	U.S.	1973	1973[2]	RP	4		853	12,797
Beattie et al.	U.K.	1998	1996	SP	4		1,351	15,264
Blomquist	U.S.	1979	1978	RP	1	530		
Blomquist and Miller	U.S.	1992	1987	RP	3		1,481	5,730
Carthy et al.	U.K.	1999	1997	SP	4		1,655	2,154
Cohen	U.S.	1980	1974	RP	1	392		
Desaigues and Rabl	France	1995	1994	SP	6		876	20,377
Dreyfus and Viscusi	U.S.	1995	1987	RP	1	4,159		
Ghosh, Lees and Seal	U.K.	1975	1974	RP	1	838		
Hansen and Scuffham	New Zealand	1994	1994[2]	RP	2		690	718
Johannesson, Johansson and O'Conor	Sweden	1996	1995	SP	2		3,130	6,156
Jondrow, Bowes and Levey	U.S.	1983	1983[2]	RP	1	2,645		
Jones-Lee, Hammerton and Abbott	U.K.	1983	1982	SP	8		585	9,998
Kidholm	Denmark	1995	1993	SP	3		781	1,163
Lanoie, Pedro and Latour	Canada	1995	1986	SP	2		1,756	3,142
Maier, Gerking and Weiss	Austria	1989	1989[2]	SP	6		1,531	4,226
McDaniels	U.S.	1992	1986	SP	2		8,507	30,838
Melinek	U.K.	1974	1974[2]	SP	1	776		
Miller and Guria	New Zealand	1991	1990	SP	5		1,094	1,749
Morrall	U.S.	1986	1984	RP	4		147	1,917
Persson and Cedervall	Sweden	1991	1987	SP	5		1,411	29,932
Persson et al.	Sweden	1995	1993	SP	2		3,514	4,012
Schwab Christe	Switzerland	1995	1993	SP	1	904		
Viscusi, Magat and Huber	U.S.	1991	1991[2]	SP	1	9,270		

[1] Most of the studies are also used in Elvik's (1995) literature review. Two additional studies are Beattie et al. (1998) and Carthy et al. (1999). Some traffic safety studies used by Elvik have been discarded the information needed for the meta-regression analysis was not available.
[2] Refers to the year of the study instead of the year of the data, due to missing information.
[3] SP refers to a stated preference study, and RP to a revealed preference study.
[4] GDP deflators (IMF, 1999) were used to calculate the VOSL in 1996 prices, and PPPs for 1996 (OECD, 1999) to translate local currencies into 1996 U.S. dollars.

Table 2: ANOVA-results for the ln value of statistical life estimates from 25 studies, unweighted.[1]

Group #	Label	Group size	Mean	Mean test[2]	Q_{w1}	Q_{w2}	Q_b	Q_t
1	Private	51	14.859	0.364	64.023*	22.821	1.900	88.742*
2	Public	20	14.495	(1.228)	(67.5)	(30.1)	(3.8)	(90.5)
1	WTP	66	14.727	-0.426	80.967*	6.930	0.844	
2	WTA	5	15.153	(-0.814)	(84.8)	(9. 49)	(3.8)	
1	SP	51	15.057	1.065***	50.116	22.319	16.308***	
2	RP	20	13.991	(3.941)	(67.5)	(30.1)	(3.8)	
1	Political	45	14.511	-0.563**	62.958*	20.563	5.221**	
2	Scientific	26	15.113	(-2.077)	(60.5)	(37.7)	(3.8)	
1	Europe	43	15.016	0.656**	43.198	38.237*	7.307***	
2	Rest	28	14.359	(2.488)	(58.1)	(40.1)	(3.8)	

[1] Significance is indicated by ***, ** and * referring to the 0.01, 0.05, and 0.10 level, respectively.
[2] Difference in mean between groups, *t*-value between brackets
[3] Critical *Q*-values at the 0.05-level in parenthesis

Table 3: ANOVA-results for the ln value of statistical life estimates from 51 SP estimates, unweighted.[1]

Group #	Label	Group size	Mean	Mean test[2]	Q_{w1}	Q_{w2}	Q_b	Q_t
1	Private	36	15.249	0.653**	29.282	16.317	4.517**	50.116
2	Public	15	14.596	(2.203)	(49.8)	(23.7)	(3.84)	(67.5)
1	WTP	46	15.047	0.476	43.135	6.930	0.051	
2	WTA	5	15.153	(-0.224)	(61.7)	(9. 49)	(3.84)	
1	Political	34	14.858	-0.595**	37.180	8.925	4.008**	
2	Scientific	17	14.453	(-2.064)	(47.4)	(26.4)	(3.84)	
1	Europe	41	15.085	0.142	38.996	10.958	0.161	
2	Rest	10	14.943	(0.398)	(55. 8)	(16.9)	(3.84)	

[1] Significance is indicated by ***, ** and * referring to the 0.01, 0.05, and 0.10 level, respectively.
[2] Difference in mean between group, *t*-value between brackets
[3] Critical *Q*-values at the 0.05-level in parenthesis

Table 4: ANOVA-results for the ln value of statistical life estimates from 51 SP estimates, weighted.[1]

Group #	Label	Group size	Mean	Mean test[2]	Q_{w1}[3]	Q_{w2}	Q_b	Q_t
1	Private	36	15.317	0.728**	8693***	14437***	2764***	25892
2	Public	15	14.589	(2.476)	(49.8)	(23.7)	(3.84)	(67.5)
1	WTP	46	14.973	0.186	25404***	467***	21***	
2	WTA	5	14.787	(0.206)	(61.7)	(9.49)	(3.84)	
1	Political	34	14.912	-0.432	24619***	807***	467***	
2	Scientific	17	15.344	(-0.937)	(47.4)	(26.3)	(3.84)	
1	Europe	41	15.015	0.491	24290***	1154***	448***	
2	Rest	10	14.524	(0.944)	(55.8)	(16.9)	(3.84)	

[1] Significance is indicated by ***, ** and * referring to the 0.01, 0.05, and 0.10 level, respectively.
[2] Difference in mean between groups, *t*-value between brackets
[3] Critical *Q*-values at the 0.05-level in parenthesis

Table 5: Estimation results for a meta-analysis of 25 studies of the VOSL, based on 71 estimates.[1]

Dependent variable/ Exogenous variable	VOSL	ln(VOSL)	ln(VOSL/GDP per capita)
CONSTANT	0.004×10^{8}	54.18	49.19
	(1.231)	(1.113)	(1.052)
ln(GDP PER CAPITA)[3]	0.485*	1.330	
	(1.860)	(1.617)	
ACTUAL RISK	-69.12	0.039	0.042
	(-0.264)	(0.894)	(0.989)
STATED PREFERENCE[2]	5524**	1.390***	1.351***
	(2.093)	(3.200)	(3.213)
POLICY PURPOSE[2]	-560.2	-0.701**	-0.737**
	(-0.288)	(-2.199)	(-2.421)
PRIVATE GOOD[2]	992.6	0.476	0.466
	(0.523)	(1.518)	(1.501)
WTP[2]	-1631	0.171	0.161
	(-0.539)	(0.342)	(0.325)
YEAR DATA	-192.7	-0.027	-0.023
	(-1.243)	(-1.059)	(-0.988)
EUROPE[2]	701.3	0.698	0.678
	(0.231)	(1.398)	(1.373)
R^2	0.151	0.327	0.331
R^2-adjusted	0.042	0.240	0.257
Overall F-value	1.383	3.761***	4.458***
Log-likelihood	-1203.51	-94.62	-94.71
AIC	34.16	2.919	2.893
SC	34.44	3.206	3.148
Jarque-Bera[4]	160.5***	3.895	4.722*
	(0.000)	(0.143)	(0.094)
White[4]	0.700	0.937	0.492
	(0.761)	(0.512)	(0.874)

[1] Regression coefficients are presented with t-values in parentheses. Significance indicated by ***, ** and * refers to the 0.01, 0.05, and 0.10-level, respectively. In the column labeled VOSL the coefficients are reported × 1,000. AIC is the Akaike Information Criterion, SC the Schwartz Criterion. The Jarque-Bera test is a test on a normally distributed error term, and the White test refers to heteroscedasticity.
[2] Dummy variable with unity for the observations reflected in the name of the variable.
[3] In the first specification GDP per capita is used, in the second specification the logarithmic value is used.
[4] Misspecification tests are presented with probabilities in parentheses.

Table 6: Meta regression (unweighted and weighted) of VOSL/GDP per capita, 14 studies, 51 estimates[1] .

Dependent variable/ Exogenous variable	Ln(VOSL/GDP per capita)	Weighted Ln(VOSL/GDP per capita)
CONSTANT	95.145	89.403
	(1.233)	(1.000)
LN (GDP PER CAPITA)[3]		
ACTUAL RISK	0.047	0.151*
	(0.801)	(1.694)
POLICY PURPOSE[2]	-0.864	-1.081
	(-1.534)	(-1.235)
PRIVATE GOOD[2]	0.730**	1.315***
	(2.129)	(3.377)
WTP[2]	0.317	0.756
	(0.620)	(0.852)

YEAR DATA	-0.046	-0.045
	(-1.181)	(-0.984)
EUROPE[2]	1.031	2.379**
	(1.520)	(2.473)
R^2	0.187	
R^2-adjusted	0.076	
Overall *F*-value	1.685	2.592
Log-likelihood	-65.455	-74.959
AIC	2.841	3.214
SC	3.107	3.479
Jarque-Bera[4]	4.951	16.439
	(0.084)	(0.000)
White[4]	0.809	2.140
	(0.598)	(0.053)

Table 7: Meta regression (unweighted and weighted) of ln VOSL/ GDP per capita, 10 studies, 33 estimates[1]

Dependent variable/ Exogenous variable	Ln(VOSL/GDP per capita)	Ln(VOSL/GDP capita)Weighted
CONSTANT	3.951**	3.616***
	(2.633)	(2.928)
LN (GDP PER CAPITA)[3]		
INITIAL RISK	0.048***	0.060***
	(3.515)	(5.220)
RISK DECLINE	-0.061**	-0.075***
	(-2.577)	(-3.981)
ACTUAL RISK	-0.022	-0.059
	(-0.370)	(-1.049)
POLICY PURPOSE[2]	-0.212	0.094
	(-0.454)	(0.171)
PRIVATE GOOD[2]	1.196**	1.393***
	(2.270)	(3.476)
EUROPE[2]	0.111	0.382
	(0.169)	(0.633)
R^2	0.422	
R^2-adjusted	0.288	
Overall *F*-value	3.158	6.922
Log-likelihood	-41.569	-38.128
AIC	2.944	2.735
SC	3.261	3.052
Jarque-Bera[4]	1.116	1.244
	(0.572)	(0.536)
White[4]	2.870	2.059
	(0.020)	(0.078)