Journal of Benefit-Cost Analysis (2022), 13: 2, 182–197
doi:10.1017/bca.2022.9

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

The Value of Statistical Life: A Meta-Analysis of Meta-Analyses

H. Spencer Banzhaf

Department of Economics, Georgia State University, Atlanta, GA, USA
Email: hsbanzhaf@gsu.edu

Keywords: benefit–cost analysis, meta-analysis, nonmarket valuation, value of statistical life

JEL classifications: J17, D61, H5, Q5

Abstract
The value of statistical life (VSL) is arguably the most important number in benefit–cost analyses of environmental, health, and transportation policies. However, agencies have used a wide range of VSL values. One reason may be the embarrassment of riches when it comes to VSL studies. While meta-analysis is a standard way to synthesize information across studies, we now have multiple competing meta-analyses and reviews. Thus, to analysts, picking one such meta-analysis may feel as hard as picking a single "best study." This article responds by taking the meta-analysis another step, estimating a meta-analysis (or mixture distribution) of seven meta-analyses. The baseline model yields a central VSL of \$8.0 m, with a 90 % confidence interval of \$2.4–\$14.0 m. The provided code allows users to easily change subjective weights on the studies, add new studies, or change adjustments for income, inflation, and latency.

1. Introduction

The value of statistical life (VSL) is arguably the single most important number used in benefit–cost analyses of environmental, health, and transportation policies. For example, the U.S. Environmental Protection Agency (EPA) found that 85 % of monetized benefits from the Clean Air Act are from mortality reductions (U.S. EPA, 2011). Given the importance of this category, overall benefit–cost evaluations and other welfare calculations involving mortality risks will be highly sensitive to the selected VSL.[1]

When choosing a VSL or range of VSLs, analysts must sift through a vast literature of hundreds of empirical studies and numerous commentaries and reviews to find estimates that are (i) up to date, (ii) based on samples representative of the relevant policy contexts, and

[1] Important reviews and discussions of the VSL literature include Ashenfelter (2006); Cropper *et al.* (2011); Kniesner and Viscusi (2019); Viscusi (2012); Viscusi and Aldy (2003). Additionally, Banzhaf (2014) discusses the VSL in historical context.

(iii) scientifically valid. When doing so, they can arrive at different judgments, a fact highlighted by practices of U.S. government agencies. For example, the U.S. Department of Health and Human Services (HHS) (2016) uses a central estimate of $10.8 m ($2019) with low and high values of $5.1 and $16.5 m, respectively.[2] These values are based on a range of estimates from three stated preference (SP) studies dated 2001–2013 and six hedonic wage studies dated 2004–2013 (Robinson & Hammitt, 2016). These values include adjustments for median labor earnings using an income elasticity of 1.0, as well as inflation. The U.S. Department of Transportation (DOT) (2016) uses a similar central estimate of $10.9 m, but one taken from an average of nine hedonic wage studies dating from 1997 to 2003, with a range ("for illustrative purposes") of $6.5–$16.0 m representing the broader literature. DOT uses the same income adjustments as HHS. The U.S. EPA is the only one of the three agencies that uses a formal meta-analysis. It uses a value of $11.2 m with a 90 % confidence interval of $1.8–$26.9 m (U.S. EPA, 2020; BenMAP, 2022).[3] However, even today, these estimates are based on very old studies published between 1974 and 1991, which were first synthesized in (U.S. EPA, 1997).[4] EPA adjusts for changes in income using an income elasticity of 0.4, which is lower than other agencies, but measuring income with per-capita GDP, which has grown much faster than the measure of median earnings used by other agencies.

In short, despite the importance of the parameter, U.S. agencies are basing their VSLs either on judgments summarizing a small number of studies or on data that are grossly out of date. This is all the more surprising given that there have been several meta-analyses published since EPA's effort, together synthesizing over 800 unique estimates from scores of studies (Mrozek & Taylor, 2002; Viscusi & Aldy, 2003; Kochi *et al.,* 2006; Viscusi, 2018). Compared to the ranges adopted by U.S. agencies, these meta-analyses either synthesize more systematically a wider range of studies or use more recent data – or both.

Perhaps one reason for this surprising gap is that we now have an embarrassment of riches when it comes to summarizing VSL studies. With so many to choose from, the process of selecting which meta-analysis to use, and defending that choice, might feel to some analysts almost like picking a single "best study." As with choosing a single study, the choice of a meta-analysis has profound implications for resulting benefit–cost estimates, with estimated mean VSLs varying across meta-analyses from $4.3 to $13.8 m – a factor of 3.2. This discrepancy is driven by differences in the statistical methods used in the meta-analyses, the set of original studies they synthesize, and choices about whether and how to correct for "best practices." Comparing these meta-analyses, many analysts may conclude that, as with the individual studies underlying them, each of them has a bit of something to offer, that no single one is best. Thus, the old problem of selecting a single best study has just been pushed back to the problem of selecting a single best meta-analysis.

[2] All values reported in this article have been converted to 2019 dollars using the US CPI-U. The Supplementary Appendix shows values as taken from the original studies along with the necessary inflation adjustments.

[3] EPA's low and high values were estimated using their procedure of VSL estimates drawn from a Weibull distribution and income elasticities drawn from a triangular distribution with parameters (0.08, 1.00, 0.40).

[4] More recently, the U.S. EPA (2016) attempted to update this meta-analysis using 18 studies published between 1999 and 2013 (with underlying data collected between 1993 and 2006). However, the EPA's Science Advisory Board recommended substantial revisions to this work and it was not adopted. Its current proposal is to continue using the older meta-analysis (U.S. EPA, 2020).

https://doi.org/10.1017/bca.2022.9 Published online by Cambridge University Press

In this article, I suggest a novel way to break this impasse: A "meta-analysis of meta-analyses" yielding a smooth mixture distribution of VSL estimates. Essentially, I place subjective mixture weights on ten models from six recent meta-analyses and reviews of VSL estimates applicable to the USA. I then derive a mixture distribution by, first, randomly drawing one of the nine models (the mixture component) based on the mixture weights and, second, randomly drawing one value from the distribution describing that component's VSL (e.g., a normal distribution with given mean and standard deviation), and, finally, repeating these draws until the simulated mixture distribution approximates its asymptotic distribution.

This approach has four advantages. First, it triangulates on the estimates coming from all these meta-analyses, encapsulating the idea that the truth is probably somewhere in the middle of all of them. Second, it makes explicit, in the form of the mixture weights, the inevitable judgments that experts must make when picking studies. Third, it easily allows sensitivity analysis through changing the weights. And, finally, it generates a smooth probability distribution representing, simultaneously, the variation within and between meta-analyses. This distribution can be incorporated into Monte Carlo simulations of uncertainty or provide percentiles that can be used as upper or lower bounds around a central estimate.

This article illustrates the approach using two sets of subjective judgments about these weights. Inevitably such judgments will depend on the policy context. The hypothetical context considered here is for U.S. policies immediately affecting the lives of average Americans. Policies applicable to international contexts, to young children or the elderly, or to latent effects might involve different judgments.

Finally, to facilitate implementation of this mixture approach and to make it applicable to a wider set of contexts, the Supplementary Appendix provides simple STATA code that users can easily adapt to reflect their own judgment or policy context or to update it with new studies.

2. Empirical estimates of the VSL

Studies empirically estimating VSLs can be grouped into three categories. Hedonic wage studies, the largest group, use labor market data to infer people's willingness to accept on-the-job risks in return for higher wages. Hedonic wage studies have two significant advantages. First, they combine solid data on household wages with objective data about on-the-job risks. Second, they infer VSLs from people's actual tradeoffs between money and mortality risks in real-world decisions, making them especially credible.

On the other hand, transferring these labor market willingness-to-pay (WTP) estimates to public policy contexts raises at least two issues. One consideration is the difference in age between individuals affected in the two contexts. For example, individuals at risk from air pollution tend to be older and in poorer health than the labor market participants in the hedonic studies. Evidence is emerging that potentially would allow analysts to adjust for differences in age (Aldy & Viscusi, 2007, 2008; Aldy & Smyth, 2014). A second consideration is the nature of the risk, whether workplace risks are qualitatively similar to those in the policy context and whether people feel similarly about facing them. If one type of risk is surrounded by a greater feeling of dread or fear, in principle the WTP to avoid it would be greater. For example, Itaoka *et al.* (2006) find evidence that Japanese households would pay more to avoid risks from nuclear power than quantitatively equivalent risks from fossil fuels. Nevertheless, at this time evidence on these issues is too limited to allow researchers to adjust for qualitative differences in the types of risk.

The second group of studies similarly uses decisions by people in consumer markets, such as their choice for automobiles with various safety features or their willingness to tradeoff time and bother to wear seatbelts against the reduced risk, or their willingness to wear bicycle helmets or change batteries in smoke detectors (e.g., Rohlfs *et al.,* 2015). Like hedonic wage studies, these studies have the advantage of being based on real-world decisions. However, one disadvantage is that important components of the data must be imputed by researchers. For example, whereas wages in a hedonic wage study are readily observed, the time taken to click a seatbelt or put on a helmet, and the value of that time, must be imputed. For this reason, such studies have generally been omitted from meta-analyses.

Third, SP studies use surveys to construct hypothetical markets for mortality risks. SP studies, in turn, generally come in one of two types. Contingent valuation studies describe a single scenario and elicit the WTP for that scenario. Choice experiments offer a series of scenarios varying in risk levels, costs, and potentially other factors and ask respondents to choose which scenario they would prefer. From the patterns in how respondents trade off money for risk levels, analysis can then infer VSLs.

Because what people say they would do does not always match what they actually do, revealed preference studies are generally preferable to SP studies, other things equal. However, SP studies are a widely accepted tool for measuring values for goods and services that are not traded in markets. Moreover, they have the potential to overcome some of the limitations of hedonic wage studies. In particular, they can focus on more policy relevant risks and avoid features peculiar to labor markets, such as worker's comp and life insurance benefits offered through work. They also can focus on the elderly or other populations that may not be earning wages in formal labor markets. Accordingly, some of the meta-analyses considered here make use of SP studies as well as labor market studies.

3. VSL meta-analyses and reviews

Just as there are a variety of ways to estimate a VSL, there are a variety of ways to synthesize multiple VSL studies. The traditional literature review is one way to do this. These reviews compile a bibliography of relevant literature, perhaps providing tables of estimates from the studies and sometimes averages or other simple summary statistics. They highlight differences among studies and critically evaluate their strengths and weaknesses using expert judgment. Formal meta-analyses play a similar role but more quantitatively. They seek to reduce the role of subjective judgment or at least to systemize it, using mathematical models to document patterns in the studies.

Although there are as many ways to conduct a meta-analysis as there are to do any other kind of modeling, we can distinguish between two main categories. First, random effects models view the individual studies as draws from an underlying "meta distribution" (or "mother distribution"). U.S. EPA's (1997) meta-analysis of older studies is one simple example of this approach, as it fitted the individual studies to an underlying Weibull distribution, giving each study equal weight. More sophisticated variants allow each individual study to have a draw from a distribution of sampling distributions, so that their variances as well as their estimates are random variables. Studies with high within-sample

https://doi.org/10.1017/bca.2022.9 Published online by Cambridge University Press

variance provide noisier signals and so are less informative about the meta-distribution than studies with low within-sample variance. The underlying distribution can then be estimated using maximum likelihood or Bayesian methods, as in Kochi *et al.* (2006). The random effects approach is suitable if the individual studies are similar enough to plausibly be draws from a single distribution, or if researchers are completely agnostic about the differences in methods (treating them as part of the randomization process).

Second, meta-regression models view the studies' outcomes as functions of their characteristics, including the characteristics of the sample, statistical procedures, and so forth. Each study's estimated VSL is an observation of the dependent variable, which is regressed on the explanatory factors. Examples include Mrozek and Taylor (2002), Viscusi and Aldy (2003), and Viscusi (2018). Meta-regression is especially useful for understanding patterns in the results across studies. It is also suitable if analysts believe some methods better than others, as it allows them to use all the studies while predicting a summary value at the preferred values of the explanatory variables.

In the remainder of this section, summarized in Table 1, I review five meta-analyses of VSL studies as well as two recent and influential literature reviews. First, in its benefit–cost analysis of the Clean Air Act, the U.S. EPA used a meta-analysis of 26 studies (U.S. EPA, 1997), the latest being from 1991. EPA fitted a Weibull distribution to these data, finding a central estimate of \$9.4 m and a 90 % confidence interval between \$1.3 and \$22.9 m.[5] More recent work by EPA has continued to use this estimate without updating it except for inflation and for growth in real incomes, and the agency plans to continue using it despite the fact that the most recent study is now nearly 25 years old (U.S. EPA, 2020, BenMAP 2022).

Meanwhile, additional meta-analyses have been conducted with U.S. data, incorporating newer studies and/or using newer methods in their analysis.[6] The second is Mrozek and Taylor (2002), who conducted a meta-regression analysis of 91 estimates from U.S.-based hedonic labor market studies. Using linear regression to estimate the effects on VSL estimates of various features of individual studies, they predict what each study's VSL would have been had it used generally accepted best practices, and then average those predicted values to derive a central VSL estimate under such best practices. In Mrozek and Taylor's view, the best practices for purposes of estimating U.S. VSL values include

(i) Controlling for other job characteristics in addition to risk;
(ii) Controlling for an individual's occupation;
(iii) Flexibly modeling the effect of risk;
(iv) Controlling for morbidity effects as well as mortality;
(v) Controlling for union status of the worker;
(vi) Controlling for features of the location of the job such as region and whether it is in a city;

[5] For transparency, in this section all VSLs reported are updated to for inflation (\$2019) but not for income growth, so that the effect of the latter may be kept as a distinct, separate issue.

[6] In this application, I preferred U.S.-only estimates for their applicability to U.S. agencies and to restrict the heterogeneity of estimates. Naturally, other research questions would lead to different judgments. For example, more international meta-analyses that include developing countries are useful for cross-country comparisons or developing income elasticities (e.g., Lindhjem *et al.,* 2011; Majumder & Madheswaran, 2017; Masterman & Viscusi, 2018; Robinson *et al.,* 2019; Viscusi & Masterman, 2017).

https://doi.org/10.1017/bca.2022.9 Published online by Cambridge University Press

Table 1. *Original meta-analyses and parameters.*

Study/Model	Baseline weights	Alternative weights	Distribution	Parameters	Notes
U.S. EPA (1997)	0	$\frac{1}{7}$	Weibull	$\alpha = 1.509588$ $\beta = 9648168$	Parameters reported in BenMAP (2022). Distribution in dollars, not millions
Mrozek and Taylor (2002) – NIOSH Data	$\frac{1}{6} \times \frac{1}{2}$	$\frac{1}{7} \times \frac{1}{2}$	Normal	Mean = 6.48 SD = 2.63	Table 4, NIOSH Panel, Col. 4, $R = 1.0$
Mrozek and Taylor (2002) – BLS Data	$\frac{1}{6} \times \frac{1}{2}$	$\frac{1}{7} \times \frac{1}{2}$	Normal	Mean = 3.70 SD = 1.25	Table 4, BLS Panel, Col. 4, $R = 1.0$
Viscusi and Aldy (2003)	$\frac{1}{6}$	$\frac{1}{7}$	Normal	Mean = 9.06 SD = 1.36	Table 8, Col. 4, prediction for U.S. sample
Kochi *et al.* (2006) – Full set (with imputed SEs)	$\frac{1}{6} \times \frac{1}{2}$	$\frac{1}{7} \times \frac{1}{2}$	Normal	Mean = 5.30 SD = 2.59	Table II, Row 14. Proportionate adjustment for studies with negative values from Ratio of Row 18 to Row 1
Kochi *et al.* (2006) – Only U.S. Hedonic Studies	$\frac{1}{6} \times \frac{1}{2}$	$\frac{1}{7} \times \frac{1}{2}$	Normal	Mean = 5.78 SD = 3.19	Weighted avg. of Table II Rows 2 and 4. Weights calibrated to Rows 1–3. See above adjustment for negative values
Viscusi (2018) – whole sample	$\frac{2}{6} \times \frac{1}{4}$	$\frac{1}{7} \times \frac{1}{4}$	Normal	Mean = 6.67 SD = 0.49	Table 5, Col. 3, Row 8
Viscusi (2018) – best estimates, best practices	$\frac{2}{6} \times \frac{1}{4}$	$\frac{1}{7} \times \frac{1}{4}$	Normal	Mean = 4.74 SD = 1.93	Table 5, Col. 4, Row 15
Viscusi (2018) – best estimates, original practices	$\frac{2}{6} \times \frac{1}{4}$	$\frac{1}{7} \times \frac{1}{4}$	Normal	Mean = 3.93 SD = 1.38	Table 5, Col. 4, Row 8

Table 1. *Continued*

Study/Model	Baseline weights	Alternative weights	Distribution	Parameters	Notes
Viscusi (2018) – whole sample, best practices	$\frac{2}{6} \times \frac{1}{4}$	$\frac{1}{7} \times \frac{1}{4}$	Normal	Mean = 12.27 SD = 0.31	Table 5, Col. 3, Row 15
Robinson and Hammitt (2016)	$\frac{1}{6}$	$\frac{1}{7}$	Triangular	$a = 4.61$ $b = 15.03$ $c = 7.02$	a, b from their suggested range. c calibrated so mean of distribution equals mean of studies synthesized
U.S. Department of Transportation (2016)	0	$\frac{1}{7}$	Triangular	$a = 6.01$ $b = 14.92$ $c = 9.46$	a, b from their suggested range. c calibrated to fit mean

(vii) Using after-tax wages;
(viii) Using a wide range of occupations;
(ix) Using objective rather than self-reported risk measures; and
(x) Using U.S. data.

Mrozek and Taylor provide estimates at different baseline levels of occupation risk. Following their recommendation, I rely on results at risks of 1.0×10^{-4} as representative. They also break down results under two different assumptions about what the best practice would be for sources of risk data, whether from the Bureau of Labor Statistics (BLS) or from the National Institute for Occupational Safety and Health (NIOSH). When controlling for the number of industry classifications controlled for in each study, the central estimate from U.S.-only studies using the BLS data is $3.7 m. Using NIOSH data, their central estimate is $6.5 m.

Third, Viscusi and Aldy (2003) published an extensive literature review of VSL studies which included a meta-analysis of 45 U.S. estimates. They used a similar meta-regression approach as Mrozek and Taylor, predicting the effects of various features of study designs on the estimates. However, they did not predict average VSL under best practice conditions, instead averaging over the actual practices used in the original studies. They estimate the mean VSL to be $9.1 m.[7]

Fourth, Kochi *et al.* (2006) collected 76 hedonic wage and SP studies estimating VSLs in the USA and other high-income countries, which they whittled down to 45 that used best practices. Specifically, they discarded studies if they used actuarial risk data (which includes risks to employees other than those on the job) and SP studies with very small sample sizes. Kochi *et al.* (2006) then synthesized these 45 studies using an empirical Bayes approach, estimating the underlying meta-distribution. In their base model, Kochi, Hubbell, and Kramer estimate the average VSL to be $8.0 m. However, this estimate is less than ideal for our purposes, for three reasons. First, it includes some hedonic wage studies from the UK and other developed countries. In a sensitivity analysis, they find that restricting the hedonic studies to the set from the USA would decrease the mean value by about 7 %. Combining this estimate with the estimate from the SP studies gives a value of $7.6 m.[8] Second, they omitted some studies based on missing standard errors. In a sensitivity analysis imputing those parameters, they obtained a VSL of $7.0 m (vs $8.0 m), suggesting the selection on this set is important. Third, and most importantly, the central estimate excludes estimates with negative VSLs. Although it is implausible that the true VSL is negative, individual estimates should be interpreted as random variables. Viewed as extremely low draws from a distribution, such estimates are not at all implausible. Indeed, they should be treated symmetrically with very high draws, which were not excluded. In a sensitivity analysis including these estimates, the central estimate fell 24 %.

[7] I use Model 4 of Table 8. This model uses robust regression, which reduces uncertainty in the estimate by putting more weight on estimates closer to the central tendency of the data. The model also does not control for practices of the original study, which adds noise and has no advantage if they are not to be controlled for by setting values at best practices.

[8] Based on data from rows 1–3 of Table 11 of Kochi *et al.* (2006)), I estimate the weight on stated preference studies relative to hedonic wage studies in their base model, solving for p in the equation $p^*2.8 + (1-p)^*9.6 = 5.4$. Keeping these weights constant, I then substituted the U.S.-only labor-market estimates (row 4) for the international labor-market estimates (row 3).

https://doi.org/10.1017/bca.2022.9 Published online by Cambridge University Press

Fifth and finally, Viscusi (2018) provides a more recent analysis of 818 estimates from 42 studies using U.S. data. Following Doucouliagos *et al.* (2012), he uses statistical techniques to account for publication bias. These techniques are based on the idea that, across study estimates, standard errors should not be systematically correlated with coefficients. If smaller coefficients observed in published studies tend to have lower standard errors, it suggests the observed data might be selected in favor of statistically significant estimates. Viscusi corrects for this selection by including the standard error as a regressor in his meta-regression and setting the coefficient to zero when predicting an appropriate value. This meta-analysis is particularly useful as it has the most up-to-date set of studies and corrects for publication bias, but the disadvantage of excluding SP studies. In models with U.S. data, Viscusi obtains a mean VSL of $6.7 m when using the entire set of estimates. In an alternative model, he uses a single "best" estimate from each study, restricts the sample to 20 using the most reliable and recent estimates of risk from the Census of Fatal Occupational Injuries (CFOI), and further predicts values under "best practices." The estimated mean VSL in this case is $4.7 m.

In addition to these formal meta-analyses, two relative recent literature reviews have been influential on the practices of U.S. agencies. Robinson and Hammitt (2016) reviewed and synthesized six hedonic wage studies and three SP studies. Their review underlies the VSL estimates used by the US Health and Human Services (HHS) (2016). Like Viscusi (2018), it not only has the advantage of using more up-to-date studies but also has the advantage of including both hedonic and SP studies, to better reflect the breadth of the literature. Based on their review and synthesis, they recommend a VSL range of $4.6–$15.0 m, before adjustments for income growth. The median VSL from their selected studies is $9.3 m and the mean is $8.9 m.

The U.S. Department of Transportation (DOT) (2016) conducted its own review of 15 hedonic wage studies published between 2003 and 2012. They excluded six of these, based on their judgment that the estimates were "implausibly high," used poor estimates of mortality risks, or provided estimates for specific subpopulations, leaving nine remaining. The average VSL from these nine studies is $10.1 m, before adjustments for income growth. Based on a recommendation from Kniesner *et al.* (2012), they suggest a range "for illustrative purposes" of $6.0–$14.9 m.

4. Potential adjustments

Because these meta-analyses and reviews were conducted at different times, with different sets of studies, and because the underlying studies themselves were conducted at different times with different samples, it is natural to consider various adjustments to put them into a common coin. Inflation adjustments are an obvious example, and as stated previously all numbers reported in this article are adjusted to 2019 dollars using the U.S. CPI-U.

Second, if the WTP to avoid mortality risks is increasing in income, then it would also be appropriate to adjust for changes in real income growth over time, using an estimated income elasticity. Although earlier studies estimated elasticities to be closer to 0.5 (Mrozek & Taylor, 2002; Viscusi & Aldy, 2003; Lindhjem *et al.,* 2011), recent research using international cross-sectional comparisons has suggested an income elasticity of 1.0 may be more appropriate, with higher elasticities for lower-income countries (Masterman & Viscusi, 2018; Robinson *et al.,* 2019; Viscusi & Masterman, 2017). How these cross-sectional

https://doi.org/10.1017/bca.2022.9 Published online by Cambridge University Press

estimates relate to temporal growth in real incomes is an open question, as is the relationship between current income and permanent income and the role of within- and between-study variation in income (Kniesner *et al.,* 2010). Costa and Kahn (2004), using U.S. time series variation from 1940 to 1980, estimate the elasticity to be closer to 1.5. Nevertheless, the literature appears to be coalescing around a rule-of-thumb of 1.0. This is a reasonable central value, as by the law of Engel aggregation the weighted average income elasticity for all goods must be 1.0. The U.S. DOT and U.S. HHS both use this elasticity in their benefit–cost analyses, while the U.S. EPA uses a central value of 0.4 and range of 0.08–1.0. In my main analysis, I use an elasticity of 1.0, while testing the sensitivity of this judgment to lower values.

A third potential adjustment is for latency. Many policies enacted at a point in time may have impacts on future mortality rates. For example, recent epidemiological evidence has stressed the importance of long-term chronic exposures to particulate pollution. Accordingly, changes in emissions *now* would show up in mortality rates over a *future* time horizon (as they are part of the lifetime exposure that households accumulate from this time forward). The best evidence suggests the majority of these effects occur within 2–5 years, but then some lingering effects probably occur many years in the future. To account for this delay in effects, the EPA's Science Advisory Board has proposed discounting VSL values by 14 %.[9] Such an adjustment would reduce any VSL estimates proportionately. However, the appropriate adjustment will depend on the specific policy context, so I omit such considerations from the base model.

Potentially, one might also consider adjustments for age and remaining life-years. A simple economic model of age-specific values for risk reduction isolates two potentially offsetting factors (Shepard & Zeckhauser, 1984). First, and most simply, as people age they have fewer life-years remaining. Other things equal, people would be expected to pay more to reduce the risk of dying earlier in their life than later in their life. Second, expenditures generally increase as people enter mid-life. This pattern would tend to increase the WTP to avoid the risk of losing a remaining life-year as people age, potentially offsetting the effect of fewer life-years remaining. While some early work imposed a constant value of a life-year, thus imposing that the VSL declines by age, more recent work has empirically estimated age effects. These empirical studies allow the WTP per life-year to vary by age as incomes and expenditures vary or as attitudes toward risk change. Aldy and Viscusi (2008) estimate VSLs by age, allowing both for changes in WTP per life-year and for changes in life-years remaining as people age. They also allow cohorts to have different WTP at the same age. That is, today's 50-year-olds, for example, need not have the same WTP as today's 60-year-olds did 10 years ago. They find that WTP per life-year increases from age 18 up to about age 55, and only slightly decreases from 55 to 62 (the highest age in their data, which is based on labor-market risks). On the other hand, naturally the number of remaining life-years declines with age. Putting the two effects together, the VSL has an inverted-U shape. It is \$4.6 m at age 18, increases to \$10.6 m at its peak at age 46, and then declines to \$6.9 m at age 62. Aldy and Smyth (2014) calibrate a model allowing one to predict how these effects extrapolate out to older ages. They find that the VSL at age 80 is about one-third that at age 50. Interestingly, evidence from a SP study finds strikingly similar patterns (Cameron & DeShazo, 2013).

[9] Specifically, EPA assumes 30 % of effects are in the 1st year after the emissions, 50 % are spread out evenly in years 2–5, and 20 % are spread out evenly in years 6–20. Applying a discount rate of 5 % leads to the overall 14 % adjustment.

https://doi.org/10.1017/bca.2022.9 Published online by Cambridge University Press

While such adjustments may be important, they will depend on the specific policy context, so I have omitted consideration of them for this work.

5. A mixture of meta-distributions

As previously discussed, a meta-analysis can represent a large range of estimates holistically while, pragmatically, reducing the pressure on an analyst to pick a single best study. However, there are now several meta-studies to choose from, some with separate analyses under different judgments, with a range of values among them from \$4.3 to \$13.8 m after adjusting for income growth. Moreover, each has its characteristic advantages and disadvantages. Thus, it appears that the problem of selecting a single best study has just been pushed back to selecting a single best meta-analysis.

To overcome this challenge, I first assign a subjective weight to each of the studies reviewed in the previous section. In the base model, I give equal weight to each of the peer-reviewed studies, dropping the EPA meta-analysis and DOT review. These weights have the advantage of relying on peer review, of avoiding the problem of reflecting the agencies own decisions back to them, of putting more weight on formal meta-analyses than qualitative reviews, and of discarding the oldest studies which form the basis of EPAs review. By the same logic, I have elected to give double weight to the most recent meta-analysis, Viscusi (2018), which includes a number of more recent original studies that have appeared since the earlier analyses.[10]

For some of these studies, I further subdivide these weights and assign them to separate models reported in the original meta-analysis. In the case of Mrozek and Taylor (2002), I give half the weight to the estimated distribution using NIOSH data on job risks and half to the distribution using BLS data. Viscusi and Aldy (2003) suggest that, for the older studies that they considered, the NIOSH data are more reliable. However, Mrozek and Taylor defend the use of the BLS data, suggesting the opposite is true because of aggregation issues. In the case of Kochi *et al.* (2006), I again assign half the weight to each of two distributions, one substituting U.S.-only hedonic wage studies for the complete set, and one based on a more complete set of studies by imputed standard errors when necessary. However, in both cases I adjust the estimates by the proportionate effect of including original studies with negative estimates ($6.1/8.0 = 0.76$). With this adjustment, the mean values for these two distributions are then \$5.8 and \$5.3 m, respectively. Lastly, in the case of Viscusi (2018), I divide the weight into four components, one using the whole sample, one using only the select “best” estimate for each study, for those studies using only the CFOI data, and in each case with and without predicting the estimates using “best practices.” Table 1 summarizes all the components underlying the resulting mixture distribution.

For each estimate from the four meta-analyses, I use the mean and standard deviation to characterize the distribution of VSL estimates. However, Robinson and Hammitt (2016) do not report these parameters and themselves prefer to think of their synthesis as providing a central estimate and range. In my meta-analysis of meta-analyses, I treat their study as having a triangular distribution, with their suggested range as the two endpoints. I calibrate the mode

[10] These include Aldy and Viscusi (2007); Evans and Schauer (2010); Hersch and Viscusi (2010); Kniesner *et al.* (2010, 2012, 2014); Kochi and Taylor (2011); Scotten (2013); Viscusi (2013); Viscusi and Gentry (2015); Gentry and Viscusi (2016). Another recent study, Lee and Taylor (2019), is included in the Robinson and Hammitt review.

https://doi.org/10.1017/bca.2022.9 Published online by Cambridge University Press

of the triangular distribution so that the mean of the resulting distribution equals the mean value of the studies in their synthesis. These resulting parameters are also summarized in Table 1.

Finally, I adjust for income growth using an income elasticity of 1.0 and using median household income in the USA.[11]

The code supplied in the Supplementary Appendix shows all calculations back to the figures reported in the original meta-analyses, including all inflation and income adjustments. Readers consulting the original studies may find this code a useful reference. The code is user friendly, with user-defined inputs for the component weights assigned to each of the underlying components. It also allows for the introduction of any additional new component, such as a new meta-analysis, so long as it is represented by a normal, Weibull, uniform, or triangular distribution. Finally, it includes parameters that users can set to allow for adjustments based on alternative income elasticities or latency between a policy and mortality response. Thus, in future work, other analysts can easily compute the sensitivity of results to some of the subjective judgments used here.

Combining the above subjective weights given to each underlying component along with the distribution of VSLs within each component, we can derive an overall mixture distribution of VSLs. In this case, the overall distribution has a mean VSL of \$8.0 m in 2019 dollars. Table 2 provides additional details about the percentiles of the distribution, and Figure 1 shows the mixture density function, with the vertical line representing the mean value of the distribution. As seen in the figure, the mixture distribution is diffuse and leptokurtic, with a high density around the mean and lower density in the shoulders.

Table 2. *Summary statistics of resulting mixture distribution.*

Parameter	Value (millions of \$2019)		
Percentiles	Base model	Alternate weights	No income adj.
1st	0.57	0.63	0.51
5th	2.45	2.53	2.17
10th	3.40	3.66	3.01
25th	5.27	6.10	4.68
Median	7.79	8.56	6.91
75th	10.39	11.20	9.29
90th	13.47	13.84	11.91
95th	13.97	15.74	12.37
99th	15.05	24.12	13.05
Moments			
Mean	7.96	8.88	7.06
Standard deviation	3.55	4.45	3.14
Skewness	0.09	1.34	0.06
Kurtosis	2.44	8.90	2.39

[11] https://fred.stlouisfed.org/series/MEHOINUSA672N.

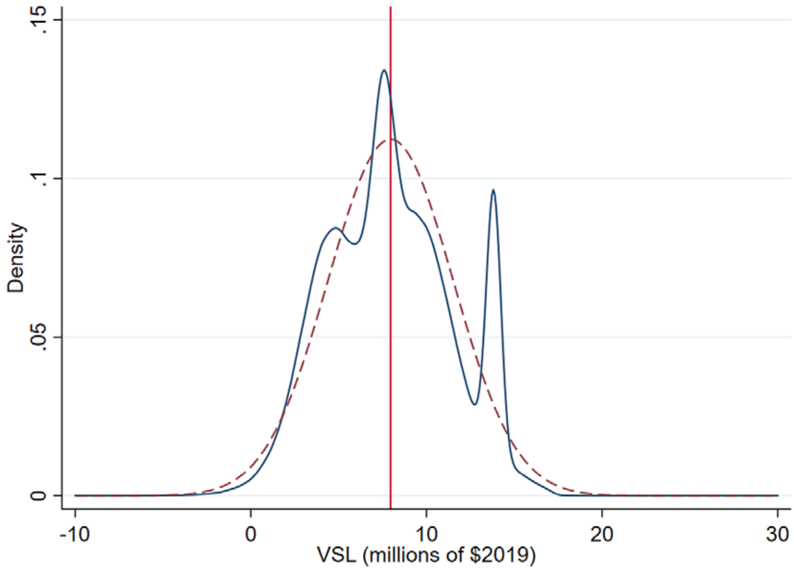

Figure 1. *Estimated mixture distribution, baseline weights.*

Moreover, it is multimodal, with different components providing some local central tendency at different points in the distribution. The 90 % confidence interval ranges from $2.4 to $14.0 m.

6. Sensitivity analyses

The weights used in the base model reflect one set of judgments about the reasonable models that one might use. To see the potential sensitivity of the results to these judgments, I consider an alternative set of weights, as shown in Table 1. For the alternative weights, I include the U.S. EPA (1997) meta-analysis, which used a Weibull distribution, and the U.S. Department of Transportation (DOT) (2016) review, bringing the number of underlying studies up to seven. I also use a more even-handed approach, assigning a weight of 1/7 to each, rather than giving more weight to the recent studies. Finally, I simplify the subdivision of the weights on the submodels presented by Viscusi (2018), using only the whole sample and the model using the "best set" and predicted at "best practices."

These alternative weights increase the mean VSL to $8.9 m, an increase of about 12 % from the base weights, indicating the EPA and DOT estimates are somewhat higher than the peer-reviewed estimates. Interestingly, as summarized in Table 2 (Col. 2) and Figure 2, the other moments of the mixture distribution are also different. The distribution now more closely resembles a normal distribution, though it remains leptokurtic and is positively skewed, with a long right tail. The 90 % confidence interval is $2.5–$15.4 m.

Finally, I consider the sensitivity of the estimates to the adjustment for income growth. In an alternative model, I use an income elasticity of zero, effectively eliminating these adjustments. Some analysts may prefer this more conservative approach, especially given the uncertainty about the role of permanent income versus current income, cross-sectional

https://doi.org/10.1017/bca.2022.9 Published online by Cambridge University Press

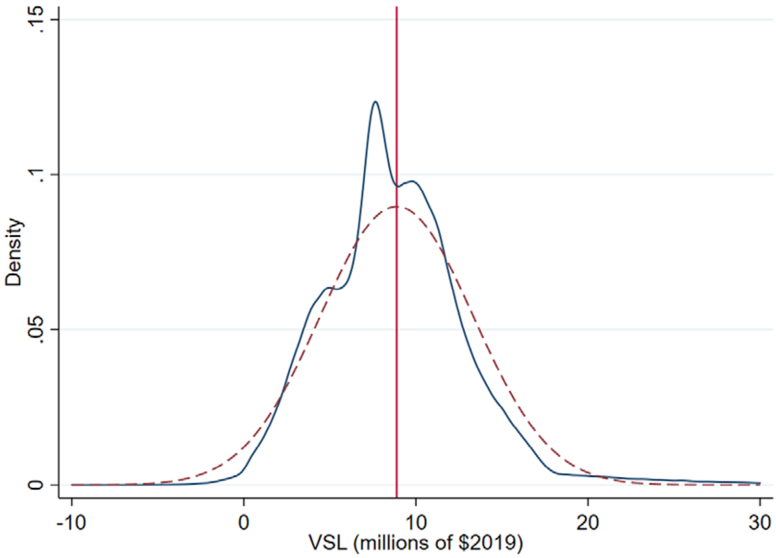

Figure 2. *Estimated mixture distribution, alternative weights.*

variation in income and time series variation, and the measurement of income. Neglecting the income adjustments, the mean VSL of the meta-distribution falls from 8.0 m in the baseline model to 7.1 m, a decrease of about 11 %. Table 2 provides additional details about the distribution.

7. Conclusions

Meta-analyses can synthesize the results from numerous underlying studies estimating common parameters like the VSL. Each meta-analysis estimates a meta-distribution of VSL values from the underlying studies, which are viewed as draws from the meta-distribution. But like the studies underlying them, these meta-analyses themselves involve subjective judgments. Nevertheless, these judgments, while differing, are well informed by the expertise of the analysts making them. This suggests no one probability distribution (as estimated from one particular meta-analysis) is solely correct, but rather each has something to offer.

In this setting, it is reasonable to estimate a mixture distribution or meta-analysis of meta-analyses. Each individual meta-analysis can be thought of as representing one distribution, and these distributions can be mixed into an overarching distribution. Taking this approach, and using the subjective weights described here, yields a central VSL of $8.0 m, with a 90 % confidence interval of $2.4–$14.0 m. The provided code allows users to easily change subjective weights on the studies, add new studies, or change adjustments for income, inflation, and latency.

Future research might consider alternative study weights, adjustments for trends in the underlying data, or adjustments for changing economic environments and preferred research methods.

Supplementary Material. To view supplementary material for this article, please visit http://doi.org/10.1017/bca.2022.9.

Acknowledgements. The idea for this work was originally developed under a contract with William H. Desvousges & Associates and Xcel Energy in 2015, as part of an externalities costing study. I thank Bill Desvousges, Chris Dockins, Lisa Robinson, the editor, and an anonymous referee for comments and discussion.

Competing Interests. The author declares no competing interest.

References

Aldy, Joseph E., and Seamus J. Smyth. 2014. *Heterogeneity in the Value of Life*. NBER Working Paper, 20206. Cambridge, MA: National Bureau of Economic Research. http://www.nber.org/papers/w20206.

Aldy, Joseph E., and W. Kip Viscusi. 2007. "Age Differences in the Value of Statistical Life: Revealed Preference Evidence." *Review of Environmental Economics and Policy*, *1*(2): 241–260.

Aldy, Joseph E., and W. Kip Viscusi. 2008. "Adjusting the Value of a Statistical Life for Age and Cohort Effects." *Review of Economics and Statistics*, *90*: 573–581.

Ashenfelter, Orley. 2006. "Measuring the Value of a Statistical Life: Problems and Prospects." *Economic Journal*, *116*: C10–C23.

Banzhaf, H. Spencer. 2014. "Retrospectives: The Cold-War Origins of the Value of Statistical Life." *Journal of Economic Perspectives*, *28*(4): 213–226.

BenMAP. 2022. *BenMAP: Environmental Benefits Mapping and Analysis Program—Community Edition.* User's Manual. January.

Cameron, Trudy Ann, and J. R. DeShazo. 2013. "Demand for Health Risk Reductions." *Journal of Environmental Economics and Management*, *65*(1): 87–109.

Costa, Dora L., and Matthew E. Kahn. 2004. "Changes in the Value of Life, 1940-1980." *Journal of Risk and Uncertainty 29*(2): 159–80.

Cropper, Maureen, James K. Hammitt, and Lisa A. Robinson. 2011. "Valuing Mortality Risk Reductions: Progress and Challenges." *Annual Review of Resource Economics*, *3*: 313–336.

Doucouliagos, Chris, T. D. Stanley, and Margaret Giles. 2012. "Are Estimates of the Value of a Statistical Life Exaggerated?" *Journal of Health Economics*, *31*(1): 197–206.

Evans, Mary F., and George Schauer. 2010. "A Quantile Estimation Approach to Identify Income and Age Variation in the Value of a Statistical Life." *Journal of Environmental Economics and Management*, *59*(3): 260–270.

Gentry, Elissa Philip, and W. Kip Viscusi. 2016. "The Fatality and Morbidity Components of the Value of Statistical Life." *Journal of Health Economics*, *46*: 90–99.

Hersch, Joni, and W. Kip Viscusi. 2010. "Immigrant Status and the Value of Statistical Life." *Journal of Human Resources*, *45*(3): 749–771.

Itaoka, K., A. Saito, A. Krupnick, W. Adamowicz, and T. Taniguchi. 2006. "The Effect of Risk Characteristics on the Willingness to Pay for Mortality Risk Reductions from Electric Power Generation." *Environmental and Resource Economics*, *33*: 371–398.

Kniesner, Thomas J., and W. Kip Viscusi. 2019. "The Value of Statistical Life." In *Oxford Research Encyclopedia of Economics and Finance*, DOI 10.1093/acrefore/9780190625979.013.138.

Kniesner, Thomas J., W. Kip Viscusi, Christopher Woock, and James P. Ziliak. 2012. "The Value of a Statistical Life: Evidence from Panel Data." *Review of Economics and Statistics*, *94*(1): 74–87.

Kniesner, Thomas J., W. Kip Viscusi, and James P. Ziliak. 2010. "Policy Relevant Heterogeneity in the Value of Statistical Life: New Evidence from Panel Data Quantile Regressions." *Journal of Risk and Uncertainty*, *40*: 15–32.

Kniesner, Thomas J., W. Kip Viscusi, and James P. Ziliak. 2014. "Willingness to Accept Equals Willingness to Pay for Labor Market Estimates of the Value of a Statistical Life." *Journal of Risk and Uncertainty*, *48*(3): 187–205.

Kochi, Ikuho, Bryan Hubbell, and Randall Kramer. 2006. "An Empirical Bayes Approach to Combining and Comparing Estimates of the Value of a Statistical Life for Environmental Policy Analysis." *Environmental and Resource Economics*, *34*: 385–406.

Kochi, Ikuho, and Laura O. Taylor. 2011. "Risk Heterogeneity and the Value of Reducing Fatal Risks: Further Market-Based Evidence." *Journal of Benefit-Cost Analysis*, *2*(3): 1–28.

Lee, Jonathan M., and Laura O. Taylor. 2019. "Randomized Safety Inspections and Risk Exposure on the Job: Quasi-Experimental Estimates of the Value of a Statistical Life." *American Economic Journal: Economic Policy*, *11*(4): 350–374.

Lindhjem, Henrik, Ståle Navrud, Nils Axel Braathen, and Vincent Biausque. 2011. "Valuing Mortality Risk Reductions from Environmental, Transport, and Health Policies: A Global Meta-Analysis of Stated Preference Studies." *Risk Analysis*, *31*(9): 1381–1407.

Majumder, Agamoni, and S. Madheswaran. 2017. "Compensating Wage Differential and the Value of Statistical Life." *Indian Journal of Labor Economics*, *60*: 527–548.

Masterman, Clayton J., and W. Kip Viscusi. 2018. "The Income Elasticity of Global Values of a Statistical Life: Stated Preference Evidence." *Journal of Benefit-Cost Analysis*, *9*(3): 407–434.

Mrozek, Janusz R., and Laura O. Taylor. 2002. "What Determines the Value of Life? A Meta-Analysis." *Journal of Policy Analysis and Management*, *21*: 253–270.

Robinson, Lisa A., and James K. Hammitt. 2016. "Valuing Reductions in Fatal Illness Risks: Implications of Recent Research." *Health Economics*, *25*: 1039–1052.

Robinson, Lisa A., James K. Hammitt, and Lucy O'Keeffe. 2019. "Valuing Mortality Risk Reductions in Global Benefit-Cost Analysis." *Journal of Benefit-Cost Analysis*, *10*(S1): 15–50.

Rohlfs, Chris, Ryan Sullivan, and Thomas Kniesner. 2015. "New Estimates of the Value of a Statistical Life Using Air Bag Regulations as a Quasi-Experiment." *American Economic Journal: Economic Policy*, *7*(1): 331–359.

Scotten, Carol R. 2013. "New Risk Rates, Inter-Industry Differentials and the Magnitude of VSL Estimates." *Journal of Benefit-Cost Analysis*, *4*(1): 39–80.

Shepard, D.S., and R.J. Zeckhauser. 1984. "Survival versus Consumption." *Management Science*, *30*: 423–439.

U.S. Department of Transportation (DOT). 2016. *Revised Departmental Guidance 2016: Treatment of the Value of Preventing Fatalities and Injuries in Preparing Economic Analyzes.* https://www.transportation.gov/sites/dot.gov/files/docs/2016%20Revised%20Value%20of%20a%20Statistical%20Life%20Guidance.pdf.

US Environmental Protection Agency (EPA). 1997. *The Benefits and Costs of the Clean Air Act from 1970 to 1990.* Report to Congress EPA-410-R-97-002.

US Environmental Protection Agency (EPA). 2011. *The Benefits and Costs of the Clean Air Act from 1990 to 2020*. Final Report, Office of Air and Radiation.

US Environmental Protection Agency (EPA). 2016. *Valuing Mortality Risk Reductions for Policy: A Meta Analytic Approach*. White Paper prepared by the US EPA's Office of Policy for review by the EPA's Science Advisory Board, Environmental Economics Advisory Committee.

US Environmental Protection Agency (EPA). 2020. *Guidelines for Preparing Economic Analyses.* Review Copy Prepared for EPA's Science Advisory Board's Economic Guidelines Review Panel.

US Department of Health and Human Services (HHS). 2016. *Guidelines for Regulatory Impact Analyses.* https://aspe.hhs.gov/system/files/pdf/242926/HHS_RIAGuidance.pdf.

Viscusi, W. Kip. 2012. "What's to Know? Puzzles in the Literature on the Value of Statistical Life." *Journal of Economic Surveys*, *26*: 763–768.

Viscusi, W. Kip. 2013. "Using data from the Census of Fatal Occupational Injuries to estimate the value of a statistical life." *Monthly Labor Review*, *136*: 1–17.

Viscusi, W. Kip. 2018. ""Best Estimate Selection Bias in the Value of a Statistical Life." *Journal of Benefit Cost Analysis*, *9*(2): 205–246.

Viscusi, W. Kip, and Joseph E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World." *Journal of Risk and Uncertainty*, *27*: 5–76.

Viscusi, W. Kip, and Elissa Philip Gentry. 2015. "The Value of a Statistical Life for Transportation Regulations: A Test of the Benefits Transfer Methodology." *Journal of Risk and Uncertainty*, *51*(1): 53–77.

Viscusi, W. Kip, and Clayton J. Masterman. 2017. "Income Elasticities and Global Values of a Statistical Life." *Journal of Benefit-Cost Analysis*, *8*(2): 226–250.

Cite this article: Banzhaf, H. S. 2022. "The Value of Statistical Life: A Meta-Analysis of Meta-Analyses." *Journal of Benefit-Cost Analysis* 13: 182–197, doi:10.1017/bca.2022.9