NCEE Working Paper

A two-stage random-effects meta-analysis of value per statistical life estimates

Stephen C. Newbold, Chris Dockins, Nathalie Simon, Kelly Maguire, and Abdullah Sakib

Working Paper 24-01
March, 2024

U.S. Environmental Protection Agency
National Center for Environmental Economics
https://www.epa.gov/environmental-economics

NCEE
NATIONAL CENTER FOR ENVIRONMENTAL ECONOMICS

A two-stage random-effects meta-analysis of value per statistical life estimates

Stephen C. Newbold
University of Wyoming, Department of Economics
snewbold@uwyo.edu

Chris Dockins, U.S. EPA, NCEE

Nathalie Simon, U.S. EPA, NCEE

Kelly Maguire, U.S.D.A., ERS

Abdullah Sakib
University of Wyoming, Department of Economics

Abstract

We demonstrate the use of a two-stage random-effects meta-analysis estimator for synthesizing published estimates of the value per statistical life (VSL). The meta-estimation approach accommodates unbalanced panels with one or multiple observations from each independent group of primary estimates, and distinguishes between sampling and non-sampling sources of error, both within and between groups. We use a series of Monte Carlo simulation experiments to examine the performance of the meta-estimator on constructed datasets. Simulation results indicate that, when applied to datasets of modest size, the approach performs best when the within-group non-sampling error variances are constrained to be equal across groups. This allows for two levels of non-sampling errors while preserving degrees of freedom and therefore increasing statistical efficiency. Simulation results also show that the performance of the estimator compares favorably to several other commonly used meta-analysis estimators, including other two-stage estimators. We illustrate the approach by applying it to a preliminary meta-dataset comprising 88 VSL estimates assembled from 9 hedonic wage and 9 stated preference studies conducted in the U.S. and published between 1999 and 2013.

Keywords: Hierarchical Meta-analysis, value of statistical life, unbalanced panels
JEL Codes: Q5, C4

Disclaimer: The views expressed in this paper are those of the author(s) and do not necessarily represent those of the U.S. EPA or U.S.D.A. In addition, although the research described in this paper may have been funded entirely or in part by the U.S. EPA, it has not been subjected to the Agency's required peer and policy review. No official Agency endorsement should be inferred.

Contents

1 Introduction

Analysts often use quantitative predictive models to aid in the design and evaluation of public policy interventions, and generally one or more key parameters of such models are not known with certainty. In some domains, many studies have reported one or more competing estimates of an important parameter using more-or-less credible research methods. In these cases, some means of synthesizing the available estimates—into a single best point estimate, or a credible range, or a probability distribution—is needed for use in quantitative policy evaluations.

Meta-analysis is a statistical approach for estimating the central tendency and examining the factors that influence the variation among multiple estimates of an unknown quantity of interest from different studies (Borenstein, Hedges, Higgins, & Rothstein, 2009; J. Nelson & Kennedy, 2009). Meta-analysis has been used to synthesize quantitative results from empirical studies in a wide variety of public policy domains, including job search and training programs (Card, Kluve, & Weber, 2010), the impacts of ethanol regulations on corn prices (Condon, Klemick, & Wolverton, 2015), the efficacy of nudges for improving public health (Arno & Thomas, 2016), the influence of education on intelligence (Ritchie & Tucker-Drob, 2018), COVID-19 infection fatality rates (Levin et al., 2020), and many more.

The "value per statistical life" (VSL) is among the most important estimates used in benefit-cost analyses of public policies related to health, safety, and the environment as reduced mortality often comprises the largest category of benefits for these actions (Arrow et al., 1996; Cropper, Hammitt, & Robinson, 2011). In the recently signed Reconsideration of the National Ambient Air Quality Standards Revision for Particulate Matter, for instance, over 98 percent of the monetized benefits were attributed to avoided statistical deaths (USEPA 2024). The VSL, used to monetize reduced mortality risk, refers to the marginal rate of substitution between income and mortality risks, i.e., an individual's marginal willingness-to-pay to reduce their risk of death (Kniesner & Viscusi, 2019). The VSL corresponds to the total dollar value associated with small changes in the risk of dying that, when aggregated over a large population, yield one statistical life. For example, if 100,000 individuals are each willing to pay, on average, $100 for a reduction in their risk of death in the coming year of 1/100,000, then the value of reducing the expected number of deaths in the group—i.e., saving one "statistical life"—equals 100,000 $\times$ $100, or $10 million (USEPA, 2014, p XV).

A comprehensive review of the theory and methods used to derive and estimate the VSL is beyond the scope of this paper. Banzhaf (2014) describes the historical origins of the VSL concept, and Cropper et al. (2011) provides a broad overview of estimation approaches and applications of the VSL in benefit-cost analysis. Other useful reviews, including discussions of open research questions, can be found in Ashenfelter (2006), Cropper et al. (2011), Viscusi (2012), Sunstein (2013), Robinson and Hammitt (2016), and Robinson, Hammitt, and O'Keeffe (2019).

Hundreds of VSL estimates have been reported in the peer-reviewed literature, and more than a dozen previous meta-analyses have been conducted to synthesize multiple estimates and examine the factors that influence their magnitudes. However, the focus and scope of previous meta-analyses have often been limited. Analyses have typically focused on a sub-set of the literature, either hedonic wage or stated preference studies, both of which can provide VSL estimates. [1] Many studies also select a single estimate per study or independent data sample, and even when multiple estimates per study are available, these

1. Kochi, Hubbell, and Kramer (2006) includes both types of studies and concludes that there are systematic differences in the VSL estimates. There are theoretic reasons why the values might differ. As noted by EPA' Science Advisory Board, hedonic wage studies may be interpreted as Marshallian willingness-to-accept values while most stated preference estimates may be interpreted as Hicksian willingness-to-pay USEPA (2017).

are frequently averaged to produce a single central study estimate before being combined with estimates from other studies. Instead, to preserve as much information as possible from the underlying studies, we use a two-stage random-effects meta-analysis estimator. This approach accommodates unbalanced panels with one or multiple observations from each independent group of primary estimates, and distinguishes between sampling and non-sampling sources of error, both within and between groups. In demonstrating this approach, we use VSL estimates from both the hedonic wage and stated preference literatures and we include tests for publication bias.

In the remainder of this introduction, we review several previous VSL meta-analyses. In subsequent sections, we describe our estimation approach in detail, examine its performance using Monte Carlo simulations, and demonstrate its application to a preliminary meta-dataset of VSL estimates assembled in an earlier U.S. EPA report (USEPA, 2016).

1.1 Previous VSL meta-analyses

In this section we briefly summarize the key features and results from previous VSL meta-analyses published in peer-reviewed journals. While there are several excellent reviews and summaries of the VSL literature (e.g., Keller, Newman, Ortmann, Jorm, & Chambers, 2021; Robinson & Hammitt, 2016), here we focus on statistical meta-analyses. In the discussion section below, we compare and contrast our estimation methods and results to some of the studies reviewed here. See Table 1 for key summary statistics from the VSL meta-analyses summarized below: the number of primary studies from which VSL estimates were drawn (I), the total number of observations in the meta-dataset (N), whether the included observations were estimated using a hedonic wage (HW) or stated preference (SP) approach, and the summary ranges of the VSL and the income elasticity estimates reported in each meta-analysis study.

Most previous VSL meta-analyses have focused on synthesizing hedonic wage-based estimates of the VSL. Mrozek and Taylor (2002) performed a meta-regression of 203 estimates from 33 hedonic wage studies. Weighted least squares was used for estimation, with weights equal to the inverse of the number of estimates from the parent study, giving each study equal weight. Precision weights were not possible because standard errors were often not reported. Viscusi and Aldy (2003) used single estimates from each of 44 to 46 studies in six meta-regression model specifications without precision weighting. Bellavance, Dionne, and Lebeau (2009) used a mixed-effects regression model to combine 39 estimates drawn from 37 hedonic wage studies. Estimates were chosen from each independent data sample (in most cases selecting a single estimate per study) based on similarity of the estimating equation with other studies, the original authors' preferred estimate, and other best-practice considerations. J. P. Nelson (2015) used the data assembled by Bellavance et al. (2009) plus additional hedonic wage observations from the U.S. Environmental Protection Agency USEPA (2010) in a "tentative and exploratory" meta-analysis of VSL estimates. After dropping outliers, single estimates from 28 primary studies were included in the final meta-dataset for four meta-regression specifications—OLS, fixed-effect, and two versions of random-effects models—which included use of the inverse standard errors as a test for publication bias.

The issue of publication bias on VSL estimates has been the focus of many meta-analyses of hedonic wage studies, first by Doucouliagos, Stanley, and Giles (2012) who found significant bias using the Bellavance et al. (2009) data set, and later in a series of articles by Viscusi and co-authors. Viscusi (2015) constructed a sample of 550 hedonic wage estimates based on 17 studies that used workplace fatality risks calculated from the

Census of Fatal Occupational Injury (CFOI) dataset, and compared VSL estimates and publication selection bias in this set to that found in other hedonic wage datasets, including that constructed by Bellavance et al. (2009). Estimates were weighted by inverse variance, and fixed- and random-effects variants of meta-regression models were estimated. CFOI-based estimates exhibited relatively little publication bias. Viscusi and Masterman (2017a) examined publication bias in U.S. and non-U.S. VSL estimates using a larger international dataset of 1,025 observations from 68 hedonic wage studies. The authors used weighted least squares with inverse variances of the VSL estimates used as observation weights. A quantile regression approach was used to examine publication bias at different levels of VSL estimates. Little evidence of publication bias was found in CFOI-based estimates, and there was evidence of strong bias among non-US studies, which the authors attributed to an anchoring effect of previously published U.S. VSL estimates. Viscusi (2018) further examined publication bias in the hedonic wage literature, comparing bias in "best-set" samples (i.e., 1 selected estimate per study) with that found when all study estimates are used. Weighted least squares results suggested that publication bias is statistically significant for both samples but is larger for the best-set sample. The central bias-adjusted VSL estimate for the all-set sample was $8.8 million (in 2020 U.S. dollars).[2]

Fewer meta-analyses of stated preference-based VSL estimates have been conducted. Dekker, Brouwer, Hofkes, and Moeltner (2011) used a Bayesian estimation approach in their meta-analysis of 77 estimates from 26 international contingent valuation studies conducted in 15 countries, with the goal of examining the effect of risk context on VSL estimates. Specifically, they empirically estimated correction factors to apply for "out of context" benefit transfers using CV studies focused on air pollution, road safety, or those considered

2. In this paper, all VSL estimates are reported in 2020 US dollars.

"context free." Lindhjem, Navrud, Braathen, and Biausque (2011) is perhaps the most comprehensive, global meta-analysis of stated preference VSL estimates, with 850 mean estimates drawn from 76 studies conducted in 38 countries. The study focused on the effects of population characteristics, risk type and context, survey format, and statistical methodological choices on the VSL estimates. Estimates were weighted by the inverse of the number of estimates selected from each study so that each survey received equal weight. For a subset of the primary studies, they also reported results using alternative weighting schemes for comparison—the inverse of the number of estimates from each study, the inverse of the standard deviation of the mean VSL estimates, and a combination of the two—and found that results were reasonably robust to the weights used. More recently, Masterman and Viscusi (2020) performed a meta-analysis of global stated preference VSL estimates, using 1148 estimates drawn from 85 studies. Using WLS weighted by inverse variance and including article controls, the authors found large and statistically significant publication biases with bias-adjusted VSLs never larger than $980,000.

Meta-analyses that examine both stated and revealed preference estimates are less common. Kochi et al. (2006) combined both hedonic wage and stated preference estimates. The authors used an empirical Bayes estimation approach in a two-stage pooling model to examine 197 estimates selected from 40 studies published in the U.S. and other high-income countries. In a first stage, the authors created subsets of estimates by the same author or groups of authors and calculated the mean value for the subset if it passed a statistical test for homogeneity. In a second stage, the authors combined the estimates from the 60 homogeneous first-stage subsets accounting for across-group variability using the Q-statistics for each group. A bootstrap approach was used to compare the distributions of VSL by study type, finding that the mean VSL from hedonic wage studies was roughly three times larger than that from stated preference studies.

Several meta-analyses have focused on the income elasticity of the VSL (IEVSL) rather than on the VSL estimates themselves. Doucouliagos, Stanley, and Viscusi (2014) assessed publication bias in 101 IEVSL estimates from 14 prior meta-analyses using a variety of estimators, including precision-effect estimate with standard error (PEESE) and with fixed- and random-effects meta analysis estimators, finding substantial publication bias effects. Viscusi and Masterman (2017b) estimated IEVSL for both US and non-US samples with publication bias correction and study controls, and also conducted quantile regression to estimate IEVSL across the national income distribution. Masterman and Viscusi (2018) used similar methods to examine IEVSL from stated preference VSL estimates based on the dataset from Lindhjem et al. (2011) supplemented by more recent studies and found similar results.

Table 1 lists key summary statistics from the VSL meta-analyses summarized above: the number of primary studies from which VSL estimates were drawn (I), the total number of observations in the meta-dataset (N), whether the included observations were estimated using a hedonic wage (HW) or stated preference (SP) approach, and the summary ranges of the VSL estimates reported in each meta-analysis study. Of the VSL meta-analyses reviewed here, only one examined both HW and SP estimates (Kochi et al., 2006) and none applied a multi-level random-effects estimator.

As a crude synthesis of the results from prior meta-analyses studies, we note that the average low, midpoint, and high ends of the ranges reported in the final column of Table 1 are 5.2, 7.4, and 9.5 million 2020 US$. In the Discussion section we will compare our results to those summarized here.

[Insert Table 1 around here.]

With so many VSL meta-analyses now available in the published literature, Banzhaf (2022) observed that "...the old problem of selecting a single best study has just been pushed back to the problem of selecting a single best meta-analysis." To consolidate this literature, Banzhaf synthesized 11 meta-estimates of the VSL from 6 prior meta-analysis studies: USEPA (1997), Mrozek and Taylor (2002) (2 estimates), Viscusi and Aldy (2003), Kochi et al. (2006) (2 estimates), Robinson and Hammitt (2016), and Viscusi (2018) (4 estimates). In his alternative model, which includes all source studies, Banzhaf specified weights for each estimate such that each study received equal weight. He then produced a mixture distribution by taking repeated random draws from the distributions defined by the means and standard errors of the constituent meta-estimates, with the pre-specified weights applied to each estimate.[3] The resulting VSL mixture distribution has a mean of \$7.6 million and 90% confidence interval from \$2.0 to \$13.1 million. We note that Banzhaf's consolidated central estimate is very close to the average of the midpoints in Table 1, and Banzhaf's range safely encompasses the range of average low and high estimates in Table 1.

Our crude summary of previous meta-analyses above and the quantitative synthesis by Banzhaf (2022) point to the same ballpark of central estimates for the VSL. These preliminaries provide the context for our main contribution in the present study, which is to describe and illustrate the use of a multi-level random-effects estimator that is more general and—at least under some circumstances, elaborated below—more precise than those

3. There is an important distinction between the conventional meta-analysis approach we use in the present study and the mixture distribution approach used by Banzhaf (2022). If we meta-analyze two independent estimates each with the same mean, μ, and variance, σ^2, the meta-estimate would be equal to μ and the variance of the meta-estimate would be equal to $\sigma^2/2$. If we mix two distributions with the same mean, μ, and variance, σ^2, the mean of the mixture distribution would be equal to μ, but the variance of the mixture distribution would be equal to σ^2. It is more appropriate to mix the meta-estimates, as Banzhaf did, rather than meta-analyze them if the constituent meta-analysis studies themselves used many of the same primary VSL estimates, and so the meta-estimates being synthesized cannot safely be treated as independent.

used in many previous VSL meta-analysis studies.

2 Methods

In this study, we demonstrate the use of a two-stage random-effects (2SRE) estimator for synthesizing and analyzing published estimates of the VSL.[4] Our estimation approach is closely related to other multi-level meta-analysis methods, including the three-level meta-analysis approach described by Konstantopoulos (2011) and the hierarchical dependence model described by L. V. Hedges, Tipton, and Johnson (2010) and Tipton (2015). Our main contributions include: 1) tailoring a multi-level random-effects estimator to conditions we expect to characterize VSL meta-datasets, 2) conducting a series of Monte Carlo simulation experiments to examine the performance of the estimator in comparison to several other commonly used meta-analysis estimators in our data environment, and 3) and applying the estimator to a preliminary meta-dataset of VSL estimates from hedonic wage and stated preference studies conducted in the United States between 1999 and 2013.

2.1 A two-stage random-effects meta-analysis estimator

In this sub-section we describe the 2SRE estimator that we propose to use for synthesizing published VSL estimates. A complete derivation is provided in the Appendix.

To begin, we decompose each observation into the sum of the true effect size and three error components,

$$y_{ij} = y + \eta_i + \mu_{ij} + \varepsilon_{ij}, \tag{1}$$

where y_{ij} is an observed VSL estimate j from group i, y is the average VSL among the

4. A Github repository containing a set of R scripts sufficient to replicate all results reported in this paper can be found at: `https://github.com/scnewbold/2SRE`.

U.S. adult general population (our target of estimation), η_i is a group-level non-sampling error, μ_{ij} is an observation-level non-sampling error, and ε_{ij} is an observation-level sampling error.[5] To clarify the distinction between sampling and non-sampling errors in our setting, note that our target of estimation, or estimand, is the average VSL among the adult U.S. population. To form our meta-dataset, we will draw from the literature primary estimates based on data samples and model specifications originally designed to identify the average VSL among the entire U.S. adult general population or a large sub-set of the general population—e.g., working adults between the ages of 18 and 65, as in many hedonic wage studies. The estimands in those primary studies with non-representative samples will differ from our estimand by an amount that depends on the degree of non-representativeness of their samples along the relevant dimensions and the association between those sample characteristics and people's marginal willingness-to-pay for mortality risk reductions. In these cases, the primary estimates would be a biased estimate of our estimand even if it were an unbiased estimate of the average VSL among the subset of the population from which the original sample was drawn. All deviations in the primary VSL estimates stemming from differences in the sampling frames, estimation approaches, functional forms of estimating equations, selection of exogenous control variables, handling of outliers, and any other idiosyncratic data cleaning and modeling choices among the primary studies—i.e., all sources of variability in the primary VSL estimates that do not arise from sampling variation per se—are subsumed in our composite "non-sampling error" terms, $\eta_i + \mu_{ij}$.

5. A note on terminology is in order. We use "non-sampling errors" to refer to what is commonly called "heterogeneity" in the meta-analysis literature. For example, L. Hedges, Shymansky, and Woodworth (1989) discussed this distinction as follows: "Sampling standard error measures the sampling variation of the estimated effect size but does not reflect non-sampling variations which would occur if the study had used a different population of students or different teachers...," and "The variation among studies is, of course, due in part to random sampling fluctuations as reflected in the sampling standard errors. However, in some cases differences between individual studies exceed several standard errors, presumably reflecting differences in the characteristics of those studies... To study this 'non-sampling' variation we use heterogeneity analysis."

Returning now to our model set-up, we decompose the composite errors such that η_i varies between but not within groups, while μ_{ij} and ε_{ij} can vary both between and within groups. The standard errors reported for each observation represent the sampling variability of the published estimates conditional on the designs of the original studies. The variances of the sampling error components, $\sigma^2_{\varepsilon,i}$, are assumed known and equal to the squared standard errors of the VSL estimates as reported in the original studies, se^2_{ij}.[6] The variances of the between- and within-group non-sampling error components, σ^2_η and $\sigma^2_{\mu,i}$, are unknown and will be estimated from the data.

For our meta-analysis estimator to be unbiased, all error components must have means of zero. This is a common assumption but its plausibility will depend in part on the selection criteria used to draw primary estimates from the published literature. In particular, at least two constituent assumptions must hold to make $\mathbb{E}[\eta_i]=0$ and $\mathbb{E}[\mu_{i,j}]=0$: (1) the non-sampling errors stemming from non-representative sampling frames and differences in study designs are idiosyncratic, and so just as likely to lead to positive as negative biases with respect to our estimand, and (2) publication bias is negligible, and so the estimates that appear in the published literature are not selected on their magnitudes. We will maintain the first assumption throughout, but we will demonstrate how to test the second assumption in a side-analysis using two conventional publication bias estimators.[7]

Conditional on the zero-mean-errors assumption, any convex combination of the observations will provide a consistent estimate of the average VSL. Our aim is to find the set of weights that gives the most efficient consistent estimator possible. The estimator can be

6. This is a simplification that is shared by all meta-analysis estimators that we are aware of. A more general approach would also account for uncertainty in the reported standard errors.

7. The findings of Viscusi and Masterman (2017a) suggest that studies using CFOI mortality risk data are the least subject to publication bias among all subsets of VSL estimates from previous hedonic wage studies that they examined. Masterman and Viscusi (2020) find that stated preference studies are subject to large publication bias.

written as a weighted average,

$$\hat{y} = \sum_{i=1}^{I} \sum_{j=1}^{J_i} w_{ij} y_{ij}, \tag{2}$$

where $\sum_{i=1}^{I} \sum_{j=1}^{J_i} w_{ij} = 1$. We derived formulas for the weights, w_{ij}, as follows. First, we found conditional observation-level weights, g_{ij}, to calculate group-level estimates $\hat{y}_i = \sum_{j=1}^{J_i} g_{ij} y_{ij}$, where $\sum_{j=1}^{J_i} g_{ij} = 1$. Second, we found group-level weights $\hat{y} = \sum_{i=1}^{I} h_i \hat{y}_i$, where $\sum_{i=1}^{I} h_i = 1$. Third, we calculated the unconditional observation-level weights as $w_{ij} = h_i g_{ij}$. Constraining the weights to sum to 1 at each level ensures that the group-level estimates are consistent and that the overall estimate is consistent.[8]

We derived the g_{ij}'s to minimize the variance of the group-level estimates, which requires estimates of $\sigma_{\mu,i}$ for each group, and we derived the h_i's to minimize the variance of the overall weighted mean, which depends on the conditional variances of the group level estimates and requires an estimate of σ_η. We used a method-of-moments approach to derive estimates of $\sigma_{\mu,i}$ and σ_η, so no assumptions about the shapes of the error distributions were used.

Some groups in the preliminary meta-dataset have only a single observation, which means $\sigma_{\mu,i}$ cannot be estimated for those groups. To proxy the within-group non-sampling error variance for singleton groups, we used the average of the $\hat{\sigma}_{\mu,i}$'s for the non-singleton groups. The alternative of assuming $\sigma_{\mu,i} = 0$ for singleton groups would have the unintended effect of penalizing primary studies that reported more than one VSL estimate. By assigning the mean non-sampling error variance to the singleton groups, non-singleton groups with

8. We could generalize further and consider an inconsistent estimator that may achieve a lower mean-squared error by shrinking the estimate towards zero (Efron & Morris, 1977; Rasmusen, 2015; Samworth, 2012). However, it appears there is little to be gained from such an approach in the present context. For example, consider perhaps the simplest possible shrinkage estimator $\tilde{y} = \alpha\hat{y}$, where $\hat{y}$ is our most efficient consistent 2SRE estimator. The optimal shrinkage factor is $\alpha^\star = \hat{y}^2/(\hat{y}^2 + se[\hat{y}]^2)$. Looking ahead to our results in Table 7, we find $\hat{y} \approx 8$ and $se[\hat{y}] \approx 1$, which gives $\alpha^\star \approx 0.985$. Considering that the shrinkage factor must be estimated from the data and so would itself add noise, this slim potential gain suggests that very little if any shrinkage would be optimal in practice in this setting.

observations that have lower than average non-sampling error variances will receive more weight than the singleton groups, and those with higher than average non-sampling error variances will receive less weight, all else equal. This gives more leverage to studies whose estimates are more robust to variations in functional form assumptions and other sensitivity tests designed to examine uncertainties unrelated to sampling variability.

The estimator also allows for correlation among sampling errors, ρ, but does not estimate this value. The analyst must specify ρ and can examine the influence of this assumption through sensitivity analysis. We investigated the effect of mis-specifying this correlation in our Monte Carlo experiments described below.

The foregoing description of the estimation approach has focused on the calculation of precision weights for the observations in a meta-analysis context, with no moderator variables included. For applications to meta-regression models, which include one or more moderator variables intended to help explain some of the systematic heterogeneity among the quantities estimated in each primary study, the same approach to calculating the optimal precision weights applies except the mean of each observation, y, is replaced with $f(x_{ij}, \beta)$—e.g., $x_{ij}\beta$ in a linear meta-regression model—in equations (1) and (2) above. All equations necessary to compute the 2SRE estimator are shown in Table 2, and the Appendix provides a full derivation.

[Insert Table 2 around here.]

In our illustrative application, we used iterated weighted least squares to estimate linear meta-regression models. This involves initializing $\hat{\beta}$ by regressing y on x with either no weighting (ordinary least squares) or precision weights based on the reported standard errors only (a fixed-effect size meta-regression model). Then $\hat{\beta}$ is used to estimate the error component variances, and the estimated error component variances are used to re-calculate

$\hat{\beta}$ using weighted least squares. The process is repeated until the joint set of estimates converge to stable values.[9] If x only includes a constant, then the estimator collapses to the simple meta-analysis model with no moderator variables described above, in which case no iteration is required.

2.2 Performance comparison using Monte Carlo experiments

To examine the performance of the 2SRE estimator, we conducted a series of Monte Carlo simulation experiments using constructed data. For each experiment, we specified the number of groups, I, the number of observations for each group, J_i, the true VSL, y, the error component variances, σ^2_η and $\sigma^2_{\mu,i}$, and the within-group sampling error correlation, ρ. For 16 combinations of the experimental design parameters, we applied several alternative meta-analysis estimators, including the 2SRE estimator, to each of 2,000 simulated meta-datasets. The estimators we compared are listed and described in Table 3.

[Insert Table 3 around here.]

The first two estimators, the simple mean and group means, make no use of the reported standard errors for each observation nor do they attempt to estimate any unobserved error components for precision weighting. The next three estimators—metafor, robumeta, and MAd—are commonly used meta-analysis packages developed for `R`. The final estimator is the the two-stage random-effects estimator tested in this study for the purpose of synthesizing published VSL estimates. We apply three versions of the 2SRE estimator. The first version (2SRE-true) uses the true error component variances to compute precision weights. This is impossible using real data, but is useful here to provide a lower bound estimate of the

9. Our stopping criteria was when the largest change among the coefficient estimates became smaller than 0.001%, which is safely below the sampling variability of these estimators.

standard errors for all feasible estimators that can take the form of an unbiased weighted mean as in equation (2). The second version (2SRE-free) allows for heterogeneous within-group non-sampling error variances. This version is the most general and should be the most efficient feasible estimator with sufficiently many groups and observations per group. The third version (2SRE-equal) is constrained by imposing a common within-group non-sampling error variance. This version may outperform the second version of the 2SRE estimator if the number of observations per group is small.

The precision of each estimator is indicated by the standard deviation of the resulting VSL weighted mean estimates among all 2,000 Monte Carlo trials. For comparison to our simulation-based estimates of standard errors, we calculated robust standard errors following L. V. Hedges et al. (2010).[10]

2.3 Detecting and correcting publication bias

A common concern in meta-analysis studies is the possibility of publication bias (Ioannidis & Doucouliagos, 2013). Though our main focus in this study is on the statistical efficiency of alternative meta-analysis estimators when applied to VSL meta-datasets, we also used two conventional methods to address publication bias: the trim-and-fill and PET-PEESE estimators.

The trim-and-fill estimator (Duval & Tweedie, 2000) is a non-parametric method based on the observation that a plot of precision estimates ($1/se^2$) versus corresponding effect size estimates—often called a "funnel plot"—should be vertically symmetric. If all estimates are equally likely to be published, then the funnel plot should be wide at the bottom (low

10. In a side analysis not reported here, we also calculated bootstrapped standard errors by re-sampling independent groups with replacement (Ren et al., 2010). This allowed us to compare the performance of robust and bootstrapped standard error estimates under a common set of experimental design settings. We found a close correspondence between the bootstrapped standard errors and the robust standard errors, so here we report only the more easily calculated robust standard errors.

precision studies) and narrow at the top (high precision studies) with roughly the same number of estimates on the left and right sides of their center of mass. On the other hand, if estimates with low t-statistics are less likely to be published, then the funnel plot will have a conspicuously lower density of estimates in the bottom-left region of the plot (assuming positive effect size estimates). The trim-and-fill estimator works by iteratively "trimming" estimates on the far right side of the plot until the trimmed funnel is no longer asymmetric, then re-calculating the mean of the remaining estimates, then "filling" the trimmed and missing estimates on both sides of the plot around the corrected mean to compute the variance of the estimator.

The PET-PEESE estimator (Stanley & Doucouliagos, 2014) uses a two-stage regression approach to detect and correct for publication bias. The first stage (the PET or "precision effect test") involves regressing the effect size estimates on a constant and the standard errors. If the coefficient on the standard errors is significantly different from zero, this is taken as evidence of publication bias. In these cases, a second stage (the PEESE or "precision-effect estimate with SE") is applied, which involves regressing the effect size estimates on a constant and the squared standard errors. The estimated constant in this regression is taken as a corrected mean effect size. Intuitively, the se^2 term controls for the influence of study precision on the reported effect size estimates, and the estimated constant extrapolates the relationship to an (hypothetical) infinitely precise study with $se^2 = 0$.

2.4 Application to a preliminary VSL meta-dataset

To demonstrate the 2SRE estimation approach using realistic data, we applied it to a preliminary meta-dataset assembled by the U.S. Environmental Protection Agency as part of a review of proposed meta-analysis methods by the Agency's Science Advisory Board

(USEPA, 2016).

The dataset contains VSL estimates (hereafter "observations") from both stated preference and hedonic wage studies. Multiple observations were drawn from studies meeting our screening criteria. Where available, this includes both mean and median VSL estimates, and their respective standard errors. The dataset includes 46 observations from 9 hedonic wage studies and 42 observations from 9 stated preference studies. Detailed information about the dataset, including the full list of studies and the screening criteria, is provided by USEPA (2016).

The EPA Science Advisory Board made a number of recommendations for altering both the dataset and methods proposed in the 2015 EPA report (EEAC, 2017). An important motivation for the present study was the board's recommendations to refine and improve the estimation approach. Updates and modifications to the dataset itself will be completed as a separate task. Therefore, our analysis of the preliminary meta-dataset is presented here for the purpose of demonstrating the proposed estimation approach using realistic data. Given the limitations of the dataset, our results should be viewed as illustrative and do not represent an official summary measure of the VSL for use in benefit-cost analysis.

3 Results

3.1 Monte Carlo experiments

Results from our Monte Carlo experiments are shown in Tables 4–7. We compared the candidate estimators under four combinations of true and assumed correlations among non-sampling errors within studies, ρ and $\hat{\rho}$. In all four tables, each row corresponds to a unique combination of the number of groups, I (20 or 60), the minimum and maximum number of observations in each group, J (drawn randomly from the range 1–5 or 1–15), the group-level (between groups) non-sampling error variability, σ_η (1.0 or 3.0), and the observation-level (within groups) non-sampling error variability, σ_μ (drawn randomly from the range 0.5–1.0 or 0.5–3.0). In all cases the true VSL was 10 and the sampling error variability, se, was drawn from the range 0.5–5.0. The following 8 columns in each table contain standard deviations of 2,000 Monte Carlo applications of each estimator to constructed datasets based on the experimental design settings in the first 4 columns.

The alternative estimators appear from left to right in the columns of Tables 4–7 in the order the estimators are listed in Table 3. In all cases, the data were constructed with σ_{μ_i} heterogeneous across groups, so the constrained 2SRE-equal estimator, with $\hat{\sigma}_{\mu_i} = \hat{\sigma}_\mu$ for all i, imposes a binding restriction on the estimating equation. This restriction will not bias the estimator but will make it less efficient than the unconstrained 2SRE-free estimator in sufficiently large samples, or more efficient in sufficiently small samples, where the large-versus-small sample size threshold will depend on all parameters of the data generating process. We have attempted to vary the experimental design settings to cover ranges that are typical for VSL meta-analyses, so the Monte Carlo comparisons among all of the estimators, including the unconstrained and constrained versions of the 2SRE estimator, should be informative for realistic VSL meta-analysis applications.

Before considering the results in detail, we note that in all four tables the standard errors are less than 1.0 for nearly all estimators under nearly all experimental design settings. This is relatively high precision considering that the true VSL was set at 10 for these numerical experiments. Therefore, using a VSL meta-dataset with characteristics within the range of sample sizes and error component variances considered here, a variety of reasonable meta-analysis estimators should produce a 95% confidence interval with a half-width less than

20% of the central estimate itself. Nevertheless, we can clearly discern systematic differences in performance among the competing estimators by considering each of the four tables of Monte Carlo simulation results in turn.

[Insert Table 4 around here.]

The results in Table 4 are based on cases with no correlation among non-sampling errors within groups, $\rho = 0$, and where the analyst has correctly set $\hat{\rho} = 0$. Here the 2SRE-equal estimator is more efficient than the other estimators in 13 out of 16 cases.

[Insert Table 5 around here.]

The results in Table 5 are based on cases with a positive correlation among non-sampling errors within groups, $\rho = 0.5$, and where the analyst has incorrectly set $\hat{\rho} = 0$. Here the 2SRE-equal estimator is more efficient than all other estimators in 12 of 16 cases, the 2SRE-free estimator performs best in one case, and the metafor estimator performs best in the three remaining cases.

[Insert Table 6 around here.]

The results in Table 6 are based on cases with a positive correlation among non-sampling errors within groups, $\rho = 0.5$, and where the analyst has correctly set $\hat{\rho} = 0.5$. Here again the 2SRE-equal estimator is more efficient than all other estimators in all 16 cases.

[Insert Table 7 around here.]

The results in Table 7 are based on cases with no correlation among non-sampling errors within groups, $\rho = 0$, and where the analyst has incorrectly set $\hat{\rho} = 0.5$. Here the 2SRE-equal estimator is as or more efficient than the other estimators in 12 out of 16 cases. Under this $(\rho, \hat{\rho})$ configuration, the robumeta-HIER estimator is most efficient in the remaining 4 cases.

The final columns in Tables 4–7 show that the average robust standard errors, $\hat{se}$ (L. V. Hedges et al., 2010), which were computed for the 2SRE-equal estimator, closely match the standard deviations of their corresponding simulated estimates, as they should. The R^2 values between the two quantities is greater than 0.99 in all four tables, and we see no apparent bias of the robust standard errors even in Tables 5 and 7 where $\hat{\rho} \neq \rho$, so it does not appear that mis-specification of $\hat{\rho}$ will substantially compromise standard errors.

[Insert Figure 1 around here.]

A visual depiction of the relative performance of all tested estimators is shown in Figure 1. The four charts in the figure contain the same information as the four tables of Monte Carlo simulation results, Tables 4-7, but presented as means and ranges of the standard errors of each estimator normalized by their theoretical minimum possible standard errors (based on the "2SRE-true" estimator, which uses the true error component variances to compute standard errors). For example, a bar with height 0.2 indicates that the standard error of the estimator is 20% larger than the minimum possible standard error. The error bars indicate the minimum and maximum normalized standard errors for each estimator across all 16 cases examined in our Monte Carlo experiments. These charts show more clearly that the 2SRE-equal estimator performs at least as well as the others on average in all four $(\rho, \hat{\rho})$ configurations. The charts also show that the 2SRE-free estimator performs poorly relative to to the constrained version, especially when $\rho = 0$. The simple mean and group mean estimators show their best performance when $\rho = 0$. The efficiency advantages of the more sophisticated meta-analysis estimators are more clearly evident when $\rho > 0$, a condition we expect to hold in most realistic meta-datasets that include multiple estimates from the same study or the same underlying primary datasets.[11]

3.2 Demonstration using realistic data

Meta-analysis results

A variety of meta-analysis estimates using several subsets of the preliminary EPA metadata are shown in Table 8. Results for seven estimators are presented: simple mean, group means, 2SRE-free, 2SRE-equal, and three modified versions of the 2SRE estimator with corrections for publication bias using on the trim-and-fill (T&F) and the PET-PEESE (P-P) methods. Each estimator is applied to data only from hedonic wage (HW) studies, only from stated preference (SP) studies, and from both HW and SP studies (pooled). The final column shows the simple average of the independent HW and SP estimates (balanced), which places equal weight on the two types of primary estimation methods regardless of the number of studies and observations of each type.

[Insert Table 8 around here.]

Comparing primary estimation approaches, HW estimates are larger than SP estimates in 12 of 14 cases, but the differences are smaller when using only mean VSL observations from the SP studies. The pooled and balanced estimates are very close to each other for all estimators that do not involve publication bias corrections. The largest difference between the pooled and balanced estimates is produced by the 2SRE-free T&F estimator using only mean VSL observations, for which the balanced estimate is nearly $1.5 million larger than the pooled estimate.

All primary studies using the hedonic wage approach reported only mean VSL obser-

11. Recall that ρ is the correlation among sampling errors within groups. Non-sampling errors within groups are correlated in all cases because each observation includes a group-specific error term, η_i. Therefore, $\rho = 0$ does *not* mean that observations within groups are completely independent.

vations, so the "mm" and "m" entries are the same for each estimator in the HW column. Primary studies using stated preference approaches reported mean or median or both types of VSL observations, so the "mm" and "m" entries are different for each estimator in the SP column. In all cases, median observations were lower than mean observations, so the "mm" estimates are lower than the "m" estimates.

Publication bias corrections have variable effects on the meta-analysis estimates. The trim-and-fill (T&F) correction reduces the 2SRE HW estimates by $0.36 and $0.57 million, and it substantially reduces most of the 2SRE SP estimates, by $2 million or more, the only exception being the 2SRE-free "m" estimate which increases slightly. The PET-PEESE (P-P) HW estimate is $1.4 and $0.83 million lower than the uncorrected 2SRE estimates, and the P-P SP estimates are $0.56 lower and $0.92 higher than the corresponding uncorrected 2SRE-equal mm and m estimates.

A broad-brush summary of the results in Table 8 is that the average of all of estimates is $8.14 million, and 38 of 49 estimates (not counting the repeated HW estimates) are between $6 and $10 million, including the four estimates with the lowest RMSE's highlighted in bold font.

Meta-regression results

In addition to the meta-analyses reported in Table 8, we also estimated a variety of meta-regression specifications with control variables for SP observations, median observations, the year of data collection, and the average U.S. income in the year of data collection. We estimated a benchmark model with no control variables plus six specifications including two or more control variables or their interactions. Beginning with Table 9, we show results for the following seven specifications:

S0. No controls

S1. SP, median

S2. SP, median, year

S3. SP, median, income

S4. SP, median, year, income

S5. SP, median, year, SP×year

S6. SP, median, income, SP×income

Table 10 shows results from seven parallel specifications where each also includes the standard error of the primary VSL observations, *se*, as an additional control variable, which implements the PET stage of the PET-PEESE publication bias estimator. Table 11 shows results for the same specifications where each also includes the squared standard error, se^2, which implements the PEESE stage of the PET-PEESE estimator. In Tables 9-11, the 2SRE-equal meta-regression estimation approach was used. Tables 12-14 show all of the same specifications as the preceding three tables but now using the 2SRE-free estimation approach.

[Insert Table 9 around here.]

In Table 9, the estimate of the constant in specification S0 matches the 2SRE-equal pooled "mm" estimate in Table 8. This occurs because the meta-regression estimator with no control variables is equivalent to the meta-analysis estimator. The standard errors are slightly different because in Table 8 we report bootstrapped standard errors while in Table 9 we report robust standard errors. All estimates of σ_μ and σ_η in Table 9 are between 2 and 3, which is within the ranges of values used in our Monte Carlo simulation experiments. Coefficient estimates for the SP dummy variables are always negative, but their magnitudes vary widely across specifications (from -0.4 to -3.4). The SP coefficient estimate is statistically significant only in specification S5. Coefficient estimates for the median dummy variable are always negative and between -1 and -3, but never statistically significant. The year time trend is between 0.4 and 0.6, and is statistically significant in all three specifications in which it appears. The income coefficient is negative in specifications S3 and S4 but not statistically significant; it is positive and statistically significant in specification S6. The corresponding value of the *IEVSL* at the means of the control variables in specification S6 is 0.333. Based on the R^2_{CV} values reported in the final row of Table 9—which were computed using leave-one-out cross-validation residuals (Hastie, Tibshirani, Friedman, & Friedman, 2009, Ch 7)—the best fitting specification is S2, and the best fitting specification that excludes the time trend variable is S6.[12]

[Insert Table 10 around here.]

In Table 10, the *se* coefficient is close to the conventional threshold for statistical significance (somewhat above or below 2) in all specifications. We view this as modest evidence for publication bias according to the PET test.

[Insert Table 11 around here.]

In Table 11, the PEESE-corrected estimates of the constant—which correspond to HW-based VSL observations at the average of the 'datayear' variable—in all specifications are lower than their counterparts in the benchmark specifications reported in Table 9, where the differences are between $0.7 and $1.8 million. However, the results in Tables 10 and 11 should be viewed in light of the relatively noisy PET-PEESE meta-analysis estimates

12. The EPA Science Advisory Board concluded that without a clear rationale for giving different weights to estimates from different years, a time trend should not be included in the specification. Instead, they recommended that the influence of the timing of the studies be explored through sensitivity analysis (EEAC, 2017).

reported in Table 8 above, as well as the apparently reduced power of the PET-PEESE estimator in random effects panel data environments reported in some simulation studies (Alinaghi & Reed, 2018; Hong, 2019; Reed, 2019).

Estimates of the constant in Table 12 are between \$0.3 and \$0.7 million lower than their counterparts in Table 9. Estimates of the coefficient on *se* in Table 13 are slightly higher than their counterparts in Table 10 and now are statistically significant in all specifications, which suggests a stronger signal of publication bias. Estimates of the constant in Table 14, which include the PEESE publication bias correction, are between \$0.5 and \$1 million lower than their uncorrected counterparts in Table 11. Estimation results shown in Tables 12–14 from the 2SRE-free model, with unconstrained group-level non-sampling error variances, $\sigma^2_{\mu,i}$, generally have lower R^2_{CV} values than their constrained counterparts in Tables 9–11. This is consistent with our simulation results, which indicate that the 2SRE-equal estimator performs better than the 2SRE-free variant in data environments similar to the preliminary EPA meta-dataset used in this demonstration application.

4 Discussion

In this study we described and demonstrated a 2-stage random effects (2SRE) meta-analysis estimation approach that accommodates unbalanced panels with single or multiple observations per group, accounts for sampling and non-sampling sources of error, and allows for correlations among non-sampling and sampling errors within groups. Our estimation approach is similar to the three-level meta-analysis approach described by Konstantopoulos (2011) and to the robust variance estimation approach described by L. V. Hedges et al. (2010) and Tipton (2015), which is operationalized in the robumeta `R` package Fisher and Tipton (2015). The primary contributions of the present study include our extensive simulation experiments and our application of the estimation approach to a realistic, albeit preliminary, VSL meta-dataset, which together provide a robust indication of the strong performance of the estimator in relevant data environments.

We examined the performance of the estimator on constructed datasets in a series of Monte Carlo simulation experiments designed to bracket the range of data features that we expect to characterize VSL meta-analyses focused on estimates from the U.S. We found that the estimator performs well in this setting compared to alternatives including three meta-analysis estimators that have been developed into commonly-used `R` packages. The strong performance of the 2SRE estimator included cases involving within-group correlations among sampling errors that the analyst may or may not correctly specify. The constrained 2SRE-equal estimator, which assumes a common non-sampling error variance among groups, outperformed the 2SRE-free variant in all of the simulation experiments we conducted. The latter is the most general version of the estimator, which in principle should perform best in large samples. This suggests that our simulated meta-datasets were too small for its potential performance advantages to emerge. The 2SRE-equal variant of the estimator performed best overall in all four cases we examined. In particular, the 2SRE-equal estimator outperformed all others when $\rho = 0.5$ and $\hat{\rho} = 0.5$. We believe $\rho = 0.5$ is more realistic than $\rho = 0$, so we would recommend $\hat{\rho} = 0.5$ as a default setting. Our simulation results also suggest that robust standard errors are (nearly) unbiased even when the analyst incorrectly specifies the correlation among non-sampling errors within groups.

We applied several variations of the 2SRE estimator to a preliminary meta-dataset of VSL estimates assembled by the U.S. EPA. Variations of the estimation approach, including un-weighted and weighted meta-analyses and meta-regressions with and without adjustments for publication bias, were applied to the full dataset and various subsets of the data and produced central estimates of the VSL between $6 and $12 million.

How might an analyst curate the many competing estimates of the VSL that appear in Tables 8–14 to select a central value or range of values for use in policy evaluations? This will require a number of judgment calls by the end-user that we cannot anticipate, but here is one possible approach that we believe would be defensible: First, consider the pooled and balanced meta-analysis estimates that do not correct for publication bias in Table 8 and the meta-regression estimates in Tables 9 and 11 because these use all of the available evidence from both hedonic wage and stated preference studies. Among these estimates in Table 8, focus on those with the lowest root mean squared errors, which in our demonstration include the group means "mm" balanced case ($\hat{y} = 9.59$, RMSE $= 0.82$), the 2SRE-free "mm" pooled case ($\hat{y} = 7.59$, RMSE $= 0.82$), and the 2SRE–free "mm" balanced case ($\hat{y} = 7.54$, RMSE $= 0.79$). Next, consider the corresponding 2SRE estimates that adjust for publication bias. The T&F estimator has substantially lower variance than the PET-PEESE estimator, and the respective T&F-adjusted VSL estimates are $6.14 and $6.66 million. This large difference suggests that publication bias may be important, so we would include these estimates within the range of values to be used for policy analysis. As noted earlier, 38 of the 49 estimates in Table 8 are between $6.0 and $10 million. Among the meta-regression estimates reported in Tables 9 and 11, the best fitting models are those with the 'datayear' variable included. However, the EPA's Science Advisory Board recommended *not* controlling for the year of data collection in VSL meta-regressions citing a lack of a clear rationale for including them (EEAC, 2017). The specifications in Table 9 that exclude 'datayear,' produce produce balanced VSL estimates between $8 and $10 million.These estimates fall safely within the central range of estimates from Table 8, so we would not expand that range based on the meta-regression specifications that do not control for publication bias or a time trend in Table 9. The best-fitting specifications among those that do control for publication bias using the PEESE estimator in Tables 11 and 14 also produce central VSL estimates within this range. Based on all of these considerations, our overall synthesis of these preliminary results is that a central estimate around \$8 million and a range for sensitivity analysis between \$6 and \$10 million would be reasonable.

We also used a model averaging approach to combine the results from our various meta-regressions (Steel, 2020). Specifically, we applied jackknife model averaging (JMA)—which computes a weighted average of model outputs where the weights are chosen to minimize the sum of squared residuals of the resulting weighted average (Hansen & Racine, 2012)—to 28 meta-estimates of the VSL in Tables 9-14 (seven specifications $\times$ constrained or unconstrained $\sigma_{\mu,i}$'s $\times$ without and with correction for publication bias). The JMA weighted average of the estimated balanced VSL's (regression constants plus one half of the SP dummy coefficient) at the means of all control variables was \$8.28 million.

Because the metadata used in this application are preliminary, the results from this application also are preliminary. We offer it as a demonstration of the general estimation approach on realistic data; it should not be construed as an official update of VSL values for use in EPA economic analyses. The natural next step would be to develop a more definitive meta-dataset of VSL estimates to which a multi-level meta-analysis estimator like the 2SRE estimation approach developed here could be applied in a follow-up study.

References

Alinaghi, N., & Reed, W. R. (2018). Meta-analysis and publication bias: how well does the FAT-PET-PEESE procedure work? *Research Synthesis Methods*, *9*(2), 285–311.

Arno, A., & Thomas, S. (2016). The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis. *BMC Public Health*, *16*(1), 1–11.

Arrow, K., Cropper, M., Eads, G., Hahn, R., Lave, L., Noll, R., ... Stavins, R. (1996). Is there a role for benefit-cost analysis in environmental, health, and safety regulation? *Science*, *272*, 221–222.

Ashenfelter, O. (2006). Measuring the value of a statistical life: problems and prospects. *The Economic Journal*, *116*, C10–C23.

Banzhaf, H. (2014). The cold-war origins of the value of statistical life. *Journal of Economic Perspectives*, *28*, 213–226.

Banzhaf, H. (2022). The value of statistical life: a meta-analysis of meta-analyses. *Journal of Benefit Cost Analysis*, *13*, 182-197.

Bellavance, F., Dionne, G., & Lebeau, M. (2009). The value of a statistical life: a meta-analysis with a mixed effects regression model. *Journal of Health economics*, *28*(2), 444–464.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. West Sussex, United Kingdom: John Wiley Sons, Ltd.

Card, D., Kluve, J., & Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal*, *120*(548), F452–F477.

Condon, N., Klemick, H., & Wolverton, A. (2015). Impacts of ethanol policy on corn prices: a review and meta-analysis of recent evidence. *Food Policy*, *51*, 63–73.

Cropper, M., Hammitt, J., & Robinson, L. (2011). Valuing mortality risk reductions: progress and challenges. *Annual Review of Resource Economics*, *3*, 313–336.

Dekker, T., Brouwer, R., Hofkes, M., & Moeltner, K. (2011). The effect of risk context on the value of a statistical life: a Bayesian meta-model. *Environmental and Resource Economics*, *49*(4), 597–624.

Del Re, A. (2015). A practical tutorial on conducting meta-analysis in R. *The Quantitative Methods for Psychology*, *11*(1), 37–50.

Del Re, A., & Hoyt, W. (2014). *Package 'mad'.*

Doucouliagos, H., Stanley, T. D., & Giles, M. (2012). Are estimates of the value of a statistical life exaggerated? *Journal of Health Economics*, *31*, 197–206.

Doucouliagos, H., Stanley, T. D., & Viscusi, W. K. (2014). Publication selection and the income elasticity of the value of a statistical life. *Journal of Health Economics*, *33*, 67–75.

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463.

EEAC. (2017). *SAB review of EPA's proposed methodology for updating mortality risk valuation estimates for policy analysis* (Tech. Rep.). Environmental Protection Agency Science Advisory Board, Environmental Economics Advisory Committee: Washington, DC.

Efron, B., & Morris, C. (1977). Stein's paradox in statistics. *Scientific American*, *236*(5), 119–127.

Fisher, Z., & Tipton, E. (2015). Robumeta: An R-package for robust variance estimation in meta-analysis. *arXiv preprint arXiv:1503.02220*.

Hansen, B. E., & Racine, J. S. (2012). Jackknife model averaging. *Journal of Econometrics*, *167*(1), 38–46.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2). Springer.

Hedges, L., Shymansky, J., & Woodworth, G. (1989). *A Practical Guide to Modern Methods of Meta-Analysis*. Washington, DC.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65.

Hong, S. (2019). Meta-analysis and publication bias: how well does the FAT-PET-PEESE procedure work? A replication study of Alinaghi & Reed (Research Synthesis Methods, 2018). *International Journal for Re-Views in Empirical Economics (IREE)*, *3*(2019-4), 1–22.

Ioannidis, J., & Doucouliagos, C. (2013). What's to know about the credibility of empirical economics? *Journal of Economic Surveys*, *27*(5), 997–1004.

Keller, E., Newman, J. E., Ortmann, A., Jorm, L. R., & Chambers, G. M. (2021). How much is a human life worth? a systematic review. *Value in Health*, *24*(10), 1531–1541.

Kniesner, T. J., & Viscusi, W. K. (2019). The value of a statistical life. In *Oxford research encyclopedia of economics and finance.*

Kochi, I., Hubbell, B., & Kramer, R. (2006). An empirical Bayes approach to combining and comparing estimates of the value of a statistical life for environmental policy analysis. *Environmental & Resource Economics*, *34*(3), 385–406.

Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods*, *2*(1), 61–76.

Levin, A. T., Hanage, W. P., Owusu-Boaitey, N., Cochran, K. B., Walsh, S. P., & Meyerowitz-Katz, G. (2020). Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *European Journal of Epidemiology*, *35*(12), 1123–1138.

Lindhjem, H., Navrud, S., Braathen, N. A., & Biausque, V. (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.

Masterman, C. J., & Viscusi, W. K. (2018). The income elasticity of global values of a statistical life: stated preference evidence. *Journal of Benefit-Cost Analysis*, *9*(3), 407–434.

Masterman, C. J., & Viscusi, W. K. (2020). Publication selection biases in stated preference estimates of the value of a statistical life. *Journal of Benefit-Cost Analysis*, *11*(3), 357–379.

Mrozek, J. R., & Taylor, L. O. (2002). What determines the value of life? a meta-analysis. *Journal of Policy Analysis and Management*, *21*(2), 253–270.

Nelson, J., & Kennedy, P. (2009). The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment. *Environmental Resource Economics*, *42*, 345–377.

Nelson, J. P. (2015). Meta-analysis: statistical methods. In *Benefit transfer of environmental and resource values* (pp. 329–356). Springer.

Rasmusen, E. B. (2015). Understanding shrinkage estimators: from zero to oracle to james-stein. *Kelley School of Business Research Paper*(15-76).

Reed, W. R. (2019). Meta-analysis and publication bias: How well does the FAT-PET-PEESE procedure work? A reply to Hong (International Journal for Re-Views in

Empirical Economics, 2019). *International Journal for Re-Views in Empirical Economics (IREE)*, *3*(2019-5), 1–4.

Ren, S., Lai, H., Tong, W., Aminzadeh, M., Hou, X., & Lai, S. (2010). Nonparametric bootstrapping for hierarchical data. *Journal of Applied Statistics*, *37*, 1487–1498.

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? a meta-analysis. *Psychological Science*, *29*(8), 1358–1369.

Robinson, L. A., & Hammitt, J. K. (2016). Valuing reductions in fatal illness risks: implications of recent research. *Health Economics*, *25*(8), 1039–1052.

Robinson, L. A., Hammitt, J. K., & O'Keeffe, L. (2019). Valuing mortality risk reductions in global benefit-cost analysis. *Journal of Benefit-Cost Analysis*, *10*(S1), 15–50.

Samworth, R. J. (2012). Stein's paradox. *Eureka*, *62*, 38–41.

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78.

Steel, M. F. (2020). Model averaging and its use in economics. *Journal of Economic Literature*, *58*(3), 644–719.

Sunstein, C. R. (2013). The value of a statistical life: some clarifications and puzzles. *Journal of Benefit-Cost Analysis*, *4*(2), 237–261.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods*, *20*(3), 375–393.

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

USEPA. (2010). Valuing mortality risk reductions for environmental policy: a white paper. (https://www.epa.gov/environmental-economics/valuing-mortality-risk-reductions-environmental-policy-white-paper-2010)

USEPA. (2014). *Guidelines for Preparing Economic Analyses.* (https://www.epa.gov/sites/default/files/2017-08/documents/ee-0568-50.pdf)

USEPA. (2016). *Valuing mortality risk for policy: a meta-analytic approach* (Tech. Rep.). (https://sab.epa.gov/ords/sab/f?p=114:0:12222657115986:APPLICATION_PROCESS=AGENCY_REVIEW:::AR_ID:2437)

USEPA. (2017). *Sab review of epa's proposed methodology for updating mortality risk valuation estimates for policy analysis* (Tech. Rep.). (https://sab.epa.gov/ords/sab/f?p=114:0:10098920973528:APPLICATION_PROCESS=REPORT_DOC:::REPORT_ID:1047)

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Viechtbauer, W. (2015). Package 'metafor'. *The Comprehensive R Archive Network*. (http://cran.r-project.org/web/packages/metafor/metafor.pdf)

Viscusi, W. K. (2012). What's to know? puzzles in the literature on the value of statistical life. *Journal of Economic Surveys*, *26*(5), 763–768.

Viscusi, W. K. (2015). The role of publication selection bias in estimates of the value of a statistical life. *American Journal of Health Economics*, *1*(1), 27–52.

Viscusi, W. K. (2018). Best estimate selection bias in the value of a statistical life. *Journal of Benefit-Cost Analysis*, *9*(2), 205–246.

Viscusi, W. K., & Aldy, J. E. (2003). The value of a statistical life: a critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, *27*(1), 5–76.

Viscusi, W. K., & Masterman, C. (2017a). Anchoring biases in international estimates of the value of a statistical life. *Journal of Risk and Uncertainty*, *54*(2), 103–128.

Viscusi, W. K., & Masterman, C. J. (2017b). Income elasticities and global values of a statistical life. *Journal of Benefit-Cost Analysis*, *8*(2), 226–250.

Figures and Tables

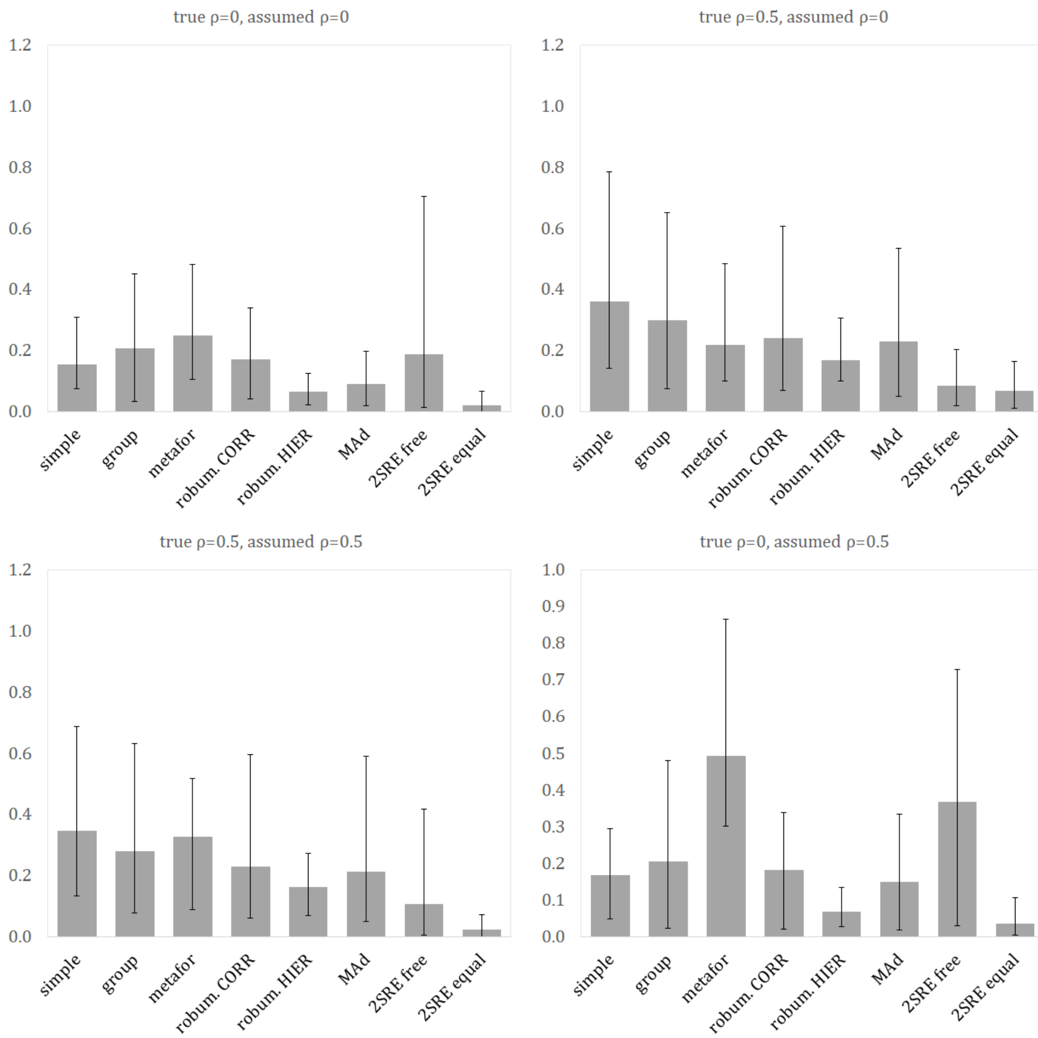

Figure 1: Estimator performance comparisons. Bar heights are average and error bars are min and max $(\hat{se} - se_{min})/se_{min}$—so lower heights indicate better performance—over the 16 experimental configurations examined in the Monte Carlo experiments listed in Tables 3–6.

Table 1: Summary statistics from 13 previous VSL meta-analysis studies. I is the number of groups, N is the number of observations, HW and SP indicate whether hedonic wage or stated preference estimates were included, and VSL range indicates the smallest and largest VSL estimates reported in the original article converted to 2020 US$.

Authors (year)	I	N	HW	SP	VSL range
Mrozek and Taylor (2002)	33	203	1	0	2.4 – 4.0
Viscusi and Aldy (2003)	49	49	1	0	7.5 – 9.3
Kochi et al. (2006)	40	197	1	1	4.2 – 14.4
Bellavance et al. (2009)	37	39	1	0	7.5 – 12.6
Lindhjem et al. (2011)	95	856	0	1	2.0 – 9.8
Dekker et al. (2011)	26	77	0	1	3.3 – 10.7
Doucouliagos et al. (2012)	37	39	1	0	1.3 – 2.8
Nelson (2015)	28	28	1	0	6.0 – 13.9
Viscusi (2015)	17	550	1	0	10.6 – 12.2
Masterman and Viscusi (2017)	68	1025	1	0	9.5 – 11.5
Viscusi (2018)	68	1025	1	0	8.8 – 12.4
Masterman and Viscusi (2020)	85	1148	0	1	0.2 – 1.0

Table 2: All equations necessary to compute the two-stage random-effects (2SRE) estimator listed in a feasible sequence. See the Appendix for a complete derivation.

$\hat{\sigma}^2_{\mu,i} = \frac{1}{J_i - 1} \sum_{j=1}^{J_i} \left(y_{i,j} - \frac{1}{J_i} \sum_{j=1}^{J_i} y_{i,j} \right) - \frac{1}{J_i} \sum_{j=1}^{J_i} \left(se^2_{i,j} - \frac{1}{J_i} \sum_{k \neq j}^{J_i} \rho_i se_{ij} se_{i,k} \right)$	2.1
$\hat{A}_{i,j} = \left(\left[\hat{\sigma}^2_{\mu,i} + (1-\rho_i)\, se^2_{i,j} \right] \sum_{j=1}^{J_i} \frac{1}{\hat{\sigma}^2_{\mu,i} + (1-\rho_i)\, se^2_{i,k}} \right)^{-1}$ $\hat{\mathbf{A}}_i = \left[\hat{A}_{i,1} \;\; \hat{A}_{i,2} \;\; \cdots \;\; \hat{A}_{i,J_i} \right]'$	2.2
$\hat{B}_{i,j,k} = \rho_i \left(\frac{\sum_{\ell=1}^{J_i} \frac{se_{i,\ell}}{\hat{\sigma}^2_{\mu,i} + (1-\rho_i) se^2_{i,\ell}}}{\sum_{\ell=1}^{J_i} \frac{\hat{\sigma}^2_{\mu,i} + (1-\rho_i) se^2_{i,j}}{\hat{\sigma}^2_{\mu,i} + (1-\rho_i) se^2_{i,\ell}}} - \frac{se_{i,j}}{\hat{\sigma}^2_{\mu,i} + (1-\rho_i)\, se^2_{i,j}} \right) se_{i,k}$ $\hat{\mathbf{B}}_i = \begin{bmatrix} \hat{B}_{i,1,1} & \hat{B}_{i,1,2} & \cdots & \hat{B}_{i,1,J_i} \\ \hat{B}_{i,2,1} & \hat{B}_{i,2,2} & \cdots & \hat{B}_{i,2,J_i} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{B}_{i,J_i,1} & \hat{B}_{i,J_i,2} & \cdots & \hat{B}_{i,J_i,J,i} \end{bmatrix}$	2.3
$\hat{\mathbf{g}}_i \equiv [\hat{g}_{i,1} \;\; \hat{g}_{i,2} \;\; \cdots \hat{g}_{i,J_i}]' = \left(\mathbf{I}_i - \hat{\mathbf{B}}_i \right)^{-1} \hat{\mathbf{A}}_i$	2.4
$\hat{y}_i = \sum_{j=1}^{J_i} \hat{g}_{i,j} y_{i,j}$	2.5
$\hat{\sigma}^2_{\eta} = \frac{1}{I-1} \sum_{i=1}^{I} \left(\hat{y}_i - \frac{1}{I} \sum_{i=1}^{I} \hat{y}_i \right)^2$ $-\frac{1}{I} \sum_{i=1}^{I} \sum_{j=1}^{J_i} \left[\hat{g}^2_{i,j} \left(\hat{\sigma}^2_{\mu,i} + se^2_{i,j} \right) \rho_i \sum_{k \neq j}^{J_i} \hat{g}_{i,j} \hat{g}_{i,k} se_{i,j} se_{i,k} \right]$	2.6
$\hat{v}_i = \hat{\sigma}^2_{\eta} + \sum_{j=1}^{J_i} \left[\hat{g}^2_{i,j} \left(\hat{\sigma}^2_{\mu,i} + se^2_{i,j} \right) \rho_i \sum_{k \neq j}^{J_i} \hat{g}_{i,j} \hat{g}_{i,k} se_{i,j} se_{i,k} \right]$	2.7
$\hat{h}_i = \frac{\hat{v}_i^{-1}}{\sum_{q=1}^{I} \hat{v}_q^{-1}}$	2.8
$\hat{w}_{i,j} = \hat{g}_{i,j} \hat{h}_i$	2.9
$\hat{y} = \sum_{i=1}^{I} \hat{h}_i \hat{y}_i = \sum_{i=1}^{I} \sum_{j=1}^{J_i} \hat{w}_{i,j} y_{i,j}$	2.10

Note: When conducting a meta-regression with moderator variables, $x_{i,j}$, the first term on the right-hand side of equation (2.1), should be replaced with $\frac{1}{J_i - 1} \sum_{j=1}^{J_i} \left[\left(y_{i,j} - \sum_{j=1}^{J_i} y_{i,j} \right)^2 - \left(x_{i,j} \hat{\beta}_j - \sum_{j=1}^{J_i} x_{i,j} \hat{\beta}_j \right)^2 \right]$. The resulting $\hat{w}_{i,j}$'s can be used as weights in a weighted-least-squares estimation of β. Start at $\hat{\beta} = 0$, iterate the entire sequence of computations updating $\hat{\beta}$ each iteration, and repeat until convergence.

Table 3: Alternative meta-analysis estimators compared in Monte Carlo simulation experiments.

Estimator	**Description**
simple mean	Unweighted mean of all observations
group means	Unweighted mean of group means
metafor	`R` package for estimating meta-regression models (Viechtbauer, 2010, 2015). Uses restricted maximum likelihood to estimate σ_η assuming normally distributed errors. Accounts for $\hat{\rho}$ in the `rma.mv` function via user-specified variance-covariance matrix of sampling errors. Does not estimate within-group non-sampling error variances, $\sigma_{\mu,i}$.
robumeta	`R` package for meta-regression with robust (Huber-White) standard errors for non-independent observations (Fisher & Tipton, 2015). We estimated two versions: 1) The correlated effects model (CORR) computes approximate inverse variance weights assuming constant error variances within groups and a user-specified common within-group correlation, ρ. 2) The hierarchical effects model (HIER) corresponds to our 2SRE estimator with $\sigma_{\mu,i} = \sigma_\mu \ \forall i$ and $\rho = 0$, but appears to use different estimators for the error component variances.
MAd	`R` package that provides a wrapper for metafor (Del Re, 2015; Del Re & Hoyt, 2014). Includes a procedure to aggregate dependent observations with a user-specified within-group correlation.
2SRE	Two-stage random-effects meta-analysis estimator, as described in the main text and Appendix. We estimate three versions: 1) "2SRE-true" using inverse variance weights computed with the true error component variances, 2) "2SRE-free" using estimated error component variances with $\sigma_{\mu,i}$ estimated freely for each group, and 3) "2SRE-equal" using estimated error component variances with $\sigma_{\mu,i}$ constrained to be equal for all groups.

Table 4: Monte Carlo simulation experiment results ($\rho = 0$, $\hat{\rho} = 0$)

I	J	σ_η	σ_μ	simple mean	group means	meta-for	robum CORR	robum HIER	MAd	2SRE true	2SRE free	2SRE equal	$\bar{se}$
20	1, 5	1.0	0.5,1.0	0.479	0.480	0.405	0.478	**0.378**	0.423	0.366	0.492	0.383	0.392
20	1, 5	1.0	0.5,3.0	0.518	0.534	0.575	0.532	**0.463**	0.499	0.436	0.623	0.465	0.477
20	1, 5	3.0	0.5,1.0	0.834	0.839	0.976	0.824	0.813	0.809	0.771	1.104	**0.774**	0.792
20	1, 5	3.0	0.5,3.0	0.898	0.907	1.176	0.864	0.881	0.842	0.813	0.856	**0.825**	0.859
20	1,15	1.0	0.5,1.0	0.350	0.398	0.319	0.376	0.299	0.338	0.282	0.369	**0.291**	0.304
20	1,15	1.0	0.5,3.0	0.396	0.483	0.438	0.448	0.367	0.387	0.350	0.389	**0.361**	0.371
20	1,15	3.0	0.5,1.0	0.811	0.740	0.837	0.740	0.799	0.727	0.710	0.721	**0.711**	0.740
20	1,15	3.0	0.5,3.0	0.862	0.831	0.940	0.828	0.854	0.804	0.781	0.814	**0.789**	0.772
60	1, 5	1.0	0.5,1.0	0.274	0.298	0.248	0.270	**0.217**	0.252	0.212	0.236	0.220	0.223
60	1, 5	1.0	0.5,3.0	0.300	0.334	0.375	0.316	0.269	0.288	0.253	0.431	**0.267**	0.268
60	1, 5	3.0	0.5,1.0	0.504	0.498	0.596	0.497	0.493	0.485	0.468	0.478	**0.469**	0.442
60	1, 5	3.0	0.5,3.0	0.501	0.505	0.616	0.493	0.489	0.480	0.464	0.482	**0.467**	0.478
60	1,15	1.0	0.5,1.0	0.208	0.248	0.191	0.229	0.177	0.202	0.171	0.182	**0.174**	0.173
60	1,15	1.0	0.5,3.0	0.219	0.245	0.242	0.249	0.204	0.212	0.193	0.245	**0.197**	0.199
60	1,15	3.0	0.5,1.0	0.463	0.431	0.524	0.432	0.462	0.426	0.412	0.418	**0.413**	0.411
60	1,15	3.0	0.5,3.0	0.465	0.446	0.523	0.455	0.467	0.440	0.432	0.455	**0.433**	0.424

Table 5: Monte Carlo simulation experiment results ($\rho = 0.5$, $\hat{\rho} = 0$)

I	J	σ_η	σ_μ	simple mean	group means	meta-for	robum CORR	robum HIER	MAd	2SRE true	2SRE free	2SRE equal	$\hat{se}$
20	1, 5	1.0	0.5,1.0	0.629	0.625	**0.464**	0.569	0.481	0.563	0.422	0.473	0.468	0.483
20	1, 5	1.0	0.5,3.0	0.622	0.639	0.530	0.558	0.503	0.571	0.445	0.535	**0.494**	0.496
20	1, 5	3.0	0.5,1.0	0.977	0.893	0.972	0.852	0.894	0.856	0.779	0.807	**0.798**	0.800
20	1, 5	3.0	0.5,3.0	1.015	0.965	1.304	0.957	0.968	0.929	0.879	0.914	**0.893**	0.918
20	1,15	1.0	0.5,1.0	0.642	0.594	**0.402**	0.578	0.470	0.552	0.360	0.425	0.419	0.415
20	1,15	1.0	0.5,3.0	0.535	0.551	0.458	0.523	0.456	0.519	0.385	0.429	**0.430**	0.440
20	1,15	3.0	0.5,1.0	0.894	0.824	0.836	0.832	0.853	0.824	0.732	0.751	**0.743**	0.742
20	1,15	3.0	0.5,3.0	0.909	0.822	0.910	0.815	0.865	0.819	0.745	0.759	**0.757**	0.788
60	1, 5	1.0	0.5,1.0	0.361	0.360	**0.270**	0.335	0.279	0.332	0.240	0.274	**0.270**	0.270
60	1, 5	1.0	0.5,3.0	0.374	0.394	0.369	0.345	0.313	0.338	0.281	0.321	**0.311**	0.320
60	1, 5	3.0	0.5,1.0	0.521	0.489	0.593	0.475	0.505	0.473	0.444	0.456	**0.449**	0.462
60	1, 5	3.0	0.5,3.0	0.547	0.515	0.616	0.515	0.528	0.503	0.479	0.492	**0.484**	0.491
60	1,15	1.0	0.5,1.0	0.333	0.288	0.224	0.283	0.243	0.284	0.191	0.215	**0.210**	0.220
60	1,15	1.0	0.5,3.0	0.339	0.335	0.261	0.317	0.279	0.311	0.226	**0.254**	0.257	0.254
60	1,15	3.0	0.5,1.0	0.537	0.483	0.533	0.484	0.511	0.482	0.430	0.440	**0.436**	0.430
60	1,15	3.0	0.5,3.0	0.555	0.502	0.540	0.495	0.528	0.495	0.452	0.465	**0.460**	0.457

Table 6: Monte Carlo simulation experiment results ($\rho = 0.5$, $\hat{\rho} = 0.5$)

I	J	σ_η	σ_μ	simple mean	group means	meta-for	robum CORR	robum HIER	MAd	2SRE true	2SRE free	2SRE equal	$\hat{se}$
20	1, 5	1.0	0.5,1.0	0.629	0.625	0.500	0.577	0.493	0.560	0.436	0.476	**0.454**	0.448
20	1, 5	1.0	0.5,3.0	0.630	0.615	0.726	0.577	0.542	0.575	0.478	0.678	**0.513**	0.510
20	1, 5	3.0	0.5,1.0	0.933	0.871	1.082	0.847	0.886	0.838	0.798	0.818	**0.798**	0.802
20	1, 5	3.0	0.5,3.0	1.017	0.985	1.219	0.973	0.961	0.964	0.898	0.938	**0.907**	0.896
20	1,15	1.0	0.5,1.0	0.598	0.579	0.398	0.566	0.452	0.564	0.355	0.399	**0.362**	0.367
20	1,15	1.0	0.5,3.0	0.620	0.587	0.613	0.550	0.517	0.540	0.406	0.520	**0.433**	0.433
20	1,15	3.0	0.5,1.0	0.909	0.826	0.986	0.822	0.846	0.819	0.731	0.736	**0.730**	0.743
20	1,15	3.0	0.5,3.0	0.971	0.865	1.209	0.863	0.946	0.855	0.803	0.903	**0.806**	0.835
60	1, 5	1.0	0.5,1.0	0.373	0.365	0.300	0.325	0.282	0.319	0.252	0.281	**0.259**	0.257
60	1, 5	1.0	0.5,3.0	0.378	0.368	0.409	0.347	0.311	0.337	0.275	0.310	**0.289**	0.297
60	1, 5	3.0	0.5,1.0	0.520	0.489	0.614	0.484	0.496	0.481	0.438	0.463	**0.439**	0.450
60	1, 5	3.0	0.5,3.0	0.558	0.530	0.723	0.520	0.541	0.519	0.490	0.527	**0.497**	0.499
60	1,15	1.0	0.5,1.0	0.341	0.309	0.220	0.299	0.254	0.290	0.202	0.216	**0.205**	0.205
60	1,15	1.0	0.5,3.0	0.361	0.366	0.313	0.349	0.296	0.339	0.246	0.271	**0.257**	0.252
60	1,15	3.0	0.5,1.0	0.525	0.496	0.496	0.486	0.499	0.483	0.430	0.439	**0.431**	0.423
60	1,15	3.0	0.5,3.0	0.564	0.507	0.592	0.505	0.540	0.504	0.453	0.473	**0.456**	0.450

Table 7: Monte Carlo simulation experiment results ($\rho = 0$, $\hat{\rho} = 0.5$)

I	J	σ_η	σ_μ	simple mean	group means	meta-for	robum CORR	robum HIER	MAd	2SRE true	2SRE free	2SRE equal	$\hat{se}$
20	1, 5	1.0	0.5,1.0	0.468	0.537	0.492	0.486	**0.375**	0.453	0.363	0.513	0.382	0.381
20	1, 5	1.0	0.5,3.0	0.503	0.513	0.726	0.498	0.421	0.477	0.398	0.669	**0.417**	0.430
20	1, 5	3.0	0.5,1.0	0.905	0.836	1.164	0.837	0.851	0.821	0.770	0.859	**0.780**	0.795
20	1, 5	3.0	0.5,3.0	0.882	0.883	1.170	0.881	0.847	0.863	0.803	1.026	**0.821**	0.835
20	1,15	1.0	0.5,1.0	0.333	0.341	0.367	0.350	0.287	0.328	0.274	0.371	**0.281**	0.292
20	1,15	1.0	0.5,3.0	0.381	0.482	0.542	0.434	**0.352**	0.436	0.327	0.565	0.362	0.366
20	1,15	3.0	0.5,1.0	0.820	0.745	0.990	0.758	0.803	0.750	0.720	0.960	**0.727**	0.726
20	1,15	3.0	0.5,3.0	0.849	0.761	1.018	0.769	0.832	0.764	0.733	0.782	**0.749**	0.746
60	1, 5	1.0	0.5,1.0	0.299	0.327	0.315	0.306	0.238	0.292	0.231	0.315	**0.235**	0.233
60	1, 5	1.0	0.5,3.0	0.314	0.341	0.479	0.337	**0.276**	0.323	0.257	0.439	0.279	0.284
60	1, 5	3.0	0.5,1.0	0.482	0.469	0.576	0.467	0.463	0.461	0.435	0.663	**0.441**	0.430
60	1, 5	3.0	0.5,3.0	0.471	0.472	0.735	0.470	0.462	0.463	0.449	0.581	**0.456**	0.466
60	1,15	1.0	0.5,1.0	0.199	0.222	0.216	0.222	0.175	0.214	0.166	0.236	**0.171**	0.170
60	1,15	1.0	0.5,3.0	0.233	0.259	0.379	0.262	**0.217**	0.250	0.206	0.312	0.223	0.207
60	1,15	3.0	0.5,1.0	0.461	0.428	0.539	0.432	0.457	0.430	0.412	0.428	**0.414**	0.411
60	1,15	3.0	0.5,3.0	0.446	0.429	0.566	0.428	0.442	0.427	0.419	0.432	**0.426**	0.418

Table 8: Meta-analysis estimates of the average VSL among U.S. adults [2020$US] based on the preliminary meta-dataset compiled by USEPA (2016). Alternative estimators are indicated by the row labels. "mm" indicates that both mean and median VSL observations were included; "m" indicates that only mean VSL observations were included. "HW" indicates hedonic wage observations, "SP" indicates stated preference observations, "pooled" indicates both HW and SP observations, and "balanced" indicates a mixture of the independent HW and SP results with each given equal weight. Numbers in parentheses are bootstrapped standard errors, and numbers in brackets are root mean squared errors (RMSE's) where the bias was estimated as the difference between the "mm" and "m" estimates. The "m" estimates are assumed to be unbiased, in which cases root mean squared errors are equal to standard errors. Four estimates with the lowest RMSE's are highlighted in bold font.

Estimator	m(m)	HW		SP			pooled			balanced		
simple mean	mm	12.47	(1.77)	7.65	(1.16)	[1.58]	10.17	(1.49)	[1.73]	10.06	(0.73)	[0.91]
	m	12.47	(1.77)	8.72	(2.12)		11.05	(1.64)		10.60	(0.97)	
group means	mm	11.05	(1.33)	8.14	(1.12)	[1.56]	9.59	(0.94)	[1.18]	**9.59**	(0.61)	[0.82]
	m	11.05	(1.33)	9.22	(1.85)		10.32	(1.08)		10.14	(0.79)	
2SRE–free	mm	8.45	(1.02)	6.64	(0.69)	[1.50]	**7.59**	(0.54)	[0.82]	**7.54**	(0.43)	[0.79]
	m	8.45	(1.02)	7.97	(1.64)		8.21	(0.85)		8.21	(0.67)	
–equal	mm	8.68	(0.94)	7.20	(1.28)	[1.94]	7.99	(0.86)	[1.25]	7.94	(0.56)	[0.92]
	m	8.68	(0.94)	8.66	(1.82)		8.90	(1.03)		8.67	(0.69)	
–free T&F	mm	8.09	(0.94)	4.65	(1.08)	[3.62]	6.66	(1.04)	[1.04]	6.37	(0.51)	[1.80]
	m	8.09	(0.94)	8.11	(2.57)		6.61	(1.72)		8.10	(0.88)	
–equal T&F	mm	8.11	(0.94)	4.35	(1.72)	[1.81]	6.14	(1.32)	[1.32]	**6.23**	(0.66)	[0.72]
	m	8.11	(0.94)	4.90	(2.22)		6.04	(1.55)		6.51	(0.79)	
–P-P	mm	7.85	(1.11)	6.28	(3.11)	[4.28]	6.34	(3.15)	[4.27]	7.06	(1.06)	[1.81]
	m	7.85	(1.11)	9.22	(3.52)		9.22	(3.39)		8.53	(1.16)	

Table 9: EPA data meta-regression results with no correction for publication bias. Seven specifications (S0-S6) of the two-stage random-effects meta-regression model with σ_μ constrained (2SRE-equal) and with income elasticity of VSL *(IEVSL)* for models including income. Numbers in parentheses are robust standard errors.

	S0	S1	S2	S3	S4	S5	S6
constant	7.986	9.275	10.923	9.027	10.840	13.665	10.027
	(0.901)	(0.979)	(0.863)	(1.373)	(1.227)	(0.448)	(0.930)
SP		-0.961	-3.394	-0.495	-3.240	-2.991	-0.378
		(2.305)	(1.390)	(2.477)	(1.789)	(1.255)	(2.235)
median		-2.152	-1.175	-2.279	-1.218	-1.038	-2.927
		(2.158)	(1.337)	(2.093)	(1.361)	(1.343)	(2.004)
year			0.610		0.609	0.421	
			(0.143)		(0.143)	(0.112)	
income				-0.190	-0.061		0.584
				(0.573)	(0.540)		(0.222)
SP×year						0.277	
						(0.216)	
SP×income							-1.747
							(0.762)
σ_μ	2.335	2.327	2.251	2.320	2.248	2.252	2.275
σ_η	2.728	2.731	2.756	2.733	2.757	2.756	2.748
IEVSL				-0.110	-0.035		0.338
				(0.332)	(0.312)		(0.129)
R^2	0.477	0.533	0.699	0.536	0.699	0.706	0.571
R^2_{CV}	0.454	0.453	0.637	0.442	0.622	0.631	0.472

Table 10: EPA data meta-regression results with the "precision-effect test" (PET) for publication bias. Seven specifications (S0-S6) of the two-stage random-effects meta-regression model with σ_μ constrained (2SRE-equal) and with *IEVSL* for models including income. Numbers in parentheses are robust standard errors.

	S0	S1	S2	S3	S4	S5	S6
constant	6.418	6.871	9.112	6.585	8.991	10.476	7.612
	(1.495)	(1.268)	(1.452)	(1.590)	(1.686)	(1.685)	(1.541)
SP		0.624	-2.145	1.157	-1.924	-1.418	0.947
		(2.110)	(1.373)	(2.245)	(1.729)	(1.266)	(2.137)
median		-2.248	-1.285	-2.393	-1.345	-1.109	-2.769
		(2.270)	(1.368)	(2.218)	(1.381)	(1.383)	(2.204)
year			0.572		0.571	0.300	
			(0.178)		(0.180)	(0.128)	
income				-0.217	-0.087		0.256
				(0.487)	(0.498)		(0.335)
SP×year						0.389	
						(0.247)	
SP×income							-1.062
							(0.821)
se	1.005	0.938	0.667	0.939	0.668	0.752	0.776
	(0.512)	(0.425)	(0.417)	(0.432)	(0.420)	(0.387)	(0.462)
σ_μ	2.169	2.159	2.105	2.145	2.095	2.103	2.134
σ_η	2.784	2.787	2.804	2.791	2.807	2.805	2.795
IEVSL				-0.125	-0.050		0.148
				(0.282)	(0.288)		(0.194)
R^2	0.572	0.599	0.736	0.602	0.738	0.750	0.613
R^2_{CV}	0.527	0.507	0.653	0.500	0.641	0.659	0.504

Table 11: EPA data meta-regression results with the "precision-effect estimate with SE" (PEESE) for publication bias. Seven specifications (S0-S6) of the two-stage random-effects meta-regression model with σ_μ constrained (2SRE-equal) and with *IEVSL* for models including income. Numbers in parentheses are robust standard errors.

	S0	S1	S2	S3	S4	S5	S6
constant	7.382	8.225	10.025	7.948	9.916	11.892	8.923
	(0.982)	(0.943)	(0.966)	(1.321)	(1.257)	(0.813)	(0.987)
SP		-0.137	-2.659	0.382	-2.459	-2.069	0.377
		(2.238)	(1.339)	(2.389)	(1.690)	(1.219)	(2.213)
median		-2.048	-1.127	-2.189	-1.181	-0.943	-2.755
		(2.210)	(1.349)	(2.150)	(1.373)	(1.351)	(2.085)
year			0.587		0.585	0.339	
			(0.153)		(0.154)	(0.111)	
income				-0.211	-0.079		0.453
				(0.549)	(0.530)		(0.292)
SP×year						0.359	
						(0.224)	
SP×income							-1.490
							(0.800)
se^2	0.122	0.103	0.108	0.103	0.083	0.090	0.092
	(0.045)	(0.038)	(0.045)	(0.038)	(0.048)	(0.042)	(0.038)
σ_μ	2.201	2.223	2.193	2.213	2.187	2.191	2.209
σ_η	2.773	2.766	2.775	2.769	2.777	2.776	2.770
IEVSL				-0.122	-0.046		0.262
				(0.318)	(0.307)		(0.169)
R^2	0.554	0.582	0.728	0.585	0.729	0.740	0.607
R^2_{CV}	0.513	0.488	0.641	0.478	0.627	0.646	0.498

Table 12: EPA data meta-regression results with no correction for publication bias. Seven specifications (S0-S6) of the two-stage random-effects meta-regression model with σ_μ unconstrained (2SRE-free) and with *IEVSL* for models including income. Numbers in parentheses are robust standard errors.

	0	1	2	3	4	5	6
constant	7.594	8.720	10.118	8.397	9.989	13.931	9.385
	(0.682)	(1.059)	(0.809)	(1.425)	(1.113)	(0.415)	(1.061)
SP		-1.625	-2.850	-1.041	-2.628	-2.950	-0.952
		(1.495)	(1.420)	(1.974)	(1.668)	(1.268)	(1.440)
median		-0.740	-1.186	-0.968	-1.263	-1.167	-1.907
		(1.394)	(1.242)	(1.353)	(1.285)	(1.253)	(0.951)
year			0.451		0.447	0.481	
			(0.097)		(0.095)	(0.111)	
income				-0.245	-0.088		0.494
				(0.520)	(0.480)		(0.294)
SP×year						-0.049	
						(0.178)	
SP×income							-1.761
							(0.538)
σ_μ	2.541	2.533	2.454	2.527	2.451	2.455	2.485
σ_η	2.183	2.187	2.195	2.190	2.196	2.194	2.209
IEVSL				-0.142	-0.051		0.286
				(0.301)	(0.277)		(0.170)
R^2	0.369	0.420	0.546	0.424	0.547	0.547	0.475
R^2_{CV}	0.352	0.355	0.493	0.341	0.477	0.485	0.397

Table 13: EPA data meta-regression results with the "precision-effect test" (PET) for publication bias. Seven specifications (S0-S6) of the two-stage random-effects meta-regression model with σ_μ unconstrained (2SRE-free) and with *IEVSL* for models including income. Numbers in parentheses are robust standard errors.

	0	1	2	3	4	5	6
constant	5.413	5.832	7.422	5.366	7.060	9.504	6.016
	(0.758)	(1.065)	(1.306)	(1.152)	(1.392)	(1.702)	(1.140)
SP		-0.266	-1.439	0.556	-0.862	-1.279	0.503
		(1.209)	(1.272)	(1.484)	(1.447)	(1.204)	(1.351)
median		-0.524	-0.876	-0.832	-1.060	-0.898	-1.241
		(1.144)	(1.070)	(1.136)	(1.111)	(1.084)	(1.092)
year			0.335		0.323	0.291	
			(0.121)		(0.128)	(0.127)	
income				-0.342	-0.221		-0.034
				(0.364)	(0.384)		(0.364)
SP×year						0.070	
						(0.174)	
SP×income							-0.742
							(0.596)
se	1.329	1.257	1.018	1.267	1.033	1.030	1.157
	(0.361)	(0.353)	(0.402)	(0.360)	(0.407)	(0.399)	(0.380)
σ_μ	2.368	2.360	2.308	2.344	2.297	2.306	2.337
σ_η	2.195	2.201	2.202	2.207	2.207	2.203	2.213
IEVSL				-0.198	-0.128		-0.019
				(0.211)	(0.222)		(0.211)
R^2	0.598	0.602	0.657	0.609	0.664	0.659	0.615
R^2_{CV}	0.572	0.549	0.601	0.547	0.595	0.593	0.552

Table 14: EPA data meta-regression results with the "precision-effect estimate with SE" (PEESE) for publication bias. Seven specifications (S0-S6) of the two-stage random-effects meta-regression model with σ_μ unconstrained (2SRE-free) and with *IEVSL* for models including income. Numbers in parentheses are robust standard errors.

	0	1	2	3	4	5	6
constant	6.884	7.635	9.047	7.239	8.811	11.763	8.083
	(0.646)	(0.962)	(0.887)	(1.218)	(1.067)	(0.814)	(0.985)
SP		-0.958	-2.152	-0.241	-1.754	-2.041	-0.189
		(1.290)	(1.299)	(1.674)	(1.489)	(1.185)	(1.315)
median		-0.485	-0.913	-0.761	-1.046	-0.930	-1.526
		(1.290)	(1.163)	(1.243)	(1.206)	(1.175)	(1.008)
year			0.397		0.389	0.365	
			(0.100)		(0.102)	(0.109)	
income				-0.302	-0.156		0.271
				(0.458)	(0.449)		(0.345)
SP×year						0.051	
						(0.171)	
SP×income							-1.383
							(0.576)
se^2	0.144	0.128	0.108	0.129	0.109	0.109	0.118
	(0.042)	(0.040)	(0.045)	(0.040)	(0.044)	(0.043)	(0.038)
σ_μ	2.407	2.428	2.392	2.418	2.386	2.390	2.413
σ_η	2.182	2.185	2.192	2.188	2.194	2.192	2.202
IEVSL				-0.175	-0.090		0.157
				(0.265)	(0.259)		(0.200)
R^2	0.518	0.532	0.620	0.539	0.622	0.620	0.567
R^2_{CV}	0.478	0.456	0.546	0.444	0.530	0.538	0.485

Appendix

In this appendix we derive the two-stage random-effects estimator used in the main text. The dataset comprises $i = 1, 2, ..., I$ groups of observations, where group i comprises $j = 1, 2, ..., J_i$ individual observations.[13] We will denote the observations as y_{ij} and their associated standard errors as se_{ij}. The estimator will take the form of a weighted mean, $\hat{y} = \sum_{i=1}^{I} \sum_{j=1}^{J_i} w_{ij} y_{ij}$, where the weights sum to one, $\sum_{i=1}^{I} \sum_{j=1}^{J_i} w_{ij} = 1$. Our task is to find the optimal weights given the structure of our data set and our assumptions about the nature of the data generating process.

4.1 Sources of error

We decompose each observation into the sum of the true mean and three error components:

$$y_{ij} = y + \eta_i + \mu_{ij} + \varepsilon_{ij}, \tag{3}$$

where y is the true value of the average VSL among the U.S. adult general population (our target of estimation), η_i is a group-level non-sampling error, μ_{ij} is an observation-level non-sampling error, and ε_{ij} is an observation-level sampling error. η_i varies among but not within groups, while both μ_{ij} and ε_{ij} vary both among and within groups. (Below we describe how to generalize this meta-analysis model to a meta-regression model, which basically involves replacing y with $x_{ij}\beta$ throughout.)

Before proceeding with the derivation, it might be helpful to explain our assumptions about the nature of the error components. Sampling errors, represented by ε_{ij}, arise from sampling variability alone. This refers to the variability of a statistic if it were calculated

13. Our primary grouping strategy groups estimates based on the same underlying dataset, but other strategies are possible, e.g., grouping by study or primary author. The aim is to group observations such that correlations across groups are eliminated or minimized.

many times repeating the same study design with the same sample size but with a different random draw of observations from the target population each time. Non-sampling errors include all other sources of deviation between the estimate calculated from the sample and the true quantity that is the target of estimation, such as measurement error, missing variable bias, other forms of model mis-specification, mis-matches between the sampling frame and the population of interest, ad hoc treatment of "outliers," and other methodological choices that may lead to biased estimates. The practical relevance of this distinction is that the standard errors reported in the original studies represent only the sampling variability of the primary VSL estimates. Therefore, these quantities can serve as estimates of the standard deviations of ε_{ij} but not the other error components in the model, the variances of which must be estimated using the meta-data itself.

We will assume that the non-sampling error components are uncorrelated with each other and with the sampling errors, but we will allow for possible correlations among sampling errors within groups. Also note that the composite non-sampling errors, $\eta_i + \mu_{ij}$, will be correlated within groups but not across groups by the assumption that η_i is common to all observations in group i. We will derive the minimum variance estimator and calculate the associated standard error based on these assumptions. We also will calculate standard errors using both a bootstrap approach and the robust standard errors proposed by L. V. Hedges et al. (2010) to avoid the bias associated with nominal standard errors when the assumed error structure does not correspond to the true error structure.

4.2 A two-stage random effects estimator

To find the optimal weights to place on each observation in the meta-dataset we proceed in two stages. In the first stage we find the optimal feasible weights, $\hat{g}_{ij}$, for calculating composite estimates for each group,[14]

$$\hat{y}_i = \sum_{j=1}^{J_i} \hat{g}_{ij} y_{ij}. \tag{4}$$

In this stage we impose the constraint that $\sum_{j=1}^{J_i} \hat{g}_{ij} = 1$, which is required to make the group-level composite estimate an unbiased estimator of the group mean, $y + \eta_i$. In the second stage we find the optimal weights $\hat{h}_i$ for calculating the overall composite estimate,

$$\hat{y} = \sum_{i=1}^{I} \hat{h}_i \hat{y}_i. \tag{5}$$

In this stage we impose the constraint that $\sum_{i=1}^{I} \hat{h}_i = 1$, which is required to make the expected value of the composite estimate equal to the mean of the group-level effects, and therefore equal to the true effect y by the assumption that the expected value of the group level non-sampling errors η_i is zero. The composite weights for each observation are then $\hat{w}_{ij} = \hat{h}_i \hat{g}_{ij}$.

4.3 Stage one

We begin by finding the optimal infeasible weights for calculating the group-level composite estimates, $\hat{y}_i$. We will denote the variance of the composite estimate for group i as v_i, which is

$$v_i = \mathbb{E}\left[\left(\sum_{j=1}^{J_i} g_{ij} y_{ij}\right)^2\right] - \mathbb{E}\left[\left(\sum_{j=1}^{J_i} g_{ij} y_{ij}\right)\right]^2. \tag{6}$$

14. We will use "ˆ" notation to indicate quantities that can be computed from our data or other estimated quantities. The same symbol without a "ˆ" indicates an infeasible estimator because it depends on one or more unknown population parameters.

Substituting $y_{ij} = y + \eta_i + \mu_{ij} + \varepsilon_{ij}$ gives

$$v_i = \mathbb{E}\left[\left(\sum\nolimits_{j=1}^{J_i} g_{ij}\left(y + \eta_i + \mu_{ij} + \varepsilon_{ij}\right)\right)^2\right] - \mathbb{E}\left[\left(\sum\nolimits_{j=1}^{J_i} g_{ij}\left(y + \eta_i + \mu_{ij} + \varepsilon_{ij}\right)\right)\right]^2. \quad (7)$$

Using the constraint $\sum_{j=1}^{J_i} g_{ij} = 1$, we can factor the y and η_i out of the summation in the first term, and we can simplify the second term to y^2, which gives

$$v_i = \mathbb{E}\left[\left(y + \eta_i + \sum\nolimits_{j=1}^{J_i} g_{ij}\left(\mu_{ij} + \varepsilon_{ij}\right)\right)^2\right] - y^2. \quad (8)$$

By assumption, η_i is not correlated with either μ_{ij} or ε_{ij}, so all terms involving products of η_i and μ_{ij} or η_i and ε_{ij} will equal zero in expectation. All terms involving products of y and any error terms also will equal zero in expectation, so we can simplify equation (8) to

$$v_i = \mathbb{E}\left[\left(y + \eta_i\right)^2 + \left(\sum\nolimits_{j=1}^{J_i} g_{ij}\left(\mu_{ij} + \varepsilon_{ij}\right)\right)^2\right] - y^2. \quad (9)$$

Completing both squares inside the expectation operator and eliminating the y^2 and $-y^2$ terms gives

$$v_i = \mathbb{E}\left[\eta_i^2 + 2y\eta_i + \sum\nolimits_{j=1}^{J_i}\left\{g_{ij}^2\left(\mu_{ij} + \varepsilon_{ij}\right)^2 + \sum\nolimits_{k\neq j}^{J_i} g_{ij}g_{ik}\left(\mu_{ij} + \varepsilon_{ij}\right)\left(\mu_{ik} + \varepsilon_{ik}\right)\right\}\right], \quad (10)$$

where $\sum_{k\neq j}^{J_i}$ indicates the sum from 1 to J_i excluding element j. Note that $\mathbb{E}\big[\big(\mu_{ij}^2 + \varepsilon_{ij}^2\big)\big]$ is the variance of observation ij, and $\mathbb{E}\big[\big(\mu_{ij} + \varepsilon_{ij}\big)\big(\mu_{ik} + \varepsilon_{ik}\big)\big]$ is the covariance between observation ij and ik (both conditional on η_i).

Next we evaluate the right hand side of equation (10) using the assumptions that all error terms are mean zero and all but the sampling errors are uncorrelated, and using the fact that the covariance between two random variables equals their correlation multiplied by their respective standard deviations. This gives

$$v_i = \sigma_\eta^2 + \sum\nolimits_{j=1}^{J_i} \left[g_{ij}^2 \left(\sigma_{\mu,i}^2 + se_{ij}^2 \right) + \rho_i \sum\nolimits_{k \neq j}^{J_i} g_{ij} g_{ik} se_{ij} se_{ik} \right], \tag{11}$$

where ρ_i is the correlation among sampling errors for observations in group i. Equation (11) is the quantity we want to minimize by choosing weights, g_{ij}, subject to the constraint that the weights sum to one. The Lagrangian is

$$\mathcal{L}_i = \sigma_\eta^2 + \sum\nolimits_{j=1}^{J_i} \left[g_{ij}^2 \left(\sigma_{\mu,i}^2 + se_{ij}^2 \right) + \rho_i \sum\nolimits_{k \neq j}^{J_i} g_{ij} g_{ik} se_{ij} se_{ik} \right] - \lambda_i \left(\sum\nolimits_{j=1}^{J_i} g_{ij} - 1 \right), \tag{12}$$

and the first-order conditions for a minimum are

$$\frac{\partial \mathcal{L}_i}{\partial g_{ij}} = 2 g_{ij} \left(\sigma_{\mu,i}^2 + se_{ij}^2 \right) + 2\rho_i \sum\nolimits_{k \neq j}^{J_i} g_{ik} se_{ij} se_{ik} - \lambda_i = 0, \tag{13}$$

for each j in group i.[15] Next, subtracting $2\rho_i \sum_{k \neq j}^{J_i} g_{ik} se_{ij} se_{ik} - \lambda_i$ from both sides of the second equality in expression (13) gives

$$2 g_{ij} \left(\sigma_{\mu,i}^2 + se_{ij}^2 \right) = \lambda_i - 2\rho_i \left(-g_{ij} se_{ij}^2 + \sum\nolimits_{k=1}^{J_i} g_{ik} se_{ij} se_{ik} \right), \tag{14}$$

and then distributing the $2\rho_i$ to the terms inside the parentheses on the right hand side of equation (14) gives

$$2 g_{ij} \left(\sigma_{ij}^2 + se_{ij}^2 \right) = \lambda_i + 2\rho_i g_{ij} se_{ij}^2 - 2\rho_i \sum\nolimits_{k=1}^{J_i} g_{ik} se_{ij} se_{ik}. \tag{15}$$

15. If it is not obvious where the 2 multiplying ρ_i in equation (13) comes from, note that in equation (12) the double summation term, $\sum_{j=1}^{J_i} \sum_{k \neq j}^{J_i} g_{ij} g_{ik} se_{ij} se_{ik}$, is the sum of all off-diagonal elements of the $J_i \times J_i$ matrix formed by cross-multiplying the vector $\mathbf{g}_i \odot \mathbf{se}_i$ by itself, where $\odot$ indicates element-by-element multiplication, i.e., $[\mathbf{g}_i \odot \mathbf{se}_i]'[\mathbf{g}_i \odot \mathbf{se}_i] = \sum_{j=1}^{J_i} \sum_{k=1}^{J_i} g_{ij} g_{ik} se_{ij} se_{ik}$. Taking the derivative of the sum of the off-diagonal terms with respect to any given element of the vector $\mathbf{g}_i$ gives 2 times the sum of all other elements of $\mathbf{g}_i$ because each of these elements appears once below and once above the diagonal of $[\mathbf{g}_i \odot \mathbf{se}_i]'[\mathbf{g}_i \odot \mathbf{se}_i]$.

Next, we subtract $2\rho_i g_{ij} se_{ij}^2$ from both sides of equation (15) to get

$$2g_{ij}\left(\sigma_{\mu,i}^2 + se_{ij}^2\right) - 2\rho_i g_{ij} se_{ij}^2 = \lambda_i - 2\rho_i \sum\nolimits_{k=1}^{J_k} g_{ik} se_{ij} se_{ik}. \tag{16}$$

Then we factor $2g_{ij}$ out of the left hand side of equation (16) to get

$$2g_{ij}\left[\left(\sigma_{\mu,i}^2 + se_{ij}^2\right) - \rho_i se_{ij}^2\right] = \lambda_i - 2\rho_i \sum\nolimits_{k=1}^{J_i} g_{ik} se_{ij} se_{ik}, \tag{17}$$

then solve for g_{ij} by dividing both sides of equation (17) by $2\left[\left(\sigma_{\mu,i}^2 + se_{ij}^2\right) - \rho_i se_{ij}^2\right]$, which gives

$$g_{ij} = \frac{\lambda_i - 2\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}}{2\left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]}. \tag{18}$$

Next, we apply the constraint that the g_{ij}'s must sum to 1 in each group to get

$$1 = \sum_{j=1}^{J_i} \frac{\lambda_i - 2\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}}{2\left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]}, \tag{19}$$

and then separate the term on the right hand side of equation (19) into two sums:

$$1 = \sum_{j=1}^{J_i} \frac{\lambda_i}{2\left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]} - \sum_{j=1}^{J_i} \frac{2\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}}{2\left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]}. \tag{20}$$

Next we factor the Lagrange multiplier out of the sum in the first term and cancel the 2's in the second term on the right hand side of equation (20) to get

$$1 = \lambda_i \sum_{j=1}^{J_i} \frac{1}{2\left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]} - \sum_{j=1}^{J_i} \frac{\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}}{\left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]}. \tag{21}$$

We solve for the Lagrange multiplier by adding the second term on the right hand side to both sides of equation (21) then dividing both sides by the term that multiplies the Lagrange multiplier:

$$\lambda_i = \frac{1 + \sum_{j=1}^{J_i} \frac{\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}}{\left[\sigma_{\mu,i}^2 + (1-\rho_i) se_{ij}^2\right]}}{\sum_{j=1}^{J_i} \frac{1}{2\left[\sigma_{\mu,i}^2 + (1-\rho_i) se_{ij}^2\right]}}. \tag{22}$$

Next, we rearrange equation (22) by factoring $\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}$ out of the sum over j in the numerator and substituting $\frac{1}{\left[\sigma_{\mu,i}^2 + (1-\rho_i) se_{ij}^2\right]} = \left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]^{-1}$ in both the numerator and denominator of (22) to get

$$\lambda_i = \frac{1 + \left(\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}\right) \sum_{j=1}^{J_i} \left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]^{-1}}{\frac{1}{2} \sum_{j=1}^{J_i} \left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]^{-1}}. \tag{23}$$

Then we plug (23) back into the expression for g_{ij} in equation (18) and divide the numerator and denominator by 2 to get

$$g_{ij} = \frac{\frac{1 + \left(\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}\right) \sum_{j=1}^{J_i} \left[\sigma_{\mu,i}^2 + (1-\rho_i) se_{ij}^2\right]^{-1}}{\sum_{j=1}^{J_i} \left[\sigma_{\mu,i}^2 + (1-\rho_i) se_{ij}^2\right]^{-1}} - \rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}}{\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2} \tag{24}$$

whih we can separate into two fractions,

$$g_{ij} = \frac{1 + \left(\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}\right) \sum_{j=1}^{J_i} \left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]^{-1}}{\left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right] \sum_{j=1}^{J_i} \left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]^{-1}} - \frac{\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}}{\left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]}, \tag{25}$$

then we separate the first term on the right hand side into two fractions and rearrange terms in the sum to get:

$$g_{ij} = \frac{1}{\sum_{k=1}^{J_i} \frac{\sigma_{\mu,i}^2 + (1-\rho_i) se_{ij}^2}{\sigma_{\mu,i}^2 + (1-\rho_i) se_{ik}^2}} + \frac{\left(\rho_i \sum_{k=1}^{J_i} se_{ik} g_{ik}\right) \sum_{k=1}^{J_i} \frac{se_{ik}}{\sigma_{\mu,i}^2 + (1-\rho_i) se_{ik}^2}}{\sum_{k=1}^{J_i} \frac{\sigma_{\mu,i}^2 + (1-\rho_i) se_{ij}^2}{\sigma_{\mu,i}^2 + (1-\rho_i) se_{ik}^2}} - \frac{\rho_i se_{ij} \sum_{k=1}^{J_i} se_{ik} g_{ik}}{\left[\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2\right]}. \tag{26}$$

Then we move $\rho_i \sum_{k=1}^{J_i} se_{ik} g_{ik}$ out of the numerator of both fractions in which it appears to get

$$g_{ij} = \frac{1}{\sum_{k=1}^{J_i} \frac{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ij}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}} + \left(\frac{\sum_{k=1}^{J_i} \frac{se_{ik}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}}{\sum_{k=1}^{J_i} \frac{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ij}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}} \right) \rho_i \sum_{k=1}^{J_i} se_{ik} g_{ik} -$$

$$\frac{se_{ij}}{\sigma^2_{\mu,i} + (1-\rho_i)\, se^2_{ij}} \rho_i \sum_{k=1}^{J_i} se_{ik} g_{ik} \tag{27}$$

Distribute terms into the summations in the second and third terms on the right hand side of equation (27) to get

$$g_{ij} = \frac{1}{\sum_{k=1}^{J_i} \frac{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ij}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}} + \rho_i \sum_{k=1}^{J_i} \left(\frac{\sum_{k=1}^{J_i} \frac{se_{ik}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}}{\sum_{k=1}^{J_i} \frac{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ij}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}} \right) se_{ik} g_{ik} -$$

$$\rho_i \sum_{k=1}^{J_i} \left(\frac{se_{ij}}{\sigma^2_{\mu,i} + (1-\rho_i)\, se^2_{ij}} \right) se_{ik} g_{ik}. \tag{28}$$

Next, we combine the second and third terms into a single summation to get

$$g_{ij} = \frac{1}{\sum_{k=1}^{J_i} \frac{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ij}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}} + \rho_i \sum_{k=1}^{J_i} \left(\frac{\sum_{k=1}^{J_i} \frac{se_{ik}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}}{\sum_{k=1}^{J_i} \frac{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ij}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}} - \frac{se_{ij}}{\sigma^2_{\mu,i} + (1-\rho_i)\, se^2_{ij}} \right) se_{ik} g_{ik}. \tag{29}$$

Equation (29) has the form of

$$g_{ij} = A_{ij} + \sum_{k=1}^{J_i} B_{ijk} g_{ik}, \tag{30}$$

where

$$A_{ij} = \frac{1}{\sum_{k=1}^{J_i} \frac{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ij}}{\sigma^2_{\mu,i}+(1-\rho_i)se^2_{ik}}} \tag{31}$$

and

$$B_{ijk} = \rho_i \left(\frac{\sum_{\ell=1}^{J_i} \frac{se_{i\ell}}{\sigma_{\mu,i}^2 + (1-\rho_i) se_{i\ell}^2}}{\sum_{\ell=1}^{J_i} \frac{\sigma_{\mu,i}^2 + (1-\rho_i) se_{ij}^2}{\sigma_{\mu,i}^2 + (1-\rho_i) se_{i\ell}^2}} - \frac{se_{ij}}{\sigma_{\mu,i}^2 + (1-\rho_i)\, se_{ij}^2} \right) se_{ik}. \tag{32}$$

Equation (30) can be written in matrix notation as

$$\mathbf{g}_i = \mathbf{A}_i + \mathbf{B}_i \mathbf{g}_i, \tag{33}$$

so, finally, we can solve for $\mathbf{g}_i$ as follows:

$$(\mathbf{I}_i - \mathbf{B}_i)\, \mathbf{g}_i = \mathbf{A}_i \Rightarrow \mathbf{g}_i = (\mathbf{I}_i - \mathbf{B}_i)^{-1} \mathbf{A}_i. \tag{34}$$

4.4 Cross checks

To provide some indirect confirmation that the above derivation is valid, we can apply the formula for the optimal weights to two simple cases where the weights are straightforward to verify independently. First, consider the case when the sampling errors in a group are uncorrelated. When $\rho_i = 0$, equation (32) implies $\mathbf{B}_i = 0$ and equation (31) implies that the elements of $\mathbf{A}_i$ simplify to

$$A_{ij} = \frac{\left(\sigma_{\mu,i}^2 + se_{ij}^2\right)^{-1}}{\sum_{k=1}^{J_i} \left(\sigma_{\mu,i}^2 + se_{ik}^2\right)^{-1}}. \tag{35}$$

Because $\mathbf{B}_i = 0$, equation (30) implies that $g_{ij} = A_{ij}$. The formula in equation (35) is consistent with the weights given by Hedges and Olkin (1985 p 199) for a traditional random effects (RE) model with one observation per group. Equation (35) also corresponds to the maximum likelihood estimate of the mean effect size given by Raudenbush (2009 p 310 eq 16.33). This shows that our more general model, allowing for the possibility of non-zero correlations among sampling errors within groups, is consistent with the traditional RE meta-analysis estimator when considering the special case where the correlations are zero.

To examine the ingredients of the formula that include the sampling error correlations, consider the case with two groups where one of the groups has one observation and the other group has two observations and there are no non-sampling errors within groups, i.e., $\sigma_{\mu,i} = \sigma_{\mu,2} = 0$. In this case, the overall summary estimate is

$$\hat{y} = hy_1 + (1-h)\hat{y}_2 = hy_1 + (1-h)\left[gy_{2,1} + (1-g)y_{2,2}\right], \tag{36}$$

where $\hat{y}_2 = gy_{2,1} + (1-g)y_{2,2}$ is the composite estimate for group 2. With only two groups and three observations, there are only two weights to determine: h and g. We will begin by determining g. The variance of the composite estimate for group 2 is

$$v_2 = g^2 \operatorname{var}\left[y_{2,1}\right] + (1-g)^2 \operatorname{var}\left[y_{2,2}\right] + 2g(1-g)\operatorname{cov}\left[y_{2,1}, y_{2,2}\right]. \tag{37}$$

The first-order condition for an optimum is

$$\frac{\partial v_2}{\partial g} = 2g \operatorname{var}\left[y_{2,1}\right] - 2(1-g)\operatorname{var}\left[y_{2,2}\right] + (2-4g)\operatorname{cov}\left[y_{2,1}, y_{2,2}\right] = 0. \tag{38}$$

substituting ρ and the se's for the covariance and variances gives

$$2gse_{2,1}^2 - 2(1-g)se_{2,2}^2 + (2-4g)\rho se_{2,1}se_{2,2} = 0, \tag{39}$$

which can be solved for g:

$$g = \frac{se_{2,2}^2 - \rho se_{2,1} se_{2,2}}{se_{2,1}^2 + se_{2,2}^2 - 2\rho se_{2,1} se_{2,2}}. \tag{40}$$

As a numerical cross-check, the following R script confirms that the g's computed using equation (40) exactly match those computed using equation (29) in the simple case of a group with two observations and no within-group (observation-level) non-sampling errors.

```
y       <- matrix(c(8,0,10,7),2,2,byrow=1)
se      <- matrix(c(2,0,3,2),2,2,byrow=1)
rho     <- matrix(c(0,0),2,1)
sig.mu  <- matrix(0,2,1)
sig.eta <- 1.5

I <- length(rho)
J <- matrix(0,I,1)

for(i in 1:I){J[i] <- sum(y[i,]!=0)}

# Calculate optimal weights using matrix formula:
g <- 0 * y
A <- matrix(0,i,max(J))
B <- array(0,dim=c(max(J),max(J),I))

for(i in 1:I){

  for(j in 1:J[i]){

    for(k in 1:J[i]){

      A[i,j] <- A[i,j] + (sig.mu[i]^2+(1-rho[i])*se[i,j]^2)/
                         (sig.mu[i]^2+(1-rho[i])*se[i,k]^2)
      num <- 0
      den <- 0
      for(kk in 1:J[i]){

        num <- num + se[i,kk]/(sig.mu[i]^2+(1-rho[i])*se[i,kk]^2)
        den <- den + (sig.mu[i]^2+(1-rho[i])*se[i,j]^2)/
                     (sig.mu[i]^2+(1-rho[i])*se[i,kk]^2)
      }
      B[j,k,i] <- rho[i]*(num/den-se[i,j]/(sig.mu[i]^2+(1-rho[i])*se[i,j]^2))*se[i,k]

    }
    A[i,j] <- 1/A[i,j]
  }
  g[i,1:J[i]] = solve(diag(J[i])-B[1:J[i],1:J[i],i]) \%*\% as.matrix(A[i,1:J[i]])
}

g.21.check <- (se[2,2]^2 - rho[2]*se[2,1]*se[2,2])/
              (se[2,1]^2 + se[2,2]^2 - 2*rho[2]*se[2,1]*se[2,2])
```

```
print(g[2,1])
print(g.21.check)
```

4.5 Stage two

Next we want to find the infeasible optimal weights, h_i, to place on the composite group-level estimates, $\hat{y}_i$. We will choose the h_i's to minimize the variance of the overall estimate subject to the constraint that the weights sum to one. The variance of the overall estimate is:

$$\mathrm{var}\left[\hat{y}\right] = \sum\nolimits_{i=1}^{I} h_i^2 v_i, \tag{41}$$

where v_i is the variance of the composite estimate for group i and can be computed using equation (11) after computing (feasible versions of) the g_{ij}'s using equation (33) from stage one and feasible estimates of the unknown quantities therein. The first-order condition for a minimum is:

$$\frac{\partial \, \mathrm{var}\left[\hat{y}\right]}{\partial h_i} = 2h_i v_i - \lambda = 0, \tag{42}$$

where λ is the Lagrange multiplier on the constraint $\sum_{i=1}^{I} h_i = 1$. Solving for h_i gives

$$h_i = \frac{\lambda}{2v_i}. \tag{43}$$

Next, we can use the constraint to write

$$\sum\nolimits_{i=1}^{I} h_i = \sum\nolimits_{i=1}^{I} \frac{\lambda}{2v_i} = 1, \tag{44}$$

then solve for the Lagrange multiplier to get

$$\lambda = \frac{1}{\sum_{i=1}^{I} \frac{1}{2v_i}}, \tag{45}$$

then plug this result back into the expression for the group-level weight, equation (43), to get

$$h_i = \frac{\frac{1}{\sum_{i=1}^{I} \frac{1}{2v_i}}}{2v_i}, \tag{46}$$

which can be simplified by cancelling the 2 and written slightly more compactly as

$$h_i = \frac{v_i^{-1}}{\sum_{i=1}^{I} v_i^{-1}}. \tag{47}$$

This formula for the second-stage group-level weights is directly analogous to that for the within-group observations derived in the first stage—i.e., weights are proportional to the inverse variances of the estimates, where the variances account for contributions from all relevant error components—and so also is consistent with the weights given by Hedges and Olkin for the traditional RE meta analysis estimator (1985 p 199).

4.6 Estimation of error variances

To construct a feasible version of the two-stage random effects estimator derived above, we require estimates of the unknown error variances. This section derives estimators for the two unknown error variance components: the within-group non-sampling errors ($\sigma^2_{\mu,i}$) and between-group non-sampling errors (σ^2_η). At this point we will generalize to a meta-regression framework, so now we assume

$$y_{ij} = x_{ij}\beta + \eta_i + \mu_{ij} + \varepsilon_{ij}. \tag{48}$$

We will begin by ignoring the correlations among sampling errors within groups, then we will generalize our result to allow for non-zero sampling error correlations. If the sampling errors are uncorrelated, then observations in group i can be viewed as drawn from a mixture distribution: with frequency $1/J_i$ the mean is $x_{ij}\beta + \eta_i$ and the variance is $\sigma^2_{\mu,i} + se^2_{ij}$. The variance of a mixture is the weighted average of the variances plus the variance of the means, therefore the expected variance of the observations in group i is

$$\mathbb{E}\left[\mathrm{var}\big[\mathbf{y}_i\big]\right] = \frac{1}{J_i}\sum_{j=1}^{J_i}\left(\sigma^2_{\mu,i} + se^2_{ij}\right) + \mathrm{var}\big[\mathbf{x}_i\beta\big], \tag{49}$$

where $\mathbf{y}_i$ denotes the vector of observations in group i, $\begin{bmatrix} y_{i,1} & y_{i,2} & \dots & y_{i,J_i} \end{bmatrix}$. We can rearrange equation (49) and use sample estimates of the unknown variance terms to derive a feasible method-of-moments estimator for the non-sampling error variance for group i,

$$\hat{\sigma}^2_{\mu,i} = \mathrm{var}\big[\mathbf{y}_i\big] - \frac{1}{J_i}\sum_{j=1}^{J_i} se^2_{ij} - \mathrm{var}\big[\mathbf{x}_i\hat{\beta}\big]. \tag{50}$$

This estimator is consistent with the estimator derived by Hedges and Olkin (1985 p 194) for the across-group non-sampling error variance (i.e., heterogeneity of true effect sizes) in a standard RE model with one observation per group and excluding the $\mathbf{x}_i$ vector. This estimator also appears to be identical to the method-of-moments estimator for the variance of the across-group random effects using OLS regression given by Raudenbush (2009 p 311). Here we are applying the same logic to estimate the variance of the within-group heterogeneity or non-sampling errors.

We derived the estimator in equation (50) under the assumption of uncorrelated sampling errors. With correlated sampling errors, the relevant expression is

$$\hat{\sigma}^2_{\mu,i} = \mathrm{var}\big[\mathbf{y}_i\big] - \frac{1}{J_i}\left[\sum_{j=1}^{J_i}\left(se_{ij}^2 - \frac{1}{J_i}\sum_{k\neq j}^{J_i}\rho_i se_{ij} se_{ik}\right)\right] - \mathrm{var}\big[\mathbf{x}_i\hat{\beta}\big]. \tag{51}$$

This is a generalization of equation (50), here accounting for the covariances among observation-level sampling errors within the group.

It is possible for the estimator in equation (51) to return a negative $\hat{\sigma}^2_{\mu,i}$ when the average of the reported standard errors is greater than the total variance of the individual observations from group i, $\mathrm{var}\left[\mathbf{y}_i\right]$. This can occur due to sampling variability alone especially when J is small. Alternatively, this might suggest a positive correlation between the within-study non-sampling errors; however, we have not included such a correlation term in our estimator. In practice we handle this by setting $\hat{\sigma}^2_{\mu,i} = 0$ in such cases.

Finally, we derive a feasible method-of-moments estimator of σ^2_η based on equation (11):

$$\hat{\sigma}^2_\eta = \frac{1}{I-1}\sum_{i=1}^{I}\left(\hat{y}_i - \tfrac{1}{I}\sum_{i=1}^{I}\hat{y}_i\right)^2 - \frac{1}{I-1}\sum_{i=1}^{I}\sum_{j=1}^{J_i}\left[\hat{g}_{ij}^2\left(\hat{\sigma}^2_{\mu,i} + se_{ij}^2\right) + \rho_i \textstyle\sum_{k\neq j}^{J_i} \hat{g}_{ij}\hat{g}_{ik} se_{ij} se_{ik}\right], \tag{52}$$

where the first term on the right-hand side of equation (52) is the sample analog of v_i in equation (11) and the $\hat{g}_{ij}$'s in the second term are computed by plugging in the estimated values for all unknown quantities in equation (34).

Note that we cannot estimate both the group-level non-sampling error variance $\sigma^2_{\mu,i}$ and the within-group sampling error correlation ρ_i simultaneously for all groups because we have only one equation, expression (52), to identify the two unknowns for each group. To address this limitation, we specify a common value for ρ_i for all groups ex ante and then estimate the $\sigma^2_{\mu,i}$'s conditional on the maintained assumption about the ρ_i's.

4.7 Iterative estimation approach

When conducting a meta-regression including one or more explanatory variables, efficient estimation of the coefficient vector $\hat{\beta}$ requires estimates of the observation weights, which in turn requires estimates of the error component variances. However, estimating the error component variances requires an estimate of β, as indicated in the Note at the bottom of Table 2. This chicken-and-egg problem can be solved using an iterative estimation approach: Begin by setting all elements of $\hat{\beta}$ to 0 and calculate estimates of the error component variances and observation weights as described in this Appendix and summarized in Table 2. Use the resulting weights to perform a weighted least squares regression to produce an updated estimate of β. Use the updated $\hat{\beta}$ to recompute the observation weights, again following the sequence of computations in Table 2 but now using the adjustment indicated in the note at the bottom of the table. Repeat this process until the estimates change by an amount smaller than a pre-defined tolerance. The stopping criterion we used in our demonstration was when the largest change among the coefficient estimates became smaller than 0.001%, which is safely below the sampling variability of these estimators.