Meta-Analysis and Publication Bias in the Hedonic Wage Literature

William Wheeler, Economist, U.S. Environmental Protection Agency

Chris Dockins, Economist, U.S. Environmental Protection Agency

Working selected paper being prepared for presentation at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4-6, 2013.

Disclaimer: *The views expressed in this presentation are those of the author[s] and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.*

Meta-Analysis and Publication Bias in the Hedonic Wage Literature

Abstract: The value of statistical life (VSL) is one of the most scrutinized and controversial parameters estimated by environmental economists (Cameron 2009, Viscusi 2010), largely due to the wide use of VSL estimates to value the mortality risk benefits of regulations that affect public health and safety (OMB 2011, Robinson and Hammitt 2010). The hedonic wage method has been a primary source of VSL estimates for use in applied benefit-cost analysis and there have been several meta-analyses of these studies, including examinations of publication bias. We build on the existing literature by focusing on the coefficient on fatal risk rather than the VSL itself. This allows for larger sample sizes and reflects more recent methods that provide a cleaner test for bias. Results suggest that publication bias is present in the full sample of hedonic wage VSL estimates and that correcting for this by using those observations with the most precise estimates results in lower mean VSL estimates.

Key Words: value of statistical life, VSL, mortality risk, benefits analysis

PRELIMINARY DRAFT, DO NOT QUOTE OR CITE WITHOUT PERMISSION

Introduction

The VSL is one of the most scrutinized and controversial parameters estimated by environmental economists (Cameron 2009, Viscusi 2010), largely due to the wide use of VSL estimates to value the mortality risk benefits of regulations that affect public health and safety (OMB 2011, Robinson and Hammitt 2010). In many cases mortality risk benefits account for the largest portion of total quantified benefits of these regulations (e.g., US EPA 2011).

VSL can be estimated through both stated and revealed preference methods, although there is evidence that these two sets of literature provide systematically different results (Kochi, et al. 2006). Both sets of literature have been used for applied policy analysis and sometimes results from both sets of studies have together informed guidance for federal agencies, including the US EPA (US EPA 2010a Guidelines). Revealed preference methods include hedonic property studies (e.g., Gayer, et al.) and defensive behavior studies (e.g., Blomquist, 2004). However, the most prevalent and influential revealed preference studies use the hedonic wage method to estimate workplace premiums for on-the-job risk of fatality.

As noted in Viscusi 2008, the "canonical" hedonic wage equation is given by

$$ln(w_i) = \alpha + X_i'\beta + \gamma_1 p_i + \gamma_2 q_i + \gamma_3 WC_i + \varepsilon_i$$

where w_i = wage

X_i = vector of worker and job characteristics

p_i = fatal job risk

q_i = fatal injury risk, and

WC_i = presence of workers' compensation.

Risk is typically reported in deaths per 10,000 or deaths per 100,000 workers per year and wage is typically measured in hourly terms. From this specification VSL is generally calculated assuming full time hours worked over the course of a year (e.g., 2,000 hours) and scaled for the denominator of the risk variable (e.g., 100,000):

$$VSL = w * \beta * 2000 \text{ hours/year} * 100{,}000.$$

There have been several systematic, quantitative reviews and meta-analyses of the hedonic wage literature VSL estimates. All of these have studied the VSL itself, many of these have focused on how VSL estimates that are sometimes widely divergent are influenced by econometric specification, sampling, and data differences across studies (e.g., Bowland and Beghin 2001, Mrozek and Taylor 2002, Viscusi and Aldy 2003). Some, but not all, of these analyses have estimated a VSL's that could be used in policy analysis. None of these studies, however, have examined the coefficient on fatal risk from the hedonic wage equation, from which the VSL is estimated.

Publication selection bias, in the form of choosing which estimates are reported, has long been recognized as a concern for statistical inferences from published literature (Card and Kreuger 1994), and several studies have noted it as a potential concern for the VSL literature (e.g., Hwang 1992, Day 1999, US EPA 2006). To date, however, only a few meta-analyses have rigorously examined potential bias in published VSL estimates. Day (1999) attempted to test for publication bias, but, as noted in Doucouliagos, et. al 2012, the meta-regression model has a poor fit and relies upon publication bias test that has since been identified as being invalid. Only Doucouliagos et al. (2012) use the most recent methods for identifying publication bias (Stanley 2005, 2008)

In this paper, we build on the existing literature on the VSL in two ways. First, all extant meta-analyses of the VSL examine the VSL itself, while we focus on the coefficient on fatal risk in hedonic wage equations. This allows for a larger sample size because many original studies do not include an estimate of the VSL for all hedonic wage regressions. (Although meta-analysts will calculate the VSL in these cases when possible, sufficient information is not always provided in the original study.) Second, because recent methods for identifying publication bias are based on testing regression coefficients, focusing on the fatal risk coefficient (rather than the VSL) yields a cleaner test for bias.

Methods for Meta-Analysis and Measuring Publication Bias from Hedonic Wage Studies

Our analysis starts with the estimation of fixed and random effect weighted means (Borenstein et al. 2010; Nelson 2013); the former assumes a common, true population effect size, and that differences in estimates are caused by within-study estimation error. The latter assumes a distribution of effects sizes and estimates the mean of those effect sizes. The random effects model will give relatively more weight to less precise estimates because of the addition of the common error component. Because the studies in our sample vary across countries, time, and data, the assumption of a true effect size is probably unrealistic (Borenstein et al. 2010).

Publication bias

Most approaches to publication bias assume a relationship between the distribution of the coefficient of interest and its standard error. The simplest way to examine this relationship is to graph the coefficient estimates versus their inverse standard errors. In the absence of publication bias, if there is a true effect, the graph will resemble an inverted funnel, with the most precise estimates forming the narrow part of the funnel around the true effect and less precise estimates forming the wide part of the funnel. The funnel shape natural arises from sampling error. With publication selection, these funnels will be truncated on one side: generally only positive, statistically significant estimates are retained and published, leading to misleading aggregate estimates.

As a first step, we construct funnel graphs (Stanley and Doucouliagos 2009) of the coefficient on fatal risk against the inverse of its standard error. If there is no publication bias we expect the funnel plot to be symmetric, as each observation is an estimate of a true effect size, or underlying value on the fatal risk coefficient. The results of the funnel graph, while suggestive, are not a fully statistically rigorous test of publication bias.

Following Stanley (2008) and Doucouliagos, et al. (2012) we next estimate regressions to directly test for publication bias and estimate the "true" effect size (fatal risk coefficient) correcting for this bias. The rationale for this approach is that researchers with large sample sizes, and therefore more precise estimates will be able to identify and report smaller estimated empirical effects. By contrast, studies with smaller sample sizes and less precision will need to find large effects for these to be statistically significant. If sufficient numbers of researchers are searching for these effects and reporting them then we would expect standard error to be related to the size of the effect and we have evidence of publication bias.

In the regression equation

$$effect_i = \alpha_1 + \alpha_0 SE_i + \varepsilon_i$$

$effect_i$ is coefficient on fatality risk in a hedonic wage equation, α_1 is an estimate of the true effect, and α_0 is an estimate of publication selection (see Stanley 2008 for details). If there is no publication bias, α_0 will equal zero. If there is publication bias, α_0 will be positive (assuming a positive effect) as researchers add statistically significant estimates with large standard errors to their publications.

Because the above equation suffers from known heteroskedasticity (Stanley 2008), a weighted least squares estimate can be used instead. Dividing through by the standard error gives:

$$t_i = \alpha_0 + \alpha_1 1/SE_i + e_i$$

We build on the analysis of means by incorporating covariates into a meta-regression model (Stanley 2008, Nelson and Kennedy 2009, Nelson 2013). We do not used fixed effect regressions, although they are often recommended in the meta-analysis literature (Nelson and Kennedy 2009; Nelson 2013) because some of the primary variables of interest—those designating the different datasets of job characteristics such as the COI—generally will not vary within a study and so will drop out. Further, the assumption of a common effect size across the disparate studies in our analysis is a strong one and fixed effects regressions are not recommended by all authors (Harbord and Higgins 2008).

Card and Krueger (1995) note that increased sample sizes should result in lower standard errors;[1] they specifically suggest that in a regression with the log of the (absolute) t-ratio as the dependent variable, the coefficient on the log of the square root of the degrees of freedom (used as an independent variable) should have a coefficient equal to 1 . Görg and Strobl (2001) use same test.

Data

Our sample of studies was based on that collected by EPA 2010b[2]. However, where the dataset of VSLs described there is based on one estimate per study, we include all estimates within a study that are specified by ln(wages) on the left-hand side of the equation. We also excluded studies or estimates that were based on aggregate, not individual, level data, or those that used measures of excess risk instead of mean risk. All variables were coded by one research assistant and re-coded by a second; the authors examined and resolved any discrepancies. Our sample consists of 35 studies and 386 coefficient estimates from those studies using ln(wage) as the dependent variable.

** Describe Final Dataset **

Results and Discussion

Our results for fixed and random effect means are shown in table XX1. The means vary by an order of magnitude: the fixed effect mean is 0.00230 (95 percent confidence interval of 0.00223-0.00237) and the random effect mean is 0.0245 (95 percent confidence of 0.0221-0.0269). The variation in means is due to the differential weighting of estimates. We test for

---

[1] Unless there is not a true effect (Stanley 2008).

[2] The sample selection criteria are explained on pp. 41-42.

homogeneity of the estimates using Cochrane's Q-test (Borenstein et al. 2010) and strongly reject it (p-value of 0.000), indicating that the fixed effects model is probably inappropriate.

Tables xx2 and xx3 present meta-regression results for two sets of samples: those studies that use a 1/10,000 risk estimate (because that is the most common approach and avoids issues with rescaling) and all studies in our sample, respectively. Reading across each table are the following two models: weighted OLS with clustered errors and weighted random effects. Inverse variances are used for the weights.

Focusing on the random effects estimates in Table xx2, which are preferred to the OLS estimates that ignore the correlation between estimates taken from the same study, we find that using CFOI data results in a statistically significant larger estimate on the fatal risk coefficient compared to NIOSH, BLS, and all other risk data sources. Both the fatal interaction and fatal quadratic dummies are statistically significant, but with opposite signs: the fatal interaction term leads to smaller fatal risk coefficients and the fatal quadratic term leads to lower coefficients. These are both consistent with our expectations. The dummy variable for male samples and blue collar occupations are also both significant; male samples have smaller values of the fatal risk coefficient while blue collar samples have larger. The dummy variable for inclusion of workers compensation is negative, indicating that specifications including this term have smaller coefficients on fatal risk.

The results in Table xx3 are fairly different, possibly because of the larger dispersion in study factors in the larger data set. None of the variables representing different data sets are significant and only the dummies for the fatal interaction term, the nonfatal dummy, and workers compensation are significant.

Turning to publication bias, Figure 1 is a funnel graph of the fatal risk coefficients with the lowest and highest five percent of the estimates eliminated so the graph is readable. In the absence of publication bias, the graph would be symmetric. However, the graph is clearly skewed to the right, with very few negative observations. Even recognizing that the coefficient on fatal risk should be positive, sampling error should lead to some negative estimates; further, in econometric specifications with interaction terms (including quadratic terms), we would fully expect some estimated coefficients on fatal risk to be negative. The skewness of the graph is suggestive of publication bias.

To calculate publication bias, we estimate our second equation above using a random effects model. The constant term, or estimated true coefficient on fatal risk, is 0.00120 (standard error=0.003, p-value < 0.0001). The coefficient on the inverse standard error is 4.49275 (standard error=2.6180, p-value = 0.08615) which Stanley notes as indicative of "severe" publication bias.

Using the Card and Krueger test, $\ln(tstat) = \alpha + \beta\ln(\sqrt{SE})$, that is regressing the natural log of the t-statistic on the natural log of the square root of the standard error,[3] we find that the coefficient β is equal to 0.31 (p-value=0.04) in a regression with fixed effects. This indicates a true effect (Stanley 2008) but the possibility of some publication bias because the coefficient is not equal to one.

The observations in our sample that report a VSL in USD have mean VSL of $8.3 million ($2000). We explore the influence of precision of the VSL by following Stanley et al. (2010) accounting for publication bias by using only the observations with the ten percent most precise estimates of the fatal risk coefficient.[4] Doing so lowers the unweighted mean to $6.3 million. These numbers are fairly similar to the VSL used in recent policy analyses in the US (Robinson and Hammitt 2011). These estimates come from a wide array of studies from different time periods. EPA, for example, uses a central estimate (based on a Weibull distribution) of the studies in their sample, none of which were published after 1991, but the EPA estimate is

[3] Card and Kruger (1994) use the degrees of freedom instead of the sample size; degrees of freedom are rarely coded but we have much better information on sample size, which will be highly correlated with the degrees of freedom.

[4] Note that we drop the observations that have standard errors that are not robust or clustered because these overestimate precision.

otherwise similar to ours although it is smaller than the most precise estimates in our sample after accounting for inflation.

Conclusions

We find evidence of publication bias in the hedonic wage-risk literature, however, this bias appears to have only a moderate effect on the estimated VSLs. Because VSLs used for policy come from a variety of sources and benefit-transfer considerations, it is difficult to compare them directly with our results. Still, the results here suggest that both the adjusted and unadjusted VSLs are generally consistent with those commonly used in policy analysis. Because the most recent hedonic wage VSLs in the US are based on risk data from CFOI, one issue to examine more fully is the relationship between data source and publication bias.

NOTES

(In general the inclusion of unpublished studies reduces the problem of publication bias (Sterne et al., 2001).) (From http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0494-01.pdf/$file/EE-0494-01.pdf Report of the EPA Work Group on VSL Meta-Analyses)

Distinction between publication bias and "reporting bias" (the exclusion or failure to report models or subpopulation results that did not reach significance or did not conform to expectations from previously published literature)

References

Ashenfelter, O and M Greenstone. 2004. Estimating the value of statistical life: the importance of omitted variables and publication bias. American *Economic Review Papers and Proceedings* 94:454-460.

Begg, CB and JA Berlin. 1988. Publication bias: a problem in interpreting medical data. *Journal of the Royal Statistical Society, Series A*. 161:95-105.

Bellavance, F, G Dionne, and M Lebeau. 2008. The value of a statistical life: a meta-analysis with a mixed effects regression model. *Journal of Health Economics* 28:444-464.

Blomquist, G. 2004. Self-protection and averting behavior, values of statistical lives, and benefit-cost analysis of environmental policy. *Review of Economics and the Household* 2(1):89-110.

Borenstein M, LV Hedges, JPT Higgins, and HR Rothstein. 2010. A basic introduction to fixed effect and random effect models for meta-analysis. *Research Synthesis Methods* 1:97-111.

Bowland BJ and JC Beghin. 2001. Robust estimates of the value of a statistical life for developing economicies. *Journal of Policy Modeling* 23:385-396.

Cameron, TA. 2010. Euthanizing the value of statistical life. *Review of Environmental Economics and Policy* 4(2):161-178.

Day B. 1999. A meta-analysis of wage-risk estimates of the value of statistical life. Centre for Social and Economic Research on the Global Environment, University College London, mimeo.

Doucouliagos, C, TD Stanley, and M Giles. 2012. Are estimates of the value of statistical life exaggerated? *Journal of Health Economics* 31:197-2006.

Gayer, T, JT Hamilton, and WK Viscusi. 2000. Private values of risk tradeoffs at superfund sites: housing market evidence on learning about risk. *Review of Economics and Statistics* 82(3):439-451.

Görg, H and E Strobl. 2002. Multinational companies and indigenous development: an empirical analysis, *European Economic Review* 46, 1305-1322.

Harbord, RM and JPT Higgins. 2008. Meta-regression in Stata. *Stata Journal* 8(4):493-519.

Hwang H-S, WR Reed, and C Hubbard. 1992. Compensating wage differentials and unobserved productivity. *The Journal of Political Economy* 100(4):835-858.

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

Mrozek JR and LO Taylor. 2002. What determines the value of life? A meta-analysis. *Journal of Policy Analysis and Management*. 21(2):253-270.

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

Robinson LA and JK Hammitt. 2011. Valuing health and longevity in regulatory analysis: current Issues and challenges in D Levi-Faur, ed. *Handbook on the Politics of Regulation* (Cheltenham and Northampton: Edward Elgar.)

Stanley, TD (2001). Wheat from chaff: meta-analysis as quantitative literature review. *Journal of Economic Perspectives* 15:131-150.

Stanley, TD. 2005. Beyond publication bias. *Journal of Economic Surveys* 19:309-345.

Stanley TD. 2008. Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics* 103-127.

TD Stanley and Doucouliagos, C. 2010. "Picture This: A Simple Graph That Reveals Much Ado About Research." *Journal of Economics Surveys* 24 (1): 170–191.

Doucouliagos, Chris and Stanley, T. D.* (2009) Publication selection bias in minimum-wage research? A meta-regression analysis, *British journal of industrial relations*, vol. 47, no. 2, pp. 406-428Stanley, TD, SB Jarrell, and C Doucouliagos. 2010. Could it be better to discard 90% of the data? A statistical paradox. *The American Statistician* 64(1):70-77.

Sterne, JAC, M Egger, and GD Smith. 2001. Investigating and dealing with publication and other biases in meta-analysis. *BioMedical Journal* 323:101.

US EPA 2006. *Report of the EPA Workgroup on VSL Meta-Analyses*. http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0494-01.pdf/$file/EE-0494-01.pdf (accessed on June 3, 2013)

US EPA 2010a. *Guidelines for Preparing Economic Analyses*. EPA 240-R-10-001.

US EPA. 2010b. Valuing Mortality Risk Reductions for Environmental Policy: A White Paper. http://yosemite.epa.gov/ee/epa/eerm.nsf/vwan/ee-0563-1.pdf/$file/ee-0563-1.pdf

US EPA 2011. *The Benefits and Costs of the Clean Air Act from 1990 to 2020*. http://www.epa.gov/cleanairactbenefits/prospective2.html (accessed June 3, 2013)

US Office of Management and Budget (OMB) 2011. *2011 Report to Congress on the Benefits and Costs of Federal Regulations and Unfunded Mandates on State, Local and Tribal Entities.*

Viscusi, WK. 2008. The value of life in SN Durlauf and E Blume, eds., *The New Palgrave Dictionary of Economics, 2nd Ed*ition (London, UK: Macmillan).

Viscusi, WK. 2010. The heterogeneity of the value of statistical life: introduction and overview. *Journal of Risk and Uncertainty* 40(1):1-13.

Viscusi WK and JE Aldy. 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.

Table X1. Fixed and Random Effect Means and Confidence Interval for Coefficient on Fatal Risk

	Mean	95 percent Confidence Interval	
Fixed Effect	0.00230	0.00223	0.00237
Random Effect	0.0245	0.0221	0.0269

Table xx2. Meta-regressions for studies using a 1/10,000 risk estimate (n=214, i=13).

	Clustered b/se/p	RE b/se/p
frse_use	-0.07688	-0.14699
	(0.3560)	(0.1743)
	0.83264	0.39899
CFOI	0.00257	0.00201
	(0.0007)	(0.0005)
	0.00193**	0.00012***
NIOSH	0.00154	0.00119
	(0.0003)	(0.0008)
	0.00014***	0.13806
BLS	0.00074	0.00070
	(0.0004)	(0.0004)
	0.09326	0.07719+
fatal interaction	-0.00152	-0.00104
	(0.0003)	(0.0003)
	0.00020***	0.00009***
fatal quad	0.12131	0.12704
	(0.0287)	(0.0710)+
	0.00117**	0.07348
nonfataldummy	-0.00000	-0.00001
	(0.0001)	(0.0002)
	0.99058	0.95215
samplegender=..0000	-0.00122	-0.00076
	(0.0002)	(0.0002)
	0.00017***	0.00037***
sampleoccupat..0000	0.00117	0.00119
	(0.0002)	(0.0004)
	0.00041***	0.00740**
workerscomp	-0.00146	-0.00133
	(0.0005)	(0.0004)
	0.01169*	0.00240**
Constant	0.00170	0.00172
	(0.0001)	(0.0002)
	0.00000***	0.00000***

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table xx3. Meta-regressions, all studies (n=386, i=35).

	Weighted&C~S b/se/p	RE b/se/p
frse_use_adj	0.99233	0.57068
	(0.4385)	(0.2339)
	0.03014*	0.01470*
CFOI	0.02657	0.02810
	(0.0086)	(3.2110)
	0.00413**	0.99302
NIOSH	0.00254	0.00349
	(0.0077)	(0.2198)
	0.74458	0.98733
BLS	0.01072	0.00078
	(0.0031)	(0.2203)
	0.00147**	0.99717
fatal interaction	0.01213	0.01532
	(0.0083)	(0.0009)
	0.15074	0.00000***
fatal quad	0.00010	0.00012
	(0.0000)	(0.0001)
	0.00000***	0.37369
nonfataldummy	0.01015	0.00609
	(0.0021)	(0.0024)
	0.00003***	0.01072*
samplegender=..0000	-0.01471	-0.00196
	(0.0044)	(0.0040)
	0.00212**	0.62591
sampleoccupat..0000	0.01106	0.01064
	(0.0043)	(0.0082)
	0.01397*	0.19385
workerscomp	-0.00436	-0.01624+
	(0.0075)	(0.0087)
	0.56697	0.06158
Constant	0.00008	0.00009
	(0.0000)	(0.0433)
	0.00000***	0.99830

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Figure 1.Funnel graph, fatal risk coefficients, 90 percent most precise estimates