Anchoring effect of VSL estimates on previous studies

Victor Champonnois

Aix-Marseille Univ. (Aix-Marseille School of Economics), CNRS and EHESS

Abstract

In this paper, I propose a test for a publication bias suggested by Viscusi and Masterman [2017]. This bias arises from the selection of estimates based on their distance from the usual range in the literature. It implies that, in a given research area, there is a progressive concentration of estimates around this range as the literature grows. I test this hypothesis using OECD meta-data on VSL elicited via stated preferences. Using interquantile regression, I am able to model semi-parametrically the dispersion of estimates as a function of the number of prior studies, controlling for confounding factors. Consistently with the hypothesis, my results show a decrease in the dispersion for the lower part of the conditional distribution. I find no significant pattern for the upper part of the conditional distribution. This result has two implications. First, there is a systematic underestimation of the variance of mean estimates in meta-analyses of VSL. Second, there is an upward bias in the mean estimate of VSL meta-analyses due to the asymmetry of the selection pattern. Both implications might also affect meta-analyses in other research areas.

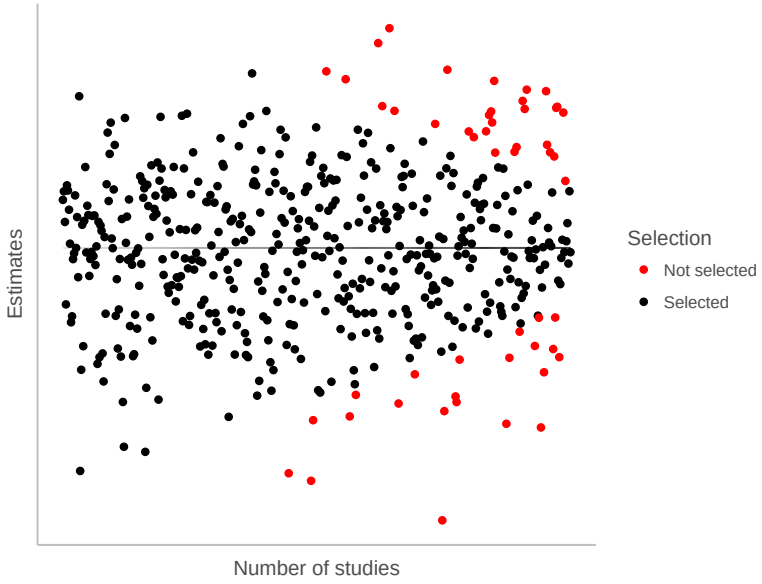
1 Introduction

Meta-analyses have been increasingly used in various areas during the past decades. This research technique allows to average key estimates from different studies to obtain a unique and more precise measure. It also enables identifying determinants of the estimate heterogeneity between studies, which can be used to provide methodological guidelines. A large share of meta-analyses have been performed on the value of a statistical life (VSL) (Ashenfelter and Greenstone [2004], Bellavance et al. [2009], Lindhjem et al. [2011], Robinson and Hammitt [2015]). One of the main issue in these studies is the selection of estimates, both by the researcher and the journals. If not accounted for, this selection creates a publication bias, which can lead to a misrepresentation of the aggregate estimate and inappropriate recommendations (Stanley et al. [2017]).

To my knowledge, three main types of publication bias have been identified so far. There is a large literature on the publication selection based on the significance of the results. This publication bias has been acknowledged for a long time and there have been many attempts to correct for it, including in VSL meta-analyses (Doucouliagos et al. [2012, 2014], Viscusi [2015]). A second type of publication bias comes from the preference towards estimates that are consistent with theory (see for example Card and Krueger [1995] on the effect of minimum wage). This bias is also present in the literature on the value of a statistical life (VSL): there is a preference for positive VSL (Doucouliagos et al. [2012]) positive income elasticities (Doucouliagos et al. [2014]).

In two recent VSL meta-analyses, Viscusi [2015], Viscusi and Masterman [2017] suggest a third type of publication bias. This bias arises "because of the efforts by researchers to provide estimates

Figure 1: Estimates w.r.t. quantity of studies in the literature



in line with the previous literature". Besides researchers, "journal editors and reviewers likewise may be more likely to favor publication of result in the usual range". This hypothesis is based on their findings of a large difference in magnitude of publication bias between US and non-US studies. This bias would be due to non-US researchers anchoring their results on US studies. However, these authors do not provide a way to test for this effect, so it has not been formally identified.

In this paper I propose a test for a broader form of "publication anchoring bias". Instead of relying on the distinction between US v.s. non-US studies, I base my approach on the size of the literature at the time of publication. My hypothesis is the following. If researchers and reviewers actually judge the quality of an estimate based on its distance from the central tendency in the literature, they would prefer an estimate close to the central range to one far from it. So if a literature is well developed, there exists a range of values seen as acceptable and the selected estimates will tend to lie in this predetermined range. This selection pattern will then lead to a concentration of estimates around the central tendency in the literature. However, when a literature emerges there is no consensus on what the magnitude of the estimates should be, so it should not be possible to discriminate estimates based on this pattern. It is only when the body of works grows that a central tendency appears and that selection can happen. Thus, when the number of studies increases there should be a decrease in the dispersion of estimates. Figure 1 illustrates this pattern using simulated data. Estimates are represented in the y-axis and the number of prior studies in the x-axis. Black dots represent selected estimates, and red dots rejected estimates. When the literature is scarce, there is no selection because of the uncertainty on what the appropriate range should be, but as the size of the literature increases, a central tendency emerges (the black line in the graph) and estimates start to be selected based on their distance from this central tendency. This selection becomes more strict as the central trend is more apparent.

Under this hypothesis, if the selection is symmetric, the concentration of estimates towards the central tendency would not be strictly speaking a bias. The selection based on the distance from this central tendency would lead to an underestimation of the variance of the mean estimates. This would only lead to an error regarding accuracy, not regarding the mean estimate. The selection could be asymmetric though, i.e. there could be a selection more strict on large estimates than on small estimates. Such a pattern would lead to a bias in the mean estimate, in addition to an underestimation of the variance.

In this paper, I propose a semi-parametric test for the anchoring of estimates on prior studies. Using a meta analysis on the VSL, I show that this publication selection process exists and that it affects the mean and the bias of the VSL estimates. Instead of using meta-data from revealed preference studies as in Viscusi and Masterman [2017], I use OECD meta-data (OECD [2012]) on stated preference studies. The stated preference literature on VSL is very different from the revealed preference literature. The dichotomy between US and non-US studies might be less pronounced. I try to provide a better understanding of this difference by studying the influence of past prior U.S. studies, and whether the anchoring bias affects differently US and non US studies.

The remainder of the paper proceeds as follows. Section 2 presents the design of the test, Section 3 describes the meta-data on the VSL, Section 4 presents and discusses the results and Section 5 concludes.

2 Methodology

In this section, I present a methodology to detect the anchoring of estimates on prior studies. The methodology is based on quantile regressions. Quantile regressions are mainly used to model the heteroskedasticity by looking at how the marginal effect of a variable varies along the conditional distribution of the dependent variable.[1] If the marginal effects are the same along the conditional distribution, then the variance does not vary with the covariate. If the marginal effect increases, then variance increases with the covariate, and reciprocally, if the marginal effect decreases along the conditional distribution, the variance decreases with the covariate (Koenker [2005]). Therefore, I can check how the number of prior studies affects the variance of the estimates by looking at the coefficients associated with this covariate for a set of quantiles of the VSL conditional distribution. It allows to control for the confounding factors that could affect the variance of the VSL. The model is written

$$Q_\tau(y_i|x_i, Nstudies_i) = \alpha_\tau + \gamma_\tau Nstudies_i + x_i'\beta_\tau + \epsilon_{\tau,i}$$

Where y_i is the estimate of study i, x_i is a set of controls associated with study i, and $Nstudies_i$ is the number of prior studies. Then, if the γ_τ decreases with the quantiles τ, it verifies the hypothesis of a concentration of estimates as the literature grows. Note however that this is not a formal test.

To test my hypothesis formally, I use interquantile range regression. I choose to measure the dispersion based on the interquartile range (the test works for any pair of quantile). Quantile regression on the 75th and 25th conditional percentiles allow to model how the conditional quartiles varies with time, while controlling for other covariates. One can then use the conditional interquartile difference $Q_{75}(y_i|x_i, Nstudies_i) - Q_{25}(y_i|x_i, Nstudies_i)$ as a measure of the conditional dispersion. Thus the regression model is the following:

[1] Quantile regressions are also robust to outliers, which is an important feature when working with VSL data, which can sometimes be extremely large.

$$Q_{75}(y_i|x_i, Nstudies_i) - Q_{25}(y_i|x_i, Nstudies_i) = \alpha + \gamma Nstudies_i + x_i'\beta + \epsilon_i$$

The parameter of interest is γ. The model above is estimated using the interquantile range regressions.[2] The estimated variance-covariance matrix of the estimators is estimated via bootstrap. If the coefficient $\hat{\gamma}$ is negative and significant, it shows that estimates become more concentrated as the literature grows, controlling for the set of observables x.

In order to choose the right set of controls, one should know which factors could have an effect on the concentration of published estimates. Two important factors should be relevant. The first one is convergence in methods. As the literature grows, a consensus often emerges on which methods are the most appropriate. Therefore there might be an homogenization of the methods, which could lead to a concentration of published estimates. If a study uses a method that is considered inappropriate by the majority, it should be rightly disregarded. Therefore, one should control for the set of methods, which includes data collection methods and statistical methods. These variables need to be reported in the meta-analysis if one wants to test for this type of publication selection.

The second factor is the increase in sample sizes through time. It is well known that larger datasets have become available in the past years. An increase in sample sizes implies an increase in the precision of the estimates. The estimates of the most recent studies should then be more concentrated.

These two factors should be the main causes of concern for the validity of the test. However, some other aspect specific to the field of study and the nature of the estimates may have an impact as well. Regarding the application on the VSL, I will discuss which other controls should be included in the next section.

3 Meta-data on VSL

I use a meta-data on stated preference surveys for mortality risk valuation. The dataset was collected by the OECD (OECD [2012]). All the surveys elicit VSL from the stated WTP for a reduction in mortality risk related to environmental, health and transport policies.[3] The first study published is dated from 1973 and the most recent ones are from 2009. Figure 2 represent the VSL estimates as a function of the number of prior studies.[4] One can indeed see that there is a convergence in the estimates. The first forty estimates have a very large variance. Then there is a sharp decrease in the variance, followed by another smaller decrease after eighty studies. As explained before, many factors may explain this evolution, the most important ones are likely to be the change in methodological guidelines and the increase in sample size. The estimates published in journals are represented in red and other studies in black. Interestingly, the variance seems to decrease faster for published studies than for unpublished ones. This could be a sign that the homogenization of the methods is faster for published studies than for unpublished ones. Another possibility, is that, there could be a selection by the journals of studies based on the distance from the central tendency.

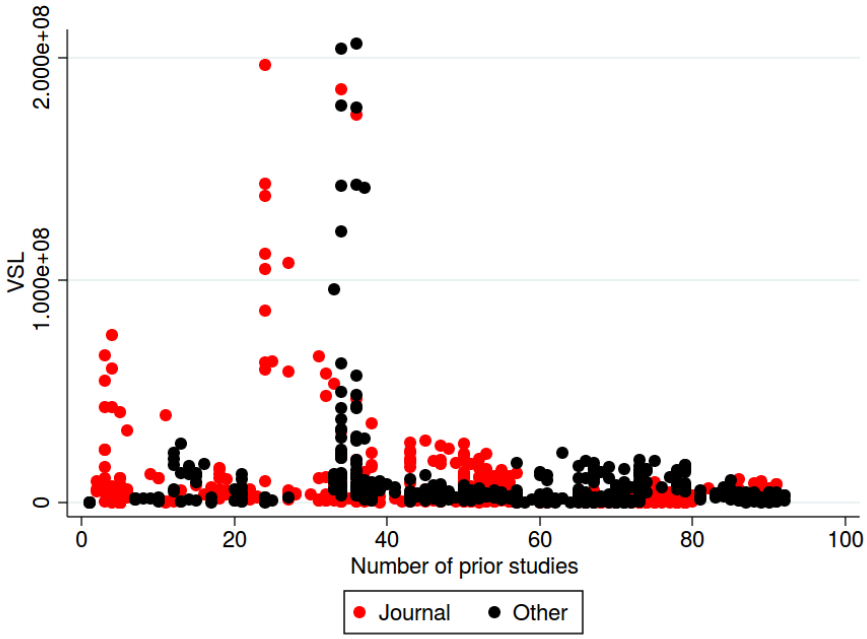
For each study, the OECD meta-data contain information about the methodology, some characteristics of the affected population and the characteristics of the risk (Lindhjem et al. [2011]).

[2] I use the program iqreg from the Stata software.

[3] VSL are converted in USD 2005 PPP adjusted based on the AIC (actual individual consumption).

[4] This variable is created using the year of Publication. Among study published in the same year I could not identify which one came first.

Figure 2: VSL estimates and quantity of studies in the literature



Variables characterizing the survey design and the statistical method are summarized in Table 1. Four variables characterize the survey design. The first one is the elicitation method, i.e. by which mechanism was the WTP elicited. The majority of studies are contingent valuation studies, using dichotomous choice, as it was prescribed by the NOAA panel. Open ended questions and payment cards are also widely used. Conjoint analyses, which also seem to include discrete choice experiments, represent a minor part of the studies. The second aspect is the survey approach. Most surveys are conducted face-to-face or are self-administrated (this include mails). Web based surveys also represent a significant part of the studies. Regarding the payment vehicle, most studies use the price of a product. Other frequently used options are cost of living and a tax. The last aspect is whether the risk change was explained in a way that makes it more intuitive to the respondents, as we see most surveys did.

Regarding the statistical method, I use a classification that depend on the flexibility of the specification. I distinguish between fully parametric models (mostly models estimated by maximum likelihood), semi parametric models (which do not need distributional assumptions for consistent estimations, mostly OLS) and non-parametric (descriptive statistics and turnbull model). Within these categories, the choice of models is mostly determined by the nature of the dependent variables (continuous, interval, binary, etc.), which depends on the elicitation methods. Therefore, more precision on the methods would be redundant with the elicitation method information.

Other relevant variables are summarized in Table 2. The average VSL is 8.7 million, the 1009 estimates range from 5 thousands to 200 millions.

Some variables characterize the magnitude and nature of the risk change. As noted by Lindhjem

Table 1: Summary of methods, N=1167

Elicitation Methods	CV - cards	CV - dicho	CV - open	Conjoint analysis	Other	Total
Freq	145	448	229	118	226	1166
Survey Approach	Face-to-face	Self-administrated	Telephone	Web-based	Other	Total
Freq	482	403	65	148	54	1152
Payment Vehicle	Cost of living	Donation	Price of product	Tax	Other	Total
Freq	146	27	730	171	75	1149
Risk Explanation	Yes	No				Total
Freq	935	186				1121
Statistical Method	Parametric	Semi-parametric	Non-Parametric			Total
Freq	815	37	256			1,108

Table 2: Summary statistics, N=1167

Variable	**Description**	**Mean**	**Std. Dev.**	**N**
vslaic	Sample mean VSL in PPP-adjusted USD 2005	8728300	21686387	1009
riskchange	Change in mortality risk on an annual basis per 1000	0.003	0.016	757
cancerrisk	=1 if reference to cancer risk in survey; 0 if not	0.147	0.354	1167
public	=1 if public good; 0 if private	0.274	0.446	1167
envir	=1 if environment-related risk change; 0 if not	0.264	0.441	1167
health	=1 if health-related risk change; 0 if not	0.422	0.494	1167
traffic	=1 if traffic-related risk change; 0 if not	0.314	0.465	1167
samplesize	Number of valid response used to estimate VSL	846.1	2099.2	848
gdpcapita	Country GDP/capita in PPP-adjusted USD 2005	26616	12215	1155
gdpgrowth	Country growth in GDP at the collection year	3.434	3.032	1165
lifeexp	Country life expectancy at the collection year	76.542	3.514	1155
num_stud	Number of studies in the literature at year of publication	49.458	24.284	1167

et al. [2011], the magnitude of the change is a significant determinant of the VSL. Besides, there might be an evolution of the risk change through time, because surveys conducted in more developed countries tend to value smaller risk changes. There is a reasonable balance between studies on environment, health and transportation policies, with a majority of a health policies. Whether there is a cancer risk may also be important, since it could be perceived as more dreadful.

Regarding the sample size, the meta data report both the total sample size and the subsample size (of the sample that was actually used to compute the VSL). I choose to include the subsample size, because it is this variable that affects the precision of the estimates. As expected, it is positively correlated with the year of data collection (corr = 0.26, p-val < 0.0001).

Some other factors might influence how VSL estimates evolve as the literature grows. In their meta-analysis on VSL from hedonic wage method, Bellavance et al. [2009] observe an increase of VSL through time. They advance several hypotheses to explain this phenomenon. While the first two are related to the hedonic wage methodology, the third is more general: an increase in the life expectancy through the year could lead to a higher VSL. Thus I control for life expectancy in the country at the collection year. Another plausible hypothesis that they do not mention is the income growth since 1970, which could also affect the VSL. Besides, if the estimates are collected in different countries, a convergence in GDP between countries may also explain a convergence in VSL. Therefore I control for the growth rate of GDP per capita of the country during the collection year of the survey and the level of GDP in 2005.

4 Results

As explained in Section 2, I first use standard quantile regressions to observe how the effect of the number of prior studies varies along the conditional distribution. If the coefficient of a covariate in a quantile regression decreases along the conditional distribution, it shows that the conditional variance is decreasing with the covariate.

I compare four model specifications. The first one is a bivariate regression of log VSL on the log of number of prior studies.[5] In the second I control for the method and the sample size. To this specification, I add the other variables that could explain a concentration in the VSL estimates that I presented in the section above (GDP in level and growth, life expectancy). In the fourth specification I add the characteristics that are found to have an impact in Lindhjem et al. [2011] (type of policy, magnitude of risk change, risk of cancer, public nature of the good).

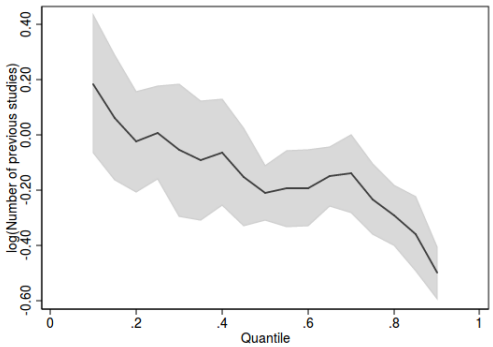
For these four specifications, I report graphs of the coefficient of log of number of prior studies as function of conditional VSL quantiles in the Figure 3.

All curves tend to decrease with VSL conditional quantiles, showing a decrease in the dispersion as the number of studies increases. The decrease is almost monotonic in the bivariate case, which shows a convergence in the VSL estimates, all factors confounded. When accounting for methods, there is a decrease of the effect in the first half of the conditional distribution, but it remains constant in the second half. This means that the reduction in variance for the upper quantiles is seemingly due to methodological guidelines. When adding GDP and life expectancy, a sharp increase in the coefficient for the conditional quantiles above 0.8 appear. With the full specification, this increase of the upper quantiles is much less important. Coefficients are still decreasing along the conditional distribution, up the the 0.8 quantile. This result suggests, that controlling for confounding factors,

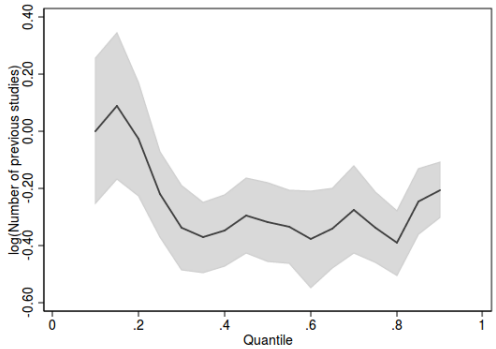
[5] Following Lindhjem et al. [2011] I chose a log-log model: all continuous variables are transformed by taking their log.

Figure 3: Quantile regressions, graphical results

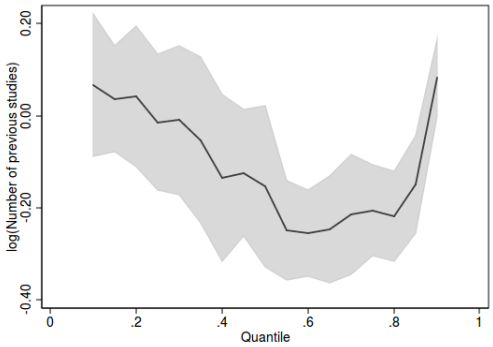
Bivariate



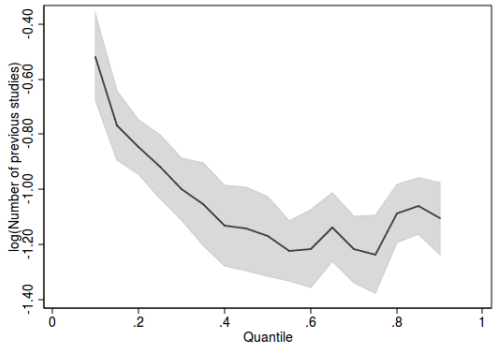
Adding methods and sample size



Adding GDP and Life expectancy



Adding other survey characteristics



there is a decrease in the major lower part of the distribution of the dispersion of VSL estimates as the number of studies increases. On the extreme upper part, there might be an increase instead. This means that the selection pattern might be asymmetric. Therefore, besides performing the test on the interquartile range, I will also test the range between the 0.1 and the 0.5 quantiles, and between the 0.5 and the 0.9 quantiles to detect an asymmetric selection.

Table 2 shows the results from interquantile range regressions for quantile ranges (0.25;0.75), (0.1;0.5), (0.5;0.9). In the (0.25;0.75) range regression, the coefficient of the log of quantity of past studies is negative but only weakly statistically significant. However it is negative and highly significant for the lower quantiles range (0.1;0.5). This confirms the previous results that the decrease in the dispersion only affects the lower part of the conditional distribution. The coefficient is positive but not significant for the (0.5;0.9) range, indicating that the effect on the (0.25;0.75) is only driven by the lower part of the conditional distribution.

The same analysis is performed in Table 3, with the difference that the size of the literature is based only on US studies, i.e. the variable of interest is the number of prior US studies. The effect on interquartile difference is still not significant, it is even smaller in magnitude. However, the effect on the (0.1;0.5) range is significant and much larger than for the first model (1.05 vs 0.65). This result is in line with those of Viscusi and Masterman [2017], of an anchoring effect on US estimates. The effect of the upper part of the conditional distribution is still non-significant, but again it is larger in size.

Finally, I perform an analysis of Viscusi and Masterman [2017]'s hypothesis that non-US studies are affected by anchoring, while US studies are not. As shown in Table 4, the baseline effect of anchoring of international studies on US studies is -0.96 in the lowest quantiles, with no significant additional or canceling effect for US studies. Therefore, US studies and international studies are affected with the same intensity by anchoring bias in stated preference studies.

Note that the distinction between US and non US studies might only be a proxy for another characteristic of the study, such as its quality. Therefore we should be cautious when interpreting these results.

5 Conclusion

In this paper, I test the hypothesis of a selection of estimates based on their distance from the usual range in the literature. I show that this selection happens for estimates in the lower part of the conditional distribution. No statistically significant effect is found for the upper part of the conditional distribution. These results have two implications. First, there is a systematic underestimation of the variance of mean estimates in meta-analyses of VSL. Second, there is a upward bias in the mean estimate of VSL meta-analyses due to the asymetry of the selection pattern. Both implications might also affect meta-analyses in other research areas.

Consistently with Viscusi and Masterman [2017], I find that the anchoring effect is larger when only US studies are considered as anchors. However I do not find that non-US studies are more anchored than US studies. Caution should be used when interpreting these last results, because they might non-causal.

Future work will be devoted to accounting for this bias in the computation of the estimates mean and variance.

Table 3: Results of the interquantile range test

VARIABLES	(1) IQRR 25-75	(2) IQRR 10-50	(3) IQRR 50-90
log(gdp)	-0.137	-0.0898	-0.0355
	(0.246)	(0.277)	(0.278)
log(riskchange)	-0.182***	-0.192***	-0.138*
	(0.0543)	(0.0614)	(0.0723)
cancerrisk	0.106	-0.121	0.767
	(0.516)	(0.486)	(0.497)
public	-0.380	0.631	-1.458*
	(0.620)	(0.786)	(0.782)
envir	0.941**	-0.723*	1.734***
	(0.438)	(0.414)	(0.479)
traffic	-0.246	-1.093***	0.542
	(0.376)	(0.375)	(0.431)
noexplan	-0.193	0.838**	-0.845**
	(0.361)	(0.339)	(0.368)
gdpgrowth	-0.195***	-0.133**	-0.115**
	(0.0415)	(0.0627)	(0.0510)
log(lifeexp)	-12.49***	-11.95***	-5.965
	(3.896)	(4.612)	(4.488)
log(num_stud)	-0.318*	-0.649***	0.0646
	(0.175)	(0.180)	(0.329)
samplesize	9.32e-05	0.000136	-3.39e-05
	(0.000141)	(0.000117)	(0.000183)
Constant	56.80***	54.79***	26.31
	(14.89)	(18.99)	(18.55)
Observations	500	500	500

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Only with prior US studies

VARIABLES	(1) IQRR 25-75	(2) IQRR 10-50	(3) IQRR 50-90
log(gdp)	0.0378	0.250	-0.463
	(0.220)	(0.312)	(0.302)
log(riskchange)	-0.165**	-0.210***	-0.170***
	(0.0660)	(0.0787)	(0.0632)
cancerrisk	0.165	-0.00200	0.753
	(0.814)	(0.371)	(0.933)
public	-0.305	0.499	-1.992**
	(0.902)	(0.745)	(0.883)
envir	0.626	-1.200**	2.494***
	(0.774)	(0.550)	(0.689)
traffic	-0.130	-1.218***	0.509*
	(0.358)	(0.365)	(0.291)
noexplan	0.0201	0.433	0.0111
	(0.437)	(0.306)	(0.401)
gdpgrowth	-0.225***	-0.0545	-0.229***
	(0.0783)	(0.0520)	(0.0579)
log(lifeexp)	-18.63***	-10.67*	-11.85***
	(5.776)	(5.578)	(4.008)
log(numstud from US)	-0.201	-1.063***	0.889
	(0.607)	(0.346)	(0.577)
samplesize	6.96e-05	0.000182**	-0.000235
	(0.000165)	(8.32e-05)	(0.000150)
Constant	81.47***	46.23**	53.64***
	(24.46)	(21.82)	(16.36)
Observations	500	500	500

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Are only non-US studies affected by selection?

VARIABLES	(1) IQRR 25-75	(2) IQRR 10-50	(3) IQRR 50-90
log(gdp)	-0.165	0.376	-0.427
	(0.416)	(0.400)	(0.623)
log(riskchange)	-0.177***	-0.249***	-0.168**
	(0.0616)	(0.0749)	(0.0761)
cancerrisk	0.250	-0.193	1.127
	(0.745)	(0.775)	(0.719)
public	-0.816	0.619	-1.682*
	(0.868)	(0.549)	(0.908)
envir	1.035**	-0.991**	2.068**
	(0.487)	(0.463)	(0.906)
traffic	-0.0413	-1.510***	0.588
	(0.434)	(0.450)	(0.409)
noexplan	-0.0873	0.736**	-0.363
	(0.402)	(0.372)	(0.369)
gdpgrowth	-0.243***	-0.0635	-0.211**
	(0.0582)	(0.0689)	(0.0857)
log(lifeexp)	-15.35**	-15.69***	-8.281
	(7.455)	(5.673)	(5.841)
log(numstud from US)	-0.346	-0.960**	0.133
	(0.491)	(0.416)	(0.697)
log(numstud from US) * US study	-0.0408	-0.176	0.0257
	(0.108)	(0.146)	(0.125)
samplesize	9.47e-05	7.78e-05	-7.06e-05
	(0.000144)	(9.33e-05)	(0.000163)
Constant	69.65**	66.11***	39.92**
	(27.99)	(23.42)	(19.37)
Observations	500	500	500

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Acknowledgement

This paper is part of my PhD thesis "Theoretical and empirical issues in non-market valuation". I would like to thank Nils-Axel Braathen and the OECD for allowing me to use their dataset.

References

Orley Ashenfelter and Michael Greenstone. Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias. *The American Economic Review*, 94(2):454–460, 2004. ISSN 0002-8282. URL http://www.jstor.org.lama.univ-amu.fr/stable/3592927.

François Bellavance, Georges Dionne, and Martin Lebeau. The value of a statistical life: A meta-analysis with a mixed effects regression model. *Journal of Health Economics*, 28 (2):444–464, March 2009. ISSN 0167-6296. doi: 10.1016/j.jhealeco.2008.10.013. URL http://www.sciencedirect.com/science/article/pii/S0167629608001549.

David Card and Alan B. Krueger. Time-Series Minimum-Wage Studies: A Meta-analysis. *The American Economic Review*, 85(2):238–243, 1995. ISSN 0002-8282. URL http://www.jstor.org/stable/2117925.

Chris Doucouliagos, T. D. Stanley, and Margaret Giles. Are estimates of the value of a statistical life exaggerated? *Journal of Health Economics*, 31(1):197–206, January 2012. ISSN 0167-6296. doi: 10.1016/j.jhealeco.2011.10.001. URL http://www.sciencedirect.com/science/article/pii/S0167629611001342.

Hristos Doucouliagos, T.D. Stanley, and W. Kip Viscusi. Publication selection and the income elasticity of the value of a statistical life. *Journal of Health Economics*, 33: 67–75, January 2014. ISSN 01676296. doi: 10.1016/j.jhealeco.2013.10.010. URL http://linkinghub.elsevier.com/retrieve/pii/S0167629613001422.

Roger Koenker. *Quantile regression*. Cambridge University Press, Cambridge ; New York, 2005. ISBN 0-511-13034-1 978-0-511-13034-2 0-511-13033-3 978-0-511-13033-5. URL http://dx.doi.org/10.1017/CBO9780511754098.

Henrik Lindhjem, Ståle Navrud, Nils Axel Braathen, and Vincent Biausque. Valuing mortality risk reductions from environmental, transport, and health policies: a global meta-analysis of stated preference studies. *Risk Analysis*, 31(9):1381–1407, September 2011. ISSN 1539-6924. doi: 10.1111/j.1539-6924.2011.01694.x.

OECD. Mortality Risk Valuation in Environment, Health and Transport Policies - OECD, 2012. URL https://www.oecd.org/env/tools-evaluation/mortalityriskvaluationinenvironmenthealthandtransportpolicies.htm.

Lisa A. Robinson and James K. Hammitt. Research Synthesis and the Value per Statistical Life: Research Synthesis and the Value per Statistical Life. *Risk Analysis*, 35(6):1086–1100, June 2015. ISSN 02724332. doi: 10.1111/risa.12366. URL http://doi.wiley.com/10.1111/risa.12366.

T. D. Stanley, Hristos Doucouliagos, and John P. A. Ioannidis. Finding the power to reduce publication bias. *Statistics in Medicine*, 36(10):

1580–1598, May 2017. ISSN 1097-0258. doi: 10.1002/sim.7228. URL http://onlinelibrary.wiley.com/doi/10.1002/sim.7228/abstract.

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

W. Kip Viscusi and Clayton Masterman. Anchoring biases in international estimates of the value of a statistical life. *Journal of Risk and Uncertainty*, 54(2):103–128, April 2017. ISSN 0895-5646, 1573-0476. doi: 10.1007/s11166-017-9255-1. URL http://link.springer.com/10.1007/s11166-017-9255-1.