THE LIMITS OF SELF-INTEREST: RESULTS FROM A NOVEL STATED-PREFERENCE SURVEY TO ESTIMATE THE SOCIAL BENEFITS OF LIFE-PROLONGING REGULATIONS

BY

ADAM M. FINKEL* & BRANDEN B. JOHNSON**

The benefits of health, safety, and environmental regulations will be underestimated if regulatory agencies produce risk estimates that are biased low,[1] *but also if they undervalue the risk reductions that regulations offer. In such cases, regulatory interventions that would generate social benefits in excess of their costs will mistakenly be turned down. Conversely, overestimation or over-valuation of benefits will result in adoption of regulations whose costs exceed their true benefits. For several decades now, both of the standard methods used to estimate the value of a statistical life (VSL)—stated-preference and revealed-preference designs—construe the tradeoffs involved at a "micro" scale: individuals are asked to (or observed as they) react to a very small probability of grave harm affecting themselves, and trade off small changes in economic welfare in response. But regulators then apply the resulting VSL estimates to a quite different set of circumstances, wherein society collectively marshals large sums of money (tens of millions of dollars or more) to avert hundreds or thousands of "statistical fatalities" over a large population, rather than a fraction of a single life.*

As part of a larger psychometric survey[2] *probing laypersons' "regulatory cost literacy" and their attitudes towards uncertain regulatory costs and benefits, we instead posed questions at a "macro"*

* University of Pennsylvania Law School and University of Michigan School of Public Health.

** Decision Research, Eugene, Oregon. We gratefully acknowledge generous support from the Human and Social Dynamics of Change Initiative of the National Science Foundation, through award SES-0756539 from the Decision, Risk, and Management Sciences program. We also thank a group of law students at the University of Pennsylvania who participated in a focus group to help us refine the wording of the survey, Adam Lake for programming the survey (including the sliders featured in Figure 2), Marcus Mayorga for overseeing the online panel and data collection, Leisha Wharfield and Andrew Quist for creating Figure 1, Laurent Abergel for research assistance, Nils Axel Braathen for information about the OECD dataset discussed at length in Section IV, and Paul Slovic for his usual insightful comments.

1 *See generally* Adam M. Finkel, *Is Risk Assessment Really Too 'Conservative'?: Revising the Revisionists*, 14 COLUM. J. ENVTL. L. 427 (1989).

2 Branden B. Johnson & Adam M. Finkel, *Public Perceptions of Regulatory Costs, Their Uncertainty and Interindividual Distribution*, 36 RISK ANALYSIS 1148 (2016).

scale that resulted in valuation estimates that may complement or challenge prior approaches to this issue. Depending on whether and how we excluded outliers, the mean response in our survey ranged from $5.5 million (in 2012 U.S. dollars, close to the roughly $9 million value currently used by many U.S. federal regulatory agencies) to $31.5 million. Our new stated-preference approach attempts to measure the "social value of life-prolonging regulatory benefits," rather than the respondent's personal valuation of a small change in her own mortality risk. Citizens may express higher implicit valuations when asked to gauge the acceptability of large life-prolonging regulatory programs because they are free to let altruism, and a concept we call here "shared purpose," affect their thinking. We regard "shared purpose" as a feature of these benefits estimates, not a "bug." Other implications of this novel approach for benefits imputation include the advantages of eliciting expressions of personal uncertainty rather than a single-point tradeoff, and the effects of anchoring tradeoff elicitation on either costs or benefits rather than solely on risk reduction benefits. We recommend more systematic reporting of inter-individual distributions of imputed valuations within micro-scale stated-preference studies. Time will tell whether this new method eventually becomes a useful complement to standard methods for informing regulatory benefit-cost analysis, or merely is a productive prompt to test the strengths and limitations of those standard methods more carefully.

I. INTRODUCTION

We may fervently believe that human life is "priceless,"[3] but society routinely trades off lives for dollars, both by inaction and by regulatory (and "soft law") action.[4] For supporters of stringent environmental, health, and safety regulation, and for believers in the principle that governmental

[3] *See, e.g.*, FRANK ACKERMAN & LISA HEINZERLING, PRICELESS: ON KNOWING THE PRICE OF EVERYTHING AND THE VALUE OF NOTHING 8 (2004).

[4] *See* Frank Ackerman & Lisa Heinzerling, *Pricing the Priceless: Cost-Benefit Analysis of Environmental Protection*, 150 U. PA. L. REV. 1553, 1563–64 (2002).

intervention is only justified when the benefits of the intervention exceed its costs, it is crucial that methods of analysis neither misestimate risks nor incorrectly value the benefits of risk reduction.

This Article reports on a new way to estimate the value that citizens place on life-prolonging regulatory benefits, one that challenges both the conventional goal of estimating the "value of a statistical life" (VSL) and the specific way economists currently estimate it.[5] VSL estimates are the primary way regulatory agencies now translate information about risk reduction into monetary benefits that can be compared with costs and used to guide decisions about whether and how stringently to regulate.[6] But there is an odd mismatch between the scenarios economists use to estimate VSL and the arenas in which lives are prolonged via regulatory intervention (or not prolonged because of decisions to eschew regulation). Simply put, estimates of VSL impute (or elicit) tradeoffs between a small probability of death to an individual and the individual's willingness to pay (WTP) to avoid it (in "stated-preference" (SP) studies), or her demands for higher wages to accept it (in the typical "revealed-preference" (RP) study).[7] The decision of whether and how stringently society should regulate, in contrast, is not a personal choice but a policy choice: should we *collectively* incur large financial costs to prolong many lives?

This question operates at a far different scale from the personal, and also quintessentially involves what we might term "shared purpose,"[8] an

---

[5] Although health, safety, and environmental regulations confer diverse benefits, including ecological improvements, increased quality of life, reduced property damage, and others, most *monetized* benefits (which by executive branch policy are what should exceed or justify monetized costs) are deemed to come from reduced mortality risks. Lisa A. Robinson, *How US Government Agencies Value Mortality Risk Reductions*, 1 REV. ENVTL. ECON. & POL'Y 283, 286 (2007); Regulatory Planning and Review, Exec. Order No. 12,866, 58 Fed. Reg. 51,735, 51,735 (Oct. 4, 1993). According to the White House Office of Information and Regulatory Affairs (OIRA), the 120 major federal regulations issued between 2004 and 2014 had combined monetized benefits of between $261 billion and $981 billion annually; of these, twenty-two were United States Environmental Protection Agency (EPA) air pollution rules which together provided $157 to $778 billion in these benefits, or roughly 70% of the total. OFFICE OF INFO. & REGULATORY AFFAIRS, OFFICE OF MGMT. & BUDGET, 2015 DRAFT REPORT TO CONGRESS ON THE BENEFITS AND COSTS OF FEDERAL REGULATIONS AND AGENCY COMPLIANCE WITH THE UNFUNDED MANDATES REFORM ACT 10–13, 15 (2015), https://perma.cc/SV2Y-J3YF; *see also* Robinson, *supra*, at 287 (noting how roughly 90% of the benefits of EPA air rules are attributed to monetized "reduced mortality risks").

[6] *See* Glenn C. Blomquist, *Value of Life, Economics of*, *in* 25 INTERNATIONAL ENCYCLOPEDIA OF THE SOCIAL & BEHAVIORAL SCIENCES 16,133, 16,138–39 (Neil J. Smelser & Paul B. Baltes eds., 2001); *see also* CASS R. SUNSTEIN, VALUING LIFE: HUMANIZING THE REGULATORY STATE 51 (2014).

[7] *See* Johnson & Finkel, *supra* note 2, at 1151, 1167 & n.19 (2016); *see also* RYAN C. BOSWORTH ET AL., THE VALUE OF A STATISTICAL LIFE: ECONOMICS AND POLITICS 8–9 (2017). However, RP estimates emerge from a market equilibrium and thus also depend on workers' knowledge of risk and their ability to bargain for their preferences, thus revealing more than simply WTP or willingness to accept risk. *Id.*

[8] We suggest that the traditional term "altruism" (also see later distinction between paternalistic and nonpaternalistic or "pure" altruism) is part of, but not as broad as, "shared purpose." In some cases (e.g., support for risk reductions accruing to individuals or groups one cares about, such as children in general), prior experiments have likely elicited both altruism (no direct benefits to self) and potential indirect benefits (e.g., people with children are more

enterprise that triggers citizens' desires to help improve the public welfare (including benefits to those close to them but not themselves), as well as any aversions they might have to paying for an intervention that will benefit people other than themselves (perhaps benefiting *only* unknown people).[9] Herein we describe a complementary approach to standard means of VSL estimation to raise broader questions about how society develops such crucial information. We raise the possibility that because traditional methods actually estimate the individual marginal rate of substitution (MRS) between wealth and mortality risk, they are not necessarily estimating the VSL, or that in any event *the VSL may not be the best unit estimator of the benefits of large-scale life-saving*. But we hasten to add that because our method never directly probes the MRS itself, it too may miss the target.

Regulatory decisions affect society as a whole, and significant regulations often involve expenditures of about $\$10^7$ to $\$10^9$ per regulation to reduce statistical fatalities on the order of about 10 to 10,000.[10] But the studies from which VSL estimates are either inferred or elicited do not involve observations or questions anywhere near this scale.[11] Instead, they

---

likely to support policies reducing children's risks). *See* Ryan Bosworth et al., *Is an Ounce of Prevention Worth a Pound of Cure? Comparing Demand for Public Prevention and Treatment Policies*, 30 MED. DECISION MAKING E40, E40, E53 (2010); *see also* ANNA ALBERINI ET AL., VALUATION OF ENVIRONMENT-RELATED HEALTH RISKS FOR CHILDREN 142 (2010). In contrast, our experiment offered subjects no information about the age of those whose lives would be prolonged via the hypothetical regulation, thereby possibly isolating the value subjects placed on helping to confer benefits on fellow citizens unknown to them. Shared purpose could also include mutual desire *not* to seek costly benefits accruing to unknown others. For example, Svensson and Johansson found that people who had negative feelings about government provision of goods and services (or the taxes they required), or preferred to provide the risk reduction themselves, had greater WTP in a private rather than public goods frame; these authors recommended the public goods frame be used if WTP responses were to be used to set public policy. Mikael Svensson & Maria Vredin Johansson, *Willingness to Pay for Private and Public Road Safety in Stated Preference Studies: Why the Difference?*, 42 ACCIDENT ANALYSIS & PREVENTION 1205, 1211 (2010).

[9] Scholars have long recognized that private and public valuations may diverge. *See, e.g.*, Blomquist, *supra* note 6, at 14 ("Value of life is about private choices that individuals make implicitly and explicitly about their own health and safety. Value of life is also about collective, public choices that societies make concerning tax and expenditure programs and especially regulatory programs that affect mortality risks."). Our concern here is that there has been little systematic investigation of how to *estimate* each type of value, and how to reconcile (conceptually and in implementation) any divergences found between the two types.

[10] Taking a simple arithmetic average of the impacts of the 120 major rules reported in the 2015 OIRA report to Congress, the typical major rule imposes costs of $600 million to $900 million to generate benefits of $2 billion to $8 billion. OFFICE OF INFO. & REGULATORY AFFAIRS, *supra* note 5, at 10–11 tbl.1-1. Hence, the base case in our survey (a hypothetical rule costing $1 billion or one that "saves" 1,000 lives) is in line with a typical and frequently-promulgated large rule.

[11] For excellent meta-analyses of RP (wage) studies, see W. Kip Viscusi & Joseph E. Aldy, *The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World*, 27 J. RISK & UNCERTAINTY 5, 5–6 (2003). *See also* W. Kip Viscusi, *The Role of Publication Selection Bias in Estimates of the Value of a Statistical Life*, 1 AM. J. HEALTH ECON. 27, 27 (2015) (finding under-estimates of VSL when earlier studies with greater measurement error in fatalities were included in meta-regression analyses). For similar meta-analyses of SP studies, see Anna Alberini, *Robustness of VSL Values from Contingent Valuation Surveys* 3–4 (U.S. Envtl. Prot.

typically involve fractions of one life "saved," and hypothetical WTP on the order of fifty dollars.[12] For example, someone willing to pay fifty dollars to eliminate a fatality risk to herself of one chance in 100,000 would be indicating a VSL of $5 million ($50 divided by 0.00001), which is close to (though below) the average value federal agencies now use.[13] Tradeoffs involving wage premia are typically of a slightly larger scale[14]—perhaps on the order of $1,000 in additional annual wages associated with an excess annual occupational risk of 1/5000 (again, yielding a VSL of $5 million). *So the scale of tradeoffs examined may differ from the scale of the actual decisions made by a factor of up to 1 billion-fold* (10^{-5} expected fatalities versus 10^{4}).

So why not ask people *directly* to evaluate tradeoffs in the (10^{9} dollars/10^{3} lives) range, rather than the (10 dollars/10^{-5} lives) range, to see if the change of scale—and the addition of shared purpose—affects VSL estimates? Scholars and regulatory agencies have been reluctant to pose these sorts of questions or utilize their results, due to concern that it is either inappropriate or quantitatively misleading to deliberately include altruism in VSL estimates.[15] We suggest that the first concern is misplaced, as policy makers at least need access both to purely self-interested and to other-regarding estimates of the value of life-prolonging interventions.

The concern about misestimating is more challenging to resolve, but researchers seem to agree that the problem here, if any, involves only one of two basic types of altruism.[16] "Paternalistic" or "safety-focused" altruism embodies concern only about the health/safety of another; a paternalistic altruist would want a fellow citizen to benefit from increased longevity even

---

Agency, Nat'l Ctr. for Envtl. Econ., Working Paper No. 05-01, 2005). *See also* Henrik Lindhjem et al., *Valuing Mortality Risk Reductions from Environmental, Transport, and Health Policies: A Global Meta-Analysis of Stated Preference Studies*, 31 RISK ANALYSIS 1381, 1381 (2011). For comparisons of the two methods, see NAT'L CTR. FOR ENVTL. ECON., OFFICE OF POLICY, U.S. ENVTL. PROT. AGENCY, VALUING MORTALITY RISK REDUCTIONS FOR POLICY: A META-ANALYTIC APPROACH 2 (2016) [hereinafter 2016 NCEE REPORT]; Paul Lanoie et al., *The Value of a Statistical Life: A Comparison of Two Approaches*, 10 J. RISK & UNCERTAINTY 235, 235 (1995).

12 Note that we placed "saved" in quotes because, in reality, regulations to reduce exposures to hazards do not eliminate anyone's risk of death, but may reduce or even eliminate the chance that one will die from one particular cause. Our survey carefully explained to subjects that "lives saved" is a misnomer, and that beneficial regulations *prolong* lives by reducing specific causes of death. Johnson & Finkel, *supra* note 2, at 1150 n.6.

13 *See* Lisa A. Robinson & James K. Hammitt, *Research Synthesis and the Value Per Statistical Life*, 35 RISK ANALYSIS 1086, 1087, 1089 & n.10 (2015) (noting the value of life that EPA was using at the time was $9.4 million, and that the United States Department of Transportation was using a range extending from $5.2 million to $13 million); *see also* Memorandum from Kathryn Thomson, Gen. Counsel, & Carlos Monje, Assistant Sec'y for Policy, to Secretarial Officers & Modal Administrators (June 17, 2015), https://perma.cc/D7B2-4BA9 (identifying a $9.4 million value of a statistical life for the Department of Transportation).

14 Robinson & Hammitt, *supra* note 13, at 1094 tbl.I.

15 *See* Kevin L. Brady, *The Value of Human Life: A Case for Altruism*, 48 NAT. RESOURCES J. 541, 542–43, 548–50 (2008).

16 *See* W. Kip Viscusi et al., *Altruistic and Private Valuations of Risk Reduction*, 7 J. POL'Y ANALYSIS & MGMT. 227, 227 (1988) (stating that altruism in risk reduction analysis ought to take account of the valuation "that other members of society place on their health").

if it came at a cost that citizen preferred not to incur.[17] In contrast, "non-paternalistic" altruism would value others' longevity only to the extent that their overall *welfare* (longevity benefits net of the welfare decrement due to increased expenditure) increased.[18] In theory, summing VSL estimates from citizens who in fact care about others' overall welfare (not just their health) will introduce upward bias (double-counting) into the equation, because each individual's estimate will include some "over-spending" on all their fellow citizens (spending that increases their health beyond that necessary to maximize their welfare). For this reason, the United States Environmental Protection Agency (EPA) has explicitly recommended that cost-benefit analysis might try to account for paternalistic altruism, and favors isolating that component via "a survey that would inform respondents about health improvements that others would experience from the policy, but also ask each respondent to assume that all others would be taxed an amount equal to their private willingness to pay for the policy, so that their utility remains unchanged."[19]

Our method as implemented to date cannot separate these two components of altruism, but it can provide an estimate of the value of life-prolonging outcomes possibly containing some double-counting that could be identified and removed.[20] We note, however, that at least one fact—United States welfare programs provide less than 10% of their value in cash, versus specific goods or vouchers for specific purchases—implies a national tendency to favor safety-focused over non-paternalistic altruism.[21]

We report here the results of an online survey administered to 744 lay Americans in mid-2012, in which implicit individual estimates of the value of life-prolonging regulatory programs were obtained as one part of a detailed "cost perception" experiment designed in large part to gauge respondents' attitudes about mean-preserving spreads of uncertainty in regulatory cost or benefit.[22] The survey is detailed elsewhere,[23] but briefly, we sought

---

17 Theodore C. Bergstrom, *Benefit-Cost in a Benevolent Society*, 96 AM. ECON. REV. 339, 340–41, 348 (2006); Kevin Lee Brady, Safety-Focused Altruism: Valuing the Lives of Others 3 & n.3 (Dec. 2008) (unpublished M.S. thesis, Utah State University), https://perma.cc/X2UQ-B3MD.

18 *See* Brady, *supra* note 17, at 3.

19 NAT'L CTR. FOR ENVTL. ECON., U.S. ENVTL. PROT. AGENCY, VALUING MORTALITY RISK REDUCTIONS FOR ENVIRONMENTAL POLICY: A WHITE PAPER 20 (2010); SCI. ADVISORY BD., U.S. ENVTL. PROT. AGENCY, REVIEW OF VALUING MORTALITY RISK REDUCTIONS FOR ENVIRONMENTAL POLICY: A WHITE PAPER (DECEMBER 10, 2010), at 13 (2011) (agreeing in a review of the 2010 White Paper that "[v]alues driven by paternalistic altruism are considered legitimate in benefit-cost analysis").

20 Others have suggested on theoretical grounds that VSL estimates including paternalistic altruism might be roughly 10% to 40% higher than purely self-interested ones. M.W. Jones-Lee, *Paternalistic Altruism and the Value of Statistical Life*, 102 ECON. J. 80, 89 (1992); *see also* Jorge E. Araña & Carmelo J. León, *Willingness to Pay for Health Risk Reduction in the Context of Altruism*, 11 HEALTH ECON. 623, 624, 630 (2002) (arriving at a similar range, 14% to 24% higher than private valuations, in a survey that estimated the effect of changing the hypothetical risk reduction from a private to a public campaign).

21 *See* Brady, *supra* note 17, at 28.

22 Johnson & Finkel, *supra* note 2, at 1149, 1154.

23 *See generally id.*

information about four basic questions: 1) How knowledgeable are laypeople about the costs of regulation(s)?; 2) Do laypeople interpret information from regulatory agencies about costs (and lives) as being exaggerated or understated?; 3) (How) do people's preferences for (hypothetical) regulations change when certainty of cost (or benefit) is replaced by a mean-preserving spread of uncertainty?; and 4) (How) do preferences change when (hypothetical) equal inter-individual distributions of regulatory costs are replaced with various unequal ones, both uncorrelated with personal income and correlated with it in various ways?[24]

II. GROPING TOWARDS SOCIAL BENEFITS VALUATION, FROM TWO DIFFERENT STARTING POINTS

We emphasize that subjects' responses to our "macro" questions do *not* yield estimates of the VSL, at least not as that quantity is currently defined. The VSL[25] is currently defined as the MRS between personal wealth and personal mortality risk,[26] intended to apply when that risk is small enough that endowment effects (inability to afford the amount that would compensate for the risk) are not at issue.[27] Here we refer to an individual's response to a micro-VSL elicitation as her MRS_i and the researcher's aggregation of those responses across a population as the MRS_p (some kind of aggregation is inevitable when individual responses differ, as they always do: should the MRS_p be defined as the mean (or median) over all responses, the maximum response recorded, or some other percentile of the distribution of responses?).

Our experiment did not elicit an MRS_i from each subject, but rather asked subjects about the desirability of their *contributing* some personal wealth to a national program costing $X (perhaps with some ambiguity about the exact size of one's personal contribution, see below) so that the national *incidence* of deaths associated with a particular hazard would decline by Y.[28] This absence of an MRS_i does not comport with the standard

---

24 *Id.* at 1149.

25 EPA has recommended that the term "VSL" be replaced with a term that limits misinterpretation that any identified persons are being "valued," or will lose their lives absent regulation. *See* SCI. ADVISORY BD., *supra* note 19, at i. The Science Advisory Board has endorsed "value of risk reduction." *Id.* at 5. Similarly, the staff of the EPA National Center for Environmental Economics has recommended "value of mortality risk" instead of "VSL." NAT'L CTR. FOR ENVTL. ECON., *supra* note 19, at 16. To the extent that these variant terms are still predicated on the MRS_i, our experiment does not yield estimates of the "value of mortality risk (reduction)" either.

26 Louis R. Eeckhoudt & James K. Hammitt, *Does Risk Aversion Increase the Value of Mortality Risk?*, 47 J. ENVTL. ECON. & MGMT. 13, 15 (2004).

27 *See* W. Kip Viscusi, *The Flawed Hedonic Damages Measure of Compensation for Wrongful Death and Personal Injury*, 20 J. FORENSIC ECON. 113, 114 (2007) ("[T]he cost of purchasing large decreases in the risk lowers one's wealth and reduces the willingness-to-pay amount below the VSL.").

28 We did not specify exactly how much money the subject would have to contribute, nor exactly how much personal risk reduction (if any) she would receive under the program, but we did show what personal cost would be under a uniform (per capita) apportionment of cost

paradigm for construing the VSL, but we argue here that our experiment can shed valuable light on the valuation of life-prolonging programs and hence of regulatory benefits.

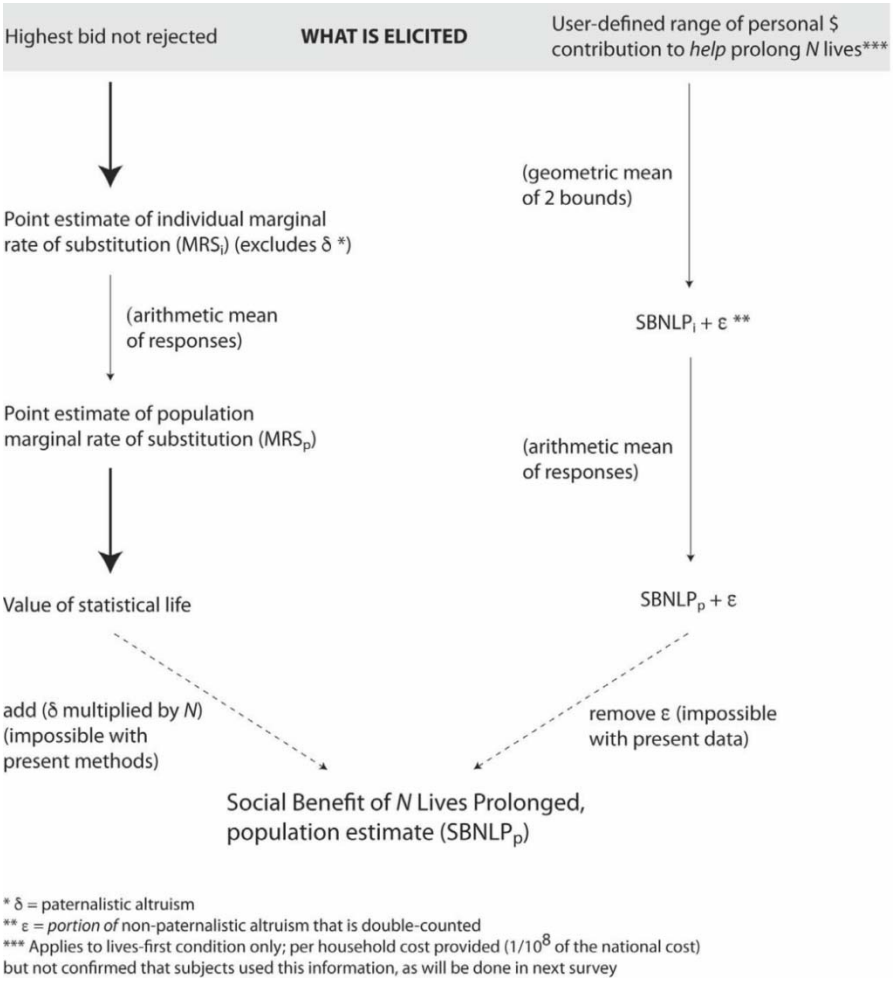
Indeed, just because a response to a particular kind of SP question yields an MRS_i, we question: 1) whether (once aggregated to yield an MRS_p) this is in fact a coherent estimate of the VSL (or of the "value of mortality risk reduction"); and 2) more importantly, whether the VSL (once multiplied by the number N of statistical fatalities averted) is in fact the desired measure of the *life-saving benefits of a national-scale risk-reduction regulation or policy.*[29] We suggest that *neither* the conventional VSL approach nor our approach yields precisely the desired measure, and thus that both are instructive.

Figure 1 shows how both traditional SP methods and our modified approach rely heavily on assumptions: the thin arrows denote algebraic or statistical transformations that may be wholly legitimate, while the thick arrows involve no computation, but simply make assertions that one quantity is synonymous with another. In particular, Figure 1 shows that traditional methods: 1) assume that the highest bid not rejected by each subject (whether in a direct question, bidding game, or payment card approach) equals her MRS_i; 2) define the mean (or trimmed mean, see Part IV below) of the individual responses as the MRS_p; and 3) crucially, assert that the MRS_p is the same as the VSL, and (tacitly) that policy makers can multiply the MRS_p by N to yield the life-saving benefits of any mortality-reducing intervention. We suggest each such assertion may be precarious, and that the end result [(N * MRS_p)] might more properly be thought of as "the average of personal valuations of 'N statistical lives like mine.'" *Instead, what decision makers really need might be quite distinct: "the social benefit of N lives prolonged," which we abbreviate here as the "SBNLP."*

while the subject considered his response in the lives-first condition, where he was defining, in dollars, acceptable and unacceptable levels of regulatory costs. *See* discussion *infra* Part III; *see also infra* fig.2. Notably, our subjects preferred and expected roughly equal cost distributions, although a subset expected to pay more than their (equal) share, with roughly equal numbers seeing this above-average share as being fair or unfair. Johnson & Finkel, *supra* note 2, at 1159.

[29] Put more simply, is the "value of a life" necessarily synonymous with the "value of a life '*saved*' or prolonged by a specific policy"?

Figure 1: Social Benefit of *N* Lives Prolonged



Our method instead assumes that the geometric mean of each subject's upper and lower bounds[30] that define the range of tradeoffs between dollars and lives she deems not-clearly-unacceptable yields the SBNLP *directly*, though we readily acknowledge that this attribute, depending on one's perspective, either "thankfully includes" or "is contaminated by" considerations of altruism and shared purpose. As discussed in Part I, paternalistic altruism is widely considered acceptable or even important to include in benefits valuation,[31] but our method as currently implemented yields an estimate of SBNLP that may well be biased high if it also includes some amount of non-paternalistic altruism.[32] Various other objections might

30 *See infra* Part III.

31 SCI. ADVISORY BD., *supra* note 19, at 13.

32 *See id.*

arise about our implementation of this method—in particular, that people lack contextual information on personal costs,[33] have no budget constraint,[34] and could engage in strategic answers,[35] but these also can be addressed in our (and, we hope, others') future research.

Again, we do not purport to estimate the VSL in this Article, but we *may* have derived an appropriate complementary estimate to the VSL, and may be on a pathway to deriving a more appropriate estimate of the desired quantity—the SBNLP—than any RP or SP estimate of MRS can provide. EPA's Science Advisory Board has recommended that cost-benefit analyses should rely if possible on studies that estimate the value of "public risk reductions."[36] We caution that the few prior studies of this type are valuable, but they do not necessarily elicit subjects' absolute willingness to contribute to regulatory programs solely as a function of the magnitude of their life-prolonging benefits, and none of them have directly probed subjects' WTP for specified increments of life-saving "purchased" collectively via costly regulation. For example, McDaniels confronted subjects with three different options to treat sewage in their community, but deliberately asked them only to choose among fixed alternatives, which the author said "require[s] less precision and [is] more readily accomplished on a large scale than

---

33 *See* Johnson & Finkel, *supra* note 2, at 1166–67. In addition to the information we provided on potential personal costs, the tradeoff task was preceded by a table providing some factual context (annual U.S. mortality from different causes of death, ranging from rare causes such as rabies to "all accidents"; annual U.S. expenditures on programs ranging from hurricane-tracking aircraft to the entire Department of Homeland Security budget; and life-saving costs on programs ranging from jet fighter ejector seats to heart transplants—including implied VSL and costs per U.S. household and per capita). This table was available as a pop-up throughout the tradeoff task. *Id.* at 1150.

34 Subjects *were* confronted as they made their choices with the per capita cost of the hypothetical program they were defining as "not unacceptable." *Id.* at 1152 fig.2. We think the parallel and largely unexamined issue is the value (and the pitfalls) of putting budget constraints into SP studies. If the research question concerns risk-reduction actions that the respondent would implement herself, *and* the answers would not be used to estimate values for cost-benefit analyses underpinning societal action, then a budget constraint is quite justified; no one is going to spend more than his net worth for his own micro-risk reduction. However, we wonder whether and on what grounds national policy tradeoffs should be constrained by the limits to private budgets, when the choice elicited has to do with the collective *national* payment needed to make a societal-level choice of risk reduction acceptable.

35 Strategic responses rather than true preferences could occur as well, if not more so, in micro-studies, whether the risk to be reduced is the respondent's own or not. *See* Richard T. Carson & Theodore Groves, *Incentive and Informational Properties of Preference Questions*, 37 ENVTL. & RESOURCE ECON. 181, 182, 188 (2007) (noting that when a private good framing is used it is always rational for the respondent to overstate WTP, because if the good is available in the market he can always choose later whether to actually purchase it). In the initial study reported here we did not specify whether the participant was included in, excluded from, or ambiguously among the regulation's beneficiaries, or ask whether the participant thought she was—omissions we will address in future work—but such omissions are if anything likely to *reduce* strategic answers relative to conventional studies that usually specify that the participant benefits from the risk reduction. *See* Johnson & Finkel, *supra* note 2, at 1149 (explaining that the survey asks for respondents' "own beliefs about tradeoffs between social costs" and "life-saving benefits" for society).

36 SCI. ADVISORY BD., *supra* note 19, at 13.

preference elicitation methods involving cardinal judgments."[37] Yet the tradeoffs McDaniels posited entailed rather impressionistic benefits compared to the precise benefits our subjects were considering: Table 3 in McDaniels, despite precise cost values, only states that the cheapest wastewater treatment option would yield "[r]eductions in some chemicals in wastewater," that a much more expensive option would give "[p]ossible further reductions in some chemicals," and that the most expensive option would also give "[p]ossible further reductions . . . ."[38] Similarly, Bergstrom et al. asked subjects if they were willing to pay one of eight specified amounts to reduce groundwater pollution from nitrates in their community (and also asked in an open-ended fashion the maximum amount the program could cost them before they would refuse to vote for it), but again, this survey did not explicitly ask for tradeoffs between cost and a precise number of lives prolonged (or any precise increment of benefit).[39] Koford used an innovative two-step approach: he asked subjects in Kentucky to allocate $100 million in new state spending across fourteen budget categories, and separately elicited private WTP for one of those categories (by asking which, if any, of eight different increments in their tax bill each subject would be willing to pay for an expansion of the state's community and technical college system).[40] Perhaps the closest analog to the work we report here was previously undertaken by Jones-Lee et al.[41] One of the many questions they asked their roughly 1,000 lay subjects was, "[i]f *you* were asked to make a single payment to help raise the money needed to avoid these 100 [annual] deaths [in England and Wales, from either motor vehicle accidents, heart disease, or cancer, whichever cause the subject previously deemed relatively most important], what is the most you personally would be prepared to pay . . . ?"[42] Although the mean response to this question was quite large relative to standard VSL estimates,[43] we note that the decision context did not necessarily have anything to do with regulation, externalities,

---

37 Timothy L. McDaniels, *The Structured Value Referendum: Eliciting Preferences for Environmental Policy Alternatives*, 15 J. POL'Y ANALYSIS & MGMT. 227, 228 (1996).

38 *Id.* at 246 tbl.3.

39 John C. Bergstrom et al., *Trading Taxes vs. Paying Taxes to Value and Finance Public Environmental Goods*, 28 ENVTL. & RESOURCE ECON. 533, 539–40 (2004).

40 Brandon C. Koford, *Public Budget Choices and Private Willingness to Pay*, PUB. BUDGETING & FIN., Fall 2010, at 47, 52–54. Koford's model posited that the WTP for any given budget category (*a*) could be approximated as the WTP for the community and technical college system (*b*) adjusted by the ratio *a/b* elicited in the relative-allocation part of the experiment. *Id.* at 52, 61, 65–66. While this elicitation did not trade off private dollars for any specific amount of public-good benefits, the two-step approach may be useful in estimating shared-purpose valuations of life-prolonging for hazards other than the ones we asked about.

41 *See generally* M.W. Jones-Lee et al., *The Value of Safety: Results of a National Sample Survey*, 95 ECON. J. 49, 49–51 (1985) (finding that concern for the safety of others increases VSL).

42 *Id.* at 51, 55–56 (emphasis in original).

43 *Id.* at 60 tbl.7 (noting the mean response was 20.3 million British pounds (1982 currency), or roughly $81 million in 2015 U.S. dollars). Note that although the question posed involved spending to avoid 100 excess deaths, the entries in these tables in Jones-Lee et al. were normalized to represent the valuation of a single statistical life. *Id.* at 58.

involuntary risks, or costs of cleaner or safer technology, and therefore is not necessarily germane to valuation for regulatory agencies.

In the remainder of this Article, we aim to further the debate on how best to elicit stated preferences for tradeoffs to advance the use of cost-benefit analysis in regulatory decision making, by reporting on four aspects of this initial implementation of our approach: 1) the method we used to elicit subject-specific bounds on personal uncertainty about the tradeoff's tolerability, and in turn to use those bounds to impute an implied valuation of life-prolonging; 2) the means and distribution of the valuation estimates obtained, primarily to raise questions about the importance of assessing the uncertainty and variability in all such estimates, given prior investigations' inconsistency about reporting them; 3) the effect of varying which attribute is traded off against which (i.e., either specifying a fixed dollar cost or a fixed number of lives prolonged and letting the respondent choose the magnitude of the other attribute); and 4) the effect of other variables on valuation, including demographic factors probed in some other SP studies, and constructed variables and attitudes gleaned from the survey itself.

We stress that we are not advocating replacing conventional micro-scale elicitations with macro-scale ones, only that findings from more use of the latter could complement the former.[44] If such results differ little from prior findings, this would suggest incorporation of shared purpose would not significantly affect valuation and hence regulatory choice, lessening the need for consensus on whether and how to include altruism in benefits estimation. But to the extent these results do diverge from the status quo, it is important to begin to understand how and why, as theoretical discussions continue about what to measure and how. Most importantly, it seems precarious to deem any phenomenon in cost-benefit analysis, including shared purpose, "important but non-quantifiable." For example, various scholars have expressed concern with the all-or-none way that society currently considers employment effects in regulatory analysis.[45] Similarly, the current interest in altruism seems to allow only one of two polar reactions: to treat it as interesting but ancillary to the main analysis (and hence unable to influence it at all), or to treat it as inchoate but let it dominate all of the quantifiable impacts that have been so carefully

---

44 Ample precedent exists not only for allowing a regulatory analysis to present two or more estimates of the same quantity (for illustrative purposes, or to probe the robustness of the decision to debated quantities), but for *requiring* multiple estimates. OIRA currently requires federal agencies to present costs and benefits whose future time-streams have been discounted both at a 7% and a 3% alternative rate. OFFICE OF INFO. & REGULATORY AFFAIRS, OFFICE OF MGMT. & BUDGET, REGULATORY IMPACT ANALYSIS: A PRIMER 11, https://perma.cc/B8GY-V2VN.

45 *See* Cary Coglianese & Christopher Carrigan, *The Jobs and Regulation Debate*, *in* DOES REGULATION KILL JOBS? 2, 2 (Cary Coglianese et al. eds., 2014) ("Only by developing better estimates of the real effects of regulation on employment can policy debate in the United States even hope to rise above the current polarized predicament where regulation's effects on jobs are too often either superficially treated or overblown by officials on both ends of the ideological spectrum.").

estimated.[46] Having access to a complementary set of valuations could allow policy makers to give altruism and other factors their *proportionate* weight, rather than zero or infinite weight.

III. IMPUTED VALUATION AS THE MIDPOINT OF TWO MORE ROBUST RESPONSES

In trying to design a meaningful and interpretable experiment on how subjects might be affected by a mean-preserving spread of uncertainty, relative to a definite cost or a definite number of lives prolonged,[47] we realized that as an intermediate step we would need to generate a new type of mortality-benefits estimate. We had initially considered the obvious path of giving all respondents a common example of a regulation with a particular cost or a particular life-saving benefit, and then permuting the quantity on the other side of the cost-benefit ledger with various uncertainty distributions. But unless the researcher "individualizes" the questions for any particular respondent, the gamble presented may seem pointless no matter how certain or uncertain the compared scenarios. For example, the question "how would you feel about buying an ice cream cone that definitely costs $100, versus one that would turn out to cost you something between $80 and $120?" would be unlikely to elicit useful information, as the answer would probably be "I am supremely uninterested in either choice, and certainly see no difference between the two" (and ditto if we had asked about a cone that might have cost two pennies, versus an uncertain gamble between one and three cents).

Therefore, to ensure that the contrast between certain and uncertain estimates would offer each respondent a meaningful choice, we allowed each subject to first set her own *personalized* range of (un)acceptable costs or lives.[48] Half of the respondents read a "lives-first" scenario, in which a hypothetical regulation (either an EPA regulation to control a carcinogen in drinking water or a National Highway Traffic Safety Administration (NHTSA) regulation to reduce fatal automobile crashes) would definitely save 1,000 lives nationwide, and were asked to think about acceptable and unacceptable ranges for the cost of that intervention.[49] To do so, the respondent was first asked to choose a cost figure that was so low she

---

46 For a related example, the National Highway Traffic Safety Administration (NHTSA) recently promulgated a rule (requiring backup cameras on passenger cars) whose quantified costs exceeded quantified benefits. Federal Motor Vehicle Safety Standards; Rear Visibility, 79 Fed. Reg. 19,178, 19,180 (Apr. 7, 2014) (to be codified at 49 C.F.R. pt. 571). NHTSA did so on the grounds of an "exceptionally high emotional cost" (unquantified) linked to the main hazard, parents unable to see their own toddlers in the path of a car they were driving. *Id.* at 19,180–81. Just as an *emotion-adjusted valuation* might have shed light on how cost-justified this rule truly was, having access to *shared-purpose valuation estimates* might enrich discussions about many regulatory choices.

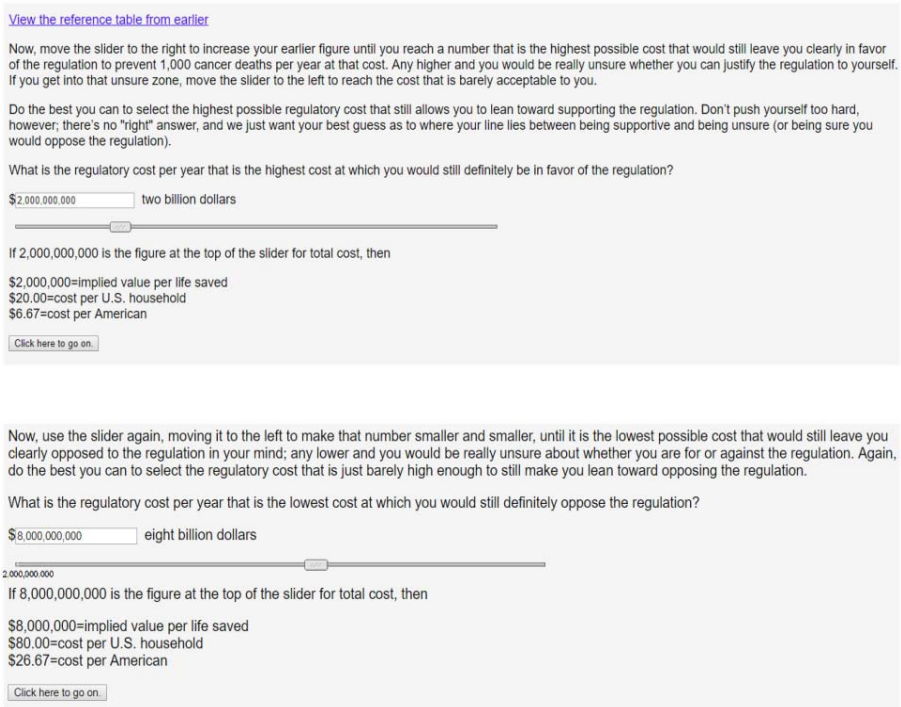
47 Johnson & Finkel, *supra* note 2, at 1149.

48 *Id.* at 1150–51 ("[W]e developed personalized uncertainty ranges off of each subject's *own* reported range where she had a definite preference either for or against a regulation with a given balance of risk reduction and cost.").

49 *Id.* at 1150–54.

would unequivocally support the regulation as a "bargain" (i.e., the cost figure she chose is low enough to make the "saving" of 1,000 lives clearly a worthwhile investment).[50] Then, using a slider on the computer screen (see Figure 2), the respondent was asked to increase that lower bound until she reached a point where she was *no longer certain* that the regulation was a bargain (we call this bound L).[51] Then, the upper bound U (the cost above which she would surely oppose the regulation as too profligate for the lives prolonged) was elicited similarly, in two steps.[52]

Figure 2: Hypothetical Subject's Response (and Revised Response) to the Lives-First Range-Finding Valuation Question



[50] *Id.* at 1149–51. A static printout of the entire instrument is available at *What Can Regulatory Economics Learn from Risk Assessment?*, SITES.GOOGLE.COM, https://sites.google.com/a/adamfinkel.com/nsfprojectoutputs/home (last visited July 14, 2018), a site that also contains other published and under-review outputs of the National Science Foundation project of which this study was a part. Readers who wish to mimic the experience of completing the survey can access one version of it at *Online Survey on Regulatory Costs and Uncertainty*, U. PENN. L. SCH., https://www.law.upenn.edu/cf/faculty/afinkel/RegCost/ (last visited July 14, 2018). No results will be captured.

[51] Johnson & Finkel, *supra* note 2, at 1151–52 & fig.2.

[52] *Id.* at 1151. Subjects who instead received a "costs-first" version (EPA or NHTSA) were asked to first choose a large enough number of lives prolonged (U) that a fixed cost of $1 billion would be a bargain in their eyes, and then asked to choose a small enough number (L) such that "too few" lives would be prolonged for $1 billion in cost. *Id.* at 1151, 1153.

The range between these bounds demarcates the region where the respondent neither definitely supports nor definitely opposes the regulation.[53] For the lives-first respondents, the quantity ($L/1000) is the lower bound on the implicit value they place on prolonging one life as part of a large-scale project; the quantity ($U/1000) is the corresponding upper bound.[54] For the "cost first" respondents, the bounds are instead ($1 billion/U) and ($1 billion/L).[55] *In the results that follow in Part V, we interpret the geometric mean of the two bounds as a respondent's implicit valuation of one life prolonged out of many.*[56]

Our approach of eliciting two bounds rather than a single "tipping point" contrasts with conventional approaches, where researchers aim for a single-number answer per respondent (e.g., the maximum value accepted in a bidding game).[57] Given inevitable constraints on the duration of an elicitation, no scholar can presume that the "accepted" value is indeed the maximum acceptable value, but in practice it is treated as such. Such methods as the double-bounded binary choice (asking for a second cost amount conditioned on the initial choice) reduce the sample size needed for a given confidence interval on overall sample responses, but still yield one cost number per respondent.[58]

While our method does not preclude a single answer (if someone reports a single value where the regulation shifts abruptly from "a bargain" right to "a clear waste"), in practice we expected (and found) that participants respond with more uncertainty than that. We believe that eliciting a range avoids various problems of eliciting overly precise point estimates (problems most severe with direct questions, but not limited to that format). The elicitation of a precise point estimate makes two presumptions: that people already have a coherent opinion on the tradeoff,

---

53 *Id.* at 1151. We later made use of these bounds not only by using their geometric mean as the point of comparison between certainty and uncertain cost (or benefit) estimates proffered by federal agencies, but by structuring uncertain estimates in two ways: a "narrow" estimate whose range of cost (or benefit) lay entirely within the respondent's own grey area, and a "wide" estimate whose range extended well above and below the subject's bounds (but still preserving the mean). *Id.* at 1152–53.

54 *See id.* at 1151.

55 *See id.*

56 For costs-first respondents, the geometric mean valuation is equal to (1 billion/$\sqrt{(U\ x\ L)}$); for lives-first respondents, it is equal to ($\sqrt{(U\ x\ L)}$/1000). For example, if a respondent said that to save 1000 lives, a cost of $1 *billion* or less would definitely be worthwhile, and a cost of $9 billion or more would definitely be wasteful, his implicit valuation would be $3 *million*. We believe the geometric mean is the proper summary statistic: if someone is not sure (say) whether one thing is "10 times better" or "1/10 as good" than another thing, 1:1 would be the central tendency of his views, not 5.05 to 1. Using arithmetic means would give large numbers far more weight than small ones, which we felt would be inappropriate in general, but even more so when employing a novel method.

57 *See* Richard T. Carson & Mikołaj Czajkowski, *The Discrete Choice Experiment Approach to Environmental Contingent Valuation*, *in* HANDBOOK OF CHOICE MODELLING 202, 204 (Stephane Hess & Andrew Daly eds., 2014); *see, e.g.*, Jones-Lee et al., *supra* note 41, at 52–53, 55–58 (using a maximum value bidding game to elicit single-number answers when surveying valuation of safety).

58 *See* Carson & Czajkowski, *supra* note 57, at 207–08.

and that they are able to access it properly when asked for it properly by researchers.[59] As outlined in a classic paper,[60] when labile (changeable or unstable) values are at stake—with "values" defined as evaluations of the "relative or absolute worth or desirability of possible events"—one or both of these conditions may fail to be met.[61] A respondent may have a coherent opinion but may access only part of it or something else entirely, may have an incoherent opinion, or may have no opinion but may try to form one based on inadequate evidence or deliberation.[62]

By relieving the pressure to come up with a single point macro-scale estimate—or even to produce a range of any particular width—*we argue that laypeople will have an easier and more reliable time demarcating unacceptable from ambiguous than they would finding the exact tipping point at which support turns to opposition, if indeed such a point even exists.*[63]

To sum up, we are inclined to believe that stability is a strength of the double-bound approach to macro-scale elicitation, though our aim here is simply to propose that it opens up a new suite of provocative and potentially fruitful questions (e.g., test-retest reliability, calibration of individual and social valuations) that are worth exploring whatever the elicitation method used.[64]

---

59 *See* Richard T. Carson et al., *Consequentiality: A Theoretical and Experimental Exploration of a Single Binary Choice*, 1 J. ASS'N ENVTL. & RESOURCE ECONOMISTS 171, 173 (2014) (claiming that use of a consequential single binary choice with incentive-compatible format can yield "truthful preference revelation," per mechanisms defined in Richard T. Carson & Theodore Groves, *Incentive and Informational Properties of Preference Questions*, 37 ENVTL. & RESOURCE ECON. 181, 198–203 (2007)).

60 *See generally* Baruch Fischhoff et al., *Knowing What You Want: Measuring Labile Values*, *in* COGNITIVE PROCESSES IN CHOICE AND DECISION BEHAVIOR 117 (Thomas S. Wallsten ed., 1980).

61 *Id.* at 117–19 (suggesting "elicitation procedures" cause respondents to answer "with the first thing that comes to mind" and then "commit . . . to maintaining that first expression and to mustering support for it, suppressing other views and uncertainties").

62 *Id.* at 120 tbl.7.1.

63 Another way to think about this is that the "ideal" single value is by definition a value for which minor deviations may not be discernible. Consider the "ideal" temperature for the water in a swimming pool—if it rises in one-degree increments, the boundary between "too cold to bear" and "not too cold to bear" can be readily ascertained, whereas the experience of "ideal minus one degree" and "ideal plus one degree" may be identical.

64 We also emphasize that the method and data we discuss in this Article are *not* related to ambiguity aversion. *See* Mary Riddel & W. Douglass Shaw, *A Theoretically-Consistent Empirical Model of Non-Expected Utility: An Application to Nuclear-Waste Transport*, 32 J. RISK & UNCERTAINTY 131, 131–32, 145–46 (2006) (discussing the negative externalities arising from ambiguity about the risks from nuclear waste transport). The tradeoff task in our survey allowed, though it did not mandate, participants' use of wide bounds to express their own uncertainty about acceptable or unacceptable tradeoffs. However, the expression of uncertainty and the aversion to uncertainty are not identical. *See* Johnson & Finkel, *supra* note 2, at 1151.

IV. RESULTS: MEANS AND DISTRIBUTIONS OF IMPUTED VALUATIONS

As we present the following summary statistics for our subjects' implicit valuations, we naturally were interested in whether they are high or low compared to the range and midpoint of those emerging from all prior studies, especially other SP studies.[65] However, such a comparison is far from straightforward, in part because of uncertainty in the mean estimates from each study and especially because of the much larger variability in inter-subject responses. Our impression of the VSL literature is that it is relatively rare for moments other than means of the distribution of elicited tradeoffs to be reported; hence there is little insight offered into the range of subjects' valuations and in turn the reliability of such estimated means. Sometimes a report of trimmed samples implies that trimming of some number or percentage of outlying responses produces greater reliability, but we have not observed any systematic discussion of when trimming should be done or how extensive it should be.

We report in Table 1 the data on means and standard deviations, medians, and fifth and ninety-fifth percentiles for our full sample, and three trimmed sub-samples, which (respectively) excluded specific outlying values (values less than or equal to $100, or greater than or equal to $1 billion), the top and bottom 1% of values regardless of their magnitude, or the top and bottom 5%. The first group of six rows reports these data for our full sample, the second for those given the costs-first anchor, and the third for those given the lives-first anchor. In the absence of reference points from the SP VSL literature standardizing how and how much to trim, we note that the current EPA standard value of $9.7 million[66] falls between our fifth and ninety-fifth percentile values in all three groups (full group, lives-first subsample, and costs-first subsample), and that for all three groups the EPA value falls between the mean when the top and bottom 1% of values are trimmed and the mean when the top and bottom 5% of responses are trimmed. Our overall grand mean when extremely low and high individual responses are trimmed ($18.4 million, in 2012 dollars), is about twice the current EPA standard value, but the corresponding median value is about 10% of the current EPA estimate (as noted above, we cannot discern how these medians relate to the medians of SP VSL distributions).

---

65 At this writing, the most recent meta-analysis of SP and RP studies comes from EPA, which considered forty-two SP and forty-six hedonic-wage (HW) estimates; EPA found an inter-study range of mean VSLs to be from $1 million to $24 million, with cross-study grand mean estimates of $8.6 million (SP) and $11.9 million (HW). 2016 NCEE REPORT, *supra* note 11, at 17–19.

66 *Id.* at 2 (this table is in 2013 dollars and uses 2013 income levels).

Full Sample	Total	Excluding ≤ $100 & ≥ $1 billion	Excluding Top & Bottom 1%	Excluding Top & Bottom 5%
N	733	637	697	665
Mean	$31,499,343	$18,416,134	$16,136,284	$5,538,754
Standard deviation	$180,126,601	$71,491,563	$67,188,408	$14,848,235
Median	$342,525	$632,456	$316,228	$316,228
5th percentile	$1	$1,080	$1	$5
95th percentile	$102,277,236	$100,000,000	$77,606,185	$31,622,777
Costs-first				
N	393	383	387	368
Mean	$34,774,369	$19,875,285	$19,320,392	$6,517,581
Standard deviation	$140,093,322	$70,641,746	$69,723,224	$16,152,844
Median	$830,455	$816,497	$766,001	$624,834
5th percentile	$1,772	$2,403	$558	$482
95th percentile	$141,421,356	$111,533,954	$110,052,004	$32,912,581
Lives-first				
N	340	254	310	297
Mean	$27,746,908	$16,215,918	$12,368,901	$4,306,156
Standard deviation	$217,273,233	$72,839,115	$63,959,372	$12,942,744
Median	$54,772	$329,141	$69,282	$90,499
5th percentile	$0	$362	$0	$1
95th percentile	$45,116,165	$68,352,063	$36,276,543	$30,000,000

Table 1. Distribution of Imputed Values of Prolonging a Statistical Life

We note[67] that for all four choices of whether and how to trim outliers, the mean valuation of the half-sample who saw the costs-first framing of the tradeoff ("How many lives would need to be prolonged to justify your contributing to a $1 billion national-scale regulation?") was about 30%–50% higher than the mean for subjects who saw the lives-first framing ("What amount of spending by yourself and others would justify a regulation that prolonged 1,000 lives nationwide?").

To summarize what one can and cannot discern about inter-subject variability in prior studies, we refer to the 1,010 conventional VSL estimates (in 2005 U.S. dollars) featured in a meta-analysis of studies for the Organisation for Economic Co-operation and Development (OECD).[68] Our assessment is that their mean *study* estimate—i.e., not the mean of

67 *See infra* Part VI.

68 Henrik Lindhjem et al., *Valuing Mortality Risk Reductions from Environmental, Transport, and Health Policies: A Global Meta-Analysis of Stated Preference Studies*, 31 RISK ANALYSIS 1381, 1384 (2011). Besides using estimates that Lindhjem et al. converted to 2005 U.S. dollars, we used the responses they subsequently normalized via "purchasing power parity" factors to represent the amount in 2005 U.S. dollars that would purchase the same amount of market goods as one unit of the respondent's home-country currency. *Id.*; *Meta-Analysis of Value of Statistical Life Estimates*, ORGANISATION FOR ECON. CO-OPERATION & DEV., http://www.oecd.org/env/tools-evaluation/env-value-statistical-life.htm (last visited July 14, 2018) (follow "Click here" hyperlink for the raw data for this study).

individual-level estimates, as in our Table 1—was $8.7 million (minimum $4,450, maximum $206 million), with a median of study estimates of $2.8 million.[69]

It is also not clear, however, whether the large coefficient of variation and mean/median ratio in our study compared to prior estimates of inter-subject variability is to some (large?) extent a function of: 1) more severe trimming of outliers by prior studies; or 2) the fact that our experiment allowed subjects to choose any non-negative value for both the lower or upper bounds of the acceptable-tradeoff region, while many prior SP studies offered only specific investigator-provided bids that subjects could only accept or reject.[70] Assuming that the inter-subject range of WTP cannot exceed the range of bids, it may be informative that the ratio of the highest to lowest possible bid in the OECD dataset was often about 10, though for some studies the ratio was as high as 100. This difference further complicates attempts to reconcile or contrast our estimates with prior work.

We offer these distributional inferences from the OECD dataset not because they offer definitive conclusions about the value of either the conventional or our novel approach to imputing values of life-saving, but because they raise intriguing questions. Why do only 42% of the OECD estimates offer a standard error, thus severely constraining the number of studies to which our distributional data could be validly compared in terms of individual-level maxima and minima? If comparing variances can indeed help identify the more reliable values or approach, the proper comparison of variances may be only with those conventional estimates that allow similar freedom to their participants to set the parameters of the task. Further research could help resolve whether any differences in imputed valuations between the two methods reflect substantive, methodological, or information-processing factors.

V. THE IMPORTANCE OF WHAT CITIZENS TRADE FOR WHAT

Conventional VSL elicitation provides subjects with a specified mortality reduction, and begins and ends by asking them to report how much they value monetarily this risk reduction.[71] As the bottom two-thirds of Table 1 shows, the converse (costs-first) approach—in effect, "how much risk reduction would be needed to justify this cost?"—yielded a significantly

---

69 A standard error was provided for 495 (42%) of these estimates, allowing for calculation of the coefficient of variation (CV; standard error divided by the mean). The average CV in the OECD dataset was 0.44, compared to ours of about 6.0, but 20% of the former exceeded 1, with a maximum of about 4.5 (and the CV of our sample reduces substantially, to about 0.5, when the top and bottom 5% of values are trimmed; again, it is possible that some studies in the OECD dataset have *already* been trimmed by the original investigators). If we measure skewness by the mean/median ratio, ours was 89 for the full sample and 16 for the 5% trimmed sample; most OECD mean/median ratios were about 4, but one was as high as 106. *See Meta-analysis of Value of Statistical Life Estimates*, *supra* note 68.

70 *See* Johnson & Finkel, *supra* note 2, at 1151.

71 *See* discussion *supra* Part I.

higher imputed valuation than this lives-first framing.[72] Again, our aim here is not to argue that basing elicitation on one of these attributes rather than the other is necessarily superior, either for national or personal tradeoffs, but to ask whether this substantive difference signals a critical issue with regard to the reliability or generalizability of previous VSL estimates based on SP studies, as these studies use only the lives-first framing.[73]

Starting with lives might fixate people upon the relatively small number involved, yielding smaller numbers for imputed VSL than when they start with the much larger numbers for costs. For example, it could be more palatable to "only" save twenty lives when $1 billion seems already committed[74] to the regulation, rather than to urge spending $50 billion of public money to save 1,000 lives at risk, even though the imputed VSL is $50 million in both cases. Alternatively, perhaps lives as an initial attribute induce more thoughts of "infinite value" than dollars, and thus yield more resistance to the notion of putting a value on lives.

Whichever of these processes, alone or in combination, is at work in this difference, it raises basic questions about the approach taken in conventional contingent valuation studies. For example, how might VSL estimates based on SP studies differ if those studies had used a costs-first framing (e.g., "It will cost you $5 to reduce your risk; how much risk reduction would you need to receive in return for this $5 to make the cost worthwhile?") instead of or in addition to the standard lives-first framing? Both framings are novel tasks for most people, so whether one is cognitively easier to process than the other needs to be tested. We also note the increasing congressional interest in imposing a "regulatory budget" upon individual agencies or upon the entire regulatory apparatus.[75] Such a budget constraint would require agencies to consider how best to confer benefits given a cap on costs,[76] which might well require them to consider how much benefit citizens expect for a given cost, *rather than the converse formulation*.

---

72 *See supra* tbl.1.

73 While standard SP studies can pose their elicitation in two *seemingly* opposite phrasings—e.g., "what are you willing to pay to eliminate a risk?" versus "what sum are you willing to accept to let me impose a risk?"—these are both "lives-first" questions wherein money values are elicited. Johnson & Finkel, *supra* note 2, at 1151.

74 It may seem odd to refer to an agency estimate of future costs associated with implementation of a regulation as implying that funds are "committed." However, subjects' explanations for why they expected the EPA to under- or over-estimate costs incurred indicated that many people made an analogy between regulatory cost estimates and cost overruns in government projects and home contracting, in terms of being under-estimated and/or commitments to "spend" the money. *Id.* at 1157, 1163.

75 Jeffrey A. Rosen & Brian Callanan, *The Regulatory Budget Revisited*, 66 ADMIN L. REV. 835, 837 (2014); *see also* Adam M. Finkel, *A Healthy Public Cannot Abide Unhealthy and Unsafe Workplaces*, 108 AM. J. PUB. HEALTH 312, 313 (2018).

76 *See id.* at 853–54.

VI. "EXPLAINING" VALUATIONS DERIVED FROM MACRO-SCALE ELICITATION

Empirical and conceptual efforts in the standard SP literature tend to assess whether and how imputed valuations differ across a limited number of explanatory variables, particularly age, income, health status, and risk type.[77] As a brief insight into how these and other variables may influence our subjects' responses, we regressed the 5% trimmed estimate against selected factors. These included income, age, and risk type (chronic or acute hazard), to parallel some factors explored in earlier studies.[78] Two experimental manipulations from the survey other than risk type were also included: 1) the focal attribute (costs or lives) used as the basis for eliciting preferred tradeoffs, and 2) whether respondents were shown a sentence summarizing the range of VSLs used by federal agencies, as a test of whether these might anchor respondents' own preferences.[79] We also used as an explanatory variable the respondent's confidence in her tradeoff preferences—defined as the ratio of the high to low bounds each reported—and used dummy variables for whether the respondent thought EPA would under-estimate regulatory costs and would be accurate about benefits (deaths postponed), both majority views in this sample.[80]

Table 2 shows that little of the total variance was explained by these factors, with the overall regression analysis insignificant. Only income and the cost-versus-lives anchor were significant predictors. The imputed valuation was higher for those with lower incomes, a finding at odds with those from most micro-VSL studies.[81] This might indicate that, in the context of shared national purpose, poorer people no longer are constrained by their own income limits in indicating their "true" valuation, or perhaps that very rich people value "shared purpose" less than others (these speculations, of course, come with all the prior caveats). The absence of age effects is consistent with some prior findings.[82] The lack of effects of most other variables tested here *might* indicate that the high valuations identified are not peculiar to a specific demographic group, methodological variation, or

---

77 *See* Anna Alberini et al., *Does the Value of a Statistical Life Vary with Age and Health Status? Evidence from the US and Canada*, 48 J. ENVTL. ECON. & MGMT. 769, 771 (2004) (showing weak effects of age on decline in VSLs, mainly for those older than seventy); Mary F. Evans & V. Kerry Smith, *Do We Really Understand the Age-VSL Relationship?*, 28 RESOURCE & ENERGY ECON. 242, 251 (2006) (finding no significant effect of age on VSL for a U.S. sample and a significant negative effect on a Canadian sample); Alan Krupnick, *Mortality-Risk Valuation and Age: Stated Preference Evidence*, 1 REV. ENVTL. ECON. & POL'Y 261, 261 (2007) (discussing the idea that older people may have a lower VSL); *see also* Maureen Cropper et al., *Valuing Mortality Risk Reductions: Progress and Challenges*, 3 ANN. REV. RESOURCE ECON., March 2011, at 313, 330 (providing mixed evidence about the valuation of cancer versus non-cancer health and safety risks); Koford, *supra* note 40, at 56 tbl.1 (offering mixed evidence for how income, education, sex, and race affected relative allocations across public spending categories).

78 *See supra* note 77 and accompanying text.

79 Johnson & Finkel, *supra* note 2, at 1154.

80 *Id.* at 1156 tbl.I.

81 *See* W. Kip Viscusi, *The Benefits of Mortality Risk Reduction: Happiness Surveys vs. The Value of a Statistical Life*, 62 DUKE L.J. 1735, 1741 (2013).

82 Alberini et al., *supra* note 77, at 769–71.

belief about agency performance—i.e., they reflect estimates from all subjects—but both the wide variance in responses in this study and our small number of predictors warrant testing of this speculation.

	Standardized Coefficient	Statistical Significance
Youth (1 = < 45 years old, 0 = others)	.02	.671
Elders (1 = > 64 years old, 0 = others)	-.01	.854
Income	-.08	.048
Risk type (1 = carcinogen, 0 = traffic accident)	.01	.825
Anchor (1 = cost, 0 = deaths postponed)	.09	.024
VSL sentence (1 = yes, 0 = no)	.03	.465
Confidence (ratio of upper to lower bound)	-.02	.599
Expect under-estimation of cost (1 = yes, 0 = no)	-.01	.766
Expect accurate estimates of deaths postponed (1 = yes, 0 = no)	-.05	.188
F, p	F(9,616) = 1.32	.226
R^2	.019	
Adjusted R^2	.005	

Table 2. Linear Regression of Imputed Valuations of Life-Saving in the 5% Trimmed Sample

VII. DISCUSSION AND CONCLUSIONS

This Article has raised various questions about how conventional SP studies impute values of a statistical life. We raise these questions despite our support for the trend in the United States and the European Union of agencies relying increasingly on SP studies over RP ones, as by definition an RP study cannot yield VSL estimates that reflect any altruism or shared purpose (no one accepts a riskier job or buys a house in a neighborhood with higher levels of environmental contaminants out of concern for society).[83] We have suggested that it may be possible to elicit valid national-

[83] *See* BOSWORTH ET AL., *supra* note 7, at 8–9 (explaining that RP studies elicit trade-offs between risk and income and could not be used to discern an individual's WTP to "buy cleaner air for their community"). A small subset of RP studies *do* attempt to infer the VSL by observing large-group behavior that may involve shared purpose, but we argue that these studies do not "reveal" individual willingness to contribute to risk reduction programs. For example, among studies of decisions by legislatures to allow more (or less) mortality risk from a particular source, the estimated VSL from 21 states opting to allow higher highway speed limits in the late 1980s ranged from a negative number (in two states) to roughly $9.7 million in 1997 dollars. Orley Ashenfelter & Michael Greenstone, *Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias*, AM. ECON. REV., May 2004, at 454, 458. These estimates could be reasonable if legislators and voters were fully aware of the expected

level tradeoffs from citizens between regulatory costs and lives prolonged, from which benefits valuations can be imputed that *might* converge with micro-level valuations, once certain methodological and conceptual issues are addressed. At the same time, we have suggested that, in the absence of more information about data distributions beyond their means, even conventional SP data may be hard to justify as providing stable imputed VSLs; the relatively narrow range of VSL estimates may be spurious, representing methodological constraints as much or more than it does public consensus about the value of a statistical life. Our results therefore may spur attempts to test the claim that "[t]here is little empirical evidence that altruistic concerns are significant drivers of values for risk reduction."[84] *That statement may be literally true simply because there have been few if any wholly relevant attempts to elicit valuations that incorporate, rather than studiously exclude, altruistic values and other considerations that go beyond individual MRS elicitation.*

We acknowledge that it is very difficult to tease out the individual effects of such simultaneous major departures from previous contingent valuation studies embedded in our initial study, including: 1) changing scale, 2) introducing the potential for shared purpose to affect responses, 3) eliciting two bounds instead of a single tipping point between acceptable and unacceptable tradeoffs, 4) eschewing fixed bids in favor of wholly user-generated bounds, and 5) framing the tradeoff as starting with costs for some respondents and with lives for others.[85] Further, the huge variance in individual estimates imputed here suggests that we need to understand better which kinds of people give answers that policy makers might not wish to weight fully. Because we cannot parse any differences in these results (compared to prior results) as due to one or more of these novel aspects of our survey, and because we believe that ours *and all prior results* are sensitive to specific choices about statistical estimators and the handling of

---

increase in fatalities in their state due to the higher speed limit, if they construed its economic benefits as the hours their citizens saved by driving faster, and if they valued each hour saved at the prevailing average wage rate in the state. We suggest all three of these assumptions are reasonable but precarious, and therefore that our approach of asking citizens directly about contributing personal dollars to national risk-reduction programs with specified results may be a more reasonable path to eliciting the SBNLP. Another subset of articles try to infer VSLs from decisions by regulatory agencies, insurance carriers, and others to require or allow certain risk-reducing interventions whose costs and benefits can be estimated. *See, e.g.*, Tammy O. Tengs et al., *Five-Hundred Life-Saving Interventions and Their Cost-Effectiveness*, 15 RISK ANALYSIS 369, 369, 371 (1995). The Tengs et al. compendium reveals, however, a gigantic variation in the efficiency of life-saving interventions, with many items on their list prolonging life and saving money simultaneously, while others "costing" upwards of $99 *billion* per life-year extended (equivalent to a VSL of more than $1 trillion). *Id.* at 373–84. Others have criticized this article for mistakenly including various interventions as "revealing" a valuation when in fact the interventions were rejected by regulatory agencies precisely because they *were* so relatively inefficient at reducing risks. Lisa Heinzerling, *Five Hundred Life-Saving Interventions and Their Misuse in the Debate over Regulatory Reform*, 13 RISK: HEALTH, SAFETY & ENV'T 151, 152, 155–56, 159 (2002).

84 SCI. ADVISORY BD., *supra* note 19, at 13.

85 Johnson & Finkel, *supra* note 2, at 1149, 1151, 1166.

outliers (choices which prior investigators often do not report transparently), we emphasize that additional work should be done to systematically compare each choice of survey design, using common methods of estimation and handling of outliers. For example, future research could include an experiment that directly contrasts the double-bound versus single tipping point approaches for both macro-tradeoffs and conventional micro-scale questions, using both costs-first and lives-first framings, among other variations. Such an experiment could help isolate the effect of each alternative we have proposed, and begin to reveal whether and how individual respondents impute different values for life-prolonging benefits depending upon whether they ponder personal or societal tradeoffs.

Another important step would be for authors of prior estimates based on SP methods to report the distribution of their subjects' responses in similar detail to the data presented here. Without such data, it is impossible for anyone to conclude whether the large inter-subject variance reported here is normal for SP research or a result unique to this study's methods, subjects, or both. We encourage future meta-analyses of SP studies, whether they include macro-scale questions or not, to synthesize distributional information from each study rather than to shoehorn a set of overconfident point estimates into a pseudo-distribution.[86]

We expect the notion of complementing micro-scale with macro-scale imputation to be controversial, and we welcome the debate. Our anticipation is primarily driven by curiosity as to whether we have made a serendipitous discovery, but we also welcome any arguments as to why analysts and decision makers should *not* make use of survey questions at policy-relevant scales, incorporating paternalistic altruism, shared purpose, or both, as a necessary feature of the estimation rather than a "bug." Perhaps the addition of more expansive distributional data to such estimates will add little insight, and perhaps additional research taking into account the psychological literature on cognitive challenges will demonstrate that eliciting national-level tradeoffs validly and reliably is even more difficult than it is at the micro-level of standard SP methods. But we have faith that the discussion itself will be fruitful for benefit-cost analysis, and helpful for decisions about the life-prolonging regulations dependent on these analyses.

---

[86] *See* Adam M. Finkel, *The Cost of Nothing Trumps the Value of Everything: The Failure of Regulatory Economics to Keep Pace with Improvements in Quantitative Risk Analysis*, 4 MICH. J. ENVTL. & ADMIN. L. 91, 127–28 (2014) (offering in Table 5 a typology that criticizes the practice of defining two or more incompatible estimates as a "range of uncertainty").