Appalachian
STATE UNIVERSITY

Department of Economics Working Paper

*Review of Contingent Valuation of
Environmental Goods: A Comprehensive
Critique.*

*Edited by Daniel McFadden and Kenneth
Train (2017): An Update*

John C. Whitehead
Appalachian State University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

Review of *Contingent Valuation of Environmental Goods: A Comprehensive Critique*.

Edited by Daniel McFadden and Kenneth Train (2017): An Update[1]

John C. Whitehead[2]
Department of Economics
Appalachian State University
Boone, NC 28608

Revised: August 8, 2024

Abstract. This paper updates the review of the 2017 BP-funded book critical of the contingent valuation method (McFadden and Train, 2017) that was published in the *Australian Journal of Agricultural and Resource Economics* (Whitehead, volume 64, number 2, pp. 710-713, 2018). In that review I noted that an expanded review, with summaries of each chapter, was available upon request. A few requests trickled in but I always responded that the expanded review was not quite ready to share. This was primarily due to a comment that I was writing on a chapter in the book and other on-going work that I wanted to include. That comment has been published and my retort to the reply was posted as a working paper in 2024. Therefore, it is high time to finish the extended review of the book. In this paper I include the original review and add two appendices. One appendix contains the short reviews of each chapter promised in 2018 and another is the 2021 proposal narrative for a forthcoming book that more fully responds to McFadden and Train (2017).

---

[1] As of this writing, a PDF copy of the McFadden and Train book can be downloaded for free by searching for "mcfadden and train contingent valuation of environmental goods a comprehensive critique" in Google Scholar.

[2] Email: whiteheadjc@appstate.edu

Review of *Contingent Valuation of Environmental Goods: A Comprehensive Critique*.

McFadden and Train, editors, (2017): An Update

John C. Whitehead

Most economists naturally prefer to use revealed preference approaches, e.g., the travel cost method or hedonics, for the valuation of environmental and natural resources. Revealed preference approaches are commonly believed to provide unbiased estimates of the economic value of changes in recreation, health and other endpoints related to environmental and resource policy, i.e., use values. Oftentimes, revealed preference approaches do not cover the full range of environmental values. It is not unusual for an environmental policy to create a behavioral context that has not been observed in the past. Researchers will not have any possibility of collecting revealed preference data for ex-ante policy analysis. For example, revealed preference methods are limited when used to estimate the recreation value of new fishing sites or pollution cleanup that enhances residential property. These situations are outside the experience of consumers and, therefore, unfamiliar environmental goods. No ex-ante revealed preference data exists with which to estimate the values.

Stated preference approaches, such as the contingent valuation method (CVM) and the closely-related discrete choice experiments (DCE), are useful when ex-ante valuation information is needed for these gaps. Using survey approaches, stated preference approaches can elicit hypothetical behavior data that can be combined and used in tandem with revealed preference data. A fairly large "joint estimation" literature has developed over the past 25 years (Whitehead Haab and Huang 2011). Another gap is due to non-use values which Krutilla (1967) first suggested could be an important part of the total value of the environment. Many

2

environmental resources may provide non-use values (aka, existence value, passive use value) and for some of these the non-use value component may dominate the use values. Since there is no observable behavior that would allow revealed preference methods to estimate non-use values, stated preference methods are the only approach capable of estimating non-use values.

The primary purpose of McFadden and Train (2017) is to discredit stated preference (both CVM and DCE) approaches for the measurement of non-use values. Not explicated stated, the context is Natural Resource Damage Assessment (NRDA) compensation for the BP/Deepwater Horizon oil spill in the United States[3]. McFadden and Train's effort is not original. After the Exxon Valdez oil spill in 1989 Exxon hired a number of consultants to examine the ability of the CVM to estimate economic values suitable for NRDA. The consultants for Exxon conducted a number of studies designed to discredit the CVM (Maas and Svorenčík, 2017). A conference was held to publicize the results of the studies undertaken on behalf of Exxon. Proponents of the CVM, many of them working as consultants for the State of Alaska, were invited to the conference and allowed to comment on the studies. A proceedings volume was published (Hausman 1993). The effort to criticize a valuation methodology, undertaken on behalf of Exxon to avoid paying damages to the State of Alaska, led to what has become known as the "CVM Debate" (Banzhaf 2017).

McFadden and Train (2017) follow this same path, sans the conference, in discrediting the CVM (and DCE). A large number of consultants from several major firms were hired, studies were conducted and another book has been published critical of the CVM (and DCE).

---

[3] In the United States, compensation to resource trustees from oil spills via Natural Resource Damage Assessment is allowed through the Oil Pollution Act of 1990.

The current book mostly follows the approach taken by Hausman (1993). Several chapters focus on the historically major weaknesses of stated preference methods and some focus on newer issues. Some of the data collected with funding from BP is of relatively low quality. Overly broad generalizations are made from these studies without placing them in the context of the literature. The studies are conducted by researchers with limited experience in stated preference methods. The data analysis for each chapter leaves many questions. A good objective round of peer-reviews would have tightened each chapter.[4]

The criticisms of the CVM are often inconsistent and contradictory. For example, using data collected with funding from BP, Chapter 7 closes the "fat tail" that Chapter 2 says is a pervasive problem. The reader is left to wonder what is driving this difference. Chapter 5, the scope test chapter, eliminates internal scope tests (i.e., using repeated valuation questions) from inclusion in the meta-data while Chapter 6 states that repeated questions are necessary for respondent learning (in the context of the use of stated preference marketing surveys by profit-maximizing firms). The reader is left wondering why an internal scope can be used for market goods but not for non-use values. Chapter 6 states that opt-in panel data is of relatively low quality and leads to marketing errors while Chapters 3 and 4 rely on BP funded opt-in panel data. None of these inconsistencies are recognized by the editors. In other words, the chapters have not been thoroughly edited to develop a cohesive message.

While CVM applications to recreation and health have been shown to have convergent validity with revealed preference approaches, this sort of application of CVM is never

---

[4] Each chapter is discussed in more detail in Appendix A.

recognized. The authors take aim at non-use values in the context of NRDA, leaving naïve readers with the impression that this is the only type of application for which CVM is applied. The editors conclude that the CVM (and DCE) is of little use in in policy making. Some of the issues raised in the chapters are legitimate concerns, but the book provides no suggestions or guidance on how to improve the state-of-the-art of the CVM. The criticism in Chapter 6 are never recognized as potential suggestions for future CVM applications.

The final nail in the coffin is the last chapter, a legal opinion that the CVM should not be used for natural resource damage assessment because it has no precedent. This capstone chapter is suggestive that the studies published in the book were the raw material for expert testimony in front of a judge presiding over BP's NRDA. The heading of the last section of this chapter, "trustees should abandon contingent valuation and similar methods as a matter of policy," makes clear the ultimate purpose of the book. The entire book describes a bleak dystopia.

McFadden and Train have missed a major opportunity following the BP/Deepwater Horizon oil spill. Instead of a balanced discussion of the advantages and disadvantages of stated preference methods in a broad range of contexts, they focus on the disadvantages in a single narrow context where are only a few studies have been conducted (NRDA). The book does not address the broader purpose of the CVM and other stated preference methods, which is to provide measures of economic value for ordinary benefit-cost analysis when there are no conceivable revealed preference approaches available to obtaining these values (or, simply in the spirit of basic research, to develop a better understanding of environmental values). The danger of the book is that it will have a chilling effect on attempts at research that has a goal of improving the accuracy of environmental valuation. In this sense, the editors have done a

disservice to the economics profession and to the general public that economics research is intended to serve.

Following the BP/Deepwater Horizon oil spill the economics profession is waiting for an objective book that recognizes the intended purpose of stated preference methods and critically evaluates strengths and weaknesses with a primary goal of improving the accuracy of environmental valuation.[5] An improvement in the accuracy of environmental values will almost certainly enhance the efficiency of modern-day economies by improving economic decision-making. "Abandon(ing) contingent valuation and similar methods as a matter of policy" will only benefit oil companies and other business firms that insult the environment and wish to avoid full compensation.

Conflict of Interest Statement: The author worked on behalf of the State of Florida conducting a CVM study of economic damages to support litigation against BP. The State of Florida ultimately joined the Federal case against BP and neither the Florida nor Federal study was used for litigation.

References

Banzhaf, H. Spencer. "Constructing markets: environmental economics and the contingent valuation controversy." History of Political Economy 49, no. Supplement (2017): 213-239.

Hausman, Jerry A., ed. Contingent Valuation: A Critical Assessment. Elsevier North Holland,

---

[5] In 2021, Tim Haab, Dan Petrolia and myself signed a contract to write such a book with Edward Elgar. The preliminaries of the book proposal are in Appendix B.

1993.

Krutilla, John V. "Conservation reconsidered." The American Economic Review 57, no. 4

(1967): 777-786.

Maas, Harro, and Andrej Svorenčík. ""Fraught with Controversy": Organizing Expertise against

Contingent Valuation." History of Political Economy 49, no. 2 (2017): 315-345.

Whitehead, John, Tim Haab, and Ju-Chin Huang, eds. Preference data for environmental

valuation: combining revealed and stated approaches. Vol. 31. Routledge, 2012.

Appendix A: Chapter Reviews

The book consists of an introduction, ten empirical chapters and a final chapter that

contains a legal opinion. Overall, these chapters raise some valid points that are worthy of future

study but the authors seem mostly satisfied that they have done their job for BP.[6] In contrast, the

book should be read with a mindset of solving these problems with better survey design, data and

empirical methods. Economic science does not advance in the context of litigation over natural

resource damage assessments (NRDA) but in the push and pull of academic research published

in peer-reviewed journals. If the book is of any use beyond internal BP legal machinations, it

would be as a motivator to develop ideas to address these issues. In the remainder of this

---

[6] Note that the vast majority of the chapter authors have not published an article on the
contingent valuation method or stated preference methods in the peer-reviewed literature since
the book was published. Several "forthcoming" papers are referenced but these can not be found
in Google Scholar. Emails to the authors requesting these papers have been either ignored or
responded to without a paper attached.

appendix I quickly review each chapter, hitting the highlights and offering paths forward.[7]

Chapter 1

Burrows et al. use choice experiment data from Wallmo and Lew (2012) and find 4 of the 44 versions of the survey are identical except for the cost amounts.[8] They split the sample into high and low-cost amount versions and find that willingness to pay is higher with higher cost amount versions. There are several problems with the analysis. First, the experiment design work done in Wallmo and Lew is tossed out. The choice experiment was designed to generate data to cover the entire range of the willingness to pay distribution so it should be estimated as a full sample. Second, the empirical models are estimated with the conditional logit, not accounting for the panel nature of the data (unless there is more to footnote 11). Third, the full sample is analyzed in Chapter 8 and partial explanations for the Chapter 1 results are provided. Finally, the authors use this result to suggest that survey respondents are not sensitive to cost amounts in spite of the overwhelming evidence in the literature that respondents, in general, are more likely to be less willing to pay higher cost amounts. Chapter 1 leaves the stench of cherry picking.

Chapter 2

Desvousges, Mathews and Train show that estimates of willingness to pay do not pass the

---

[7] Another (negative) review of this book was written by Ståle Navrud and published in *Economics of Energy & Environmental Policy* in March 2018. Navrud includes summaries and critiques of each chapter.

[8] The data was obtained by an anonymous Freedom of Information Act request instead of the more usual request to the authors. The authors were surprised to learn that their data were used in this book. I cry foul.

adding up test. This chapter was originally published in *Land Economics* in 2015. I wrote a comment on this article that was eventually published in *Ecological Economics* in 2020. Desvousges, Mathews and Train's published reply pointed out an embarrassing statistical error and a number of other "mistakes". In a recent working paper I correct the error and argue that the other mistakes aren't mistakes at all (Whitehead 2024). In short, the adding up test used by Desvousges, Mathews and Train with the Turnbull estimate of willingness to pay does not replicate with parametric estimates of willingness to pay. The data used by Desvousges, Mathews and Train is of such poor quality that the key willingness to pay confidence interval has an enormous upper tail: the willingness to pay point estimate is $434 and the confidence interval ranges from $234 to $2062. In contrast, the Chapman et al. (2009) study, from which Devousges, Mathews and Train claim they use the same survey instrument, has a confidence interval of $245 to $328 with a point estimate of $279. One is correct to wonder why the estimates differ so much.

Chapter 3

Parsons and Myers address the problem that survey respondents tend to be insensitive to the cost amount in the tail of the distribution. This suggests that respondents are unwilling to indicate that they won't support the policy, engaging in so-called "yea-saying" behavior at some high-cost amount threshold, and thereafter. This is a troubling result because data that suffers from fat tails will lead to upward biased willingness to pay estimates. The Parsons and Myers study is well-done, with little to quibble about. Respondents may have considered cost amounts at the higher end of the range to be less than credible, but it is not clear when that might be the case. This should have led to ongoing research to address the fat tails problem. But, nothing

empirical has appeared in the literature until recently (see Lewis, Richardson and Whitehead 2024). In a forthcoming paper presented at the 2022 SBCA conference[9], I designed a survey that included increased cost amounts past what may have been wise, similarly to Parsons and Myers, but not to the point of incredibility. I find that there is a fat tail problem in the upper third of the cost amounts. Employing an attribute non-attendance model, I find that there is a 36% probability that respondents do not pay attention to the cost amount. Once accounting for cost nonattendance, willingness to pay falls by 50%. See also Anonymous (forthcoming).

Chapter 4

Parsons and Myers re-employ their survey and estimate implicit discount rates from willingness to pay estimates from different payment schedules (one-time payment vs in-perpetuity payments). They find unreasonably large discount rates. There are many studies in the literature that demonstrate this result and, also, many that estimate reasonable discount rates (e.g, Howard, Whitehead and Hochard 2021). The question is: what are the factors that affect respondent sensitivity to the payment schedule? Examining Parsons and Myers Table 3, it is obvious that the survey respondents did not differentiate between the two payment schedules. In a forthcoming study, I employ an experimental design similar to Parsons and Myers. Employing an attribute non-attendance model, I find that non-attendance increases with the length of the payment schedule and implicit discount rates are reasonable, 7%, once attribute non-attendance is accounted for.

---

[9] Whitehead, John C., "Estimating the Economic Valuation of the Crab Bank Seabird Sanctuary," Paper presented at the 2022 Society for Benefit-Cost Analysis Conference, Washington DC, March 2023. The paper is being written as a chapter in the book described in Appendix B.

In the longest chapter of the book, by far[10], Burrows, Newman, Genser and Plewes review the split-sample scope test literature and conduct a meta-analysis on the determinants of scope elasticity. The first part of the chapter updates the review from Desvousges, Mathews and Train (2012) with a focus on environmental goods that provide nonuse value. In this review the authors try to determine if an article passes the scope test, fails the scope test or if the results are mixed.[11] One criticism of this categorization of scope test results is the same as I made in Whitehead (2016). Some researchers present a variety of models and the purpose of the comparison is to provide insights on when a scope test might pass or fail. Most studies that find mixed results actually find results that pass the scope test under certain conditions. Another concern with this part of the chapter is that Burrows et al. exclude internal scope tests. In an internal scope test, survey respondents are presented with more than one valuation question. Burrows et al. reject this because of the purported anchoring effect. Respondents might demonstrate scope with their survey responses to be internally consistent. While this may be true, the economics literature does not care very much, as every article that uses discrete choice experiments to estimate willingness to pay is an internal scope test (also see McFadden's chapter 6). Therefore, this part of the Burrows et al. review is limited to a narrow slice of the valuation

---

[10] Chapter 5 is 71 pages long. The next longest chapter is 36 pages. One might have thought to split the chapter into two chapters of more reasonable length (in the same way I should have this into paragraph in two).

[11] It is not clear how much the reader should trust Burrows et al.'s pass and fail grades on these studies. The authors find that the Banzhaf et al. (2006) study fails the scope test. Yet, the abstract in Banzhaf et al. (2006) concludes with "The instrument passes external scope, sensitivity to bid, and sample selection tests."

literature. Burrows et al. then attempt a meta-analysis with "pass" as the dependent variable. The models are not very informative with only a few statistically significant coefficients. The majority of the chapter the focus is on statistical significance. In a later section of the chapter, Burrows et al. consider economic significance by calculating scope elasticities (and finally address the scope "adequacy" that is promised in the title). Whitehead (2016) shows that a scope elasticity is likely to range between zero and one, increasing with the sensitivity of willingness to pay to the scope of the policy. Burrows et al. are obsessed with categories of elasticity, claiming that 0.2 and 0.5 thresholds might be considered useful for determining scope adequacy or plausibility. Unlike quantitative thresholds for own-price and income elasticities, there is not objective justification for asserting thresholds for adequate or plausible scope elasticities.

Chapter 6

McFadden begins the chapter with a thirty-one year old quote from Rick Freeman in his chapter of the 1986 book on the contingent valuation method (Cummings, Brooshire and Schulze 1986). As it turns out, Freeman had changed his mind about the CVM before the McFadden and Train book was published, coming around to the idea that recent advances have improved the validity and reliability of contingent valuation data in the second and, especially, third editions of his RFF book on valuation (Freeman, Herriges and Kling 2014) (pages 411-412): "Our own assessment of SP methods is cautiously optimistic. … the preponderance of evidence suggests that a carefully executed SP study provides valuable insights into the tradeoffs that individuals are willing to make to secure or avoid changes to the environment." More substantively, McFadden goes on to argue that single-bound willingness to pay questions are lacking in that survey respondents need to gain experience through repeated questioning. I agree and (since a

disaster of a survey) in 2017 I have abandoned the single bound question and have been asking

repeated referendum questions in a number of studies. This is the type of data used in the models

that I describe as forthcoming in the reviews of Chapters 3, 4, and 5 above.

Chapter 7

In what may be the most interesting chapter in the book, Kemp, Leamer, Burrows and

Dixon consider a "'composite good' approach to contingent valuation". The chapter is a

continuation of the Kemp and Maxwell (1993) study where the authors elicit willingness to pay

in a "top-down" approach. Willingness to pay is first elicited for a several public goods and then

the willingness to pay for each good is elicited so that the values add up to total willingness to

pay. The authors are negative about the validity of the resulting willingness to pay estimates. It is

therefore interesting that Kemp et al. took another look at the top-down approach in this book.

This is especially true considering the cost of the study. The sample is from the GfK probability-

based Knowledge Panel which around this time cost about $50 per completed survey. With a

sample of 4000, the total cost of the data would be $200,000 in 2017 dollars. The conclusions of

this chapter are largely negative about contingent valuation as in Kemp and Maxwell. For

example, one of the authors' conclusions is that "the similarity of these two response curves

reveals a colossal scope failure." The primary takeaway from this chapter is to avoid spending

$200,000 of your research budget on top-down willingness to pay questions. Researchers should

be encouraged to try various approaches beyond dichotomous choice/referendum valuation

questions as recommended by Johnston et al. (2017), but my advice is to do this with less

expensive opt-in panel data (as was done in Chapters 2 and 4).[12]

Chapter 8

Leamer and Lustig confront the problem that not all stated preference survey respondents pay attention to every thing in the survey when answering valuation questions. The attribute non-attendance literature was underway and this was a well-known issue at the time (see footnote 3 on page 226). The authors find that only about a quarter of the survey respondents in the Wallmo and Lew (2012) discrete choice experiment data, used also in Chapter 1, paid full attention to the costs and benefits of the survey scenarios. Another part of the attribute non-attendance story is that respondent sensitivity to attributes (cost, payment schedule, scope) is better measured after non-attendance is accounted for[13]. This result could have been applied to the anomalies found in other Chapters of the book, but this approach was not considered by any of the other authors or recognized by the editors of the book.

Chapter 9

Myers, MacNair, Tomasi and Schneider consider researcher deletion of observations based on certain rationality criteria elicited through willingness to pay follow-up questions. Although not a widespread practice at the time of the book, the authors are motivated by some

---

[12] Another interesting aspect of the chapter is that the authors seem to be excited about their work. In footnote 3 they report that "It is anticipated that these aspects of the overall study will be reported in other papers." As of August 8, 2024, this chapter has been cited once in the peer-reviewed literature according to Google Scholar suggesting that the authors have not published any extensions of this work. Emails asking about these papers have gone unanswered.

[13] See Whitehead, Cornicelli and Howard (2024) for a recent example.

researchers' decisions to delete "protest no" responses, increasing willingness to pay estimates. These are survey respondents who respond "no" to a willingness to pay question because they reject the valuation exercise. The authors, admittedly, include a large number of follow-up questions and multiple thresholds to determine rejection (stringent, lenient). They suggest that only 2 (!) of 1224 respondents form a "rational core" of respondents who are worthy of analysis. The purpose of the chapter is consistent with the purpose of the book and is summarized by a question posed by the authors (page 267): "Should policy decisions and legal damages be assessed using information obtained from people who appear to give invalid responses to follow-up questions such as these?" The authors refuse to provide an answer except implicitly. Better guidance on this issue is found in Johnston et al. (2017) who suggest that, what they call, auxiliary questions can be important for better understanding willingness to pay estimates but "responses to auxiliary questions may be endogenous to valuation responses and while a useful part of the survey, may have limited use in the estimation of valuation response equations."

Chapter 10

Foster and Burrows conduct a meta-analysis of hypothetical bias studies with a sample size that is larger than in previous studies. The sample size is 432 pairs of hypothetical and real willingness to pay estimates from 77 studies. They assert that their sample includes hypothetical and real willingness to pay values that contain non-use value, but it is not clear how this is accomplished. Elsewhere in the book, various authors state that non-use values cannot be connected with behavior so it is surprising to see that 197 of the 432 estimates contain non-use values. The authors argue that not enough variation in the dependent variable (hypothetical WTP divided by real WTP) is explained by a small set of explanatory variables (10) in order to use the

ratio as an adjustment factor for policy or NRDA. Note that R^2 values in previous meta-analyses are 2 to 4 times higher than found here. They achieve a similarly low R^2 value in their regressions with an expanded set of explanatory variables. As with many conclusions in the chapters of this book, a reasonable person might consider the glass half-full and find these results useful for policy analysis.[14]

Chapter 11

Israel, Martin, Fayne and Daniel provide a legal analysis of the use of the contingent valuation method in the NRDA courtroom. They review three NRDA cases that used contingent valuation estimates of willingness to pay and find that the courts rejected its use as evidence in two of the cases. In the third case, the defense presented a lower willingness to pay estimate from a contingent valuation study but the jury decided on the plaintiffs higher value, from a benefit transfer based on the travel cost method. Chapman and Hanemann (2001) describe these two estimates and speculate on why the jury may have rejected the CVM estimate (in short, it had problems). Notably, the Exxon Valdez oil spill case is not included in this review. Carson (2012) describes how the CVM estimate was used in an out of court settlement with Exxon. A non-legal reason that the authors give for not using the contingent valuation method is the expense of a study (page 305): "CV methods are extraordinarily expensive." This may be beside the point, but, I submit that the expense of a valid contingent valuation study is increasing in the number of attorneys interested in the results.

---

[14] Penn and Hu published their first hypothetical bias meta-analysis one year later (Penn and Hu 2018). The sample size in Penn and Hu (2018) is 908 pairs of hypothetical and real willingness to pay estimates from 132 studies.

References

Anonymous, "Mitigating cost vector effects in stated choice experiments using cheap talk and opt-out reminders," Journal of the Association of Environmental and Resource Economists, forthcoming.

Banzhaf, H. Spencer, Dallas Burtraw, David Evans, and Alan Krupnick. "Valuation of natural resource improvements in the Adirondacks." Land Economics 82, no. 3 (2006): 445-464.

Carson, Richard T. "Contingent valuation: A practical alternative when prices aren't available." Journal of economic perspectives 26, no. 4 (2012): 27-42.

Chapman, David J., and W. Michael Hanemann. "Environmental damages in court: the American Trader case." Chapter 15 in The Law and Economics of the Environment, Edited by Anthony Heyes (2001): 319-368.

Chapman, David, Richard Bishop, Michael Hanemann, Barbara Kanninen, Jon Krosnick, Edward Morey and Roger Tourangeau. 2009. Natural Resource Damages Associated with Aesthetic and Ecosystem Injuries to Oklahoma's Illinois River System and Tenkiller Lake.

Cummings, Ronald G., David S. Brookshire, William D. Schulze, "Valuing Environmental Goods: An Assessment of the Contingent Valuation Method, Rowman & Allanheld: Totowa, NJ, 1986.

Desvousges, William, Kristy Mathews, and Kenneth Train. "Adequate responsiveness to scope in contingent valuation." Ecological Economics 84 (2012): 121-128.

Freeman III, A. Myrick, Joseph A. Herriges, and Catherine L. Kling. The measurement of
environmental and resource values: theory and methods. Routledge, 2014.

Howard, Gregory, John C. Whitehead, and Jacob Hochard. "Estimating discount rates using
referendum-style choice experiments: An analysis of multiple methodologies." Journal of
environmental economics and management 105 (2021): 102399.

Johnston, Robert J., Kevin J. Boyle, Wiktor Adamowicz, Jeff Bennett, Roy Brouwer, Trudy Ann
Cameron, W. Michael Hanemann et al. "Contemporary guidance for stated preference
studies." Journal of the Association of Environmental and Resource Economists 4, no. 2
(2017): 319-405.

Kemp, Michael A., and Christopher Maxwell. "Exploring a budget context for contingent
valuation estimates." Contingent valuation: A critical assessment, Jerry Hausman, editor
(1993): 217-269.

Lewis, Lynne Y., Leslie Richardson, and John C. Whitehead. "Fat Tails, Flat Tails, and
Willingness to Pay: Kriström Revisited." Land Economics (2024).

Penn, Jerrod M., and Wuyang Hu. "Understanding hypothetical bias: An enhanced meta-
analysis." American Journal of Agricultural Economics 100, no. 4 (2018): 1186-1206.

Wallmo, Kristy, and Daniel K. Lew. "Public willingness to pay for recovering and downlisting
threatened and endangered marine species." Conservation Biology 26, no. 5 (2012): 830-
839.

Whitehead, John C. "Plausible responsiveness to scope in contingent valuation." Ecological

Economics 128 (2016): 17-22.

Whitehead, John C. "They doth protest too much, methinks: Reply to 'Reply to Whitehead'".
No. 24-04, Working Papers from Department of Economics, Appalachian State
University (2024).

Whitehead, John C., Louis Cornicelli and Gregory Howard, "Total Economic Valuation of Great
Lakes Recreational Fisheries: Attribute Non-attendance, Hypothetical Bias and
Insensitivity to Scope," No 24-10, Working Papers from Department of Economics,
Appalachian State University (2024).

Appendix B: Proposal for the book titled *The Contingent Valuation Method: A Modern Approach* (Petrolia, Haab and Whitehead)

Interest in the contingent valuation method (CVM) has been intense since the Exxon Valdez oil spill. Since then, there has been much attention focused on the validity and reliability of the CVM including such issues as hypothetical bias, the scope test, temporal embedding and the fat tails problem. This interest has peaked during two distinct time periods triggered by environmental disasters -- the Exxon Valdez oil spill and the BP/Deepwater Horizon oil spill. This peak in interest was primarily due to the contentious nature of the natural resource damage assessment process in the United States. In both cases, economists were hired by the "potentially responsible parties" with a primary purpose of disparaging the CVM to avoid paying additional damages due to lost passive use values. In this context, this book is an attempt to "set the record straight" and help motivate researchers to recognize that the CVM is no different methodologically than the less heavily scrutinized discrete choice experiment (DCE) valuation

approach. The book begins with several chapters that provide the necessary background for a reader unfamiliar with the issues. The second section of the book presents several case studies where application of the CVM had demonstrable problems but, with the benefit of hindsight and state-of-the-art methods, demonstrates how these problems were overcome. The next sections demonstrate the state-of-the-art approach to designing CVM surveys and data analysis. The final section applies these same techniques to DCE data to illustrate the convergence of a modern approach to the CVM and DCEs.

Two parallel trends in the CVM research literature motivate this book. Since 1979, the dominant paradigm in the contingent valuation method (CVM) literature has been that the single bound dichotomous choice question format is preferred in terms of incentive compatibility. This notion, along with the split sample scope test (i.e., the "more is better" test), was calcified by the NOAA Blue Ribbon Panel on CVM in 1993 (the NOAA Panel was formed in response to the CVM controversy following the Exxon Valdez oil spill). But, this question format provides only a limited amount of information which leads to low powered scope and other validity tests. The second trend is the discrete choice experiment literature. Following the Exxon Valdez oil spill, discrete choice experiments (aka, choice modelling, conjoint analysis) made their way into the environmental valuation literature as an alternative to the CVM. Discrete choice experiments are similar to the CVM but their proponents suggest that they are superior because the CVM values a bundled environmental good or service while a DCE can value individual attributes within this bundle.

Both of these strands of literature lead us to argue that the CVM and DCEs should not be considered as separate valuation methods. These are the same thing. Hence, a modern approach

to the contingent valuation method recognizes the similarities between the CVM and DCEs. In this book we will explore the range of approaches to stated preference valuation beginning with the single-bound CVM valuation question and ending with multiple-choice DCEs. Innovations to the CVM from the DCE literature include using repeated referendum questions and attribute non-attendance methods. The primary goal of the book is to move the CVM literature away from the recommendations of NOAA panel and towards a more modern approach.

There are five sections to the book. The first section provides an introduction to the issues with chapters that review the literature (and preview of the rest of the book), present the theory that establishes incentive compatibility of the single bound valuation question and its variants and provide a number of empirical examples illustrating the problems with single bound data in the CVM literature. The second section includes chapters that illustrate some of the problems with single bound CVM data. Each chapter provides an illustration when the CVM did not pass validity tests and the solutions found for these problems. The third and fourth sections contain six chapters in total that illustrate the gains from using repeated single bound valuation questions ("the modern approach"). The third section contains chapters that use data that estimates total economic value, including the most troublesome passive use value component. These chapters address estimating the implicit discount rate, scope effects, fat tails and hypothetical bias. The fourth section contains three chapters that use focus on use values (recreation was the first and most often application of the CVM). The last section concludes with chapters that illustrate that the issues raised by CVM critics are no different than those encountered in DCE studies. These chapters illustrate the similarities and differences between CVM and DCE studies. The book concludes with a chapter describing how these data should be used for policy analysis.