Title
Reply to Whitehead

Authors
Desvousges, William H
Mathews, Kristy E
Train, Kenneth E

Publication Date
2020-11-01

DOI
10.1016/j.ecolecon.2020.106769

Peer reviewed

Ecological Economics 177 (2020) 106769

Contents lists available at ScienceDirect

Ecological Economics

journal homepage: www.elsevier.com/locate/ecolecon

Commentary

Reply to Whitehead

William H. Desvousges[a,*], Kristy E. Mathews[b], Kenneth E. Train[c]

[a] W.H. Desvousges & Associates, Inc, St. Augustine, FL, United States of America
[b] Independent Consultant, Cary, NC, United States of America
[c] Adjunct Professor Emeritus of Economics and Public Policy, University of California – Berkeley, The Brattle Group, San Francisco, CA, United States of America

Whitehead (2020, "Whitehead") made numerous mistakes in his calculations, which we enumerate below. When these errors are corrected, adding-up fails for each theoretically valid parametric model that Whitehead used. Adding-up also fails with the Kristrom non-parametric estimator, which he had suggested in a previous critique of our paper (Whitehead, 2016a). Whitehead's analyses provide no instance, when his errors are corrected, of adding-up passing.

Whitehead argues that non-monotonicity, flat portions of the response curve, and fat tails are problems with our data. These problems are typical of contingent valuation (CV) data, and not unique to our study. As examples, we show below that Whitehead's past CV studies have the same problems, as does the Chapman et al. (2009) study that underlies ours. These data issues constitute another reason, along with failure of adding-up, to be concerned about the general unreliability of CV.

Whitehead made the following mistakes:

(1) His claim that adding-up passes for the log-linear model is incorrect. His estimated log-linear models have infinite mean willingness-to-pay (WTP), making it impossible to perform adding-up tests on the means, as well as rendering the models meaningless. He used medians for the tests instead of means, assuming – incorrectly – that the sum of medians is the median of the sum. The difference is especially pronounced for skewed distributions, like the log-normal distributions that he used for these models. To correct this error, we calculated the median of the sum of WTPs through simulation based on the estimated parameters of his log-linear model in Table 4, assuming independence across distributions. The median of the sum of the WTPs is $4904, which is 13 times larger than the $359 that he reported in his Table 5, and 24 times larger than his estimated median WTP for the whole – clearly a violation of adding-up. His claim that adding-up passes is based on his incorrect calculation of the median of a sum.

(2) Whitehead's claim that adding-up passes empirically for some linear models is incorrect. The linear models were estimated on data that covered, by definition, only positive WTPs. He assumed– against logic and without any data-– that a portion of the population has negative WTPs and found that when this assumption is imposed post-estimation, the adding-up test passes. He reports that, when this assumption is not imposed, adding-up fails with his linear models. This means that adding-up passed in his calculations on linear models not because of the data but because of his implausible additional assumption that many people have a negative WTP for the environmental programs.

(3) Whitehead's claim that adding-up passes when respondents with missing demographics are dropped is incorrect. In describing his Table 9, he states that "The 95% confidence intervals for these estimates overlap," and concludes – incorrectly – that adding-up passes. The t-statistic for the adding-up test is 2.44,[1] which is significant at the 95% level and indicates that adding-up fails (i.e., the whole is not equal to the sum of the parts.) His claim that adding-up passes is based on his mistaken use of overlapping confidence intervals.

(4) Whitehead's claim that weighted estimation passes the adding-up test is incorrect. His weighted model contains an incorrect sign for a cost coefficient,[2] and yet he mechanically calculates WTP and performs adding-up tests with this invalid model. WTP is undefined when the cost coefficient is positive and adding-up tests cannot be performed when WTP is undefined. We re-estimated his model with one cost coefficient, instead of separate coefficients for each increment; the estimated cost coefficient takes the correct sign and adding-up fails.

(5) Whitehead failed to report relevant findings that contradict his conclusions. In an earlier critique of our paper, Whitehead (2016a) used the non-parametric Kristrom estimator and claimed that adding-up passed with this estimator. On examination of his computer programs, we found that he had inadvertently dropped observations from his calculations. When these observations are included, the adding-up test fails with the Kristrom estimator. In his current Comment, he does not mention the Kristrom estimator

* Corresponding author.
E-mail address: william.desvousges@whdesvousgesassociates.com (W.H. Desvousges).

[1] Using the numbers from Table 9, the t-statistic is $\frac{1079-445}{SQRT(174^2+193^2)} = 2.44$.

[2] His estimates are in Table 6. (An alert to readers: Whitehead's Table 8 also shows an incorrect sign for one of the cost coefficients. However, this is a typo. The estimated slope coefficient for the second increment is −0.00425).

https://doi.org/10.1016/j.ecolecon.2020.106769
Received 22 May 2020; Received in revised form 15 June 2020; Accepted 22 June 2020

rather than report the correct results. It is useful for readers to know that adding-up fails under the non-parametric Kristrom estimator.

(6) Whitehead discussed non-monotonicity, flat areas of demand, and fat tails as problems in our dataset. As we have said in our papers, these problems are typical of CV studies, and not, as Whitehead suggests, unique to our data. All three of these problems are evident in the original study's data (Chapman et al., 2009), which was a multimillion-dollar study using extensive focus groups, pretesting, and in-person interviews. Also, we examined the papers authored or co-authored by Whitehead that are cited in the recent reviews by Burrows et al. (2017) and Parsons and Myers (2016).[3] These papers provide 15 CV datasets. Each of the three problems that Whitehead identified for our paper is evidenced in these datasets:
   - Non-monotonicity: 12 of the 15 datasets exhibit non-monotonicity.
   - Flat portions of the response curve: All 15 datasets have flat areas for at least half of the possible adjacent prompts, and 4 datasets have flat areas for all adjacent prompts.
   - Fat tails: In our data, the yes-share at the highest cost prompt ranged from 15 to 45%, depending on the program increment. In Whitehead's studies, the share ranged from 14 to 53%.

   If Whitehead's data are no worse than typical CV studies, then his papers indicate the pervasiveness of these problems in CV studies.

(7) Whitehead incorrectly claims that our procedure does not allow for substitution effects. Survey respondents were informed about past increments being already provided when asked about a new program. Any substitution induced by the prior increments is reflected in the respondents' answers about the new program. Readers can verify this by reading Desvousges et al. (2015, hereafter "DMT"). In his discussion of substitution, Whitehead describes his interpretation of our wording for the scenarios; and his own interpretation allows for substitution.

(8) Whitehead incorrectly states that we made the "implicit claim" that income effects are small. We tested for the impact of income effects and found them to be too small to affect the test results. This is an empirical finding, not an implicit claim. Again, readers can verify this in DMT.

(9) Whitehead reports that he obtained a significant income coefficient in some cases and incorrectly suggests that "[DMT] are using an inappropriate income coefficient for their income effect simulations." He did not re-run the simulations using his estimated income coefficients. We performed the simulations ourselves and found that the income effects from his estimated income coefficients are too small to affect the adding-up test.[4]

(10) Whitehead incorrectly claims that a one-tailed test is appropriate since substitution and income effects are not included. As stated above, substitution effects were included, and income effects were found empirically to be negligible. But it is important for readers to realize that Whitehead's proposed version of the test renders the test useless: any CV study that evidences inadequate response to scope will "pass" his one-tailed adding-up test.

(11) Whitehead incorrectly claims that the concept of scope elasticity, which he introduced in an earlier paper (Whitehead, 2016b), addresses the issue of inadequate scope. It does not. It simply changes the question of "Is the response to scope adequately large?" to "Is the scope elasticity adequately large?"

(12) Whitehead calculated standard errors for the Turnbull estimates using a formula from Haab and Haab and McConnell (2002) and reports: "the WTP estimates fail the adding-up test, replicating the result in [DMT]." However, even though the result is the same, Whitehead's calculation of the standard errors is incorrect. Haab and McConnell's formula is only applicable when monotonicity holds in all possible samples; it does not reflect the pooling that arises from non-monotonicities. The adding-up test in our paper was based on a bootstrap that accounts for the fact that different samples could (and do) result in different points of non-monotonicity.

Declaration of Competing Interest

None.

References

Burrows, James, Newman, Rebecca, Genser, Jerry, Plewes, Jeffrey, 2017. Do contingent valuation estimates of willingness to pay for non-use environmental goods pass the scope test with adequacy? In: McFadden, Daniel, Train, Kenneth (Eds.), Contingent Valuation of Environmental Goods: A Comprehensive Critique. Edward Elgar, Cheltenham UK, pp. 82–152.

Caudill, Steven, Groothius, Peter, Whitehead, John, 2011. The development and estimation of a latent choice multinomial logit model with application to contingent valuation. Am. J. Agric. Econ. 93 (4), 983–992.

Chapman, David, Bishop, Richard, Hanemann, Michael, Kanninen, Barbara, Krosnick, Jon, Morey, Edward, Tourangeau, Roger, 2009. Natural Resource Damages Associated with Aesthetic and Ecosystem Injuries to Oklahoma's Illinois River System and Tenkiller Lake. Expert Report for State of Oklahoma. vol. I available at. https://pcl.uscourts.gov/search (*Oklahoma v. Tyson Foods Inc.*, No. 4:05-cv-329 (N.D. Okla. Feb. 13, 2009), Docket No. 1853-4, exhibit D). The public-use copy from this site is also available at. http://elsa.berkeley.edu/~train/chapman.pdf.

Desvousges, William, Mathews, Kristy, Train, Kenneth, 2015. An adding up test on contingent valuations of river and lake quality. Land Econ. 91, 556–571.

Haab, Timothy, McConnell, Kenneth, 2002. Valuing Environmental and Natural Resources: The Econometrics of Non-market Valuation. Edward Elgar Publishing.

Huang, Ju-Chin, Haab, Timothy, Whitehead, John, 1997. Willingness to pay for quality improvements: should revealed and stated preference data be combined? J. Environ. Econ. Manag. 34 (3), 240–255.

Parsons, George, Myers, Kelley, 2016. Fat tails and truncated bids in contingent valuation: an application to an endangered shorebird species. Ecol. Econ. 129, 210–219.

Whitehead, John, 1992. Ex ante willingness to pay with supply and demand uncertainty: implications for valuing a sea turtle protection programme. Appl. Econ. 24 (9), 981–988.

Whitehead, John, 2002. Incentive incompatibility and starting-point bias in iterative valuation. Land Econ. 78 (2), 285–297.

Whitehead, John, 2016a. Comment on Desvousges, Mathews and Train (2015) continued: WTP estimates. Environ. Econ. 2016 26 September, available at. https://www.env-econ.net/2016/09/comment-on-desvousges-mathews-and-train-2015-continued-wtp-estimates-.html (accessed 5 June 2020).

Whitehead, John, 2016b. Plausible responsiveness to scope in contingent valuation. Ecol. Econ. 128, 17–22.

Whitehead, John, 2020. Comment on Desvousges, Mathews and Train (2015): an adding up test on contingent valuations of river and lake quality. Ecol. Econ (in this issue).

Whitehead, John, Cherry, Todd, 2007. Willingness to pay for a green energy program: a comparison of ex-ante and ex-post hypothetical bias mitigation approaches. Resour. Energy Econ. 29 (4), 247–261.

Whitehead, John, Finney, Suzanne, 2003. Willingness to pay for submerged maritime cultural resources. J. Cult. Econ. 27, 231–240.

Whitehead, John, Haab, Timothy, Huang, Ju-Chin, 1998. Part-whole bias in contingent valuation: will scope effects be detected with inexpensive survey methods? South. Econ. J. 65 (1), 160–168.

Whitehead, John, Groothius, Peter, Southwick, Rob, 2007. Linking recreation demand and willingness to pay with the inclusive value: Valuation of Saginaw Bay coastal marsh. In: Presented at the US Environmental Protection Agency Workshop (Valuation for Environmental Policy: Ecological Benefits Workshop).

Whitehead, John, Groothius, Peter, Southwick, Rob, Foster-Turley, Pat, 2009. Measuring the economic benefits of Saginaw Bay coastal marsh with revealed and stated preference methods. J. Great Lakes Res. 35 (3), 430–437.

[3] Specifically, Whitehead (1992), Huang et al. (1997), Whitehead et al. (1998), Whitehead (2002), Whitehead and Finney (2003), Whitehead and Cherry (2007), Whitehead et al. (2007), Whitehead et al. (2009), Caudill et al. (2011).

[4] In this section on income effects, Whitehead suggests that respondents may have a binding constraint for environmental goods. This suggestion violates utility maximization theory, and without utility maximization, WTP is not defined for welfare analysis.