Fat Tails, Flat Tails, and Willingness to Pay: Kriström Revisited

Lynne Y. Lewis, Leslie Richardson and John C. Whitehead

Corresponding Author

Lynne Y. Lewis
Professor, Department of Agricultural and Resource Economics
Colorado State University
And Elmer W. Campbell Professor of Economics, Emeritus
Bates College
lylewis@colostate.edu
207-939-5364

Co-authors

Leslie Richardson
Economist, National Park Service
1201 Oakridge Dr.
Fort Collins, CO 80525
leslie_a_richardson@nps.gov
970-821-5352

John C. Whitehead, Professor
Appalachian State University
Boone, NC 28608
whiteheadjc@appstate.edu
828-773-5049

doi:10.3368/le.100.4.060823-0052R

Abstract

This paper explores the limitations of two commonly used nonparametric approaches, the Turnbull estimator and the Kriström estimator. Fat tails and flat tails in the distribution of bid responses can result in a large divergence in willingness to pay estimates across different estimation methods. Since it is extremely rare to capture the tail of the distribution, we highlight the merit of an alternative approach using 120 data sets drawn from the literature. We calculate 4 non-parametric estimates of willingness to pay for each data set. Meta analysis is used to help determine when the Turnbull and Kriström estimators diverge.

Introduction

Willingness to pay (WTP) estimates can be extremely sensitive to the empirical model used (Bengochea-Morancho et al. 2005, Hannemann and Kanninen 1996, Kriström 1990). As a result, best practices for estimating WTP from stated preference survey data should include sensitivity analysis and the presentation of WTP values calculated using different methods (e.g. Bengochea-Morancha et al. 2005). While both parametric and nonparametric methods are commonly used, non-parametric methods are appealing since they are a less restrictive alternative to parametric models and do not rely on the analyst knowing or assuming the distribution of WTP. Most commonly used non-parametric approaches for discrete choice contingent valuation data are the Turnbull estimator (Haab and McConnell, 2002) and the Kriström estimator (Kriström; 1990; Boman, Bostedt, Kriström, 1999). Both of these approaches have limitations.

The Turnbull nonparametric estimator develops a step function with the proportion of the sample that indicated a willingness to pay (hereafter, "yes") the offered bid amount (WTP $\geq$ that amount) and truncates the distribution of the responses at the highest bid, regardless of how many said yes to the highest bid. Thus the Turnbull is a lower bound estimate of WTP (Haab and McConnell, 2002, Turnbull, 1976). The Kriström (1990) estimator uses linear interpolation to characterize the distribution between bid amounts and smooth the step function. This approach assumes that the distribution (i.e., survival) function is piecewise linear between bid amounts (Haab & McConnell, 2002). This approach is appealing in that it is not a lower bound and allows for the estimation of the right tail of the WTP distribution. The challenge, however, is knowing what assumption to make regarding the choke price, or the point at which the probability of a yes response falls to zero. This choke price is often calculated using either linear interpolation from the highest two bid amounts, which is the approach taken in Kriström (1990), or simply by truncating at the highest bid amount. These assumptions can have a significant effect on WTP estimates, especially with data sets for which the percentage of respondents saying yes to the highest bid amounts is relatively high. This problem has been termed "fat tails." Parsons and Myers (2016) find that 60% of contingent valuation studies from 1995 to 2014 exhibit bid functions with a yes-response rate of 20% or greater at the highest bid amount. They note,

> "Perhaps our most startling finding is the sensitivity of mean willingness to pay to the largest bid. This is because so much of the willingness to pay is captured in the high-end tail of the yes-response function (or demand function over high prices). One can easily double or triple a mean willingness to pay by simply picking a larger bid. This lack of robustness is troubling."

Fat tails can result in a large divergence in WTP estimates across different estimation methods, including the two commonly used non-parametric approaches. Compounding this problem is the existence of "flat" tails resulting from a slowly declining yes-response rate at the highest bid amounts. While the Turnbull estimator is a consistent estimate of the lower bound on expected WTP, the Kriström estimator can result in an upper bound on WTP when the choke price is calculated based on a flat slope between the two highest bid amounts. The combination of fat and flat tails becomes especially problematic.[i] This raises questions about which estimate, if any, to use for benefit-cost analyses and other policy applications where the magnitude of WTP matters.

This paper explores the limitations of these commonly used approaches, and in particular, assumptions about the choke price. Since it would be extremely rare, if not impossible, to capture the tail of the distribution exactly (the bid at which 0% of the respondents say yes), it is important to explore whether or not an alternative approach is merited. Table 1 of Parsons and Myers lists 86 contingent valuation WTP studies published in seven field journals in agricultural, environmental and resource economics from 1995 to 2014. We reconstruct as many of these datasets as possible and review the literature to incorporate additional datasets from studies published through 2023. Using a total of 51 articles comprising 120 datasets, we calculate WTP based on the Turnbull estimator, the Kriström estimator (1990), the Kriström estimator truncated at the highest bid amount, and explore the merits of an alternative approach, the Richardson and Lewis (2022) Kriström adjustment. For each dataset we calculate ratios of the various WTP estimates to demonstrate the magnitude of differences across the estimators and evaluate cases where there is a large divergence. We then use meta-analysis to further evaluate the factors that influence the divergence of WTP estimates across these non-parametric approaches.

Our findings demonstrate that, in many cases, the Richardson and Lewis (2022) Kriström adjustment is a preferable measure of WTP, relative to the Kriström estimator. In the presence of flat tails, and especially both fat and flat tails, the adjusted Kriström estimate reduces the range between the Turnbull and Kriström WTP estimates, resulting in a superior estimate for benefit-cost analysis.

Nature of the Problem

The Kriström (1990) nonparametric estimator assumes a hundred percent yes response at a zero bid, linear interpolation to estimate the percentage yes between bids, and linear interpolation between the highest two bid amounts to estimate the choke price. This last procedure can result in an extremely high choke price when the dataset exhibits a flat tail (small differences in percentage yes responses at the highest two bid amounts). Richardson and Lewis (2022) present a combined approach for the calculation of the choke price. Using a large and very well-behaved data set (see Figure 1 - the percentage of yes responses decreases as the bid amount increases and very few respondents said yes to the highest bid amount) they utilize four different methods to calculate the choke price. First, they follow the standard Kriström approach of simply calculating the slope based on data from the two highest bid amounts, at which point the slope is very flat and results in a high choke price. They then use a linear probability model to estimate a choke price, using the slope over the entire range of the bid curve (as suggested by Whitehead, 2017). Interestingly, this method results in a choke price that is lower than their highest bid. The third option is to simply truncate the Kriström estimate at the highest bid amount. These methods result in choke prices of $2,300, $450, and $500, respectively. In addition to those approaches, they present a unique, slightly modified Kriström approach. They use the slope from the linear probability model but extend it from the highest bid amount to calculate a choke price. By utilizing additional information from the bid responses, this results in a choke price of $567. This choke price is significantly smaller than one based on simply calculating the slope using responses to the two highest bid amounts, at which point the slope is flat and results in a choke price nearly five times the highest bid amount (despite the fact that only 6% of respondents said yes to the highest bid amount).

Figure 1. Percentage of yes responses to each bid amount from Richardson and Lewis (2022)

This adjustment (referred to hereafter as the Kriström adjustment (KA)) results in a significantly different WTP estimate compared to that calculated with the standard Kriström ($84 vs. $140), but does not exhibit a significant difference from the WTP calculated with a truncated Kriström ($84 vs. $82), simply because so few respondents said yes to the highest bid. But it begs the question: at what point might it make a difference?

To illustrate the problem first consider some textbook dichotomous choice data (Figure 2). We construct a data set from question number 3 from Boardman et al.'s (2015) contingent valuation chapter (p. 398). Students are told to "consider a project that would involve purchasing marginal farmland that would then be allowed to return to wetlands capable of supporting migrant birds. Researchers designed a survey to implement the dichotomous choice method. They reported the following data." In the data table there are ten bids that range from $0 to $50 and the percentage of those who are willing to pay each bid falls from 98% to 2%. Students are asked "What is the mean WTP for the sampled population?" With these well-behaved data students could take a number of approaches and provide an acceptable answer to their professor.

Figure 2. Boardman et al. (2017) textbook data

The WTP estimates are $18 and $21 with the Turnbull and Kriström estimators, a 13% difference (without rounding).[ii] Estimation of WTP with these data is robust to discarding high bid amounts, low bid amounts, and both high and low bid amounts with the Kriström (and parametric) estimators. The Turnbull WTP estimates are more sensitive to bid design, especially when discarding higher bids due to the truncation of the distribution.

Unfortunately, data collected from real samples of the population are not so well-behaved. Problems have been found with negative WTP estimates, non-monotonicity of bid functions, fat tails, and flat tails. Negative WTP estimates will result when the predicted probability of a yes/for response is less than 50% at a bid amount of zero if logit models are used for estimation (Hanemann, 1984; Haab and McConnell, 1997). The Turnbull and Kriström models provide more likely estimates of mean WTP as they are constrained to be positive.[iii]

Non-monotonicity of bid functions results when the probability of being willing to pay for the policy rises when the bid amount rises. Haab and McConnell (2002) call this "difficult data." This violation of rational choice theory can be the result of small random samples of the population. The Turnbull and Kriström nonparametric approaches handle this by pooling bid amounts and yes/no responses until the bid function is monotonically decreasing over pooled bid amounts. Non-monotonicity can lead to flat bid functions and upward biased mean WTP estimates when Kriström choke prices are estimated by the slope of the bid function.[iv]

The fat tails problem exists when the probability of a yes response is relatively high, say 20% or more at the highest bid amount as defined by Parsons and Myers (2016). Fat tails leaves the researcher uncertain about a potentially large portion of the WTP distribution. They note,

> "Truncating high-end bids is a tempting response to fat tails. If the tail of the yes-response surface is ignored over its high end, the analyst may offer truncated values using a lower-bound nonparametric estimator as a conservative value. But, this is not a real fix to the underlying problem of hypothetical bias, nor is the resulting willingness to pay truly conservative. Indeed, it 'hides' the effects of fat tails. One may falsely believe that he or she has a reasonable estimate of value when in fact the survey instrument could produce vastly different values with only modest changes in the bid levels offered. Truncating offers nothing new for understanding underlying preferences, explaining why contingent valuation data yield fat tails, or dealing with hypothetical bias." (Parsons and Myers 2016).

The Turnbull estimator truncates willingness to pay at the highest bid leading to estimates that are biased downwards. These biased downward WTP amounts may be appropriate for natural resource damage assessments where a conservative estimate is easier to defend in court, but little else.

As illustrated by the results from the textbook data above, fat tails combined with a sufficient negative slope of the bid function will not lead to upward biased WTP estimates (Figure 3). However, flat tails could lead to a wide range between the lower bound WTP estimate produced by the Turnbull and the upper bound WTP estimate produced by the Kriström estimator.

When the fat and flat tails problems are combined, problems worsen. For example, if flat bid functions are due to non-monotonicities or yes responses that are insensitive to higher bids, then the Kriström choke price can become implausibly large (Figure 3). This can lead to wide ranges between the lower bound WTP estimate produced by the Turnbull and the upper bound WTP estimate produced by the Kriström estimator. The question then becomes, which number should be used in benefit-cost analysis? Or should there be an alternative?

Figure 3. Fat and Flat Tails

Data

To thoroughly review the literature, we collected the summary data from all of the papers cited in Parsons and Myers (2016). Their review of seven journals[v] identified 86 contingent valuation studies that reported the percentage of yes responses at the highest bid amount. Approximately sixty percent of these studies exhibit fat tails (as defined by Parsons and Myers as a yes-response rate of 20% or greater at the highest bid) in at least one scenario. After reviewing these articles, we find that 46 provide the necessary information to reconstruct the data (e.g., the number of yes/no responses at all bid amounts). Appendix Table A1 describes the reasons why one cannot reconstruct the data in the other 40 articles. The primary reason is that the crosstab table of yes/no responses by bid amount presented in the article does not contain the sample sizes at each bid amount.

Table 1 in Parsons and Myers (2016) contains the magnitude of the height of the tail of the distribution (i.e., percentage yes at the highest bid amount) and a range when the experimental design of the article has multiple treatments. In order to compare the height of the tail in the usable and excluded articles, we compute the midpoint of the range for those articles with a range presented. The average height of the tail is 23% for the 46 articles that are usable and 29% for those 40 articles that are excluded from our analysis. The t-statistic for the difference in the height of the tails is t=1.92 (p=0.03). We conclude that studies that thoroughly report their data and are included in our analysis will have a relatively less severe fat tails problem. We can reach no conclusions about the flat tails problem when comparing the two samples.

In addition to these studies, we searched the same set of journals for articles published through 2023 and found five additional articles that contain the necessary information to reconstruct the data. The average height of the tail in the six data sets in these five articles is 27%. The total number of studies for our analysis is n=51.

The data summary for studies included in our analysis (n=51) is presented in Appendix Table A2. The articles were published between 1990 and 2022 with all but three published between 1995 and 2018. Appendix Figure A1 illustrates the frequency of studies over the time period we examined. Sixty-one percent (31) of the studies are U.S. based, and an additional 5 studies are based in Sweden, 3 in Spain, 2 in England, and 1 each in Australia, Austria, China, Ireland, Kuwait, Mexico, the Philippines, Taiwan, Uruguay, and Vietnam. Twenty-two percent of the articles use a donation or voluntary contributions payment vehicle. There are five survey modes represented in the sample with the percentages adding up to more than one due to mixed modes being used in three studies. Forty-seven percent of the studies used a mail survey contact mode, 25% used an in-person contact mode, 14% are laboratory experiments (with student samples), 14% use the telephone survey mode, and 6% are online surveys. Seventy-one percent of the studies are valuing public goods, relative to private goods. Fifty-three percent have one-time payment schedules. The average number of years in each payment schedule is 8 with a range of 1 (for one-time payments) to 30, where in perpetuity payment schedules are coded as 30.

Of these 51 articles, 21 have one data set and the remainder have between 2 and 9 data sets. Twelve articles have 2 data sets, 10 articles have 3 data sets, 4 articles have 4 data sets, 2 articles have 6 data sets, 1 article has 8 data sets and another has 9 data sets. In total, there are 120 data sets available for analysis. In those articles that present multiple data sets, the source could be an experimental treatment or samples of different populations. The experimental treatments include tests for hypothetical bias, sensitivity to scope, payment schedules, payment vehicles, visual aids, and ordering effects.

A summary of the data at the data set level (n=120) is presented in Appendix Table A3. The mean of the sample size variable is 433 with a range of 47 to 4361. The average number of bid amounts presented to respondents is 7 with a range of 3 to 21. The mean of the sample size per bid amount variable is 71 with a range of 7 to 396. The bid amounts are left in the home countries currency and are not adjusted for inflation, so the bid amounts themselves contain a limited amount of information. In order to make the bid amounts comparable, we divide each bid by the maximum bid amount so that the standardized bid amounts can range from zero to one. The mean of the minimum standardized bid is 0.10. The two bids that form the slope for the tail of the Kriström WTP estimator are the two highest bids with and without pooling of bids for non-monotonicity. Forty-four percent of the data sets have pooled bids for one of the bid amounts used to calculate this slope. The standardized lowest bid in the slope (Sbid1) is 0.56 and the standardized highest bid for the slope (Sbid2) is 0.88. The average percentage yes response at Sbid1 (Pctyes1) is 35% and the average percentage yes response at Sbid2 (Pctyes2) is 23%. The absolute value of the Kriström slope with the standardized bids is 0.48 with a range of 0.01 to 4.02.

Results

Using the re-created yes/no data for our final sample, we are able to calculate WTP using the Kriström estimator (1990), the Kriström adjustment (Richardson and Lewis, 2022), the Kriström with truncation at the highest bid, and the Turnbull (See Appendix Table A4 for all results).[vi] Table 1 highlights summary statistics for the ratios of these estimates. The mean of the Kriström to Turnbull ratio (K/T) is 2.14 with a range of 1.04 to 36.59. This suggests that the high end of the range of the nonparametric WTP estimates is 114% greater than the Turnbull WTP. The mean of the Kriström adjustment to Turnbull ratio (KA/T) is 1.66 with a range of 1.09 to 11.79. The high end of the range is a more reasonable 66% higher than the Turnbull. The mean of the truncated Kriström to Turnbull ratio (TK/T) is 1.24 with a range of 1.02 to 3.32. This indicates that the average amount of the K/T and KA/T ratios that are contained in the Kriström triangles above the Turnbull step function is 24% of the Turnbull WTP estimate. Considering this, 90% of the K/T ratio and 42% of the KA/T ratio is due to the tail of the Kriström distribution.

The data contain two outliers with K/T ratios greater than 28. Removing these outliers leaves a mean of the Kriström to Turnbull ratio (K/T) of 1.62 (Table 1). The mean of the Kriström adjustment to Turnbull ratio (KA/T) is 1.57. The mean of the truncated Kriström to Turnbull ratio (TK/T) is 1.24. Considering the data with the outliers removed, 58% of the data sets have Kriström adjustment WTP estimates less than the Kriström WTP and the K/KA ratio is 1.17 for this subsample (n=68). Of course, outlying K/T ratios are situations where an adjustment to the Kriström estimator is most beneficial. But, we find this is rare in our data (recall that our sample has smaller tails than the full data).

We next consider the factors that affect the K/T ratio. We first begin with the variables which must be related to the ratio by construction, i.e. the measures of fat tails (Pctyes2) and flat tails (Slope). The ratio is expected to increase as the percentage yes response at the highest bid increases, and decrease as the absolute value of the Kriström slope increases (gets steeper). The models are estimated with standard errors clustered at the study level.

The results of these models are presented in Table 2. Each of the variables have the expected sign and are statistically significant. The model fit improves from the linear to double-log model with the double-log functional form explaining more than 3 times the variation in the K/T ratio in the model that includes outliers and 1.5 times in the model that excludes outliers.

The simple linear models can be used to simulate the effects on the K/T ratio of changes in the height and slope of the tail of the Kriström estimator. In Figure 4 we show that the K/T ratio increases from 1.28 when the height of the tail is 0.10 to 3.78 when the height of the tail is 0.50, holding the slope of the tail constant at 0.50. Removing outliers (n=118), the range is a K/T ratio of 1.34 when the height of the tail is 0 and 1.95 when it is 0.50.

In Figure 5 we show that the K/T ratio decreases from 2.74 when the slope of the tail is 0.01 to 1.17 when the slope of the tail is 1.5, holding the height of the tail constant at 0.25. Removing outliers (n=118), the K/T ratio decreases from 1.83 when the height of the tail is 0 to 1.27 when it is 1.50. The double-log model coefficients from Table 2 can be interpreted as elasticities. A 10% increase in Pctyes2 leads to a 13% increase in the K/T ratio. A 10% decrease in the slope leads to a 7% increase in the K/T ratio. Removing outliers (n=118), a 10% increase in Pctyes2 leads to a 7% increase in the K/T ratio. A 10% decrease in the slope leads to a 52% increase in the K/T ratio.

In models 3 and 6 we include characteristics of each of the studies using the data with outliers removed and clustered standard errors. We find that the effects of fat and flat tails are similar to models 2 and 4 when including these controls. In both models, the ratio is larger when donations are used as the payment vehicle.

Figure 4. K/T Ratio Holding the Slope of the Tail Constant at 0.5

Figure 5. K/T Ratio Holding the Height of the Tail Constant at 0.25

Next, we consider similar models where the dependent variable is the Kriström to Kriström adjustment ratio with the sample where outliers are removed (Table 3). In this analysis we find that the Kriström adjustment WTP is much closer to the Kriström WTP when outliers are removed. We estimate four models considering both functional forms and with and without controls. We find that the height of the tail is not a factor in the K/KA ratio. But, as the absolute value of the Kriström slope increases (gets steeper) the Kriström WTP estimate decreases relative to the Kriström adjustment WTP. Setting the slope equal to zero, the simple linear model suggests that the Kriström adjustment will reduce the ratio by 13%. When the Kriström slope rises to 0.75 the difference between the Kriström and Kriström adjustment WTP is zero. The elasticity in the double log model says that a 10% decrease in the slope leads to a 3.5% increase in the ratio. In both models, other factors that affect this difference are the year of the study, lab data, and whether a public good is being valued. The ratio gets smaller as publication year gets closer to the present. Experimental lab data and public goods reduces the ratio. A donations payment vehicle increases the ratio in the linear model.

For the majority of the data sets in our sample (58%), the WTP estimate using the Kriström adjustment falls between the lower-bound Turnbull (and truncated Kriström) and the traditional Kriström (1990) estimator. The exceptions are those where the slope between the highest two bids is *steeper* than the slope over the entire function. If the absolute value of the Kriström slope is smaller (i.e., flatter) than the slope over the entire bid range, then the Kriström adjustment estimate is less than the Kriström estimate, and vice versa.

Table 4 illustrates the range of K/T, KA/T, and K/KA ratios for a sample of our data sets. We select illustrative data for four specific cases: datasets that exhibit both fat *and* flat tails; datasets that exhibit flat, but not fat tails; datasets that exhibit fat, but not flat tails; and datasets that have neither flat nor fat tails. In each category we present four examples, as well as the two outliers in the fat and flat tails category. Figure 6 illustrates the bid functions using an example from each of these four cases. As can be seen, survey data is imperfect and bid functions often do not come close to capturing the tail of the distribution. This leads to several interesting implications regarding the divergence of non-parametric estimators. First, if a dataset exhibits fat, but not flat tails (e.g., NahuelhualA in Figure 6), the Kriström and Kriström adjustment will yield similar results, whereas the Turnbull (or truncation at the highest bid) will result in a significantly lower WTP estimate. If, however, a dataset exhibits flat, but not fat tails (e.g., IvehammerC in Figure 6), the Kriström adjustment and truncation at the highest bid will typically yield similar results, yet the adjustment produces what could be viewed as a more reasonable choke price (since there are still yes responses to the highest bid). In these cases, the traditional Kriström (1990) may result in a much higher estimate of WTP based on a high choke price relative to the Turnbull estimate of WTP. Finally, for datasets that exhibit both flat and fat tails, there will frequently be a large divergence between the Turnbull and the traditional Kriström WTP estimates. WTP estimates based on the latter are highly dependent on the slope between the two highest bid amounts. In these especially problematic cases, the Kriström adjustment provides a useful measure of WTP between the Turnbull and Kriström WTP estimates, providing a choke price that incorporates information over the entire bid function and avoiding the loss of information resulting from truncation.[vii]

Figure 6. Percentage yes responses at bid amounts for selected studies

In general, survey data is nuanced and there are various reasons a researcher might see large differences across non-parametric (and parametric) estimators. While the Turnbull provides a consistent lower bound estimate of expected WTP, the traditional Kriström (1990) is frequently biased upwards. Importantly, the Kriström estimate is extremely sensitive to the percentage yes response to the two highest bid amounts. The slope between the last two bids becomes especially arbitrary when data needs to be pooled due to non-monotonicity in the bid function. For example, in the dataset Richardson22 from Appendix Table A4, the exact same percentage of respondents said yes to the highest and second highest bid amounts (6% at both $350 and $500). The traditional Kriström (1990) estimate is forced to rely on the third highest bid amount to calculate a choke price, resulting in a relatively steep slope and similar WTP estimates across the approaches (Kriström choke price = $612; Kriström adjustment choke price = $567; K/T = 1.22; KA/T = 1.20; K/KA = 1.02). However, if just one less respondent had said yes to the highest bid amount, the Kriström estimate would result in a choke price more than seven times higher than the highest bid, resulting in a large divergence between the estimates (Kriström choke price = $3,650; Kriström adjustment choke price = $564; K/T = 2.56; KA/T = 1.20; K/KA = 2.13). By incorporating additional information over the entire bid function to estimate a more realistic choke price, the Kriström adjustment in many cases produces a better estimate of willingness to pay, especially in those cases of both fat and flat tails.

Discussion

For willingness to pay studies, Haab and McConnell (2002) note: “The set of offered bids should be designed to ensure that the tails of the distribution are well defined. Undefined tails can lead to unreliable measures of central tendency of WTP.” But the criteria for valid measurement of WTP is that “estimation and calculation are accomplished with no arbitrary truncation” (Haab and McConnell, 2002; 1998). However, in much of the literature we see intentional bid truncation or somewhat arbitrary choke prices. If this is the case, as Parsons and Myers (2016) point out, “binary choice models smooth out the tails in a statistically satisfying way but do so by censoring response data over the very range where we would like to know more about true behavioral response.”

How much does this matter? It turns out that it matters a lot in certain situations which we have illustrated in this paper. “Real” data is riddled with complexity and many data sets exhibit the types of problems that we have discussed here. In these cases, the researcher often turns to the Turnbull estimator since it avoids negative (parametric) willingness to pay, is a lower bound, and may be convenient for policy applications such as natural resource damage assessments. We question this common practice.

One way to think about this question is in the context of sensitivity analysis. There are a number of ways to conduct sensitivity analysis with the lower and upper bounds of a distribution, but each assumes that the midpoint of the distribution is the best estimate of central tendency (or imposes ad-hoc assumptions). For example, a researcher could assume that the true value is drawn from a uniform distribution between the lower and upper bounds. The mean of a uniform distribution is the midpoint between the lower and upper bounds. If this midpoint is considered a more likely measure of central tendency than any randomly drawn value between the bounds, then a normal or triangular distribution could be assumed. Both of these will produce a greater mass of drawn values near the midpoint.

These midpoint estimates could be used as the base case WTP estimate in a benefit cost analysis. But, it is not clear if the midpoint is the best estimate of the central tendency of the distribution bounded by the Turnbull and Kriström estimators. The greater the fat tail, the lower the Turnbull WTP estimate and greater the chance that midpoint WTP is downward biased. In the presence of both fat and flat tails, we see the greatest discrepancy between the Turnbull and Kriström estimators, and in these cases, the Kriström adjustment may provide a better estimate of WTP. With neither fat nor flat tails, the range between the Kriström and Turnbull narrows considerably and the Kriström adjustment and the Kriström estimators will converge. If nonparametric estimators are used for benefit-cost analysis, then an approach for developing a WTP measure that is better than the midpoint is needed. With a wide range between the Turnbull and Kriström (1990) estimators, policy analysts need a more accurate point estimate of WTP. While the current candidate is the truncated Kriström, this estimate is very close to the Turnbull, which is biased downwards. In this paper we have proposed another measure of WTP, a "Kriström adjustment." By incorporating additional information from responses over the entire bid range to calculate the choke price, this approach may result in a more accurate measure of WTP for many data sets.

In the presence of both fat and flat tails, there can be a large difference between the Turnbull and Kriström estimates. In these cases, the Kriström adjustment reduces the range and hence the uncertainty over WTP estimates. Reductions in uncertainty are helpful for benefit-cost analysis. Decreasing the range of uncertainty using the Kriström adjustment decreases the uncertainty for the decision-maker.

An obvious criticism of our recommendation is that it is ad-hoc. However, it is less ad-hoc than assuming the midpoint between Turnbull and Kriström WTP estimates and can easily be adjusted for irregularities in different data sets. And it is certainly less sensitive than the Kriström to the highest two bid amounts. Given the frequency of WTP data exhibiting fat or flat tails (or both), we thus suggest consideration of this Kriström adjustment for benefit-cost analysis and use of the Turnbull and Kriström (lower and upper bounds) for sensitivity analysis. Using the adjustment to the original Kriström estimator, we think we offer an improvement to the estimation of willingness to pay for benefit cost analysis.

In terms of future research, continued effort should be directed towards identifying the tail of the distribution in dichotomous choice willingness to pay data. Deep-pocketed researchers could better pretest bid designs. Researchers could also consider more extensive sensitivity analyses and alternative strategies to those presented here.

Acknowledgements

The authors would like to thank Heidi Nydam, Cerise Stanley and Drew Williams, students at Bates College, for their valuable research assistance. We would also like to thank participants in the W4133 Nonmarket Valuation workshop 2023 for valuable feedback on this paper. We also thank two anonymous reviewers for very helpful comments on this paper.

Disclaimer

Views and conclusions in this article are those of the authors and do not necessarily reflect the opinions or policies of the National Park Service.

References

Alolayan, Mohammad A., John S. Evans, and James K. Hammitt. 2017. "Valuing mortality risk in Kuwait: Stated-preference with a new consistency test." *Environmental and Resource Economics* 66: 629-646.

Azevedo, João Pedro. 2010. TURNBULL: Stata module to estimate the Turnbull empirical distribution estimator of willingness to pay," Statistical Software Components S457125, Boston College Department of Economics.

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

Bengochea-Morancho, Aurelia, Ana Ma Fuertes-Eugenio, and Salvador del Saz-Salazar. 2005. "A comparison of empirical models used to infer the willingness to pay in contingent valuation." *Empirical Economics* 30, no. 1: 235-244.

Berrens, Robert P., Philip Ganderton, and Carol L. Silva. 1996. "Valuing the protection of minimum instream flows in New Mexico." *Journal of Agricultural and Resource Economics*: 294-308.

Bishop, Richard C., Kevin J. Boyle, Richard T. Carson, David Chapman, W. Michael Hanemann, Barbara Kanninen, Raymond J. Kopp et al. 2017."Putting a value on injuries to natural assets: The BP oil spill." *Science* 356, no. 6335: 253-254.

Boardman, Anthony E., David H. Greenberg, Aidan R. Vining, and David L. Weimer. 2017. Cost-benefit analysis: concepts and practice. Cambridge University Press.

Boman, Mattias, Göran Bostedt, and Bengt Kriström. 1999. "Obtaining welfare bounds in discrete-response valuation studies: A non-parametric approach." *Land Economics* 284-294.

Champ, Patricia A., Nicholas E. Flores, Thomas C. Brown, and James Chivers. 2002. "Contingent valuation and incentives." *Land Economics* 78, no. 4:: 591-604.

Desvousges, W., Mathews, K., & Train, K. 2015. An adding-up test on contingent valuations of river and lake quality. *Land Economics*, *91*(3), 556-571.

Egan, Kevin J., Jay R. Corrigan, and Daryl F. Dwyer. 2015. "Three reasons to use annual payments in contingent valuation surveys: Convergent validity, discount rates, and mental accounting." *Journal of Environmental Economics and Management* 72: 123-136.

Haab, Timothy C., and Kenneth E. McConnell. 1997. “"Referendum models and negative willingness to pay: alternative solutions." *Journal of environmental economics and management* 32, no. 2: 251-270.

Haab, Timothy C., and Kenneth E. McConnell.1998. "Referendum models and economic values: theoretical, intuitive, and practical bounds on willingness to pay." *Land Economics*: 216-229.

Haab, Timothy C., and Kenneth E. McConnell. 2002. *Valuing environmental and natural resources: the econometrics of non-market valuation*. Edward Elgar Publishing.

Hanemann, W. Michael. 1984. "Welfare evaluations in contingent valuation experiments with discrete responses." *American Journal of Agricultural Economics* 66, no. 3 : 332-341.

Kriström, Bengt. 1990. "A non-parametric approach to the estimation of welfare measures in discrete response valuation studies." *Land Economics* 66, no. 2: 135-139.

Labao, Rex, Herminia Francisco, Dieldre Harder, and Florence Ivy Santos. 2008. "Do colored photographs affect willingness to pay responses for endangered species conservation?." *Environmental and Resource Economics* 40, no. 2: 251-264.

Mwebaze, Paul, Jeff Bennett, Nigel W. Beebe, Gregor J. Devine, and Paul De Barro. 2018."Economic valuation of the threat posed by the establishment of the Asian Tiger Mosquito in Australia." *Environmental and Resource Economics* 71: 357-379.

Parsons, George, and Kelley Myers. 2016. "Fat tails and truncated bids in contingent valuation: An application to an endangered shorebird species." *Ecological Economics* 129: 210-219.

Petrolia, Daniel R., and Tae-Goun Kim. 2009. "What are barrier islands worth? Estimates of willingness to pay for restoration." *Marine Resource Economics* 24, no. 2: 131-146.

Ready, Richard C., and Dayuan Hu. "Statistical approaches to the fat tail problem for dichotomous choice contingent valuation." Land Economics (1995): 491-499.

Richardson, Leslie, and Lynne Lewis. 2022. "Getting to know you: individual animals, wildlife webcams, and willingness to pay for brown bear preservation." *American Journal of Agricultural Economics* 104, no. 2: 673-692.

Turnbull, Bruce W. 1976. "The empirical distribution function with arbitrarily grouped, censored and truncated data." *Journal of the Royal Statistical Society: Series B (Methodological)* 38, no. 3: 290-295.

Whitehead, John C. 2017."Who knows what willingness to pay lurks in the hearts of men? A rejoinder to Egan, Corrigan, and Dwyer." Econ Journal Watch 14, no. 3: 346.

Whitehead, John C., William P. Anderson, Jr, Dennis Guignet, Craig Landry and O. Ashton Morgan, "Sea-Level Rise, Drinking Water Quality and the Economic Value of Coastal Tourism in North Carolina," Working paper number 24-01, Department of Economics, Appalachian State University.

Table 1. Kriström to Turnbull WTP Ratios

	Full Data (n=120)			
Variable	Mean	Std. Dev.	Minimum	Maximum
Kriström/Turnbull	2.14	4.05	1.04	36.59
Kriström Adjustment/Turnbull	1.66	1.14	1.09	11.79
Truncated Kriström/Turnbull	1.24	0.29	1.02	3.32
	Outliers Removed (n=118)			
Variable	Mean	Std. Dev.	Minimum	Maximum
Kriström/Turnbull	1.62	0.60	1.04	4.47
Kriström Adjustment/Turnbull	1.57	0.66	1.09	6.98
Truncated Kriström/Turnbull	1.24	0.29	1.02	3.32

Table 2. Determinants of K/T Ratio

	K/T Ratio			ln(K/T Ratio)		
		Outliers removed			Outliers removed	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	1.178	1.524	-2.582	0.481	0.481	-2.550
	(2.01)	(15.89)	(-0.15)	(6.03)	(9.10)	(-0.30)
PCTYES2	6.257	1.212	1.544			
	(1.67)	(3.12)	(3.54)			
SLOPE	-1.045	-0.369	-0.406			
	(-1.65)	(-2.85)	(-4.05)			
LN(PCTYES2+1)				1.256	0.717	0.900
				(2.72)	(2.96)	(3.51)
LN(SLOPE+1)				-0.700	-0.520	-0.526
				(-3.76)	(-4.78)	(-5.66)
N_BIDS			0.000			-0.000
			(-0.75)			(-1.03)
YEAR			0.002			0.0015
			(0.23)			(0.35)
US			-0.100			-0.042
			(-0.93)			(-0.73)
DONATION			0.429			0.163
			(3.09)			(2.54)
LAB			0.068			0.069
			(0.31)			(0.63)
ONLINE			0.142			0.070
			(0.70)			(0.65)
MAIL			0.171			0.081
			(1.61)			(1.52)
INPERSON			0.137			0.070
			(1.14)			(1.08)
ONETIME			0.140			0.033
			(1.22)			(0.53)
YEARS			-0.005			-0.003
			(-0.98)			(-1.07)
PUBLIC			-0.114			-0.048
			(-1.03)			(-0.85)
R^2	0.091	0.229	0.334	0.294	0.363	0.446
Sample size	120	118	118	120	118	118

Note: t-statistic in parentheses

Table 3. Determinants of K/KA Ratio

	K/KA Ratio		Ln(K/KA Ratio)	
	Model 1	Model 2	Model 3	Model 4
Constant	1.126	15.106	0.191	12.763
	(23.58)	(1.81)	(3.52)	(2.21)
PCTYES2	0.061	0.252		
	(0.34)	(0.87)		
SLOPE	-0.164	-0.167		
	(-1.97)	(-2.29)		
LN(PCTYES2+1)			-0.193	-0.021
			(-1.07)	(-0.09)
LN(SLOPE+1)			-0.348	-0.348
			(-2.73)	(-2.76)
N_BIDS		0.000		0.000
		(0.00)		(-0.36)
YEAR		-0.007		-0.006
		(-1.67)		(-2.18)
US		-0.007		0.011
		(-0.20)		(0.37)
DONATION		0.121		0.060
		(1.97)		(1.27)
LAB		-0.163		-0.098
		(-2.07)		(-1.65)
ONLINE		0.056		0.079
		(0.63)		(1.07)
MAIL		-0.002		-0.004
		(-0.03)		(-0.09)
INPERSON		0.043		0.037
		(0.52)		(0.58)
ONETIME		0.011		-0.022
		(0.24)		(-0.53)
YEARS		-0.004		-0.003
		(-1.29)		(-1.35)
PUBLIC		-0.120		-0.072
		(-1.94)		(-1.76)
R^2	0.071	0.140	0.158	0.214
Sample size	118	118	118	118

Table 4. Non-parametric WTP estimates and ratios for a sample of studies

Reference	% yes at second highest bid	% yes at highest bid	Kriström (1990)	Kriström Adjustment	Kriström Truncated	Turnbull	Ratio K/T	Ratio KA/T	Ratio K/KA
Fat and flat tails									
AlolayanA	39%	38%	50,828	2,011	1,501	1,389	36.59	1.45	25.28
Brown03A	70%	69%	160	66	6	6	28.41	11.79	2.41
GerkingD	45%	42%	4,846	1,874	1,282	1,182	4.10	1.59	2.59
HammitC	35%	33%	422	121	101	94	4.47	1.28	3.48
LongoD	47%	42%	234	137	99	91	2.57	1.50	1.71
WhiteheadA	52%	50%	201	94	70	65	3.10	1.45	2.14
Flat, not fat tails									
HammitG	18%	15%	354	280	247	210	1.69	1.33	1.26
IvehammarC	12%	8%	941	708	671	569	1.65	1.24	1.33
PetroliaC	22%	18%	461	414	366	287	1.61	1.44	1.11
WhiteheadC	15%	13%	53	45	44	36	1.45	1.26	1.16
Fat, not flat tails									
AlberiniB	50%	34%	83	83	61	53	1.57	1.57	1.00
ChienB	63%	50%	3,307	3,128	1,813	1,743	1.90	1.79	1.06
Groothuis	39%	23%	71	71	59	50	1.41	1.42	1.00
NahuelhualA	47%	33%	199	190	152	140	1.43	1.36	1.05
Neither flat nor fat tails									
Brown96B	14%	4%	7	7	6	4	1.61	1.59	1.01
HammitB	22%	10%	46	45	44	35	1.32	1.29	1.02
IvehammarA	13%	5%	692	665	654	481	1.44	1.38	1.04
TuanA	46%	11%	8	8	8	6	1.28	1.29	1.00

Footnotes

[i] Parsons and Myers (2016) introduce the concept of fat tails as "…a yes-response function having a high and slowly-declining yes-response rate at high bid levels..." However, their summary of studies focuses exclusively on the % yes response at the highest bid. We intentionally parse out both "fat" and "flat" tails in an effort to characterize and evaluate these issues separately.

[ii] Turnbull and Kriström WTP are estimated using the DCchoice package in R (Aizaki, Nakatani, and Sato 2014).

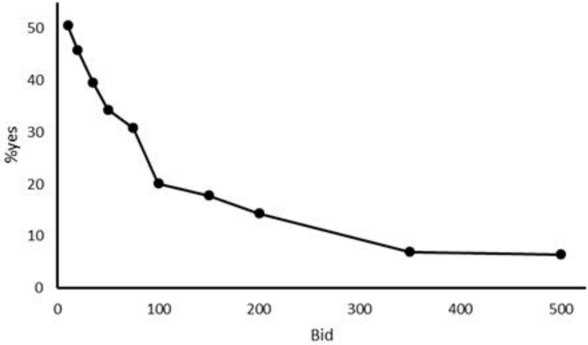
[iii] Hanemann (1989) provides a formula for estimating WTP with the logit model as the area under the positive portion of the probability distribution.

[iv] Researchers often construct standard errors and conduct hypothesis tests with these smoothed data. Such a practice should be reconsidered since it amounts to recoding of the variables under consideration, effectively reducing standard errors and narrowing confidence intervals.

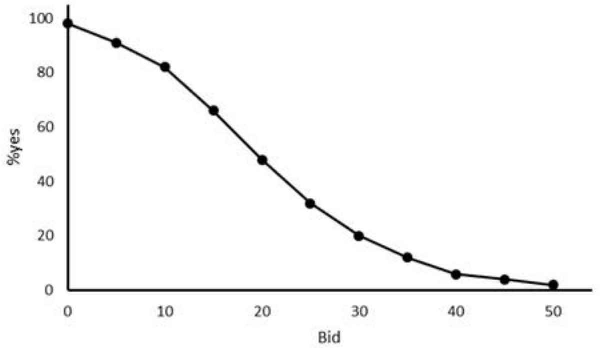
[v] In alphabetical order: AJAE, ARER, ERE, JARE, JAERE, JEEM, and MRE.

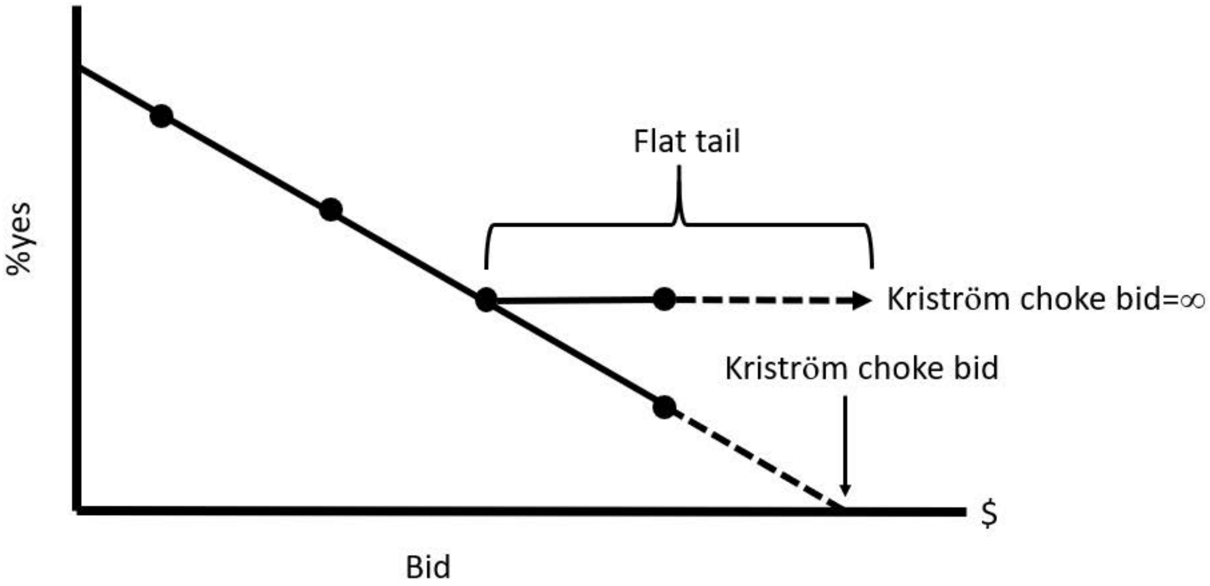
[vi] The DCchoice package in R was used to calculate the Turnbull WTP estimates. It should be noted that, when there are non-monotonicities in the dataset, this package calculates the Turnbull slightly differently than the approach outlined in Haab and McConnell (2002) and used in Stata's turnbull package (Azevedo, 2010). The approach in R results in a slightly higher estimate.

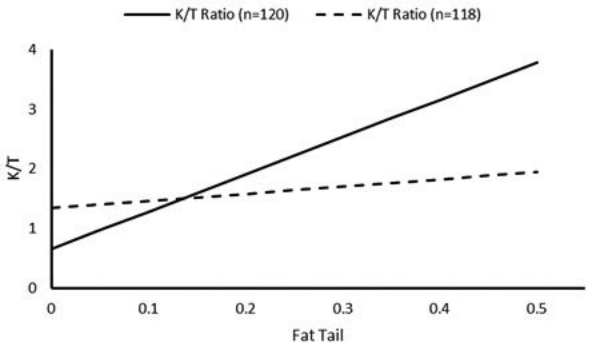
[vii] Interestingly, when we look at subsamples of our data, the more "well-behaved" the data (i.e., it does not need to be pooled), the larger the divergence between the Kriström and Turnbull estimates (K/T = 1.7 for non-pooled data sets vs. 1.5 for pooled datasets) and the more important the Kriström adjustment likely is. When we run regressions on this ratio, the coefficient on the Pctyes2 variable is larger for the non-pooled data and the coefficient on the Slope variable is negative for both the pooled and non-pooled data, but much larger in absolute value for the non-pooled data.

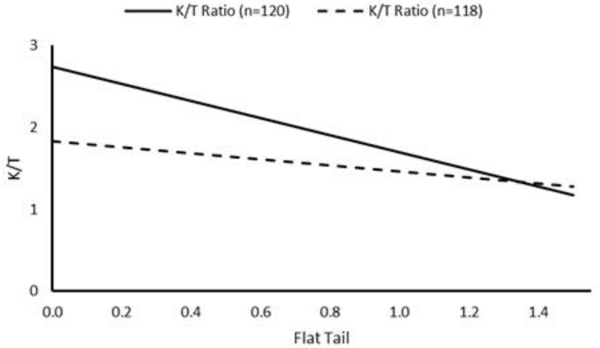


50
40
30
20
10
0
%yes
0
100
200
300
400
500
Bid


100
80
60
40
20
0
%yes
0
10
20
30
40
50
Bid

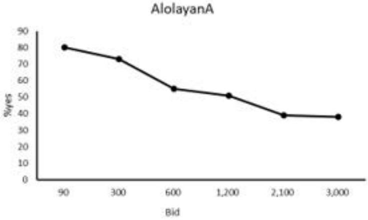

Flat tail
Kriström choke bid=∞
Kriström choke bid
%yes
$
Bid


K/T Ratio (n=120)
K/T Ratio (n=118)
K/T
4
3
2
1
0
0
0.1
0.2
0.3
0.4
0.5
Fat Tail

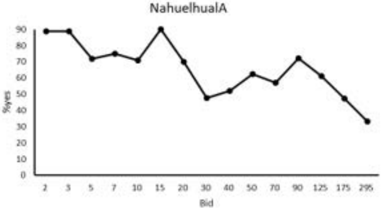


K/T Ratio (n=120)
K/T Ratio (n=118)
3
2
1
0
K/T
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Flat Tail

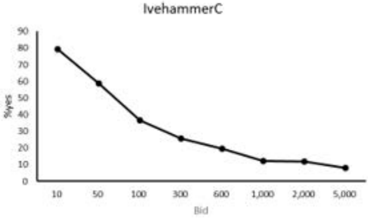
Fat and flat tails



Fat, not flat tails



Flat, not fat tails



Neither fat nor flat tails