Nonparametric Bounds on Welfare Measures: A New Tool for Nonmarket Valuation

by

John Crooker and Catherine L. Kling

Department of Economics
Iowa State University
Ames, Iowa 50014

November 9, 1998

CARD Working Paper 99-WP 208
Center for Agricultural and Rural Development
Iowa State University

Nonparametric Bounds on Welfare Measures:
A New Tool for Nonmarket Valuation

In a series of influential papers, Varian (1982, 1985) extended and refined the work of Afriat (1967, 1976), Samuelson (1948), Houthakker (1950), and Richter (1966), among others, to form the basis for a series of empirically testable hypotheses known generally as the theory of revealed preference. This work demonstrates how observed demand behavior can be used to recover information about an individual's preference ordering without resorting to parametric assumptions regarding the form of the consumer's underlying demand or utility function. Revealed preference theory has been influential in developing empirical tests of utility theory (Varian (1982, 1983)), investigating issues of changes in consumer's tastes (Chalfant and Alston (1988)), testing whether firms behave as profit maximizers (Varian (1984)), as well as a variety of other applications. The general framework has also been extended to account for stochastic elements (Varian (1985)), Sakong and Hayes (1993)). The ability to characterize information about consumer's preferences without imposing a specific functional form for utility or demand is intuitively appealing and has provided a rich base for empirical research in consumer and firm theory.

The issue of parametric specification has been of widespread concern in nonmarket valuation. Most nonmarket valuation methods require the analyst to specify a particular functional form for an estimating equation. It may be a demand, bid, or utility function (or hedonic price function). Although the analyst may perform goodness of fit tests or use other tools to choose among functional forms, there remains a great deal of arbitrariness and researcher judgment in the choice of functional form.

In the travel cost model, it has long been understood that the choice of functional form for either the demand function or the indirect utility function can have significant consequences for the magnitude of the resulting welfare estimates (Ziemer, Musser, and Hill (1980), Kling (1989), Ozuna, Jones, and Capps (1993)). The same has been found in random utility models of recreation demand with respect to the choice of functional form and the assumed error structure (Morey, Rowe, and Watson (1993), Kling and Thomson (1996), Herriges and Kling (1997)). Hedonic housing models used to value air quality are subject to similar concerns (Cropper, Deck, and McConnell (1989)). Finally, the contingent valuation literature has found that changes in either the error structure or the assumed bid function's form can yield large differences in valuation estimates from discrete choice formats (Hanemann (1994)).

Given the empirically observed sensitivity of welfare estimates to functional form, it is natural to consider whether nonparametric methods such as those refined and developed by Varian might be of value in nonmarket welfare analysis. In this research, we first adapt Varian's work on bounding welfare measures to the task of valuing nonmarket commodities. We show how Varian's bounds can be constructed with a single price/quantity combination for each individual in the sample. Unfortunately, these bounds may not be very tight. To remedy this, we develop narrower bounds that can be derived if the analyst has additional data on optimal market bundles at new prices. To derive these bounds, we show how the theoretical relationship between compensating variation and equivalent variation can be exploited to further tighten the bounds. The exciting aspect of this work is that these bounds are derived using only quantity and price information and without any parametric assumption on demand or utility.[1]

The nonparametric bounds thus developed will only prove useful if they are fairly tight. To investigate their potential empirical value, we conduct a Monte Carlo experiment. In this experiment, the nonparametric lower and upper bounds are compared to simulated "true" values of WTP using simulated data sets. Additionally, a natural comparison is to consider how well the bounds perform in estimating welfare relative to traditional parametric approaches. To consider this, traditional travel cost type models are estimated on the simulated data sets and point estimates and confidence intervals are constructed from these models which are then compared to the nonparametric bounds.

In the recreation demand literature, Boxall, Adamowicz, and Tomasi, (1996) and Larson, Klotz, and Chien (1991) have used Varian's methodology to test for consistency between contingent valuation (stated preference) models and recreation demand (revealed preference) models. Here, we use and extend the methodology to actually provide information on the magnitude of welfare changes for nonmarket goods.

Using Observed Data to Compute Bounds on the WTP for Price Changes

Bounds Based on One Data Point for Each Individual

To begin, assume that the analyst has a single price/quantity observation for each sample observation. Varian's seminal work demonstrates how bounds on each

[1] Note, however, that because the bounds rely upon the theoretical relationship between compensating and equivalent variation, and because this relationship itself reverses if the good is inferior, the analyst must assume that the good is normal (or inferior and make the appropriate changes in the derivations). This then is a caveat on the bounds and their nonparametric nature. We comment further on this issue in the theoretical section of the paper.

individual's compensating variation for a proposed price change can be constructed. Consider a simple budget constraint for an individual choosing between recreation visits (v) and a composite commodity (z). In Figure 1, $X_0 \equiv (v_0, z_0)$ denotes the chosen commodity bundle at the initial price vector (denoted P_0 in the figure) and M is the consumer's income. Let $\mathrm{X} \equiv \{(v, z) : v, z \in R_+\}$ be the set of all possible bundles.

In order to calculate the exact compensating variation (CV) associated with a particular price change, we would need to determine the amount of money the individual is willing to give up to receive the price change. Formally,

$$\begin{aligned} CV &= e(P_0, U_0) - e(P_N, U_0) \\ &= M - e(P_N, U_0), \end{aligned} \tag{1}$$

where $e(P, U)$ denotes the individual's expenditure function, $U_0 \equiv U(v_0, z_0)$ denotes the level of utility at X_0, and P_0 and P_N are the prices before and after the price change. The first term $e(P_0, U_0)$ is exactly the initial income of the consumer (M). If we can provide bounds on the second term, $e(P_N, U_0)$, we can also bound CV. Thus, we seek to compute lower and upper bounds on the expenditure that would be necessary for the consumer after the price change to obtain the original utility level.

We now ask the question: What is the most amount of income we can take away from or give to this individual after a price change to be sure that he or she can attain the original level of utility? Suppose, as depicted in Figure 1, we are interested in the CV for a price decrease from P_0 to P_N where P_0 represents the budget constraint at the initial prices and P_N represents the new budget constraint.

We know the individual can at least attain his initial level of utility if he can afford his initial bundle. Thus, that amount of expenditure is the most he would ever need after the price change. In Figure 1, this upper-bound on expenditure is:

$$\overline{M}_{CV}^{O} = P_N X_O . \qquad (2)$$

Graphically, $\overline{M}_{CV}^{O}$ can be identified as the vertical intercept of a straight line parallel to P_N that intersects X_0 (the dashed line through X_0 in Figure 1). If the consumer views v and z as perfect complements, $\overline{M}_{CV}^{O}$ is exactly equal to the expenditure necessary to attain the original level of utility at the new prices. However, if there is at least some substitution possible between v and z, the consumer could attain his initial utility level with less income than $P_N X_0$, thus $\overline{M}_{CV}^{O}$ represents an upper bound on necessary expenditure.

Following this logic, the least expenditure that could possibly be required to keep the consumer at the original level of utility after the price change would occur if the goods were perfect substitutes (i.e., straight line indifference curves). In this case, income can be taken away from or given to the consumer until he would pick the corner solution that minimizes expenditures. Graphically, the lower bound on expenditure can be identified by drawing a straight line parallel to P_N that intersects the vertical intercept of P_0, denoted $\underline{M}_{CV}^{O}$ in Figure 1.

Combining the upper and lower bounds on expenditure, we get bounds on CV:

$$B_{CV}^{O} \equiv \left[M - \overline{M}_{CV}^{O} , M - \underline{M}_{CV}^{O} \right] . \qquad (3)$$

The superscript on the LHS and the expenditure bounds reflects the fact that these bounds are constructed knowing only a single data point (the original commodity bundle). Note that the lower bound on expenditures determines the upper bound on CV and vice versa. The proximity of CV to the bounds depends upon the degree of substitutability between the goods. If the goods are perfect substitutes, CV will exactly equal the upper bound. Conversely, if the goods are perfect complements, CV is exactly the lower bound.

Although it is clear that this procedure can be used to compute bounds on individual CV, such bounds will only be of interest if they are fairly narrow. The next section describes how the addition of a second data point (price/quantity observation) can narrow these bounds.

Bounds Based on Two Data Points for Each Individual

In this section, we demonstrate how Varian's bounds can be improved upon with additional data and by appealing to the properties of Hicksian welfare measures. Suppose that in addition to knowing the optimal bundle chosen by the consumer at the original prices, the analyst also knows the optimal bundle chosen by the individual at the new prices. A second price/quantity vector might be obtained for an actual sample in at least two different ways. First, analysts might collect data on use over two seasons or time periods. In this case, the analyst would have two consumption bundles at two sets of prices based on revealed preference data. Alternatively, contingent behavior (stated preference) data could be combined with the revealed preference data to generate the second data point. In fact, a series of price/quantity combinations could be collected in a survey where respondents are asked how many visits they would take under a range of different prices of access to the good.

Regardless of the source of this second data point, the question of interest is: does the addition of this information help us tighten the bounds on CV for a price change from

P_0 to P_N? The answer is yes, but the link is indirect and requires us to consider the equivalent variation (EV) for the price decrease. In particular, suppose that the consumer reveals to the researcher that X_N is (or would be) his chosen commodity bundle at prices P_N. This information allows us to compute bounds on the EV for the price change from P_0 to P_N. By appealing to the fact that the equivalent variation for a price decrease for a normal good is greater than or equal to the compensating variation for the same price decrease, we can potentially tighten the upper bound on CV by using the upper bound on EV in its place.

Equivalent variation for the price decrease is defined as

$$\begin{aligned} EV &= e(P_0, U_N) - e(P_N, U_N) \\ &= e(P_0, U_N) - M. \end{aligned} \tag{4}$$

The second term on the RHS of (4) equals the consumers income so, again, if we can bound the first term, we can bound the equivalent variation.

To do so, again consider Figure 1. The exact EV could be obtained if we knew exactly how much money we would need to give the consumer at the initial prices (P_0) to achieve the utility at X_N. Now, the most that would be required to achieve this utility level is if the consumer could obtain bundle X_N at the original prices. Thus, if the consumer were given $\overline{M}_{EV}^{N} - M$ instead of the price change, we can be certain that he could achieve at least the same level utility as if the price change had occurred. Thus, $\overline{M}_{EV}^{N} - M$ provides an upper bound on the necessary compensation.

However, unless the consumer is unwilling to substitute any z for v, the consumer will be able to achieve the same level of utility as X_N provides at less than this level of compensation. What is the least amount of compensation that might allow the consumer to obtain the same utility as provide by X_N? If z and v are perfect substitutes and an interior solution is observed, the indifference curve between them would be a straight line and would be identical to the budget line defined by P_N. In this case, the consumer would need only his original income to achieve the new utility level. Thus, the lower bound on EV is simply $\underline{M}^{N}_{EV} - M = 0$. Unfortunately, a lower bound of zero is not particularly informative. Nevertheless, we can now bound EV as follows:

$$B^{N}_{EV} \equiv \left[\underline{M}^{N}_{EV} - M, \overline{M}^{N}_{EV} - M \right] = \left[0, \overline{M}^{N}_{EV} - M \right], \tag{5}$$

where the superscript "N" indicates that only the second data point is used to construct these bounds. We now use the bounds on EV to potentially help tighten the bounds on CV. Since EV for a price decline is greater than CV, we know that an upper bound on EV must also be an upper bound on the CV. Thus, we can use the lower of the two upper bounds derived via nonparametric methods to provide an upper bound on CV. The bounds on CV derived using information from both data points can be written

$$B^{ON}_{CV} \equiv \left[M - \overline{M}^{O}_{CV}, Min(M - \underline{M}^{O}_{CV}, \overline{M}^{N}_{EV} - M) \right]. \tag{6}$$

The superscripts on B indicate that both points are used in inferring the bounds.

As pointed out initially, this methodology is valid only for a non-inferior good. That is, the income effect must be non-negative. In some cases, this may be problematic as empirical research on recreation goods has found evidence of negative income effects for certain resources. However if the analyst knows that the good is inferior, the relationship between CV and EV can still be used to tighten the bounds. In this case, CV exceeds EV so the EV provides a tighter **lower** bound.

Finally, note that bounds on EV can be similarly constructed and tightened by using information about CV. Specifically,

$$B_{EV}^{ON} \equiv \left[Max(\underline{M}_{EV}^{N} - M, M - \overline{M}_{CV}^{O}), \overline{M}_{EV}^{N} - M \right]. \qquad (7)$$

The improvement of the lower bound in this case also follows from the fact that the EV for a price decrease equals or exceeds the CV. Clearly this might tighten the bounds significantly as the lower bound of $\underline{M}_{EV}^{N} - M = 0$ is uninformative.

Both commodity bundles considered thus far have been located on one of the budget constraints corresponding to the two price vectors for which the welfare change is being assessed. In the next section, we consider whether further tightening of the bounds is possible if the analyst also knows what choices the consumer would make at intermediate price ratios.

Bounds Based on Three or More Data Points for Each Individual

Now suppose that the analyst knows yet a third price-quantity combination for each individual and suppose that that combination corresponds to a price ratio that lies between the initial and proposed price change. Can information about the commodity bundle that the consumer chooses at such a price ratio be used to narrow the bounds on CV (or EV)? The answer is yes: it can raise the lower bound under some circumstances and lower the upper bound in all cases.

To see how this point may raise the lower bound, turn to Figure 2 where we have depicted the original and new budget constraints (P_0 and P_N) and the corresponding optimal commodity bundles (X_0 and X_N). We have also drawn an intermediate budget constraint and an associated optimal bundle labeled X_1. Recall that to provide a lower bound on CV, we want to know what amount of income we can take away from the consumer and be sure that he can still attain the same level of utility with the new prices as at the original commodity bundle.

As drawn, knowledge that X_1 is the optimal commodity bundle at prices P_1 allows us to increase the amount of income that can be taken away from the consumer and still be sure that the original utility level is obtained, thus increasing the lower bound on CV. To see this, note that since X_1 is chosen at P_1 when X_0 was affordable, we know that X_1 represents a higher level of utility than X_0 and lies on a higher indifference curve than X_0. This, in turn, implies that if income were taken away from the consumer at the new set of prices (P_N) until the consumer could afford X_1, they would still be obtaining at least as much utility as at X_0. Thus, an expenditure level of $\overline{M}_{CV}^{1}$ is sufficient to ensure that the consumer is no worse off than the original utility level. Thus, we have an improved lower bound on CV and we can write our newly formulated lower bounds that are based on information from three data points as

$$LB_{CV}^{ON1} \equiv \left\{ Max(M - \overline{M}_{CV}^{O}, M - \overline{M}_{CV}^{1}) \right\}. \qquad (8)$$

Thus, we have succeeded in further decreasing the interval over which the true CV is contained. In like manner, lower bounds on the EV can be written

$$LB_{EV}^{ON1} \equiv \left\{ Max(\underline{M}_{EV}^{N} - M, M - \overline{M}_{CV}^{O}, M - \overline{M}_{CV}^{1}) \right\} \qquad (9)$$

At this point, it is important to point out that not all intermediate price ratios will provide information that can be used to raise the lower bounds. Graphically, the optimal commodity bundle associated with P_1 (X_1) must lie to the left of the line through X_0 with a price ratio of P_N. Otherwise, no improvement on the bound generated by $\overline{M}_{CV}^{O}$ can be computed. Consumption bundles that will tighten the welfare bounds will be generated only when the consumer's preferences generate backward bending offer curves such that the new consumption bundle is cheaper than the original bundle at the new prices.

The addition of this third data point can also lower the upper bound on CV. Specifically, with a third data point, the new upper bound can be written

$$\left(P_O - P_1\right)v_1 + \left(P_1 - P_N\right)v_N . \tag{10}$$

To demonstrate that (10) constitutes an upper bound, appeal again to the fact that the EV for a price decrease is greater than the CV for the same price decrease. From this fact follows the first inequality in (11)

$$\begin{aligned} e(P_0,U_0) - e(P_1,U_0) &\leq e(P_0,U_1) - e(P_1,U_1) \\ &\leq P_0 v_1 + z_1 - M \\ &= P_0 v_1 + z_1 - (P_1 v_1 + z_1) \\ &= (P_0 - P_1)v_1. \end{aligned} \tag{11}$$

The second inequality in (11) follows from the fact that the expenditure necessary to achieve U_1 at the initial prices (P_0) must be less than or equal to the expenditure that would be required to allow the consumer to purchase the commodity bundle that achieves U_1 at prices P_1. Based on identical reasoning, the following inequalities hold

$$\begin{aligned} e(P_1,U_0) - e(P_N,U_0) &\leq e(P_1,U_N) - e(P_N,U_N) \\ &\leq P_1 v_N + z_N - M \\ &= P_1 v_N + z_N - (P_N v_N + z_N) \\ &= (P_1 - P_N)v_N. \end{aligned} \tag{12}$$

Summing (11) and (12) yields

$$e(P_0,U_0) - e(P_N,U_0) \leq (P_0 - P_1)v_1 + (P_1 - P_N)v_N , \tag{13}$$

which establishes the new upper bound. The reasoning can be extended indefinitely so that all additional data points will also lower this upper bound.

This new upper bound is strictly less than the potential upper bound determined by EV i.e.,

$$(P_O - P_1)v_1 + (P_1 - P_N)v_N < (P_O - P_N)v_N = P_O v_N - P_N v_N + z_N - z_N = \overline{M}_{EV}^{N} - M. \qquad (14)$$

It is now possible to write lower and upper bounds on CV associated with three data points

$$B_{CV}^{ON1} \equiv \left[Max(M - \overline{M}_{CV}^{O}, M - \overline{M}_{CV}^{1}), Min(M - \underline{M}_{CV}^{O}, (P_O - P_1)v_1 + (P_1 - P_N)v_N \right] \qquad (15)$$

Adding information on individual's optimal commodity bundles at a variety of price ratios can tighten the nonparametric bounds on CV or EV for a price change. Again, the analyst must know whether the good is normal or inferior, but other than that, there are no parametric assumptions necessary: regardless of the preferences of the individual, as long as they conform to the basic postulates of neo-classical consumer theory, the bounds must contain the true WTP.[2]

Although their accuracy is certain (subject only to error in the underlying data), the ultimate value of these bounds depends on their width. Bounds that are very wide will provide too little information for a policy analyst and will likely be passed over in favor of parametric estimates that provide at least the appearance of precision to those who use this information. If parametric methods can be accurately estimated and/or if the nonparametric bounds are quite wide, there is little reason to pursue research employing the nonparametric bounds. Alternatively, if nonparametric bounds are found to have the potential to be relatively narrow in practice and/or if parametric methods generate significant error in welfare measurement then nonparametric bounds may have an important role to play in welfare analysis. Thus, we undertake a simulation exercise in an effort to gauge the likely value of additional research on these bounds.   We use Monte Carlo simulation techniques to explore the improvement to the bounds that additional data points generate and also how the width of the nonparametric bounds compare to point estimates and confidence intervals generated by traditional parametric approaches to welfare estimation.

A Monte Carlo Study

Design of the study

The Monte Carlo experiment is designed with these three questions in mind:

- How narrow can we expect the nonparametric bounds to be?
- How much does the addition of data points improve (tighten) the bounds?
- How do the nonparametric bounds compare to welfare estimates generated by parametric estimators?

In the previous section, the lower bound on CV was seen to exactly equal the true CV when the two goods are perfect complements and the upper bound was exactly the true CV when the two goods are perfect substitutes. These results make clear that the accuracy of the bounds are affected by the degree of substitutability between the good whose price change is being evaluated and the numeraire. To consider alternative degrees of substitution possibilities easily, we employ a Constant Elasticity of Substitution (CES) utility function which allows a wide range of substitution possibilities. Consider the CES utility framework

---

[2] Although we abstract from considering error terms here to concentrate on the fundamentals of the theory, it may be necessary to worry about the implications of errors in consumer's optimization behavior or measurement error when applying the bounds.

$$U(v,z) = \left(\alpha z^{\rho} + (1-\alpha)v^{\rho}\right)^{1/\rho}, \quad s \equiv \frac{1}{1-\rho}, \tag{16}$$

where, as before, z is the numeraire, v is the quantity of the environmental good, and s, ρ, and α are parameters. The CES is a convenient utility function to work with since the single parameter, s, determines the degree of substitutability between the goods.

An error term is introduced into the CES additively via the "α" parameter:

$$U = \left[(1-\alpha-\eta)v^{\rho} + (\alpha+\eta)z^{\rho}\right]^{1/\rho}, \tag{17}$$

where η~Uniform(-0.25,0.25). Then, the true form of demand is given by

$$v = \frac{(1-\alpha-\eta)^{s} M}{(\alpha+\eta)^{s} P^{s} + (1-\alpha-\eta)^{s} P}. \tag{18}$$

We set the parameter α=0.75 . To examine the sensitivity of the results to the degree of substitution, we investigate four different values of s: s=0.5,2, 5, and 20.

Using this utility function and parameter values, we generate 1,000 samples of 300 observations each. For each observation, the simulated price is randomly drawn from the uniform distribution on the interval (5,55). Also, income is randomly drawn from the uniform distribution on the interval (5000, 85000).

How Tight Are the Nonparametric Bounds and How Much Do Additional Data Points Improve the Bounds?

As demonstrated in the theoretical sections above, bounds on welfare measures can be constructed with a single data point, two data points, and three or more points for each individual. In the first part of the Monte Carlo experiment, we investigate how the addition of data points (observations) for each individual in the sample can narrow the bounds. As mentioned earlier, one possible source for such additional observations is via contingent behavior. Although those who are suspicious of contingent valuation as a reliable valuation method may discount such data, some analysts may be more comfortable with behavioral contingent data than willingness to pay questions. For example, Bockstael and McConnell (forthcoming) have recently argued that:

> Such contingent behavior studies might not suffer from many of the problems encountered when asking values and they would be targeted towards people who "behave" in the context of the problem and who would presumably not find it difficult to imagine the behavioral changes they would make when faced with different prices, different qualities, different alternatives (page 29).

If contingent behavior is viewed as a reliable source of data and if nonparametric bounds can be constructed from this data that are sufficiently narrow to be of practical use, there might be a potentially compelling case for their use in place of parametric estimates. A Monte Carlo experiment where there is assumed to be no measurement error associated with the data is an ideal environment to shed light on this question. For, if the nonparametric bounds are too wide to be of policy interest in this setting, they can almost certainly be ruled out as a viable valuation strategy when the vagaries of real data are considered.

To assess the gains from adding contingent behavior data to a single observed data point for each observation (such as might be collected in a typical recreation demand study), we compute the nonparametric bounds for each Monte Carlo sample and average the lower and upper bounds. This process is repeated for each of the samples.

First, the upper and lower bound on CV for a price decrease associated with a single data point is computed.[3] This is equivalent to using the information an analyst

[3] Since the CV for a price decrease is identical to the EV for the inverse price increase, the values in the tables can be interpreted as bounds on either measure; however, we will refer to it as a bound on CV for simplicity.

might typically have from a travel cost type recreation demand study. For each individual in the sample, the analyst would know only how many trips the individual took during the time period and at what price. In the rows marked "Point O" of Tables 1a and 1b we report the bounds generated by this procedure. Results are presented for two different price changes: a 25% decrease and a 80% decrease and four different values of s (the substitutability parameter).

As can quickly be seen, the range between the lower and upper bound is enormous in all cases and thus of no real value from an applied policy perspective. This is not surprising as a single data point per individual provides little information. In the rows marked "Point N", a second data point for each individual is used (along with the first) to form the bounds. This point corresponds to the quantity chosen by the individual at the "new" price, i.e., it corresponds to point "N" in Figures 1 and 2 from the theoretical discussion. With the introduction of this second point, the upper bound on the compensating variation drops dramatically in all cases.

In the row marked "Point 1", the third data point is used to raise the lower bound and lower the upper bound as described in the theoretical section above. The third data point is generated by determining the quantity consumed at the midpoint price between the initial and final price in the welfare change. Although the gains in tightening the interval are not nearly as large as the addition of the second point, it is clear that valuable gains are possible. In a number of cases, the conditions necessary to raise the lower bound are present, thus the addition of the third data point both raises the lower bound and lowers the upper bound. However, even when the lower bound remains unchanged, the range between the lower and upper bound is small enough to be of use in certain policy situations.

In the rows marked "Point 2" and "Point 3", two additional price/quantity combinations are used to tighten the bounds. These combinations are determined by computing the midpoints between the point 1 price and the initial price and the final and in initial price, respectively. Again, the nonparametric bounds are potentially tightened by this additional information. The gains come primarily from lowering the upper bound on WTP. As noted above, each new data point will necessarily lower the upper bound as we are able to trace out the individuals demand function. If we learn of every commodity bundle the individual would choose for all intermediate prices, our upper bound on WTP would be precisely the individual's Marshallian consumer surplus. The individual's consumer surplus is the best we can do in deriving an upper bound on WTP for the price decrease in the nonparametric setting.

Our ability to raise the lower bound hinges on the shape of the individual's offer-curve. Specifically, if the offer curve is backward bending for some intermediate price changes and we learn of commodity bundles chosen at these prices, then we may raise the lower bound on WTP. A backward bending offer curve is a necessary but not sufficient condition for raising the lower bound.

These Monte Carlo results strongly suggest that with the addition of at least one more, and possibly several, data points, nonparametric bounds can be constructed that are narrow enough to be truly informative to a policy maker. Next, we consider how these bounds compare to parametric estimates generated by the same amount of information.

How Do the Nonparametric Bounds Compare to Standard Parametric Estimates?

For purposes of this portion of the Monte Carlo study, we assume that the researcher has access to a data set with three data points for each individual in the sample, corresponding to points O, N, and 1 from the previous section. Again, we have in mind that the researcher may have undertaken a contingent behavior survey to collect such data and we will again abstract from measurement error or other problems potentially associated with such data. Here we ask how well the researcher could do with such a data set in estimating CV using the nonparametric bounds relative to employing a parametric demand model (such as a typical travel cost type model).

For each sample, we estimate each of three parametric demand functions:

$$\begin{aligned} &\text{Log-linear:} && \ln(\mathrm{v}) = a + b\ln(P) + g\ln(M) + e, \\ &\text{Semi-log:} && \ln(v) = a + bP + gM + e, \text{ and} \qquad (19) \\ &\text{Linear:} && v = a + bP + gM + e, \end{aligned}$$

where the greek letters again correspond to parameters. These demand functions were chosen due to their common use in recreation demand modeling. To estimate the models, we include all three data points for each individual that are used in constructing the nonparametric bounds. Thus, the original point plus the "contingent behavior" data are used in constructing both the nonparametric bounds and the parametric estimates. In this way, the parametric and nonparametric methods are both confronted with the same amount of information. To incorporate the fact that the three observations for each individual are not independent[4] (that is, $E(e_{ij}e_{ij}) \neq 0, \;\; j = 1,2,3$ where i indexes individuals and j indexes observations), we estimate the models in (22) using a standard Feasible Generalized Least Square Estimators to capture this correlation.[5]

After estimating each model, we calculate the average estimated CV for each functional form and do so for each of the 1000 repetitions. Next, we order the respective averages from smallest to largest and construct empirical 95% confidence intervals for each method.

To provide a benchmark against which to compare both the nonparametric bounds and the parametric estimates, we compute the true compensating variation for a proposed price decrease and average these over all individuals in the simulated samples and over the 1000 Monte Carlo trials. We also order the distribution of the 1000 sample average true CV's from highest to lowest and identify the fifth and ninety-fifth percentile of that distribution. This provides the 95% confidence interval for the true distribution against which the parametric confidence intervals and the nonparametric bounds can be assessed.

Tables 2a and 2b contain the point estimates, confidence intervals, nonparametric bounds and true CV for the simulated data for a 25% and 80% price reduction, respectively. We also report the average R^2's for the parametric estimates to provide a sense of the goodness-of-fit of the parametric models to the data (and thus how "typical" these scenarios might be). First note that the nonparametric bounds are always (by construction) true bounds on the true intervals. In contrast, the parametric bounds (confidence intervals) are not.

---

[4] In real data, the correlation across individuals may arise from omitted variables specific to individuals or any number of measurement problems. In our simulated data, correlation across individuals arises from the fact that individuals have different true parameters values from one another.

In this case, there is no parametric demand function that is an exact match for the true demand function, although the log-linear represents a special case of the CES demand. In fact, the situation where the "true" demand functions are not an exact match to the parametric specification strikes us as the most accurate representation of the typical study.

Not surprisingly, there are a number of estimated confidence intervals that lie within the true intervals (i.e., they are too narrow). Even more strikingly, in some cases (identified in the table in italics) the point estimates themselves lie outside of the true interval. Thus, by using a parametric point estimate an analyst might actually be reporting a welfare measure that is not even within the true 95% confidence interval. This of course is not news to applied researchers: incorrect functional forms are well known to potentially generate welfare measures with large error. More to the point is that an alternative that <u>does not</u> require the assumption of a particular functional form exists and generates ranges that, at least in some cases, are likely to be narrow enough for policy making.

Nonparametric Bounds and Standard Parametric Estimators When the Population Preference Structure is Heterogeneous.

An even more realistic situation than one in which recreationists have random parameters, but share the same functional form for utility is one in which the population consists of individuals with different utility structures. To consider this situation, we allow the population we are sampling from to consist of individuals with both semilog

---

[5] Specific details on the GLS estimator are available from the authors upon request.

demand utility and CES utility. Each type comprises 50% of the population. The semi-log demand utility function is

$$U = \frac{\beta + \gamma}{-\gamma\beta}\exp\left[\frac{(\gamma(\alpha+\varepsilon)v - \beta z - v\ln(v))}{\beta+\gamma}\right], \tag{17}$$

where, greek letters indicate parameters. The parameter values in the semilog utility function are set at α=2, β=-0.04, and γ=-0.00002.[6] The stochastic error component is distributed $N(0, \sigma_\varepsilon^2)$ and three different dispersion levels are examined: σ_ε = 0.015625, 0.0625, and 0.125. For the CES framework the parameters are α=0.55, s=2.5 and η~U[-0.00125,0.00125].

Table 3 contains the results of this simulation experiment. The numbers reported correspond to WTP for a 25% price reduction. The results are fairly striking: despite the relatively high values of R^2, the parametric model's confidence intervals do not contain any of the true mean values of WTP. This is particularly interesting in the case of the semilog demand specification where the average values estimated are quite close to the simulated "truth", but the confidence intervals still exclude the mean. In contrast, the nonparametric bounds are true bounds and for this particular parameterization are quite tight.

As a final measure of the value of the nonparametric bounds, we compute the mean percent error associated with using the midpoint of the nonparametric bounds as an estimate of the average WTP. These statistics range from -10.2% to +6.2% with an

[6] These parameter values were chosen because they were employed in a previous Monte Carlo study (Kling, 1997) and they produce "sensible" looking numbers of visits in our application.

average of -4.4%. We believe that these results provide a rather compelling case for further investigation of nonparametric methods in nonmarket valuation.

Final Remarks on the Value of Nonparametric Bounds on Welfare Measures

In this paper, we have presented simple methods for constructing nonparametric bounds on compensating or equivalent variation for price changes based on nonparametric methods. We began with the methods developed by Varian and derived additional results allowing significant tightening of the bounds. These bounds have the potential to provide an alternative valuation method to standard parametric estimation of recreation demand.

The ultimate usefulness of the bounds derived here will depend upon how tight the bounds can be constructed for real data and on whether the data necessary to compute such bounds can be obtained and deemed reliable. In our Monte Carlo analysis, we have demonstrated that there are situations under which the first of these conditions will hold: bounds constructed without reference to parametric demand specifications can yield intervals that are narrow enough for policy purposes. However, questions concerning the reliability of contingent behavior data or the possibilities of collecting time series data must await the confrontation of a real data set.

As noted in the derivation of the bounds, the analyst must be sure that he knows whether the good is inferior or normal. This is a potential limitation of the approach, especially given that many recreation goods for which nonmarket values are sought may have negative income effects. However, the issue appears to be no less troublesome for parametric models, which generally impose and estimate a single parameter value for the income effect on the entire sample.

Nonparametric bounds on welfare measures are appealing in that they require no assumptions about utility functions or error structures. Equally importantly, they also do not require assuming that all individuals in a sample have the same preference structures or parameter values. Such liberty is heartening, but comes at a cost. Rather than being able to report precise-sounding estimates of welfare, bounds convey uncertainty. However, as the results of these Monte Carlo experiments suggest, the "certainty" conveyed by point estimates from traditional parametric estimators may be misleading.

The results using nonparametric bounds developed here constitute a first look at applying nonparametric methods to bound welfare measures for nonmarket goods. Based on the theoretical and simulated results presented here, we are optimistic that additional work in this area will yield substantial returns. The ability to provide policy makers with tight bounds on welfare measures for nonmarket goods that are free of functional form assumptions is an appealing proposition.

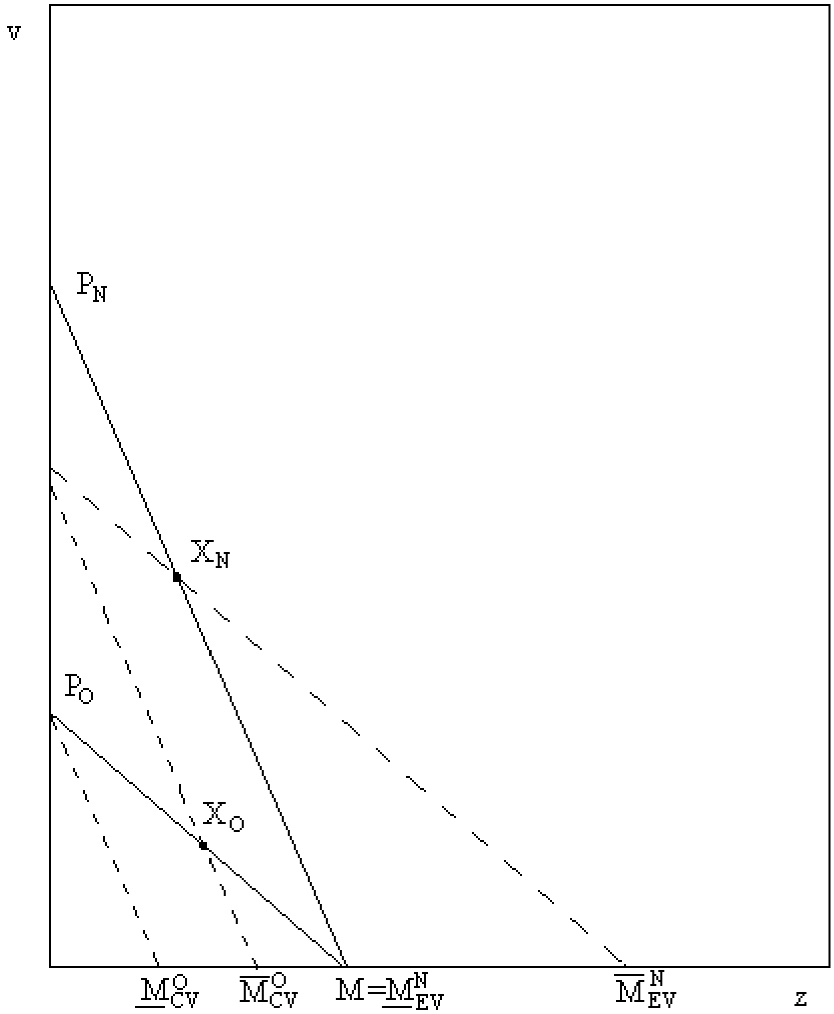
v
P_N
X_N
P_O
X_O
$\underline{M}^O_{CV}$
$\overline{M}^O_{CV}$
$M=\underline{M}^N_{EV}$
$\overline{M}^N_{EV}$
z

Figure 1

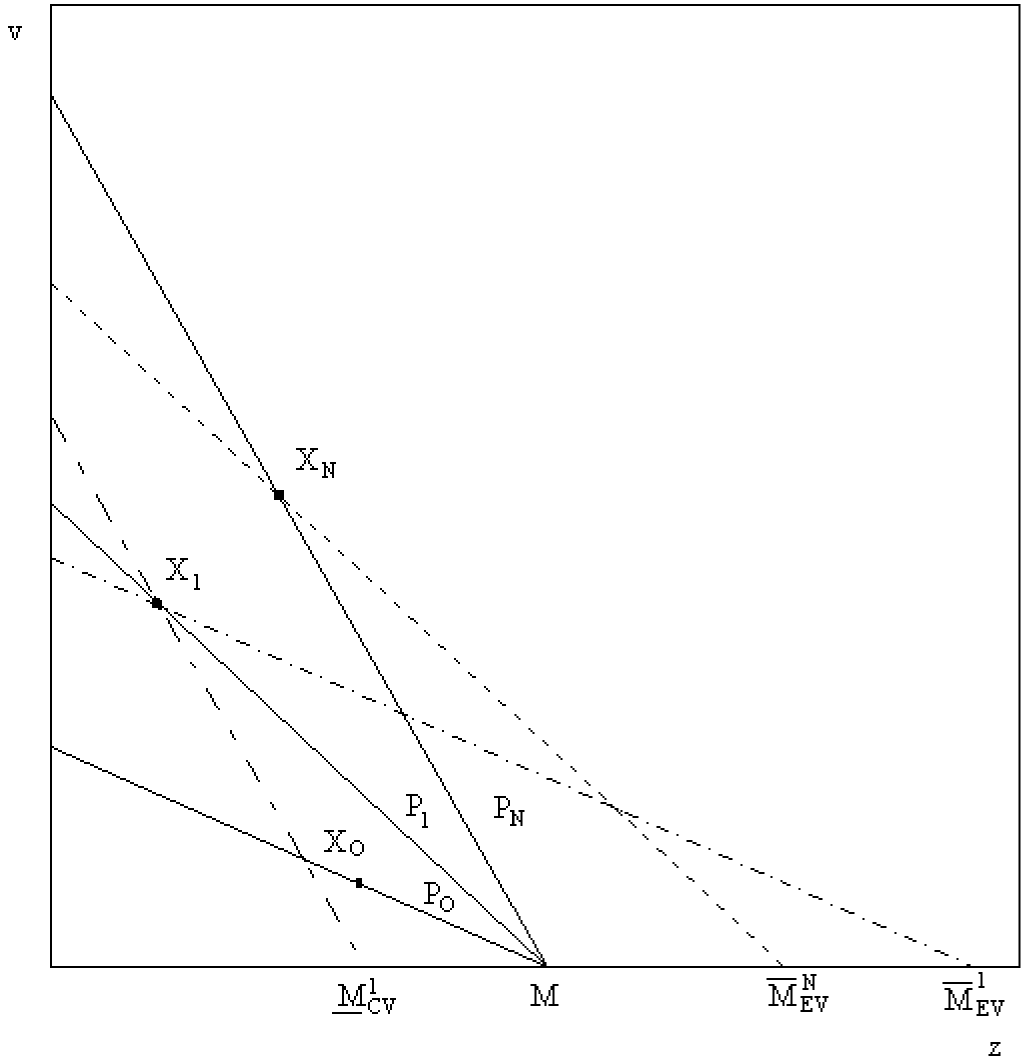
v
X_N
X_1
P_1
P_N
X_O
P_O
$\underline{M}^1_{CV}$
M
$\overline{M}^N_{EV}$
$\overline{M}^1_{EV}$
z

Figure 2

Table 1a: **WTP for a 25% reduction in price: CES Preferences**

Theoretical Bounds	**s=0.5**		**s=2**		**s=5**	
	I_L	**I_H**	**I_L**	**I_H**	**I_L**	**I_H**
Point O	7882	11450	98	11450	0.02	11450
Point N	7882	10563	98	251	0.02	0.54
Point 1	7882	10035	98	224	0.02	0.41
Point 2	7882	9730	98	210	0.02	0.35
Point 3	7882	9581	98	204	0.02	0.33
True	8040	8389	130	188	0.04	0.28

Table 1b: **WTP for a 80% reduction in price: CES Preferences**

Theoretical Bounds	**s=0.5**		**s=2**		**s=5**	
	I_L	**I_H**	**I_L**	**I_H**	**I_L**	**I_H**
Point O	25408	36882	313	36882	0.05	36882
Point N	25408	36882	313	9569	0.05	627
Point 1	25408	36882	468	4330	1.26	174
Point 2	25408	36882	468	3628	1.26	164
Point 3	25408	36379	468	3500	1.26	163
True	28115	29205	1408	1914	9.96	42

Table 2a: **WTP for a 25% reduction in price: CES Preferences**

Models	s=0.5				s=2				s=5			
	I_L	AVG	I_H	R^2	I_L	AVG	I_H	R^2	I_L	AVG	I_H	R^2
Linear	6347	6519	6674	0.67	121	160	204	0.18	0.06	0.21	0.41	0.04
Log-Linear	3786	5109	6981	0.92	70	101	160	0.25	0.02	0.06	0.11	0.21
Semi-Log	5791	5924	6038	0.82	81	99	117	0.23	0.01	0.04	0.08	0.19
Nonparametric	7687		9352		89		181		0.02		0.22	
True	7839	8021	8183		118	142	167		0.03	0.10	0.19	

Table 2b: **WTP for an 80% reduction in price: CES Preferences**

Models	s=0.5				s=2				s=5			
	I_L	AVG	I_H	R^2	I_L	AVG	I_H	R^2	I_L	AVG	I_H	R^2
Linear	25864	26686	27469	0.48	2697	3640	4655	0.07	111	490	998	0.01
Log-Linear	15183	16601	18086	0.92	538	633	747	0.37	0.74	2.74	5.58	0.34
Semi-Log	19754	20214	20672	0.80	632	771	912	0.31	0.22	0.47	0.82	0.28
Nonparametric	23845		34185		450		3437		1.48		289	
True	26406	26936	27453		1355	1613	1872		11.57	39.64	76.20	

Table 3: **WTP for a 25% reduction in price: Mixed Semilog, CES Preferences**

Models	**s_e=0.015625**				**s_e=0.0625**				**s_e=0.125**			
	I_L	**AVG**	**I_H**	**R^2**	**I_L**	**AVG**	**I_H**	**R^2**	**I_L**	**AVG**	**I_H**	**R^2**
Linear	170.3	170.7	171.1	0.11	181.2	182.0	182.8	0.19	171.1	172.5	173.9	0.20
Log-Linear	-1049	-823	-648	0.50	77.27	80.0	82.78	0.54	151.4	159.7	168.6	0.56
Semi-Log	130.3	130.6	130.8	0.49	130.9	131.2	132.9	0.49	127.9	129.8	131.6	0.54
Nonparametric	108.7	120.1	131.5		107.0	119.5	132.1		104.2	119.3	134.4	
True	125.4	125.7	125.9		124.5	125.4	126.2		124.6	126.2	127.6	

References

Afriat, S., "The Construction of a Utility Function from Expenditure Data," *International Economic Review*, 8(1967): 67-77.

Afriat, S., *The Combinatorial Theory of Demand*, London: Input-Output Publishing Company, 1976.

Bockstael, N. and K. McConnell. "The Behavioral Basis of Non-Market Valuation," Chapter 1 in *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*, edited by J. Herriges and C. Kling, Edward Elgar Press, forthcoming 1999.

Boxall, P., W. Adamowicz, and T. Tomasi, "A Nonparametric Test of the Traditional Travel Cost Model," *Canadian Journal of Agricultural Economics* 44 (1996): 183-93.

Chalfant, J. A., and J. M. Alston, "Accounting for Changes in Tastes," *Journal of Political Economy*, 96(1988): 391-410.

Cropper, M. L. L. B. Deck, and K. E. McConnell, "On the Choice of Functional Form for Hedonic Price Functions," *Review of Economics and Statistics*, 70(1989): 668-75.

Hanemann, M. " Valuing the Environment through Contingent Valuation," *Journal of Economics Perspectives*, 8(1994): 19-43.

Herriges, J. A., and C. L. Kling, "The Performance of Nested Logit Models When Welfare Estimation is the Goal," *American Journal of Agricultural Economics*, 79(1997): 792-802.

Houthakker, H., "Revealed Preference and the Utility Function," *Economica*, 17(1950): 159-174.

Kling, C. L., "The Importance of Functional Form in Estimation of Welfare," *Western Journal of Agricultural Economics*, 14(1989): 168-174.

Kling, C. L., and C. J. Thomson "The Implications of Model Specification for Welfare Estimation in Nested Logit Models," *American Journal of Agricultural Economics*, 78(1996): 103-14.

Larson, D. M., C. Klotz, and Y. L. Chien, "Nonparametric Consistency Tests for Nonmarket Data," manuscript, University of California, Davis, (1991).

Morey, E., R. Rowe, and M. Watson. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing," *American Journal of Agricultural Economics*75(1993): 578-592.

Ozuna, T., L. Jones, and O. Capps, "Functional Form and Welfare Measures in Truncated Recreatioin Demand Models," *American Journal of Agricultural Economics*, 75(1993): 1030-1035.

Richter, M., "Revealed Preference Theory," *Econometrica*, 34(1966): 635-645.

Sakong, Y., and D. J. Hayes, "Testing the Stability of Preferences," *American Journal of Agricultural Economics*, 75(1993): 269-277.

Samuelson, P., "Consumption Theory in Terms of Revealed Preference," *Economica*, 15(1948): 243-253.

Varian, H. R., "The Nonparametric Approach to Demand Analysis," *Econometrics*, 50(1982): 945-973.

Varian, H.R., "Nonparametric Tests of Consumer Behavior," *Review of Economic Studies*, 50(1983): 99-110.

Varian, H.R., "Nonparametric Tests of Models of Investor Behavior," *Journal of Financial and Quantitative Analysis*, 18(1983): 269-278.

Varian, H.R., "The Nonparametric Approach to Production Analysis," *Econometrica*, 52(1984): 579-597.

Varian, H. R., "Nonparametric Analysis of Optimizing Behavior with Measurement Error," *Journal of Econometrics* 30(1985): 445-458.

Ziemer, R. F., W.N. Musser and R. C. Hill, "Recreation Demand Equations: Functional Form and Consumer Surplus," *American Journal of Agricultural Economics*, 62(1980): 136-41.