Bias Correction of Welfare Measures in Non-Market Valuation: Comparison of the Delta Method, Jackknife and Bootstrap

Rui Zhang and J. Scott Shonkwiler

Department of Agricultural and Applied Economics

University of Georgia

rui.zhang25@uga.edu

Selected Paper prepared for presentation at the 2017 Agricultural & Applied Economics Association Annual Meeting, Chicago, Illinois, July 30 - August 1

ABSTRACT

Welfare measures in non-market are biased when they are nonlinear functions of parameter maximum likelihood estimates and when the sample size is small. This paper compares the performances of the delta method, jackknife, and bootstrap to correct this bias. The performance of the approaches is compared using Monte Carlo simulation for models including Poisson, logit, probit, and misspecified probit. The results indicate that the delta method and jackknife can correct the bias of welfare measures for the models listed in small samples. They have similar expected bias as well as mean square error in the single parameter case. With distributional misspecification, the delta method outperforms the jackknife. The bootstrap result has more expected bias, but also the smallest mean square error, especially when the sample size is large enough.

1. INTRODUCTION

1.1 Background

Welfare measures based on observed behavior or stated preferences are rarely estimated in a straightforward fashion. Haab and McConnell (2002) have summarized this as "The need for statistical inference and econometrics arises because individual actions, whether behavior that is observed in quasi-market settings or responses to hypothetical questions, almost never reveal precisely the economic value that a researcher wishes to measure."

Bias of welfare measures can arise at the data collection stage (i.e. hypothetical bias) or the model estimation stage. Although the former bias source has been extensively discussed, there is far less attention about statistical methods of bias correction during and after parameter estimation in non-market valuation. Welfare measures are typically biased because they are calculated as non-linear function of estimated parameters. Besides, when the sample size is small, even a consistent parameter estimate can be biased. Unlike the hypothetical bias being conceptual, this statistical bias can be rigorously identified and corrected through statistical techniques.

1.2 Motivation and Major Objectives

Several bias correction estimators have been proposed for welfare measures using the delta method, jackknife, bootstrap and distributional-free nonparametric method. In practice, the delta method is the most commonly used bias correction method, but the reliability of these methods has not been validated and compared. The delta method can be difficult to apply in that the statistical differentials (for univariate), or variance-covariance matrix (for multivariate) specifications are not readily available. And for many large and complicated models of the type encountered in non-market valuation, the bootstrap is not a feasible alternative. Thus it should be in the practitioner's interest to see if the jackknife can be a reliable alternative for the delta method.

Most of the statistical literature in this field compares the asymptotic properties of bias correction estimators for simple statistics like sample mean or correlation coefficient with these bias correction methods (Efron,1977,1981,1983; Parr,1983; Chow, 1985; Withers, 2013, 2014), which cannot be directly applied to models used in non-market valuation, especially for small sample estimation. Only a few studies compared these methods for welfare measures in non-market valuation (Cooper, 1994; Langford, 1998; Bliemer, 2013). These studies used Monte Carlo simulation to show the small sample performance of bias correction methods, but they only used limited models with little variation of circumstances, which can hardly represent versatile cases in practice with different model specifications, sample sizes, and possible distributional misspecification. Besides, in these studies the Monte Carlo simulation results have not been linked to the asymptotic properties of these bias correction methods, and no mathematic explanations for the simulation results were proposed, although there is extensive theoretical literature about resampling methods.

We compare the performance of these bias correction methods for welfare measures in non-market valuation using Monte Carlo simulation with variations of sample size, model specification, and distributional assumption. The conclusions of this study can provide some rules of thumb for the practitioner to select the optimal bias correction method.

2. LITERATURE REVIEW

2.1 Bias Correction Method

In CV models the welfare measures are often obtained by nonlinear transformation of estimated parameters, which are known to be biased due to nonlinearity. Also, as summarized by Chen (2012), the maximum likelihood parameter estimates can be biased in finite samples. A problem of interest is to reduce the bias as well as MSE of such estimators. The most common way is to apply Taylor's series expansion to analytically derive bias at the $O(n^{-1})$ corrected estimator, namely the Delta method. Oehlert (1992) gave a review about the Delta method and concluded that the Taylor series approximation for expected value of the function is consistent provided that the underlying sequences of random variables have enough bounded moments.

Alternatively, resampling procedures like jackknife and bootstrap can also remove the $O(n^{-1})$ bias, which entails a negligible increase in variance as shown by Quenouille (1956), or even decreases the variance as shown by Durbin (1959). Besides, Efron (1979) showed that the jackknife is a linear approximation for the bootstrap. The "delete-one" jackknife was extended by Jaeckel (1972) to the infinitesimal jackknife, and by Gray (1975) and Sharot (1976) to the generalized jackknife. Gray (1975) and Efron (1979) pointed out that the bias and variance expressions suggested by the infinitesimal jackknife transform equivalently to the Delta method, and the Delta method is a limiting case of a generalized jackknife. Kim (1998) derived conditions under which bias corrected estimates for smooth function and non-smooth function statistics with the jackknife and bootstrap method lead to increase in MSE in terms of second-order asymptotic superiority, which can be extended to regression models. Withers (2013) assessed through a simulation study the performance of the Delta method bias corrected estimators for odds ratio in the Binomial (n, θ), including the maximum likelihood estimator, the delta estimators up to the third order. Bias and MSE of the four estimators are computed in 10,000 simulations. The steps are repeated for sample size 1,2, … ,100, and θ = 0.3, 0.5, 0.7, 0.9. Results showed that the delta method could reduce both bias and MSE in multinomial odds ratio estimation, and the results are consistent under the different value of θ.

2.2 Comparison of Bias Correction Methods

Efron (1981) used bootstrap, jackknife and infinitesimal jackknife and other nonparametric methods in a Monte Carlo experiment for calculating the correlation coefficient. The point estimates, variances, and MSEs are compared. The simulation results showed that jackknife has less bias than the delta method; the bootstrap generally performs the best among all methods at the cost of less computational efficiency.

Parr (1983) showed that the Delta method, jackknife, and bootstrap coincide at least to the first order term, and the absence of skewness leads to the second-order coincidence of the jackknife and delta methods.

Shein-Chunq Chow (1985) studied the Delta method, jackknife and bootstrap estimators for nonlinear functions of local parameters and derived asymptotic expansions of the mean squared error up to the third order. There is no consistent relationship among these estimators; the performance depends on the underlying distribution and particular form of the nonlinear function. A least MSE estimator was proposed and compared with these three estimators, in simulations for several commonly used nonlinear functions, logx, 1/x, and $\sqrt{x}$ of sample mean estimates, and distributions including gamma, Poisson, and lognormal, with sample size n=10 and n=20. A total of 1000 runs were made for each combination. This study provides a direct basis for our study, but the sample size is too small for any non-market valuation practice, and only a nonlinear function of a single parameter estimate was studied. So we will extend to nonlinear functions of more than one estimated parameter, and increase the simulation sample size range, as well as functional forms, according to the importance and usage frequency in non-market valuation practice.

The comparison of these methods in the field of non-market valuation was firstly studied by Cooper (1994). A Monte Carlo simulation technique is used to assess the reliability of the Krinsky and Robb, jackknife, bootstrap, and the Cameron confidence interval routines as applied to the logit regression model with simulated sample size N=100, N=500 and N=1000. The simulated data sets are created using both the logistic and Weibull distributions. Thus the robustness of these methods under misspecification of skewed distribution can be tested. Bias, standard deviation, and frequency of inclusion of true mean in the estimated confidence interval were used as comparison criterions. All methods produced similar results with a larger sample, but for N=1000 and logistic distribution, bootstrap performs the best. An analytical approach like the Cameron method performs the best when N=100, but tends to underestimate the standard deviation. Under falsified Weibull distribution, all methods perform worse than when the logistic is the true underlying distribution, but the jackknife performs the best when N=500 and N=1000.

Langford (1998) compared bootstrap and delta method bias corrected estimations of WTP median in a logit model and showed that with bid level effect included (for random-effect model), the bootstrap result is more skewed and with more kurtosis than the Delta method results, and the variances of two results are also significantly different.

Withers (2014) compared the computational efficiency of bias reduction estimators using the delta method, jackknife, and bootstrap. It is shown that the p_{th} order delta estimator requires ~n calculations, the p_{th} order "drop-one-observation" jackknife estimator requires ~n calculations if p = 1 and ~n^{p-1} calculations if p ≥ 2, the p_{th} order "drop-m-observation" jackknife estimator requires ~n^{ms+1} calculations or ~n^{ms} calculations if s = p - 1 > 0; and bootstrap is second order and requires ~n^2 calculations.

3. METHOD

3.1 Bias of welfare measures from parameter estimation and nonlinear transformation

In the process of obtaining welfare measures in non-market valuation practices, the findings can be biased for several reasons. Firstly, parameter estimation bias, namely the estimation results of underlying parameters are biased due to the insufficient sample size. This type of bias can be alleviated by applying the jackknife method. In non-market valuation practice, this source of bias is entangled with other sources of bias, which complicates the analysis of the performance of bias correction methods. That is a reason why we tend to rely on simulation results instead of an analytical approach when comparing bias correction methods.

Secondly, non-linear transformation bias, namely the welfare measures obtained by unbiased parameter estimation can still be biased because the welfare measure is usually calculated as a non-linear function of the estimated parameters, and it is well-known that the expected value of a non-linear function of unbiased parameter estimates is itself generally biased. Welfare measures in non-market valuation are usually derived with a nonlinear function of parameters. If $\boldsymbol{\beta}$ is a vector of k parameters, under maximum likelihood estimation

$$\widehat{\boldsymbol{\theta}} \ \dot{\sim} \ N(\boldsymbol{\theta}, V(\boldsymbol{\theta})) \tag{1}$$

where $\dot{\sim}$ means approximately distributed as, $\mathrm{V}(\boldsymbol{\theta})$is the variance-covariance matrix. Assume welfare measure $W = h(\boldsymbol{\theta})$, for which the usually estimator is $\widehat{W} = h(\widehat{\boldsymbol{\theta}})$. Here $\widehat{\boldsymbol{\theta}}$ usually takes the form of sample mean $\overline{\boldsymbol{\theta}}$**.** When h is a nonlinear function, the function of parameter estimates $h(\widehat{\boldsymbol{\theta}})$ will be biased estimate for WTP. If h is a convex function, then $E\left(h(\widehat{\boldsymbol{\theta}})\right) > E(\widehat{h(\boldsymbol{\theta})})$; if h is a concave function, then $E\left(h(\widehat{\boldsymbol{\theta}})\right) < E(\widehat{h(\boldsymbol{\theta})})$. With sample size increasing, the expected bias will decrease, but will always have the same sign.

Thirdly, distributional misspecification bias, namely the data generating process assumed by the statistical model can be different from the economic model specified. When applying random utility models, we derive the model specification by taking up a distribution of the random part of the utility or corresponding probability distribution. For instance, the logit model with logistic probability distribution is obtained by assuming that each random part of utility is independent, identically distributed extreme value, namely Gumbel and type I extreme value distribution; while a probit model is used with normal distribution specified.

Common probability distributions are the normal, logistic, log-logistic, Weibull and exponential (Kerr, 2000). With real-world datasets, the difficulty is that the underlying true distribution cannot be revealed even if we compare results with different possible distributional specifications. However, simulated data can be generated by setting up a distribution, and by comparing results under “true” and alternative models for the influence of distributional misspecification. The effects of skewness or kurtosis from the “true” distribution can be driving the bias in different ways.

We will illustrate these sources of bias with frequently used models in non-market valuation practice as examples, including Poisson and negative binomial for travel cost models and binary probit and logit for random utility models. We will also consider more complicated models including double bounded models, random effect models, and scaled or drifted distributional specifications as suggested by Kerr (2000). We will show in simulations how large and significant the bias can be with different combinations of sources of bias as well as under different sample size.

These different sources of bias may become the major driving factors for different models at various sample sizes. Basically the parameter estimation bias is prominent when sample size is small, and it can also complement the nonlinear transformation bias at first order. As shown in (33), when the sample size is large enough to neglect the parameter estimation bias, the nonlinear transformation bias becomes the primary source of bias. The asymptotic result can be seen as the case when sample size approaches infinity, and the consistent parameter estimation has zero bias. Although in practice we care more about small sample results than asymptotic results, the asymptotic result can help explain the dynamics of these driving factors.

3.2 Bias Correction Methods for Parameter Nonlinear Transformation

3.2.1 Analytical Bias Correction Method for Maximum Likelihood Parameter Estimation

The maximum likelihood parameter estimation bias can be corrected by computing the bias of parameter estimate at first or higher order, and then subtracting the bias from the original estimate. Here we give several examples of correcting the bias of the ML estimator $\widehat{\boldsymbol{\beta}}$.

For a logit model, assume $Pr(y=1)=\pi$, with which a set of covariates $\mathbf{X}$ is associated. And the log odds ratio is a linear function of $\boldsymbol{X}$, specified as $\boldsymbol{X}'\boldsymbol{\beta}$. To estimate $\widehat{\boldsymbol{\beta}}$, we need to solve the equation

$$\frac{\partial l(\boldsymbol{\beta})}{\partial \beta_r} \stackrel{\text{def}}{=} g(\beta_r) = \sum (y_i - \hat{\pi}_i)\, \boldsymbol{X_{ir}} = 0 \qquad (2)$$

where $l(\boldsymbol{\beta})$ is the log-likelihood function, and $g(\beta_r)$ is the score function. In matrix notation $\boldsymbol{\pi} = (1+e^{-\boldsymbol{X}'\boldsymbol{\beta}})^{-1}$, and $g(\boldsymbol{\beta}) = \boldsymbol{X}'(y - \widehat{\boldsymbol{\pi}})$. The fisher information $\mathrm{F}(\boldsymbol{\beta}) = \boldsymbol{X}'\boldsymbol{W}\boldsymbol{X}$, where $\boldsymbol{W} = diag(\widehat{\boldsymbol{\pi}}(1-\widehat{\boldsymbol{\pi}}))$. Then $\widehat{\boldsymbol{\beta}}$ is obtained by applying the iterative Newton-Raphson method:

$$\boldsymbol{\beta}^{t+1} = \boldsymbol{\beta}^{t} + F(\boldsymbol{\beta}^{t})^{-1}\, g(\boldsymbol{\beta}) \qquad (3)$$

where β^t is the estimate at t-th stage. This step is repeated until convergence criterion is reached. This method is called iterative maximum likelihood estimates.

Firth (1993) suggested that the bias of MLE can be reduced by modifying the score function (41) as

$$g_f(\beta_r) = \sum ((y_i - \hat{\pi}_i) + h_i(0.5 - \hat{\pi}_i))\boldsymbol{X_{ir}} \qquad (4)$$

where h_i is the i-th diagonal element of the matrix $\mathbf{H} = W^{\frac{1}{2}}X(X'WX)^{-1}X'W^{\frac{1}{2}}$. Then use $g_f(\beta_r)$ for the iterative Newton-Raphson method.

Cordeiro (1991) proposed a bias correction estimator for the generalize linear model

$$\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}} - (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}\boldsymbol{\xi} \tag{5}$$

For logit model $\boldsymbol{\xi} = diag[\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'](\widehat{\boldsymbol{\pi}} - 0.5)$.

3.2.2 Delta Method

The Delta method can correct the nonlinear transformation bias by approximating the nonlinear function with a Taylor series. When the function h has continuous second derivatives, by taking a second order Taylor series approximation of $h(\widehat{\boldsymbol{\theta}})$ from (1), we get

$$h(\widehat{\boldsymbol{\theta}}) \cong h(\boldsymbol{\theta}) + \sum_{i=1}^{k} E\big[(\hat{\theta}_i - \theta_i)\big]\frac{\partial h(\boldsymbol{\theta})}{\partial \theta_i} + \frac{1}{2}\sum_{i,j=1}^{k} V_{ij}(\widehat{\boldsymbol{\theta}})\frac{\partial^2 h(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \tag{6}$$

Taking expectations of both sides, we get a bias corrected estimator

$$\widehat{W}_D \cong \mathrm{E}\big[h(\widehat{\boldsymbol{\theta}})\big] - \sum_{i=1}^{k} E\big[(\hat{\theta}_i - \theta_i)\big]\frac{\partial h(\widehat{\boldsymbol{\theta}})}{\partial \hat{\theta}_i} - \frac{1}{2}\sum_{i,j=1}^{k} V_{ij}(\widehat{\boldsymbol{\theta}})\frac{\partial^2 h(\widehat{\boldsymbol{\theta}})}{\partial \hat{\theta}_i \partial \hat{\theta}_j} \tag{7}$$

This first-order approximation is usually referred to as the delta method. Although equation (34) can be evaluated at every data point, typically interest is focused on the sample mean $\mathrm{h}(\overline{\boldsymbol{\theta}})$. The delta method can eliminate the 1/n bias leading term from the nonlinear function (Miller, 1964).

If $\hat{\theta}_i$ is treated as an unbiased estimator of θ_i, then the first order term in (34) can be eliminated, then we get the most commonly used version of delta method estimator. For small samples, the first order bias is prominent, which weakens the performance of delta method if this term is omitted. For certain models, the bias of $\hat{\theta}_i$has been derived, so that a more precise result can be attained.

3.2.3 Jackknife

The basic idea behind the jackknife bias corrected estimators lies in recomputing the statistic of interest using all but one of the observations. Suppose $\widehat{\boldsymbol{\theta}}$ is estimated from a dataset $X_1, X_2, X_3, \dots, X_n$. And welfare measure $\widehat{W} = h(\widehat{\boldsymbol{\theta}})$, denote

$$\widehat{W}_{-i} = h(\widehat{\boldsymbol{\theta}}_{-\boldsymbol{i}}) \tag{8}$$

Where $\widehat{\boldsymbol{\theta}}_{-\boldsymbol{i}}$ is estimated with all but i_{th} observation X_i for $i = 1,2, \dots, \mathrm{n}$. Denote the i_{th} pseudo-value by $\widehat{W}_i = n\widehat{W} - (n-1)\,\widehat{W}_{-i}$. Then the jackknife bias corrected estimator is

$$\widehat{W}_J = \frac{1}{n}\sum_{i=1}^{n}\widehat{W}_i = n\,h(\widehat{\boldsymbol{\theta}}) - \frac{n-1}{n}\sum_{i=1}^{n}h(\widehat{\boldsymbol{\theta}}_{-\boldsymbol{i}}) \tag{9}$$

The Jackknife can also be defined as a form of weighted observations. If we attach a weight to each observation, then leaving one observation out is the same as giving an observation a weight of zero. Instead of doing that for ordinary jackknife, we give the observation only a slightly less weight than the others, and then we get the infinitesimal jackknife. We assign weights $w_1, w_2, w_3, \dots, w_n$ to observations $X_1, X_2, X_3, \dots, X_n$ respectively. Then $\widehat{W}$ can be written as a function of 2*n variables $T(\,X_1, X_2, X_3, \dots, X_n; w_1, w_2, w_3, \dots, w_n\,)$, since all observations have equal weight of 1/n, we get

$$\widehat{W} = T(\,X_1, X_2, X_3, \dots, X_n; \frac{1}{n}, \dots, \frac{1}{n}) \tag{10}$$

For $\widehat{W}_{-i}$, the weight of X_i becomes 0, and all other observations have same weight, so we have

$$\widehat{W}_{-i} = T(\,X_1, X_2, X_3, \dots, X_n; \frac{1}{n-1}, \dots, \frac{1}{n-1}, 0, \frac{1}{n-1}, \dots, \frac{1}{n-1}\,) \tag{11}$$

To simplify the algebra, we extend the definition of T as follows. If G is any probability distribution for which T is defined, and c is a positive constant, then we let $T(cG) =$

$T(G)$.Thus, if we are considering discrete distributions with possible values X_i, i=1, …, n. we can assign to each X_i the weight w_i without requiring that $\sum_{i=1}^{n} w_i = 1$.

If we reduce w_i by ϵ and let $\epsilon \to 0$. We have

$$\widehat{W}_{-i}(\epsilon) = T(\,X_1, X_2, X_3, \dots, X_n; \frac{1}{n}, \dots, \frac{1}{n}, \frac{1}{n} - \epsilon, \frac{1}{n}, \dots, \frac{1}{n}\,) \tag{12}$$

Let $\widehat{D}_i = \frac{\partial T}{\partial w_i}$, $\widehat{D}_{ii} = \frac{\partial^2 T}{\partial {w_i}^2}$, since $T(cG) = T(G)$, it can be proved that $\frac{1}{n}\sum \widehat{D}_i = 0$. We can form the Taylor series expansion

$$\widehat{W}_{-i}(\epsilon) - \widehat{W} = T\left(\frac{1}{n}, \dots, \frac{1}{n} - \epsilon, \dots, \frac{1}{n}\right) - T\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = -\epsilon\widehat{D}_i + \frac{\epsilon^2}{2}\widehat{D}_{ii} - \cdots \tag{13}$$

Suppose $\widehat{W}_{(\cdot)}(\epsilon) = \frac{1}{n}\sum \widehat{W}_{-i}(\epsilon)$, we can define a bias estimate $\hat{B}(\epsilon)$ by

$$n^2\epsilon^2\hat{B}(\epsilon) = n(1-\epsilon)(\widehat{W}_{(\cdot)}(\epsilon) - \widehat{W}) \tag{14}$$

Letting $\epsilon \to 0$, we have

$$\hat{B}(0) = \frac{1}{2n^2}\sum \widehat{D}_{ii} \tag{15}$$

Finally, we get the infinitesimal jackknife estimate

$$\widehat{W}_{IJ} = \widehat{W} - \hat{B}(0) = \widehat{W} - \frac{1}{2n^2}\sum \widehat{D}_{ii} \tag{16}$$

which is exactly the result of applying delta method to $\widehat{W}$(Gray,1975).

The usage of jackknife can be extended to many aspects of estimation. Kézdi (2002) proposes a jackknife minimum distance estimator designed to reduce the finite-sample bias of the optimal minimum distance estimator. Lee (2012) showed that the jackknife maximum likelihood estimator is consistent, and the jackknife estimate of the log likelihood is asymptotically unbiased.

3.2.4 Bootstrap

An alternative resampling method, the bootstrap procedure can also be used for bias correction in welfare measures with nonlinear transformation. The procedures are as follows:

Step 1. Construct the empirical probability distribution $\widehat{F}$ of each dataset with n observations by putting probability mass n^{-1} on each observation X_i.

Step 2. Draw a random sample of size n from $\widehat{F}$ as $X_1^*, X_2^*, X_3^*, \dots, X_n^* \sim \widehat{F}$, and get $\widehat{\boldsymbol{\theta}}^* =$ f($\bar{X}^*$), where $\bar{X}^* = \frac{1}{n}\sum_{i=1}^{n} X_i^*$.

Step 3. Repeat step 2 with a large number T times, get $\widehat{\boldsymbol{\theta}}_{\mathbf{1}}^*, \widehat{\boldsymbol{\theta}}_{\mathbf{2}}^*, \widehat{\boldsymbol{\theta}}_{\mathbf{3}}^*, \dots, \widehat{\boldsymbol{\theta}}_{\boldsymbol{T}}^*$. With $\widehat{W} = h(\widehat{\boldsymbol{\theta}})$, get $\widehat{W}_1^*, \widehat{W}_2^*, \widehat{W}_3^*, \dots, \widehat{W}_B^*$. Then we get the estimate of bias

$$\widehat{bias}(\widehat{W}) = \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{W}_t^* - \widehat{W}\right) \qquad (17)$$

So the bootstrap bias-corrected estimator is

$$\widehat{W}_B = 2\widehat{W} - \frac{1}{T}\sum_{t=1}^{T}\widehat{W}_t^* \qquad (18)$$

The Bootstrap can take different forms by various schemes of the random drawing, here for simplicity, we only consider the most commonly used method.

3.3 Comparison of bias correction methods through Monte Carlo simulation

Monte Carlo Simulation analysis is started by assuming a distribution, which can be the true underlying distribution of the model to be estimated or a wrong distribution. The reason for the distributional divergence is to assess the bias under distributional misspecification. Then a large number S simulated datasets are created based on the assumed distribution. With each simulated dataset, $\widehat{W}_D, \widehat{W}_J$ and $\widehat{W}_B$ are calculated. With these S estimators based on the simulated distribution of $\widehat{W}$, expected bias, variance, mean square error, confidence intervals can be calculated and compared for $\widehat{W}_D, \widehat{W}_J$ and $\widehat{W}_B$.

Simulations need to be run under a wide range of variations to represent the possible alternative cases. We believe the key variations include: model specification/function form; sample size; distributional assumption.

Firstly, different bias correction methods may have optimal performance with different model specifications. For instance, the welfare measures may be nonlinear functions of one parameter, multiple parameters, or estimated parameters with observed values. We will start with commonly used models like logit, probit, and Poisson; and extend to more advanced ones like double bounded probit and the Box-Cox model.

Secondly, different bias correction methods may perform best under different sample size. For instance, as the sample size in each repetition increases, the variance of estimates in each repetition gets smaller, and the nonlinear transformation becomes closer to linear when it happens within a narrower interval so that the nonlinear transformation bias becomes smaller when sample size increases. On the other hand, the delta method requires sample size to be above a certain threshold to produce a robust result. However, for the jackknife, the performance has no such correlation to sample size. The simulation will be done under sample sizes of 50,100,200 and 400.

Thirdly, under distributional misspecification, the welfare measure will be further biased, for which different bias correction methods may show varying levels of robustness. The misspecification may be about skewness as in Weibull /logistic, or kurtosis as in normal/logistic. Besides, several nonparametric/semi-nonparametric methods can be alternative solutions for distributional misspecification, although they are in general less efficient than MLE.

4 SIMULATION RESULT

In this section, a Monte Carlo simulation technique is applied to compare the performances of the delta method, jackknife, and bootstrap. Simulated datasets are generated with underlying model assumed and true parameters set. And then we use the true or misspecified model to estimate the parameters. Each model was then estimated by applying each of the bias correction methods, and the results are compared in terms of expected bias as well as MSE. The models used include Poisson, logit, misspecified probit, and double bound probit.

4.1 Poisson

The travel cost model of recreational demand can take the form of a Poisson specification

$$\lambda = e^{\beta_0 + \beta_1 x} \tag{19}$$

with λ representing the number of trips to a recreation site and the variable x representing travel costs. If we assume $\beta_0 = 2$, $\beta_1 = 0.25$, and x being a positive valued uniform variable with mean 4 and minimum 1.5. Then $-1/\beta_1$ is the per-trip consumer surplus, the true value of which under the population model should be 4.

The model is simulated 1,000 times for each estimation. Then the 1000 outcomes are treated as a new sample with 1,000 observations, with which the mean, mean of bias, standard deviation, and MSE are calculated. For standard deviation we use the one calculated over 1,000 simulations, instead of the average of the 1,000 standard deviations computed in each simulation; because what we are interested is the precision and reliability of the bias estimates, instead of the precise estimates of standard deviation.

Using (34), we derive the delta method estimator as

$$\widehat{W}_D \cong -\frac{1}{\beta_1} + \frac{V(\beta_1)}{{\beta_1}^3} \tag{20}$$

For the bootstrap in each simulation we use a sample with the same size as the original one, so that it can remain at a comparable computational efficiency level as another method. The estimates without any bias correction procedure are summarized in “original” row.

Here the point estimate results (with true value being 4) is listed in TABLE 1, and standard deviation and MSE results are listed in TABLE 2. The “original” bias declines when sample size increases. The nonlinear transformation always caused the estimates to be biased upwards because of the convexity of the function $-\frac{1}{\beta_1}$ with β_1 being negative.

The delta method and jackknife both decrease bias as well as a standard deviation with sample size level 50~1000, and they have very close results at all sample size levels. A converging trend can be observed for the bias to shift from downwards to upwards when sample size increases over 100. The jackknife has smaller bias roughly with sample size from 70~100; while the delta method has less bias roughly with sample size above 200. The reason for the superiority of delta method at larger sample size may be that the primary source of bias is the bias of $\widehat{\beta_1}$ when sample size is small; when sample size is large enough, the bias of $\widehat{\beta_1}$ become negligible compared to the second order expansion of $\widehat{W}_D$ as calculated in (50). When the standard deviation from these two methods are compared, consistent results cannot be reached at different sample size; so it cannot with the MSE. Thus by the standard of MSE, the delta method and the jackknife can be seen similar to each other for the Poisson model when sample size is below 1000.

The reason for the delta method and jackknife to have similar performance may be that the delta method is equivalent to infinitesimal jackknife, which is a smoothed version of the ordinary leave-one-out jackknife. The gap between the delta method and the jackknife can be seen more clearly when the sample size is large enough. In spite of this difference, when the precision requirement is not so high, or when the delta method is not convenient to apply due to a complex second derivative, the jackknife can be an acceptable alternative for delta method for Poisson model.

The bias of the bootstrap result is larger than the original result at sample size 50,70, and 1000. And when the sample size is sufficiently large, the bootstrap estimate has the smallest MSE, but still more substantial bias. One advantage of the bootstrap is that even for small sample estimation, we can always improve the precision of the bootstrap by increasing of random draws in each simulation. And it is noticeable that the sign of bias is always upward.

TABLE 1

MONTE CARLO ANALYSIS COMPARING SEVERAL METHODS FOR POINT ESTIMATE OF WELFARE MEASURE FOR POISSON MODEL

Sample size	**50**	**70**	**100**	**200**	**400**	**1000**
Original	4.224	4.185	4.137	4.080	4.048	4.013
Delta	3.840	3.964	3.994	4.016	4.017	4.001
Jackknife	3.830	3.967	3.999	4.019	4.019	4.002
Bootstrap	4.270	4.185	4.083	4.051	4.015	4.015

TABLE 2

MONTE CARLO ANALYSIS COMPARING SEVERAL METHODS FOR VARIANCE & MEAN SQUARE ERROR OF WELFARE MEASURE FOR POISSON MODEL

Sample size	**50**	**70**	**100**	**200**	**400**	**1000**
STD.DEV						
Original	1.355	0.938	0.781	0.512	0.354	0.215
Delta	0.858	0.762	0.692	0.488	0.346	0.214
Jackknife	0.870	0.764	0.693	0.487	0.346	0.214
Bootstrap	1.316	1.080	0.722	0.513	0.337	0.212
MSE						
Original	1.8860	0.9142	0.6284	0.2688	0.1272	0.04651
Delta	0.7613	0.5820	0.4795	0.2381	0.1198	0.04558
Jackknife	0.7859	0.5842	0.4800	0.2380	0.1199	0.04560
Bootstrap	1.8048	1.1999	0.5283	0.2654	0.1139	0.04508

4.2 Logit

A binary model is assumed with a binary choice variable y representing whether to accept the bid, with x representing conditioning variables, and A representing the bid amount offered. By assuming the latent unobserved WTP as either logistic or Weibull, we can get the true data generating process specified as, respectively

$$Prob(y=1) = \frac{1}{1+e^{\beta_0+\beta_1 x+\beta_2 A}} \tag{21}$$

$$Prob(y=1) = e^{-e^{\beta_0+\beta_1 x+\beta_2 A}} \tag{22}$$

If we assume logistic distribution as in (46), the expected WTP is

$$E(WTP|x) = \frac{\beta_0+\beta_1\bar{x}}{-\beta_2} \tag{23}$$

where $\bar{x}$ is the sample mean of x. If we assume WTP is non-negative, then

$$E(WTP|x) = \frac{\ln\left(1+e^{\beta_0+\beta_1\bar{x}}\right)}{-\beta_2} \tag{24}$$

If we assume Weibull distribution as in (55), the expected WTP is

$$E(WTP|x) = \int_0^\infty [1-e^{e^{\beta_0+\beta_1+\beta_2 A}}]dA \tag{25}$$

For Monte Carlo, we generate simulated data using both (54) (55), and fit the data using the logit model, so the estimates for the Weibull dataset can be assumed as a distributional misspecification case. We assume $\beta_0 = 1.4136$, $\beta_1 = -0.008561$, and $\beta_2 = 0.00372$. Consumption variable x is generated with random drawing from a normal distribution with mean 95.78 and standard deviation 119.226. A is generated by a repeating sequence (2.5, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 200, 250, 300, 400, 500, 700). The dependent variable y is generated by estimating $P_i = \widehat{Prob}(y_i = 1)$ for each observation with (54) and (55), and using a Bernoulli distribution with parameter P_i to generate a "1 or 0" response. Following these procedures, for logistic and Weibull distribution, the true $E(WTP)$ are $206.74 and $274.22. For non-negative WTP restriction model with logistic distribution, $E(WTP)$is $225.12. Because x is randomly generated, for every simulation we plug the $\bar{x}$ calculated from the generated x into (56) (57) (58) to get the distribution of true WTP for each sample size.

For simulated datasets, we choose sample sizes of 40, 80, 100, 200, 400 to ensure an equal number of each bid amount is included for each sample size. For each sample size, the WTP estimate is derived with direct nonlinear parameter transformation, delta method, and jackknifing. For each analysis, 1,000 simulations are conducted to obtain 1,000 point estimates of WTP forming up the simulated distribution, of which the mean, expected bias, standard deviation, MSE are reported.

For each sample size, there is a separate true mean because although the $\beta_i, i = 1,2,3$ do not change with sample size, x changes for each random draw at different sample size. Thus for each sample size the expected $\overline{W}$ is calculated as the true mean, and then the bias for each method at each sample size is calculated by $\widehat{W}_i - \overline{W}, i = D, J, B$.

From (30), the delta method estimator is

$$\widehat{W}_D \cong \frac{\widehat{\beta_0} + \widehat{\beta_1}\bar{x}}{-\widehat{\beta_2}} - \frac{Cov(\widehat{\beta_0}, \widehat{\beta_2}) + Cov(\widehat{\beta_1}, \widehat{\beta_2})\bar{x}}{\widehat{\beta_2}^2} + \frac{(\widehat{\beta_0} + \widehat{\beta_1}\bar{x}) * V(\widehat{\beta_2})}{\widehat{\beta_2}^3} \quad (26)$$

TABLE 3

MONTE CARLO ANALYSIS COMPARING SEVERAL METHODS FOR POINT ESTIMATE OF WELFARE MEASURE FOR LOGIT MODEL

Sample size	40	80	100	200	400	1000
MEAN						
True	206.736	206.932	206.694	206.916	206.721	206.766
Original	213.014	207.411	208.778	208.326	206.723	207.160
Delta	199.452	202.043	204.464	206.215	205.687	206.744
Jackknife	204.879	206.682	208.212	208.133	206.643	207.129
Bootstrap	207.993	207.413	207.609	207.308	207.314	206.489
BIAS						
Original	6.278	0.479	2.084	1.411	0.002	0.394
Delta	-7.284	-4.889	-2.230	-0.701	-1.035	-0.022
Jackknife	-1.857	-0.249	1.518	1.217	-0.079	0.363
Bootstrap	1.258	0.481	0.916	0.392	0.592	-0.277

Note: true values are different for each sample size due to randomly drew variable x.

TABLE 4

MONTE CARLO ANALYSIS COMPARING SEVERAL METHODS FOR VARIANCE & MEAN SQUARE ERROR OF WELFARE MEASURE FOR LOGIT MODEL

Sample size	**40**	**80**	**100**	**200**	**400**	**1000**
STD.DEV						
Original	66.723	39.652	36.732	24.388	16.674	10.536
Delta	57.756	37.852	35.364	23.971	16.537	10.502
Jackknife	101.144	38.684	35.948	24.178	16.612	10.521
Bootstrap	58.159	38.548	35.118	23.715	16.733	11.075
MSE						
Original	4491.311	1572.489	1353.598	596.766	278.026	111.162
Delta	3388.809	1456.645	1255.567	575.081	274.536	110.296
Jackknife	10233.516	1496.499	1294.593	586.045	275.955	110.829
Bootstrap	3384.098	1486.141	1234.144	562.536	280.346	122.722

4.3 Misspecified Probit

To test the robustness of bias correction methods to distributional misspecification, we randomly generated the data using logistic assumptions, but estimate the data using a probit model, The simulation and bias correction procedures are essentially the same as what were done with the logit in the last section, except the misspecification, and we only used the delta method and jackknife for bias correction.

As is shown in the simulation result the misspecification results in inconsistent estimation. Even at the sample size of 1000, there is still systematic bias after any bias correction. However, in general, the delta method is more robust to this misspecification. And for jackknife, in one simulation the estimate is so severely biased that it has a tremendous effect on the average value. In a future study, we will use different misspecification with multiple scales of deviation of skewness and kurtosis to compare the performances of bias correction methods.

TABLE 5

MONTE CARLO ANALYSIS COMPARING SEVERAL METHODS FOR POINT ESTIMATE OF WELFARE MEASURE FOR MISSPECIFIED PROBIT MODEL

Sample size	**40**	**80**	**100**	**200**	**400**	**1000**
MEAN						
True	206.733	206.931	206.696	206.916	206.718	206.766
Original	213.882	209.031	210.868	210.781	209.668	210.251
Delta	200.735	204.085	206.919	208.861	208.732	209.876
Jackknife	207.356 (210.231)	209.606	211.472	211.197	209.938	210.374
BIAS						
Original	7.149	2.100	4.172	3.865	2.950	3.485
Delta	-5.998	-2.846	0.224	1.945	2.014	3.110
Jackknife	0.620 (3.496)	2.674	4.779	4.282	3.217	3.609

Note: the value in the parenthesis under jackknife is the value calculated after eliminating one extreme estimate.

TABLE 6

MONTE CARLO ANALYSIS COMPARING SEVERAL METHODS FOR VARIANCE & MSE OF WELFARE MEASURE FOR MISSPECIFIED PROBIT MODEL

Sample size	**40**	**80**	**100**	**200**	**400**	**1000**
STD.DEV						
Original	66.728	39.513	36.844	24.480	16.803	10.542
Delta	58.260	37.815	35.565	24.096	16.680	10.511
Jackknife	106.997	38.719	36.277	24.327	16.768	10.538
	(56.52)					
MSE						
Original	4503.703	1565.662	1374.910	614.211	291.052	123.279
Delta	**3430.171**	**1438.069**	**1264.893**	**584.424**	**282.263**	**120.159**
Jackknife	11448.739	1506.292	1338.851	610.140	291.522	124.077
	(3206.719)					

Note: the value in the parenthesis under jackknife is the value calculated after eliminating one extreme estimate.

5 CONCLUSION

The delta method, jackknife and bootstrap can effectively correct the bias of welfare measures caused by nonlinear transformation in small samples (less than 400). When sample size is too small, the bias of parameters is ineligible and uncertain, which can disturb and weaken the bias correction method. When sample size is large enough (over 1000), the results converge to the asymptotic situation, which almost eliminates both the parameter estimation bias and nonlinear transformation bias, but not the systematic misspecification bias.

The delta method generally performs best in the sense that it usually has the smallest expected bias and MSE among these three, although the delta method and jackknife have similar performance for Poisson model. With distributional misspecification, the delta method can still effectively correct the bias, but the jackknife method cannot. When the model is correctly specified, jackknife can work as a substitute for the delta method given that it is very easy to apply and computationally inexpensive. Bootstrap performs the best when the sample is sufficiently large (more than 1,000), which is consistent with asymptotic result.

REFERENCE

Alberini, A. (1995). Efficiency vs Bias of Willingness-to-Pay Estimates: Bivariate and Interval-Data Models. *Journal of Environmental Economics and Management, 29*(2), 169-180. doi:http://dx.doi.org/10.1006/jeem.1995.1039

Alberini, A. (1995). Optimal Designs for Discrete Choice Contingent Valuation Surveys: Single-Bound, Double-Bound, and Bivariate Models. *Journal of Environmental Economics and Management, 28*(3), 287-306. doi:http://dx.doi.org/10.1006/jeem.1995.1019

Álvarez Díaz, M., González Gómez, M., Saavedra González, Á., & De Uña Álvarez, J. (2010). On dichotomous choice contingent valuation data analysis: Semiparametric methods and Genetic Programming. *Journal of Forest Economics, 16*(2), 145-156. doi:http://dx.doi.org/10.1016/j.jfe.2009.02.002

Amemiya, T. (1980). The n^{-2} Order Mean Squared Errors of the Maximum Likelihood and the Minimum Logit Chi-Square Estimator. *The Annals of Statistics, 8*(3), 488-505.

Amemiya, T. (1981). Qualitative response models: a survey. *Journal of Economic Literature, 19*, 1483-1536.

Araña, J. E., & León, C. J. (2005). Flexible mixture distribution modeling of dichotomous choice contingent valuation with heterogenity. *Journal of Environmental Economics and Management, 50*(1), 170-188. doi:http://dx.doi.org/10.1016/j.jeem.2004.05.009

Belluzzo, W. (2004). Semiparametric Approaches to Welfare Evaluations in Binary Response Models. *Journal of Business & Economic Statistics, 22*(3), 322-330.

Bengochea-Morancho, A., Fuertes-Eugenio, A. M., & Del Saz-Salazar, S. (2005). A comparison of empirical models used to infer the willingness to pay in contingent valuation. *Empirical Economics, 30*(1), 235. doi:10.1007/s00181-005-0236-x

Bishop, R. C., & Heberlein, T. A. (1979). Measuring Values of Extramarket Goods: Are Indirect Measures Biased? *American Journal of Agricultural Economics, 61*(5), 926-930. doi:10.2307/3180348

Bliemer, M. C. J., & Rose, J. M. (2013). Confidence intervals of willingness-to-pay for random coefficient logit models. *Transportation Research Part B: Methodological, 58*, 199-214. doi:http://dx.doi.org/10.1016/j.trb.2013.09.010

Boman, M., Bostedt, G., xf, ran, Kristr, xf, & m, B. (1999). Obtaining Welfare Bounds in Discrete-Response Valuation Studies: A Non-Parametric Approach. *Land Economics, 75*(2), 284-294. doi:10.2307/3147011

Calia, P., & Strazzera, E. (2000). Bias and efficiency of single versus double bound models for continent valuation studies : a Monte Carlo analysis. *Applied economics, 32*(10), 1329-1336.

Cameron, T. A. (1988). A new paradigm for valuing non-market goods using referendum data: Maximum likelihood estimation by censored logistic regression. *Journal of Environmental Economics and Management, 15*(3), 355-379. doi:http://dx.doi.org/10.1016/0095-0696(88)90008-3

Cameron, T. A., & James, M. D. (1987). Efficient Estimation Methods for 'Closed-ended' Contingent Valuation Surveys. *Review of Economics and Statistics, 69*(2), 269-276. doi:http://www.mitpressjournals.org/loi/rest

Chen, H. Z., & Randall, A. (1997). Semi-nonparametric estimation of binary response models with an application to natural resource valuation. *Journal of Econometrics, 76*(1–2), 323-340. doi:http://dx.doi.org/10.1016/0304-4076(95)01794-1

Chen, Q., & Giles, D. (2012). Finite-sample properties of the maximum likelihood estimator for the binary logit model with random covariates. *Statistical Papers, 53*(2), 409. doi:10.1007/s00362-010-0348-z

Chow, S.C. (1985). Resampling procedures for the estimation of nonlinear functions of parameters. (Doctoral dissertation)

Cooper, J. C. (1994). A Comparison of Approaches to Calculating Confidence Intervals for Benefit Measures from Dichotomous Choice Contingent Valuation Surveys. *Land Economics, 70*(1), 111-122. doi:10.2307/3146445

Cooper, J. C. (2000). Nonparametric and Semi-nonparametric Recreational Demand Analysis. *American Journal of Agricultural Economics, 82*(2), 451-462. doi:http://ajae.oxfordjournals.org/content/by/year

Cooper, J. C. (2002). Flexible Functional Form Estimation of Willingness to Pay Using Dichotomous Choice Data. *Journal of Environmental Economics and Management, 43*(2), 267-279. doi:http://dx.doi.org/10.1006/jeem.2000.1177

Cordeiro, G. M., & McCullagh, P. (1991). Bias Correction in Generalized Linear Models. *Journal of the Royal Statistical Society. Series B (Methodological), 53*(3), 629-643.

Creel, M. (1998). A Note on Consistent Estimation of Mean WTP Using a Misspecified Logit Contingent Valuation Model. *Journal of Environmental Economics and Management, 35*(3), 277-284. doi:http://dx.doi.org/10.1006/jeem.1998.1030

Creel, M., & Loomis, J. (1997). Semi-nonparametric Distribution-Free Dichotomous Choice Contingent Valuation. *Journal of Environmental Economics and Management, 32*(3), 341-358. doi:http://dx.doi.org/10.1006/jeem.1997.0972

Crooker, J. R., & Herriges, J. A. (2004). Parametric and Semi-Nonparametric Estimation of Willingness-to-Pay in the Dichotomous Choice Contingent Valuation Framework. *Environmental and Resource Economics, 27*(4), 451-480. doi:10.1023/B:EARE.0000018518.55067.b2

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. 1-26. doi:10.1214/aos/1176344552

Efron, B. (1981). Nonparametric Estimates of Standard Error: The Jackknife, the Bootstrap and Other Methods. *Biometrika, 68*(3), 589-599. doi:10.2307/2335441

Efron, B., & Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician, 37*(1), 36-48. doi:10.2307/2685844

Efron, B., & Stein, C. (1981). The Jackknife Estimate of Variance. 586-596. doi:10.1214/aos/1176345462

Firth, D. (1993). Bias Reduction of Maximum Likelihood Estimates. *Biometrika, 80*(1), 27-38. doi:10.2307/2336755

Giraud, K. L., Loomis, J. B., & Cooper, J. C. (2001). A Comparison of Willingness to Pay Estimation Techniques from Referendum Questions. *Environmental and Resource Economics, 20*(4), 331-346. doi:10.1023/a:1013025120987

Gray, H. L., Schucany, W. R., & Watkins, T. A. (1975). On the Generalized Jackknife and its Relation to Statistical Differentials. *Biometrika, 62*(3), 637-642. doi:10.2307/2335521

Haab, T. C., & McConnell, K. E. (1998). Referendum Models and Economic Values: Theoretical, Intuitive, and Practical Bounds on Willingness to Pay. *Land Economics, 74*(2), 216-229. doi:10.2307/3147052

Hanemann, M., Loomis, J., & Kanninen, B. (1991). Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation. *American Journal of Agricultural Economics, 73*(4), 1255-1263. doi:10.2307/1242453

Hanemann, W. M. (1989). Welfare Evaluations in Contingent Valuation Experiments with Discrete Response Data: Reply. *American Journal of Agricultural Economics, 71*(4), 1057-1061.

Hofler, R. A., & List, J. A. (2004). Valuation on the Frontier: Calibrating Actual and Hypothetical Statements of Value. *American Journal of Agricultural Economics, 86*(1), 213-221. doi:10.1111/j.0092-5853.2004.00573.x

Huang, J.-C., Nychka, D. W., & Smith, V. K. (2008). Semi-parametric discrete choice measures of willingness to pay. *Economics Letters, 101*(1), 91-94. doi:http://dx.doi.org/10.1016/j.econlet.2008.06.010

*Jaeckel, L. A. The Infinitesimal Jackknife. Bell Laboratories

Kanninen, B. J. (1995). Bias in Discrete Response Contingent Valuation. *Journal of Environmental Economics and Management, 28*(1), 114-125. doi:http://dx.doi.org/10.1006/jeem.1995.1008

Kerr, G. N. (2000). Dichotomous choice contingent valuation probability distributions. *Australian Journal of Agricultural and Resource Economics*,*44*(2).

Kézdi, G., Hahn, J., & Solon, G. (2002). Jackknife minimum distance estimation. *Economics Letters, 76*(1), 35-45. doi:http://dx.doi.org/10.1016/S0165-1765(02)00016-2

Kim, Y., & Singh, K. (1998). Sharpening estimators using resampling. *Journal of Statistical Planning and Inference, 66*(1), 121-146. doi:http://dx.doi.org/10.1016/S0378-3758(97)00079-7

Kosmidis, I. (2007). Bias Reduction in Exponential Family Nonlinear Models.

Langford, I. H., Bateman, I. J., Jones, A. P., Langford, H. D., & Georgiou, S. (1998). Improved Estimation of Willingness to Pay in Dichotomous Choice Contingent Valuation Studies. *Land Economics, 74*(1), 65-75. doi:10.2307/3147213

Lee, H., Jogesh Babu, G., & Rao, C. R. (2012). A jackknife type approach to statistical model selection. *Journal of Statistical Planning and Inference, 142*(1), 301-311. doi:http://dx.doi.org/10.1016/j.jspi.2011.07.017

MacKinnon, J. G., & Smith, A. A., Jr. (1998). Approximate Bias Correction in Econometrics. *Journal of Econometrics, 85*(2), 205-230. doi:http://www.sciencedirect.com/science/journal/03044076

MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics, 29*(3), 305-325. doi:http://dx.doi.org/10.1016/0304-4076(85)90158-7

Maiti, T., & Pradhan, V. (2008). A comparative study of the bias corrected estimates in logistic regression. *Statistical Methods in Medical Research, 17*(6), 621.

McLeod, D. M., & Bergland, O. (1999). Willingness-to-Pay Estimates Using the Double-Bounded Dichotomous-Choice Contingent Valuation Format: A Test for Validity and Precision in a Bayesian Framework. *Land Economics, 75*(1), 115-125. doi:10.2307/3146996

Miller, R. G. (1964). A Trustworthy Jackknife. *The Annals of Mathematical Statistics, 35*(4), 1594-1605.

Oehlert, G. W. (1992). A Note on the Delta Method. *The American Statistician, 46*(1), 27-29. doi:10.2307/2684406

Parr, W. C. (1983). A Note on the Jackknife, the Bootstrap and the Delta Method Estimators of Bias and Variance. *Biometrika, 70*(3), 719-722. doi:10.2307/2336511

Parr, W. C., & Schucany, W. R. (1980). The Jackknife: A Bibliography. *International Statistical Review / Revue Internationale de Statistique, 48*(1), 73-78.

Sharot, T. (1976). The Generalized Jackknife: Finite Samples and Subsample Sizes. *Journal of the American Statistical Association, 71*(354), 451-454. doi:10.2307/2285332

Sharot, T. (1976). Sharpening the Jackknife. *Biometrika, 63*(2), 315-321. doi:10.2307/2335625

SOWDEN, R. R. (1972). On the first-order bias of parameter estimates in a quantal response model under alternative estimation procedures. *Biometrika, 59*(3), 573-579. doi:10.1093/biomet/59.3.573

Vasconcellos, G. M. C. K. L. P. (1997). Bias correction for a class of multivariate nonlinear regression models. *Statistics and Probability Letters, 35*(2), 155.

Venkatachalam, L. (2004). The contingent valuation method: a review. *Environmental Impact Assessment Review, 24*(1), 89-124. doi:http://dx.doi.org/10.1016/S0195-9255(03)00138-0

Watanabe, M., & Asano, K. (2009). Distribution Free Consistent Estimation of Mean WTP in Dichotomous Choice Contingent Valuation. *Environmental & Resource Economics, 44*(1), 1-10. doi:10.1007/s10640-008-9255-3

Weber, N. C. (1986). The jackknife and heteroskedasticity: Consistent variance estimation for regression models. *Economics Letters, 20*(2), 161-163. doi:http://dx.doi.org/10.1016/0165-1765(86)90165-5

Withers, C. S., & Nadarajah, S. (2013). Delta and jackknife estimators with low bias for functions of binomial and multinomial parameters. *Journal of Multivariate Analysis, 118*, 138-147. doi:http://dx.doi.org/10.1016/j.jmva.2013.02.006

Withers, C. S., & Nadarajah, S. (2014). BIAS REDUCTION: THE DELTA METHOD VERSUS THE JACKKNIFE AND THE BOOTSTRAP. *Pakistan Journal of Statistics, 30*(1), 143-151.

Magazine Article

Kling, C. L., & Thomson, C. J. (1996). The Implications of Model Specification for Welfare Estimation in Nested Logit Models*,* 103.

Kriström, B. (1990). A Non-Parametric Approach to the Estimation of Welfare Measures in Discrete Response Valuation Studies*,* 135.

Book

Haab, T.C and McConnell, K.E. (2002) Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation. Edward Elgar Publishing, Northampton. http://dx.doi.org/10.4337/9781843765431