Developing Confidence in Your Net-to-Gross Ratio Estimates

Bruce Mast and Patrice Ignelzi, Pacific Consulting Services

Some authors have identified the lack of an accepted methodology for calculating the precision of net-to-gross ratio estimates derived using qualitative choice analysis (QCA) as a limitation to the widespread application of QCA. Without this methodological tool, the researcher lacks a solid basis for assessing whether the net-to-gross estimate is reliable in a particular case where only a single estimate is available.

This paper compares four previously published methods for calculating QCA net-to-gross ratio confidence intervals that have been applied to either simulated or real utility program datasets. As a benchmark, the authors developed a rigorous, albeit computationally intensive, simulation method for estimating the true net-to-gross ratio distribution. The benchmark approach and the four published approaches were all applied to a set of on-site and telephone survey data developed in conjunction with a recent utility DSM impact evaluation.

Of the four methods compared, one method in particular produces results that are quite consistent with the benchmark. However, a second approach also produces quite reasonable results and is virtually trivial to compute. Both of these methods appear to underestimate the precision of the net-to-gross ratio estimate. The remaining two methods ignore numerous factors that contribute to variation in the net-to-gross ratio estimate and thus appear to significantly overestimate the ratio's precision.

INTRODUCTION

Recent research into the estimation of net impacts for utility DSM programs has produced promising new techniques for estimating net-to-gross ratios. One of the more promising approaches has been to model the customer decision process for both program participants and nonparticipants using qualitative choice analysis (Train et al. 1994; Train & Paquette, 1995). However, Goldberg and Kademan (1995) identify the lack of an accepted methodology for calculating the precision of the resulting net-to-gross ratio estimates as a limitation to the widespread application of qualitative choice analysis (QCA). Without this methodological tool, the researcher lacks a solid basis for assessing whether the net-to-gross estimate is reliable in a particular case where only a single estimate is available.

This paper compares four previously published methods for calculating QCA net-to-gross ratio confidence intervals that have been applied to either simulated or real utility program datasets. As a benchmark, the author developed a rigorous, albeit computationally intensive, simulation method for estimating the true net-to-gross ratio distribution. The benchmark approach and the four published approaches were all applied to a set of on-site and telephone survey data developed in conjunction with a recent utility DSM impact evaluation. A comparison of the results clearly illustrates the strengths and weaknesses of each approach.

The five approaches discussed here are all based on a nested logit framework appropriate for incentive and other type programs that require customer implementation of a targeted measure as a condition for program participation. The paper does not discuss confidence interval calculation for QCA analysis of audit-type programs in which customers may be defined as participants independent of their implementation decision.

METHODOLOGY

Summary of QCA Estimation Procedure

The basic method for deriving a point estimate of the net-to-gross ratio using QCA has been described in some detail by Train et al. (1994). For purposes of this paper, the following definitions and equations are used.

The probability that customer i participates in the program, given that he implements a measure, is estimated as:

$$PP_i = \frac{\exp(\beta Z_i)}{1 + \exp(\beta Z_i)} \tag{1}$$

where Z_i is a vector of explanatory variables that include factors affecting customer i's decision, and β is a vector of estimated coefficients that maximizes PP_i. The inclusive value term is calculated as:

$$IVi = \log_e(1 + \exp(\beta Z_i)) \tag{2}$$

The probability that customer i implements a measure, given program existence, is estimated as:

$$PI_i = \frac{\exp(\delta Z_i + \alpha IV_i)}{1 + \exp(\delta Z_i + \alpha IV_i)} \tag{3}$$

where Z_i is a vector of explanatory variables that affect the outcome of the choice; IV_i is the inclusive value term, indicating customer i's perception of the difference between options 1 and 2; and α and δ are estimated coefficients that maximize PI_i. The predicted probability of implementation in the absence of the program, PA_i, is calculated as:

$$PA_i = \frac{(\exp(\delta Z_i)}{1 + \exp(\delta Z_i)} \tag{4}$$

To calculate the net-to-gross ratio, one calculates the gross, naturally occurring, total, and net impacts of the program. In equation form, these quantities are calculated as:

$$\text{Total savings}_i = PI_i \bullet \text{Measure Savings}_i \tag{5}$$

$$\text{Naturally occurring savings}_i = PA_i \bullet \text{Measure Savings}_i \tag{6}$$

$$\text{Gross savings}_i = PP_i \bullet PI_i \bullet \text{Measure Savings}_i \tag{7}$$

$$\text{Net savings}_i = \text{Total savings}_i - \text{Naturally occurring savings}_i \tag{8}$$

$$NTG = \frac{\Sigma\ \text{Net savings}_i}{\Sigma\ \text{Gross savings}_i} \tag{9}$$

It can be shown that, if measure savings can be taken as constant across customers, then the net-to-gross ratio reduces to a function of probabilities, that is:

$$NTG = \frac{\Sigma\ PI_i - PA_i}{\Sigma\ PP_i \bullet PI_i} \tag{10}$$

Theoretical Considerations for Estimating Confidence Intervals

The most theoretically rigorous approach to estimating the net-to-gross ratio and its variance would be to estimate equations 1 and 3 as a set of simultaneous equations. This approach would produce a single variance-covariance matrix for the parameter vectors α, β, and δ. With this matrix in hand, one could derive an analytical solution for the variance of NTG as a function of vectors α, β, δ, the variance-covariance matrix, and the explanatory variable matrices Z. However, some commonly used statistical software packages, such as SAS, do not offer a convenient method for simultaneously solving a system of logit equations. In the absence of this capability, researchers must resort to the sequential estimation approach described above.

To date, no one has succeeded in developing an analytical solution for the variance of the net-to-gross ratio estimate in the case of sequentially estimated models. This calculation is complicated by the nonlinear and nested functional relationship between NTG, α, β, and δ. To illustrate, NTG is a nonlinear function of PI_i, PP_i, and PA_i. The variance of PA_i is a nonlinear function of the variance-covariance of δ, the variance of PI_i is a nonlinear function of the variance-covariance of α and δ, and the variance of PP_i is a nonlinear function of the variance-covariance of β. To complicate matters further, the *point estimates* of α and δ, and by extension the point estimates of PI_i and PA_i, are a function of the variance of IV_i, itself a nonlinear function of the variance-covariance of β.

Benchmark Approach to Confidence Interval Estimation

In the absence of an analytical solution, the most rigorous alternative is a Monte Carlo simulation-type solution, that is, a solution based on repeated random draws for δZ_i, $\delta Z_i + \alpha IV_i$, and βZ_i, based on the variance-covariance of α, β, and δ. Since the net-to-gross ratio point estimate is a function of two nested logit models, the random draw approach requires two nested Monte Carlos.

One first calculates the maximum likelihood estimate of β and the standard deviation of the exponential term, βZ_i, using equation 1. The standard deviation of the exponential term, βZ_i, is calculated as:

$$\sigma_{i1} = \sqrt{[(1,Z') \bullet V_b \bullet (1,Z')']_{ii}} \tag{11}$$

where V_b is the estimated covariance matrix of the estimated parameter vector β. One then draws a set of m random values from the distribution of βZ. This distribution is taken to be normal with mean equal to βZ and variance equal to σ_1^2. For each draw, the probability of participation, PP, and the corresponding inclusive value term, IV, is recalculated.

One then estimates equation 3 m times, once for each value of IV, producing m sets of parameter estimates, α and δ, and m covariance matrices. For each set of parameter estimates, one calculates the standard deviation of the exponential terms, $\delta Z + \alpha IV$ and δZ. The standard deviations are calculated as:

$$\sigma_{i2} = \sqrt{[(1,Z',IV') \bullet V_{\delta\alpha} \bullet (1,Z',IV')']_{ii}} \tag{12}$$

$$\sigma_{i3} = \sqrt{[(1,Z') \bullet V_{\delta} \bullet (1,Z')']_{ii}} \tag{13}$$

One then draws m sets of n random values from the distributions of $\delta Z + \alpha IV$ and δZ, one for each set of parameter estimates. These distributions are taken to be normal with means equal to $\delta Z + \alpha IV$ and δZ and variances equal to σ_2^2 and σ_3^2, respectively. For each draw, the probability of implementation with and without the program, PP and PA, are recalculated along with the corresponding net-to-gross ratio. The result is m x n values for the net-to-gross ratio, distributed as a function of the variance-covariance matrices for α, β, and δ. Finally, the confidence interval is tabulated on the basis of this approximate distribution. For example, if a 90% confidence level is required, the corresponding confidence interval is obtained by removing the highest 5% and the lowest 5% of the values; the range from the lowest to highest remaining value represents the 90% confidence interval around the net-to-gross ratio point estimate.

A rigorous random draw approach is clearly very computationally intensive. Researchers have explored various approximations and proxies that require much less effort to compute. Four such approaches are described below, beginning with the conceptually most straightforward approach and proceeding to the more complicated approaches. Two approaches have been applied to simulated datasets and described in published conference proceedings. The other two have been applied to utility data and described in impact evaluation reports to those utilities.

Goldberg and Train (1995) Approach

The most straightforward approach described in the literature is that put forward by Goldberg and Train (1995). They used the standard error of the inclusive value parameter, α, as a likely proxy for the standard error of the estimated net-to-gross ratio. For a desired level of confidence, say $100 \bullet (1 - \phi)\%$, the upper and lower bounds for the net-to-gross ratio are thus calculated as:

$$NTG_{upper} = NTG + z_{\phi/2}\, \sigma_\alpha \tag{14}$$

$$NTG_{lower} = NTG - z_{\phi/2}\, \sigma_\alpha \tag{15}$$

where $z_{\phi/2}$ is the $100(1 - \phi/2)$ percentile point of the standard normal distribution. Since the method involved simulated data with a known true net-to-gross ratio and a known true confidence interval, they were able to demonstrate that this approach produces reasonable results.

PCS (1993) Approach

In this study, we calculated a confidence interval around the net-to-gross ratio point estimate based on extreme values of the inclusive value parameter, α. As a first step using this method, one derives maximum likelihood estimates for α, β, and δ and then one calculates minimum and maximum values for α, based on the corresponding standard error and the required confidence level restrictions. Plugging these values into the equation for PI_i produces two extreme values:

$$PI_{i\ upper} = \frac{\exp((\alpha + z_{\phi/2} \bullet \sigma_\alpha)IV_i + \delta Z_i)}{1 + \exp((\alpha + z_{\phi/2} \bullet \sigma_\alpha)IV_i + \delta Z_i)} \tag{16}$$

$$PI_{i\ lower} = \frac{(\exp((\alpha - z_{\phi/2} \bullet \sigma_\alpha)IV_i + \delta Z_i)}{1 + \exp((\alpha - z_{\phi/2} \bullet \sigma_\alpha)IV_i + \delta Z_i)} \tag{17}$$

One then calculates a net-to-gross ratio estimate for each extreme value, thus producing upper and lower bounds for the point estimate. This range is taken to be the confidence interval around the point estimate.

Train et al. (1994) Approach

This study took the PCS (1993) logic one step further, applying a random draw approach based on the variance of the inclusive value parameter, α. The application of this approach to the particular program evaluated was complicated by the fact that the research team estimated measure-specific net-to-gross ratios and then calculated weighted averages to arrive at a program-level net-to-gross ratio. Applied to the simple case of one program-specific set of models (equivalent to the case of one program measure), this method of determining an appropriate confidence interval proceeds as follows.

As in the PCS (1993) method, one derives maximum likelihood estimates for α, β, and δ, only now one calculates a distribution of values for the inclusive value parameter, α. The distribution of a around its estimated value is taken to be normal with variance equal to the square of the standard error. Random draws are taken from this distribution. For each draw, that is for each value of a, the probability of participation, PI, and the corresponding net-to-gross ratio is recalculated. This process produces a distribution of net-to-gross ratios arising from the distribution of α. The confidence interval is tabulated on the basis of this approximate distribution.

PCS (1995) Approach

This alternative accounts for the variance and covariance of α, β and δ. However, rather than run m x n Monte Carlo iterations to derive the complete net-to-gross ratio distribution, it focuses on the extreme values, as in the PCS (1993) approach.

The basic method is to calculate standard deviations for the product βZ from the participation model (equation 1); propagate the resulting extreme values for IV to the implementation model (equation 3); estimate standard deviations for δZ and $\alpha IV + \delta Z$ corresponding to each extreme value of IV; and then propagate the standard deviations to the net-to-gross ratio. The result is a set of minimum and maximum values of the net-to-gross ratio that reflect the desired level of confidence.

One first calculates point estimate for the probability of participation, the standard deviation of the exponential term, βZ_i, and the upper and lower bounds for the probability of participation, PP_i, and the inclusive value term, IV_i. Equations for the point estimate and standard deviation are given above.

The upper and lower bounds for customer i are then given by the following equations:

$$PP_{i\ upper} = \frac{\exp(\beta Z_i + z_{\phi/2} \bullet \sigma_{i1})}{1 + \exp(\beta Z_i + z_{\phi/2} \bullet \sigma_{i1})} \tag{18}$$

$$PP_{i\ lower} = \frac{\exp(\beta Z_i - z_{\phi/2} \bullet \sigma_{i1})}{1 + \exp(\beta Z_i - z_{\phi/2} \bullet \sigma_{i1})} \tag{19}$$

$$IV_i\ upper = Log_e(1 + \exp(\beta Z_i + z_{\phi/2} \bullet \sigma_{i1})) \tag{20}$$

$$IV_i\ lower = Log_e(1 + \exp(\beta Z_i - z_{\phi/2} \bullet \sigma_{i1})) \tag{21}$$

One then estimates the probability of participation, PI_i, three times, using the point estimate of IV_i, the upper bound of IV_i, and the lower bound. The model specification remains unchanged across estimates; only the value for IV_i changes. The purpose of the first model, using the point estimate of IV_i, is to derive the point estimate for PI_i that will be used to derive the point estimate of the net-to-gross ratio. The purpose of the other two models is to calculate extreme values for PI_i, given extreme values for IV_i. Thus, for each extreme value of IV_i, one calculates upper and lower bounds for PI_i. These bounds are given for customer i by the following equations, where σ_{i2} is the standard deviation of the exponent, $\alpha IV_i + \delta Z_i$ and $z_{\phi/2}$ is again the $100(1-\phi/2)$ percentile point of the standard normal distribution.

$$PI_{i\ upper\ given\ IV\ upper} = \frac{\exp(\alpha IV_i\ upper + \delta Z_i + z_{\phi/2} \bullet \sigma_{i2})}{1 + \exp(\alpha IV_i\ upper + \delta Z_i + z_{\phi/2} \bullet \sigma_{i2})} \tag{22}$$

$$PI_{i\ lower\ given\ IV\ upper} = \frac{\exp(\alpha IV_i\ upper + \delta Z_i - z_{\phi/2} \bullet \sigma_{i2})}{1 + \exp(\alpha IV_i\ upper + \delta Z_i - z_{\phi/2} \bullet \sigma_{i2})} \tag{23}$$

$$PI_{i\ upper\ given\ IV\ lower} = \frac{\exp(\alpha IV_i\ lower + \delta Z_i + z_{\alpha/2} \bullet \sigma_{i2})}{1 + \exp(\alpha IV_i\ lower + \delta Z_i + z_{\alpha/2} \bullet \sigma_{i2})} \tag{24}$$

$$PI_{i\ lower\ given\ IV\ lower} = \frac{\exp(\alpha IV_i\ lower + \delta Z_i - z_{\alpha/2} \bullet \sigma_{i2})}{1 + \exp(\alpha IV_i\ lower + \delta Z_i - z_{\alpha/2} \bullet \sigma_{i2})} \tag{25}$$

Corresponding values for PA_i are calculated by setting a equal to zero in equations 22 through 25 and substituting σ_{i3} for σ_{i2}. For simplicity of notation, the above equations do not distinguish between α, δ, and σ_{i2} for the model using IV_i upper and the model using IV_i lower. In reality, α and δ are estimated independently for each value of IV_i, resulting in different values for σ_{i2}.

Finally, one calculates five values for the net-to-gross ratio: a point estimate and an estimate for each of the four extreme values of PI_i. Since the net-to-gross ratio is a function of both PI_i and PP_i, the upper bound for PP_i is used for the two net-to-gross ratio values based on IV_i upper and the lower bound for PP_i is used for the two net-to-gross ratio values based on IV_i lower. It can be shown that the net-to-gross ratio is a decreasing function of IV_i; that is, replacing IV_i with its upper bound results in a lower value for the net-to-gross ratio. For this reason, the maximum value of the net-to-gross ratio occurs using the upper bound of PI_i given the lower bound of IV_i. Conversely, the minimum value of the net-to-gross ratio occurs using the lower bound of PI_i given the upper bound of IV_i. The other two combinations produce intermediate values of the net-to-gross ratio around the point estimate.

RESULTS

We performed a direct comparison of these different confidence interval calculation methods using an analysis dataset constructed to evaluate a DSM incentive program for a major utility. In this dataset of 299 customers, 189 reported having installed some form of lighting efficiency measure. Of these, 112 did so through the incentive program and received a rebate. A discrete choice analysis of net impacts for this program produced a net-to-gross ratio estimate of 0.68. The corresponding 90% confidence intervals using the various calculation approaches are shown in Table 1. The benchmark results were derived with the iteration indices m and n both set to 50, producing 2,500 NTG values.

Table 1. *Comparison of Confidence Intervals Using Different Calculation Methods*

Method	Minimum NTGR	Maximum NTGR
Benchmark	0.49	0.89
Goldberg & Train (1995)	0.39	0.95
PCS (1993)	0.60	0.73
Train et al. (1994)	0.60	0.74
PCS (1995)	0.46	0.97

The PCS (1995) approach appears to produce the most reliable confidence interval estimate, short of generating the entire net-to-gross ratio distribution through repeated sampling and reestimation. It properly acknowledges that all parameters in both the participation and implementation models are estimated with uncertainty. However, it only accounts for the variance of the participation model by propagating two extreme values through the calculation process. As a result, this method oversimplifies the functional relationship between the variance of NTG and the variance of βZ, producing a confidence interval that is too wide.

The Goldberg and Train (1995) approach estimates the upper confidence interval bound at least as well as the PCS (1995) approach and offers the added advantage of requiring no additional computer programming beyond the basic point estimation procedure. However, it estimates the lower bound with much less accuracy, probably because it treats the net-to-gross ratio as a normal distribution, whereas the actual distribution is asymmetric about the median value.

The PCS (1993) and Train et al. (1994) produce virtually identical results. Both approaches are based on the observation that the inclusive value parameter, α, is the single most informative parameter in the nested logit system of equations and that the net-to-gross ratio is a nonlinear function of α. The PCS (1993) approach is somewhat easier to implement since it calculates only two extreme values rather than the entire NTG distribution. Both methods ignore the variance and covariances of β and δ. By assuming that the probability of participation has no variance and that the variance of the probability of implementation depends only on a, these methods significantly underestimate the resulting variance of the net-to-gross ratio.

Post Script

Further research into estimating confidence intervals has suggested that a variation on the benchmark methodology may provide a more rigorous estimate of the true confidence interval. This revised approach consists of a Monte Carlo simulation based on random draws of α, β, and δ rather than δZ_i, $\delta Z_i + \alpha IV_i$, and βZ_i. As of the time of publication, the authors have not yet completed programming and testing this alternative.

ACKNOWLEDGMENTS

The authors wish to acknowledge the help of Dr. Kenneth Train, Professor of Economics at the University of California, who provided invaluable insights into the methodological issues discussed in this paper. We would also like to thank Susan Buller of Pacific Gas & Electric and Marian Brown of Southern California Edison, for whom much of this research was originally conducted.

REFERENCES

Goldberg, M.L. and E. Kademan. 1995. ''Is It Net or Not? A Simulation Study of Two Methods.'' *In Proceedings of the 1995 Energy Program Evaluation Conference*: 459-465.

Goldberg, M.L. and K. Train. 1995. *Net Savings Estimation: An Analysis of Regression and Discrete Choice Approaches*. Report submitted by XENERGY, Inc. to CADMAC Subcommittee on Base Efficiency.

Hartman, R. 1988. ''Self-Selection Bias in the Evaluation of Energy Conservation Programs.'' *Review of Economics and Statistics*: 448-459.

McFadden, D. 1973. ''Conditional Logit Analysis of Qualitative Choice Behavior,'' in P. Zarembka, ed. *Frontiers in Econometrics*. New York: Academic Press, New York.

Pacific Consulting Services. 1993. *1990 Southern California Edison Energy Management services and Hardware Rebate Program Evaluation*, *Vol.6*. Submitted to Southern California Edison.

Pacific Consulting Services. 1995. *1992-93 Nonresidential New construction Programs: Impact Evaluation Final Report*. Submitted by SBW Inc. to Pacific Gas and Electric Company.

Train, K. 1986. *Qualitative Choice Analysis*. Cambridge, Massachusetts: MIT Press.

Train, K. 1993. *Estimation of Net Savings from Energy Conservation Programs*. Submitted to the Southern California Edison Company.

Train, K. 1994. ''Estimation of Net Savings from Energy Conservation Programs.'' *Energy*. 19(4):423-441.

Train, K. 1995. ''Net Savings Estimation.'' *EPRI Non-Residential Impact Evaluation Guidebook*, EPRI Report, Chapter 5.

Train, K., S. Buller, B. Mast, K. Parikh, and E. Paquette. 1994. "Estimation of Net Savings for Rebate Programs: A Three-Option Nested Logit Approach.'' *In Proceedings of the ACEEE 1994 Summer Study*: 7:239-247.

Train, K. and E. Paquette. 1995. ''A Discrete Choice Method to Estimate Free-Ridership, Net-to-Gross Ratios, and the Effect of Program Advertising.'' *Energy Services Journal*:1(1).