

A Comparison of Resampling Techniques when Parameters are on a
Boundary:

The Bootstrap, Subsample Bootstrap, and Subsample Jackknife

by:

Christiana E. Hilmer and Matthew T. Holt¹

Selected Paper

American Agricultural Economics Association Annual Meeting
Tampa, Florida
July 30-August 2, 2000

Abstract:

This paper compares the finite sample performance of subsample bootstrap and subsample jackknife techniques to the traditional bootstrap method when parameters are constrained to be on some boundary. To assess how these three methods perform in an empirical application, a negative semi-definite translog cost function is estimated using U.S. manufacturing data.

Copyright 2000 by Christiana E. Hilmer and Matthew T. Holt. All rights reserved. Readers may make verbatim copies of the document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

¹ Christiana E. Hilmer is a graduate student in the Department of Agricultural and Resource Economics and the Department of Statistics, North Carolina State University. Matthew T. Holt is the Professor in the Department of Agricultural and Resource Economics and holds the Cardon Chair in Agribusiness at the University of Arizona.

1. Introduction

Empirical economists frequently estimate models in which parameters are constrained to be on a boundary of the parameter space. The need to do so usually arises when *a priori* theoretical restrictions require a certain estimated parameter to be of a specific sign. Examples of these types of restrictions include traditional demand analysis where the income effect for a normal good is constrained to be positive while the own-price effect is constrained to be negative; cost function analysis where curvature constraints imply that second-order price terms satisfy concavity conditions; and time series models for conditional heteroskedasticity where the GARCH parameters are constrained to be non-negative. Such cases are potentially problematic for the researcher, as standard error estimates and confidence bounds are difficult to compute using classical statistical inference. In short, traditional distributional assumptions no longer apply when the parameter space is truncated because the asymptotic distribution is no longer normal.

Two primary strategies have been utilized in order to account for the problems presented by models in which parameters are constrained to be on a boundary. The first strategy is to simply refuse to compute standard errors or confidence bounds because the procedure required to do so would be too complicated (Diewert and Wales [1988], O'Donnell *et al.* [1999], Barnett and Yue [1988]). The second strategy is to impose the inequality constraints implied by the *a priori* theoretical restrictions and use traditional methods such as the delta method and the Wald test to compute standard errors and test hypotheses (Appelbaum and Ulah [1997], Ryan and Wales [1999]). This approach is not recommended considering that a common regularity condition for use of the Wald and Likelihood Ratio tests is that the true parameter must be in the interior of the parameter space (Demos and Sentana, 1998).² Therefore, any conclusions researchers draw from this type of analysis are, at best,

questionable. Indeed, both of these strategies are unsatisfactory, as the first amounts to not having a measure of confidence in the reported estimates and the second amounts to reporting incorrect estimates of standard errors or test statistics.

The reason researchers have relied heavily on these two unsatisfactory approaches is that most proposed classical methods for dealing with the problem have proven to be exceedingly complex. For instance, Goueiroux, *et al.* (1982), Self and Liang (1987), and Wolak (1989a, 1989b, 1991) find that for models with inequality constraints the correct testing procedures involve Kuhn-Tucker-like tests in which the asymptotic distribution of the null hypothesis is a weighted mixture of chi-square distributions. Specifically, the appropriate distribution for a model with q restrictions is a weighted average of $q+1$ chi-square distributions, $\chi_0^2, \chi_1^2, \chi_2^2, \dots, \chi_q^2$, which is more concentrated toward the origin than the traditional tabled χ_q^2 distribution. See Demos and Sentana (1998) for additional details. More recently, Andrews (1998) developed a weighted average power optimality criterion that indexes the amount of weight that is placed on alternatives depending on their distance from the null hypothesis. These tests involve a weighted combination of truncated normal distributions. The advantage of this testing approach over that of Goueiroux *et al.* (1982) is that it can also be applied to nonlinear regression problems; however, the practical problem with utilizing an approach that involves a mixture of distributions is determining how much weight to put on each distribution. For example, Self and Liang (1987) show that the test statistic in a linear model with one linear inequality restriction is an average of the χ_0^2 and χ_1^2 distributions. Because the correct weighting factors depend on the structure of the inverse information matrix and the number of restrictions, they will differ greatly across applications.

Alternatives to classical approaches include mathematical programming and Bayesian methods. Gallant and Golub (1984) use a mathematical programming routine to impose curvature constraints directly during estimation in a flexible functional form. These curvature constraints are

imposed on the eigenvalues of the matrix of Allen elasticities of substitution. An approach adopted by Chalfant and White (1988), Chalfant, Gray, and White (1994), Koop, Osiewalski and Steel (1994), and Terrell (1996) is to rely on Bayesian techniques to impose curvature restrictions. Bayesian methodology treats parameters not as fixed values but as items that can be continually updated as one's subjective views of the world change.

It is evident that each of the methods listed above are either quite complicated to compute, apply to fairly specific cases, or rely on Bayesian methodology. There is clearly a place in the classical statistics literature for a test that applies in both linear and nonlinear applications *and* is easy to implement when parameters are constrained to be on a boundary. A recent theoretical work by Andrews (1999) explores the use of re-sampling techniques to calculate standard errors and confidence intervals when a parameter is on a boundary. Since being developed by Efron (1979), the bootstrap has become a popular method to calculate standard errors. As demonstrated by Andrews (1999), however, this procedure is not asymptotically correct to the first order when parameters are on a boundary. For this reason, Andrews (1999) proposes using subsample bootstrap or subsample jackknife methods, in lieu of the traditional bootstrap. This dissertation is the first attempt to empirically address the important theoretical work of Andrews. We start by comparing the finite sample performance of the subsample bootstrap and subsample jackknife methods to the traditional bootstrap method preferred by previous researchers. An important concern when using subsample techniques is the correct subsample size to use for different sample sizes. Because previous theoretical work does not address this issue, Monte Carlo simulations are conducted for various sample and subsample sizes to help determine the relationship between the two. In addition to comparing the finite sample performance of the different techniques, it is necessary to examine how the three methods perform in actual empirical situations where parameters are inequality constrained. To this end, we estimate a translog cost function, applied to Berndt-Woods manufacturing data, in which concavity requires the matrix of parameter estimates to be negative semi-definite.

2. Theory

Re-sampling techniques, such as the bootstrap or the jackknife, are often used when the underlying distribution of the parameter is either unknown, mathematically difficult to compute, or has no analytic distribution. These techniques rely on computer power instead of distributional assumptions or mathematical rigor. The jackknife is the older of the two techniques, and was first used by Quenouille (1949) to reduce bias in estimators and later by Tukey (1958) as a method for estimating variances. The standard (or delete one) jackknife estimate is computed by deleting one observation from the sample data set, calculating the parameter of interest for the remaining $n-1$ observations, and repeating this process until each of the n observations has been deleted. The jackknife estimator is then the average across all of the parameter estimates from the n different jackknife samples. A shortcoming of this technique is that it fails if the parameter of interest is discontinuous, such as the median. For example, Miller (1974) demonstrated that jackknife estimates of sample quantiles do not perform well in finite samples. Because of this shortcoming, Efron (1979) developed the bootstrap as an alternative. The goal of the bootstrap is to use the sample data to mimic the overall population distribution and to use resampling to calculate an estimated sampling distribution. For an initial sample data set of size n , this is accomplished by drawing a random sample of size n with replacement from the initial sample data set, calculating the parameter of interest for the sample drawn, and repeating the process many times. The bootstrap estimator is then the average across all of the parameter estimates from the different bootstrap samples. The intuition behind this methodology is that the distribution of the parameter estimates from the bootstrap samples mimics the traditional asymptotic sampling distribution of parameter estimates for samples drawn from the population distribution. Hence, the bootstrap technique allows the researcher to generate an estimated sampling distribution in cases in which she/he only has access to a single sample rather than the entire population.

These types of re-sampling techniques hold promise in our quest to derive a comparatively easy and general method for conducting statistical inference when parameters are on a boundary of a

parameter space. This is because the bootstrap doesn't require excessive mathematical rigor and modern computer power makes re-sampling relatively painless. In order for the bootstrap to provide consistent estimates of the asymptotic distribution of a parameter, however, certain conditions must be met. Bickel and Freedman (1981) define these necessary conditions as:

- (1) weak convergence of the statistic when the sample data has distribution G , for all G in a neighborhood of the true distribution, F .
- (2) the convergence of G to F is uniform in a neighborhood of F .
- (3) continuity of the mapping from G to the asymptotic distribution of the statistic.

Unfortunately, when parameters are on a boundary of the parameter space, the mapping from the sample data to the true distribution is no longer continuous, thereby violating condition (3). Hence, as Andrews (1997a) demonstrates, the bootstrap method for obtaining confidence intervals and standard errors when a parameter is on a boundary is theoretically invalid. Intuitively, this is because the bootstrap yields an estimated sampling distribution that puts too much mass below the cutoff for the parameter and therefore does a poor job of mimicking the true population distribution.

Because the traditional bootstrap yields inconsistent standard error estimates when parameters are on a boundary, alternative techniques must be developed. Andrews (1999) offers three different possibilities for obtaining consistent estimators of the entire asymptotic distribution in such cases: the law of the iterated logarithm; the subsample jackknife; and the subsample bootstrap. For discussion purposes assume we are trying to estimate $N_T(\hat{\theta} - \theta)$ where $\hat{\theta}$ is an estimator that maximizes a function over the parameter space, θ is the true value of the parameter, and N_T is a normalizing constant.³

The first method, the law of the iterated logarithm, is based on a rule that decides whether or not any of the inequality constraints are binding. This rule involves comparing the parameter estimate to a cutoff point derived by multiplying the critical value from a Wald test by the estimated standard error obtained by assuming the parameter is not on the boundary. If the parameter is above the cutoff

point, one concludes that the true parameter is not on the boundary, implying the distribution is not truncated. It is therefore possible to proceed by using standard asymptotic results. If the estimated parameter is below the cutoff point, however, the inequality constraint is binding and the asymptotic distribution must be simulated by a truncated distribution. This method is unattractive because, among other things, it is highly complicated and computationally much more difficult to employ than the latter two methods. Consequently, it will not be addressed in this dissertation.

The second method, the subsample jackknife, differs from the standard jackknife by deleting more than one observation. Specifically, to perform the subsample jackknife, d (greater than 1) observations are dropped, the parameters are calculated using the remaining m (where $m = n-d$) observations, and the process is repeated until all possible samples of size m have been drawn. Hence, to cover all possible subsets of the data with d observations deleted, $\binom{n}{m}$ statistics should be computed. Simple inspection reveals that as d increases, $\binom{n}{m}$ increases. This presents a practical problem for the researcher in that the potential number of subsamples to be drawn is likely far too large to allow for efficient calculation of each of the possible subsamples. Instead, the researcher will only want to take a random sample of the possible subsamples. Accordingly, for the subsample jackknife, the asymptotic distribution of $N_T(\hat{\theta} - \theta)$ is approximated by $\sqrt{m}(\hat{\theta}_{sj} - \hat{\theta})$, where $\hat{\theta}_{sj}$ is the subsample jackknife estimator and $\hat{\theta}$ is some maximizer of the objective function.

The third method, the subsample bootstrap, corrects the problems inherent in the standard bootstrap by drawing, with replacement, repeated samples of size q (where q is less than n) from the initial sample of size n . Andrews (1999) demonstrates that basing the bootstrap on these smaller samples yields a consistent asymptotic distribution. For the subsample bootstrap $N_T(\hat{\theta} - \theta)$ is approximated by $\sqrt{q}(\hat{\theta}_{sb} - \hat{\theta})$, where $\hat{\theta}_{sb}$ is the subsample bootstrap estimator and $\hat{\theta}$ is again some maximizer of the objective function.

Although subsampling techniques have not been applied to economic problems where a parameter is on a boundary, the literature does contain examples where they have been applied to other situations. The subsample jackknife was first proposed by Wu (1989) to rectify the deficiencies of the standard jackknife for non-smooth estimators. Those results were extended by Politis and Romano (1994) to show that neither asymptotic normality nor iid data are needed to accurately estimate a sampling distribution. The latter work is the basis of many papers (e.g., Li and Maddala [1997], Politis *et al.* [1997]) that extend re-sampling techniques to stationary time series data. The subsample bootstrap was first used in applied economic research by Ziari et al. (1997) who utilized a mathematical programming model to produce estimates without any statistical properties. Comparing bootstrap, subsample bootstrap, and subsample jackknife techniques to identify statistical parameter estimates, those authors found that the subsample jackknife method performed best within their application. A final application of the subsample bootstrap method appears in a paper by Bickel and Ren (1996) who set critical values for Cramer-von Mises goodness-of-fit tests with doubly censored data.⁴ They first prove that the standard bootstrap fails to estimate the null distribution of the test statistic and then proceed to show that the subsample bootstrap has the correct asymptotic power function and is asymptotically consistent.

3. Simulation Study

The theoretical work of Andrews demonstrates that the subsample bootstrap and subsample jackknife provide consistent standard error estimates in a model where parameters are constrained to be on a boundary. In practice, however, researchers do not have access to infinite samples. It is therefore important to consider the finite sample performance of the various re-sampling techniques. To this end, we perform Monte Carlo studies to compare how each method performs for samples of different sizes. In addition, we conduct traditional bootstrap estimation to assess whether there is any finite sample gain in using the re-sampling techniques over the traditional bootstrap methodology. In other words, even though the traditional bootstrap is theoretically inconsistent in the limit, we would like to know the degree to which a researcher is in error if she/he uses the traditional bootstrap in lieu

of one of the subsampling methods in finite samples. A collateral concern when using the two subsample techniques is that previous theoretical work has given no indication as to the optimal subsample size. Consequently, within our Monte Carlo studies we explore the effects that different subsample sizes have on the relative performance of the estimators.

To replicate situations in which economists may use these techniques, Monte Carlo studies are conducted for one and two equation systems. This paper presents results from sample sizes of 50 for both the one equation and two equation systems in order to determine how subsampling techniques behave as sample size changes. As we are also interested in how results vary with different subsample sizes, results are calculated for a number of different subsampling sizes, ranging from \sqrt{n} to $n - \sqrt{n}$.

3.1 One-equation Model

The first application we consider is a simple one-equation model in which we wish to estimate two parameters. The model is specified as

$$(1) \quad y_t = \alpha + \beta x_t + \varepsilon_t$$

where *a priori* theoretical restrictions constrain β to be non-negative and the usual Gauss Markov assumptions hold. A simplistic macroeconomic version of equation (1) is the principles-level consumption function that relates consumption to disposable income with y_t and x_t representing consumption and disposable income respectively in period t . In this formulation, the intercept α is autonomous spending and β is the marginal propensity to consume. Basic macroeconomics principles suggest that β is positive, meaning that if disposable income increases then consumption increases by some fraction of that amount. Therefore, due to *a priori* restrictions, β is constrained to be greater than or equal to zero.

To examine the finite sample performance of the bootstrap, subsample bootstrap, and subsample jackknife techniques Monte Carlo studies are conducted using a nested loop. The first step in the nested loop is to assign values to the coefficients in equation (1). Because we are trying to

assess finite sample performance when a parameter is on a boundary, we set β equal to the lowest value under the constraint, which is zero. α is set equal to two. Following Godfrey and Veall (1998), we randomly construct the independent variable, which is fixed in repeated samples, using the model $x_t = .75x_{t-1} + \eta_t$, where η_t is a normal random variable with mean zero and variance two. The error term in (1) is randomly generated from a normal distribution with mean zero and variance one. The Monte Carlo strategy we employ nests an inner loop within an outer loop. The outer loop, which we will call a Monte Carlo run, begins by creating dependent variables, y_t , using our fully specified model in equation (1). After generating the data, we employ maximum likelihood estimation. In this outer loop, or Monte Carlo run, we denote the estimates of α and β from equation (1) as A_{ml} and B_{ml} . To impose the non-negativity constraint on β , we use the technique of Barnett and Yue (1988) to actually estimate c , where by definition $c = \beta$. The immediate implication of this construct is that even though c may be negative, an estimate of β that is less than zero will never be obtained.

The inner loop, or bootstrap run begins at this point. Using one of the three resampling techniques, we obtain a new set of data by taking random draws from the residuals of the Monte Carlo estimation to acquire a set of resampled residuals, e_i^* . A new dependent variable, y_t^* , is then obtained by combining A_{ml} , B_{ml} , x_t , and e_i^* as prescribed by equation (1). The model is then re-estimated using the new data obtained from the re-sampling methods. We denote the resulting estimates as A_{rs} , c_{rs} , and $B_{rs} = c_{rs}c_{rs}$, where rs indicates that a re-sampling estimator has been used. An asymptotic estimate of the distribution $N_T(B_{rs} - B_{ml})$ is obtained by subtracting the Monte Carlo estimate from the resampling estimate and multiplying by the weighting factor N_T . The inner loop, or bootstrap run, is performed 1000 times and the outer loop, or Monte Carlo run, is also performed 1000 times. Therefore, equation (1) is estimated a total of 1,000,000 times for each of the five different sample sizes. At the end of each inner loop, both 90 and 95 percent confidence intervals of the asymptotic distribution are obtained.

It is instructive to carefully consider the method used for calculating confidence intervals. Figure 1 depicts a histogram of the asymptotic distribution resulting from a single Monte Carlo run for the subsample bootstrap for a sample size of size fifty with a subsample size of forty. These figures are consistent with the other sample sizes and re-sampling methods. As illustrated by this figure, the re-sampling distribution is skewed towards the left. When constructing confidence intervals, it is therefore necessary to account for skewness of the distributions. The usual approach to calculating confidence intervals for symmetric distributions is the percentile method, under which the lower limit of a 95% confidence interval with 1000 bootstrap runs is the average of the 25th and 26th ordered observations and the upper limit is the average of the 975th and 976th observations. This method places a total of 5 percent of the distribution equally in the two exterior tails. This method is not appropriate for skewed distributions, however, as it leads to confidence intervals that are wider than desired. The wider intervals result in a loss of precision because they place less of the mass of the confidence interval in the tightest part of the distribution. To calculate more precise intervals for our skewed distributions, we use the modified percentile method as detailed in Davidson and MacKinnon (1993, pp. 766). The modified percentile method creates narrower confidence intervals by minimizing the length of the confidence interval. The intuition behind this method is that the confidence interval is moved away from the side of the distribution that has the least amount of information.⁵ In analyzing the precision of our estimators we will be interested in those with the shortest average confidence interval length as those will contain the most information.

To measure the accuracy of our re-sampling estimators, we calculate the coverage probability, which is the number of times the true value of the estimate, $\sqrt{n} (B_{ml} - 0)$, falls within the estimated confidence intervals, $N_T(B_{rs} - B_{ml})$, for each of the 1,000 Monte Carlo runs. Recall that $N_T(B_{rs} - B_{ml})$ is the estimate of the asymptotic distribution, where B_{rs} is the re-sampling estimator from the inner loop, B_{ml} is the Monte Carlo estimator from the outer loop, and 0 is the true value of the parameter. We then compare the percentage of times that the estimated confidence intervals contain the true value of

the estimate to the nominal confidence level of either 90 or 95 percent. Perfect coverage probabilities would be ones that contain $\sqrt{n}(\hat{B}_{ml} - 0)$ 90 and 95 percent of the time, respectively. We would therefore conclude that the best-performing estimator would be the one that provides coverage probabilities closest to those percentages.

Table 1 presents the estimated coverage probabilities for the 90 and 95 percent confidence intervals for the sample size of fifty and subsample sizes of 40, 25, and 10 in the one equation system. As discussed above, to explore the issue of appropriate subsample size, the Monte Carlo studies are run for a variety of different subsample sizes. Recall that in analyzing the performance over estimators, we are looking for the shortest average confidence lengths *and* for coverage probabilities that are closest to the nominal confidence level. The bootstrap, subsample jackknife, and subsample bootstrap provide virtually identical coverage probabilities for each of the three subsample sizes. The coverage probabilities generally lie a little below the nominal coverage levels, except for the subsample jackknife for the subsample size of 10. In this case, the coverage probability is just above the nominal level. The average lengths for the various subsample sizes are fairly similar. The subsample jackknife is the most precise for the subsample sizes of 40 and 25 whereas the subsample bootstrap is the most precise for the subsample size of 10. These results suggest that the three re-sampling methods perform similarly for the sample size of 50. The Monte Carlo results suggest that coverage probabilities and levels of precision are very similar for the bootstrap, subsample bootstrap, and subsample jackknife. Thus, even though the bootstrap is asymptotically inconsistent, it seems to perform well in finite samples.

3.2 Two-equation System

The next step in assessing the relative performance of the three estimators is to examine a more advanced two-equation system that might be more relevant to empirical economists. For example, in production or demand analysis a researcher might want to estimate the following model

$$(2) \quad y_{1t} = A_1 + B_{11}x_{1t} + B_{12}x_{2t} + \mu_{1t}$$

$$(3) \quad y_{2t} = A_2 + B_{12}x_{1t} + B_{22}x_{2t} + \mu_{2t}.$$

where the error terms μ_{1t} and μ_{2t} are assumed to be correlated. A production analogy to (2) and (3) might have representing share equations from a standard three input cost function model (one share equation is dropped due to adding up restrictions) with the usual *a priori* assumption that the model is concave. This implies that the matrix of parameter estimates, B_{11} , B_{12} , B_{21} , and B_{22} is negative semi-definite. We impose this constraint using the Cholesky decomposition and assuming symmetry between the cross-partial. Hence, we end up estimating c_{11} , c_{12} , and c_{22} , where $B_{11} = -(c_{11} \ c_{11})$, $B_{12} = -(c_{11}c_{12})$, and $B_{22} = -(c_{12}c_{12} + c_{22}c_{22})$. Once again, we are interested in evaluating the performance of the three estimators when the parameter is on a boundary. Therefore, when constructing the data, the values for each of the B_{ij} are set to zero, A_1 is set to 2, and A_2 is set to 3. The x_{it} are created in a manner similar to the method detailed above, except that the error terms now have a bivariate normal

distribution with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix}$.

Because we have a two-equation system with correlated error terms, the model is estimated using seemingly unrelated regression. The process for the two-equation Monte Carlo study is very similar to that of the one-equation Monte Carlo study detailed above. The only difference is that instead of estimating the two parameters, A and B , from the one-equation system we are now estimating the five parameters, A_1 , A_2 , B_{11} , B_{12} , and B_{22} , from the two-equation system. Therefore, we are interested in the asymptotic distribution of the three parameters that lie on the boundary, B_{11} , B_{12} , and B_{22} .

Once again, we ultimately want to calculate coverage probabilities and measure of precision to compare and contrast how well the three re-sampling estimators perform in finite samples. These results, contained in Table 2 for a sample size of 50 are fairly similar to the one-equation results. The main difference between the one and two equation results is the dramatic increase in the coverage probabilities for the two subsampling methods as the subsample size decreases. The three methods in

Table 2 have similar coverage probabilities for the subsample size of 40 but for the subsample size of 10 the two re-sampling methods over cover.

General conclusions can be drawn about the relative performances of the three techniques after examining the results in the one and two equation simulation studies. The main implication of these results is that even though the bootstrap is asymptotically inconsistent, it performs very similarly to the two consistent techniques in finite samples. These results also suggest that the subsampling techniques perform similarly for various subsample sizes for the one-equation model and tend to have increased coverage probabilities as sample sizes decrease for the two-equation model. Now that we have analyzed these methods in a Monte Carlo study, we will now discuss how to implement them into economic applications.

4. Empirical Application – Translog Cost Function

Inequality constraints are used extensively in the economics literature in supply and demand analysis to impose well-known microeconomic conditions such as convexity, concavity, and quasi-convexity. The specific classes of supply and demand specifications we focus on in this dissertation are those with flexible functional forms. The idea behind these specifications is to assign to a consumer, firm, or industry a function that requires relatively few restrictions to attain arbitrary elasticities at any given point. These functional forms are usually second-order (or Diewert) flexible and are local approximations to any arbitrary twice continuously differentiable function at any given point in the data space. Examples of local and global flexible functional forms include the translog, generalized Leontief, Box-Cox, Almost Ideal, Asymptotically Ideal, Laurent, Muntz-Szatz, and Fourier flexible forms. In this paper, we focus on the translog model.

The translog model was proposed by Christensen, Jorgenson, and Lau (1975) as a model that is a second-order locally flexible functional form. Here we plan to estimate a translog cost function similar to that used by Diewert and Wales (1987):

$$(4) \quad \ln C(p, y, t) = B_0 + A_y \ln y + \sum_{i=1}^L B_i \ln p_i + .5 \sum_{i=1}^L \sum_{j=1}^L D_{ij} \ln p_i \ln p_j + \sum_{i=1}^L D_{iy} \ln p_i \ln y$$

$$+.5 A_{yy} \ln y \ln y + \sum_{i=1}^L D_{it} \ln p_i t + A_{tt} + .5A_{tt}t^2 + A_{yt} \ln y t$$

where C is cost, p is price, y is output, t is time, and L is the number of inputs. We assume that the cost function is symmetric and linearly homogeneous in factor prices.

In this study we are interested in imposing proper curvature on equation (4). *A priori* restrictions constrain this function to be concave in prices, which is equivalent to requiring the Hessian matrix to be negative semi-definite. A procedure due to Moschini (1998) Ryan and Wales (1998) is to impose concavity at a reference point using the Cholesky decomposition.⁶ This procedure entails redefining the Hessian matrix, H, so that it is equal to B + C, where B is a matrix of parameters of the same order as H and C is a matrix of some or all of the other parameters in the model. After B and C are determined, the Hessian matrix is set equal to -AA', where A is a lower triangular matrix. The equation is then solved for B = -AA' - C. The elements of A and C, rather than the matrix B, are used in estimation. For well-behaved data, this procedure guarantees that the Hessian matrix is negative semi-definite and the model is concave at the reference point. Applying this technique to equation (5), yields $D_{ij} = (-AA')_{ij} - B_i B_j + \delta_{ij} B_i$, where δ_{ij} equals one if $i=j$, zero otherwise, and

$$(5) \quad A = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{12} & a_{22} & 0 \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

so that $(-AA')_{11} = -a_{11}a_{11}$, $(-AA')_{12} = -a_{11}a_{12}$, and so on. The a_{ij} 's are used in estimation in lieu of the D_{ij} 's.

Our application of the translog cost model uses the familiar 1947-71 U.S. manufacturing data from Berndt and Wood (1975). These data contain price and input information about capital (K), labor (L), energy (E), and manufacturing (M) and output of U.S. manufacturing. Therefore, we are interested in estimating the share equations derived from the cost function found in equation (4):

$$(6) \quad s_i(p, y, t) = B_i + \sum_{j=1}^L D_{ij} \ln p_j + D_{iy} \ln y + D_{it} t.$$

Due to adding up restrictions (shares sum to one) one share equation is dropped in estimation, manufacturing, to prevent a singular covariance matrix. Because the error terms may be contemporaneously correlated, we will employ seemingly unrelated estimation.

The elasticities derived from the estimation of the above model contain the economic information of interest. The price-elasticities derived from the translog cost function specified in equation (4) are

$$(7) \quad \epsilon_{ij} = \frac{D_{ij}}{\hat{s}_i} + \hat{s}_j - \delta_{ij} \quad \text{where } \delta_{ij} = 1 \text{ if } i=j \text{ and } 0 \text{ otherwise}$$

where \hat{s}_i is the predicted share from input i .

As mentioned above, a similar model has previously been estimated by Diewert and Wales (1987). However, those authors reported only the elasticities and chose not to report confidence intervals or standard errors. Our analysis corrects this major shortcoming by employing the resampling techniques in order to obtain consistent standard errors and to conduct hypothesis tests on the estimated elasticities in the model.

Estimated price elasticities are contained in Tables 3 for the bootstrap, subsample bootstrap, and subsample jackknife with a sample size of 25 and a subsample size of 18. Point estimates of the elasticities as well as 95% confidence intervals are presented. Elasticities are the same across the three methods because they are computed from the initial data set. It is worth noting that these elasticities have the same sign and are very similar in magnitude to those obtained by Diewert and Wales (1987). Confidence intervals are computed using the modified percentile method described in the simulation study. As is typical in studies that use bootstrap or jackknife methodologies, hypothesis tests are conducted using the empirical distribution generated from the resampling methods rather than relying on asymptotic theory such as t-tests. In the present case confidence intervals are used to perform hypothesis tests in the following standard manner. If the confidence interval contains zero then we conclude that the elasticity is not significantly different from zero. Conversely, if the estimated confidence interval does not contain zero, we conclude that the estimated elasticity is statistically

significant at the 5% level and denote this with an asterisk in the table. The bootstrap and subsample bootstrap arrive at the same conclusions concerning the statistical significance of the elasticities except in the case of the elasticity of labor with respect to materials. The subsample jackknife differs on two occasions from the bootstrap. Specifically, the own-price capital and own-price materials elasticities are insignificant using the subsample bootstrap where they are significant at the 5% method with the other two methods. This implies that conclusions about significance will differ depending on the resampling techniques that are used.

5. Conclusion

We have explored two relatively simple methods, the subsample bootstrap and subsample jackknife, to calculate measures of precision within a classical statistics framework when parameters are inequality constrained. Monte Carlo studies were conducted for one and two equation models to gauge finite sample performance of the bootstrap, subsample bootstrap, and sample jackknife techniques. The three methods were then applied to an inequality constrained cost function with U.S. manufacturing data. Results suggest that the subsample bootstrap, subsample jackknife, and traditional bootstrap perform similarly in both the Monte Carlo and empirical applications.

Future research will entail applying the subsampling methods to additional empirical applications in which the models are curvature constrained such as the Asymptotically Ideal Model as the ultimate goal of this research is to utilize the proposed estimation techniques in meaningful empirical applications. The subsampling techniques will facilitate the computation of standard errors and, as well, will provide a direct way of performing hypothesis tests on point estimates. An implication of these applications is that we will be able to perform statistical tests on the elasticities from a model that conforms to economic theory, something that heretofore has proved challenging if not impossible to do.

Figure 1. Representative Subsample Bootstrap Histogram of the Estimated Asymptotic Distribution.
 Sample size – 50, Subsample size - 40

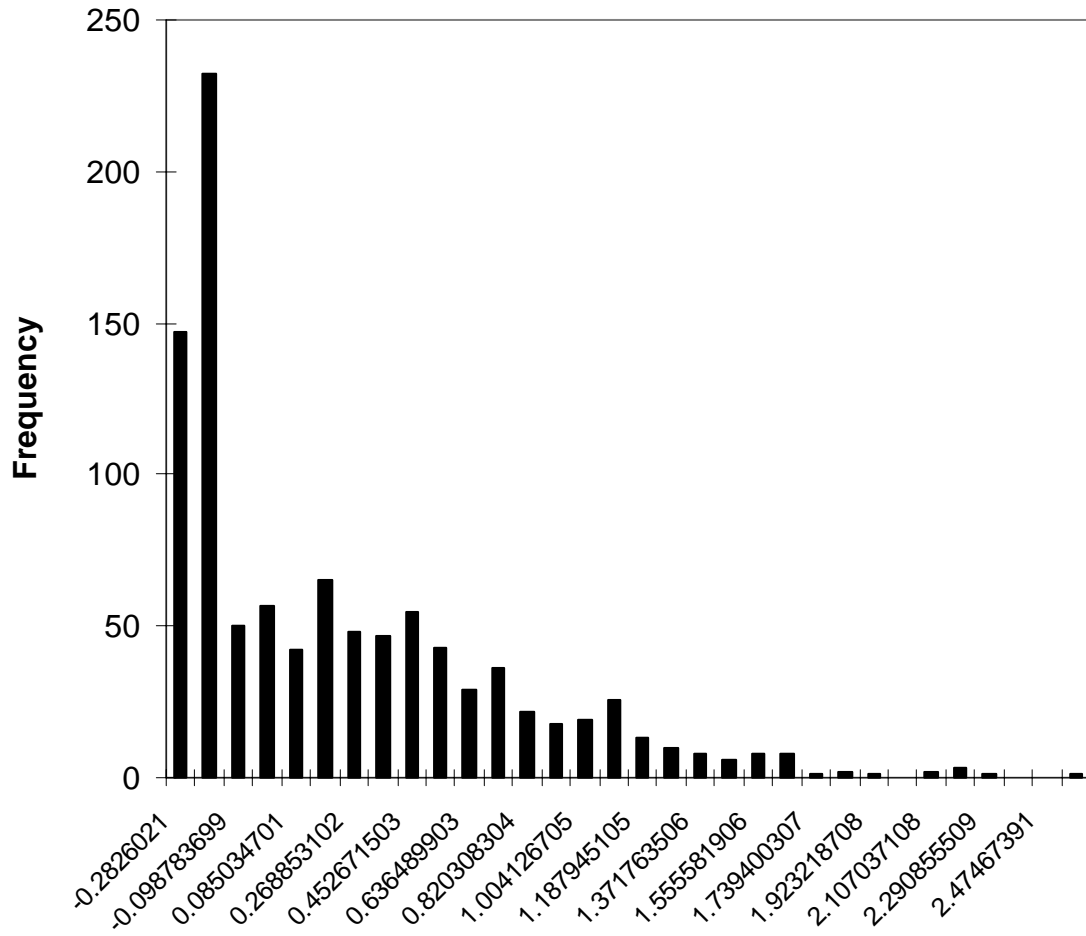


Table 1. Hypothesis Tests Results – Single Equation Model
 Non-Symmetric Interval
 Sample Size 50

Subsample Size: 40

	90%		95%	
	Coverage	Avg. Length	Coverage	Avg. Length
Bootstrap	88.4	1.2971	93.8	1.5867
Subsample Jackknife	89.3	1.2943	93.8	1.5848
Subsample Bootstrap	89.6	1.3056	94.1	1.6024

Subsample Size: 25

	90%		95%	
	Coverage	Avg. Length	Coverage	Avg. Length
Bootstrap	88.4	1.2971	93.8	1.5867
Subsample Jackknife	89.2	1.2552	93.9	1.5488
Subsample Bootstrap	89.3	1.2551	94.3	1.5513

Subsample Size: 10

	90%		95%	
	Coverage	Avg. Length	Coverage	Avg. Length
Bootstrap	88.4	1.2971	93.8	1.5867
Subsample Jackknife	90.6	1.2293	95.5	1.5598
Subsample Bootstrap	87.7	1.1024	93.0	1.3757

Table 2. Hypothesis Tests Results – Two Equation Model
Non-Symmetric Interval
Sample Size 50

Subsample Size: 40

	90%			95%		
	B ₁₁	B ₁₂	B ₂₂	B ₁₁	B ₁₂	B ₂₂
Bootstrap	89.0	90.2	88.2	94.2	95.6	94.6
Subsample Jackknife	90.2	91.6	90.6	94.8	96.8	96.0
Subsample Bootstrap	89.0	90.4	88.6	94.8	95.4	93.4

Subsample Size: 25

	90%			95%		
	B ₁₁	B ₁₂	B ₂₂	B ₁₁	B ₁₂	B ₂₂
Bootstrap	89.0	90.2	88.2	94.2	95.6	94.6
Subsample Jackknife	90.2	91.6	90.6	94.8	96.8	96.0
Subsample Bootstrap	89.0	90.4	88.6	94.8	95.4	93.4

Subsample Size: 10

	90%			95%		
	B ₁₁	B ₁₂	B ₂₂	B ₁₁	B ₁₂	B ₂₂
Bootstrap	89.0	90.2	88.2	94.2	95.6	94.6
Subsample Jackknife	92.8	95.0	95.6	96.8	98.6	99.2
Subsample Bootstrap	92.8	98.4	99.2	97.2	100.0	99.8

Table 3. Translog Cost Function Estimated Price Elasticities and Confidence Intervals.

		Bootstrap	Subsample Bootstrap	Subsample Jackknife
ϵ_{KK}	Estimate	-0.315*	-0.315*	-0.315
	C. I.	(-3.613, -0.304)	(-1.781, -0.304)	(-3.384, 0.176)
ϵ_{KL}	Estimate	0.550	0.550	0.550
	C. I.	(-1.509, 0.941)	(-1.816, 0.793)	(-5.973, 0.956)
ϵ_{KE}	Estimate	-0.102	-0.102	-0.102
	C. I.	(-1.600, 0.616)	(-0.879, 0.428)	(-0.947, 2.668)
ϵ_{KM}	Estimate	-0.133	-0.133	-0.133
	C. I.	(-0.352, 5.894)	(-0.460, 3.509)	(-0.677, 8.906)
ϵ_{LK}	Estimate	0.107	0.107	0.107
	C. I.	(-0.264, 0.150)	(-0.360, 0.177)	(-0.997, 0.171)
ϵ_{LL}	Estimate	-0.359*	-0.359*	-0.359*
	C. I.	(-5.339, -0.292)	(-7.064, -0.222)	(-4.972, -0.36)
ϵ_{LE}	Estimate	0.052	0.052	0.052
	C. I.	(-0.515, 0.569)	(-0.585, 1.115)	(-0.090, 1.779)
ϵ_{LM}	Estimate	0.200*	0.200	0.200*
	C. I.	(0.151, 5.064)	(-0.015, 6.188)	(0.110, 4.862)
ϵ_{EK}	Estimate	-0.122	-0.122	-0.122
	C. I.	(-2.085, 0.574)	(-0.913, 0.494)	(-1.168, 3.158)
ϵ_{EL}	Estimate	0.318	0.318	0.318
	C. I.	(-3.671, 4.482)	(-2.614, 6.490)	(-0.879, 12.464)
ϵ_{EE}	Estimate	-0.592*	-0.592*	-0.592*
	C. I.	(-5.656, -0.102)	(-4.280, -0.161)	(-12.167, -1.186)
ϵ_{EM}	Estimate	0.396	0.396	0.396
	C. I.	(-0.965, 7.536)	(-2.227, 4.543)	(-4.780, 4.304)

Notes: * denotes significance at the 5% level. The interval is at 95% confidence.

Table 3 Cont. Translog Cost Function Estimated Price Elasticities and Confidence Intervals.

		Bootstrap	Subsample Bootstrap	Subsample Jackknife
ε_{MK}	Estimate	-0.011	-0.011	-0.011
	C. I.	(-0.031, 0.535)	(-0.050, 0.306)	(-0.050, 0.637)
ε_{ML}	Estimate	0.087*	0.087*	0.087*
	C. I.	(0.061, 2.259)	(0.005, 3.036)	(0.043, 1.722)
ε_{ME}	Estimate	0.028	0.028	0.028
	C. I.	(-0.088, 0.552)	(-0.176, 0.408)	(-0.235, 0.279)
ε_{MM}	Estimate	-0.104*	-0.104*	-0.104
	C. I.	(-2.627, -0.108)	(-3.286, -0.062)	(-2.080, -0.034)

Notes: * denotes significance at the 5% level. The interval is at 95% confidence.

References

- Appelbaum, E. and A. Ulah (1997): "Estimation of Moments and Production Decisions Under Uncertainty," *Review of Economics and Statistics*, 79 (4), 631-637.
- Andrews, D.W.K. (1999): "Estimation when a Parameter is on a Boundary," *Econometrica*, 67 (6), 1341-1383.
- Andrews, D.W.K. (1998): "Hypothesis Testing with a Restricted Parameter Space," *Journal of Econometrics*, 84, 155-199.
- Andrews, D. W. K. (1997a): "A Simple Counterexample to the Bootstrap," unpublished manuscript, Cowles Foundation for Research in Economics, Yale University.
- Andrews, D. W. K. (1997b): "Estimation When a Parameter Is on a Boundary: Theory and Applications," unpublished manuscript, Cowles Foundation for Research in Economics, Yale University.
- Barnett, W.B., J. Geweke, and M. Wolfe (1991): "Semiparametric Bayesian Estimation of the Asymptotically Ideal Production Model," *Journal of Econometrics*, 49, 5-50.
- Barnett, W.B. and A.B. Jonas (1983): "The Muntz-Szatz Demand System: An Application of a Globally Well Behaved Series Expansion," *Economics Letters*, 11, 337-342.
- Barnett, W.A. and Y. Lee (1985): "The Global Properties of the Miniflex Laurent, Generalized Leontief, and Translog Flexible Functional Forms," *Econometrica*, 53(6), 1421-1437.
- Barnett, W.B. and P. Yue (1988): "Semiparametric Estimation of the Asymptotically Ideal Model: The AIM Demand System," *Advances in Econometrics*, Vol. 7, 229-251.
- Berndt, E.T., and M.S. Khaled (1979): "Parametric Productivity Measurement and Choice Among Flexible Functional Forms," *Journal of Political Economy*, 87, 1220-1245.
- Bickel, P.J. and D. Freeman (1981): "Some Asymptotic Theory for the Bootstrap," *Annals of Statistics*, 9, 1196-1217.
- Bickel, P.J., F. Gotze, and W.R. van Zwet (1997): "Resampling Fewer than n Observations: Gains, Losses, and Remedies for Losses," *Statistica Sinica*, 7, 1-31.
- Bickel, P.J. and J.J. Ren (1995): "The m out of n Bootstrap and Goodness of Fit Tests with Double Censored Data," *Robust Statistics, Data Analysis, and Computer Intensive Methods*, Ed. H. Rieder Lecture Notes in Statistics, Springer-Verlag.
- Chalfant, J.A. (1987): "A Globally Flexible, Almost Ideal Demand System," *Journal of Business and Economic Statistics*, 5(2), 233-242.

- Chalfant, J.A. and A.R. Gallant (1985): "Estimating Substitution Elasticities with the Fourier Cost Function," *Journal of Econometrics*, 28, 205-222.
- Chalfant, J.A., R.S. Gray, and K.J. White (1991): "Evaluating Prior Beliefs in a Demand System: The Case of Meat Demand in Canada," *American Journal of Agricultural Economics*, May, 476-490.
- Chalfant, J.A. and R.S. White (1988): "Estimation of Demand Systems with Concavity and Monotonicity Constraints," Working Paper, Dept. of Agr. And Res. Econ., University of California, Berkeley.
- Demos, A. and E. Sentana (1998): "Testing for GARCH Effects: A One-Sided Approach," *Journal of Econometrics*, 86, 97-127.
- Diewert, W.E. and T. J. Wales (1987): "Flexible Functional Forms and Global Curvature Conditions," *Econometrica*, 55, 43-68.
- Diewert, W.E. and T. J. Wales (1998): "Normalized Quadratic Systems of Consumer Demand Functions," *Journal of Business and Economic Statistics*, 6(3), 303-312.
- Efron, B. (1979): "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7, 1-26.
- Fleissig, A.R. and J.L. Swofford (1997): "Dynamic Asymptotically Ideal Models and Finite Approximation," *Journal of Business and Economic Statistics*, 15(4), 482-492.
- Gallant, A.R. and G.H. Golub (1984): "Imposing curvature restrictions on flexible functional forms," *Journal of Econometrics*, 26, 295-321.
- Godfrey, L.G. and M.R. Veal (1998): "Bootstrap based critical values for tests of common factor restrictions," *Economics Letters*, 59, 1-5.
- Gourieroux, C., A. Holly, and A. Monfort (1982): "Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters," *Econometrica*, 50 (1), 63-80.
- Havener, A. and A. Saha. (1999): "Globally Flexible Asymptotically Ideal Models," *American Journal of Economics*, 81, 703-710.
- Hogny, L. and G.S. Maddala (1996): "Bootstrapping Time Series Models," *Econometric Reviews*, 15(2), 115-158.

- Koop, G., J. Osiewalski, and M.F.J. Steel (1994): "Bayesian Efficiency Analysis with a Flexible Form: The AIM Cost Function," *Journal of Business and Economic Statistics*, 12(3), 339-346.
- O'Donnell, C.J., C.R. Shumway, and V.E. Ball (1999): "Input Demands and Inefficiency in US Agriculture," *American Journal of Agricultural Economics*, 81(4), 865-880.
- Politis, D. N. and J. P. Romano (1994): "Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions," *Annals of Statistics*. 22, 2031-2050.
- Politis, D.N., J.P. Romano, and M. Wolf (1997): "Subsampling for heteroskedastic time series," *Journal of Econometrics*, 81, 281-317.
- Quenouille, M.H. (1949): "A Note on Bias in Estimation," *Biometrika*, 43, 353-60.
- Ryan, D.L. and T.J. Wales (1999): "Flexible and Semi-flexible Consumer Demands with Quadratic Engel Curves," *Review of Economics and Statistics*, 81(2), 277-87.
- Ryan, D.L. and T.J. Wales (1998): "A Simple Method for Imposing Local Curvature in Some Flexible Consumer Demand Systems," *Journal of Business and Economics Statistics*, 16(3), 331-38.
- Self, S. G. and K. Liang (1987): "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions," *Journal of the American Statistical Association*, 82 (398), 605-610.
- Terrell, D. (1995): "Flexibility and Regularity Properties of the Asymptotically Ideal Production Model," *Econometric Reviews*, 14(1), 1-17.
- Thornton, D. and P. Yue (1992): "An Extended Series of Divisia Monetary Aggregates," *Quarterly Review of Federal Reserve Bank of St. Louis*, Nov/Dec., 35-52.
- Tukey J.W. (1958): "A Bias and Confidence in Not-quite Large Samples," *Annals Mathematical Statistics* (abstract), 29, 614
- Wales, T.J. (1977): "On the Flexibility of Flexible Functional Forms: An Empirical Approach," *Journal of Econometrics*, 5, 183-193.
- Wolak, F. (1989a): "Local and Global Testing of Linear and Nonlinear Inequality Constraints in Nonlinear Econometric Models," *Econometric Theory*, 5, 1-35.
- Wolak, F. (198b): "Testing Inequality Constraints in Linear Econometric Models," *Journal of Econometrics*, 41, 205-235.
- Wolak, F. (1991): "The Local Nature of Hypothesis Testing Involving Inequality Constraints in Nonlinear Models," *Econometrics*, 59, 981-995.

Wu, C. F. J. (1990): "On the Asymptotic Properties of the Jackknife Histogram," *Annals of Statistics*, 18, 1438-1452.

Ziari, H.A., D.J. Leatham, and P.N. Ellinger (1997): "Development of Statistical Discriminant Mathematical Programming Model via Resampling Estimation Techniques," *American Journal of Agricultural Economics*, 79, 1352-1362.

² Note that this regularity condition does not apply to the Lagrange multiplier test because it is not affected by the fact that the parameters of interest lie at a boundary of the parameter space under the null hypothesis. However, it is obvious that this is not the optimal test as it does not take into account the directional nature of the hypothesis, which implies that the Lagrange multiplier test is not the most powerful option.

³ Andrews (1999) provides assumptions and sufficient conditions under which the asymptotic results hold. The proof of the validity of these methods is an extension of a method pioneered by Chernoff (1954) that approximates the parameter space by a cone in the context of likelihood ratio tests for iid data. The cone approximation allows a boundary to be linear, curved, or kinked. Andrews extends Chernoff's results to allow for trends in the data. The results are general enough to encompass a variety of objective functions, including log-likelihood, quasi-log likelihood, least squares, and GMM estimators. The results hold whether or not the true parameter lies on a boundary of the parameter space.

⁴ Doubly censored data occur in studies of age-dependent, life-threatening diseases. The Cramer-von Mises goodness-of-fit test is conducted in order to gauge whether the assumptions that are made on the underlying lifetime distributions are legitimate.

⁵ The modified percentile method for a $1 - \alpha$ confidence interval, with B bootstrap replications, and μ^* as the bootstrap distribution, is computed by minimizing $.5(\mu^*(1+(1-\alpha)B) + \mu^*(1+(1-\alpha)B+1)) - .5(\mu^*(1-1) + \mu^*(1))$ with respect to l which is less than $.05B$. This method finds the shortest interval that includes $(1 - \alpha)\%$ of the observations.

⁶ Lau (1978) proposed the use of the Cholesky decomposition method to impose and test for curvature constraints in flexible functional forms.