Economics Bulletin

Volume 30, Issue 2

A note on nonidentification in truncated sampling distribution estimation

William Barnett
University of Kansas

Ousmane Seck
University of Texas at El Paso

Abstract

Theoretical constraints on economic model parameters often are in the form of inequality restrictions. For example, many theoretical results are in the form of monotonicity or nonnegativity restrictions. Inequality constraints can truncate sampling distributions of parameter estimators, so that asymptotic normality no longer is possible. Sampling theoretic asymptotic inference is thereby greatly complicated or compromised. In Barnett and Seck (2009), which will be appear in volume 1 number 1 of the new journal, Journal of Statistics: Advances in Theory and Applications, we use numerical methods to investigate the resulting sampling properties of estimation with inequality constraints, with particular emphasis on the method of squaring, which is the most widely used method in applied literature on estimating integrable neoclassical systems of equations. In this note, we make our most important results more widely and easily available.

Citation: William Barnett and Ousmane Seck, (2010) "A note on nonidentification in truncated sampling distribution estimation", *Economics Bulletin*, Vol. 30 no.2 pp. 1670-1679.
Submitted: Sep 26 2009. **Published:** June 19, 2010.

1. Introduction

Using Monte Carlo experiments, we investigate the possible bias in the asymptotic standard errors of estimators of inequality constrained estimators, when the constraint is imposed by the popular method of squaring. That approach is known to violate a regularity condition in the available asymptotic proofs regarding the unconstrained estimator, since the sign of the unconstrained estimator, prior to squaring, is nonidentified. Most existing theoretical results on asymptotics subject to inequality constraints condition upon linearity of the model, while most integrable neoclassical demand and supply system models are nonlinear. See Barnett and Binner (2004). But even in the case of linear models, the regularity conditions used in the existing asymptotic proofs are violated by the nonidentification of the sign of the transformed parameter in the method of squaring. See. e.g., Gourieroux and Monfort (1982), Gourieroux and Monfort (1995, p. 247), Rothenberg (1971), and Silvapulle and Sen (2005, section 4.9).

Consider the following transformation approach, widely used to impose inequality constraints in econometrics. If g is a continuous function of θ, and β is the constrained parameter, each approach acquires point estimates of β from the transformation $\beta = g(\theta)$, where g is chosen such that $g(\theta)$ satisfies the relevant inequality constraint for all unconstrained values of θ. The constrained parameter β is replaced within the regression by $\beta = g(\theta)$, and the parameter θ is estimated without constraints. The unconstrained parameter can be estimated by maximum likelihood, and the constrained parameter estimate can be recovered from the invariance property of maximum likelihood estimator.[1] No compromise in the approach to point estimation is implied by truncation of the sampling distribution, but computation of the standard error of the constrained estimator presents problems. We focus on the delta method, the Jackknife, and the Bootstrap, among the most popular sampling theoretic approaches used to address problems stemming from truncation of sampling distributions.

The most widely used method for imposing nonnegativity constraints is the method of squaring. However the non-identification of the sign of the unconstrained parameter inherent to the method of squaring is a major potential problem, when the delta method approach is employed to estimate the standard errors of the constrained parameter. When using the method of squaring to impose nonnegativity on $\beta_i = g_i(\theta_i)$, the estimation of $g_i(\hat{\theta}_i)$ cannot distinguish between $-\hat{\theta}_i$ and $+\hat{\theta}_i$. Hence, one of the regularity conditions is violated in the asymptotic proof with the delta method. We investigate the extent of the damage by using the delta method, when the sign of θ_i is nonidentified.

Our primary objective is to determine whether $Y = \sqrt{N}\left[g(\hat{\theta}) - E\, g(\hat{\theta})\right]$ has a limiting distribution providing accurate measures of its standard deviation. Other properties of the limiting distribution are not relevant to this study, and limiting normality is impossible for Y with the distribution of $g(\hat{\theta})$ being truncated at the origin. Nevertheless, it is possible that enough properties of the limiting distribution may be undamaged so that limiting normality of Y cannot be rejected empirically. Since we are only concerned with the first two moments, the unavoidable errors in the higher order moments (that do not exist with the normal distribution) need not concern us. In fact our objective is focused solely on convergence of the standard

---

[1] The maximum likelihood estimator of $\beta = g(\theta)$ is $g(\hat{\theta}_{ML})$.

deviation, which remains possible, even if the distribution cannot converge to a limiting normal. We provide the most important results from Barnett and Seck (2009).

It should be observed that the delta method usually is often used, with $\hat{\boldsymbol{\theta}}$ assumed to be asymptotically normal and the stronger conclusion than we use is that $\hat{\boldsymbol{\beta}} = \mathbf{g}(\hat{\boldsymbol{\theta}})$ is asymptotically normal. But since we are exploring the implications of truncation of the distribution of $\hat{\boldsymbol{\beta}} = \mathbf{g}(\hat{\boldsymbol{\theta}})$, asymptotic normality is not possible. Our concern is only with the first two moments of the limiting distribution.[2]

2. Monte Carlo Experiment

In this section, we describe the process that led to the generation of data with known characteristics. In particular, we want the true value of our parameter of interest to satisfy the nonnegativity constraint which we investigate. The two-good Constant Elasticity of Substitution (CES) utility function is a typical model having the ability to estimate the elasticity of substitution ($\sigma > 0$) between two goods, and is suitable choice for our illustration. It is globally flexible and globally regular. That model will be used to provide parameter values used as a "norm" for illustration. But results with only one vector of parameter values are of limited value, without confirmation that the results are robust to the parameter value choices. In fact, we ran our Monte Carlo simulations with different values of the parameters. Since we found our results to be robust to different parameter settings, we reported the results only for our one (admittedly arbitrary, but currently interesting) calibrated "norm" settings of model parameters.[3]

In producing our parameter setting norm, we started with real world data by looking at the relationship between two monetary assets . We then simulated two goods assumed to be substitutable to some degree, so that they are subject to the inequality constraint $\sigma > 0$ (perhaps monetary assets, but only used as an illustration in the one calibrated case). With the simulated data described below, we estimate the demand model with the simulated data subject to that inequality constraint, using the method of squaring by applying the reparameterization, $\sigma = 10^{-20} + 0.01\theta^2$, while alternatively the exponential transformation approach is implemented by applying the reparameterization, $\sigma = 0.00001e^{\theta}$.

The next section describes the data generation process and the estimation results. There are two objectives of our Monte Carlo experiment: (1) assess the potential damage to the asymptotic standard errors of $\widehat{\beta} = \widehat{\sigma} = g(\widehat{\theta})$, resulting from the indeterminacy of the sign of the squared parameter $\widehat{\theta}$ in the method of squaring[4] and (2) determine the asymptotic properties of the constrained parameter $\widehat{\beta}$, when the jackknife and the bootstrap are used to calculate the finite sample standard errors, with sample sizes permitted to increase to large values.
In our simulations, the model parameters σ, and γ are set at various values, but since our results were robust to the setting of those parameters, we provided illustrative figures only for the case calibrated to have $(\sigma,\gamma) = (0.37, 2.8)$ (see footnote 6 for the definition of γ).

2.1. Data Generation Process

[2] As we discuss below, problems with higher order moments are unavoidable.
[3] The SAS code and outputs with other parameter settings are available upon request.
[4] In this context, $g(\theta) = 10^{-20} + 0.01\theta^2$

The data generation process proceeds in six steps as shown in figure 1, following the setting of the values of the parameters. Our data set will consist of three variables: the user costs of the two assets and the expenditure share of one of the asset (w_1).

Step 1: Generate three series of 100,000 random numbers that will be the seeds for generating two user costs series and the white noise errors.
Step 2: Generate two stationary series containing *S* observations and representing the user costs of two categories of assets { $\pi_t^{(1)}$ and $\pi_t^{(2)}$: *t* = 1, 2, 3,…, *S*]. We generated that data from the following simple stationary specifications: $\pi_t^{(1)} = 2 + 6v_1$ and $\pi_t^{(2)} = 2 + 6v_2$, where v_1 and v_2 are uniformly distributed between 0 and 1.[5]
Step 3: Use the demand function in expenditure share [6] to generate a series of expenditure shares of asset 1, $w_t^{(1)}$, with the true values of the parameters set at $\sigma = 0.37$, $\gamma = 2.8$. The expenditure share of asset 2 are then derived from $w_t^{(1)} + w_t^{(2)} = 1$.
Step 4: Generate a white noise error term series with mean zero and standard deviation equal to 0.04.
Step 5: Add the errors created in step 4 to the series of expenditure shares of asset 1 from step 3. The resulting realized stochastic shares are designated by *fw1*.
Step 6: The set of increasing sample sizes are chosen to be $S \in$ {30, 45, 60, 100, 200, 400, 800, 1000, 2000, 3000, 4000 ,..., 100000}.

2.2. Estimation Results

We employ maximum likelihood to estimate the demand function in footnote 6 with w_t^1 replaced by the noise augmented data generated in section 2.1. By construction, the true value of our parameter of interest, σ, is positive, and it is estimated by imposing the positivity constraint using the method of squaring with $\sigma = 10^{-20} + 0.01\theta^2$ and alternatively by using the exponential transformation, $\sigma = 0.00001e^{\theta}$.
For every generated sample of size S, we estimate the model using the method of squaring first and then by using the exponential transformation. If the parameter estimation converges as *S* increases with the method of squaring, we consider the trial to be successful. This procedure is repeated 1000 times and the parameter estimates from the first 220 successful experiments are

[5] We considered using simulated autogressive price data, but the nature of those stochastic processes seems unrelated to the truncation and sign-identification issues that are our focus.

[6] The demand function for asset 1 in expenditure share form is as follows: $w_t^1 = \frac{\alpha_1^\sigma (\pi_t^{(1)})^{1-\sigma}}{\alpha_1^\sigma (\pi_t^{(1)})^{1-\sigma} + \alpha_2^\sigma (\pi_t^{(2)})^{1-\sigma}}$
where the elasticity of substitution between the two goods is σ, with $\sigma = 1/(1-\rho)$. The constraint $\rho < 1$ on the parameter of the CES utility function implies $\sigma > 0$. The subscript *t* represents time, and $\pi_t^{(1)}$ and $\pi_t^{(2)}$ are the user costs of assets 1 and 2 respectively. With the parameter ${\alpha_2}^{\sigma}$ normalized to be 1, we change the notation for ${\alpha_1}^{\sigma}$ to *γ*, leaving two parameters to be estimated: *γ* and *σ*.

collected to compute $\sqrt{N}\left[g(\hat{\theta})-Eg(\hat{\theta})\right]$, with N being the sample size, set at the increasing values of S.[7]

The reported results pertain to the asymptotic properties of $\sqrt{N}\left[g(\hat{\theta})-Eg(\hat{\theta})\right]$. The method of squaring was implemented by defining $g(\theta)=10^{-20}+0.01\theta^2$ and the exponential transformation by defining $g(\theta)=0.00001.\exp(\theta)$.[8] We plotted the estimated standard deviation of the limiting distribution of $\sqrt{N}\left[g(\hat{\theta})-Eg(\hat{\theta})\right]$ with the two reparameterizations (method of squaring and exponential transformation). These results were acquired from the delta method's asymptotic distribution theory, but with increasing simulated sample sizes. The results were almost identical, which demonstrates that the estimated asymptotic standard errors do not depend on the transformation used to impose the inequality constraint, or the nonidentification of the sign of the unconstrained parameter with the method of squaring. The exponential transformation and the method of squaring perform equally well. As the sample size increases, the estimated standard deviation of $\sqrt{N}\left[g(\hat{\theta})-Eg(\hat{\theta})\right]$ converges to approximately 0.42 in both cases. This convergence tends to support the use of the asymptotic theory.

These results are consistent with the directly computed finite sample estimated standard deviation of $\sqrt{N}\left[g(\hat{\theta})-Eg(\hat{\theta})\right]$ from the Monte Carlo simulation results. The standard error again converges to approximately 0.42 as the sample size increases. We view 0.42 thereby as being the correct limiting standard deviation against which all other computations should be compared.

The second objective of our research was pursued by analyzing the evolution of the finite sample estimated standard deviation of $\sqrt{N}\left[g(\hat{\theta})-Eg(\hat{\theta})\right]$ for increasing sample size, when the bootstrap and the jackknife are utilized. The jackknifed standard deviation appears to be stationary around 0.22, which is almost half the standard deviation of the limiting distribution.

The bootstrap performs better than the jackknife, since the bootstrapped standard deviation does converge to the estimated standard deviation of the limiting distribution of *Y*, as the sample size increases, while the jackknifed standard deviations are consistently lower than the bootstrapped standard deviation. Figure 2 shows that this result is a consequence of the relatively small proportion, *k*, of jackknife observations deleted. After almost 90 percent of the sample is deleted, the jackknifed finite-sample standard deviation of *Y* does converge to the estimated standard deviation of the limiting distribution of *Y*. These results strongly argue against the jackknife, in such applications as consumer demand modeling, where very large sample size is the exception rather than the rule.

The bootstrap standard deviation of *Y* performs very similarly to the estimated standard deviation from the theoretical limiting distribution, as figure 3 shows. Not only are the two very similar to each other at all sample sizes, but converge to each other as sample size grows.

---

[7] This number of replications, 1000, is arbitrary but its only importance is to guarantee that each sample of parameter estimates will have 220 observations.
[8] As mentioned in a prior footnote above, we also ran our model with different values of the constrained parameter (elasticities of substitution), and those results are available upon request.

While we know that limiting normality is impossible for a truncated distribution, our normality tests failed to reject normality. However, we cannot take seriously limiting normality with truncation, since the normal distribution has no moments higher than the second moment, while a truncated distribution does. Nevertheless, empirical inability to reject limiting normality could strengthen our ability to use the first two moments from the limiting distribution in producing asymptotic inferences, since the first two moments have particularly heavy influence on normality tests.

We were only concerned in this paper about whether or not the asymptotic theory is adequate for certain properties --- in particular standard errors. Our numerical experiments demonstrate that the asymptotic theory, using the delta method, is undamaged by the sign of the unconstrained parameter being nonidentified. Our results with tests of limiting normality suggest that there are properties of the limiting distribution that also are undamaged, at least approximately, but we do not pursue the implications for other properties of the limiting distribution. Clearly higher order limiting moments cannot be used, since the normal distribution has no moments higher than the second moment, while the truncated distribution caused by inequality constraint on the parameters displays existence of higher order moments, such as skewness towards the right.

3. Conclusion

In this paper, our goal is to investigate the empirical implication of inequality constraints imposed on the parameters of a regression. In particular, we are interested in knowing the asymptotic implications of the nonidentified sign of the unconstrained parameter in the method of squaring. While that nonidentified sign violates the regularity conditions of the currently available asymptotic proofs with the delta method, we cannot rule out the possibility that the usual asymptotic properties of the constrained parameter still apply, despite the unavailability of a theoretical proof. As a result, we explore that issue using numerical Monte Carlo methods. Results with the popular method of squaring were compared to results with the exponential transformation, which violates different regularity conditions of available theoretical asymptotic proofs.[9]

We find that the theoretical regularity conditions violations do not affect the usefulness of existing asymptotic theory in determining standard errors of the constrained parameter estimates by the delta method. In addition, the results were not sensitive to the functional form used to impose the inequality constraint. We have not attempted to weaken the existing asymptotic proofs for the delta method to permit the nonidentified sign of the unconstrained parameter estimates. But our Monte Carlo results demonstrate that the nonidentified sign does not compromise the asymptotic standard errors. It should be emphasized that the regularity assumptions in the existing proofs are sufficient but not necessary for the results on the variance of the limiting distribution.

Our second result compares the estimated standard errors from the jackknife and the bootstrap. We find that the finite sample bootstrapped standard errors and the estimated standard errors from the limiting distribution of the constrained parameter estimate converge to each other. However, the finite sample jackknifed standard errors is an increasing function of the

[9] Any transformation that produces truncated sampling distribution for the transformed parameters inherently must violate the existing proofs, which produce the excessively strong result of asymptotic normality.

proportion of the sample deleted within that procedure. For that reason, the bootstrap dominates the jackknife, even though the finite sample jackknifed standard errors are lower than the finite sample bootstrapped standard errors.

Figure 1: Data generation steps

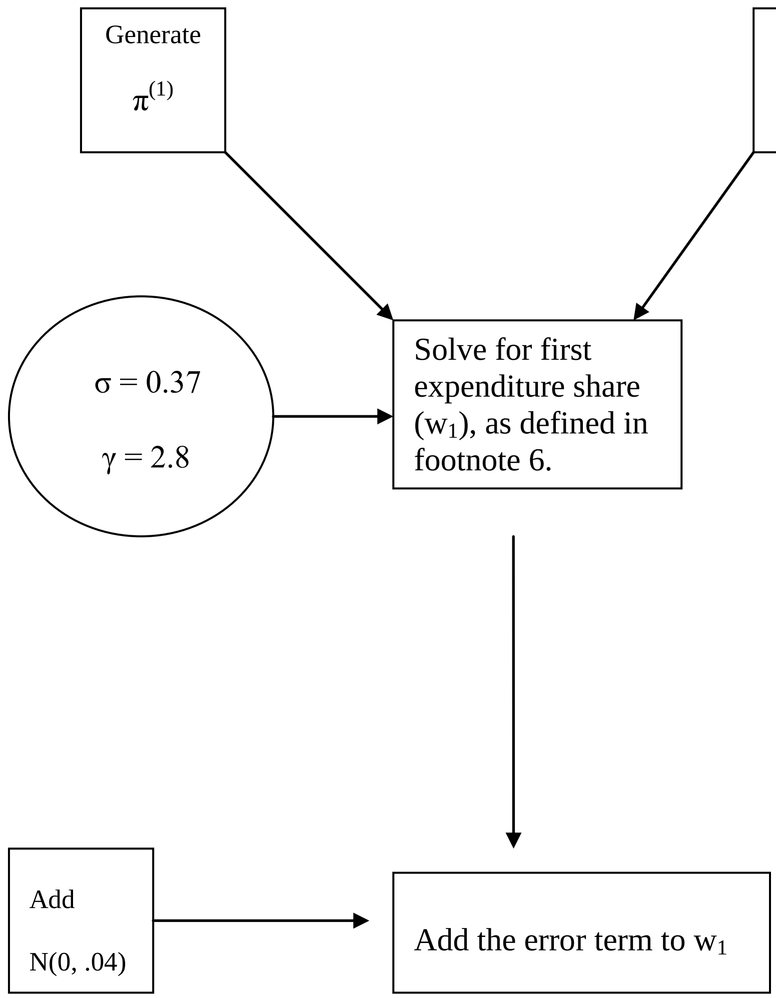


Figure 2: Finite sample estimated standard deviation of $\sqrt{N}[g(\hat{\theta}) - Eg(\hat{\theta})]$ where N=800, as the percentage of the sample deleted, k, increases (Jackknife)

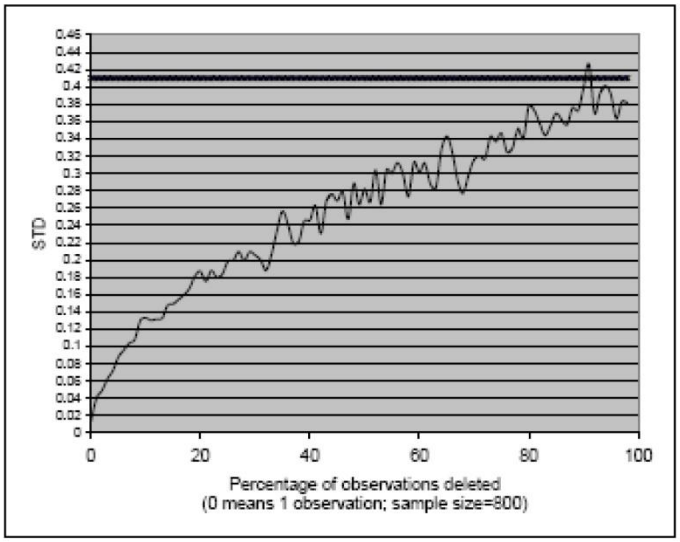
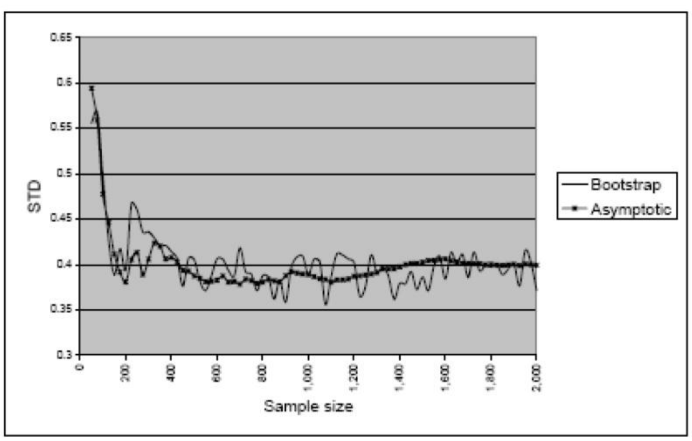


Figure 3: Bootstrapped versus asymptotic standard deviation of the limiting distribution of $\sqrt{N}[g(\widehat{\theta}) - Eg(\widehat{\theta})]$, as N increases to 2000



6. References

Barnett, WA, Binner, JM. (2004) *Functional structure and approximation in econometrics*, Amsterdam: Elsevier.

Barnett, WA, Seck, O. (2009) "Estimation with Inequality Constraints on Parameters and Truncation of the Sampling Distribution", *Journal of Statistics: Advances in Theory and Applications*, **vol 1, no 1**, pp. 65-88.

Gourieroux, C., Monfort, A. (1982) "Likelihood ratio, Wald and Kuhn-Tucker tests in linear models with inequality constraints", *Econometrica* **50**, pp. 43-62.

Gourieroux, C., Monfort, A. (1995) *Statistics and econometric methods*, **Vol 2**, Cambridge, MA: Cambridge University Press.

Rothenberg, T. J. (1971), "Identification of parametric models", *Econometrica*, **29**, pp. 577-592.

Silvapulle, M. and Sen, P.K. (2005) *Constrained statistical inference: inequality, order, and shape restrictions*, John Wiley Publishers, Hoboken, NJ.