Journal of Economics and Econometrics Vol. 53, No. 2, 2010 pp. 75-93

ISSN 2032-9652 E-ISSN 2032-9660

Frontier Techniques: Contrasting the Performance of (Single) Truncated Order Regression Methods and Replicated Moments

ANA PAULA MARTINS[1]

ABSTRACT

This research contrasts three econometric alternatives for stochastic efficiency frontier analysis: order - inter-quantile - and inverse order regression under the assumption of truncated error term distribution, and replicated moment estimation. The demonstration departs from a simple linear regression form of the effective frontier; truncated (at zero) errors are then added to it for simulation purposes. For order regression, experiments with the standard normal, uniform, exponential, Cauchy and logistic error terms are provided. For complex error structures we rely on normal distributions only. The three alternatives would perform satisfactorily for simple error disturbances, specially if they are normal. With more than one residual added to the dependent variable, the weight of the unrestricted range one can blur the conclusions regarding observation efficiency.

JEL Classification: C10, C24.
Keywords: Stochastic Frontier Model, Generalised Method of Order Statistics, Minimum Distance Method of Order Statistics, Inverse Order Regression, Replicated Moments, Linear Models.

---

1 Faculdade de Ciências Económicas e Empresariais, Universidade Católica Portuguesa, Cam. Palma de Cima, 1600 Lisboa, Portugal. Email: apm@europa.fcee.ucp.pt.

1 INTRODUCTION

1.1 Background

Efficiency analysis is *ex-post* tested with the use of a variety of quantitative methods.[2] These sometimes rely on the estimation of a efficient frontier, with distance to it being an indicator of the observation performance. It is the purpose of this note to compare the results of three efficiency frontier estimators: straightforward least squares, adding the minimum or subtracting the maximum estimated error – according to whether the efficiency being measured is, say, cost or revenue - to the estimated model residuals, the method of order statistics towards a truncated error distribution and a replicated moment one.

The analysis relies on simulation, departing from a simple regression model to both illustrate and compare the performance of the methods. Two environments are staged: in one, a truncated at zero normal, uniform, exponential, Cauchy and logistic error terms are then to a deterministic linear model, providing a simple linear regression departure. In a second attempt, an extra normal untruncated random error is additionally included.

For all series, methods evaluate both a lower as an upper truncation hypothesis and none at all. We would hope that the true assumption would emerge with the best performance.

In the order – interquantile – estimation we considered only three alternatives for the null hypothesis: the normal, the exponential, and the uniform itself. The method rely on a two step estimation procedure, departing from rankings of the (first step) OLS residual estimates.[3] It suggests direct inference, and an indirect approach relying on inverting the direct form. The latter is also subject to estimation by the method of moments in a "replicated" version.

Replicated moment estimation was previously forwarded in the literature.[4] One can justify it in linear regression if we note that for a

---

[2] See Koop (2003), p. 147 and 168-177, for references on stochastic frontier modelling; Murillo-Zamorano (2004), for a recent survey of both non-parametric – as Data Envelope Analysis – as parametric and stochastic – as in the present research; and Greene (1997).

[3] See Martins (2005).

[4] Martins (2003).

model with k parameters we have in fact nxk statistics – k, the dependent variable and k-1 independent variables, for each observation. We essayed with the straight-forward replication, with a weighted least squares and a generalised least squares one.

For simplicity, efficiency is modelled additively and the simulations depart from a linear deterministic counterpart. The methods can easily be adapted to apply to nonlinear frontiers, the logarithm of which could be subject to the proposed procedures; inefficiency would then be multiplicative towards the deterministic optimum.

The exposition proceeds as follows: section 1 describes the generated basic random series used in the simulations, summarising briefly theoretical foundations behind the data generating procedures under truncated assumptions. Section 2 applies the "method of order statistics"[5] – in minimum distance versions – an inter-quantilic inference method. The inverse approach is forwarded in section 3. Section 4 illustrates results from replicated method of moments applied to the pillar equation the last procedure. Some concluding remarks end the research.

1.2 Data and Data Generating Procedures

The generation of random samples started by independent draws from the uniform distribution,[6] W_i's, inverted according to the required cdf – a procedure justified by a well-known:

Theorem. For any random variable Z with uniform distribution in the (0, 1) interval – i.e., with cdf U(z) = z and pdf u(z) = U'(z) = 1, for $0 < z < 1$ -, $X = F^{-1}(Z)$ – where F(z) exhibits properties of a cdf in the appropriate domain of z and $F^{-1}(x)$ denotes the inverse function of F(x) - has pdf f(x) = F'(x), and cdf F(X).

---

[5] Alternatively, we would name it rank regression… Yet, rank estimation seems a term more closely (and already) associated to robust – non-parametric – methods in the econometrics literature.

[6] Using the RAND(.) function of EXCEL.

Proof: Using variable transformation, the pdf of X = G(Z) = F^{-1}(Z), which implies, Z = F(X) = G^{-1}(X), is u[G^{-1}(x)] $\frac{dG^{-1}(x)}{dx}$, with z = G^{-1}(x) = F(x); then, it equals 1 $\frac{dF(x)}{dx}$ = F'(x) = f(x).

Corollary: If X is distributed according to a cdf F(X), W = H^{-1}[F(X)] has cumulative distribution function H(w), provided that H(w) is an appropriate cdf.

Then, testing the validity of a distribution function F(.), for a random iid sample, is equivalent to test if W_k = H^{-1}(Z_k) = H^{-1}[F(X_k)] have distribution function H(.).

For example, testing that X_k come from the cdf F(x) is equivalent to test – using the same sample…- if the transformed sample values W_k = $\Phi^{-1}[F(X_k)]$, where $\Phi(.)$ denotes the cdf of the standard normal, come from the standard normal.

Three uniform (0, 1) random series (of size 100 each) were created.[7]

One was used to provide the exogenous variable X, with normal distribution of mean 7 and 3 standard deviations: X_i = 7 + 3 $\Phi^{-1}(W_i)$, where $\Phi(z)$ denotes the cumulative standard normal and $\Phi^{-1}(z)$ its inverse.

One of the other series was (invariably) used to create the residuals, E, to form the dependent variable, Y, of our baseline (simple) regression model, always created as

(1.1) $Y_i = 5 + 6\ X_i + E_i$

The E_i's are eventually generated by a truncated (at zero) distribution.

Finally, we experimented adding an unrestricted standard normal residual, V_i, built by using the inverse normal of the third uniform random series that would appear as:

---

[7] We use the same series as in Martins (2005).

(1.2) $Y_i = 5 + 6 X_i + E_i + V_i$

This is the formulation usually encountered in stochastic frontier modelling.

Let a be the lower truncation point and b the upper truncation point of a given cdf F(e). We rely on the fact that being W_i, uniform, it should equal the truncated cdf according to the definition:

(1.3) $F[e \mid a < e < b] = \frac{F(e) - F(a)}{F(b) - F(a)}$, $a < e < b$

Obviously, the problem could interchangeably have been defined in terms of the limiting probabilities, i.e., note that $a = F^{-1}(\alpha)$ and $b = F^{-1}(\beta)$ - i.e., for α - the probability left-out to the left of the original F(e) - replaced by $\alpha = F(a)$ and β - with $1 - \beta$ being the upper portion of the original distribution that was cut-off - by $\beta = F(b)$.

(1.4) $F[e \mid F^{-1}(\alpha) < e < F^{-1}(\beta)] = \frac{F(e) - \alpha}{\beta - \alpha}$, $F^{-1}(\alpha) < e < F^{-1}(\beta)$

The pdf obeys

(1.5) $f[e \mid a < e < b] = \frac{\frac{f(e)}{F(b) - F(a)}}{} = \frac{f(e)}{\beta - \alpha}$, $F^{-1}(\alpha) = a < e < b = F^{-1}(\beta)$

Say, an efficiency production frontier model would specify the addition to its deterministic – efficient - counterpart of an error e such that $F[e \mid -\infty < e < 0]$; a minimum cost boundary model, $F[e \mid 0 < e < \infty]$.

We start from our uniform (0,1) series, W_i. We postulate that:

(1.6) $W_i = \frac{F(E_i) - F(a)}{F(b) - F(a)} = \frac{F(E_i) - \alpha}{\beta - \alpha}$

Therefore, after choosing F(.) and the truncation limits, we can generate a series of a truncated distribution considering:

(1.7) $E_i = F^{-1}\{W_i\,[F(b) - F(a)] + F(a)\} = F^{-1}[W_i\,(\beta - \alpha) + \alpha]$

And of course $a < E_i < b$. We considered three cases, of general form $E_i = F^{-1}\{F(a) + [F(b) - F(a)]W_i\}$.

For a standard distribution at $b = \infty$, b is replaced by ∞ and a by $-\infty$: $E_i = F^{-1}(W_i)$.

For a truncated above distribution at $b = 0$, b is replaced by 0 and a by $-\infty$: $E_i = F^{-1}[W_i\ F(0)]$.

For a truncated below distribution at $b = \infty$, a is replaced by 0 and a by 0: $E_i = F^{-1}\{W_i\ [1 - F(0)] + F(0)\}$.

We depart from several hypothesis concerning the error term distribution to be added to the deterministic part of the model: standard normal, uniform, exponential, Cauchy and logistic; the error terms other than the normal were further transformed so as to generate for the standard series a null expected value and unitary variance one; for the Cauchy the series were multiplied 0.67449037, the inverse standard deviation of the zero-mean normal with the same quartiles. Samples had always size $n = 100$. In the philosophy of the estimated regressions, these distributions are then a function of the regression residual divided by a standard error, also subject to estimation.[8]

We report below, in I. – V., for each case the generic distribution used, the procedures taken to generate the truncated series, and tables with a summary of the descriptive statistics of the created input error series, E_i, of the output series, Y_i, and the OLS results of the regression of Y_i on X_i. $\hat{a}_1$ and $\hat{a}_2$ denote, respectively, the intercept and slope estimates of the linear model, $\hat{E}_i$, the estimated OLS residuals; SD refers standard deviation; as usual, a bar indicates a mean, except for the adjusted R^2, $\overline{R}^2$. The first table in Martins (2010a) also contains information on X_i, and the "mother" uniform (0, 1) random series, W_i

[8] In our case, 1 is always the expected value of the estimates of this parameter. But we could have multiply our theoretical error terms by some other factor before addition to the deterministic model.

– the cumulative distribution function - from which all error series were created.

Two types of series were created in Martins (2010a). Tables 1.1.1 to 1.1.5 correspond to the simple form (1.1); Tables 1.2.1 to 1.2.5 to form (1.2) – the previous Y_i series are added of an extra standard normal error, V_i.

I. Standard Normal, $F(e) = \Phi(e)$, $-\infty < e < \infty$. $F(0) = \frac{1}{2}$.

The E_i's come from the normal – $\Phi(z)$ denotes the standard normal. Then:

1. $E_i = \mu + \sigma \; \Phi^{-1}(W_i)$, $-\infty < E_i < \infty$ (i.e., $F(a) = \alpha = 0$, $F(b) = \beta = 1$).

2. a truncated above normal at $\frac{1}{2}$ ($F(a) = \alpha = 0$, $F(a) = \beta = \frac{1}{2}$). Then $E_i = \mu + \sigma \; \Phi^{-1}(\frac{1}{2} W_i)$ and $-\infty < E_i < \mu + \sigma \; \Phi^{-1}(\frac{1}{2})$.

3. a truncated from below normal at 50% ($F(a) = \alpha = \frac{1}{2}$, $F(a) = \beta = 1$).[9] Then $E_i = \mu + \sigma \; \Phi^{-1}(\frac{1}{2} + \frac{1}{2} W_i)$ and $\mu + \sigma \; \Phi^{-1}(\frac{1}{2}) < E_i < \infty$.

μ and σ were always fixed to 0 and 1 respectively (we rely on the standard normal). Other values for σ could have been used instead: for our purposes, it is irrelevant - it is a parameter that will be subject to estimation inquiry.

From our exercise, for example, we conclude that for standard normals, the results reported in Martins (2010a) in Table 1.1.1 for $\overline{E}$ and SD_E are the approximations of the mean and standard deviation of the truncated distributions.[10] The OLS estimate of the intercept, a_1, varies

[9] Notice that even if $\mu = 0$, we generate a different sample than the symmetric of the first median one.

[10] See, for example, Johnston, Kotz and Balakrishnan (1994), p. 159, Table 13.10 for comparable tabulations.

across the three columns – i.e., samples – capturing the bias induced by truncation – the truncated errors do not have a zero mean any longer. Also, a reduced estimated standard error is captured in the two truncated samples. This pattern is reproduced in all the 5 cases.

With the doubled error structure, we confirm – see column 1 of Table 1.2.1 in Martins (2010a) – an increase in the standard error: the variance of the error doubled and the standard error becomes 1.4.

II. Uniform of mean 0 and 1 standard deviation. $F(e) = \frac{z}{2\sqrt{3}} + \frac{1}{2}$, $-\sqrt{3} < e < \sqrt{3}$. $F(0) = \frac{1}{2}$.

1. $E_i = (W_i - \frac{1}{2})\, 2\sqrt{3}$, $-\sqrt{3} < E_i < \sqrt{3}$ (i.e., $F(a) = \alpha = 0$, $F(b) = \beta = 1$).

2. a truncated above error: $E_i = (W_i \frac{1}{2} - \frac{1}{2})\, 2\sqrt{3}$, $-\sqrt{3} < E_i < 0$. ($(F(a) = \alpha = 0$, $F(a) = \beta = 0.5$.)

3. a truncated from below uniform at 50% ($F(a) = \alpha = 0.50$, $F(a) = \beta = 1$).[11] Then $E_i = (W_i \frac{1}{2} + \frac{1}{2} - \frac{1}{2})\, 2\sqrt{3}$ and $0 < E_i < \sqrt{3}$.

III. Exponential of mean 0 and 1 standard deviation. $F(e) = 1 - \exp[-(e + 1)]$, $-1 < e < \infty$. $F(0) = 1 - \exp(-1)$.

1. $E_i = -\ln(1 - W_i) - 1$, $-1 < E_i < \infty$ (i.e., $F(a) = \alpha = 0$, $F(b) = \beta = 1$).

2. a truncated above distribution $E_i = -\ln\{1 - W_i\,[1 - \exp(-1)]\} - 1$, $-1 < E_i < 0$. ($F(a) = \alpha = 0$, $F(a) = \beta = 1 - \exp(-1)$.)

3. a truncated from below at 50% ($F(a) = \alpha = 1 - \exp(-1)$, $F(a) = \beta = 1$).[12] Then $E_i = -\ln\{1 - W_i\,[\exp(-1) + 1 - \exp(-1)]\} - 1$ and $0 < E_i < \infty$.

[11] Notice that even if $\mu = 0$, we generate a different sample than the symmetric of the first quartile one.

IV. Cauchy. $F(e) = \frac{1}{2} + \frac{1}{\pi} \arctan(\frac{e}{0.67449037})$, $-\infty < e < \infty$. $F(0) = \frac{1}{2}$

1. $E_i = 0.67449037 \tan[\pi (W_i - \frac{1}{2})]$, $-\infty < E_i < \infty$ (i.e., $F(a) = \alpha = 0$, $F(b) = \beta = 1$).

2. truncated above ($F(a) = \alpha = 0$, $F(a) = \beta = 0.5$) $E_i = 0.67449037 \tan[\pi (W_i \frac{1}{2} - \frac{1}{2})]$ $-\infty < E_i < 0$.

3. a truncated from below at 50% ($F(a) = \alpha = 0.50$, $F(a) = \beta = 1$).[13] Then $E_i = 0.67449037 \tan\{\pi (W_i \frac{1}{2} + \frac{1}{2} - \frac{1}{2})\}$, and $0 < E_i < \infty$.

V. Logistic $F(e) = \frac{1}{1+e^{-\frac{\pi e}{\sqrt{3}}}}$, $-\infty < e < \infty$. $F(0) = \frac{1}{2}$.

1. $E_i = -\ln(\frac{1}{W_i} - 1) \frac{\sqrt{3}}{\pi}$, $-\infty < E_i < \infty$ (i.e., $F(a) = \alpha = 0$, $F(b) = \beta = 1$).

2. $E_i = -\ln(\frac{1}{W_i \frac{1}{2}} - 1) \frac{\sqrt{3}}{\pi}$, $-\infty < E_i < 0$. ($F(a) = \alpha = 0$, $F(a) = \beta = 0.5$)

3. $E_i = -\ln(\frac{1}{W_i \frac{1}{2} + \frac{1}{2}} - 1) \frac{\sqrt{3}}{\pi}$, and $0 < E_i < \infty$. ($F(a) = \alpha = 0.5$, $F(a) = \beta = 1$)

Our main objective of the following sections will be to obtain estimates of the linear model:

(1.8) $Y_i = a_1 + a_2 X_i + E_i$, i = 1,2,..., 100

---

[12] As above, even if $\mu = 0$, we generate a different sample than the symmetric of the first quartile one.

[13] Notice that even if $\mu = 0$, we generate a different sample than the symmetric of the first quartile one.

or

(1.9) $Y_i = a_1 + a_2 X_i + E_i + V_i$, i = 1,2,..., 100

under the assumption that the E_i's have a truncated (-∞, 0) or (0, ∞) distribution or unrestricted for each of the fifteen series Y_i that were created.

Notice also that (1.8) and (1.9) are equivalent to:

(1.10) $Y_i = a_1' + a_2 X_i + RESI_i$, i = 1,2,..., 100

where $E[RESI_i] = 0$, $Var(RESI_i) = Var(E_i)$, or $VAR(E_i) + Var(V_i)$, with a_1' capturing bias of the estimation.

For each of the (2 times) fifteen Y series, OLS residuals were thus generated, $\hat{E}_i$. The ascending rank of each residual series was recorded as a variable R_i – the ranking of the order statistics of the estimated residual series. In general, and relying on well-known results, whatever the distribution, F(.) of the E_i's, we expect that $F(O_j) = \frac{j}{n+1} = \frac{j}{101}$ where O_j denotes the j-th order statistic. We create in accordance $S_i = \frac{R_i}{n+1} = \frac{R_i}{101}$. Then, we hope to approach $F(E_i)$ in inter-quantile or truncated estimation

We used TSP 4.4 and EXCEL for computation.[14] From the former, we relied more heavily on OLSQ LSQ (FRML, EQSUB) and matrix routines.

2 ORDER ESTIMATION

In section 1, we established the required principles to generate estimation strategies after a first-step OLS run:

For a truncated distribution at known truncation points, inter-quantile inference would rely on the adjustment by NLS for example of:

[14] Hall and Cummins (1997) and (1998).

$$(2.1)\ \ S_i \approx \frac{F(\frac{Y_i - a_1 - a_2X_i}{\sigma}) - F(a)}{F(b) - F(a)}$$

where $S_i = \frac{R_i}{n+1} = \frac{R_i}{101}$ denotes the rank of estimated residual i over the sample size plus 1. a and b are fixed to reproduce the three cases: (-∞, ∞), (-∞, 0) or (0, ∞). Form (1.8) rather than (1.9) would appear to suggest (2.1) – with an error added to the right hand-side at the rankings approximation (2.1).

The null hypothesis distributions F(.) considered were the standard normal, the uniform ($-\sqrt{3}$, $\sqrt{3}$), and the transformed exponential (-1, ∞). Given the way the series were built, we would hope to recover an estimate of 1 for σ in all cases of simple error structure for thr true distribution and 5 and 6 for the linear parameters, intercept and slope, if we stage the appropriately – to the sample – truncated cdf.

With respect to (2.1), "generalised" nonlinear least squares minimising $e'W^{-1}e$, where e denote the difference of S_i minus the inferred $\frac{F(\frac{Y_i - a_1 - a_2X_i}{\sigma}) - F(a)}{F(b) - F(a)}$, for a variance-covariance matrix W – inferring the variance-covariance matrix of the vector representing the right hand-side of (2.1) - given by

$$(2.2)\ \ W = [\ \frac{Min(S_i, S_j)[1 - Max(S_i, S_j)]}{n+2}\ ]$$

off it would also be a possibility – that would simply extend the GMM – generalised method of moments - estimation principle[15] to a GMOS – general method of order statistics.[16]

We present below only the minimum distance method of order-statistics (MDMOS) estimators of equation (2.1). Let F(θ, X) denote the vector of cdf functions for each observation j, F(θ, X_j), and S of the

[15] See Greene (2003), ps. 543-544, for example.

[16] See Martins (2005).

ranks over n + 1 deducted from the first step OLS regression. Then, MDMOS estimators are obtained from

(2.3) $\underset{\hat{\theta}_{MDMOS}}{Min}$ [S - F(θ, X)]' [S - F(θ, X)]

which just requires applying (nonlinear) least squares to S = F(θ, X) + u, where u denotes a vector of residuals.

Standard errors were obtained according to MM principles:[17] letting G(θ) = $\frac{\partial F(\theta, X)}{\partial \theta}$, containing in the j-th row the derivative of F(θ, X_j) with respect to each of the parameters, $G_j(\theta) = \frac{\partial F(\theta, X_j)}{\partial \theta}$:

(2.4) $C\hat{o}v(\hat{\theta}_{MDMOS})$ = [G($\hat{\theta}$)' G($\hat{\theta}$)]$^{-1}$ G($\hat{\theta}$)' W G($\hat{\theta}$) [G($\hat{\theta}$)' G($\hat{\theta}$)]$^{-1}$

In general, for a truncated version of a "standard" (or fixed, specific) cdf F(.), i.e., $F(\theta, X_j) = \frac{F(\frac{Y_j - a_1 - a_2 X_j}{\sigma}) - \alpha}{\beta - \alpha}$, $G_j(a_1, a_2, \sigma)$ can be computed from.

(2.5) $G_j(a_1, a_2, \sigma) = [- \frac{f(\frac{Y_j - a_1 - a_2 X_j}{\sigma})}{\sigma(\beta - \alpha)} ; - \frac{f(\frac{Y_j - a_1 - a_2 X_j}{\sigma}) X_j}{\sigma(\beta - \alpha)} ;$

$- \frac{f(\frac{Y_j - a_1 - a_2 X_j}{\sigma}) \frac{Y_j - a_1 - a_2 X_j}{\sigma^2}}{(\beta - \alpha)}]$

f(.) denotes the density function associated to the (untruncated) cdf F(.).

We expect that for any consistent estimator of θ, $\hat{\theta}$, the identifying restriction test that relies on

(2.6) [S - F(θ, X)]' W^{-1} [S - F(θ, X)]

exhibits under the correct cdf an asymptotic distribution:

---

[17] See Greene (2003), for example, p. 544. Hansen (1982) establishes large sample properties of GMM estimators.

(2.7) $[S - F(\hat{\theta}, X)]' W^{-1} [S - F(\hat{\theta}, X)] \sim X$[18]

where n is the number of observations and k is the number of estimated parameters (at – upper tail - 5%, X = 120.9896; therefore, the order restrictions are not rejected at that significance level if the test statistic exhibits a lower value than the theoretical one).

Minimum distance estimators for fixed α and β performed very well. Results are presented in Martins (2010a) in Tables 2.1.1 to 2.1.5. Each Table contains the results of the simulation for a particular set of the endogenous variables, assuming three particular – normal, uniform and exponential - error term distribution as the null hypothesis – each column has information for a particular sample and truncated hypothesis. We compute, by nonlinear least squares, the parameter estimates, use formulas (2.4) and (2.5) to generate the appropriate standard error, and (2.7) the identifying restricted test statistic. e'e are the reported some of square errors reported by the routine.

We note that for the standard case ($\alpha = 0$, $\beta = 1$, first row block for first column series), the s.e. are smaller than those obtained for OLS, reported in Martins (2010a) in Tables 1.1.1-1.1.3, which points to the quality of the MDMOS method – however, the latter requires knowledge of the adequate cdf, being fully parametric in spirit.

From the restriction test, we conclude that the true normals are identified by the results of Table 2.1.1 (Martins 2010a). That uniform rather than the other alternatives applies to the error structure is patent in Table 2.1.2; but the method does not allow us to identify which truncated case. The exponential exhibits high test statistics but the true truncation appears with the best statistic value – Table 2.1.3 (Martins 2010a). The logistic approaches a significance similar to the normal – Table 2.1.5. Cauchy disturbances lead to too large values of the restriction test statistic.

The next Tables in Martins (2010a) apply to the double error structures. The complexification of the error structure led to the disapearance of the truncated effects: Under all series, the smaller test statistic always points to an untruncated distribution. Notice that the variance of the extra noise is always one – where the mother untruncated distribution has also the same variance. It is possible that

[18] Which should also be useful as a rank test of a particular cdf form…

with a larger value for the variance of the latter would result in recogizable truncated effects.

3 INVERSE ORDER – INTER-QUANTILE – REGRESSION

In section 2, we established the required ordered principles to generate estimation strategies after a first-step OLS run. For the truncated normal, the order regression approach relies on the fact that:

$$(3.1)\quad \frac{Rank(\frac{Y_i - a_1 - a_2 X_i}{\sigma})}{n+1} = \frac{Rank(Y_i - a_2 X_i)}{n+1} \approx \frac{F(\frac{Y_i - a_1 - a_2 X_i}{\sigma}) - \alpha}{\beta - \alpha}$$

Observation rankings, S_i, were inferred from the OLS errors. We can invert the approximation to obtain:

$$(3.2)\quad Y_i \approx a_1 + a_2 X_i + \sigma \ F^{-1}[\alpha + S_i(\beta - \alpha)]$$

An inverse order approach would use those same rankings, to build, for given, fixed, α and β:

$$(3.3)\quad \hat{E}_i \approx F^{-1}[\alpha + S_i(\beta - \alpha)] = F^{-1}\{F(a) + S_i[F(b) - F(a)]\}$$

and regress by OLS, Y_i on X_i and the "theoretical" error, $\hat{E}_i$:

$$(3.4)\quad Y_i \approx a_1 + a_2 X_i + \sigma \ \hat{E}_i + v_i$$

The quality of the fit would guide us to the true distribution. The variance of the estimates have to be scaled relative to the OLS formula – the standard errors of the parameter estimates multiplied by the square root of *(*the variance of $\hat{E}_i$ times 99 times the OLS regression coeficient estimate squared, plus the sum of squares of the regression, and then the sum divided by 97*)*, rather than multiplied by the standard error of the OLS regression as directly reported by the software.

Also, emcompassing tests could easily be constructed relying on the inclusion of more than one transformation – say, use the inverse normal and the inverse exponential - in the right hand-side of the regression (3.4) performed.

At this stage we caution the reader to the fact that the expected value of Y_i has a bias relative to $a_1 + a_2 X_i$ that may not be well aproximated through the inverse structure (3.4) in the term $\sigma \hat{E}_i$. Yet assyntotically, (3.4) should result in an adequate framing.

Results depicted in Martins (2010a) in Tables 3.1.1 to 3.1.5 exhibit similar patterns to the previous section: the truncated normal is well approximated, exhibiting the smaller sum of squares residuals e'e – that includes the variance accounted by the last term of the regression) - as v'v. Uniform disturbances are identified as such but not the truncated case. Exponential case is correctly identified. Cauchy and Logistic appear as normal cases.

In the next Table in Martins (2010a), we repeat the procedure for normal disturbances and hypothesis with the double error structure. Again the "best" results fall on the untruncated cases for each of the series.

4 MOMENT REPLICATION

The deterministic part of a linear model is added of an error which is range restricted, ε_i, and a conventional one, v_{0i}. We can unfold k general equation blocks to the model by multiplying for each block each equation by the observations of one of the k + 1 (including the constant term) explanatory variables:

(4.1) $$Y_i = \beta_0 + X_{1i}\beta_1 + X_{2i}\beta_2 + \ldots + X_{ki}\beta_k + \varepsilon_i + v_{0i}$$

$$X_{1i}Y_i = X_{1i}\beta_0 + X_{1i}^2\beta_1 + X_{1i}X_{2i}\beta_2 + \ldots + X_{1i}X_{ki}\beta_k + X_{2i}\varepsilon_i + v_{1i}$$

$$X_{2i}Y_i = X_{2i}\beta_0 + X_{2i}X_{1i}\beta_1 + X_{2i}^2\beta_2 + \ldots + X_{2i}X_{ki}\beta_k + X_{2i}\varepsilon_i + v_{2i}$$

...

$$X_{ki}Y_i = X_{ki}\beta_0 + X_{ki}X_{1i}\beta_1 + X_{ki}X_{2i}\beta_2 + \ldots + X_{ki}^2\beta_k + X_{ki}\varepsilon_i + v_{ki}$$

$$i = 1,2,\ldots,n$$

We want to restrict $\varepsilon_i > 0$, or $\varepsilon_i < 0$, to denote the effective inefficiency. Then we could hope to parametrise our simple linear regression as a nonlinear one with 2n observations:[19]

(4.2) $Y_i = 1\, a_1 + X_i\, a_2 + 1\, \sqrt{\varepsilon_i^2} + v_{0i}$, i = 1,2,...,n

$X_i Y_i = X_i\, a_1 + X_i^2\, a_2 + X_i \sqrt{\varepsilon_i^2} + v_{1i}$, i = 1,2,...,n

$\sqrt{\varepsilon_i^2}$ would be n parameter to be estimated along with a and b. The variable that multiplies each of the $\sqrt{\varepsilon_i^2}$'s has zeros all over with the exception of the ith observation – 1 – and the n+i-th observation – that contains X_i. The previous case contemplates minimum cost structures; for a production or revenue frontier we can use:

(4.3) $Y_i = 1\, a + X_i\, b - 1\, \sqrt{\varepsilon_i^2} + v_{0i}$, i = 1,2,...,n

$X_i Y_i = X_i\, a + X_i^2\, b - X_i \sqrt{\varepsilon_i^2} + v_{1i}$, i = 1,2,...,n

To account for potential heteroscedasticity, the data of the second equation block was divided by the square root of the mean of X_i^2 [20] in a weighted least squares – WLS – version of the double equation model. We then considered a covariance matrix of the two equation system typical error that off of he diagonal has the mean of X divided be the square root of the mean of X_i^2 and applied GLS in accordance to the previous system (WLS) – a GLS refinement of the procedure.

The first comment one can make is that even if $v_{ki} = 0$, least squares procedures cannot solve for the k + 1 parameters and the ε_i's – nor the $\sqrt{\varepsilon_i^2}$ - directly as parameters of an equation system – we attain singularity. Therefore we present estimates of the model

---

[19] Method accuracy should increase with k – and the number of observations.

[20] Alternativelly, restricted SUR could be preformed on the system of two equations. The covariance structure would then allow for the mean of X to factor the covariance between the same observations of the two blocks…

(4.4) $Y_i = 1\ a + X_i\ b + X_i\ b + \sigma\ \hat{E}_i + v_{0i}$, i = 1,2,...,n

$X_iY_i = X_i\ a + X_i^2\ b + \sigma\ \hat{E}_i\ X_i + v_{1i}$, i = 1,2,...,n

where

(4.5) $\hat{E}_i \approx F^{-1}[\alpha + S_i(\beta - \alpha)]$

and S_i the frequency estimated from the rankings of the first step OLS regression. I.e., we apply the method of replicated moments to the first equation of form (4.4).

We also present the results for the standard linear regression model – (1.10). We experimented adding the 98 quasi-dummies – we discarded the two (true) median observations – to the simple linear regression. We discard additionally the last observation error when $\hat{E}_i$ is included in the regression.

We restricted the results to the normal case. Martins (2010a), Table 4.1 refers to single error structure. The correct truncation is invariably identified.

The second table used the the double-error series. As in previous sections, the correct truncation hypothesis does not exhibit the smaller sum of squares, classifying incorrectly the true truncation case. The added error dummies originated nonsensical parameters estimates, even for a_2, the slope coefficient, that is always well captured by the different procedures.

5 CONCLUSIONS

Estimation procedures of a linear regression model under truncated residual distribution assumptions have been proposed, with illustration for different density families. We focussed on truncation at zero residuals, an assumption usually encountered in stochastic frontier models.

1. Albeit the variety of procedures tested, the methods only give acceptable answer to a single residual environment – or with the majority of residual randomness coming from the truncated residuals.

2. All the methods perform satisfactorilly in identifying whether (single) truncation exists or not and of which type for normal,

exponential and sometimes also the uniform residuals. Inverse order estimation performed as accurately as the direct method in model identification – but standard errors appear smaller for the former if one is willing to accept the standard ones.

3. For the normal, the application of replicated moments to the inverse order method formulation outperformed – exhibits smaller standard errors of the parameter estimates – the (simple) inverse order regression procedure.

REFERENCES

Amemyia, Takeshi. 1985. *Econometric Analysis.* Cambridge: Harvard University Press.

Cohen, A.C. Jr. 1950. "Estimating the Mean and Variance of Normal Populations from Singly and Doubly Truncated Samples." *Annals of Mathematical Statistics.* 21(4): 557-569.

Cohen, A.C. Jr. 1949. "On Estimating the Mean and Standard Deviation of Truncated Normal Distributions." *Journal of the American Statistical Association.* 44: 518-525.

David, H. A. 1970. *Order Statistics.* New York: John Wiley.

Galambos, Janos. 1984. "Order Statistics." In *Handbook of Statistics*, Vol. 4, edited by P.R. Krishnaiah and P.K. Sen: 359-382. Elsevier Science Publishers.

Greene, William H. 1997. "Frontier Production Functions" In *Handbook of Applied Econometrics. Vol. II: Microeconomics.* Edited by M. Hasham Pesaram and Michael R. Wickens. Oxford: Blackwell.

Greene, William H. 2003. *Econometric Analysis.* 5th Edition. New Jersey: Prentice-Hall International.

Hájek, Jaroslav, Zbyněk Šidák and Pranab K. Sen. 1999. *Theory of Rank Tests.* 2nd Edition, San Diego, Academic Press.

Hall, Bronwyn H. and Clint Cummins. 1998. *TSP 4.4 Reference Manual.* TSP International.

Hall, Bronwyn H. and Clint Cummins. 1997. *TSP 4.4 User´s Guide.* TSP International.

Hansen, Lars Peter. 1982. "Large Sample Properties of Generalised Method of Moments Estimators." *Econometrica.* 50(4): 1029-1054.

Hogg, Robert V., Allen T. Craig. 1978. *Econometric Analysis.* 4th Edition. New York: Macmillan Publishing Co.

Johnson, Norman L., Samuel Kotz and N. Balakrishnan. 1994. *Continuous Univariate Distributions - 1.* 2nd Edition. New York: Wiley.

Johnson, Norman L. and Samuel Kotz. 1970a. *Continuous Univariate Distributions - 1.* New York: Wiley.

Johnson, Norman L. and Samuel Kotz. 1970b. *Continuous Univariate Distributions - 2.* New York: Wiley.

Judge, George G., W.E. Griffiths, R. Carter Hill, Helmut Lutkepohl and Tsoung-Chao Lee. 1985. *The Theory and Practice of Econometrics.* 2th Edition. New York: Wiley.

Koop, Gary. 2003. *Bayesian Econometrics.* Chichester: Wiley.

Martins, Ana Paula. 2005. "On Ordered Principles: Order Regression, Interquantile Inference and Truncated Distributions." Mimeo.

Martins, Ana Paula. 2003. "Social Disruption: Explorations on the Pattern of Criminal and Violent Occurrences in Response to Macroeconomic Conditions." Mimeo.

Martins, Ana Paula. 2010a. "Frontier Techniques: Contrasting the Performance of (Single) Truncated Order Regression Methods and Replicated Moments,". EERI Research Paper Series No 2010/36, Economics and Econometrics Research Institute, Brussels.

Martins, Ana Paula. 2010b. "Splitting Games: Nash Equilibrium and the Optimisation Problem," *Journal of Economics and Econometrics*. 53(1): 1-28.

Mittelhammer, Ron C., George G. Judge and Douglas J. Miller. 2000. *Econometric Foundations.* Cambridge: Cambridge University Press.

Mosteller, Louis A. 1946. "On Some Useful "Inefficient" Statistics." *Annals of Mathematical Statistics.* 17(5): 377-408.

Murillo-Zamorano, Luis R. 2004. "Economic Efficiency and Frontier Techniques." *Journal of Economic Surveys.* 18(1): 33-77.

Pestana, Dinis Duarte e Sílvio Filipe Velosa. 2002. *Introdução à Probabilidade e à Estatística. Vol. I.* Lisboa: Fundação Calouste Gulbenkian.