Efficient Semiparametric Estimation in Stochastic Frontier Model

Byeong U. PARK* **and Léopold SIMAR****

February 1992

ABSTRACT

This paper considers the semiparametric stochastic frontier model with panel data which arises in the problem of measuring technical inefficiency in production processes. We assume a parametric form for the frontier function, which is linear in production inputs. The density of the individual firm-specific effects is considered unknown. We construct an efficient estimator of the slope parameters in the frontier function. We also give an estimator of the level of the frontier function and its asymptotic properties are investigated. Furthermore, we provide a predictor of the individual effects which can be directly translated to firm-specific technical inefficiencies. Finally, we illustrate our methods through a real data example.

AMS 1991 Subject Classification: 62 G05, 62 G20

Keywords and phrases: Stochastic frontier model, semiparametric model, efficient estimation, information bound, panel data.

Acknowledgments. The authors would like to thank Berwin Turlach for his contribution in the computational aspects of the example.

This work was financed, in part, by contract No TR/B3/007, Programme d'impulsion "Transport et mobilité" of the Belgian Government.

---

* CORE and Institute of Statistics, Université Catholique de Louvain.
Research of the first author done while on leave from Seoul National University and partially supported by KOSEF Postdoctoral Fellowship 91.

** CORE and Institute of Statistics, Université Catholique de Louvain.

ISSN-0771 3894

1. Introduction

The model we analyse in this paper assumes N independent observations (X_i, Y_i), which are written as

$$Y_{it} = X_{it}'\beta + \alpha_i + \epsilon_{it}, \quad t = 1, \ldots, T \tag{1.1}$$

where $X_i = (X_{i1}', \ldots, X_{iT}')'$ and $Y_i = (Y_{i1}, \ldots, Y_{iT})'$. The X_i are i.i.d. (dxT)-dimensional random vectors with unknown density function g, and β is a d-dimen sional unknown vector. The ϵ_{it} are considered i.i.d. random variables from $N(0, \sigma^2)$ with unknown σ^2, and the α_i are i.i.d. from a unknown density h whose support is bounded above(below). It is assumed that the ϵ_{it}, α_i and X_{it} are independent.

This work is motivated from the problem of measuring production inefficiency which has been a subject of large econometric literature. The original work on this problem dates back to Farrell(1957). From a statistical point of view, one of the most satisfactory model to analyse the problem is the *stochastic frontier model* introduced by Aigner, Lovell and Schmidt(1977), and Meeusen and van den Broeck(1977), which is written as follows:

$$\begin{aligned} Y_i &= f(X_i, \beta) + u_i \\ u_i &= \alpha_i + \epsilon_i, \quad \alpha_i \leq 0, \quad i = 1, \ldots, N \end{aligned} \tag{1.2}$$

where Y_i is an observation on the output of the i-th firm, X_i is vector of observations on the d inputs, β is vector of unknown production parameters, ϵ_i is a random symmetric noise(typically normal) and α_i is an unobservable random variable which represents firm specific technical inefficiency. The density of α_i is considered to have the support $(-\infty, 0)$. This means that, neglecting the noise, f(x,β) is the maximal attainable output with input x, called *frontier function.*

In the analysis of the model (1.2), a parametric functional form for f, which is linear in β, is usually assumed. This specific form of f is preferred since it allows richer economic interpretations of production process under study and several important functions arising in production frontiers(such as *translog* or *Cobb-Douglas*) can be reduced to this by suitable transformations. The main interest, then, may be to estimate the parameter β and to identify the true realization of α_i for a specific firm(we will call the latter "prediction" below). The difficulty, however, lies in the fact that α_i is not identifiable from the knowledge of (Y_i, X_i, β) since it is confounded with ϵ_i. Jondrow, Lovell, Materov and Schmidt(1982) considered

$E(\alpha_i|u_i)$ as an alternative to α_i, giving a predictor of this quantity but without any statistical properties.

This identifiability problem is overcomed when one has several observations for each firm(*panel data*). The model (1.1) is an obvious extension of (1.2) in the case that panel data are available. For the model (1.1), the upper boundary of the support of the density h, denoted by $B(h)$, will be the level of the frontier function. The frontier function is then $B(h) + x'\beta$ and $\alpha_i - B(h)$ corresponds to the technical inefficiency of the i-th firm.

In this paper, we derive an asymptotic lower bound for estimating β in the presence of the nuisance parameters σ^2 and h. Then, an efficient estimator of β which attains this lower bound is presented. Furthermore, predictors of individual effects α_i are constructed utilizing the efficient estimator of β, and their asymptotic properties are investigated. Finally, an estimator of the support boundary $B(h)$ is provided together with its rate of convergence to the desired quantity. This boundary estimator yields estimators of the frontier function and of the technical inefficiencies.

Stochastic frontier models with panel data has been analysed by several authors. Schmidt and Sickles(1984), and Cornwell, Schmidt and Sickles(1990) compared several conventional methods such as GLS, *within*(which is OLS with within deviations, $Y_{it} - \bar{Y}_i$ and $X_{it} - \bar{X}_i$) and maximum likelihood method(with a parametric specification for the distributions of α_i). Without giving any statistical properties, Simar(1992) introduced an interesting approach which combines a nonparametric approach called FDH(Free Disposal Hull), proposed by Deprins, Simar and Tulkens(1984), and OLS method. Hausman and Taylor(1981) considered a panel model which allows some of covariates to be time invariant(independent of t) and/or correlated with the individual effect α_i. They proposed an estimator of β that is, in some sense, a hybrid of GLS and the within estimator. However, none of all the previously proposed estimators of β are fully efficient when the density h is unknown. Moreover, to the best of our knowledge, no asymptotic results beyond mere consistency for estimators of the level of the frontier function have been obtained so far.

This paper is organized as follows. Section 2 gives the asymptotic lower bound for estimating β and the efficient estimator of β. Section 3 considers the problems of predicting each individual effect α_i and estimating the support boundary of h. In Section 4, our methods are illustrated through a real data example. All the proofs are included in Section 5.

2. Efficient estimation of slope parameters

Throughout this section except Theorem 2.3, the time period T is considered fixed. The notion of efficient estimation in semiparametric models is well established in Begun, Hall, Huang and Wellner(1983), and Bickel, Klaassen, Ritov and Wellner(1992). Below we briefly outline the basic ideas in our context.

Write (X, Y) for "generic" observations. Let $\mathbf{P}$ be the set of all possible joint distributions of (X, Y). One calls $\mathbf{P}_0$ a regular parametric submodel of $\mathbf{P}$ if $\mathbf{P}_0(\subset\mathbf{P})$ can be represented as $\{\mathrm{P}_{(\beta,\eta)} : \beta \in \mathbf{R}^d,\ \eta \in \mathrm{S} \text{ open} \subset \mathbf{R}^k\}$ and at every (β_0, η_0) the mapping $(\beta, \eta) \to \mathrm{P}_{(\beta,\eta)}$ is continuously Hellinger differentiable(see Ibragimov and Has'minskii 1981, Section 1.7). Suppose $\mathrm{P}(= \mathrm{P}_{(\beta_0,\eta_0)})$ belongs to a regular parametric submodel $\mathbf{P}_0$ of $\mathbf{P}$. Let $L(X, Y, \beta, \eta)$ denote the log likelihood of an observation from $\mathrm{P}_{(\beta,\eta)}$ and let $\ell_\beta(X, Y) = \partial\ell/\partial\beta|_{(\beta_0,\eta_0)}$ and $\ell_{\eta_j}(X, Y) = \partial l/\partial\eta_j|_{(\beta_0,\eta_0)}$ where $\eta = (\eta_1, \ldots, \eta_k)$. Let $[\ell_\eta]$ denote the linear span generated by $\{\ell_{\eta_j}\}_{j=1}^k$. The information bound for β is, then, given by

$$I(\mathrm{P}; \beta, \mathbf{P}_0) = E_P \ell^* \ell^{*'}(X, Y)$$

where

$$\ell^* = \ell_\beta - \Pi(\ell_\beta|[\ell_\eta])$$

and the notation $\Pi(\ell|S)$ denotes the vector of projections of each component of ℓ into the space S in $L_2(\mathrm{P})$. This definition of information bound is extended to the case where we only assume that $\mathrm{P} \in \mathbf{P}$. Consider the class of all regular parametric submodels with k=1 containing P, and denote it by $\mathbf{C}$. Let $[\ell_\eta(\mathbf{P}_0)]$ denote the linear span generated by ℓ_η for a submodel $\mathbf{P}_0$. Then the information bound for β under $\mathbf{P}$ is defined by

$$I(\mathrm{P}; \beta, \mathbf{P}) = E_P \ell^* \ell^{*'}(X, Y)$$

where

$$\ell^* = \ell_\beta - \Pi(\ell_\beta|V)$$

and V is the closed linear span(called *tangent space*) of the union of $[\ell_\eta(\mathbf{P}_0)]$ over the class $\mathbf{C}$. The random variable ℓ^* is called the *efficient score function.* It is well known that $I(\mathrm{P}; \beta, \mathbf{P}_0)$ - $I(\mathrm{P}; \beta, \mathbf{P})$ is positive semidefinite for any regular parametric submodel $\mathbf{P}_0$(e.g. Bickel, Klaassen, Ritov and Wellner 1992). An estimator $\hat{\beta}_N$ is called efficient in $\mathbf{P}$ if

$$L_P(\sqrt{N}(\hat{\beta}_N - \beta)) \to N(0, I^{-1}(\mathrm{P}; \beta, \mathbf{P})).$$

Now we will exhibit the information bound for the model (1.1). To this end, let $Y = (Y_1, \ldots, Y_T)'$ and $X = (X_1', \ldots, X_T')'$ for the generic observations (X, Y). Let $Z_t(\beta) = Y_t - X_t'\beta$ and $U_t(\beta) = Z_t(\beta) - \bar{Z}(\beta)$ where $\bar{Z}(\beta) = T^{-1} \sum_{t=1}^{T} Z_t(\beta)$. Let

$$w(z) = \int \phi_{\bar{\sigma}}(z-u)h(u)du = \phi_{\bar{\sigma}} * h(z)$$

be the density of $\bar{Z}(\beta)$ where $\bar{\sigma} = \sigma/\sqrt{T}$, and let

$$I_0 = \int \frac{(w')^2}{w}(z)dz$$

be the Fisher information for location of w. Define

$$\Sigma_W = E_P T^{-1} \sum_{t=1}^{T} (X_t - \bar{X})(X_t - \bar{X})'$$

$$\Sigma_B = E_P(\bar{X} - E_P(\bar{X}))(\bar{X} - E_P(\bar{X}))'$$

where $\bar{X} = T^{-1} \sum_{t=1}^{T} X_t$. Throughout this paper we will assume that $I_0 < \infty$ and, Σ_W and Σ_B exist and are nonsingular.

Letting $I = I(\mathrm{P}\ ; \beta, \mathbf{P})$ and dropping the argument β in $U_t(\beta)$ and $\bar{Z}(\beta)$, we have the following theorem.

Theorem 2.1. *The efficient score function and the information bound for estimating β for the model (1.1) are given by*

$$\ell^* = \sigma^{-2} \sum_{t=1}^{T} U_t X_t - (\bar{X} - E_P(\bar{X}))\frac{w'}{w}(\bar{Z}),$$

$$I = \bar{\sigma}^{-2}\Sigma_W + I_0\Sigma_B.$$

Remark 2.1. When h and σ^2 are known, the Σ_B in the information bound is replaced by $E_P\bar{X}\bar{X}'$.

Remark 2.2. It can be seen that the information bound I in Theorem 2.1 coincides with $I(\ \mathrm{P}; \beta, \mathbf{P}_0)$ where $\mathbf{P}_0$ is a regular parametric submodel obtained by varying the location of h but with fixed shape. This is called *least favorable* regular parametric submodel of $\mathbf{P}$.

We now construct an efficient estimator of β. The idea is to proceed as in the classical estimation of the location problem: Let $\tilde{\ell} = \ell^{-1}l^*$. This is called the *efficient influence function* having the properties $E_P\tilde{\ell} = 0$ and $E_P\tilde{\ell}\tilde{\ell}' = I^{-1}$.

(a) Find a good estimator $\tilde{\beta}_N$ of β.

(b) Consider $\tilde{\ell}$ as $\tilde{\ell}(x, y, \beta, \sigma^2, h)$. Construct a suitable estimator $\hat{\tilde{\ell}}(\cdot, \cdot, \beta; X_1, Y_1, \ldots, X_N, Y_N)$ of $\tilde{\ell}(\cdot, \cdot, \beta, \sigma^2, h)$.

(c) Form

$$\hat{\beta}_N = \tilde{\beta}_N + N^{-1} \sum_{i=1}^{N} \hat{\tilde{\ell}}(X_i, Y_i, \tilde{\beta}_N; X_1, Y_1, \ldots, X_N, Y_N)$$

as the efficient estimator.

For a preliminary estimator $\tilde{\beta}_N$ of β, we use the within estimator obtained from regressing $Y_{it} - \bar{Y}_i$ on $X_{it} - \bar{X}_i$ by the ordinary least squares method. In fact,

$$\tilde{\beta}_N = (NT\hat{\Sigma}_W)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \bar{X}_i)(Y_{it} - \bar{Y}_i)$$

where $\hat{\Sigma}_W$ is an estimator of Σ_W defined by

$$\hat{\Sigma}_W = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \bar{X}_i)(X_{it} - \bar{X}_i)'.$$

It is straightforward to show that, as $N \to \infty$(with fixed T), we have

$$L_P(\sqrt{N}(\tilde{\beta}_N - \beta)) \to N(0, \bar{\sigma}^2 \Sigma_W^{-1}). \tag{2.1}$$

In view of Theorem 2.1, the within estimator is not efficient. However, it already lies in a $N^{1/2}$-neighborhood of β, which is enough for the final estimator given below to be efficient.

To construct the efficient estimator, define an estimator of w by the kernel estimator,

$$\hat{w}(t, \beta) = N^{-1} \sum_{i=1}^{N} K_{s_N}(t - \bar{Z}_i(\beta)) + c_N$$

where $K_s(t) = K(t/s)/s$, $K(t) = e^{-t}(1 + e^{-t})^{-2}$, and s_N, c_N tend to zero at some proper rates as specified below. Define

$$\hat{\sigma}^2(\beta) = \sum_{i=1}^{N} \sum_{t=1}^{T} U_{it}^2(\beta)/N(T-1),$$

$$\hat{\Sigma}_B = \sum_{i=1}^{N} (\bar{X}_i - \bar{\bar{X}})(\bar{X}_i - \bar{\bar{X}})'/N$$

where $\bar{\bar{X}} = \sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}/(NT)$. The efficient estimator is now defined by

$$\hat{\beta}_N = \tilde{\beta}_N + N^{-1}\hat{I}^{-1} \sum_{i=1}^{N} [\sum_{t=1}^{T} \tilde{U}_{it} X_{it}/\tilde{\sigma}^2 - (\bar{X}_i - \bar{\bar{X}}) \frac{\hat{w}'}{\hat{w}} (\tilde{Z}_i, \tilde{\beta}_N)] \tag{2.2}$$

where

$$\hat{I} = T\hat{\Sigma}_W/\tilde{\sigma}^2 + \hat{I}_0 \hat{\Sigma}_B,$$

$$\hat{I}_0 = N^{-1} \sum_{i=1}^{N} \Big(\frac{\hat{w}'}{\hat{w}}\Big)^2 (\tilde{Z}_i, \tilde{\beta}_N),$$

and $\tilde{\sigma}^2$, $\tilde{U}_{it}$, $\tilde{Z}_i$ are used for $\hat{\sigma}^2(\tilde{\beta}_N)$, $U_{it}(\tilde{\beta}_N)$, $\bar{Z}_i(\tilde{\beta}_N)$.

Let $\{c_N\}$, $\{s_N\}$ be such that

$$c_N \to 0, \;\; s_N \to 0, \;\; N c_N^2 s_N^6 \to \infty$$

as $N \to \infty$.

Theorem 2.2. *Assume $\int e^{t|\bar{x}|} g(x)dx < \infty$ for some $t > 0$, and $\int |u|^2 h(u)du < \infty$. Then,*

$$L_P(\sqrt{N}(\hat{\beta}_N - \beta)) \to N(0, I^{-1}).$$

Remark 2.3. The assumption on g in the above thorem is far from being optimal for the asymptotic efficiency of the estimator. In fact, existence of moments $E_P|\bar{X}|^r$ up to some order would be sufficient with some changes for the rates of c_N and s_N. But we choose to use this mostly for the simplicity of the presentation.

It would be interesting to investigate the asymptotic properties of $\hat{\beta}_N$ when T also goes to infinity. For this, we need to assume that X_{it} are i.i.d. d-dimensional random vectors from unknown density function g_1. Append the subscript T(in addition to N) in all the corresponding sequences of random vectors and numbers. The following theorem will be useful in the next section.

Theorem 2.3. *Let X denote the generic observations X_{it}. Suppose $E_P|X|^2 < \infty$ and $E_P(X - E_P(X))(X - E_P(X))'$ is nonsingular. Then,*

(i) when N fixed, $T \to \infty$ and $Ts_{N,T}^2 \to \infty$,

$$\sqrt{T}(\hat{\beta}_{N,T} - \beta) = O_p(1)$$

(ii) when $N, T \to \infty$, $Ts_{N,T}^2 \to \infty$ and $NT^{-1}s_{N,T}^{-2} = O(1)$,

$$\sqrt{NT}(\hat{\beta}_{N,T} - \beta) = O_p(1).$$

3. Individual effects and level of frontier function

In this section, we append the subscript T in all the related quantities to stress their dependence on T. Given the efficient estimator $\hat{\beta}_{N,T}$ at hand, it would be natural to predict α_i by

$$\hat{\alpha}_i = \bar{Z}_i(\hat{\beta}_{N,T}). \tag{3.1}$$

Since there will be no asymptotics for this predictor if T is fixed, we will let T go to infinity to investigate its asymptotic properties. As in Theorem 2.3, we need to assume that X_{it} are i.i.d. d-dimensional vectors from unknown density g_1. In the following theorem, N may be fixed or tend to infinity.

Theorem 3.1. *Under the assumptions of Theorem 2.3, as $T \to \infty$ and $Ts_{N,T}^2 \to \infty$,*

$$L_P(\sqrt{T}(\hat{\alpha}_i - \alpha_i)) \to N(0, \sigma^2).$$

Remark 3.1. When one is interested in predicting the differences $\alpha_i - \alpha_j$(the relative technical inefficiencies of the i-th firm with respect to the j-th firm), one can predict this by $\hat{\alpha}_i - \hat{\alpha}_j$. This predictor has the asymptotic $N(0, 2\sigma^2)$ distribution when normalized by $\sqrt{T}$ under the same assumptions of the above theorem.

As mentioned in the introduction, the frontier function is $B(h) + x'\beta$. Since we already have an efficient estimator of β, it remains to estimate $B(h)$ to get an estimator of this function. A natural estimator of this quantity would be

$$\hat{B}(h) = \max_{1 \le i \le N} \bar{Z}_i(\hat{\beta}_{N,T}).$$

The following two theorems will be useful to investigate the asymptotic property of this estimator. For this let $\alpha_{(N)} = \max_{1 \le i \le N} \alpha_i$.

Theorem 3.2. *Under the assumptions of Theorem 2.3 and if $\int e^{t|x|} g_1(x)dx < \infty$ for some $t > 0$, as $T \to \infty$ and $Ts_{N,T}^2 \to \infty$,*

(i) when N is fixed,

$$\sqrt{T}(\hat{B}(h) - \alpha_{(N)}) = O_p(1),$$

(ii) when N goes to infinity,

$$\sqrt{T}(\hat{B}(h) - \alpha_{(N)}) = O_p(logN).$$

How close $\alpha_{(N)}$ to $B(h)$ is illustrated by the following theorem. For this, suppose the density h satisfies

$$h(B(h) - x) \geq Cx^{\gamma} \tag{3.2}$$

as $x \downarrow 0$ for some constant $C > 0$ and $\gamma \geq 0$. This condition requires that the density h should have certain amount of mass near the boundary point $B(h)$.

Theorem 3.3. *Under the assumptions of Theorem 2.3 and the condition (3.2), as* $N \to \infty$,

$$\alpha_{(N)} - B(h) = O_p(N^{-\frac{1}{\gamma+1}}).$$

The implication of Theorem 3.2 and 3.3 is clear. For example, when $\gamma = 0$(the case where the density at the boundary stays away from zero, such as shifted half-normal and exponential),

$$\hat{B}(h) - B(h) = O_p(T^{-1/2}logN + N^{-1})$$

if both N and T go to infinity.

4. A real data example

We consider the analysis of labor efficiencies of 17 railway companies observed over a period of 14 years(T=14, N=17). Data on the activity of the main international railway companies can be found in the annual reports of the *Union Internationale des Chemins de Fer* (U.I.C.). The railway companies retained for the illustration are given in Table 1.

Table 1: *Rail Way Companies for Illustration*

Network	Country	Network	Country
BR	Great Britain	NS	Netherlands
CFF	Switzerland	NSB	Norway
CFL	Luxemburg	OBB	Austria
CH	Greece	RENFE	Spain
CP	Portugal	SJ	Sweden
DB	Germany	SNCB	Belgium
DSB	Denmark	SNCF	France
FS	Italy	VR	Finland
JNR	Japan		

The data are available for each network on annual base. In this study we used the period from 1970 to 1983. This provides 238 observations on the whole set of variables. This data set has been analysed by Hall, Härdle and Simar(1991) where the performance of the bootstrap in frontier models was analysed. The variables we use in this example are:

$$
\begin{aligned}
Y_{it} = &\ \text{labor(total number of employees) / total length of the network(in kms);} \\
X_{1it} = &\ \text{total distance covered by trains(in}10^3 \text{ kms);} \\
X_{2it} = &\ \text{ ratio for passenger trains in } X_{1it} \text{ (in \%);} \\
X_{3it} = &\ \text{density of the network(kms of lines by } 100km^2\text{);} \\
X_{4it} = &\ \text{mean distance covered by a passenger(in kms);} \\
X_{5it} = &\ \text{mean distance covered by a ton of freight tranported(in kms) ;} \\
X_{it} = &\ (X_{1it}, \ldots, X_{5it})'.
\end{aligned}
$$

Note that all the data are in logarithms and have been adjusted for time trend effect.

Since we analyse a labor function, the support of h is bounded from below(the most technically efficient railway company uses the least labor), so that the estimator of the lower boundary $B(h)$ will be given by

$$\hat{B}(h) = \min_{1\le i\le N} \bar{Z}_i(\hat{\beta}_{N,T}).$$

Furthermore, in order to compute the efficient estimator of β, we need to choose a bandwidth s. In this paper, we propose a bootstrap procedure: For a given bandwidth s, compute $\hat{\beta}_{N,T}(s)$. Then by resampling of size T in each (X_i, Y_i), $i = 1, \ldots, N$, construct M pseudo-samples $\{(X_i^*, Y_i^*)^{(m)}\}_{i=1}^N$, $m = 1, \ldots, M$. Each pseudo sample provides the bootstrap version of $\hat{\beta}_{N,T}(s)$, denoted by $\hat{\beta}_{N,T}^{*(m)}(s)$. Compute the criterion value

$$C(s) = \frac{1}{M} \sum_{m=1}^{M} [\hat{\beta}_{N,T}^{*(m)}(s) - \hat{\beta}_{N,T}(s)]' [\hat{\beta}_{N,T}^{*(m)}(s) - \hat{\beta}_{N,T}(s)].$$

The bootstrap optimal choice is then given by

$$s^* = \arg\min_s C(s).$$

In our example, we chose M=1000 and carried out a grid search to find s^*. The grid used consists of 10 equally spaced points in the interval [0.1, 1.0]. The value s^*=0.308 was calculated by fitting a parabola passing through the minimal and the two adjacent grid points. Figure 1 shows the shape of C(s).

The estimation procedure is then performed with the selected bandwidth s^*, on the original data. The results are illustrated in Table 2 where the point estimates $\hat{\beta}_{N,T}(s^*)$ along with their estimated standard errors (computed from $\hat{I}^{-1}(\hat{\beta}_{N,T})$) are displayed.

Table 2: *Efficient Estimation of β*

Variable	Estmate	Standard error
Total Distance	.117	.069
Ratio Passengers	-.097	.137
Density	-.658	.131
Mean distance passengers	.017	.067
Mean distance freight	-.065	.055

It may be pointed out that the coefficients which are significantly different from zero have the expected signs: the elasticity of labor with respect to the main measure of the output X_1 is positive and the elasticity of the labor with respect to the density(complexity) of network is $(1+\beta)$ which is also positive as expected.(Adding one for the latter is justified in Hall, Härdle and Simar 1991). Note also that the expected negative signs for β_2 and β_5 although not significant at the usual levels. Finally, β_4 has an unexpected sign but its standard error shows that it could be near zero.

Table 3: *Prediction of α's and Inefficiencies*

Network	$\hat{\alpha}$	$\hat{\alpha} - \hat{B}(h)$	Network	$\hat{\alpha}$	$\hat{\alpha} - \hat{B}(h)$
BR	.407	1.983	NS	.462	2.038
CFF	.670	2.247	NSB	-1.576	.000
CFL	1.293	2.869	OBB	.593	2.168
CH	-.957	.618	RENFE	-.889	.687
CP	-.193	1.383	SJ	-1.447	.128
DB	.725	2.301	SNCB	1.119	2.694
DSB	.184	1.760	SNCF	-.151	1.425
FS	.409	1.984	VR	-1.347	.229
JNR	.697	2.273			

The technical inefficiencies of the railway networks are also of interest and may be predicted from the estimated frontier. The estimated value of $B(h)$ is -1.576. Table 3 gives, for each network, the predicted value of α_i and the predicted labor technical inefficiency$(\hat{\alpha}_i - \hat{B}(h))$. It appears that the most efficient network is NSB(Norway). This confirms the findings of Hall, Härdle and Simar(1991). Note the poor performances of CFL(Luxemburg) and of SNCB(Belgium). Using the estimated value $\sqrt{2}\hat{\sigma}/\sqrt{T} = 0.01764$ for the asymptotic standard deviation of $(\hat{\alpha}_i - \hat{\alpha}_j)$, one can compare efficiencies between networks. For instance, the relative ranking of BR and FS is certainly not significant, and NSB is significantly more efficient than any other networks.

5. Proofs

Proof of Theorem 2.1. We can write the log likelihood of (X, Y) from $\mathrm{P}_{(\beta,\sigma^2,h,g)}$,

$$L(X, Y; \beta, \sigma^2, h, g) = \log[g(X)] - T\log(2\pi\sigma^2)/2 - \sum_{t=1}^{T}(Y_t - X_t'\beta)^2/2\sigma^2$$

$$+(\bar{Y} - \bar{X}'\beta)^2/2\bar{\sigma}^2 + \log\int \exp[-(\bar{Y} - \bar{X}'\beta - u)^2/2\bar{\sigma}^2]h(u)du.$$

Thus,

$$l_\beta = \frac{\partial L}{\partial\beta} = \sigma^{-2}\sum_{t=1}^{T} U_t(\beta)X_t - \bar{X}\Big(\frac{w'}{w}\Big)(\bar{Z}(\beta)),$$

$$l_{\sigma^2} = \frac{\partial L}{\partial\sigma^2} = (2\sigma^2)^{-1}\{\sum_{t=1}^{T} U_t^2(\beta)/\sigma^2 + \bar{\sigma}^{-1}\int[\bar{\sigma}^{-2}(\bar{Z}(\beta) - u)^2 - T]$$

$$\times\phi(\bar{\sigma}^{-1}(\bar{Z}(\beta) - u))h(u)du/w(\bar{Z}(\beta))\}.$$

Now, by the standard arguments in calculating tangent spaces(see, for example, Bickel, Klaassen, Ritov and Wellner 1992), V is given by

$$V = V_1 + V_2 + V_3$$

where

$$V_1 = \{a(X) : a \in L_2(\mathrm{P}),\, E_P a(X) = 0\},$$

$$V_2 = \{b(\bar{Z}) : b \in L_2(\mathrm{P}),\, E_P b(\bar{Z}) = 0\}$$

and V_3 is $[\ell_{\sigma^2}]$. Since $E[\ell_\beta|X] = 0$, ℓ_β is perpendicular to V_1, in other words $\Pi(\ell_\beta|V_1) = 0$. Hence ℓ^* can be calculated by first projecting ℓ_β on V_2 and then on the orthogonal complement of V_2 in $V_2 + V_3$, yielding

$$\ell^* = \ell_\beta - \Pi(\ell_\beta|V_2) - \Pi(\ell_\beta - \Pi(l_\beta|V_2)|W)$$

where

$$W = [\ell_{\sigma^2} - \Pi(\ell_{\sigma^2}|V_2)].$$

Moreover,

$$\Pi(\ell_\beta|V_2) = E(\ell_\beta|\bar{Z}) = -E(\bar{X})\Big(\frac{w'}{w}\Big)(\bar{Z})$$

since $E(\sum_{t=1}^T U_t X_t|\bar{Z}) = 0$ and $E(\bar{X}|\bar{Z}) = E(\bar{X})$ by the independence of U_t, X_t and $\bar{Z}$.

The proof will be completed if we show $\ell_\beta - \Pi(\ell_\beta|V_2)$ is perpendicular to W. For this, note that

$$\ell_{\sigma^2} - \Pi(\ell_{\sigma^2}|V_2) = \ell_{\sigma^2} - E(\ell_{\sigma^2}|\bar{Z}) + c(\bar{Z})$$

and

$$\begin{aligned}\ell_{\sigma^2} - E(\ell_{\sigma^2}|\bar{Z}) &= [\sum_{t=1}^T U_t^2 - E(\sum_{t=1}^T U_t^2|\bar{Z})]/2\sigma^4 \\ &= [\sum_{t=1}^T U_t^2 - \sigma^2(T-1)]/2\sigma^4.\end{aligned}$$

Furthermore,

$$E(\ell_\beta - E(\ell_\beta|\bar{Z}))(\ell_{\sigma^2} - E(\ell_{\sigma^2}|\bar{Z})) = 0$$

since $E(\sum_{t=1}^T U_t X_t)(\sum_{t=1}^T U_t^2) = 0$ by the symmetry of U_t and

$$E[\sum_{t=1}^T U_t^2 - \sigma^2(T-1)][(\bar{X} - E(\bar{X}))\frac{w'}{w}(\bar{Z})] = 0$$

by the independence of U_t, $\bar{X}$ and $\bar{Z}$. The theorem now follows since

$$E[(l_\beta - E(l_\beta|\bar{Z}))c(\bar{Z})] = 0$$

which can be easily seen by first taking conditional expectation given $\bar{Z}$.

Proof of Theorem 2.2. Let $w_N(t) = K_{s_N} * w(t) + c_N$, $I_{0N} = \int (w_N'/w_N)^2 w$ and

$$I_N = \bar{\sigma}^{-2}\Sigma_W + I_{0N}\Sigma_B.$$

First, we give a lemma which will be repeatedly used later. The proof of this lemma is omitted since it follows from parallel arguments with those for the proof of Proposition 4.5 in Bickel and Ritov(1987).

Lemma 5.1. Write $\bar{Z}$ for $\bar{Z}(\beta)$. If $c_N, s_N \to 0$ and $Nc_N^2 s_N^4 \to \infty$, then

$$E\Big(\frac{\hat{w}'}{\hat{w}}(\bar{Z},\beta) - \frac{w_N'}{w_N}(\bar{Z})\Big)^2 \to 0, \tag{5.1}$$

$$E\Big(\frac{w_N'}{w_N}(\bar{Z}) - \frac{w'}{w}(\bar{Z})\Big)^2 \to 0. \tag{5.2}$$

Furthermore, if $c_N, s_N \to 0$ and $Nc_N^2 s_N^6 \to \infty$, then

$$E\Big|\frac{\hat{w}''}{\hat{w}}(\bar{Z},\beta) - \frac{w_N''}{w_N}(\bar{Z})\Big| \to 0, \tag{5.3}$$

$$E\Big|\frac{w_N''}{w_N}(\bar{Z}) - \frac{w''}{w}(\bar{Z})\Big| \to 0. \tag{5.4}$$

Now, note that if we show

$$N^{1/2}[\hat{\beta}_N - (\beta + I_N^{-1}N^{-1}\sum_{i=1}^{N}(\sum_{t=1}^{T} U_{it}(\beta)X_{it}/\sigma^2 - (\bar{X}_i - \bar{\bar{X}})\frac{w_N'}{w_N}(\bar{Z}_i(\beta))))]$$

$$\xrightarrow{p} 0, \tag{5.5}$$

then the proof of the theorem is reduced to showing

$$I_N \xrightarrow{p} I, \tag{5.6}$$

$$(\bar{\bar{X}} - E(\bar{X}))N^{-1/2}\sum_{i=1}^{N}[\frac{w_N'}{w_N}(\bar{Z}_i(\beta))] \xrightarrow{p} 0. \tag{5.7}$$

However, the two convergences, (5.6) and (5.7), are immediate consequences of (5.2). The following lemma is the major step of the proof of (5.5). For this, write $\beta' = \beta + \Delta N^{-1/2}$ and let

$$Q_N(\Delta) = N^{-1/2}\sum_{i=1}^{N}[\sum_{t=1}^{T} U_{it}(\beta')X_{it}/\hat{\sigma}^2(\beta) - (\bar{X}_i - \bar{\bar{X}})\frac{w_N'}{w_N}(\bar{Z}_i(\beta'))],$$

$$W_N(\Delta) = N^{-1/2}\sum_{i=1}^{N}[\frac{\hat{w}'}{\hat{w}}(\bar{Z}_i(\beta'),\beta') - \frac{w_N'}{w_N}(\bar{Z}_i(\beta'))](\bar{X}_i - \bar{\bar{X}}).$$

Lemma 5.2. Under the assumptions of Theorem 2.2,

$$\sup\{|Q_N(\Delta) - Q_N(0) + I_N\Delta| : |\Delta| \le M\} \xrightarrow{p} 0 \tag{5.8}$$

$$\sup\{|W_N(\Delta)| : |\Delta| \le M\} \xrightarrow{p} 0. \tag{5.9}$$

Proof. For (5.8), it suffices to check that

$$\sup\{|Q_N(\Delta) - Q_N(0) - \frac{\partial Q_N}{\partial \Delta}(0) \cdot \Delta| : |\Delta| \le M\} \xrightarrow{p} 0, \tag{5.10}$$

$$|\frac{\partial Q_N}{\partial \Delta}(0) + I_N| \xrightarrow{p} 0. \tag{5.11}$$

For (5.11), note that

$$\begin{aligned}\frac{\partial Q_N}{\partial \Delta}(0) = &- T\hat{\Sigma}_W/\hat{\sigma}^2(\beta) + N^{-1}\sum_{i=1}^{N}\Big(\frac{w_N'}{w_N}\Big)'(\bar{Z}_i(\beta))(\bar{X}_i - E(\bar{X}))\bar{X}_i' \\ &- N^{-1}\sum_{i=1}^{N}\Big(\frac{w_N'}{w_N}\Big)'(\bar{Z}_i(\beta))(\bar{\bar{X}} - E(\bar{X}))\bar{X}_i'.\end{aligned} \tag{5.12}$$

The second term above tends to $-\Sigma_B I_0$ by the Triangular array WLLN, (5.2), (5.4) and the fact that

$$L_P\Big(\Big(\frac{w_N'}{w_N}\Big)'(\bar{Z}(\beta))(\bar{X} - E(\bar{X}))\bar{X}'\Big) \to L_P\Big(\Big(\frac{w'}{w}\Big)'(\bar{Z}(\beta))(\bar{X} - E(\bar{X}))\bar{X}').$$

Similarly the third term tends to zero. Since $\hat{\sigma}^2(\beta)\xrightarrow{p}\sigma^2$ and $\hat{\Sigma}_W\xrightarrow{p}\Sigma_W$, (5.11) is established.

Let $v^{<j>}(\Delta)$ denote the j-th component of a d-dimensional vector v. This convention applies to the rest of this paper. Note that, by the usual Taylor expansion,

$$Q_N^{<j>}(\Delta) = Q_N(0)^{<j>} + \Big[\frac{\partial Q_N^{<j>}}{\partial \Delta}(0)\Big]'\Delta + \frac{1}{2}\Delta'\Big[\frac{\partial^2 Q_N^{<j>}}{\partial \Delta^2}(\Delta^*)\Big]\Delta$$

for $j = 1, \ldots, d$ where $|\Delta^*| \le |\Delta|$. For (5.10), we need only check that

$$sup\Big\{\Big|\frac{\partial^2 Q_N^{<j>}}{\partial \Delta^{<k>}\partial \Delta^{<l>}}(\Delta)\Big| : |\Delta| \le M\Big\} \xrightarrow{p} 0,\ 1 \le j, k, l \le d.$$

But, the fact that

$$\Big|\frac{\partial^2 Q_N^{<j>}}{\partial \Delta^{<k>}\partial \Delta^{<l>}}(\Delta)\Big| \le N^{-1/2}s_N^{-3}\Big[N^{-1}\sum_{i=1}^{N}|\bar{X}_i^{<j>} - \bar{\bar{X}}^{<j>}||\bar{X}_i^{<k>}||\bar{X}_i^{<l>}|\Big]$$

establishes this.

For the proof of (5.9), first note that

$$W_N(\Delta) = W_{1N}(\Delta) - W_{2N}(\Delta)[N^{1/2}(\bar{\bar{X}} - E(\bar{X}))]$$

where

$$W_{1N}(\Delta) = N^{-1/2}\sum_{i=1}^{N}\Big[\frac{\hat{w}'}{\hat{w}}(\bar{Z}_i(\beta'),\beta') - \frac{w_N'}{w_N}(\bar{Z}_i(\beta'))\Big](\bar{X}_i - E(\bar{X})),$$

$$W_{2N}(\Delta) = N^{-1}\sum_{i=1}^{N}\Big[\frac{\hat{w}'}{\hat{w}}(\bar{Z}_i(\beta'),\beta') - \frac{w_N'}{w_N}(\bar{Z}_i(\beta'))\Big].$$

We will show that

$$\sup\{|W_{iN}(\Delta)| : |\Delta| \le M\} \xrightarrow{p} 0,\ \ i = 1, 2. \tag{5.13}$$

First, for the case i=2, note that it suffices to show that

$$E\sup\{|V_N(\Delta)| : |\Delta| \le M\}^2 \to 0$$

where

$$V_N(\Delta) = \frac{\hat{w}'}{\hat{w}}(\bar{Z}_1(\beta'),\beta') - \frac{w_N'}{w_N}(\bar{Z}_1(\beta')).$$

By (5.1), $E(V_N(0))^2 \to 0$. Hence we need only check that

$$E\sup\{|\frac{\partial}{\partial\Delta}V_N(\Delta)| : |\Delta| \le M\}^2 \to 0. \tag{5.14}$$

Observe that $\bar{Z}_i(\beta') = \bar{Z}_i(\beta) - \bar{X}_i'\Delta/\sqrt{N}$ and

$$\frac{\partial}{\partial\Delta}V_N(\Delta) = N^{-1/2}\Bigg\{\frac{\sum_{j=1}^{N}(\bar{X}_j - \bar{X}_1)K_{s_N}''(\bar{Z}_1(\beta') - \bar{Z}_j(\beta'))}{Nc_N + \sum_{j=1}^{N}K_{s_N}(\bar{Z}_1(\beta') - \bar{Z}_j(\beta'))}$$

$$-\frac{\hat{w}'}{\hat{w}}(\bar{Z}_1(\beta'),\beta') \times \frac{\sum_{j=1}^{N}(\bar{X}_j - \bar{X}_1)K_{s_N}'(\bar{Z}_1(\beta') - \bar{Z}_j(\beta'))}{Nc_N + \sum_{j=1}^{N}K_{s_N}(\bar{Z}_1(\beta') - \bar{Z}_j(\beta'))} + \bar{X}_1\Big(\frac{w_N'}{w_N}\Big)'(\bar{Z}_1(\beta'))\Bigg\}.$$

Since $|K_{s_N}^{(i)}/K_{s_N}| \le s_N^{-i}$, the right hand side of (5.14) is bounded by

$$L_3N^{-1}s_N^{-4}E[\max_{1\le j\le N}|\bar{X}_j - \bar{X}_1|^2 + |\bar{X}_1|^2] = O(N^{-1+\delta}s_N^{-4})$$

where δ is an arbitrarily small positive number. Hence we get (5.13) for i=2.

Now, for the case i=1, note that

$$E|W_{1N}(0)|^2 = E\Big[\frac{\hat{w}'}{\hat{w}}(\bar{Z}(\beta),\beta) - \frac{w_N'}{w_N}(\bar{Z}(\beta)\Big]^2 \cdot E|\bar{X} - E(\bar{X})|^2$$
$$\to 0$$

by (5.1). Hence we need only check that

$$Esup\Big\{\Big|\frac{\partial}{\partial\Delta^{<\ell>}} W_{1N}^{<k>}(\Delta)\Big| : |\Delta| \le M\Big\} \to 0,\ 1 \le \ell, k \le d.$$

Let $D_j = \bar{X}_j - E(\bar{X})$. We can write

$$\frac{\partial}{\partial\Delta^{<\ell>}} W_{1N}^{<k>}(\Delta) = -\frac{1}{N}\sum_{j=1}^{N} D_j^{<k>} \bar{X}_j^{<\ell>} \Big[\Big(\frac{\hat{w}'}{\hat{w}}\Big)'(\bar{Z}_j(\beta'),\beta') - \Big(\frac{w_N'}{w_N}\Big)'(\bar{Z}_j(\beta'))\Big]$$
$$+\frac{1}{N}\sum_{j=1}^{N} D_j^{<k>} \left[\frac{\sum_{m=1}^{N} \bar{X}_m^{<\ell>} K_{s_N}''(\bar{Z}_j(\beta') - \bar{Z}_m(\beta'))}{Nc_N + \sum_{m=1}^{N} K_{s_N}(\bar{Z}_j(\beta') - \bar{Z}_m(\beta'))}\right]$$
$$+\frac{1}{N}\sum_{j=1}^{N} D_j^{<k>} \Big[\frac{\hat{w}'}{\hat{w}}(\bar{Z}_j(\beta'),\beta')\Big] \times \left[\frac{\sum_{m=1}^{N} \bar{X}_m^{<\ell>} K_{s_N}'(\bar{Z}_j(\beta') - \bar{Z}_m(\beta'))}{Nc_N + \sum_{m=1}^{N} K_{s_N}(\bar{Z}_j(\beta') - \bar{Z}_m(\beta'))}\right]$$
$$= R_{1N}(\Delta) + R_{2N}(\Delta) + R_{3N}(\Delta),$$

say. Then, by (5.1) and (5.3), $|R_{1N}(0)| \xrightarrow{p} 0$. Since $E|R_{2N}(0)|^2 \le L_1 N^{-1} s_N^{-4}$ and $E|R_{3N}(0)|^2 \le L_2 N^{-1} s_N^{-4}$ for some positive numbers L_1 and L_2, we have $|R_{2N}(0)| \xrightarrow{p} 0$ and $|R_{3N}(0)| \xrightarrow{p} 0$. By the same arguments as in the proof of (5.14), we can show

$$E\sup\Big\{\Big|\frac{\partial}{\partial\Delta} R_{jN}(\Delta)\Big| : |\Delta| \le M\Big\}^2 \to 0,\ j = 1, 2, 3.$$

This completes the proof of (5.13).

Now, we go back to the proof of Theorem 2.2. The above lemma reduces the proof of (5.5) to establishing

$$\sup\{|\hat{\sigma}^2(\beta') - \hat{\sigma}^2(\beta)| : N^{1/2}|\beta' - \beta| \le M\} \xrightarrow{p} 0 \tag{5.15}$$

and

$$\hat{I} \xrightarrow{p} I, \tag{5.16}$$

$$N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} U_{it}(\tilde{\beta}_N) X_{it} = O_p(1). \tag{5.17}$$

This can be seen by first replacing $\hat{\beta}_N$ in (5.15) by its definition given in (2.2) and then decomposing the result into several relevant terms.

We begin with (5.15). Write $\hat{\sigma}^2(\Delta) = \hat{\sigma}^2(\beta')$. Then (5.15) is equivalent to

$$\sup\{|\hat{\sigma}^2(\Delta) - \hat{\sigma}^2(0)| : |\Delta| \le M\} \xrightarrow{p} 0. \tag{5.18}$$

Now (5.18) is an immediate consequence of the fact that

$$\begin{aligned}\frac{\partial \hat{\sigma}^2}{\partial \Delta}(\Delta) = -2N^{-1/2}(T-1)^{-1}[N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it}-\bar{Y}_i)(X_{it}-\bar{X}_i)\\ - T\hat{\Sigma}_W\beta - N^{-1/2}T\hat{\Sigma}_W\Delta].\end{aligned}$$

Now we prove (5.16). Since $\hat{\sigma}^2(\tilde{\beta}_N)\xrightarrow{p}\sigma^2$ by (5.15), and $\hat{\Sigma}_W\xrightarrow{p}\Sigma_W$, $\hat{\Sigma}_B\xrightarrow{p}\Sigma_B$, we need only check that

$$\hat{I}_0\xrightarrow{p}I_0. \tag{5.19}$$

Using the same arguments as in the proof of Lemma 5.2, we can show

$$N^{-1}\sum_{i=1}^{N}\Big[\Big(\frac{\hat{w}'}{\hat{w}}\Big)^2(\bar{Z}_i(\tilde{\beta}_N),\tilde{\beta}_N) - \Big(\frac{w_N'}{w_N}\Big)^2(\bar{Z}_i(\tilde{\beta}_N))\Big] \xrightarrow{p} 0. \tag{5.20}$$

Also note that

$$\begin{aligned}N^{-1}\sum_{i=1}^{N}\Big|\Big(\frac{w_N'}{w_N}\Big)^2(\bar{Z}_i(\tilde{\beta}_N)) - \Big(\frac{w_N'}{w_N}\Big)^2(\bar{Z}_i(\beta))\Big|\\ \le L_4 s_N^{-3}|\tilde{\beta}_N - \beta|N^{-1}\sum_{i=1}^{N}|\bar{X}_i| \qquad (5.21)\\ = O_p(N^{-1/2}s_N^{-3}).\end{aligned}$$

Combining (5.20) and (5.21), we get (5.19).

Finally, since

$$\begin{aligned}N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}U_{it}(\tilde{\beta}_N)X_{it} = -N^{1/2}T\hat{\Sigma}_W(\tilde{\beta}_N - \beta)\\ + N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}U_{it}(\beta)X_{it},\end{aligned} \tag{5.22}$$

we get (5.17).

Proof of Theorem 2.3. Useful facts about $\tilde{\beta}_{N,T}$ when X_{it} are i.i.d. are

$$T^{1/2}(\tilde{\beta}_{N,T} - \beta) = O_p(1) \tag{5.23}$$

when N is fixed and $T \to \infty$, and

$$(NT)^{1/2}(\tilde{\beta}_{N,T} - \beta) = O_p(1) \tag{5.24}$$

when both $N, T \to \infty$. We will show (ii) since the proof of (i) is virtually the same.

Drop the subscript (N,T) in $\tilde{\beta}_{N,T}$, $\hat{\beta}_{N,T}$ and $s_{N,T}$. Observe that

$$\begin{aligned}
\hat{\beta} - \beta &= (\tilde{\beta} - \beta) + \hat{I}^{-1} N^{-1} \sum_{i=1}^{N} \left[\sum_{t=1}^{T} U_{it}(\tilde{\beta}) X_{it} / \hat{\sigma}^2(\tilde{\beta}) \right] \\
&\quad - \hat{I}^{-1} N^{-1} \sum_{i=1}^{N} [(\bar{X}_i - \bar{\bar{X}}) \frac{\hat{w}'}{\hat{w}} (\bar{Z}_i(\tilde{\beta}), \tilde{\beta})] \\
&= S_1 + S_2 + S_3,
\end{aligned}$$

say. Now $\hat{I} = O_p(T)$ since $\hat{\sigma}^2(\tilde{\beta}) = O_p(1)$. Hence $|S_3|$ is bounded by $O_p(T^{-1}s^{-1})$. Note that, by (5.22) and the fact that $T\hat{I}^{-1}\hat{\Sigma}_W/\hat{\sigma}^2(\tilde{\beta}) = J + O_p(T^{-1}s^{-2})$ where J denotes the identity matrix, we have

$$S_1 + S_2 = (\tilde{\beta} - \beta) O_p(T^{-1}s^{-2}) + \hat{I}^{-1} N^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} U_{it}(\beta) X_{it} / \hat{\sigma}^2(\tilde{\beta}). \tag{5.25}$$

Combining (5.24), (5.25) and the fact that $(NT)^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} U_{it}(\beta) X_{it} = O_p(1)$, we get $(NT)^{1/2}(S_1 + S_2) = O_p(1)$. This establishes (ii).

Proof of Theorem 3.1. We can write

$$T^{1/2}(\hat{\alpha}_i - \alpha_i) = -T^{1/2} \bar{X}_i'(\hat{\beta} - \beta) + T^{1/2}[\bar{Z}_i(\beta) - \alpha_i].$$

The first term tends to zero by Theorem 2.3. The second term is equal to $T^{1/2}\bar{\epsilon}_i$ which has the $N(0, \sigma^2)$ distribution.

Proof of Theorem 3.2. The case (i) is obvious from the proof of Theorem 3.1. The case (ii) follows since both $E \max_{1 \le i \le N} |\bar{X}_i|$ and $E \max_{1 \le i \le N} |T^{1/2}\bar{\epsilon}_i|$ are $O(logN)$.

Proof of Theorem 3.3. Observe that, for any given $M > 0$,

$$\begin{aligned} \mathrm{P}[N^{\frac{1}{\gamma+1}}|\alpha_{(N)} - B(h)| > M] &= \mathrm{P}[\alpha_{(N)} < B(h) - MN^{-\frac{1}{\gamma+1}}] \\ &= [1 - h(B(h) - \delta MN^{-\frac{1}{\gamma+1}})MN^{-\frac{1}{\gamma+1}}]^N, \end{aligned}$$

where $0 \leq \delta \leq 1$. Using the condition (3.2), we can bound the probability by $(1 - D_1 M^{\gamma+1} N^{-1})^N$ for some $D_1 > 0$. Since this further can be bounded by $\exp(-D_2 M^{\gamma+1})$ for some $D_2 > 0$, the theorem follows.

References

[1] Aigner, D.J., Lovell, C.A.K. and Schmidt, P.(1977). Formulation and estimation of stochastic frontier models. *Journal of Econometrics*, 6, 21-37.

[2] Begun, J.M., Hall, W.J., Huang, W.M. and Wellner, J.A.(1983). Information and asymptotic efficiency in parametric-nonparametric models. *Annals of Statistics*, 11, 432-452.

[3] Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J.A.(1992). *Efficient and adaptive estimation in non- and semi-parametric models.* Johns Hopkins University Press. To appear.

[4] Bickel, P.J. and Ritov, Y.(1987). Efficient estimation in the errors in the variables model. *Annals of Statistics*, 15, 513-540.

[5] Cornwell, C., Schmidt, P. and Sickles, R.C.(1990). Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal pf Econometrics*, 46, 185-200.

[6] Deprins, D., Simar, L. and Tulkens, H.(1984). Measuring labor inefficiency in post offices. *The performance of Public Enterprises : Concepts and measurements*, M. Marchand, P. Pestieau and H. Tulkens (eds.), 243-267. North-Holland, Amsterdam.

[7] Farrell, M.J.(1957). The measurement of production efficiency. *Journal of the Royal Statistical Society A*, 120, 253-281.

[8] Hall, P., Härdle, W. and Simar, L.(1991). Iterated bootstrap with applications to frontier models. Unpublished manuscript.

[9] Hausman, J.A. and Taylor, W.E.(1981). Panel data and unobservable individual effects. *Econometrica*, 49, 1377-1398.

[10] Ibragimov, I.A. and Has'minskii, R.Z.(1981). *Statistical Estimation: Asymptotic Theory.* Springer, New York.

[11] Jondrow, J., Lovell, C.A.K., Materov, I.S. and Schmidt, P.(1982). On the estimation of technical inefficiency in stochastic frontier production model. *Journal of Econometrics*, 19, 233-238.

[12] Meeusen, W. and van den Broeck, J.(1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Econometric Review*, 18, 435-444.

[13] Schmidt, P. and Sickles, R.C.(1984). Production frontiers and panel data. *Journal of Business and Economic Statistics*, 2, 367-374.

[14] Simar, L.(1992). Estimating efficiencies from frontier models with panel data: A comparison of parametric, nonparametric and semi-parametric methods with bootstrapping. *Journal of Productivity Analysis*. To appear.

[15] U. I. C.(1970-1983). *Statistiques Internationales des Chemins de Fer*, Union Internationale des Chemins de Fer, Paris.