STOCHASTIC FRONTIERS: A SEMIPARAMETRIC APPROACH

by

B. Park*, R.C. Sickles and L. Simar*****

Revised November 1992

Abstract

This paper generalizes the results of Hausman and Taylor (1981), Schmidt and Sickles (1984), Cornwell, Schmidt and Sickles (1990) and Park and Simar (1992) to the efficient IV estimation of panel models in which the random effects are correlated with a subset of the regressors. The model in which this estimator has particular promise is the stochastic frontier model in which it is posited that inefficiency is correlated with certain characteristics of the determinants of technology, or observable proxies for heterogeneity in the application of that technology, which renders the random components treatment of efficiency inconsistent. In the spirit of Robinson (1988) our semiparametric model assumes a particular form for the frontier production function while considering the joint density of the individual firm-specific effects and those regressors with which they are potentially correlated as unknown. Efficiency of the slope parameters and the asymptotic properties of the level of the frontier function are explored. We illustrate our new estimator in an analysis of productive efficiency between selected European and American airlines after domestic deregulation in the U.S. and prior to recent European reforms implemented in the course of EC integration.

Key words: Stochastic frontier model, semiparametric model, efficient estimation, information bound, panel data.

---

* CORE and Institute of Statistics, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

** Department of Economics, Rice University, U.S.A.

** CORE and Institute of Statistics, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

1. Introduction

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

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

where $X_i = (X_{i1}', \ldots, X_{iT}')'$, $Z_i = (Z_{i1}, \ldots, Z_{iT})'$ and $Y_i = (Y_{i1}, \ldots, Y_{iT})'$. The X_i are i.i.d. $(d-1) \times T$-dimensional random vectors with unknown density function g, and β is a $(d-1)$-dimensional unknown vector. The ε_{it} are considered i.i.d. random variables from $N(0, \sigma^2)$ with unknown σ^2, and the $(\alpha_i, Z_i')'$ are i.i.d. from a unknown density h. The support of the marginal density of α is considered bounded above (below). It is assumed that the ε_{it}, (α_i, Z_{it}) and X_{it} are independent. Let $(\alpha, Z')'$ denote the generic observations $(\alpha_i, Z_i')'$ and $\overline{Z} = T^{-1} \sum_{t=1}^{T} Z_t$. Throughout this paper, it is assumed that the density h is of the form,

$$h(\alpha, z) = h_M(\alpha, \overline{z}) p(z)$$

where h_M is the joint density of $(\alpha, \overline{Z})$ and p is an arbitrary function. This assumption essentially tells us that Z depends on α only through $\overline{Z}$, i.e. through longrun movements in Z_i.

Although equation (1.1) is the generic panel data model, an appealing empirical motivation for the appropriate treatment of (1.1) is to estimate firm specific efficiency levels in a stochastic panel production frontier model (c.f. Schmidt and Sickles, 1984; Cornwell, Schmidt and Sickles, 1990). The availability of panel data allows one to identify the realization of α_i for a particular firm and thus overcome the limitation of single cross-sections (or time series) which allow only the identification of the expectation of α_i conditional on stochastic noise (Jondrow, Lovell, Materov and Schmidt, 1982).

In this paper we derive the asymptotic lower bound for estimating β in the presence of nuisance parameters σ^2 and h, while allowing for dependencies between tha random effects and the regressors. The efficient semiparametric estimator of β attains this lower bound. We present the estimator, as well as predictors of the individual effects whose construction utilizes the efficient estimator of β and provide proofs of their asymptotic properties. We provide an estimator of the level of the frontier function $B(h)$, which is the upper boundary of the support of h_M, and examine its rate of convergence to the desired quantity. It is this boundary estimator on which is based the estimates of firm specific relative technical efficiency. We focus on the case in which the random effects are correlated with one regressor. The generalization to an arbitrary number of regressors which leads naturally to a general semiparametric test of the orthogonality of regressors and effects is left to future research. We are unaware of any asymptotic results (beyond consistency) involving the level of the frontier function when such correlations exist (c.f. Schmidt and Sickles, 1984).

In section 2 we derive the asymptotic lower bound for estimating β and the efficient estimator of β for the generic panel data model (1.1) with correlation between α_i and the regressor Z_{it}. Section 3 considers the frontier estimation of the boundary of the support of the density of α_i and its asymptotic behavior. Section 4 discusses the use of this estimator in the modeling of efficiency differentials between selected European and American air carriers using a newly constructed data set whose use has been limited to modeling the provision of revenue service (Good, Nadiri, Roeller and Sickles, 1992). We consider here the provision of capacity service on domestic and international routes and discover a substantial gap in efficiency levels in favor of the American carriers. The data are from 1976-1986 and cover the four largest European and eight largest American carriers. Our results point to the practically and usefulness of the semiparametric frontier estimator when efficiencies are potentially correlated with incentive incompatible factors which, e.g. allow for substantial subsidies to European airlines which carry excessively bloated labor forces, ostensively to expedite the service competition objectives of the national flag carrier. They also point to the applicability of the semiparametric frontier estimator for a wide class of empirical modeling scenarios which nests most linear panel data problems. Section 5 concludes. Proofs of major theorems are in the appendix.

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). An excellent survey found in Newey (1990). Below we briefly outline the basic ideas in our context.

Write (X, Z, Y) for "generic" observations. Let $\mathbf{P}$ be the set of all possible joint distributions of (X, Z, Y). Let $\theta = (\beta', \gamma)'$. One calls $\mathbf{P}_0$ a regular parametric submodel of $\mathbf{P}$ if $\mathbf{P}_0(\subset \mathbf{P})$ can be represented as $\{P_{(\theta,\eta)} : \theta \in \mathbf{R}^d, \eta \in S$ open $\subset \mathbf{R}^k\}$ and at every (θ_0, η_0) the mapping $(\theta, \eta) \to P_{(\theta,\eta)}$ is continuously Hellinger differentiable (see Ibragimov and Has'minskii 1981, Section 1.7). Suppose $P(= P_{(\theta_0,\eta_0)})$ belongs to a regular parametric submodel $\mathbf{P}_0$ of $\mathbf{P}$. Let $L(X, Z, Y, \theta, \eta)$ denote the log likelihood of an observation from $P_{(\theta,\eta)}$ and let $\ell_\theta(X, Z, Y) = \partial\ell/\partial\theta|_{(\theta_0,\eta_0)}$ and $\ell_{\eta_j}(X, Z, Y) = \partial\ell/\partial\eta_j|_{(\theta_0,\eta_0)}$ 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 θ is, then, given by

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

where

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

and the notation $\Pi(\ell|S)$ denotes the vector of projections of each component of ℓ onto the space S in $L_2(P)$. This definition of information bound is extended to the case where we only assume that $P \in \mathbf{P}$. 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 θ under $\mathbf{P}$ is defined by

$$I(P; \theta, \mathbf{P}) = E_p \ell^* \ell^{*'}(X, Z, Y)$$

where

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

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 $I(P; \theta, \mathbf{P}_0) - I(P; \theta, \mathbf{P})$ is positive semidefinite for any regular parametric submodel $\mathbf{P}_0$ (e.g. Bickel, Klaassen, Ritov and Wellner 1992). An estimator $\hat{\theta}_N$ is called efficient in $\mathbf{P}$ if

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

Now we will exhibit the information bound for the model (1.1). To this end, let $Y = (Y_1, \ldots, Y_T)', X = (X_1', \ldots, X_T')'$ and $Z = (Z_1, \ldots, Z_T)'$ for the generic observations (X, Z, Y). Let $S_t(\theta) = Y_t - X_t'\beta - Z_t\gamma$ and $U_t(\theta) = S_t(\theta) - \overline{S}(\theta)$ where $\overline{S}(\theta) = T^{-1}\sum_{t=1}^T S_t(\theta)$. Let

$$w(z_1, z_2) = \int \phi_{\overline{\sigma}}(z_1 - u) h_M(u, z_2) du = \phi_{\overline{\sigma}} * h_M(\cdot, z_2)(z_1)$$

be the density of $(\overline{S}(\theta), \overline{Z})$ where $\overline{\sigma} = \sigma/\sqrt{T}$ and $\overline{Z} = T^{-1}\sum_{t=1}^T Z_t$, and let

$$I_0 = \int \frac{(w')^2}{w}(z_1, z_2) dz_1 dz_2$$

where and below w' denotes the first derivative of w with respect to the first component. Define

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

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

$$\Sigma_W(Z) = E_P T^{-1} \sum_{t=1}^T (Z_t - \overline{Z})^2$$

$$\Sigma_W(X, Z) = E_P T^{-1} \sum_{t=1}^T (Z_t - \overline{Z})(X_t - \overline{X})$$

$$\Sigma_W = \begin{pmatrix} \Sigma_W(X) & \Sigma_W(X, Z) \\ \Sigma_W(X, Z)' & \Sigma_W(Z) \end{pmatrix}$$

where $\overline{X} = T^{-1}\sum_{t=1}^T x_t$. Throughout this paper we will assume that $I_0 < \infty$, and $\Sigma_W, \Sigma_B(X)$ exist and are nonsingular.

Letting $I = I(P; \theta, \mathbf{P})$ and dropping the argument θ in $U_t(\theta)$ and $\overline{S}(\theta)$, we have the following theorem.

Theorem 2.1. Let $\ell^* = (\ell_\beta^{*'}, \ell_\gamma^*)'$. The efficient score function and the information bound for estimating θ for the model (1.1) are given by

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

$$\ell_\gamma^* = \sigma^{-2} \sum_{t=1}^{T} U_t Z_t,$$

$$I = \begin{pmatrix} \overline{\sigma}^{-2}\Sigma_W(X) + I_0 \Sigma_B(X) & 0 \\ 0 & \overline{\sigma}^{-2}\Sigma_W(Z) \end{pmatrix}.$$

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

Remark 2.2. It can be seen that the information bound I in Theorem 2.1 coincides with $I(P; \theta, \mathbf{P}_0)$ where $\mathbf{P}_0$ is a regular parametric submodel obtained by parametrizing $h_M(\alpha, \overline{z})$ by $h_{\mu,r}(\alpha, \overline{z}) = h_0(\alpha - \mu - r\overline{z}, \overline{z})$ for fixed h_0. This is called the *least favorable* regular parametric submodel of $\mathbf{P}$.

We now construct an efficient estimator of θ. The idea is to proceed as in the classical estimation of the location problem: Let $\tilde{\ell} = I^{-1}\ell^*$. 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 consistent estimator $\tilde{\theta}_N$ of θ.
(b) Consider $\tilde{\ell}$ as $\tilde{\ell}(x, z, y, \theta, \sigma^2, h)$. Construct a suitable estimator $\hat{\tilde{\ell}}(\cdot, \cdot, \cdot, \theta; X_1, Z_1, Y_1, \ldots, X_N, Z_N, Y_N)$ of $\tilde{\ell}(\cdot, \cdot, \cdot, \theta, \sigma^2, h)$.
(c) Form

$$\hat{\theta}_N = \tilde{\theta}_N + N^{-1} \sum_{i=1}^{N} \hat{\tilde{\ell}}(X_i, Z_i, Y_i, \tilde{\theta}_N; X_1, Z_1, Y_1, \ldots, X_N, Z_N, Y_N)$$

as the efficient estimator.

For a preliminary estimator $\tilde{\theta}_N$ of θ, we use the within estimator obtained from regressing $Y_{it} - \overline{Y}_i$ on $X_{it} - \overline{X}_i$ and $Z_{it} - \overline{Z}_i$ by the ordinary least squares method. In fact,

$$\tilde{\theta}_N = \begin{pmatrix} \tilde{\beta}_N \\ \tilde{\gamma}_N \end{pmatrix} = (NT\hat{\Sigma}_W)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \begin{pmatrix} X_{it} - \overline{X}_i \\ Z_{it} - \overline{Z}_i \end{pmatrix} (Y_{it} - \overline{Y}_i)$$

where $\hat{\Sigma}_W$ is the sample version of Σ_W. It is straightforward to show that, as $N \to \infty$ (with fixed T), we have

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

In view of Theorem 2.1, the within estimator is not efficient. However, it already lies in a $N^{1/2}$-neighborhood of θ, 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,

$$\hat{w}(t_1, t_2, \theta) = N^{-1} \sum_{i=1}^{N} K_{s_N}(t_1 - \overline{S}_i(\theta)) K_{s_N}(t_2 - \overline{Z}_i) + c_N$$

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. Note that we use the same kernel function K and the bandwidth s_N for both $\overline{S}_i$ and $\overline{Z}_i$. One may use two different kernels and bandwidths here, but we avoid this just for the simplicity of presentation. A straightforward extension to the latter case is possible.

Define

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

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

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

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

where

$$\hat{I} = \begin{pmatrix} T\hat{\Sigma}_W(X)/\tilde{\sigma}^2 + \hat{I}_0 \hat{\Sigma}_B(X) & 0 \\ 0 & T\hat{\Sigma}_W(Z)/\tilde{\sigma}^2 \end{pmatrix},$$

$$\hat{I}_0 = N^{-1} \sum_{i=1}^{N} \left(\frac{\hat{w}'}{\hat{w}} \right)^2 (\tilde{S}_i, \overline{Z}_i, \tilde{\theta}_N),$$

and $\tilde{\sigma}^2, \tilde{U}_{it}, \tilde{S}_i$ are used for $\hat{\sigma}^2(\tilde{\theta}_N), U_{it}(\tilde{\theta}_N), \overline{S}_i(\tilde{\theta}_N)$.

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

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

as $N \to \infty$. Let h_{M1} and h_{M2} denotre the marginal density of α and $\overline{Z}$ respectively.

Remark 2.3. Alternative root-n consistent first step estimators could be used. For example, the efficient instrumental variables estimator for panel frontier models introduced by Cornwell, Schmidt, and Sickles (1990) provides a more efficient first step estimator of β and γ when X_{it} is orthogonal to the effects. We use the within IV estimator here in part to highlight the role of our estimator in testing arbitrary orthogonality conditions between the random effects and selected sets of regressors. If one was uncertain about which regressors were orthogonal to the effects, then a more robust test would rely on differences among the within estimates and our efficient semiparametric estimates instead of possibly more powerful tests involving differences between the efficient IV estimates and the efficient semiparametric estimates using effiicent IV in the first step. Moreover, the use of within estimates to form our efficient semiparametric estimator points to an intuitive result that within estimates of γ in (1.1) are equivalent to the efficient semiparametric estimator of γ using within in the first step. This can be seen by inspection of (2.2). Since the least squares residuals $\hat{U}_{it}$ are orthogonal to Z_{it} by construction and since $\hat{I}$ is block diagonal, the last entry in $\hat{\theta}_N$, $\hat{\gamma}$, is equivalent to the last entry in $\tilde{\theta}_N$, $\tilde{\gamma}$, which is the within estimate of γ.

Theorem 2.2. Assume $\int e^{t|\overline{x}|} g(x)dx < \infty$ for some $t > 0$, and $\displaystyle\int |u|^2 h_M(u,z)dudz < \infty$. If h_M, h_{M1} and h_{M2} are bounded, then

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

Remark 2.4. The assumption on g in the above theorem is far from being optimal for the asymptotic efficiency of the estimator. In fact, existence of moments $E_P|\overline{X}|^r$ 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 $\hat{\theta}_N$ when T also goes to infinity. For this, we need to assume that $R_{it} = (X_{it}', Z_{it})'$ 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. It needs also some additional asumptions on the consistency of the first step estimator $\tilde{\theta}_{N,T}$ which are obviously shared by our *within* estimator of θ.

Theorem 2.3. Let R denote the generic observations R_{it}. Suppose $E_P|R|^2 < \infty$ and $E_P(R - E_P(R))(R - E_P(R))'$ is nonsingular.

(i) If

$$\sqrt{T}(\tilde{\theta}_{N,T} - \theta) = O_p(1) \quad \text{when } N \text{ is fixed and } T \to \infty \tag{2.3}$$

then, when N fixed, $T \to \infty$ and $Ts^2_{N,T} \to \infty$,

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

(ii) Further if

$$\sqrt{(N,T)}(\tilde{\theta}_{N,T} - \theta) = O_p(1) \text{ when both } N, T \to \infty \tag{2.4}$$

then, when $N, T \to \infty$, $Ts^2_{n,T} \to \infty$ and $NT^{-1}s^{-2}_{N,T} = O(1)$,

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

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 $\hat{\theta}_{N,T}$ at hand, it would be natural to predict α_i by

$$\hat{\alpha}_i = \overline{S}_i(\hat{\theta}_{N,T}).$$

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 R_{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 $Ts^2_{N,T} \to \infty$,

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

Remark 3.1. When one is interested in predicting the difference $\alpha_i - \alpha_j$ (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.

If we let B denote the upper boundary of the support of h_{M1}, then the frontier function will be $B + x'\beta + z\gamma$. Since we already have an efficient estimator of θ, it remains to estimate B to get an estimator of this function. A natural estimator of this quantity would be

$$\hat{B} = \max_{1 \le i \le N} \overline{S}_i(\hat{\theta}_{N,T}).$$

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 $T s_{N,T}^2 \to \infty$,

(i) when N is fixed,

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

(ii) when N goes to infinity,

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

How close $\alpha_{(N)}$ to B is illustrated by the following theorem. For this, suppose the density h_{M1} satisfies

$$h_{M1}(B - x) \ge C x^{\gamma} \tag{3.2}$$

as $x \downarrow 0$ for some constant $C > 0$ and $\gamma \ge 0$. This condition requires that the density h_{M1} should have a certain amount of mass near the boundary point B.

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

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

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),

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

if both N and T go to infinity.

4. Data and Empirical results

Our airline data set consists of a panel of the four largest European carriers and the seven largest American airlines. These four European carriers supply approximately half of the international traffic of all airlines based in Europe. The seven U.S. carriers supply virtually all of the scheduled international traffic of the U.S. airline industry. We follow these carriers with annual observations during the time period 1976 to 1986. The primary source for our data is the *The Digest of Statistics* from the International Civil Aviation Organization (ICAO). Using this source, we construct a set of four airlines inputs: Labor, Energy, Materials and Aircraft Fleet. In addition, aggregate airline outputs and several of its characteristics are constructed. Details of data construction can be found in Good, et al. (1992).

Our labor input is an aggregate of five separate categories of employement used in the production of air travel. These categories include pilots (as well as copilots and other cockpit crew), flight attendants, mechanics, ticketing and passenger handlers, and other employees. A Tornqvist multilateral index number is used to aggregate these subcomponents.

Fuel expenses are given for each carrier in ICAO's *Financial Data Series*. The airline's fuel price is estimated as a weighted average of the domestic fuel price (weighted by domestic available ton-kilometers), and regional prices (weighted by international available ton-kilometers in the relevant region).

Expenditures on supplies, services and ground based capital equipment are combined into a single input aggregate called materials. Purchasing power parity exchange rates are adjusted by allowing for changes in market exchange rates and changes in price levels. These subcomponents are aggregated using a multilateral index number procedure and are termed materials.

Detailed information for aircraft fleets is provided in ICAO's *Fleet and Personnel Series*. In addition to counts of the total number of aircraft, we construct the percentage of the fleet which is widebodied (a measure of average equipment size) and the percentage of the fleet which is turboprop (a measure of aircraft speed).

Information regarding the carrier's physical outputs are obtained from ICAO's

Commercial Airline Traffic Series. While this source permits several possible disaggregations, we consider three components of airline output: passenger service, cargo operations, and incidental services (which includes equipment leasing, maintenance provided to other carrier's equipment, etc.). Revenues for passenger, cargo and incidental outputs are obtained in ICAO's *Financial Data Series*. The country's purchasing power parity is used as a price deflator for incidental output. Revenue ton-kilometers is used as a quantity deflator for passenger and cargo outputs. As with the inputs, these three components are aggregated using a multilateral index number procedure. Capacity ton miles is then derived as the ratio of revenue ton mile and passenger load factor.

Two characteristics of airline output are also calculated. An average stage length which provides a measure of the length of individual route segments in the carrier's network and the number of route kilometers which provides a measure of total network size.

Efficient semiparametric and within parameter estimates and standard errors for the stochastic panel frontier production function are provided in table 1. We base the efficient estimator of β on a bandwidth s chosen by the bootstrap procedure proposed by Park and Simar (1992). For a given bandwith s, compute $\hat{\beta}_{N,T}(s)$. Then resample in each (X_i, Y_i) to construct M pseudo-samples $\{(X_i^*, Y_i^*)^{(m)}\}_{i=1}^N, \quad m = 1, \ldots, M$, of size T. Let $\hat{\beta}_{N,T}^{*(m)}(s)$ be the pseudo sample bootstrap version of $\hat{\beta}_{N,T}(s)$. The bootstrap optimal choice is then given by $s^* = \arg\min_s C(s)$ where $C(s) = (1/M)\sum_{m=1}^M [\hat{\beta}^{*(m)}(s) - \hat{\beta}_{N,T}(s)]'[\hat{\beta}_{N,T}^{*(m)}(s) - \hat{\beta}_{N,T}(s)]$. Our estimates are based on $M = 1000$ and consisted of a grid search in the interval $[0.1, 2.0]$. The optimal value of s^* was 0.13. We then reestimated β using this selected bandwith on the original data.

As table 1 indicates, output elasticities of the factor inputs vary somewhat, from 0.26 for capital to 0.46 for labor. Efficiency gains from our efficient semiparametric frontier estimator vis-à-vis the within estimator range from a high of about 55% (for labor) to less than 1% (for logarithm of network size). The technology is assumed to have long run constant returns to scale with respect to these factor inputs. The t-statistics for the test of constant returns to scale in the provision of capacity service is -0.98. Average annual productivity growth over the sample was a relatively healthy 3.1% for these firms. Owing to their clear size disadvantage, fleets of turboprop aircraft were unable to provide the route capacity enjoyed by jet aircraft. Of the two network characteristics, only the network variable appears to have statistically significant explanatory power. A ten percent increase in density (ten percent reduction in number of route kilometers flown) would tend to increase the provision of route capacity by about 1.24 percent. The model fits for the two treatments for random efficiency effects are quite comparable. Hausman-Wu specification tests on the orthogonality conditions imbedded in our efficient semiparametric estimator yield a $\mathcal{X}^2$-statistic of 8.858 which is significant at only the 19% level and indicates that the orthogonality of firm efficiency levels with all inputs (other than long-run movements in the work force), capital characteristics and output characteristics cannot be rejected at conventional significance levels. Table 2 presents the firm specific relative efficency levels and rankings based on the efficient semiparametric and within estimates. The relative efficiencies and rankings are quite comparable across the two statistical treatments. An efficiency gap of approximately 15% appears to exist between the European and American carriers. These relative efficiencies and rankings are in rough agreement with those found in Good, Nadiri, Roeller, and Sickles (1992) using revenue service output, the within estimator and a different temporal parameterization of efficiency.

5. Conclusions

This paper has introduced a new estimator which can be used to model random effects when the effects are orthogonal to selected regressors. If the effects are interpreted as firm inefficiencies in a panel stochastic frontier, then our new estimator can be viewed as a semiparametric generalization of the efficient IV estimator introduced by Hausman and Taylor (1981) and Cornwell, Schmidt and Sickles (1990). We have illustrated our estimator by considering a newly constructed panel of European airlines and their American competitors to ascertain efficiency differentials in the provision of capacity service while allowing for inefficiency to be correlated with long-run adjustements in the personnel of national flag carriers. Often the national flag carriers are viewed by their governments as an extension of government and, as such, the potential for labor slack is quite real. Efficiency differences between the European and American carriers in our sample point to a gap of approximately 15% over the sample period 1976-86.

References:

[1] 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.

[2] 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.

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

[4] Good, D., Nadiri, M.I., Roller, Lars-Hendrik, and Sickles, R.C. (1992), Some Implications for the Efficiency and Welfare Gains in the European Airline Industry. *Journal of Productivity Analysis*, special issue edited by J. Mairesse and Z. Griliches, forthcoming.

[5] Hardy, G.H., Littlewood, J.E., and Polya, G. (1952), Inequalities. 2nd ed., Cambridge University Press.

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

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

[8] 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.

[9] Newey, W.K. (1990), Semiparametric efficiency bounds. *Journal of Applied Econometrics*, 5, 99–136.

[10] Park, B.U. and Simar, L. (1992), Efficient semiparametric estimation in stochastic frontier models. Unpublished manuscript.

[11] Robinson, P. (1988), Root-N-consistent semiparametric regression. *Econometrica 56*, 931–954.

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

Table 1

Efficient Semiparametric and Within Parameter Estimates And Standard Errors

	Efficient Semiparametric		Within	
Variable	Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
lnCapital	0.2717	0.05903	0.2744	0.07222
lnLabor	0.4611	0.04130	0.4611	0.07512
LnMaterials	0.2672	0.07204	0.2645	0.07464
Percent Turbo	−0.1603	0.13566	−0.1814	0.14307
Percent Wide	−0.0867	0.17453	−0.1189	0.21618
lnStagelength	0.0011	0.07774	−0.0084	0.07815
lnNetworksize	−0.1210	0.04881	−0.1248	0.04935
time	0.0312	0.00283	0.0315	0.00342
Adjusted R^2	0.994		0.999	

Table 2

Relative Efficiency Levels and Rankings of European and American Airlines in the Provision of Capacity Service (1976-1986)

	Efficient Semiparametric		Within	
	Level	Ranking	Level	Ranking
European Carriers				
Air France	0.695	7	0.700	5
Alitalia	0.546	11	0.543	11
British Air	0.640	9	0.644	9
Lufthansa	0.630	10	0.628	10
American Carriers				
American	0.726	4	0.722	4
Continental	0.702	5	0.695	6
Delta	0.700	6	0.693	7
Eastern	0.670	8	0.665	8
Northwest	1.000	1	1.000	1
TWA	0.761	3	0.759	3
United	0.769	2	0.766	2

Appendix

Proof of Theorem 2.1. We can write the log likelihood of (X, Z, Y) from $P_{(\theta,\sigma^2,h,g)}$,

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

$$(\overline{Y}-\overline{X}'\beta-\overline{Z}\gamma)^2/2\overline{\sigma}^2 + \log\int \exp[-(\overline{Y}-\overline{X}'\beta-\overline{Z}\gamma-u)^2/2\overline{\sigma}^2]h_M(u,\overline{Z})du + \log p(Z).$$

Thus,

$$\ell_\beta = \frac{\partial L}{\partial \beta} = \sigma^{-2}\sum_{t=1}^{T} U_t(\theta)X_t - \overline{X}\left(\frac{w'}{w}\right)(\overline{S}(\theta),\overline{Z}),$$

$$\ell_\gamma = \frac{\partial L}{\partial \gamma} = \sigma^{-2}\sum_{t=1}^{T} U_t(\theta)Z_t - \overline{Z}\left(\frac{w'}{w}\right)(\overline{S}(\theta),\overline{Z}),$$

$$\ell_{\sigma^2} = \frac{\partial L}{\partial \sigma^2} = (2\sigma^2)^{-1}\{\sum_{t=1}^{T} U_t^2(\theta)/\sigma^2 + \overline{\sigma}^{-1}\int[\overline{\sigma}^{-2}(\overline{S}(\theta)-u)^2 - T]$$
$$\times\,\phi(\overline{\sigma}^{-1}(\overline{S}(\theta)-u))h_M(u,\overline{Z})du/w(\overline{S}(\theta),\overline{Z})\}.$$

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

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

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

and V_3 is $[\ell_{\sigma^2}]$. Let $\ell = (\ell_\beta', \ell_\gamma)'$. By the prescription described in Section 2, ℓ^* can be calculated by projecting ℓ onto V. This can be done by first projecting ℓ onto $V_1 + V_2$ and then onto the orthogonal complement of $V_1 + V_2$ in $V_1 + V_2 + V_3$, yielding

$$\ell^* = \begin{pmatrix} \ell_\beta^* \\ \\ \ell_\gamma^* \end{pmatrix} = \ell - \Pi(\ell|V_1+V_2) - \Pi(\ell - \Pi(\ell|V_1+V_2)|W_1)$$

where

$$W_1 = [\ell_{\sigma^2} - \Pi(\ell_{\sigma^2}|V_1+V_2)].$$

Calculation of ℓ^*_β is very similar to calculation of ℓ^* in Park and Simar (1992) and hence ommitted. For calculating ℓ^*_γ, first note that

$$\begin{aligned}\ell_\gamma - \Pi(\ell_\gamma|V_1+V_2) = \ell_\gamma &- \Pi(\ell_\gamma|V_1) \\ &- \Pi(\ell_\gamma - \Pi(\ell_\gamma|V_1)|W_2)\end{aligned}$$

where $W_2 = \{b(\overline{S},\overline{Z}) - Eb(\overline{S},\overline{Z}) : b \in L_2(P)\}$. Furthermore,

$$\begin{aligned}\ell_\gamma - \Pi(\ell_\gamma|V_1) &= \ell_\gamma - E(\ell_\gamma|X) \\ &= \sigma^{-2}\sum_{t=1}^{T} U_t(\theta)Z_t - \overline{Z}\frac{w'}{w}(\overline{S},\overline{Z}) \\ &\quad + E[\overline{Z}\frac{w'}{w}(\overline{S},\overline{Z})]\end{aligned}$$

Now,

$$\begin{aligned}\Pi[\ell_\gamma - \Pi(\ell_\gamma|V_1)|W_2] &= -\overline{Z}\frac{w'}{w}(\overline{S},\overline{Z}) + E[\overline{Z}\frac{w'}{w}(\overline{S},\overline{Z})] \\ E(U_tZ_t|\overline{S},\overline{Z}) &= E(U_t)E(Z_t|\overline{S},\overline{Z}) = 0\end{aligned}$$

The first equality above follows from the fact that U_t and $(\overline{S},\overline{Z},Z_t)$ are independent. Hence

$$\ell_\gamma - \Pi(\ell_\gamma|V_1+V_2) = \sigma^{-2}\sum_{t=1}^{T} U_t(\theta)Z_t.$$

Now we need only verify that $\sum_{t=1}^{T} U_t(\theta)Z_t$ is perpendicular to W_1. But this follows from the facts

$$\ell_{\sigma^2} - \Pi(\ell_{\sigma^2}|V_1+V_2) = (2\sigma^4)^{-1}\sum_{t=1}^{T}[U_t^2 - E(U_t^2)],$$

$$E(\sum_{t=1}^{T} U_t \cdot \sum_{t=1}^{T} U_t^2) = 0.$$

Proof of Theorem 2.2. Let $f_1 * f_2$ denotes the convolution of two dimensional density functions f_1 and f_2, i.e.,

$$f_1 * f_2(x_1,x_2) = \int f_1(x_1-y_1, x_2-y_2)f_2(y_1,y_2)dy_1dy_2.$$

Then we can write

$$\begin{aligned}w_N(t_1,t_2) = E\hat{w}(t_1,t_2) &= L_{s_N} * w + c_N \\ &= \int \phi_{\overline{\sigma}}(t_1-u)h_N(u,t_2)du + c_N\end{aligned} \qquad (A.1)$$

where $h_N = h_M * L_{s_N}(t_1, t_2) = K_{s_N}(t_1)K_{s_N}(t_2)$. The proof is then very similar to the proof of Theorem 2.2 in Park and Simar (1992). One thing that is not so straightforward to prove is

$$E\left[\frac{w'_N}{w_N}(\overline{S}, \overline{Z}) - \frac{w'}{w}(\overline{S}, \overline{Z}))\right]^2 \to 0. \tag{A.2}$$

We will show (A.2).

From (A.1), we can write

$$w'_N(t_1, t_2) = \overline{\sigma}^{-3} \int (u - t_1)\phi[\overline{\sigma}^{-1}(t_1 - u)]h_N(u, t_2)du.$$

The absolute value of the above term is now bounded by

$$\overline{\sigma}^{-2} \int |u - t_1|h_N(u, t_2)du \cdot w_N(t_1, t_2)$$

by an inequality of Chebyshev [Hardy, Littlewood and Pólya (1952), page 43]. Hence

$$\left(\frac{w'_N}{w_N}(t_1, t_2)\right)^2 \leq 2\overline{\sigma}^{-4}[\int |u|^2 h_N^2(u, t_2)du + \int |t_1|^2 h_N^2(u, t_2)du].$$

The fact that $w'_N/w_N(t_1, t_2) \to w'/w(t_1, t_2)$ and $h_N(t_1, t_2) \to h_M(t_1, t_2)$ for every (t_1, t_2) reduces, by dominated convergence, the proof of (A.2) to showing

$$E\int |u|^2 h_N^2(u, \overline{Z})du \to E\int |u|^2 h_M^2(u, \overline{Z})du \tag{A.3}$$

$$E\int |\overline{S}|^2 h_N^2(u, \overline{Z})du \to E\int |\overline{S}|^2 h_M^2(u, \overline{Z})du \tag{A.4}$$

For (A.3), by exchanging the order of integrations,

$$E\int |u|^2 h_N^2(u, \overline{Z})du = \int\int[\int |u|^2 h_N^2(u, t_2)w(t_1, t_2)dt_1]dudt_2.$$

Since h_N and the marginal densities of w are bounded, the integral inside the bracket is bounded by $C|u|^2 h_N(u, t_2)$. Furthermore,

$$\int\int |u|^2 h_N(u, t_2)dudt_2 = E|\alpha|^2 + s_N^2 \int x^2 K(x)dx$$

$$\to E|\alpha|^2 = \int\int |u|^2 h_M(u, t_2)dudt_2.$$

By dominated convergence, (A.3) is verified. Finally, (A.4) is an immediate consequence of boundedness of h_N and dominated convergence.

Proof of Theorem 2.3. For simplicity drop the subscript (N, T) in $\tilde{\theta}_{N,T}, \hat{\theta}_{N,T}$ and $s_{N,T}$. From (2.2) we have

$$\begin{aligned}\hat{\theta} - \theta = (\tilde{\theta} - \theta) + N^{-1}\hat{I}^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} U_{it}(\tilde{\theta})R_{it}/\hat{\sigma}^2(\tilde{\theta}) \\ - N^{-1}\hat{I}^{-1}\begin{pmatrix}\sum_{i=1}^{N}(\overline{X}_i - \overline{X})\frac{\hat{w}'}{\hat{w}}(\overline{S}_i(\tilde{\theta}), \overline{Z}_i, \tilde{\theta}) \\ 0\end{pmatrix} \\ = A_1 + A_2 + A_3\end{aligned}$$

Since $\hat{\sigma}^2(\tilde{\theta}) = O_p(1)$ we have $\hat{I} = O_p(T)$, hence $|A_3|$ is bounded by $O_p(T^{-1}s^{-1})$. Observe that

$$\sum_{i=1}^{N}\sum_{t=1}^{T} U_{it}(\tilde{\theta})R_{it} = -NT\hat{\Sigma}_W(\tilde{\theta} - \theta) + \sum_{i=1}^{N}\sum_{t=1}^{T} U_{it}(\theta)R_{it} \qquad (A.5)$$

Noting that $T\hat{I}^{-1}\hat{\Sigma}_W/\hat{\sigma}^2(\tilde{\theta}) = J + O_p(T^{-1}s^{-2})$ where J is the identity matrix, we have by (A.5)

$$A_1 + A_2 = (\tilde{\theta} - \theta)O_p(T^{-1}s^{-2}) + N^{-1}\hat{I}^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} U_{it}(\theta)R_{it}/\hat{\sigma}^2(\tilde{\theta})$$

Combining (2.4), (A.6) and the fact that $(NT)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} U_{it}(\theta)R_{it} = O_p(1)$, we get $(NT)^{1/2}(A_1 + A_2) = O_p(1)$. This establishes *(ii)*.

The proof of *(i)* is virtually the same, hence omitted.

Remark: Note that when $\tilde{\theta}$ is the *within* estimator, A_2 is identically equal to zero (see remark 2.3).

Proof of Theorem 3.1 Since $\alpha_i = \overline{S}_i(\theta) - \overline{\varepsilon}_i$ and $\hat{\alpha}_i = \overline{S}_i(\hat{\theta})$ we have:

$$T^{1/2}(\hat{\alpha}_i - \alpha_i) = -T^{1/2}\overline{R}_i'(\hat{\theta} - \theta) + T^{1/2}(\overline{S}_i(\theta) - \alpha_i)$$

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

Proof of Theorem 3.2. Case *(i)* is obvious from Theorem 3.1. Case *(ii)* follows since $E \max_{1 \le i \le N} |\overline{X}_i|, E \max_{1 \le i \le N} |\overline{Z}_i|$ and $E \max_{1 \le i \le N} |T^{1/2}\overline{\varepsilon}_i|$ are $O(\log N)$.

Proof of Theorem 3.3. For any given $M > 0$,

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

where $0 \le \delta \le 1$. Using 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 can be bounded by $\exp(-D_2 M^{\gamma+1})$ for some $D_2 > 0$, the theorem follows.