

# Semiparametric Efficient Estimation of Dynamic Panel Data Models

Byeong U. Park\*  
Department of Statistics  
Seoul National University

Robin C. Sickles  
Department of Economics  
Rice University

Léopold Simar†  
Institut de Statistique  
Université Catholique de Louvain

January 29, 2002

## Abstract

This paper extends the semiparametric efficient treatment of panel data models pursued by Park and Simar (1994) and Park, Sickles, and Simar (1998, 2002) to a dynamic panel setting. We develop a semiparametric efficient estimator under minimal assumptions when the panel model contains a lagged dependent variable. We apply this new estimator to analyze the structure of demand between city pairs for selected U. S. airlines during the period 1979 I to 1992 IV.

---

\*Research supports by the Brain Korea 21 Project are acknowledged

†Research support from “Projet d’Actions de Recherche Concertées” (No. 98/03–217) from the Belgian Government is acknowledged.

# 1 Introduction

Our paper extends the semiparametric efficient treatment of panel data models pursued by Park and Simar (1994) and Park, et al. (1998, 2002) to a dynamic panel setting. We develop a semiparametric efficient estimator under minimal assumptions when the panel model contains a lagged dependent variable. We apply this new estimator in an analysis of the structure of dynamic demand in markets (city-pairs) for selected U. S. airlines during the period 1 I to 1992 IV. Our results suggest that this estimator may have advantages over parametric estimators in regard to efficiency gains in parameter estimates in general in the autoregressive parameter in particular.

## 2 Main Results

The model we analyze in this paper is a dynamic panel data model that can be written as:

$$Y_{it} = \gamma Y_{i,t-1} + \beta' X_{it} + \alpha_i + \varepsilon_{it}; \quad i = 1, \dots, n; t = 1, \dots, r \quad (2.1)$$

where  $X_{it} \in \mathbb{R}^d$ ,  $\beta \in \mathbb{R}^d$  and  $\varepsilon_{it}$  are iid random variables from a  $\mathcal{N}(0, \sigma^2)$  with an unknown  $\sigma^2$ . We assume  $|\gamma| < 1$  and  $Y_{i,0} = 0$ . The random effects  $\alpha_i$  are assumed to be independent and have an unknown common density function  $h$ . Write  $\varepsilon_i \equiv (\varepsilon_{i1}, \dots, \varepsilon_{ir})'$ ,  $X_i \equiv (X_{i1}', \dots, X_{ir}')'$ , and  $Y_i \equiv (Y_{i1}, \dots, Y_{ir})'$ . The random covariates  $X_i$  are independent and identically distributed with an unknown density function  $g$  defined on  $\mathbb{R}^{dr}$ . It is assumed that  $\varepsilon$ 's,  $\alpha$ 's and  $X$ 's are independent. In this section we address efficient estimation of the parameters  $\beta$  and  $\gamma$  in the presence of the nuisance parameters  $\sigma^2$ ,  $h$  and  $g$ . Note that the parameter spaces for  $h$  and  $g$  are of infinite dimension while those for  $\beta$ ,  $\gamma$  and  $\sigma^2$  are of finite dimension, so the model (2.1) is semiparametric.

We speak of efficiency as  $n$  tends to infinity with the time period  $r$  being fixed. The notion of efficiency in the semiparametric world is well explained in Bickel *et al.* (1993). There is a Fisher-like information matrix, say  $I$ , such that all *regular* estimators have asymptotic covariance matrices that are greater than or equal to  $I$  (Hájek-Le Cam's Convolution Theorem, see Theorem 2.3.1 of Bickel *et al.*, 1993). Here, we say an estimator  $\hat{\delta}_n$  of  $q(\theta)$  is regular if the law of  $\sqrt{n}(\hat{\delta}_n - q(\theta_n))$  under  $P_{\theta_n}$  converges to a limit law whenever  $\sqrt{n}|\theta_n - \theta|$  stays bounded, and if the limit distribution does not depend on the choice of  $\{\theta_n\}$ . We call  $\hat{\delta}_n$  *efficient* if its limit law is  $\mathcal{N}(0, I^{-1})$ . In the next subsection we exhibit the information matrix  $I$  for estimating  $(\beta', \gamma)'$  in the presence of the nuisance parameters  $\sigma^2$ ,  $h(\cdot)$  and  $g(\cdot)$ , and then in the second subsection we construct an efficient estimator of  $(\beta', \gamma)'$ .

## 2.1 Information bound

Let  $Z_{1t} \equiv Z_{1t}(\beta, \gamma) = Y_{1t} - \gamma Y_{1,t-1} - \beta' X_{1t}$  and  $\bar{Z}_1 \equiv \bar{Z}_1(\beta, \gamma) = \sum_{t=1}^r Z_{1t}(\beta, \gamma)/r$ . Define  $\bar{\sigma}^2 = \sigma^2/r$ . Then we can write  $Z_{1t} = \alpha_1 + \varepsilon_{1t}$ ,  $\bar{Z}_1 = \alpha_1 + \bar{\varepsilon}_1$  and  $Z_{1t} - \bar{Z}_1 = \varepsilon_{1t} - \bar{\varepsilon}_1$ . The probability density function for  $\bar{Z}_1$  is given by

$$w(z) \equiv w(z; \sigma^2, h(\cdot)) = \int (2\pi\bar{\sigma}^2)^{-1/2} \exp\{-(z-u)^2/(2\bar{\sigma}^2)\} h(u) du.$$

Thus, the log-likelihood with a single observation  $(X_1, Y_1)$  is given by

$$\begin{aligned} L(\beta, \gamma, \sigma^2, h(\cdot), g(\cdot); X_1, Y_1) &= \log g(X_1) - \frac{r}{2} \log(2\pi\sigma^2) - \sum_{t=1}^r \frac{Z_{1t}^2}{2\sigma^2} + \frac{\bar{Z}_1^2}{2\bar{\sigma}^2} \\ &\quad + \log w(\bar{Z}_1) + \frac{1}{2} \log(2\pi\bar{\sigma}^2). \end{aligned} \quad (2.2)$$

Write  $P_{\beta, \gamma, \sigma^2, h, g}$  for the probability distribution of  $(X_1, Y_1)$  corresponding to  $\beta, \gamma, \sigma^2, h$  and  $g$ . Let  $\beta_0, \gamma_0, \sigma_0^2, h_0$  and  $g_0$  be the true values and the true functions, thus the true probability distribution is  $P_0 = P_{\beta_0, \gamma_0, \sigma_0^2, h_0, g_0}$ . For the time being, let us suppose the model (1), denoted by  $\mathcal{P}$ , is parametric. Let  $\mathcal{P} = \{P_{\beta, \gamma, \sigma^2, h(\cdot; \eta_1), g(\cdot; \eta_2)} : \beta \in \mathbb{R}^d, \gamma \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+, \eta_1 \in S_1, \eta_2 \in S_2\}$  for some open  $S_1, S_2 \subset \mathbb{R}$  where  $h(\cdot; \eta_1)$  and  $g(\cdot; \eta_2)$  are known except  $\eta_1$  and  $\eta_2$ . If the maps  $\eta_1 \rightarrow h^{1/2}(\cdot; \eta_1)$  and  $\eta_2 \rightarrow g^{1/2}(\cdot; \eta_2)$  from  $\mathbb{R}$  to  $L_2(\mu)$  ( $\mu$  is the Lebesgue measure) are “smooth”, then the model  $\mathcal{P}$  is regular. (See Ibragimov & Has'minskii, 1981, Section 1.7, or Bickel *et al.*, 1993, Section 2.1). For this regular parametric model  $\mathcal{P}$ , the information matrix, denoted by  $I(P_0 | \beta, \gamma, \mathcal{P})$ , for estimating  $\beta$  and  $\gamma$  in the presence of the nuisance parameters  $\sigma^2, \eta_1$  and  $\eta_2$  is well defined and can be computed in the following way.

Write  $L = L(\beta, \gamma, \sigma^2, h(\cdot; \eta_1), g(\cdot; \eta_2); X_1, Y_1)$  and define  $\ell_\beta = \partial L / \partial \beta |_{\beta_0, \gamma_0, \sigma_0^2, \eta_{10}, \eta_{20}}$  where  $\eta_{10}$  and  $\eta_{20}$  are the parameter values such that  $h_0 = h(\cdot, \eta_{10})$  and  $g_0 = g(\cdot, \eta_{20})$ . Define  $\ell_\gamma, \ell_{\sigma^2}, \ell_{\eta_1}$  and  $\ell_{\eta_2}$ , likewise. Let  $[\ell_{\sigma^2}, \ell_{\eta_1}, \ell_{\eta_2}]$  be the linear span generated by  $\ell_{\sigma^2}, \ell_{\eta_1}$  and  $\ell_{\eta_2}$ . Define  $\ell_\beta^* = \ell_\beta - \Pi(\ell_\beta | [\ell_{\sigma^2}, \ell_{\eta_1}, \ell_{\eta_2}])$ , and likewise define  $\ell_\gamma^*$ , where  $\Pi(u | \mathcal{S})$  denotes the vector of projections of each component of  $u$  onto the space  $\mathcal{S}$  in  $L_2(\mu)$ . Write  $\ell^* = (\ell_\beta^*, \ell_\gamma^*)'$ . The information matrix is then given by

$$I(P_0 | \beta, \gamma, \mathcal{P}) = E_{P_0} \ell^* \ell^{*'} \quad (2.3)$$

It is known that the right hand side of (2.3) equals the inverse matrix of the  $(d+1) \times (d+1)$  left-top block of the matrix  $\{E_{P_0} \ell \ell'\}^{-1}$  where  $\ell = (\ell'_\beta, \ell'_\gamma, \ell'_{\sigma^2}, \ell'_{\eta_1}, \ell'_{\eta_2})'$ . It is also known that if  $\delta$  is a Gaussian regular estimator of  $(\beta'_0, \gamma_0)'$  with asymptotic covariance  $\Sigma(P_0, \delta)$  then

$$\Sigma(P_0, \delta) \geq I^{-1}(P_0 | \beta, \gamma, \mathcal{P}),$$

where  $A \geq B$  for matrices  $A$  and  $B$  means that  $A - B$  is nonnegative definite (Bickel *et al.*, 1993, Section 2.3).

Now we go back to the original semiparametric model where the spaces for  $h$  and  $g$  are of infinite dimension. Consider classes of functions  $h(\cdot; \eta_1)$  and  $g(\cdot; \eta_2)$  indexed by  $\eta_1, \eta_2 \in \mathbb{R}$  such that  $h(\cdot; 0) = h_0$  and  $g(\cdot; 0) = g_0$ . Form a parametric submodel  $\mathcal{P}_0 = \{P_{\beta, \gamma, \sigma^2, h(\cdot; \eta_1), g(\cdot; \eta_2)} : \beta \in \mathbb{R}^d, \gamma \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+, \eta_1 \in \mathbb{R}, \eta_2 \in \mathbb{R}\}$ . If we choose  $h(\cdot; \cdot)$  and  $g(\cdot; \cdot)$  so that the maps  $\eta_1 \rightarrow h^{1/2}(\cdot; \eta_1)$  and  $\eta_2 \rightarrow g^{1/2}(\cdot; \eta_2)$  from  $\mathbb{R}$  to  $L_2(\mu)$  are “smooth”, then  $\mathcal{P}_0$  is a regular parametric submodel of  $\mathcal{P}$  and the information matrix  $I(P_0 | \beta, \gamma, \mathcal{P}_0)$  can be defined as at (2.3). Consider the class of all such regular parametric submodels, and write it  $\mathcal{C}$ . Suppose an estimator  $\delta$  of  $(\beta', \gamma)'$  is Gaussian regular on  $\mathcal{P}$ . Then it is Gaussian regular on every regular parametric submodel  $\mathcal{P}_0$ , too. So, it satisfies

$$\Sigma(P_0, \delta) \geq I^{-1}(P_0 | \beta, \gamma, \mathcal{P}_0), \quad (2.4)$$

for every regular parametric submodel  $\mathcal{P}_0$ . In view of (2.4) it is natural to define the information bound for estimating  $(\beta', \gamma)'$  in the semiparametric model by

$$I^{-1}(P_0 | \beta, \gamma, \mathcal{P}) = \sup\{I^{-1}(P_0 | \beta, \gamma, \mathcal{P}_0) : \mathcal{P}_0 \in \mathcal{C}\}. \quad (2.5)$$

A method of calculating  $I(P_0 | \beta, \gamma, \mathcal{P})$  can be found in Bickel *et al.* (1993). The main tasks are to find the *tangent space* of  $\mathcal{P}_{\text{nu}} = \{P_{\beta_0, \gamma_0, \sigma^2, h, g} : \sigma^2 \in \mathbb{R}^+, \int h = 1, \int g = 1, h, g \geq 0\}$  at  $(\sigma_0^2, h_0, g_0)$ , and to calculate the orthogonal projection of the scores  $\ell_\beta$  and  $\ell_\gamma$  onto the tangent space. Let  $\ell_{\text{nu}}(\mathcal{P}_0) = (\ell_{\sigma^2}, \ell_{\eta_1}, \ell_{\eta_2})$  be the vector of scores for the nuisance parameters  $\sigma^2$ ,  $\eta_1$  and  $\eta_2$ . We introduce  $\mathcal{P}_0$  here to stress its dependence on the choice of parametric submodel  $\mathcal{P}_0$ . Then, the tangent space of  $\mathcal{P}_{\text{nu}}$  at  $(\sigma_0^2, h_0, g_0)$  is nothing else than the closed linear span of the union of  $[\ell_{\text{nu}}(\mathcal{P}_0)]$  as  $\mathcal{P}_0$  ranges over  $\mathcal{C}$ . Write the tangent space  $\dot{\mathcal{P}}_{\text{nu}}$ . Define  $\ell_\beta^* = \ell_\beta - \Pi(\ell_\beta | \dot{\mathcal{P}}_{\text{nu}})$  and  $\ell_\gamma^*$ , likewise. These are called the *efficient score functions*. Writing  $\ell^* = (\ell_\beta^*, \ell_\gamma^*)'$ , the information matrix in the semiparametric model is given by

$$I(P_0 | \beta, \gamma, \mathcal{P}) = E_{P_0} \ell^* \ell^{*'}.$$

In the discussion that follows we omit the subscript “0” in  $\beta_0$ ,  $\gamma_0$ ,  $\sigma_0^2$ ,  $h_0$  and  $g_0$  which has been used to indicate they are the true values and functions. Also, we suppress the subscript “ $P_0$ ” in  $E_{P_0}$ .

The following theorem exhibits  $\ell_\beta^*$  and  $\ell_\gamma^*$  for estimating  $\beta$  and  $\gamma$ . To state the theorem, let  $c_t \equiv c_t(\gamma) = \sum_{j=0}^{t-1} \gamma^j$  and  $\tilde{c} \equiv \tilde{c}(\gamma) = \sum_{t=1}^{r-1} c_t(\gamma)/r$ . Write  $X_{1t}^w \equiv X_{1t}^w(\gamma) = \sum_{j=0}^{t-1} \gamma^j X_{1,t-j}$  and  $\tilde{X}_1^w \equiv \tilde{X}_1^w(\gamma) = \sum_{t=1}^{r-1} X_{1t}^w(\gamma)/r$ . Similarly, let  $Z_{1t}^w \equiv Z_{1t}^w(\beta, \gamma) = \sum_{j=0}^{t-1} \gamma^j Z_{1,t-j}$  and  $\tilde{Z}_1^w \equiv \tilde{Z}_1^w(\beta, \gamma) = \sum_{t=1}^{r-1} Z_{1t}^w(\beta, \gamma)/r$ . Let  $\bar{X}_1 = \sum_{t=1}^r X_{1t}/r$ .

**Theorem 2.1** *The efficient score functions for estimating  $\beta$  and  $\gamma$  are given by*

$$\begin{aligned}\ell_\beta^* &= \sum_{t=1}^r (Z_{1t} - \bar{Z}_1) X_{1t} / \sigma^2 - \{w^{(1)}(\bar{Z}_1) / w(\bar{Z}_1)\} (\bar{X}_1 - E\bar{X}_1) \\ \ell_\gamma^* &= \sum_{t=1}^r (Z_{1t} - \bar{Z}_1) Y_{1,t-1} / \sigma^2 + \{\tilde{c} / (r-1)\} \sigma^2 \sum_{t=1}^r (Z_{1t} - \bar{Z}_1)^2 \\ &\quad - \{w^{(1)}(\bar{Z}_1) / w(\bar{Z}_1)\} \{\beta'(\tilde{X}_1^w - E\tilde{X}_1^w) + \tilde{Z}_1^w - \tilde{c}\bar{Z}_1\}\end{aligned}$$

where  $w^{(1)}$  denotes the first derivative of  $w$ .

The information matrix  $I(P_0 | \beta, \gamma, \mathcal{P})$  can be calculated by using Theorem 2.1. Let  $\Sigma_{\text{wt}n} = \sum_{t=1}^r E(X_{1t} - \bar{X}_1)(X_{1t} - \bar{X}_1)'$  and  $\Sigma_{\text{bt}n} = \text{var}(\bar{X}_1)$ . Define  $I_w = \int \{(w^{(1)}(z))^2 / w(z)\} dz$ . Then

$$E \ell_\beta^* \ell_\beta^{*'} = \sigma^{-2} \Sigma_{\text{wt}n} + I_w \Sigma_{\text{bt}n}. \quad (2.6)$$

Define  $\xi \equiv \xi(\gamma) = \sum_{t=1}^{r-1} \sum_{s=1}^{r-1} \sum_{j=1}^{t \wedge s-1} \gamma^{|t-s|+2j}$ . It can be shown from a lengthy and cumbersome calculation that

$$\begin{aligned}E \ell_\gamma^{*2} &= \beta' E \left\{ \sum_{t=1}^{r-1} (X_{1t}^w - \tilde{X}_1^w)(X_{1t}^w - \tilde{X}_1^w)' \right\} \beta / \sigma^2 + 2\beta' \sum_{t=1}^{r-1} (c_t - \tilde{c}) E(X_{1t}^w - \tilde{X}_1^w) E(\bar{Z}_1) / \sigma^2 \\ &\quad + \sum_{t=1}^{r-1} (c_t - \tilde{c})^2 E(\bar{Z}_1^2) / \sigma^2 + I_w \{(\xi - r\tilde{c})\sigma^2 / r^2 + \beta' \text{var}(\tilde{X}_1^w) \beta\} \\ &\quad + (1 - r^{-1}) \sum_{t=1}^{r-1} \sum_{j=0}^{t-1} \gamma^{2j} - \sum_{t=1}^{r-1} c_t^2 / r - 2\tilde{c}^2 / (r-1).\end{aligned} \quad (2.7)$$

$$\begin{aligned}E \ell_\beta^* \ell_\gamma^* &= \sigma^{-2} \left\{ \sum_{t=1}^{r-1} \beta' E X_{1t}^w (X_{1,t+1} - \bar{X}_1) + \sum_{t=1}^{r-1} c_t E(X_{1,t+1} - \bar{X}_1) E(\bar{Z}_1) \right\} \\ &\quad + I_w \beta' E(\tilde{X}_1^w - E\tilde{X}_1^w)(\bar{X}_1 - E\bar{X}_1).\end{aligned} \quad (2.8)$$

The information matrix is readily obtained from (2.6), (2.7) and (2.8).

## 2.2 Construction of efficient estimators

Let  $\theta = (\beta', \gamma)'$ . Write  $I = I(P_0 | \beta, \gamma, \mathcal{P})$ . We construct an estimator  $\hat{\theta}_n$  of  $\theta$  such that  $\sqrt{n}(\hat{\theta}_n - \theta)$  converges in distribution to  $\mathcal{N}(0, I^{-1})$ . Define  $Z_{it}$  as we define  $Z_{1t}$  but replacing the subscript “1” by “i”, i.e.  $Z_{it} \equiv Z_{it}(\theta) = Y_{it} - \gamma Y_{i,t-1} - \beta' X_{it}$ . Likewise, define  $\bar{Z}_i$ ,  $X_{it}^w$ ,  $\tilde{X}_i^w$ ,  $Z_{it}^w$  and  $\tilde{Z}_i^w$ . Replace the subscript “1” by “i” in the formula for  $\ell_\beta^*$  and  $\ell_\gamma^*$  given at Theorem 2.1, and denote them by  $\ell_{\beta,i}^*$  and  $\ell_{\gamma,i}^*$ , respectively. Define  $\ell_i^* = (\ell_{\beta,i}^{*'}, \ell_{\gamma,i}^*)'$ . Instead of writing just  $\ell_i^*$  we will write  $\ell_i^*(\theta)$  to stress its dependence on  $\theta$  and for notational convenience in description of the efficient estimator given below. Note particularly that  $\ell_i^*(\theta)$  depends on

other parameters  $\sigma^2$ ,  $h$  and  $g$ , too. Efficient estimators  $\hat{\theta}_n$  are characterized by the following stochastic expansion:

$$\hat{\theta}_n = \theta + n^{-1}I^{-1} \sum_{i=1}^n \ell_i^*(\theta) + o_p(n^{-1/2}). \quad (2.9)$$

We follow the usual one-step procedure for constructing an efficient estimator: (i) Find a  $\sqrt{n}$ -consistent estimator  $\tilde{\theta}_n$  of  $\theta$ . (ii) Assuming the true parameter value  $\theta$  is known, find a reasonable estimator of  $\sigma^2$ , and using this construct an estimator of the density function  $w(\cdot)$ . (iii) Substitute the estimators obtained at (ii) into  $\ell_i^*(\theta)$ , and call it  $\hat{\ell}_i^*(\theta)$ . Also, construct an estimator of  $I$  using the estimators obtained at (ii), and denote it by  $\hat{I}(\theta)$ . (iv) Construct  $\hat{\theta}_n$  by  $\hat{\theta}_n = \tilde{\theta}_n + n^{-1}\hat{I}^{-1}(\tilde{\theta}_n) \sum_{i=1}^n \hat{\ell}_i^*(\tilde{\theta}_n)$ .

First, we construct an initial estimator of  $\theta$  which is  $\sqrt{n}$ -consistent. We take, as an initial estimator  $\tilde{\theta}_n$ , the minimizer of  $\sum_{i=1}^n \sum_{t=1}^r \left\{ (Y_{it} - \bar{Y}_i) - \gamma(Y_{i,t-1} - \bar{Y}_i) - \beta'(X_{it} - \bar{X}_i) \right\}^2$  with respect to  $\beta$  and  $\gamma$  where  $\bar{X}_i = \sum_{t=1}^r X_{it}/r$  and  $\bar{Y}_i = \sum_{t=1}^r Y_{it}/r$ . Write  $v_{it} = (X'_{it}, Y_{i,t-1})'$ ,  $m = \sum_{i=1}^n \sum_{t=1}^r v_{it} Y_{it}$  and  $\mathcal{M} = \sum_{i=1}^n \sum_{t=1}^r v_{it} v'_{it}$ . Then, the least squares initial estimator can be written as

$$\tilde{\theta}_n = \mathcal{M}^{-1}m. \quad (2.10)$$

It can be shown that  $\tilde{\theta}_n$  is  $\sqrt{n}$ -consistent.

Given the true value  $\theta$ , we define  $\tilde{\sigma}_n^2(\theta)$  by

$$\tilde{\sigma}_n^2(\theta) = \sum_{i=1}^n \sum_{t=1}^r \left\{ (Y_{it} - \bar{Y}_i) - \gamma(Y_{i,t-1} - \bar{Y}_i) - \beta'(X_{it} - \bar{X}_i) \right\}^2 / n(r-1).$$

Next, we construct a density estimator  $\hat{w}(\cdot; \theta)$ . Recalling that  $w$  is the density of  $\bar{Z}_i(\theta)$ , we estimate it by a kernel estimator

$$\hat{w}(z; \theta) = n^{-1} \sum_{i=1}^n K_{b_n}(z - \bar{Z}_i(\theta)) + c_n$$

where  $K_{b_n}(u) = (1/b_n)K(u/b_n)$ ,  $K(u) = e^{-u}(1 + e^{-u})^{-2}$  and  $b_n$  is a constant converging to zero at an appropriate rate to be described later. The constant  $c_n$  is introduced here to avoid technical difficulties due to zero denominators arising otherwise, and is taken to converge to zero too as  $n$  tends to infinity, whose rate is also to be specified below.

Now, define  $\hat{\ell}_i^*(\theta) = (\hat{\ell}_{\beta,i}^*(\theta), \hat{\ell}_{\gamma,i}^*(\theta))'$  where

$$\begin{aligned} \hat{\ell}_{\beta,i}^*(\theta) &= \sum_{t=1}^r \left\{ Z_{it}(\theta) - \bar{Z}_i(\theta) \right\} X_{it} / \tilde{\sigma}_n^2(\theta) \\ &\quad - \left\{ \hat{w}^{(1)}(\bar{Z}_i(\theta); \theta) / \hat{w}(\bar{Z}_i(\theta); \theta) \right\} (\bar{X}_i - \sum_{i=1}^n \bar{X}_i / n) \\ \hat{\ell}_{\gamma,i}^*(\theta) &= \sum_{t=1}^r \left\{ Z_{it}(\theta) - \bar{Z}_i(\theta) \right\} Y_{1,t-1} / \tilde{\sigma}_n^2(\theta) + \left\{ \tilde{c}(\gamma) / (r-1) \tilde{\sigma}_n^2(\theta) \right\} \sum_{t=1}^r \left\{ Z_{it}(\theta) - \bar{Z}_i(\theta) \right\}^2 \end{aligned} \quad (2.11)$$

$$\begin{aligned}
& - \left\{ \hat{w}^{(1)}(\bar{Z}_i(\theta); \theta) / \hat{w}(\bar{Z}_i(\theta); \theta) \right\} \left\{ \beta' \left( \tilde{X}_i^w(\gamma) - \sum_{i=1}^n \tilde{X}_i^w(\gamma) / n \right) \right. \\
& \quad \left. + \tilde{Z}_i^w(\theta) - \tilde{c}(\gamma) \bar{Z}_i(\theta) \right\}. \tag{2.12}
\end{aligned}$$

One may estimator  $I$  by  $n^{-1} \sum_{i=1}^n \hat{\ell}_i^*(\theta) \hat{\ell}_i^{*'}(\theta)$ , or by substituting the unknown quantities, except  $\theta$ , in the expressions given at (2.6), (2.7) and (2.8). It is well known that the latter approach yields more stable estimators, and so we proceed in that direction here. Denote by  $I_{11}$  the  $d \times d$  left-top block of the information matrix  $I$ , and by  $I_{12}$  and  $I_{22}$ , the  $d \times 1$  right-top and  $1 \times 1$  right-bottom blocks, respectively. Let  $\hat{\Sigma}_{\text{wtb}} = n^{-1} \sum_{i=1}^n \sum_{t=1}^r (X_{it} - \bar{X}_i)(X_{it} - \bar{X}_i)'$  and  $\hat{\Sigma}_{\text{btb}} = n^{-1} \sum_{i=1}^n \{ \bar{X}_i - \sum_{i=1}^n \bar{X}_i / n \} \{ \bar{X}_i - \sum_{i=1}^n \bar{X}_i / n \}'$ . Let  $\hat{I}_w(\theta) = n^{-1} \sum_{i=1}^n \left\{ \hat{w}^{(1)}(\bar{Z}_i(\theta); \theta) / \hat{w}(\bar{Z}_i(\theta); \theta) \right\}^2$ . Define

$$\hat{I}_{11}(\theta) = \tilde{\sigma}^{-2}(\theta) \hat{\Sigma}_{\text{wtb}} + \hat{I}_w(\theta) \hat{\Sigma}_{\text{btb}}.$$

We estimate  $I_{12}$  by

$$\begin{aligned}
\hat{I}_{12}(\theta) &= \tilde{\sigma}^{-2}(\theta) \left\{ \beta' n^{-1} \sum_{i=1}^n \sum_{t=1}^{r-1} X_{it}^w(\gamma) (X_{i,t+1} - \bar{X}_i) \right. \\
&\quad \left. + \left( n^{-1} \sum_{i=1}^n \sum_{t=1}^{r-1} c_t(\gamma) (X_{i,t+1} - \bar{X}_i) \right) \left( n^{-1} \sum_{i=1}^n \bar{Z}_i(\theta) \right) \right\} \\
&\quad + \beta' \hat{I}_w(\theta) n^{-1} \sum_{i=1}^n \left\{ \tilde{X}_i^w(\gamma) - n^{-1} \sum_{i=1}^n \tilde{X}_i^w(\gamma) \right\} \left\{ \bar{X}_i - n^{-1} \sum_{i=1}^n \bar{X}_i \right\}.
\end{aligned}$$

Finally, given the true value of  $\theta$ , we construct an estimator of  $I_{22}$  by

$$\begin{aligned}
\hat{I}_{22}(\theta) &= \tilde{\sigma}_n^{-2}(\theta) \beta' \left\{ n^{-1} \sum_{i=1}^n \sum_{t=1}^{r-1} (X_{it}^w(\gamma) - \tilde{X}_i^w(\gamma)) (X_{it}^w(\gamma) - \tilde{X}_i^w(\gamma))' \right\} \beta \\
&\quad + 2 \tilde{\sigma}_n^{-2}(\theta) \beta' \left\{ n^{-1} \sum_{i=1}^n \sum_{t=1}^{r-1} (c_t(\gamma) - \tilde{c}(\gamma)) (X_{it}^w(\gamma) - \tilde{X}_i^w(\gamma)) \right\} \left\{ n^{-1} \sum_{i=1}^n \bar{Z}_i(\theta) \right\} \\
&\quad + \tilde{\sigma}_n^{-2}(\theta) \left\{ n^{-1} \sum_{i=1}^n \sum_{t=1}^{r-1} (c_t(\gamma) - \tilde{c}(\gamma))^2 \bar{Z}_i^2(\theta) \right\} \\
&\quad + \hat{I}_w(\theta) \left\{ (\xi(\gamma) - r \tilde{c}(\gamma)) \tilde{\sigma}_n^2 / r^2 \right. \\
&\quad \left. + \beta' n^{-1} \sum_{i=1}^n (\tilde{X}_i^w(\gamma) - n^{-1} \sum_{i=1}^n \tilde{X}_i^w(\gamma)) (\tilde{X}_i^w(\gamma) - n^{-1} \sum_{i=1}^n \tilde{X}_i^w(\gamma))' \beta \right\} \\
&\quad + (1 - r^{-1}) \sum_{t=1}^{r-1} \sum_{j=0}^{t-1} \gamma^{2j} - \sum_{t=1}^{r-1} c_t^2(\gamma) / r - 2 \tilde{c}^2(\gamma) / (r - 1).
\end{aligned}$$

Plugging the initial estimator  $\tilde{\theta}_n$  into  $\hat{\ell}_i^*(\theta)$  and  $\hat{I}(\theta)$ , we obtain the following estimator of  $\theta$ :

$$\hat{\theta}_n = \tilde{\theta}_n + n^{-1} \hat{I}^{-1}(\tilde{\theta}_n) \sum_{i=1}^n \hat{\ell}_i^*(\tilde{\theta}_n) \tag{2.13}$$

The following theorem demonstrates that the estimator defined at (2.13) is a semiparametric efficient estimator of  $\theta$ .

**Theorem 2.2** *Assume that  $E(e^{t\bar{X}_1}) < \infty$  for some  $t > 0$  and that  $\int |u|^2 h(u) du < \infty$ . If  $b_n \rightarrow 0, c_n \rightarrow 0$  and  $nc_n^2 b_n^6 \rightarrow \infty$  as  $n \rightarrow \infty$ , then*

$$\sqrt{n}(\hat{\theta}_n - \theta) \rightarrow \mathcal{N}(0, I^{-1})$$

*in distribution as  $n$  tends to infinity.*

## 3 Empirical Illustration

### 3.1 Data

The primary data set used is the Department of Transportation's Origin and Destination Data Bank 1 A (DB1A) which consists of a continuous one in ten sample of all tickets sold in the United States. The data is collected on a quarterly basis and the sample runs from 1979I to 1992IV.

This study considers a market (route) to be a trip between origin and destination cities. Considering a route to be an airport pair would suggest that Southwest does not compete with United and American for flights between Houston and Chicago.

There are a number of factors for which controls other than standard demand variables such as own price, price of competitors, income, etc. are necessary in order to model the dynamic demand for airline travel. These are measured imperfectly. We imbed a number of these in the construction of the price index itself, following the methods outlined by Good et al. (2001). The factors can be categorized into five broad groups: Route specific effects, ticket restrictions, yield management, zero coupon tickets and network effects.

#### 3.1.1 Route Specific factors

There are clearly other variables which many have attempted to incorporate into modeling the demand side of long distance travel. These include factors which are weather related, such as mean temperature difference, in an attempt to capture vacation travel in the winter months. Others have included variables which attempt to capture the demand for business travel such as the number of white collar jobs in an area. We assume that these factors are either very slow to change or that they are strongly correlated with other factors in the model. These slowly moving factors are captured with fixed route specific effects which describe the origin-destination pair.



### **3.1.2 Ticket restrictions**

A major feature of airline fare structures is ticket restrictions. These either increase the risk of travel for consumers (non-refundability) or provide the airlines with improved predictability about demand (advanced booking) and enhance their ability to provide price discrimination information by separating price sensitive consumers from business travelers with more inelastic demands (Saturday night stay-overs). The major liability of using of DOT's DB1A as the primary source of ticket information is that it includes very incomplete information on ticket restrictions. There is typically a lag between fare type innovations and the way they are reported in DB1A. This makes it difficult to identify a consistent set of conditions under which service was accepted.

### **3.1.3 Yield management**

There is a great deal of competition in published fares. It is not at all uncommon for different airlines providing service on the same route to offer similar fare classes (sets of fare restrictions) at an identical price. However, fare structures may not correspond to published fares, in part due to yield management practices. We attempt to capture the effect of yield management by controlling for the percentage of first class, first class restricted, and coach restricted tickets.

### **3.1.4 Zero coupon tickets**

Frequent flyer miles were introduced in the mid 1980's. The practice has proven so successful that it has proliferated to other industries, even grocery stores offer discounts for frequent shoppers. To control for the effects of zero coupon tickets markups above marginal cost, the percentage of zero coupon tickets sold by the carrier for a particular route is controlled for in the construction of the price index.

### **3.1.5 Network Configuration**

Much has been made out of changes in airline networks by increased use of hub-and-spoke type networks. Airlines find these network configurations useful because they allow for higher passenger densities on individual routes.

Indirect routing of passengers clearly benefits the airlines because they can provide travel to passengers with fewer flights, potentially taking advantage of economies of equipment size (larger aircraft tend to have lower costs per passenger mile) and higher load factors (filling otherwise empty seats on an aircraft cost the airline very little).

Many of the different network characteristics can be measured at the individual ticket level. The DB1A database allows identification of many of the characteristics of the trip. Most fundamentally, the origin of a trip can be identified as well as the ultimate destination as indicated by a trip break. Approximately 95% of trips are either one way or round trip (depending on the year) with a small number of multi-break tickets involving as many as 23 different flights. More complex routings tend to be slightly more prevalent in later years than in earlier ones. In order to gain an understanding of the bulk of trips, attention is limited to either one way or round trip tickets which are weighted by travel distance. Information from the DB1A also allows measurement of the number of segments in a ticket. To control for the effect of the number of segments in the itinerary, we also control for the percentage of tickets with any number of stops up to 5 stops.

The minimum number of segments for a one way ticket is one. By 1984, this number fell to 25%. A very different pattern emerges for round trip tickets which have a minimum of two segments. In 1979-1 the average number of segments was 2.8, this increased somewhat to 3.05 by 1992-4. At 3.0 it suggests that approximately half of the itineraries involved a change of planes on the outbound and inbound portions of the trip. The rationale behind the difference in the one way and round trip ticket patterns is not clear. It may suggest a correlation between one way and full fare tickets which have a higher quality of demanded service for the large premium in price. On the other hand, while the presumption behind round trip tickets is that they describe the full trip, that is not the case for one way tickets since the passenger will require, at the minimum, an additional ticket for the return flight. Consequently the presumption that a full fare ticket involves the ultimate destination seems less well founded.

## 4 Results and Concluding Remarks

We have used these data to analyze upwards of 500 city-pair demands for a selected number of air carriers in the United States over periods of about 60 quarters. The cross-sectional unit is the airline market, that is, the city-pair. Our preliminary results for the semiparametric efficient dynamic panel estimators are presented in the tables below. The variables determining the demand for air travel,  $\ln(demand)$ , are  $\ln(demand)_{-1}$ ,  $\ln(own\ price)$ ,  $\ln(price\ of\ competitors)$ ,  $\ln(percapita\ income)$ ,  $\ln(population)$ ,  $\ln(unemployment\ rate)$ . City-pair characteristics as well as those portions of demand characteristics that are stable over the time series are modeled random effects. The carriers are American, Continental, Delta, and U. S. Airways. Our results are reasonable with little evidence that the roots of the dynamic equation are unstable. Demand appears to adjust sluggishly and markets are contestable.

Moreover, the U. S. airline markets (city-pairs) appear to evidence short run unit elasticities and are highly competitive in the long-run.

# Appendix

## A.1 Proof of Theorem 2.1.

The score functions are given by

$$\begin{aligned}\ell_\beta &= \sum_{t=1}^r (Z_{1t} - \bar{Z}_1) X_{1t} / \sigma^2 - \{w^{(1)}(\bar{Z}_1) / w(\bar{Z}_1)\} \bar{X}_1 \\ \ell_\gamma &= \sum_{t=2}^r (Z_{1t} - \bar{Z}_1) Y_{1,t-1} / \sigma^2 - \{w^{(1)}(\bar{Z}_1) / w(\bar{Z}_1)\} \tilde{Y}_1 \\ \ell_{\sigma^2} &= (2\sigma^2)^{-1} \left\{ \sum_{t=1}^r (Z_{1t} - \bar{Z}_1)^2 / \sigma^2 \right. \\ &\quad \left. + \int (\bar{\sigma}^{-2} (\bar{Z}_1 - u)^2 - r) \bar{\sigma}^{-1} \phi((\bar{Z}_1 - u) / \bar{\sigma}) h(u) du / w(\bar{Z}_1) \right\}\end{aligned}$$

where  $\tilde{Y}_1 = \sum_{t=1}^{r-1} Y_{1t} / r$  and  $\phi(\cdot)$  is the density function of the standard normal distribution. The tangent space  $\dot{\mathcal{P}}_{\text{nu}}$  may be decomposed into  $V_1$ ,  $V_2$  and  $V_3$ , i.e.  $\dot{\mathcal{P}}_{\text{nu}} = V_1 + V_2 + V_3$ , where  $V_1 = [\ell_{\sigma^2}]$  and

$$V_2 = \{a(\bar{Z}_1) \in L_2(P_0) : Ea(\bar{Z}_1) = 0\}, \quad V_3 = \{b(X_1) \in L_2(P_0) : Eb(X_1) = 0\}.$$

The following lemma shows that  $\ell_\beta$  and  $\ell_\gamma$  are perpendicular to  $V_3$ .

**Lemma A.1**  $E(\ell_\beta | X_1) = 0$  and  $E(\ell_\gamma | X_1) = 0$ .

*Proof.* Note that  $\{Z_{1t} - \bar{Z}_1\}$ ,  $\bar{Z}_1$  and  $X_1$  are independent. Since  $E(Z_{1t} - \bar{Z}_1) = 0$  and  $E\{w^{(1)}(\bar{Z}_1) / w(\bar{Z}_1)\} = \int w^{(1)}(u) du = 0$ , we obtain  $E(\ell_\beta | X_1) = 0$ . Next, we prove the second part. We can write  $Y_{1t} = \beta' X_{1t}^w + c_t \alpha_1 + \sum_{j=0}^{t-1} \gamma^j \varepsilon_{1,t-j}$ . Thus

$$E(\ell_\gamma | X_1) = \sigma^{-2} \sum_{t=1}^{r-1} \sum_{j=0}^{t-1} \gamma^j E\{\varepsilon_{1,t-j} (\varepsilon_{1,t+1} - \bar{\varepsilon}_1)\} - E\{\tilde{Y}_1 w^{(1)}(\bar{Z}_1) / w(\bar{Z}_1) | X_1\}. \quad (\text{A.1})$$

The first term in (A.1) equals  $\sigma^{-2} \sum_{t=1}^{r-1} \sum_{j=0}^{t-1} \gamma^j (-\sigma^2 / r) = -\tilde{c}$ . For the second term, note  $\tilde{Y}_1 = \beta' \tilde{X}_1^w + \tilde{c} \bar{Z}_1 + r^{-1} \sum_{j=0}^{t-1} \gamma^j (Z_{1,t-j} - \bar{Z}_1)$ . By this and the facts that  $E\{w^{(1)}(\bar{Z}_1) / w(\bar{Z}_1)\} = 0$  and that  $E(Z_{1,t-j} - \bar{Z}_1) = 0$ , the second term equals

$$\tilde{c} E\{\bar{Z}_1 w^{(1)}(\bar{Z}_1) / w(\bar{Z}_1)\} = \tilde{c} \int u w^{(1)}(u) du = -\tilde{c}. \quad (\text{q.e.d.})$$

Lemma A.1 implies that writing  $W = [\ell_{\sigma^2} - \Pi(\ell_{\sigma^2} | V_2)]$

$$\begin{aligned}\ell_\beta^* &= \ell_\beta - \Pi(\ell_\beta | V_2) - \Pi\{\ell_\beta - \Pi(\ell_\beta | V_2) | W\} \\ \ell_\gamma^* &= \ell_\gamma - \Pi(\ell_\gamma | V_2) - \Pi\{\ell_\gamma - \Pi(\ell_\gamma | V_2) | W\}.\end{aligned}$$

We compute  $\ell_\beta^*$  first. Note that  $\Pi(\ell_\beta | V_2) = E(\ell_\beta | V_2) = -E(\bar{X}_1)w^{(1)}(\bar{Z}_1)/w(\bar{Z}_1)$ . Thus

$$\ell_\beta - \Pi(\ell_\beta | V_2) = \sigma^{-2} \sum_{t=1}^r (Z_{1t} - \bar{Z}_1) X_{1t} - \{w^{(1)}(\bar{Z}_1)/w(\bar{Z}_1)\}(\bar{X}_1 - E\bar{X}_1). \quad (\text{A.2})$$

Since  $E \sum_{t=1}^r (Z_{1t} - \bar{Z}_1)^2 = (r-1)\sigma^2$ , we obtain

$$\ell_{\sigma^2} - \Pi(\ell_{\sigma^2} | V_2) = (2\sigma^4)^{-1} \left\{ \sum_{t=1}^r (Z_{1t} - \bar{Z}_1)^2 - (r-1)\sigma^2 \right\}. \quad (\text{A.3})$$

Now by symmetry of the distribution of  $(Z_{1t} - \bar{Z}_1)$  and by independence of  $Z_{1t} - \bar{Z}_1$ ,  $\bar{Z}_1$  and  $X_1$ , it follows that  $E\{\sum_{t=1}^r (Z_{1t} - \bar{Z}_1) X_{1t}\} \{\sum_{t=1}^r (Z_{1t} - \bar{Z}_1)^2\} = 0$  and

$$E\{w^{(1)}(\bar{Z}_1)/w(\bar{Z}_1)\}(\bar{X}_1 - E\bar{X}_1) \left\{ \sum_{t=1}^r (Z_{1t} - \bar{Z}_1)^2 - (r-1)\sigma^2 \right\} = 0.$$

Thus,  $\ell_\beta - \Pi(\ell_\beta | V_2)$  is perpendicular to  $\ell_{\sigma^2} - \Pi(\ell_{\sigma^2} | V_2)$ , which implies that  $\ell_\beta^* = \ell_\beta - \Pi(\ell_\beta | V_2)$ . The formula for  $\ell_\beta^*$  follows from (A.2).

Next, we compute  $\ell_\gamma^*$ . By independence of  $Z_{1t} - \bar{Z}_1$ ,  $\bar{Z}_1$  and  $X_1$ , we have

$$\begin{aligned} E(\ell_\gamma | V_2) &= \sigma^{-2} \sum_{t=1}^{r-1} \sum_{j=0}^{t-1} \gamma^j E(\varepsilon_{1,t-j} - \bar{\varepsilon}_1)(\varepsilon_{1,t+1} - \bar{\varepsilon}_1) - \{w^{(1)}(\bar{Z}_1)/w(\bar{Z}_1)\} \left\{ \beta' E(\tilde{X}_1^w) + \tilde{c}\bar{Z}_1 \right\} \\ &= -\tilde{c} - \{w^{(1)}(\bar{Z}_1)/w(\bar{Z}_1)\} \left\{ \beta' E(\tilde{X}_1^w) + \tilde{c}\bar{Z}_1 \right\}. \end{aligned}$$

Thus, we obtain

$$\begin{aligned} \ell_\gamma - \Pi(\ell_\gamma | V_2) &= \sigma^{-2} \sum_{t=2}^r (Z_{1t} - \bar{Z}_1) Y_{1,t-1} + \tilde{c} \\ &\quad - \{w^{(1)}(\bar{Z}_1)/w(\bar{Z}_1)\} \left\{ \beta' (\tilde{X}_1^w - E\tilde{X}_1^w) + \sum_{t=1}^{r-1} \sum_{j=0}^{t-1} \gamma^j (Z_{1,t-j} - \bar{Z}_1)/r \right\}. \end{aligned}$$

To calculate  $\Pi(\ell_\gamma - \Pi(\ell_\gamma | V_2) | W)$ , we find

$$E\{\ell_\gamma - \Pi(\ell_\gamma | V_2)\} \{\ell_{\sigma^2} - \Pi(\ell_{\sigma^2} | V_2)\} = -\tilde{c}/\sigma^2 \quad (\text{A.4})$$

$$E\{\ell_{\sigma^2} - \Pi(\ell_{\sigma^2} | V_2)\}^2 = (r-1)/(2\sigma^4). \quad (\text{A.5})$$

Denote the left hand sides of (A.4) and (A.5) by  $\zeta_{12}$  and  $\zeta_{22}$ , respectively. Then from (A.3), (A.4) and (A.5), we obtain

$$\begin{aligned} \Pi(\ell_\gamma - \Pi(\ell_\gamma | V_2) | W) &= (\zeta_{12}/\zeta_{22}) \{\ell_{\sigma^2} - \Pi(\ell_{\sigma^2} | V_2)\} \\ &= -\{\tilde{c}/(r-1)\sigma^2\} \sum_{t=1}^r (Z_{1t} - \bar{Z}_1)^2 + \tilde{c}, \end{aligned}$$

which leads to the formula for  $\ell_\gamma^*$ .

## A.2 Proof of Theorem 2.2

Define

$$w_n(z) \equiv w_n(z; \sigma^2) \equiv w_n(z; \sigma^2, h) = K_{b_n} * w(z; \sigma^2, h) + c_n$$

where  $*$  denotes the convolution. We write  $\hat{r}$ ,  $r_n$  and  $r$  for  $\hat{w}^{(1)}/\hat{w}$ ,  $w_n^{(1)}/w_n$  and  $w^{(1)}/w$ , respectively. Define  $I_{w,n} = \int r_n^2(z)w(z) dz$ . Define  $I_n$  as in the definition of the information matrix  $I$  but with  $I_w$  being replaced by  $I_{w,n}$ . Following the arguments for the proof of (B.9) in Park, Sickles and Simar (1998), one can verify

$$E \{r_n(\bar{Z}_1(\theta)) - r(\bar{Z}_1(\theta))\}^2 \rightarrow 0. \quad (\text{A.6})$$

It follows from (A.6) that  $I_n \rightarrow I$  as  $n$  tends to infinity.

Now, it may be proved that

$$\left( n^{-1} \sum_{i=1}^n \bar{X}_i - E\bar{X}_1 \right) n^{-1/2} \sum_{i=1}^n r_n(\bar{Z}_i(\theta)) \rightarrow 0 \quad (\text{A.7})$$

$$n^{-1/2} \sum_{i=1}^n (\bar{X}_i - E\bar{X}_1) \{r_n(\bar{Z}_i(\theta)) - r(\bar{Z}_i(\theta))\} \rightarrow 0, \quad (\text{A.8})$$

both in the sense of convergence in probability. They follow since the left hand sides of (A.7) and (A.8) have zero means by independence of  $\bar{X}_i$  and  $\bar{Z}_i(\theta)$ , and variances bounded by  $n^{-1} \text{var}(\bar{X}_1) I_{w,n}$  and  $\text{var}(\bar{X}_1) E \{r_n(\bar{Z}_1(\theta)) - r(\bar{Z}_1(\theta))\}^2$ , respectively, both of which converge to zero as  $n$  tends to infinity. Similarly, it can be shown that

$$\left( n^{-1} \sum_{i=1}^n \tilde{X}_i^w(\gamma) - E\tilde{X}_1^w(\gamma) \right) n^{-1/2} \sum_{i=1}^n r_n(\bar{Z}_i(\theta)) \rightarrow 0 \quad (\text{A.9})$$

$$n^{-1/2} \sum_{i=1}^n (\tilde{X}_i^w(\gamma) - E\tilde{X}_1^w(\gamma)) \{r_n(\bar{Z}_i(\theta)) - r(\bar{Z}_i(\theta))\} \rightarrow 0, \quad (\text{A.10})$$

both in the sense of convergence in probability.

Define  $\check{\ell}_{\beta,i}^*(\theta)$  and  $\check{\ell}_{\gamma,i}^*(\theta)$  as in the definitions of  $\hat{\ell}_{\beta,i}^*(\theta)$  and  $\hat{\ell}_{\gamma,i}^*(\theta)$  at (2.11) and (2.12), respectively, with  $\hat{w}(\bar{Z}_i(\theta); \theta)$  being replaced by  $w_n(\bar{Z}_i(\theta); \sigma^2)$  and  $\tilde{\sigma}_n^2$  by  $\sigma^2$ , and let  $\check{\ell}_i^*(\theta) = (\check{\ell}_{\beta,i}^*(\theta), \check{\ell}_{\gamma,i}^*(\theta))'$ . Then, (A.7)  $\sim$  (A.10) imply

$$n^{-1/2} I_n^{-1} \sum_{i=1}^n \check{\ell}_i^*(\theta) \rightarrow \mathcal{N}(0, I^{-1}) \quad (\text{A.11})$$

in distribution as  $n$  tends to infinity. Now, it can be shown that as in the proofs of Lemma A.2 and (A.16) of Park and Simar (1994)

$$\hat{I}(\tilde{\theta}_n) - I_n \rightarrow 0, \quad (\text{A.12})$$

$$n^{-1/2} \sum_{i=1}^n \left( \bar{X}_i - n^{-1} \sum_{i=1}^n \bar{X}_i \right) \left\{ \hat{r}(\bar{Z}_i(\tilde{\theta}_n); \tilde{\theta}_n) - r_n(\bar{Z}_i(\tilde{\theta}_n); \sigma^2) \right\} \rightarrow 0, \quad (\text{A.13})$$

$$n^{-1/2} \sum_{i=1}^n \left( \tilde{X}_i^w(\tilde{\gamma}_n) - n^{-1} \sum_{i=1}^n \tilde{X}_i^w(\tilde{\gamma}_n) \right) \left\{ \hat{r}(\bar{Z}_i(\tilde{\theta}_n); \tilde{\theta}_n) - r_n(\bar{Z}_i(\tilde{\theta}_n); \sigma^2) \right\} \rightarrow 0, \quad (\text{A.14})$$

$$n^{-1/2} \sum_{i=1}^n \left( \tilde{Z}_i^w(\tilde{\theta}_n) - \tilde{c}(\tilde{\gamma}_n) \bar{Z}_i(\tilde{\theta}_n) \right) \left\{ \hat{r}(\bar{Z}_i(\tilde{\theta}_n); \tilde{\theta}_n) - r_n(\bar{Z}_i(\tilde{\theta}_n); \sigma^2) \right\} \rightarrow 0, \quad (\text{A.15})$$

all in the sense of convergence in probability. Since  $\tilde{\sigma}_n^2(\tilde{\theta}_n)$  converges to  $\sigma^2$  in probability, (A.13)  $\sim$  (A.15) imply

$$n^{-1/2} I_n^{-1} \sum_{i=1}^n \left\{ \hat{\ell}_i^*(\tilde{\theta}_n) - \check{\ell}_i^*(\tilde{\theta}_n) \right\} \rightarrow 0 \quad (\text{A.16})$$

in probability. The theorem follows then from (A.11), (A.12) and (A.16) since for any  $C > 0$

$$\sup \left\{ \left| n^{-1/2} \sum_{i=1}^n \left\{ \check{\ell}_i^*(\theta') - \check{\ell}_i^*(\theta) + I_n(\theta' - \theta) \right\} \right| : n^{1/2} |\theta' - \theta| \leq C \right\} \rightarrow 0$$

in probability which can be proved as in Park and Simar (1994).

## A References

### References

- [1] Alam, I. Semenick, and R. C. Sickles (2000), "A Time Series Analysis of Deregulatory Dynamics and Technical Efficiency: the Case of the U. S. Airline Industry," *International Economic Review*," 41, 2000, 203-218.
- [2] Bickel, P.J., Klaassen, C.A.J., Ritov, Y., and Wellner, J.A. (1993), *Efficient and Adaptive Estimation in Non- and Semi-parametric Models*. Baltimore: Johns Hopkins University Press.
- [3] Good, D., R. C. Sickles, and J. C. Weiher (2001) "On a New Hedonic Price Index for Airline Travel," mimeo, Rice University.
- [4] Ibragimov, I.A. and Has'minskii, R.Z. (1981), *Statistical Estimation: Asymptotic Theory*. Springer, New-York.
- [5] Park, B.U. and Simar, L. (1994), "Efficient Semiparametric Estimation in Stochastic Frontier Models," *Journal of the American Statistical Association* 89, 929-36.
- [6] Park, B.U., Sickles, R.C. and Simar, L (1998), "Stochastic Frontiers: A Semiparametric Approach," *Journal of Econometrics* 84, 273-301.
- [7] Park, B.U., Sickles, R.C. and Simar, L (2002), "Semiparametric Efficient Estimation of AR(1) Panel Data Models," mimeo, 2002



*Semiparametric Efficient Dynamic Airline Demand Equations*

**Table 1**  
**American Airlines**

Variable	<u>Coefficient</u>	<u>Standard Error</u>
ln(demand) <sub>-1</sub>	0.628	0.0111
ln(own price)	-0.523	0.0395
ln(comp. price)	0.016	0.0486
ln(pcapinc)	0.618	0.1060
ln(pop)	0.734	0.1011
ln(unrate)	-.0395	0.0190

R<sup>2</sup>: within = 0.6751  
between = 0.6597  
overall = 0.7993

**Table 2**  
**Continental Airlines**

Variable	<u>Coefficient</u>	<u>Standard Error</u>
ln(demand) <sub>-1</sub>	0.550	0.0124
ln(own price)	-1.104	0.0671
ln(comp. price)	0.216	0.1095
ln(pcapinc)	1.562	0.2841
ln(pop)	0.271	0.2197
ln(unrate)	-.0534	0.0452

R<sup>2</sup>: within = 0.6081  
Between= 0.8732  
overall = 0.8076

**Table 3**  
**Delta Airlines**

<u>Variable</u>	<u>Coefficient</u>	<u>Standard Error</u>
ln(demand) <sub>-1</sub>	0.642	0.0082
ln(own price)	-0.915	0.0440
ln(comp. price)	0.243	0.0483
ln(pcapinc)	0.339	0.1093
ln(pop)	0.637	0.1094
ln(unrate)	0.011	0.0206

R<sup>2</sup>: within = 0.7945  
 Between= 0.7746  
 overall = 0.7773

**Table 4**  
**U. S. Airlines**

<u>Variable</u>	<u>Coefficient</u>	<u>Standard Error</u>
ln(demand) <sub>-1</sub>	0.566	0.01367
ln(own price)	-0.557	0.06823
ln(comp. price)	-0.454	0.1016
ln(pcapinc)	-0.055	0.2408
ln(pop)	1.833	0.2218
ln(unrate)	-0.267	0.0414

R<sup>2</sup>: within = 0.5905  
 between= 0.2871  
 overall = 0.4417