ECONOMIC RESEARCH REPORTS

PRODUCTION FRONTIERS WITH CROSS-SECTIONAL AND TIME-SERIES VARIATION IN EFFICIENCY LEVELS

by

Christopher Cornwell

Peter Schmidt

and

Robin C. Sickles

R.R. #89-18 July 1989

C. V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

PRODUCTION FRONTIERS WITH CROSS-SECTIONAL AND TIME-SERIES VARIATION IN EFFICIENCY LEVELS

Christopher Cornwell
University of Georgia

Peter Schmidt
Michigan State University

Robin C. Sickles
Rice University

November 1987

Revised September, 1988

Earlier versions of this paper were given at the 1986 Winter Meetings of the Econometric Society, the 1987 TIMS/ORSA Meetings, the 1987 American Statistical Association Meetings, and the National Bureau of Economic Research Conference on Productivity in the Service Sector, July 1987. Comments by Robert Gordon, Zvi Griliches, V. Kerry Smith, and M. Ishaq Nadiri strengthened the paper considerably. Schmidt and Sickles are grateful to the National Science foundation for its support. The authors also gratefully acknowledge support from the C.V. Starr Center for Applied Economics.

ABSTRACT

PRODUCTION FRONTIERS WITH CROSS-SECTION AND TIME-SERIES VARIATION IN EFFICIENCY LEVELS

In this paper we consider the efficient instrumental variables estimation of a panel data model in which coefficients in addition to the intercept vary over individuals, and we apply the methodology we develop to a model in which there is cross-sectional and temporal variation in productivity levels (or, equivalently, in levels of technical efficiency), using data on U.S. airlines. We relax the assumption that technical inefficiency is time invariant, but in such a way as to not lose the advantages of panel data. One of our more interesting empirical findings is that efficiency levels are increasing over the period 1970I-1981III and the empirical fact that appears to drive this increase is that the firms' productivity levels are becoming more similar over the period in which the industry was deregulated.

1. Introduction

In this paper we consider the efficient instrumental variables estimation of a panel data model in which coefficients in addition to the intercept vary over individuals, and we apply the methodology we develop to a model in which there is cross-sectional and temporal variation in productivity levels (or, equivalently, in levels of technical efficiency), using data on U.S. airlines. We therefore extend the current literatures on panel data, productivity measurement, and frontier production functions.

The early literature on stochastic frontier production functions (e.g., Aigner, Lovell and Schmidt (1977)) assumed the existence of data on a single cross-section of firms, and the separation of technical inefficiency from random noise required strong assumptions about their distributions. More recently, Schmidt and Sickles (1984) considered the case in which panel data are available. In their model only the intercept varied over firms; differences in the intercept were interpreted as differing efficiency levels, with the level of efficiency for each firm assumed to be time-invariant. The Schmidt and Sickles model does not require strong distributional assumptions about technical inefficiency or random noise, nor is the assumption of independence between technical inefficiency and the explanatory variables (inputs) needed. However, the assumption that technical inefficiency is time-invariant is very strong, and depending on the data, may prove unrealistic.

In this paper we seek to relax the assumption that technical inefficiency is time invariant, but in such a way as to not lose the advantages of panel data. We do so by introducing into the production function a flexible (e.g. quadratic) function of time, with coefficients varying over firms. This function can be thought of as representing productivity growth, at a rate that varies over firms, and it implies that levels of inefficiency for each firm vary over time. This model is similar to the model of Sickles, Good and Johnson (1986), who considered the measurement of efficiency growth using a profit function which included a flexible function of time, but assumed that efficiency growth was the same for all firms. Our model generalizes their treatment by allowing for cross-sectional variation in productivity growth rates. However, the model still imposes enough structure on the way in which productivity levels change over time that strong distributional assumptions are avoided.

Previous treatments of the linear model with panel data, such as Hausman and Taylor (1981) and Amemiya and MaCurdy (1986), have dealt with the case in which only the intercept varies across individuals (firms). We extend the analysis of Hausman and Taylor to the above model in which there is cross-sectional heterogeneity in slopes as well as (or instead of) intercepts. This case has previously been treated in the random coefficients literature (for example, see Swamy (1971, 1974)), but under the assumption that the variation in coefficients is independent of the regressors; like Hausman and Taylor, we allow some or all of the regressors to be correlated with the cross-sectional variation in coefficients.

The plan of the paper is as follows. Section 2 presents some statistical background on panel data models with heterogeneous intercepts. It also fixes notational conventions used in subsequent sections and in the statistical appendices. Section 3 introduces our more general model in which variables other than the intercept can vary across individuals (firms). Section 4 provides our empirical model, and Section 5 gives our empirical results for productivity growth and efficiency for U.S. airlines. Section 6 concludes.

Proofs of theorems and other statistical details are given in Appendices A and B.

2. Background

A standard form for econometric models that assume the availability of panel data is

$$(2.0) \qquad y_{it} = X_{it}'\beta + Z_i'\gamma + \alpha_i + \epsilon_{it}, \qquad i = 1,\ldots,N;\ t = 1,\ldots,T.$$

where X_{it} is a K-dimensional vector of time-varying explanatory variables, Z_i is a J-dimensional vector of time-invariant explanatory variables, and β and γ are conformably dimensioned parameter vectors. For the purpose of discussion we can think of the data set being comprised of N individuals (firms) and T time periods per individual. The disturbances ϵ_{it} are assumed uncorrelated with X_{it}, Z_i, and α_i, and iid with a zero mean and variance σ^2. The α_i are individual effects, taken to be random variables and iid with a zero mean and variance σ^2_α.

Various estimators of the model (2.0) exist, and the choice between them depends primarily on whether the individual effects (α_i) are correlated with the explanatory variables (X_{it} and Z_i). For example, when the effects are correlated with some of the regressors, ordinary least squares (OLS) and generalized least squares (GLS) are biased and inconsistent. The usual solution to this problem is to difference away the effects by transforming the data into deviations from individual means and then to apply least squares to the transformed data. Commonly known as "within" estimation, this procedure has two significant drawbacks. First, γ cannot be estimated since all time-invariant explanatory variables are eliminated by this transformation. Secondly, the within estimator is not fully efficient since it ignores "between" (across individuals) variation.

An innovative response to the defects of the within estimator is provided by Hausman and Taylor (1981), henceforth H-T. Taking an instrumental variables (IV) approach, they exploit assumptions about explanatory variables that are uncorrelated with the effects to derive a simple consistent estimator and an asymptotically efficient estimator. The extent to which their estimators represent an improvement over the within estimator depends on the number of exogeneity restrictions one is willing to impose. Noting that the within estimator of β always exists and is consistent, H-T use it as a basis of comparison, presenting clear conditions under which their IV estimators are different.

In order to review the results that we will extend to our more general model, we need to define some notation. In particular, we will use the following notation for projections. For any matrix A, let P_A be the projection onto the column space of A. If A is of full column rank, therefore,

$$(2.1) \qquad P_A = A(A'A)^{-1}A.$$

Similarly, let $M_A = I-P_A$ be the projection on the null space of A.

Consider the model (2.0) in matrix form:

$$(2.2) \qquad y = X\beta + Z\gamma + v$$

$$v = D\alpha + \epsilon,$$

where y and ϵ are NTx1, X is NTxK, Z is NTxJ, and D is an NTxN matrix of dummy variables ($D = I_N \otimes e_T$, where e_T is a T x 1 vector of ones.) We define three standard estimators of this model.

First, the within estimator is:

$$(2.3) \quad \hat{\beta}_W = (X'M_DX)^{-1} X'M_Dy.$$

Note that only β is estimated, since $M_DZ = 0$. The within estimator has several equivalent interpretations. It is an IV estimator with instruments equal to M_D. It is also an IV estimator with instruments equal to (M_DX) (the deviations from means of the time-varying variables). Alternatively, within is OLS after a transformation to deviations from means; that is, OLS of (M_Dy) on (M_DX). It is consistent (as either N or $T \rightarrow \infty$) without assumptions on the relationship between X and the effects.

Second, the GLS estimator is

$$(2.4) \quad \begin{bmatrix} \hat{\beta}_G \\ \hat{\gamma}_G \end{bmatrix} = [(X,Z)'\Omega^{-1}(X,Z)]^{-1}(X,Z)'\Omega^{-1}y$$

where $\Omega = \text{cov}(v)$. This is equivalent to OLS on

$$(2.5) \quad \Omega^{-1/2}y = \Omega^{-1/2}X\beta + \Omega^{-1/2}Z\gamma + \Omega^{-1/2}v$$

where

$$(2.6) \quad \Omega^{-1/2} = M_D + \theta P_D = I_{NT} - (1-\theta)P_D, \ \theta = [\sigma^2/(\sigma^2 + T\sigma_\alpha^2)]^{1/2}.$$

This is OLS after "(1-θ) differences"; e.g., $(\Omega^{-1/2}y)_{it} = y_{it} - (1-\theta)\,\bar{y}_i$. GLS is consistent as $N\to\infty$ provided that (X,Z) is uncorrelated with v (and hence with the effects $D\alpha$). For fixed T, it is more efficient than within.

Third, we consider the H-T estimators. We distinguish variables assumed to be correlated with the effects from those uncorrelated with the effects. Thus we partition X and Z as:

$$(2.7) \qquad X = [X_1, X_2], \; Z = [Z_1, Z_2],$$

where X_1 is $NTxk_1$, X_2 is $NTxk_2$, Z_1 is $NTxj_1$, and Z_2 is $NTxj_2$ (and $k_1 + k_2 = K_1$, $j_1 + j_2 = J$). Assume, for fixed T,

$$(2.8) \qquad \frac{1}{N} X_1' D\alpha \to 0, \qquad \frac{1}{N} Z_1' D\alpha \to 0$$

$$\frac{1}{N} X_2' D\alpha \to h_X \neq 0, \quad \frac{1}{N} Z_2' D\alpha \to h_Z \neq 0.$$

(Thus X_2 and Z_2 are correlated with the effects; X_1 and Z_1 are not.)

The H-T "simple consistent" estimator proceeds in two steps. First, estimate β by within and calculate the within residuals:

$$(2.9) \qquad (y - X\hat{\beta}_W) = Z\gamma + [v + X(\beta - \hat{\beta}_W)]$$

Then, define $\hat{\gamma}_W$ as the IV estimator of (2.9), using as instruments $B = (X_1, Z_1)$. This estimator is defined if $\text{rank}(B) \geq J$, which requires $k_1 \geq j_2$. It is consistent provided the exogeneity assumptions in (2.8) are correct, and that $Z'P_BZ$ is (asymptotically) of full rank.

We note for future reference that H-T actually define their simple estimator as IV, using instruments B, of (2.9) after it has been transformed by P_D. Similarly, we could transform (2.9) by $\Omega^{-1/2}$ and do IV using B as instruments. Neither of these transformations matters, in the sense that the same estimator results in each case. However, if we transform (2.9) by $\Omega^{-1/2}$ and then use $(\Omega^{-1/2}B)$ as instruments, we get an estimator that is different from the H-T simple estimator (except in the "exactly identified" case, $k_1 = j_2$, when it is still the same); see Breusch, Mizon and Schmidt (1987). Which estimator is more efficient depends on whether Z_2 is better explained by B or by $\Omega^{-1/2}B$. This depends on the "reduced form" with which one completes the system.

On the other hand, the H-T "efficient" estimator is IV applied to (2.5), using as instruments $A = (M_D, X_1, Z_1)$. We will denote the resulting estimates as $\hat{\beta}^*$, $\hat{\gamma}^*$. Although H-T do not point it out, the same estimator would result if we used as instruments $A^* = (M_D X, X_1, Z_1)$, or $(\Omega^{-1/2}A)$, or $(\Omega^{-1/2}A^*)$. The estimator is consistent as $N\to\infty$ provided that the exogeneity assumptions are correct and $[(X,Z)'\ P_A\ (X,Z)]$ is (asymptotically) of full rank. The nonsingularity of $[(X,Z)'P_A(X,Z)]$ requires an identifiability condition which is discussed below.

The relationship between the H-T simple consistent and efficient estimators depends on the relationship between the dimensions of X_1 and Z_2. H-T distinguish three cases: under-identification, exact-identification, and over-identification. First, in the under-identified case $(k_1 < j_2)$, $\hat{\beta}^* = \hat{\beta}_W$ and $\hat{\gamma}^*$ does not exist. This is the case in which there are more regressors in (2.5) than instruments in A, so IV encounters a singularity. Second, in the case of exact-identification $(k_1 = j_2)$, $\hat{\beta}^* = \hat{\beta}_W$ and $\hat{\gamma}^* = \hat{\gamma}_W$. Finally, if the model is over-identified $(k_1 > j_2)$, $(\hat{\beta}^*, \hat{\gamma}^*) \neq (\hat{\beta}_W, \hat{\gamma}_W)$ and $(\hat{\beta}^*, \hat{\gamma}^*)$ is more efficient than $(\hat{\beta}_W, \hat{\gamma}_W)$. The efficiency gain in estimation of β occurs because of the use of the extra instruments in (X_1, Z_1) in addition to M_D (or $M_D X$), while the efficiency gain in estimation of γ occurs because of the use of a more efficient estimate of β.

3. The General Model

We now proceed to the more general model in which variables other than the intercept vary across individuals. Specifically, our model is:

$$(3.1) \qquad y_{it} = X_{it}'\beta + Z_{it}'\gamma + W_{it}'\delta_i + \epsilon_{it},$$

where evidently the variables in W have individually-varying coefficients. (In our particular application W_{it} will contain an intercept plus time and time squared.)

Let

$$(3.2) \qquad \delta_i = \delta_o + u_i,$$

where the u_i are assumed to be random variables with a zero mean and covariance matrix Δ. We can then write

$$y_{it} = X_{it}'\beta + Z_{it}'\gamma + W_{it}'\delta_o + v_{it},$$

$$(3.3) \qquad v_{it} = W_{it}'u_i + \epsilon_{it}.$$

The matrix form of (3.1) is

$$(3.4) \qquad y = X\beta + Z\gamma + Q\delta + \epsilon$$

while the matrix form of (3.3) is

$$y = X\beta + Z\gamma + W\delta_o + v,$$

(3.5) $$v = Qu + \epsilon$$

where W is NT x L (L being the dimension of W_{it})

(3.6) $$Q = \begin{bmatrix} W_1 & 0 & \cdot & \cdot & 0 \\ 0 & W_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & W_N \end{bmatrix}$$

is NT x NL, and δ and u are NL x 1 vectors containing δ_i (or u_i), i = 1, 2,...,N.

We assume $L \leq T$, so that Q is of full column rank. This is not necessary for identifiability of β. However, it is necessary for estimation of the individual δ_i. Also, if $L > T$, some of the matrices which we must invert would be singular. This is not really a substantive matter, since the projections involved are still well defined, but the algebra would become more complicated.

The distinction between X and Z is that we assume that $P_Q Z = Z$ (or, equivalently, $M_Q Z = 0$). In practice, Z will contain time-invariant variables. However, a variable also belongs in Z if its temporal variation, at the individual level, is perfectly explained by the temporal variation in W.

The generalization of the within estimator is OLS on (3.4). Since Z is collinear with Q, we can not estimate γ. The within estimator of β can be written as

$$(3.7) \quad \hat{\beta}_W = (X'M_QX)^{-1}X'M_Qy.$$

The within estimator is an IV estimator, with instruments M_Q (or, equivalently, M_QX). Its consistency does not depend on assumptions of uncorrelatedness of (X,Z) and (Qu).[2]

The GLS estimator of $(\beta, \gamma, \delta_o)$ is

$$(3.8) \quad [(X,Z,W)'\Omega^{-1}(X,Z,W)]^{-1} (X,Z,W)'\Omega^{-1}y.$$

where

$$(3.9) \quad \Omega = \text{cov}(v) = \sigma^2 I_{NT} + Q(I_N \otimes \Delta) Q'$$

While Ω is a large matrix, it is block diagonal, with blocks of the form $\sigma^2 I_N + W_i \Delta W_i'$; thus its inversion is practical.

Alternatively, GLS is OLS applied to the transformed equation

$$(3.10) \quad \Omega^{-1/2}y = \Omega^{-1/2}X\beta + \Omega^{-1/2}Z\gamma + \Omega^{-1/2}W\delta_o + \Omega^{-1/2}v.$$

This expression is not of much actual computational use, however, since $\Omega^{-1/2}$ is harder to calculate than Ω^{-1}; we have

$$(3.11) \quad \Omega^{-1/2} = \frac{1}{\sigma} M_Q + F$$

where

$$(3.12) \quad F = Q(Q'Q)^{-1/2} [\sigma^2 I_{NL} + (Q'Q)^{1/2} (I_N \otimes \Delta) (Q'Q)^{1/2}]^{-1/2} (Q'Q)^{-1/2}Q'$$

(This formula follows from a straightforward application of Wansbeek and Kapteyn (1982).)

The consistency of GLS hinges on the uncorrelatedness of (X,Z,W) and Qu. For fixed T, it is more efficient than the within estimator (3.7). This is exactly as in the simple case of the last section; an explicit proof can be found in Cornwell (1985, section 3.3).

Following H-T, we next wish to consider the case in which some of the regressors are correlated with the effects. Assume that (X_1, Z_1, W_1) are uncorrelated with the effects, in the sense that plim $(NT)^{-1}X_1'Qu = 0$, and similarly for Z_1 and W_1, while (X_2, Z_2, W_2) are correlated with the effects. Let the dimensions of X_1, Z_1, W_1, X_2, Z_2 and W_2 be k_1, j_1, l_1, k_2, j_2, and l_2 (with $k_1 + k_2 = K$, $j_1 + j_2 = J$, and $l_1 + l_2 = L$).

As in the simple model, we begin with the within estimator, in this case (3.7). The within residuals are

$$(3.13) \qquad (y-X\hat{\beta}_W) = Z\gamma + W\delta_o + [Qu + \epsilon + X(\beta - \hat{\beta}_W)].$$

We transform (3.13) by premultiplying by $\Omega^{-1/2}$:

$$(3.14) \qquad \Omega^{-1/2}(y-X\hat{\beta}_W) = \Omega^{-1/2}Z + \Omega^{-1/2}W\delta_o + \Omega^{-1/2}[Qu + \epsilon + X(\beta - \hat{\beta}_W)]$$

The simple consistent estimator is then defined as IV of (3.14), using as instruments

$$(3.15) \qquad B^* = \Omega^{-1/2}B = \Omega^{-1/2}(X_1,Z_1,W_1).$$

This yields the estimator

$$(3.16)\qquad \begin{bmatrix} \hat{\gamma}_W \\ \hat{\delta}_{oW} \end{bmatrix} = [(Z,W)'\Omega^{-1/2}P_{B^*}\Omega^{-1/2}(Z,W)]^{-1}(Z,W)'\Omega^{-1/2}P_{B^*}\Omega^{-1/2}(y-X\hat{\beta}_W).$$

The estimator will exist if we have enough instruments; i.e. if the order condition $k_1 + j_1 + l_1 \geq J + L$, or equivalently $k_1 \geq j_2 + l_2$, is satisified. The corresponding rank condition is that the matrix to be inverted in (3.16) be (asymptotically) of full rank. If it holds, the estimator will be consistent.

As in the simpler model, we could define an alternative simple consistent estimate by using untransformed instruments. In this model, this would make a difference even in the exactly identified case, $k_1 = j_2 + l_2$. We will not pursue these possibilities here, however. Following White (1983), the use of untransformed instruments is clearly suboptimal, if we assume "reduced form" equations for (Z_2, W_2) which are linear in (X_1, Z_1, W_1).

To define our efficient IV estimator, we estimate (3.10) by IV, using as instruments

$$(3.17)\qquad A^* = \Omega^{-1/2} A = \Omega^{-1/2} (M_Q, X_1, Z_1, W_1).$$

This yields

$$(3.18)\qquad \begin{bmatrix} \tilde{\beta}^* \\ \tilde{\gamma}^* \\ \tilde{\delta}_o^* \end{bmatrix} = (G'\Omega^{-1/2}P_{A^*}\Omega^{-1/2}G)^{-1}G'\Omega^{-1/2}P_{A^*}\Omega^{-1/2}y.$$

where $G = (X, Z, W)$.

In Appendix A, we state the necessary conditions for the existence of (3.18), and give the relationship between the efficient estimates (3.18) and the simple consistent estimates (3.16). This discussion can be summarized as follows. If $k_1 < j_2+l_2$, $\hat{\beta}^* = \hat{\beta}_W$ and $(\tilde{\gamma}^*, \tilde{\delta}_o^*)$ does not exist. If $k_1=j_2+l_2$, $\tilde{\beta}^* = \hat{\beta}_W$ and $(\tilde{\gamma}^*, \tilde{\delta}_o^*) = (\hat{\gamma}_W, \hat{\delta}_{oW})$, where $(\hat{\gamma}_W, \hat{\delta}_{oW})$ is defined in (3.16). And, if $k_1 > j_2 + l_2$, $(\tilde{\beta}^*, \tilde{\gamma}, \tilde{\delta}_o^*) \neq (\hat{\beta}_W, \hat{\gamma}_W, \hat{\delta}_{oW})$ with the former being more efficient than the latter. The analogy between these results and the corresponding H-T results for the simple model is evident.

Both (3.16) and (3.18) involve $\Omega^{-1/2}$, which is computationally bothersome. However, these estimates can also be calculated using only Ω^{-1} (or, equivalently, F^2). We have

$$(3.19) \qquad \Omega^{-1} = \frac{1}{\sigma^2} M_Q + Q(Q'Q)^{-1}\Gamma^{-1}(Q'Q)^{-1}Q' = \frac{1}{\sigma^2} M + F^2$$

where $\Gamma = \sigma^2(Q'Q)^{-1} + I_N \otimes \Delta$. So, in the simple consistent estimator,

$$(3.20) \qquad \Omega^{-1/2}P_{B^*}\Omega^{-1/2} = \Omega^{-1}B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1},$$

and (3.16) becomes

$$(3.21) \qquad \begin{bmatrix} \tilde{\gamma}_W \\ \tilde{\delta}_{oW} \end{bmatrix} = [(Z,W)'F^2B(B'\Omega^{-1}B)^{-1}B'F^2(Z,W)]^{-1} (Z,W)'F^2B(B'\Omega^{-1}B)^{-1}B'F^2(y-X\hat{\beta}_W).$$

And, since $P_A^* = M_Q$ plus the projection onto $P_Q\Omega^{-1/2}B$, the efficient estimator (3.18) can be computed as

$$
(3.22)\quad \begin{bmatrix} \tilde{\beta}^* \\ \tilde{\gamma}^* \\ \tilde{\delta}_o^* \end{bmatrix} = \{G'[\frac{1}{\sigma^2} M_Q + F^2B(B'F^2B)^{-1}B'F^2]G\}^{-1}G'[\frac{1}{\sigma^2} M_Q + F^2B(B'F^2B)^{-1}B'F\ ]y
$$

4. <u>A Frontiers Model with Time-Varying Inefficiency</u>

Schmidt and Sickles (1984) consider the estimation of a stochastic frontier production function with panel data, using the model

$$
(4.1)\qquad y_{it} = \alpha + X_{it}'\beta + v_{it} - u_i,
$$

where y = output, X = inputs, v = statistical noise, and u > 0 is a firm effect representing technical inefficiency. This model can obviously be put in the form

$$
(4.2)\qquad y_{it} = \alpha_i + X_{it}'\beta + v_{it},
$$

where $\alpha_i = \alpha - u_i$. The model (4.2) is of the standard form found in the panel data literature, and β can be estimated by standard methods such as "within", GLS or the H-T instrumental variables estimator. It can also be estimated by MLE, assuming a particular distribution for the one-sided error u_i in (4.1). Schmidt and Sickles apply (4.2) to a panel of airlines for the period 1970I - 1977IV (the period prior to deregulation), assuming a Cobb-Douglas technology. Results from the use of "within", GLS, and MLE (assuming a half-normal distribution for the firm effects) are compared, and a Hausman-Wu specification error test is carried out to test the null hypothesis that firm-specific effects are uncorrelated with the regressors.

The great benefit of panel data is that one can choose whether to assume particular distributions of v and u, or whether to assume that technical inefficiency is uncorrelated with the inputs, and that therefore these assumptions are testable. However, these benefits come at the cost of the assumption that the firm effects are constant over time. This is a very strong assumption, and probably would be unrealistic in many potential applications. In terms of the Schmidt and Sickles application, as the airline industry moved into the deregulatory transition and beyond, the potential for unstable productivity patterns (reflected in the firm effects) should be clear. Firms within the industry would be expected to respond differently to the new regulatory environment. Although this issue has been dealt with in part by Sickles, Good and Johnson (1986), the model introduced therein was highly parameterized and required maximum likelihood on a highly nonlinear model. The model we propose here is more parsimoniously parameterized and can be estimated in straightforward ways.

In order to relax the assumption that the firm effects are time-invariant, but in such a way that the advantages of panel data are preserved, we will replace the firm effect (α_i) in (4.2) by a flexibly-parameterized function of time, with parameters that vary over firms. The functional form chosen in this paper is a quadratic:

$$(4.3) \qquad \alpha_{it} = \theta_{i1} + \theta_{i2}\, t + \theta_{i3}\, t^2.$$

Since (4.3) is linear in the elements of θ_{ij} (j=1,2,3), we have exactly the type of model considered in Section 3.

In terms of the notation of Section 3, we have:

(4.4) $W_{it}' = [1, t, t^2]$, $\delta_i' = [\theta_{i1}, \theta_{i2}, \theta_{i3}]$

and with this notation the model (4.2) can be written:

(4.5) $y_{it} = X_{it}'\beta + W_{it}'\delta_i + v_{it}$

Clearly the specification (4.3) implies that output levels vary both over firms and over time. Efficiency measurement focuses on the cross-sectional variation, and the model allows efficiency levels to vary over time. Conversely, the measurement of productivity growth focuses on the temporal variation, and the model allows the rate of productivity growth to vary over firms.

Time-varying firm productivity and efficiency levels and rates of productivity growth can be derived from the residuals based on the within, GLS, and efficient instrumental variables estimators presented in Section 3.[3] In Schmidt and Sickles (1984), using the model (4.1), the residuals $(y_{it} - X_{it}'\hat{\beta})$ are an estimate of $(v_{it} - u_i)$, and the firm effect (for a given firm) is estimated by averaging its residuals over time. Specifically, the estimate of α_i is

(4.6) $\hat{\alpha}_i = \bar{y}_i - \bar{x}_i\hat{\beta}.$

This estimate is consistent as $T\to\infty$. The analogous procedure for the present model is to estimate δ_i by regressing the residuals $(y_{it} - X_{it}'\hat{\beta})$ for firm i on W_{it}; that is, on a constant, time and time-squared. The fitted values from this regression provide an estimate of α_{it} in (4.3) that is consistent (for all i and t) as $T \to \infty$. Finally, in Schmidt and Sickles the frontier intercept α is estimated as the maximum of the $\hat{\alpha}_i$:

$$(4.7) \qquad \hat{\alpha} = \max_j (\hat{\alpha}_j)$$

and the firm-specific level of inefficiency for firm i is estimated as

$$(4.8) \qquad \hat{u}_i = \hat{\alpha} - \hat{\alpha}_i .$$

The analogous procedure here is to estimate the frontier intercept at time t as

$$(4.9) \qquad \hat{\alpha}_t = \max_j (\hat{\alpha}_{jt})$$

and the firm-specific level of technical inefficiency of firm i at time t as

$$(4.10) \quad \hat{u}_{it} = \alpha_t - \hat{\alpha}_{it} .$$

5. Empirical Results

Our data are on U.S. airlines over the time period 1970I-1981IV, so that T = 48. The data follow certificated carriers that existed throughout the study period and that accounted for over 80 percent of domestic air traffic. Information on output and input prices and quantities was obtained from over 250 accounts from the CAB Form-41. These accounts were aggregated into the four broad input measures of capital, labor, energy, and materials; one output measure, available ton miles; and two output attributes, average stage length (thousands of miles) and service quality. Service quality is based on the number of complaints received by the CAB's Office of Consumer Affairs and is normalized by the number of passenger enplanements for that quarter. The output and input quantities and prices are constructed as Tornqvist indices. We examined the following airlines: American, Allegheny, Delta, Eastern, North Central, Ozark, Piedmont, and United, so that N=8. We control for seasonal factors with three dummy variables (with fall the omitted category), and condition on two service attributes, average stage length and quality. For a further discussion of data construction see Sickles (1985) and Sickles, Good, and Johnson (1986).

The functional form that we use for (4.1) is a special case of the transcendental logarithmic function (Christensen, Jorgenson, and Lau, 1973). We assume that the average technology is given by a first-order approximation in the logarithms of input quantities, and a second order approximation in the logarithms of output attributes. In addition, we make the assumptions that input quantities and output characteristics are separable in production, that productivity levels and growths are disembodied, and that seasonal factors are neutral. This reduces the possible number of unrestricted parameters from 66 to 15, a manageable number given the time series nature of our data, the typical collinearity problems associated with data of this sort, and the use of no additional restrictions embodied in the first-order conditions for output maximization, cost minimization, or profit maximization. The average production technology under consideration is therefore

$$(5.1) \qquad \ln Q = \ln \alpha_o + \alpha_k \ln K + \alpha_L \ln L + \alpha_E \ln E + \alpha_M \ln M + \sum_i \text{Season}_i + \sum_i \ln \text{Attribute}_i + \sum_{i \leq j} \sum \ln \text{Attribute}_i \ln \text{Attribute}_j$$

where Q is available ton miles, K, L, E, M are capital, labor, energy, and material input quantities, the seasons are indexed from winter through summer, and where the attributes are average stage length and our service quality index. Summary statistics for the variables in (5.1) are given in Table 1.

Estimation results are given in Tables 2 and 3. Table 2 displays benchmark GLS and within estimates that are comparable to those given in Schmidt and Sickles (1984) in that only the intercept is allowed to vary across firms. Productivity, however, is allowed to vary over the period. The results of GLS and within are comparable, with energy having the largest output elasticity, followed by labor, materials, and capital. Returns to scale are not significantly different from unity for both estimates at the 95% level, and annual productivity growth is about 1.5% in the median period, 1975I. The $\bar{R}^2$ for both sets of results is above 0.999. Table 3 presents the within, GLS and efficient IV estimates given in (3.7), (3.8), and (3.18) with σ_ϵ^2 and Δ estimated from (B.2) and (B.8). Consider first the GLS and within-estimates. The output elasticities do change somewhat across estimation procedures (GLS versus within) as well as across specifications (Table 2 versus Table 3). The within-estimated capital elasticity in Table 3 is considerably higher than either estimate in Table 2, while the within-estimated materials elasticity is considerably lower. Returns to scale are still insiginificantly different from unity at the 95% level.

The consistency of the GLS estimates depends on the effects being uncorrelated with all of the explanatory variables. As explained in Schmidt and Sickles (1984), this assumption can be tested using a Hausman-Wu test based on the significance of the differences between the GLS and within estimates. This test statistic equals 17.2. Its asymptotic distribution is chi-squared with 12 degrees of freedom, and a value of 17.2 is significant only at about the .15 level. Thus there is some evidence against the exogeneity assumptions underlying the GLS estimator, but it is not significant at usual confidence levels such as .05, although this may reflect the low power of the test against nonlocal alternatives.

Despite the insignificance of the evidence against the GLS estimator's exogeneity assumptions, it is reasonable to ask if there is a subset of the explanatory variables for which uncorrelatedness with the effects is more strongly supported by the data. If so, we can impose these uncorrelatedness assumptions using the efficient IV estimator of Section 3. For this purpose we will assume that the seasonal dummy variables and the intercept and time trend variables are uncorrelated with the effects, while the output attribute variables will be treated as correlated with the effects. Correlation patterns between the effects and the input variables were harder to assign a priori, but we decided to treat capital and energy as correlated with the effects, and labor and materials as uncorrelated with the effects. We did this for several reasons. The labor input index is based on headcount data. Since adjustment costs for numbers of employees are typically much higher than for hours (which are not measured in the CAB Form-41), any short-run (quarterly) firm shock will likely result in reduced hours or overtime, not in numbers of employees (Schultze (1985); Shapiro (1986)). Furthermore, since union contracts cover approximately 70% of the employees in our sample airlines, rational expectations would suggest that any information available to the contracting parties when the contract was made would have been conditioned on, and therefore any unforseen firm-specific supply shifts would be orthogonal to employment variation while the contract was in force (Sargent (1978)). The other input which is assumed to be uncorrelated with firm effects is the materials index. This is a residual category, roughly 70% of which is for professional services contracted outside the firm. These include advertising, charter travel bookings, unplanned maintenance of firm's flight equipment by another carrier, and catering services. These data came to us in expenditure form and a Tornqvist index was constructed using a variety of price deflators such as the McCann Erickson Advertising index, the producer price index for miscellaneous business services, and the producer price index for processed foods. The aggregate price indices would have no correlation with airline specific productivity changes unless firms had a substantial degree of monopsony power in those markets. There is no evidence that this is the case. Whatever weak correlation might have existed between the materials expenditure data and firm productivity effects would be mitigated by the index construction.

The efficient IV estimates based on these exogeneity assumptions are given in Table 3. The coefficient estimates are fairly similar to the within estimates, and there is a slight improvement in the precision of the estimates. Furthermore, the eight uncorrelatedness assumptions that underly the efficient IV estimator are testable, and the Hausman-Wu statistic (based on differences between the within and efficient IV estimates) is only 1.08. Thus there is no evidence in the data to make us doubt these exogeneity assumptions.

Table 4 presents the relative efficiency levels derived from our estimates for the carriers at three points in time: 1970I, 1975I, 1981IV. The efficiencies are calculated using the GLS, within, and efficient IV estimates. As expected, the within and efficient IV results are quite similar, while GLS efficiency levels and rankings are quite different from within and efficient IV. In either case there is evidence of considerable change in the efficiency rankings over time; for example, American and United show large improvements in their efficiency rankings from 1970 to 1980.[5].

Growth rates in productivity can be calculated by examining the time derivative of the estimate of (4.2). Although these estimates were quite unstable when evaluated period by period we can compare the average values between the first and last period and calculate simple annualized percent rates of growth in total factor productivity (TFP). These calculations are summarized in Table 5. Below the rates of total factor productivity growth are the output share weighted averages, which are comparable to the estimates from the naive model with heterogeneity in the intercept only. We can see that, although magnitudes are not equal, the estimates based on within and efficient IV are of the same sign (except for Piedmont which is quite small) and roughly the same magnitude. TFP growth rates calculated from the (probably misspecified) GLS estimates are quite different from the within and efficient IV TFP growth rates and suggest an industry average growth rate of 0.44, versus the 1.22-1.85 implied by the consistent within and efficient IV estimates.

It is obvious in Table 4 that, on average, the firms in our sample became more efficient over time. The change in the average efficiency level is illustrated in Figure 1. The average level of efficiency for our eight firms is roughly 82% in 1970I and grows to almost 95% in 1980 before dropping slightly in 1981. It is important to stress that this increase in efficiency levels is not just a reflection of the fact that there was productivity growth over the sample period. A firm's efficiency level for a given time period is calculated by comparing the firm's output to the frontier level calculated using the production function of the most efficient firm (the one with the highest intercept u_{it} in equation (4.2)); see equation (4.5). Thus the empirical fact that drives an increase in efficiency levels over time is that the firms' productivity levels are becoming more similar over time. It is easy to conjecture that this is due to increasing competitive pressures in the airline industry over the sample period, although in fact most of the increase in average efficiency levels occurred before the formal passage of the air deregulation act in late 1978.

The temporal pattern of changes in efficiency levels displayed in Table 4 and Figure 1 is of obvious interest. It indicates exactly the kind of detail available in the present model and not available in the simpler model of Schmidt and Sickles (1984).

6. Conclusions

In this paper we have specified a simple model which, in the presence of panel data, allows us to estimate time-varying efficiency levels for individual firms, without making strong distributional assumptions for technical inefficiency or random noise. We do so by including in the production function a flexible function of time, with parameters that differ across firms. We also generalize the earlier econometric results of Hausman and Taylor (1981) to develop an econometric technique that allows us to choose how many explanatory variables we wish to assume to be uncorrelated with the firm's temporal pattern of productivity growth. We have used this model and these estimators to analyze the U.S. airline industry during two periods of regulation and obtained results that are quite intuitive, reasonable results, including believable evidence on the pattern of changes in efficiency across regulatory environments.

Table 1

Summary Statistics
(48 quarters; 8 Airlines)

Variable	Mean	Standard Deviation
ln Q	19.04	1.38
ln K	16.84	1.11
ln L	17.54	1.15
ln E	16.10	1.27
ln M	16.91	1.14
ln stage length	-1.08	0.65
ln quality	-3.36	0.55
$(\ln \text{stage length})^2$	1.59	1.45
$(\ln \text{quality})^2$	11.57	3.79
(ln stage length)* ln quality	3.60	2.18

Table 2

Heterogeneity in Intercept Only

Variable	GLS		Within	
	Estimate	S.E.	Estimate	S.E.
ln K	.183	.027	.169	.027
ln L	.242	.030	.243	.030
ln E	.502	.025	.500	.025
ln M	.203	.028	.203	.028
Winter	.00198	.0064	.00151	.0060
Spring	.0223	.0066	.0229	.0062
Summer	.0284	.0066	.0303	.0062
ln stage length	.221	.054	.101	.076
ln quality	.0073	.041	.0122	.040
$(\ln \text{stage length})^2$	.0434	.016	.0103	.0213
$(\ln \text{quality})^2$	-.00370	.0058	-.00355	.0058
ln stage length* ln quality	.0251	.0081	.0261	.0081
Intercept	.0205	.290	---	---
Time	.000591	.00084	.0000743	.00083
Time^2	.000065	.000017	.0000875	.00000191
σ_u^2	.00180		---	
σ_ϵ^2	.00166		.00169	

Table 3

Heterogeneity in Intercept, Time, $Time^2$

Variable	GLS		Within		EffIV	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
ln K	.193	.0303	.233	.0321	.221	.0276
ln L	.317	.0326	.300	.0273	.300	.0270
ln E	.466	.0286	.498	.0259	.494	.0245
ln M	.147	.0300	.139	.0249	.137	.0242
Winter	.00189	.00613	.00280	.00491	.00218	.00484
Spring	.0202	.00629	.0207	.00509	.0211	.00495
Summer	.0264	.00625	.0252	.00515	.0263	.00496
ln stage length	.0780	.0666	−.0608	.0902	−.102	.0788
ln quality	−.0383	.0429	−.0237	.0349	−.0172	.0342
$(\ln \text{stage length})^2$	−.0247	.0255	−.0874	.0365	−.105	.0335
$(\ln \text{quality})^2$	−.00714	.00628	−.00464	.00516	−.00384	.00505
ln stage length* ln quality	.0110	.00883	.00941	.00724	.00992	.00705
Intercept	−.0404	.396	---	---	−.407	.406
Time	.00104	.00171	---	---	.000224	.000214
$Time^2$	.0000464	.0000314	---	---	.0000465	.0000372
Λ	.00407			.0179		
	−.000211	.0000180		−.000561	.0000328	
	$.264 \times 10^{-5}$	$-.291 \times 10^{-6}$	$.552 \times 10^{-8}$	$.448 \times 10^{-5}$	$-.474 \times 10^{-6}$	$.904 \times 10^{-8}$
σ^2_ϵ	.00166			.00100		

Table 4

Efficiency Levels (%)
For Selected Time Periods
(1970I, 1975I, 1980IV)

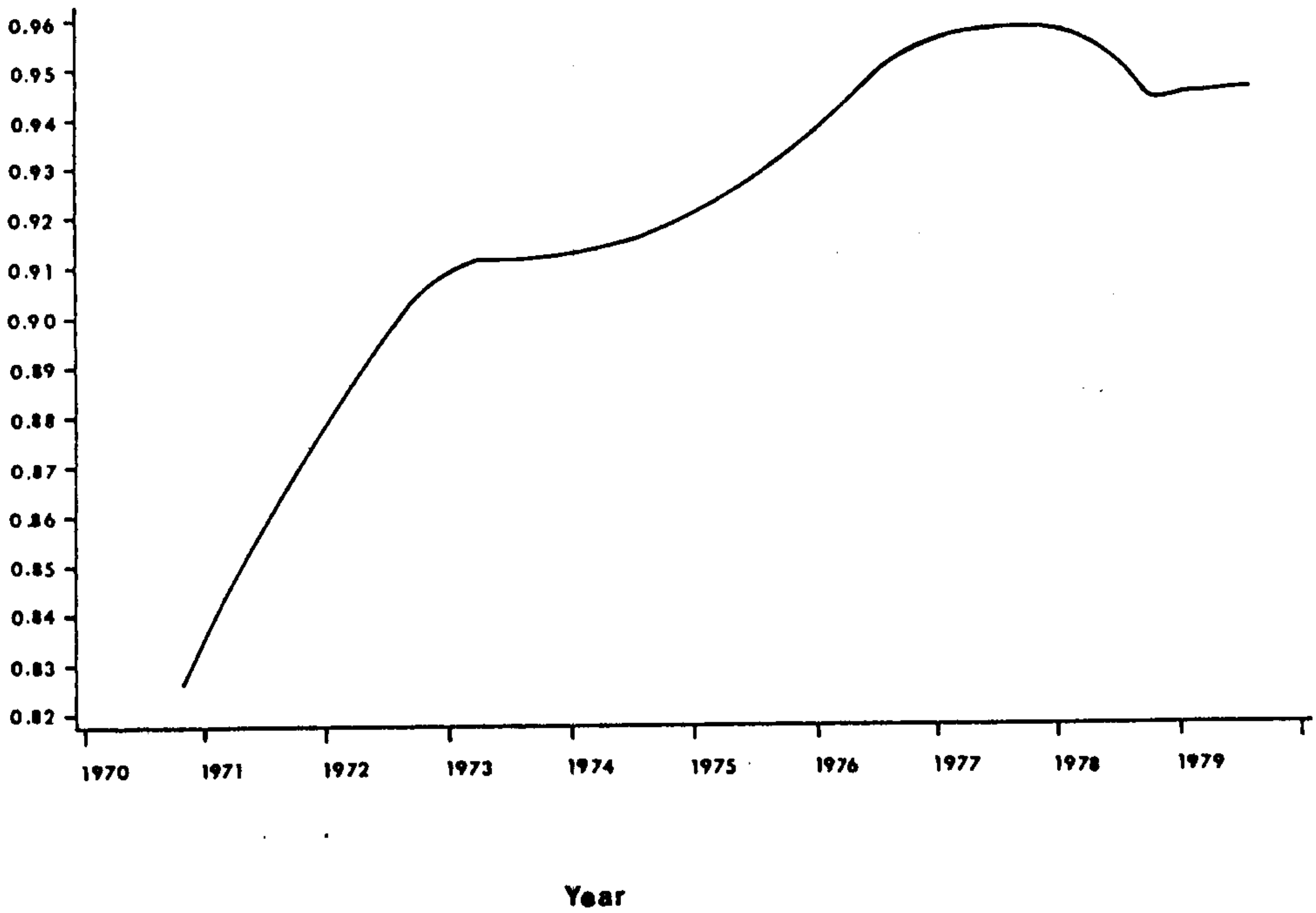
Carrier	GLS	Within	EffIV
American	81, 95, 93	65, 90, 93	72, 93, 94
Alleghany	92, 88, 83	85, 86, 83	86, 86, 80
Delta	92, 99, 99	78, 91, 94	81, 93, 92
Eastern	74, 92, 92	60, 81, 88	64, 84, 87
North Central	86, 100, 88	85, 100, 84	86, 100, 82
Ozark	100, 96, 65	100, 97, 99	100, 96, 93
Piedmont	88, 93, 97	89, 98, 100	90, 97, 94
United	87, 92, 100	66, 84, 100	72, 88, 100

Table 5

Annual Productivity Growth
Rates (%) From 1970I-1981IV

Carrier	GLS	Within	EffIV
American	0.45	2.08	1.13
Alleghany	-0.05	-0.42	-1.27
Delta	-0.62	1.08	2.21
Eastern	0.86	2.08	1.24
North Central	-0.38	-0.33	-1.07
Ozark	-3.55	-0.33	-1.30
Piedmont	0.12	0.64	-0.30
United	1.08	2.55	1.60
Output Share Weighted Average	0.44	1.85	1.22

Figure 1. Average Relative Productivity



APPENDIX A

To derive the characteristics of $\tilde{\beta}^*$ and $\tilde{\gamma}^*$, and $\tilde{\delta}^*$ when $k_1 \leq j_2$, we will make use of the following lemma[6]:

Lemma: Let H and C be nxm and nxp matrices respectively, such that H and H C both have rank m. Then H and $H(H'H)^{-1}H'C$ have the same column space.

Case I (under-identification): If $k_1 < j_2 + l_2$, $\tilde{\beta}^* = \tilde{\beta}_W$ and $\begin{bmatrix} \tilde{\gamma}^* \\ \tilde{\delta}^*_o \end{bmatrix}$ does not exist.

First, consider β. The "efficient" estimator of β, $\tilde{\beta}^*$, is obtained by OLS of $P_{A^*}\Omega^{-1/2}y$ on the part of $P_{A^*}\Omega^{-1/2}X$ orthogonal to $P_{A^*}\Omega^{-1/2}(Z,W)$. Now, since $A^* = \Omega^{-1/2}(M_Q,X_1,Z_1,W_1) = (M_Q,B^*)$, $P_{A^*} = M_Q$ + projection onto the part of B^* orthogonal to M_Q:

$$P_{A^*} = M_Q + P_QB^*(B^{*\prime}P_QB^*)^{-1}B^{*\prime}P_Q$$

$$\text{(A.1)} \qquad = M_Q + P_Q\Omega^{-1/2}B(B'\Omega^{-1/2}P_Q\Omega^{-1/2}B)^{-1}B'\Omega^{-1/2}P_Q$$

Hence, $P_{A^*} = M_Q + FB(B'F^2B)^{-1}B'F$ (Since $P_Q\Omega^{-1/2} = F$), and

$$\text{(A.3)} \qquad P_{A^*}\Omega^{-1/2} = \frac{1}{\sigma} M_Q + FB(B'F^2B)^{-1}B'F^2.$$

Therefore,

$$\text{(A.4)} \quad P_{A^*}\Omega^{-1/2}X = \frac{1}{\sigma} M_Q X + FB(B'F^2B)^{-1}B'F^2X$$

$$\text{(A.5)} \quad P_{A^*}\Omega^{-1/2}(Z,W) = FB(B'F^2B)^{-1}B'F^2(Z,W)$$

$$\text{(A.6)} \quad P_{A^*}\Omega^{-1/2}y = \frac{1}{\sigma} M_Q Y + FB(B'F^2B)^{-1}B'F^2y.$$

Given the above lemma (with H = FB and C = F(Z,W)), when $k_1 < j_2 + l_2$, the rank of B determines the rank of $P_{A^*}\Omega^{-1/2}(Z,W)$. Thus, $P_{A^*}\Omega^{-1/2}(Z,W)$ and FB share the same column space and null space.

Hence, the part of $P_{A^*}\Omega^{-1/2}X$ orthogonal to $P_{A^*}\Omega^{-1/2}(Z,W)$ is exactly the part which is orthogonal to FB. This part of $P_{A^*}\Omega^{-1/2}X$ is $\frac{1}{\sigma} M_Q X$. So, when $k_1 < j_2 + l_2$,

$$\text{(A.7)} \quad \begin{aligned} \tilde{\beta}^* &= [(\tfrac{1}{\sigma} M_Q X)' \tfrac{1}{\sigma} M_Q X]^{-1}(\tfrac{1}{\sigma} M_Q X)'[\tfrac{1}{\sigma} M_Q + FB(B'F^2B)^{-1}B'F^2]y \\ &= (X'M_Q X)^{-1}X'M_Q y \\ &= \hat{\beta}_W. \end{aligned}$$

Now, consider γ and δ_o. When $k_1 < j_2 + l_2$, $P_{A^*}\Omega^{-1/2}(Z,W)$ is not of full column rank. So, a (J+L)-dimensional vector ξ such that $P_{A^*}\Omega^{-1/2}(Z,W)\,\xi = 0$ exists, and $\begin{bmatrix} \gamma \\ \delta_o \end{bmatrix}$ cannot be distinguished from

$\begin{bmatrix} \gamma \\ \delta_o \end{bmatrix} + \xi$. Therefore, $\begin{bmatrix} \tilde{\gamma}^* \\ \tilde{\delta}_o^* \end{bmatrix}$ does not exist in the under-identified case.

Case II (exact-identification): If $k_1 = j_2 + l_2$, $\tilde{\beta}^* = \tilde{\beta}_W$ and

$$\begin{bmatrix} \tilde{\gamma}^* \\ \tilde{\delta}_o^* \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_W \\ \hat{\delta}_{oW} \end{bmatrix}$$

Again, rank $[F(Z,W)]$ = rank (FB). So, following the argument in case I, $\tilde{\beta}^* = \hat{\beta}_W$ when $k_1 = j_2$.

Since $\tilde{\beta}^* = \hat{\beta}_W$ in this case,

$$\begin{bmatrix} \tilde{\gamma}^* \\ \tilde{\delta}_o^* \end{bmatrix} = \text{OLS of } P_{A^*}\Omega^{-1/2}(y - X\tilde{\beta}^*) \text{ on } P_{A^*}\Omega^{-1/2}(Z,W)$$

whereas

$$\text{(A.8)} \quad \begin{bmatrix} \hat{\gamma}_W \\ \hat{\delta}_{oW} \end{bmatrix} = \text{OLS of } P_{B^*}\Omega^{-1/2}(y - X\tilde{\beta}^*) \text{ on } P_{A^*}\Omega^{-1/2}(Z,W)$$

Now,

$$\text{(A.9)} \quad P_{A^*}\Omega^{-1/2}(y - X\tilde{\beta}^*) = \frac{1}{\sigma} M_Q(y - X\tilde{\beta}^*) + FB(B'F^2B)^{-1}B'F^2(y - X\tilde{\beta}^*)$$

and

$$\text{(A.10)} \quad P_{B^*}\Omega^{-1/2}(y - X\tilde{\beta}^*) = \Omega^{-1}B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}(y - X\tilde{\beta}^*)$$

$$= \Omega^{-1/2}B(B'\Omega^{-1}B)^{-1}B'F^2(y - X\tilde{\beta}^*)$$

since $B'M_Q(y - X\tilde{\beta}^*) = 0$. Then note

$$\text{(A.11)} \quad P_{B^*}\Omega^{-1/2}(Z,W) = \Omega^{-1/2}B(B'\Omega^{-1}B)^{-1}B'F^2(Z,W).$$

$$\text{(A.12)} \quad \begin{bmatrix} \hat{\gamma}_W \\ \hat{\delta}_{oW} \end{bmatrix} = [(Z,W)'F^2B(B'F^2B)^{-1}B'F^2(Z,W)]^{-1}(Z,W)'F^2B(B'F^2B)^{-1}B'F^2(y-X\tilde{\beta}^*)$$

(A.13) $$\begin{bmatrix} \hat{\gamma}_W \\ \hat{\delta}_{oW} \end{bmatrix} = [(Z,W)'F^2B(B'\Omega^{-1}B)^{-1}B'F^2(Z,W)]^{-1}(Z,W)'F^2B(B'\Omega^{-1}B)B'F^2(y-X\tilde{\beta}^*),$$

which are not generally equivalent since $B'F^2B \neq B'\Omega^{-1}B$ (specifically, $X'F^2X \neq X'\Omega^{-1}X$). However, when $k_1 = j_2$, $B'F^2(Z,W)$ is nonsingular and

(A.14) $$\begin{bmatrix} \tilde{\gamma}^* \\ \tilde{\delta}_{oW} \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_W \\ \hat{\delta}_{oW} \end{bmatrix} = [B'F^2(Z,W)]^{-1}B'F^2(y - X\tilde{\beta}^*)$$

<u>Case III</u> (over-identification): If $k_1 > j_2 + l_2$, $\begin{bmatrix} \tilde{\beta}^* \\ \tilde{\gamma}^* \\ \tilde{\delta}^*_{oW} \end{bmatrix} \neq \begin{bmatrix} \hat{\beta}_W \\ \hat{\gamma}_W \\ \hat{\delta}_{oW} \end{bmatrix}$ and the former is more efficient.

If $k_1 > j_2 + l_2$, rank (FB) > rank F(W,Z). Then, the column space of $P_{A^*}\Omega^{-1/2}(Z,W)$ is smaller than the column space of FB. This means that there are parts of $P_{A^*}\Omega^{-1/2}X$ orthogonal to $P_{A^*}\Omega^{-1/2}(Z,W)$, even though they are not orthogonal to FB. Hence $\tilde{\beta}^* \neq \hat{\beta}_W$.

Since $\tilde{\beta}^* \neq \hat{\beta}_W$ in this case, $(y-X\tilde{\beta}^*) \neq (y-X\hat{\beta}_W)$. Additionally, there is the general nonequivalence of $B'\Omega^{-1}B$ and $B'F^2B$ (which is not mentioned by H-T). So, for two reasons, $\begin{bmatrix} \tilde{\gamma}^* \\ \tilde{\delta}^*_o \end{bmatrix} \neq \begin{bmatrix} \hat{\gamma}_W \\ \hat{\delta}_{oW} \end{bmatrix}$. And because $\begin{bmatrix} \tilde{\gamma}^* \\ \tilde{\delta}^*_o \end{bmatrix}$ is asymptotically efficient, $\begin{bmatrix} \hat{\gamma}_W \\ \hat{\delta}_{oW} \end{bmatrix}$ is not in this case.

APPENDIX B

In this Appendix, we consider the consistent (as $N\to\infty$) estimation of σ^2 and Δ, the unknown parameters in Ω.

As in the H-T model, $\hat{\sigma}^2$ is derived from the within residuals. Let $\tilde{y} = M_Q y$, $\tilde{X} = M_Q X$, $\tilde{\epsilon} = M_Q \epsilon$, and $M_{\tilde{X}} = I - \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'$. Then, the SSE from the within regression may be written as

$$\text{(B.1)} \qquad \tilde{y}'M_{\tilde{X}}\tilde{y} = \tilde{\epsilon}'\tilde{\epsilon} - \tilde{\epsilon}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{\epsilon},$$

and therefore

$$\text{(B.2)} \qquad \hat{\sigma}^2 = \frac{1}{N(T-L)}\,\tilde{y}'M_{\tilde{X}}\tilde{y} = \frac{1}{N(T-L)}\,y'M_Q(y-X\tilde{\beta}_W).$$

This estimator is consistent:

$$\text{(B.3)} \qquad \text{plim}\,\hat{\sigma}^2 = \text{plim}\,\frac{1}{N(T-L)}\,\tilde{y}'M_{\tilde{X}}y = \text{plim}\,\tilde{\epsilon}'\tilde{\epsilon} = \sigma^2,$$

since $\tilde{\epsilon}'\tilde{\epsilon} = \epsilon'\epsilon - \epsilon Q(Q'Q)^{-1}Q'\epsilon$.

A consistent estimator of Δ may be constructed as follows. Perform instrumental variables on

$$\text{(B.4)} \qquad (y-X\hat{\beta}_W) = Z\gamma + W\delta_o + (Qu + \epsilon),$$

using $B = (X_1, Z_1, W_1)$ as instruments. From the IV residuals, we can form

$$\hat{\Omega} = \begin{bmatrix} e_1 e_1' & & \\ & \ddots & \\ & & e_N e_N' \end{bmatrix}. \tag{B.5}$$

where e_i denotes a Tx1 vector of IV residuals. Obviously, $\hat{\Omega}$ is not a consistent estimator of $\hat{\Omega}$. However, following White (1980), $Q'\Omega Q$, can be regarded as a consistent estimator of $Q'\Omega Q$. Furthermore,

$$Q'\Omega Q = \sigma^2 Q'Q + Q'Q(I_N \otimes X\Delta)Q'Q, \tag{B.6}$$

so that

$$I_N \otimes \Delta = (Q'Q)^{-1}Q'\Omega Q(Q'Q)^{-1} - \sigma^2(Q'Q)^{-1}, \tag{B.7}$$

where each matrix on the right-hand-side of (B.7) is block diagonal. This suggests the estimator

$$\hat{\Delta} = \frac{1}{N} \sum_{i=1}^{N} [(W_i'W_i)^{-1}W_i'e_i e_i'W_i(W_i'W_i)^{-1} - \hat{\sigma}^2(W_i'W_i)^{-1}]. \tag{B.8}$$

A direct calculation reveals that this estimator is consistent (Cornwell (1985, Appendix B to Chapter 4)).

FOOTNOTES

[1]Amemiya and MaCurdy (1986) introduce an alternative IV estimator that, under stronger assumptions, is more efficient than the H-T estimator. For a clear exposition of the relationship between the two estimators, see Breusch, Mizon, and Schmidt (1985).

[2]More details on the fixed effects treatment of (3.4) can be found in Cornwell (1985).

[3]For a discussion of maximum likelihood estimators for stochastic panel frontiers which treat time-varying inefficiency see Kumbhakar (1988).

[4]We attempted to include second order terms for the inputs but the almost perfect collinearity in the moment matrix prevented us from obtaining unique parameter estimates. Within results using the generalized inverse gave us an F-statistic of 8.75 for the test of the joint insignificance of the second order effects of the logarithms of input quantities. The joint insignificance of these parameters is thus not rejected at reasonable significance levels. Instead of dealing with the problem by imposing more structure, e.g. adding optimizing assumptions in the form of first order conditions to increase the degrees of freedom, we decided let the data and its limitations speak. We simply cannot identify second order input effects using our data set and largely (or completely) the within variation in variables.

[5]Productivity levels (%) derived from the within results of Table 2 for American, Alleghany, Delta, Eastern, North Central, Ozark, Piedmont, and United are: 96, 80, 100, 89, 78, 82, 81, 97. Since parameter heterogeneity is allowed for only the constant term, productivity levels are constant over the sample period, an assumption which is clearly rejected at any reasonable level of significance (F-statistic = 17.06; $F_{.05,14,348} = 1.65$). Although

the constancy of the time and time firm dummies is rejected, there is still the possibility that productivity rankings may not be affected a great deal. This is not the case, although there does appear to be more concordance between rankings in the later periods. Spearman rank correlations between the productivities based on (2.0) and (3.18) are -.539, -.476, and .428.

[6]We are indebted to Trevor Breusch for pointing this result out to us.

REFERENCES

Aigner, D. J., C. A. K. Lovell and P. Schmidt, "Formulation and Estimation of Stochastic Frontier Production Function Models," Journal of Econometrics, 6, 21-37.

Amemiya, T. and MaCurdy (1986), "Instrumental Variables Estimation of an Error Components Model," Econometrica, 54, 869-880.

Breusch, T., G. Mizon, and P. Schmidt (1987), "Efficient Estimation Using Panel Data," Econometrics and Economic Theory Working Paper No. 8608, Michigan State University.

Christensen, L. R. , D. W. Jorgenson, and L. J. Lau, "Transcendental Logarithmic Production Frontiers, Review of Economics and Statistics, 55, 28-45, 1973.

Cornwell, C. (1985), Panel Data With Cross - Sectional Variation in Slopes as Well as Intercept, unpublished doctoral dissertation, Michigan State University.

Kumbhakar, S. C. (1988). "Production Frontiers, Panel Data and Time Variant Technical Inefficiency." mimeo.

Hausman, J. A. and Taylor, W. (1981), "Panel Data and Unobservable Individual Effects," Econometrica, 49, 1377-1399.

Schmidt, P. (1985), "Frontier Production Functions," Econometric Reviews, Vol. 4, No. 2, 289-328.

Schmidt, P. and R. E. Sickles (1984), "Production Frontiers and Panel Data," Journal of Business and Economic Statistics, 2, 367-374.

Sargent, T. J. (1978), "Estimation of Dynamic Labor Demand Schedules under Rational Expectations," Journal of Political Economy, 86, 1009-1044.

Schultze, C. (1986) "Microeconomic Efficiency and Nominal Wage Stickiness," American Economic Review, 75, 1-15.

Shapiro, M. D. (1986), "The Dynamic Demand for Capital and Labor," Quarterly Journal of Economics, 101, 513-542.

Sickles, R. C. (1985), "A Nonlinear Multivariate Error-Components Analysis of Technology and Specific Factor Productivy Growth with an Application to the U.S. Airlines," Journal of Econometrics, 27, 61-78.

Sickles, R., D. Good, and R. Johnson (1986), "Allocative Distortions and the Regulatory Transition of the Airline Industry," Journal of Econometrics, 33, 143-163.

Swamy, P.A.V.B. (1974), "Linear Models with Random Coefficients; in Frontiers of Econometrics, P. Zarembka, editor, New York: Academic Press.

Swamy, P.A.V.B. (1971), Statistical Inference in Random Coefficient Regression Models; New York: Springer-Verlag.

Wansbeek, T. and Kapteyn, A. (1982), "A Class of Decompositions of the Variance-Covariance Matrix of a General Error Components Model," Econometrica, 50, 713-724.

White, H. (1983), "Instrumental Variables Analogues of Generalized Least Squares Estimators", working paper 83-22, University of California - San Diego.

White, H. (1980), "A Heteroskedasticty - Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," Econometrica, 48, 817-838.