Semiparametric Estimation of Stochastic Production Frontier Models

Yanqin FAN
Department of Economics, University of Windsor, Windsor, Ontario, N9B 3P4, Canada

Qi LI
Department of Economics, University of Guelph, Guelph, Ontario, N1G 2W1, Canada

Alfons WEERSINK
Department of Agricultural Economics and Business, University of Guelph, Guelph, Ontario N1G 2W1, Canada

This article extends the linear stochastic frontier model proposed by Aigner, Lovell, and Schmidt to a semiparametric frontier model in which the functional form of the production frontier is unspecified and the distributions of the composite error terms are of known form. Pseudolikelihood estimators of the parameters characterizing the two error terms of the model are constructed based on kernel estimation of the conditional mean function. The Monte Carlo results show that the proposed estimators perform well in finite samples. An empirical application is presented. Extensions to a partially linear frontier function and to more flexible one-sided error distributions than the half-normal are discussed.

KEY WORDS: Kernel estimation; Partially linear model; Production efficiency; Pseudo maximum likelihood estimation.

Two different approaches to the specification and estimation of production-frontier functions have been used in the literature on frontier analysis, the mathematical programming or *Data Envelopment Analysis* (DEA) and econometric approaches. The DEA approach, first used by Farrell (1957), has the advantage that no explicit functional form need be imposed on the data. The calculated frontier may be warped, however, if the data are contaminated by statistical noise (e.g., Bauer 1990). For detecting outliers using the DEA approach, see Wilson (1993). For a complete discussion on recent developments in DEA, see Seiford and Thrall (1990). The econometric approach to frontier analysis was first proposed by Aigner, Lovell, and Schmidt (1977), Meeusen and van den Broeck (1977), and Battese and Corra (1977). It uses a parametric representation of technology along with a two-part composed-error term. One component, often assumed to follow a normal distribution, captures the random nature of production, and the other, following a particular one-sided distribution, represents technical inefficiency. Although the econometric approach accommodates statistical noise, measurement error, and exogenous shocks, it has imposed restrictive assumptions on functional forms representing both the production technology and the composed-error terms. Econometricians in the past decade have devoted most of their attention to relaxing these restrictions. As a result, panel-data frontier models have been developed that can be implemented without having to impose strong distributional assumptions for technical inefficiency or statistical noise; see Schmidt and Sickles (1984) and Cornwell, Schmidt, and Sickles (1990). Simar (1992) introduced an interesting approach that combines a nonparametric approach called FDH (Free Disposal Hull), proposed by Deprins, Simar, and Tulkens (1984) and the ordinary least squares (OLS) method. Park and Simar (1994) proposed an efficient semiparametric estimation procedure when the one-sided error has an unknown distribution. Hausman and Taylor (1981) considered a panel-data model that allows some of the covariates to be time-invariant. Without panel data, Kopp and Mullahy (1990) were able to relax the assumptions on the independent and identically distributed (iid) normal character of the noise component by using the generalized method of moment technique, and Greene (1990) allowed the one-sided component to have a more general distribution than the less flexible half-normal proposed by Aigner et al. (1977).

Although the aforementioned works were successful in relaxing some or all of the distributional assumptions on the composed-error terms, the functional form representing the production technology is still assumed to be known apart from a finite number of unknown parameters. To the best of our knowledge, no work has been done to relax such parametric restrictions. It is thus the objective of this article to attempt to bridge this gap. This is an important issue because misspecification in the functional form of the frontier function may lead to erroneous conclusions drawn from the resulting frontier even if the distributions of the composed errors are correctly specified.

The method proposed in this article makes use of nonparametric regression techniques to relax parametric restrictions on the functional form representing technology, which seems to be quite natural in view of the rapid development in nonparametric regression analysis (see Härdle 1990). The semiparametric stochastic frontier model proposed in this article is different from the standard nonparametric regression model, however, in that the conditional

Journal of Business & Economic Statistics
October 1996, Vol. 14, No. 4

expectation of the composed error term is not 0 due to the presence of the one-sided error. Thus, the nonparametric function representing the production technology cannot be consistently estimated by the existing nonparametric regression techniques, and certain modifications must be made. We propose a way that overcomes this problem.

Our technique can be viewed as a semiparametric extension of the pseudolikelihood approach originally proposed by Aigner, Amemiya, and Poirier (1976). To construct the likelihood function, we adopt the standard assumptions on the distributions of the composed-error terms. Therefore, this article can be viewed as a complement to those that avoid restrictive assumptions on the composed-error terms. Another advantage of our approach is that the likelihood function is maximized over one parameter as opposed to several under the stochastic approach of Aigner et al. (1977).

The rest of the article is organized as follows. In Section 1, we propose our semiparametric stochastic frontier model and construct pseudolikelihood estimators of the parameters characterizing the two error terms of the model based on kernel estimation of the conditional expectation of y, the dependent variable, on x, the vector of independent variables. These estimators are $\sqrt{n}$-consistent; that is, they have the same convergence rate as the corresponding parametric estimators based on a correct parametric specification of the frontier. The nonparametric frontier itself can be consistently estimated at the typical convergence rate of kernel regression estimators. The finite-sample performance of the pseudolikelihood estimators proposed in Section 1 is investigated in Section 2 by simulation experiments. Section 3 presents an empirical application of the proposed estimation method using data on Quebec dairy farms. Section 4 discusses extensions to a partially linear frontier function and to more flexible one-sided error distributions than the half-normal proposed by Aigner et al. (1977). Section 5 summarizes the findings and offers suggestions for future research.

1. THE MODEL

We consider the following semiparametric stochastic production frontier model:

$$y_i = g(x_i) + \varepsilon_i, \qquad i = 1, \ldots, n, \tag{1}$$

with

$$\varepsilon_i = v_i - u_i, \tag{2}$$

where x_i is a $p \times 1$ vector of random regressors, $g(x_i)$ is a smooth function unknown to the researcher, ε_i is the two-part composed-error term in which the second term u_i is one-sided representing technical inefficiency and the first term v_i is two-sided representing statistical noise. Note that it differs from the standard nonparametric regression model because the conditional expectation of ε_i given x_i is not 0. The random vector x_i and the error ε_i are assumed to be independent. (For some industries, it may be more reasonable to assume that the technical inefficiency level depends on the input level. In that case, one needs to have knowledge of the conditional distribution of ε given x.) Following Aigner et al. (1977), we will assume that u_i is iid derived from an $N(0, \sigma_u^2)$ distribution truncated at 0 and v_i is iid $N(0, \sigma_v^2)$. The decomposition of the error term proposed by Aigner et al. (1977) allows a frontier that is stochastic. The distribution function of the sum of a symmetric normal random variable and a truncated normal random variable can be found in the work of Weinstein (1964); see also Aigner et al. (1977). The probability density function of the composite disturbance ε is

$$f(\varepsilon) = \frac{2}{\sigma}\,\phi\left(\frac{\varepsilon}{\sigma}\right)[1 - \Phi(\varepsilon\lambda\sigma^{-1})], \qquad -\infty \le \varepsilon \le +\infty, \tag{3}$$

where $\sigma^2 = \sigma_u^2 + \sigma_v^2, \lambda = \sigma_u/\sigma_v$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and distribution functions, respectively. The mean and the variance of ε are given by

$$\begin{aligned} E(\varepsilon) &= -E(u) = -\frac{\sqrt{2}}{\sqrt{\pi}}\,\sigma_u \equiv -\mu \\ \sigma_\varepsilon^2 &\equiv V(\varepsilon) = V(u) + V(v) = \left(\frac{\pi - 2}{\pi}\right)\sigma_u^2 + \sigma_v^2. \end{aligned} \tag{4}$$

From (3), it follows that the log-likelihood function is

$$\ln l(y|\lambda, \sigma^2) = \frac{n}{2}\ln(2/\pi) - n\ln\sigma + \sum_{i=1}^{n}\ln[1 - \Phi(\varepsilon_i\lambda\sigma^{-1})] - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^2. \tag{5}$$

If $g(x_i)$ is completely known, one can get the first-order conditions (F.O.C.) for σ^2 and λ conditional on $\varepsilon_i = y_i - g(x_i)$:

$$\frac{\partial \ln l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(y_i - g(x_i))^2 + \frac{\lambda}{2\sigma^3}\sum_{i=1}^{n}\frac{\phi_i}{(1 - \Phi_i)}(y_i - g(x_i)) = 0 \tag{6}$$

and

$$\frac{\partial \ln l}{\partial \lambda} = -\frac{1}{\sigma}\sum_{i=1}^{n}\frac{\phi_i}{(1 - \Phi_i)}(y_i - g(x_i)) = 0. \tag{7}$$

From (6) and (7), one can concentrate out σ^2 to get

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - g(x_i))^2. \tag{8}$$

Thus one only needs to maximize the log-likelihood function over the single parameter λ.

However, $g(x_i)$ is generally unknown. If it belongs to a class of parametric functions, say $g_0(x_i, \beta)$, for some known smooth function $g_0(\cdot)$ and a finite number of unknown parameters β, one can then derive another F.O.C. for the unknown parameter β:

$$\frac{\partial \ln l}{\partial \beta} = \frac{1}{\sigma^2} \sum_{i=1}^{n} [y_i - g_0(x_i, \beta)] \frac{\partial g_0(x_i, \beta)}{\partial \beta} + \frac{\lambda}{\sigma} \sum_{i=1}^{n} \frac{\phi_i}{(1-\Phi_i)} (y_i - g_0(x_i, \beta)). \quad (9)$$

Replacing $g(x_i)$ by $g_0(x_i, \beta)$ in (6) and (7) and then solving (6), (7), and (9) simultaneously leads to the maximum likelihood estimators of σ^2, λ, and β.

The preceding method proposed by Aigner et al. (1977) requires choosing a specific parametric functional form for the production frontier, which may impose unwarranted structure on the technology. If we do not assume that $g(\cdot)$ is completely known nor do we assume that $g(\cdot)$ has a known parametric functional form, the preceding procedures are not feasible. One alternative is to try to estimate $g(x_i)$ by some nonparametric estimator. Existing nonparametric regression techniques cannot be applied directly to the estimation of $g(x_i)$, however, because $g(x_i)$ is not the conditional expectation of y_i given x_i: $E(y_i|x_i) = g(x_i) - E(u_i|x_i) \equiv g(x_i) - u \neq g(x_i)$. Indeed one cannot separate $g(x_i)$ from $E(y_i|x_i)$ using nonparametric estimation.

This dilemma can be resolved by writing $g(x_i) \equiv E(y_i|x_i) + \mu$, where $\mu = (\sqrt{2}\sigma_u)/\sqrt{\pi} = (\sqrt{2}\sigma\lambda)/[\pi(1+\lambda^2)]^{1/2}$. Now $E(y_i|x_i)$ can be consistently estimated by some nonparametric technique. To proceed, we note that the first-order conditions for σ^2 and λ given in (6) and (7) are valid even if $g(x_i)$ is unknown provided that it does not depend on either σ^2 or λ. Therefore, we can substitute $E(y_i|x_i) + \mu$ for $g(x_i)$ into (6) and (7) to obtain the following two equations equivalent to (6) and (7), respectively:

$$\frac{\partial \ln l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (y_i - E(y_i|x_i) - \mu)^2 + \frac{\lambda}{2\sigma^3} \sum_{i=1}^{n} \frac{\phi_i}{(1-\Phi_i)} (y_i - E(y_i|x_i) - \mu) = 0 \quad (10)$$

and

$$\frac{\partial \ln l}{\partial \lambda} = -\frac{1}{\sigma} \sum_{i=1}^{n} \frac{\phi_i}{(1-\Phi_i)} (y_i - E(y_i|x_i) - \mu) = 0. \quad (11)$$

Note that, although $\mu = \mu(\lambda, \sigma^2)$, the F.O.C. in (11) should not contain an extra term of $(\partial\mu/\partial\lambda)$. This is because $E(y_i|x_i) = g(x_i) - \mu$; hence, $\partial[E(y_i|x_i) + \mu]/\partial\lambda \equiv \partial g(x_i)/\partial\lambda = 0$. From (11), we have

$$\sum_{i=1}^{n} \frac{\phi_i}{(1-\Phi_i)} (y_i - E(y_i|x_i) - \mu) = 0.$$

Substituting this into (10) yields

$$-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (y_i - E(y_i|x_i) - \mu)^2 = 0. \quad (12)$$

Replacing μ in (12) by $\mu = (\sqrt{2}\lambda\sigma)/[\pi(1+\lambda^2)]^{1/2}$ leads to a quadratic equation in σ, which gives a unique positive solution for σ: $\tilde{\sigma}_{\text{ml}} = (b + [b^2 + 4|ac|]^{1/2})/(2|a|)$, where $a = -[1 - 2\lambda^2/\{\pi(1+\lambda^2)\}], b = \{2\sqrt{2}\lambda/[\pi(1+\lambda^2)]^{1/2}\}[n^{-1}\sum_{i=1}^{n} (y_i - E(y_i|x_i))]$, and $c = n^{-1}\sum_{i=1}^{n} [y_i - E(y_i|x_i)]^2$. Obviously $b = O_p(n^{-1/2})$. Hence we can also estimate σ^2 by

$$\tilde{\sigma}^2 = \frac{c}{|a|} = \frac{1}{n} \sum_{i=1}^{n} [y_i - E(y_i|x_i)]^2 \Big/ \left[1 - \frac{2\lambda^2}{\pi(1+\lambda^2)}\right]. \quad (13)$$

It is straightforward to show that $\tilde{\sigma}^2$ is a $\sqrt{n}$-consistent estimator of σ^2 by virtue of the law of large numbers. However, $\tilde{\sigma}^2$ given in (13) is not operational, not only because λ is unknown but also because $E(y_i|x_i)$ is unknown. The conditional expectation $E(y_i|x_i)$ can be estimated by using any of the existing nonparametric regression techniques such as the well-known kernel estimator, the nearest neighbor, and the series estimators, among others (see Härdle 1990). In this article, we use the kernel estimator because it is one of the most studied nonparametric estimators. Let $\hat{E}(y_i|x_i)$ denote the kernel estimator of $E(y_i|x_i)$. It is given by $\hat{E}(y_i|x_i) = \sum_{j=1}^{n} y_j K((x_i - x_j)/h)/[\sum_{j=1}^{n} K((x_i - x_j)/h)]$, where $K(\cdot)$ is the kernel function and $h = h_n$ is the smoothing parameter. Replacing $E(y_i|x_i)$ by $\hat{E}(y_i|x_i)$ in (13) leads to

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} [y_i - \hat{E}(y_i|x_i)]^2 \Big/ \left[1 - \frac{2\lambda^2}{\pi(1+\lambda^2)}\right]. \quad (14)$$

Although $(\hat{E}(y_i|x_i) - E(y_i|x_i))$ has an order larger than $O_p(n^{-1/2})$—that is, the pointwise convergence rate of the kernel regression estimator is slower than the rate of $n^{-1/2}$—the average quantity $\{n^{-1}\sum_{i=1}^{n} [y_i - \hat{E}(y_i|x_i)]^2\} - \{n^{-1}\sum_{i=1}^{n} [y_i - E(y_i|x_i)]^2\} = O_p(n^{-1/2})$ under very weak conditions; see Härdle and Stoker (1989) and Fan and Li (1992) for details on how to use trimming and linearization techniques to obtain this result. For expositional simplicity, we omit the trimming factor in $\hat{E}(y_i|x_i)$. Thus we have, under quite weak conditions, $\hat{\sigma}^2 - \tilde{\sigma}^2 = O_p(n^{-1/2})$ and hence $\hat{\sigma}^2 - \sigma^2 = O_p(n^{-1/2})$.

Now we can substitute $\hat{\sigma}^2$ and $\hat{E}(y_i|x_i)$ into the log-likelihood function, which is then maximized over a single parameter λ. In accordance with Aigner et al. (1976), we call the resulting estimators of λ and σ^2 semiparametric pseudolikelihood estimators. In view of the $\sqrt{n}$-consistency of the kernel residual sum of squares discussed previously, one would expect the accuracy of our semiparametric pseudolikelihood estimators based on kernel estimation of $E(y_i|x_i)$ to be similar to those based on parametric estimation of a correctly specified $g_0(x_i, \beta)$, at least when the sample size n is large. The Monte Carlo results in Section 2 indicate that this is indeed the case.

We summarize the semiparametric stochastic frontier estimation method:

Step 1: Compute the kernel estimate of the conditional expectation $E(y_i|x_i)$ by $\hat{E}(y_i|x_i) = \sum_{j=1}^{n} y_j K((x_i -$

$x_j)/h)/\sum_{j=1}^n K((x_i - x_j)/h)$ and concentrate out σ^2 as in (14) [an estimate of σ^2 is obtained later from (14) after the semiparametric pseudolikelihood estimate of λ is found in Step 2].

Step 2: Maximize the concentrated log-likelihood function $\ln l(\lambda)$ over the single parameter λ, where $\ln l(\lambda)$ is the same as in (5) except that ε_i is replaced by $\hat{\varepsilon}_i = y_i - \hat{E}(y_i|x_i) - \mu(\hat{\sigma}^2, \lambda)$ and σ is replaced by $\hat{\sigma}$ given by the square root of (14); that is,

$$\max_\lambda \ln l(\lambda)$$

$$= \max_\lambda \left\{ -n \ln \hat{\sigma} + \sum_{i=1}^n \ln\left[1 - \Phi\left(\frac{\hat{\varepsilon}_i}{\hat{\sigma}}\lambda\right)\right] - \frac{1}{2\hat{\sigma}^2}\sum_{i=1}^n \hat{\varepsilon}_i^2 \right\}$$

with $\hat{\varepsilon}_i = y_i - \hat{E}(y_i|x_i) - \mu(\hat{\sigma}^2, \lambda)$ and

$$\hat{\sigma} = \left\{ \frac{1}{n}\sum_{i=1}^n [y_i - \hat{E}(y_i|x_i)]^2 \bigg/ \left[1 - \frac{2\lambda^2}{\pi(1+\lambda^2)}\right] \right\}^{1/2}. \quad (15)$$

Let $\hat{\lambda}$ denote the resulting pseudolikelihood estimate of λ—that is, the solution of (15). Substituting $\hat{\lambda}$ into (14), we obtain the corresponding estimate of σ^2 and denote it by $\hat{\sigma}^2$ with slight abuse of notation. The nonparametric production frontier $g(x_i)$ can be consistently estimated by $\hat{g}(x_i) = \hat{E}(y_i|x_i) + \hat{\mu}, i = 1, \ldots, n$, where $\hat{\mu} = \sqrt{2}\hat{\sigma}\hat{\lambda}/[\pi(1+\hat{\lambda}^2)]^{1/2}$. Although $\hat{\sigma}^2$ and $\hat{\lambda}$ converge to σ^2 and λ, respectively, at rate $n^{-1/2}$, the estimator $\hat{g}(x_i)$ converges to $g(x_i)$ for each i at a slower rate than $n^{-1/2}$.

2. MONTE CARLO EXPERIMENTS

To study the finite-sample behavior of the proposed pseudolikelihood estimators, we conducted some Monte Carlo experiments using the following models:

$$\text{Model 1:} \quad y_i = 1 + \varepsilon_i,$$

$$\text{Model 2:} \quad y_i = 1 + x_i + \varepsilon_i,$$

$$\text{Model 3:} \quad y_i = 1 + \ln(1 + x_i) + \varepsilon_i,$$

$$\text{Model 4:} \quad y_i = 1 - \frac{1}{1 + x_i} + \varepsilon_i,$$

where $\varepsilon_i = v_i - u_i, u_i$ is iid $|N(0, \sigma_u^2)|$, and v_i is iid $N(0, \sigma_v^2)$ for $i = 1, \ldots, n$. The single regressor, $x_i, i = 1, \ldots, n$, in Models 2, 3, and 4 is iid χ_1^2. In computing $\hat{E}(y_i|x_i)$, we choose a normal kernel function, and the smoothing parameter h is chosen according to $h = x_{\mathrm{sd}} n^{-1/(p+4)} = x_{\mathrm{sd}} n^{-1/5}$, where x_{sd} denotes the sample standard deviation of $x_i, i = 1, \ldots, n$.

The estimation method is discussed in Section 3. After replacing $E(y_i|x_i)$ by its kernel estimate $\hat{E}(y_i|x_i)$ and σ^2 by $\hat{\sigma}^2$ given by (14), we maximize the log-likelihood function as given in (15) with respect to the single parameter λ. We use a numerical search procedure to find the local maximum of (15) as follows: We first evaluate the likelihood function at the point $\lambda_0 = 0$; then we increase λ by a positive increment, say $d\lambda_1$, and evaluate the likelihood function at the new point $\lambda_1 = \lambda_0 + d\lambda_1$. If the latter has a larger likelihood value, we increase λ_1 by another $d\lambda_1$ and compare the likelihood function at λ_1 and $\lambda_2 = \lambda_1 + d\lambda_1$. We continue this procedure until $\ln l(\lambda_i) \le \ln l(\lambda_{i-1})$. We then let $d\lambda_2 = -(d\lambda_1)/2$; that is, we search backward with half the increment. This procedure is repeated until a convergent value of λ is found. If the convergent value of λ is negative, we replace it by 0. The maximum search value of λ is 1,000. In all the Monte Carlo experiments reported in this section, this maximum value of λ is never reached; that is, λ always converges at a value less than 1,000. Or in other words, there is no type 2 failure [type 2 failure occurs if $\lambda = \infty$; see Olson, Schmidt, and Waldman (1980)].

We have two objectives in mind in carrying out this Monte Carlo study. First, because the main motivation for our semiparametric pseudolikelihood estimation strategy is to avoid misspecification bias that may arise from assuming an incorrect parametric functional form for the production frontier function, we want to see if our pseudolikelihood estimators $\hat{\sigma}^2$ and $\hat{\lambda}$ are robust to functional-form specification (linear and nonlinear) and how well they do for moderate samples in comparison with the corresponding parametric estimators using correctly specified functional forms. The second objective is to investigate the effect of sample size and that of the relative ratio of the variance components λ on $\hat{\sigma}^2$ and $\hat{\lambda}$. We showed in Section 1 that our estimators are $\sqrt{n}$-consistent in spite of the fact that a nonparametric kernel regression estimator is used. For a finite sample, however, the slow convergence rate of $\hat{E}(y_i|x_i)$ may affect the performance of $\hat{\sigma}^2$ and $\hat{\lambda}$. It is therefore of interest to see how large an n is needed for our estimators to have similar accuracy in terms of mean squared error (MSE) as their parametric counterparts.

We investigate the first objective by estimating σ^2 and λ in Models 1–4 using our semiparametric pseudolikelihood estimators proposed in Section 1 for different values of σ^2 and λ. In particular, we use the same sets of values of (σ^2, λ) as listed in table 2 of Aigner et al. (1977) to facilitate comparison; that is, $(\sigma^2, \lambda) = (1.88, 1.66), (1.63, 1.24), (1.35, .83)$.

The results are given in Tables 1–4. The sample size is 100, and the number of replications is 1,000 for all cases considered.

Table 1 reports the results for Model 1, an intercept-only model. Note that, for Model 1, $\hat{E}(y_i|x_i) = \bar{y}$ $(= n^{-1} \sum_{i=1}^n y_i)$; that is, the nonparametric estimator reduces to the parametric estimator. We see that the results are in general

Table 1. $y_i = 1 + \varepsilon_i$

	$(\sigma^2, \lambda) = (1.88, 1.66)$		$(\sigma^2, \lambda) = (1.63, 1.24)$		$(\sigma^2, \lambda) = (1.35, .83)$	
	MSE	Bias	MSE	Bias	MSE	Bias
σ^2	.3025	−.0838	.2740	−.0701	.2090	.0061
λ	.9405	.0116	.8266	−.0611	.6787	−.0459
σ_u^2	.5250	−.1112	.5080	−.0882	.4139	.0329
σ_v^2	.0525	.0274	.0614	.0181	.0638	−.0269

Table 2. $y_i = 1 + x_i + \varepsilon_i$

	$(\sigma^2, \lambda) = (1.88, 1.66)$		$(\sigma^2, \lambda) = (1.63, 1.24)$		$(\sigma^2, \lambda) = (1.35, .83)$	
	MSE	Bias	MSE	Bias	MSE	Bias
σ^2	.2975	−.1533	.2597	−.1073	.1978	−.0018
λ	.8006	−.1163	.7306	−.1074	.6227	−.0372
σ_u^2	.5196	−.2035	.4782	−.1298	.3877	.0374
σ_v^2	.0528	.0502	.0585	.0226	.0607	−.0392

Table 3. $y_i = 1 + \ln(1 + x_i) + \varepsilon_i$

	$(\sigma^2, \lambda) = (1.88, 1.66)$		$(\sigma^2, \lambda) = (1.63, 1.24)$		$(\sigma^2, \lambda) = (1.35, .83)$	
	MSE	Bias	MSE	Bias	MSE	Bias
σ^2	.3010	−.1883	.2599	−.1455	.1884	−.0425
λ	.8119	−.0932	.7476	−.1051	.6363	−.0461
σ_u^2	.5066	−.2191	.4629	−.1505	.3637	.0129
σ_v^2	.0480	.0309	.0548	.0051	.0593	−.0554

comparable with those of Aigner et al. (1977). Our estimators seem to be more stable than theirs in the sense that, as the true variance parameters change, the change in MSE is smaller than that of Aigner et al. (1977).

Table 2 gives the results of a linear specification of a frontier function. One can see that, by adding an extra regressor, the estimates of the variance components basically remain the same as in the first model in terms of the MSE criterion. Moreover, this result does not depend on whether the extra regressor is orthogonal to the intercept as opposed to the result of Olson et al. (1980).

Tables 3 and 4 report the results for the nonlinear Models 3 and 4. Comparing these results to Table 2, one can see that the results are quite similar to the linear production-frontier case. This shows that our pseudolikelihood estimators of the variance components based on nonparametric estimation of the production frontier are robust to the functional-form specification of the frontier as promised by the nonparametric estimation approach.

The preceding results are only for $n = 100$ and the variance ratio λ between .83 and 1.66. Tables 5 and 6 report the performance of the pseudolikelihood estimators for different sample sizes and for a wide range of the variance ratio λ.

To investigate the effect of sample size, following Olson et al. (1980), we fix $\lambda = 1, \sigma_\varepsilon^2 = \sigma_v^2 + (\pi - 2)\sigma_v^2/\pi = 1$ and choose $n = 25, 50, 100, 200, 400, 800$. To investigate the effect of the ratio of the variance component, λ, also following Olson et al. (1980), we fix $n = 50$ and choose $\lambda = 10^{-1}, 10^{-3/4}, 10^{-2/4}, 10^{-1/4}, 1, 10^{1/4}, 10^{2/4}, 10^{3/4}, 10$. Because our approach gives quite similar results for all the four different models, we only report the results of Model 2 (i.e., $y_i = 1 + x_i + \varepsilon_i$). These are given in Tables 5 and 6.

Table 5 uses the same variance component parameters as table 1 of Olson et al. (1980), except that their model was $y_i = \alpha + \varepsilon_i$. The number of replications is 1,000 for $n = 25$, 250 for $n = 800$, and 500 for the other four cases.

Olson et al. (1980) also reported the percentage of type 1 ($\lambda < 0$) and type 2 ($\lambda = \infty$) failure of the corrected ordinary least squares (COLS) estimators. We found percentages of type 1 failure in our pseudolikelihood estimation approach similar to those of Olson et al. (1980); however, no type 2 failure occurred in all the Monte Carlo experiments reported here. This is because we maximize a concentrated log-likelihood function with respect to a single parameter λ, and there is always a finite value of λ that maximizes the likelihood function.

Comparing our results to those of Olson et al. (1980), one can see that the finite-sample performance of our nonparametric-based estimators is quite satisfactory in the sense that they give quite similar results to those of Olson et al. (1980) using maximum likelihood estimation (MLE) and COLS estimation technique (both are based on a parametric specification of the frontier). In addition, the MSE of each of our estimators decreases rapidly as sample size n increases, which supports our arguments in Section 1 that the semiparametric pseudolikelihood estimators have the same rate of convergence $(n^{-1/2})$ as the parametric-based estimators such as maximum likelihood or COLS estimators.

Table 6 uses the same parameter values as table 3 of Olson et al. (1980). The number of replications is 1,000 for all cases considered. The results are similar to those of Olson et al. (1980) for most cases. This shows that, for $n = 50$, for a wide range of λ $(.1 \leq \lambda \leq 10)$, our pseudolikelihood estimators based on a nonparametric estimation of the frontier give results of the variance components quite similar to those based on a parametric estimation of the frontier. Combining the results of Table 5 and Table 6, we see that our pseudolikelihood estimation method compares well to the parametric MLE or COLS estimation method for different sample sizes and variance ratios.

3. AN EMPIRICAL APPLICATION

In this section, we present an empirical application of the proposed semiparametric estimation method using data on 471 Quebec dairy farms. The dependent variable is the production of milk per cow, and the independent variables are (a) the quantity of forage consumed per cow, (b) the quantity of grain and concentrate per cow, (c) capital stock (building, machinery, and dairy equipment) per cow, and (d) the number of labor-person units per cow. Romain and Lambert (1995) used this data to estimate a parametric stochastic production frontier to assess technical efficiency of Quebec dairy farms. Further details on the source of the data and definitions of the variables were provided in their study.

We considered two specifications for the production function: The first one is the nonparametric regression function

Table 4. $y_i = 1 - (1 + x_i)^{-1} + \varepsilon_i$

	$(\sigma^2, \lambda) = (1.88, 1.66)$		$(\sigma^2, \lambda) = (1.63, 1.24)$		$(\sigma^2, \lambda) = (1.35, .83)$	
	MSE	Bias	MSE	Bias	MSE	Bias
σ^2	.3028	−.1978	.2612	−.1567	.1860	−.0544
λ	.8249	−.0811	.7620	−.1029	.6422	−.0457
σ_u^2	.5026	−.2213	.4593	−.1553	.3564	.0070
σ_v^2	.0467	.0235	.0539	−.0014	.0592	−.0615

Table 5. $y_i = 1 + x_i + \varepsilon_i$

	σ_u^2			σ_v^2			σ^2		
n	Bias	Var	MSE	Bias	Var	MSE	Bias	Var	MSE
25	.1319	1.0694	1.0868	−.0219	.1593	.1597	.1100	.6212	.6333
50	.0848	.6887	.6959	−.0662	.0947	.0991	.0186	.3968	.3972
100	−.0024	.4358	.4358	−.0157	.0619	.0621	−.0182	.2271	.2275
200	−.0534	.2963	.2992	.0129	.0406	.0408	−.0406	.1462	.1478
400	−.0696	.1921	.1970	.0207	.0255	.0260	−.0489	.0914	.0938
800	−.0738	.1034	.1088	.0260	.0138	.0144	−.0478	.0482	.0505

as described in (1) and the second one is a linear frontier function. We used the pseudo-MLE method as described in (14) and (15). For the nonparametric kernel estimation, a product normal kernel function was used with smoothing parameter chosen by $h_l = cx_{l,\mathrm{sd}}n^{-/(4+p)}$ (see Härdle 1990), where c is a constant, and $x_{l,\mathrm{sd}}$ is the sample standard deviation of x_l $(l = 1, 2, 3, 4)$. We used $c = .8, 1.0, 1.2$. Because the results are quite similar for these different choices of c, we will only report the results for $c = 1$ to save space. The estimation results are $\hat{\sigma}^2 = .373$, $\hat{\lambda} = 1.250$, the sum of squared residuals is .224, and the (goodness of fit) R square is .670. In contrast, for the linear specification, the estimation results are $\hat{\sigma}^2 = .743$, $\hat{\lambda} = 1.478$, the sum of squared residuals is .414, and the R square is .385. We see that the semiparametric estimation gives estimation results quite different from the parametric approach, suggesting that the frontier regression function is likely to be nonlinear in x.

With little prior knowledge about the functional form, we will use the estimation results from the semiparametric pseudolikelihood approach to analyze the efficiency levels of the Quebec dairy farms. Once the estimates of σ^2 and λ are obtained, farm-level efficiency can be estimated by the summary measure $E(u_i)$, which for obvious reasons is not particularly satisfactory. Jondrow, Lovell, Materov, and Schmidt (1982) showed that the conditional distribution of u_i given ε_i is an $N(\mu_i^*, \sigma^{*2})$ random variable truncated at 0, where $\mu_i^* = -\varepsilon_i\sigma_u^2/\sigma^2$ and $\sigma^{*2} = \sigma_u^2\sigma_v^2/\sigma^2$. They suggested more appropriate firm-specific efficiency estimates—the conditional expectation of u_i given ε_i,

$$E[u_i|\varepsilon_i] = \frac{\sigma\lambda}{1+\lambda^2}\left[\frac{\phi(\varepsilon_i\lambda/\sigma)}{1-\Phi(\varepsilon_i\lambda/\sigma)} - \frac{\varepsilon_i\lambda}{\sigma}\right], \quad (16)$$

and the mode of the conditional distribution, $M(u_i|\varepsilon_i) = \varepsilon_i(\sigma_u^2/\sigma^2)$ for $\varepsilon_i \le 0$ and $M(u_i|\varepsilon_i) = 0$ if $\varepsilon_i > 0$. Similarly to Jondrow et al. (1982), we will mainly use $E(u_i|\varepsilon_i)$ and $M(u_i|\varepsilon_i)$ to discuss the efficiency of these Quebec farms. We also computed the technical-efficiency index defined by $y_i/\hat{g}(x_i)$, where $\hat{g}(x_i) = \hat{E}(y_i|x_i) + \hat{\mu}$ is the estimated frontier. Note that, for cross-section data, only descriptive comments on the efficiencies are allowed; see Simar (1992) for more discussion on this. See also Schmidt and Sickles (1984) and Simar (1992) on how to analyze a firm's efficiencies when panel data are available. We will omit the subscript i to simplify the notations. We do not present the estimation results for all 471 observations, but rather we give some summarizing results:

1. The farm-specific efficiency estimates $\hat{E}(u|\varepsilon)$ range from .089 to 1.348 with a mean of .359. The estimated $\hat{M}(u|\varepsilon)$ is between 0 and 1.348 with a mean of .2534.
2. There are 15 observations of $\hat{\varepsilon}$ that are 2 to 3.2 standard deviations above the mean (the most technically efficient farms); the estimated values of $\hat{\varepsilon}$ and $\hat{E}(u|\varepsilon)$ are in the ranges of .601 to 1.157 and .089 to .13, respectively. All of them have $\hat{M}(u|\varepsilon) = 0$ and have technical-efficiency indexes larger than 1. The range for $\hat{\mu}^*$ is $-.705$ to $-.367$ $(\hat{\sigma}^* = .278)$, and the conditional distributions of u given ε are all within the 10% right tail of the corresponding $N(\hat{\mu}^*, \hat{\sigma}^{*2})$ normal distribution; that is, less than 10% of the area of $N(\hat{\mu}^*, \hat{\sigma}^{*2})$ lies to the right of 0.
3. The most negative 17 observations of $\hat{\varepsilon}$ (the most inefficient farms) are in the range of -2.21 to -1.39 (they are two to three standard deviations below the mean of $\hat{\varepsilon}$), which yields $\hat{\mu}^*$ in a range of .86 to 1.35 $(\hat{\sigma}^* = .278)$. Hence, the conditional distributions of u given ε are basically untruncated normal distributions. Moreover, all of these farms have technical-efficiency indexes less than .80.
4. There are 84 observations with $\hat{M}(u|\varepsilon) = 0$ (including the 15 observations in 2) and all of them have small

Table 6. $y_i = 1 + x_i + \varepsilon_i$

	σ_u^2			σ_v^2			σ^2		
λ	Bias	Var	MSE	Bias	Var	MSE	Bias	Var	MSE
.100	.5455	.5034	.8010	−.2415	.0864	.1448	.3041	.2742	.3667
.178	.5257	.5064	.7827	−.2342	.0864	.1413	.2915	.2763	.3612
.316	.4649	.5171	.7332	−.2121	.0872	.1322	.2529	.2819	.3458
.562	.3129	.5602	.6581	−.1568	.0907	.1153	.1561	.3041	.3285
1.000	.0236	.6590	.6595	−.0519	.0920	.0947	−.0283	.3685	.3693
1.778	−.2778	.8060	.8832	.0570	.0808	.0840	−.2209	.4925	.5413
3.162	−.4056	.7492	.9137	.1023	.0483	.0588	−.3033	.5273	.6193
5.623	−.4507	.6749	.8780	.1178	.0280	.0418	−.3329	.5243	.6352
10.00	−.4639	.6256	.8408	.1220	.0197	.0346	−.3419	.5105	.6274

values of $\hat{E}(u|\varepsilon)$. The technical-efficiency indexes for these farms are larger than .95. There are, however, 46 farms with technical-efficiency indexes less than .85 (including the 17 observations in 3). The efficient farms tended to feed a ration with a higher ratio of protein supplement and concentrate to forages. Efficiency also tended to increase with the labor-to-capital ratio, suggesting that inefficient farms may be overcapitalized for the size of their operation. Because the inefficient farms were generally small (herd size less than 30 cows) and given the lumpy nature of capital assets in dairy farming, the results suggest that a minimum size is required to achieve technical efficiencies. Efficiency was also positively associated with membership in farm-management clubs. The clubs provide information on technical innovations and are thus expected to increase the efficiency of participating farmers.

4. EXTENSIONS

In this section we present two extensions of the proposed pseudo-MLE method: (1) The frontier function is semiparametric and (2) the one-sided error is permitted to have a more flexible distribution than the half-normal distribution.

In the first extension, the frontier function $g(x_i)$ has a partially linear (semiparametric) functional form: $g(x_i) = w_i'\beta + m(z_i)$, where β is a $q \times 1$ unknown parameter, $x_i = (w_i', z_i')', w_i \in R^q$ and $z_i \in R^{p-q}$, and the functional form of $m(\cdot)$ is unknown. The model becomes

$$\begin{aligned} y_i &= g(x_i) + \varepsilon_i \equiv w_i'\beta + m(z_i) + \varepsilon_i \\ &= w_i'\beta + \theta(z_i) + \mu + \varepsilon_i \equiv E(y_i|x_i) + \mu + \varepsilon_i, \end{aligned} \quad (17)$$

where $\theta(z_i) = m(z_i) - \mu$. Note that one cannot treat β as a parametric component in computing the first-order conditions of the likelihood function. This is because $m(z_i)$ cannot be estimated [unlike $E(y_i|x_i)$] before an estimator for β is obtained. The estimation procedure is basically the *same* as the two-step procedure discussed earlier: $g(x_i)$ is replaced in the likelihood function by $[E(y_i|x_i) + \mu]$, but the F.O.C.'s for σ^2 and λ are the same as those given in (10) and (11), which leads to (14) and (15). The only difference is that the estimator for $E(y_i|x_i)$ is now obtained by

$$\hat{E}(y_i|x_i) = w_i'\hat{\beta} + \hat{\theta}(z_i), \quad (18)$$

where $\hat{\beta} = S^{-1}_{w-\hat{E}(w|z),w-\hat{E}(w|z)} S_{w-\hat{E}(w|z),y-\hat{E}(y|z)}$ is the semiparametric estimator of β based on (17), $\hat{\theta}(z_i) = \hat{E}(y_i|z_i) - \hat{E}(w_i|z_i)\hat{\beta}, \hat{E}(y_i|z_i) = \sum_{j=1}^n y_j K((z_i - z_j)/h)/\sum_{j=1}^n K((z_i - z_j)/h)$, and $\hat{E}(w_i|z_i) = \sum_{j=1}^n w_j K((z_i - z_j)/h)/\sum_{j=1}^n K((z_i - z_j)/h)$. Note that we followed Robinson (1988) in defining $S_{A,B} = n^{-1}\sum_{i=1}^n A_i B_i'$ for column vectors or scalars A_i, B_i. For details see Robinson (1988) or Fan, Li, and Stengos (1994). Once $\hat{\sigma}^2$ and $\hat{\lambda}$ are obtained [from (14) and (15)], we get $\hat{\mu} = \mu(\hat{\sigma}^2, \hat{\lambda})$, and the frontier function is estimated by $\hat{g}(x_i) = w_i'\hat{\beta} + \hat{\theta}(z_i) + \hat{\mu}$.

It is interesting to note that, if $g(x_i)$ is linear, then $\hat{E}(y_i|x_i)$ in (14) and (15) can be replaced by the least squares prediction of y_i given x_i. Thus, our estimation method becomes a pseudolikelihood estimation approach for the linear frontier model given by

$$y_i = x_i'\beta + \varepsilon_i \equiv x_i'\delta + (\varepsilon_i + \mu), \quad (19)$$

where $\beta = (\beta_0, \beta_1, \ldots, \beta_p)'$ is a $(p+1) \times 1$ vector of unknown coefficients and $\delta = (\delta_0, \delta_1, \ldots, \delta_p)' \equiv (\beta_0 - \mu, \beta_1, \ldots, \beta_{p-1})'$; that is, the nonzero mean of the disturbance is incorporated into the intercept term of δ. The new parameter vector δ can be consistently estimated by OLS regression, and the OLS prediction of y_i given x_i is $x_i'\hat{\delta}_{\text{ols}}$, where $\hat{\delta}_{\text{ols}}$ denotes the OLS estimator of δ. Again the estimation procedure is the *same* as the two-step procedure [see (14) and (15)] described earlier but with $\hat{E}(y_i|x_i) = x_i'\hat{\delta}_{\text{ols}}$.

The second extension assumes that the one-sided error u has a general truncated normal or a gamma distribution. Stevenson (1980) generalized the model of Aigner et al. (1977) to the case in which the one-sided error u (≥ 0) has a truncated normal distribution with the mode at ν. The joint pdf of ε $(= v - u)$ is

$$f(\varepsilon) = \sigma^{-1}\phi\left(\frac{\varepsilon - \nu}{\sigma}\right)\left[1 - \Phi\left(-\frac{\nu}{\sigma\lambda} - \frac{\varepsilon\lambda}{\sigma}\right)\right] \times \left[1 - \Phi\left(-\frac{\nu}{\sigma_u}\right)\right]^{-1},$$

where σ and λ are defined in the same way as before. Stevenson (1980) also gave the F.O.C. of $\partial \ln L/\partial\lambda, \partial \ln L/\partial\sigma^2$, and $\partial \ln L/\partial\nu$. In our case, the F.O.C. for λ, σ^2, and ν are the same as (11)–(13) of Stevenson (1980) except that the $x_i'\beta$ in his (11)–(13) is replaced by $g(x_i) = E(y_i|x_i) + E(u_i)$, where $E(u_i) \equiv \mu = (\nu a)/2 + [\sigma_u a/(2\pi)^{1/2}]\exp[-\nu^2/(2\sigma_u^2)], a = [1 - \Phi(-\nu/\sigma_u)]^{-1}$, and $\sigma_u^2 = \lambda^2\sigma^2/(1+\lambda^2)$. Then $E(y_i|x_i)$ is replaced by $\hat{E}(y_i|x_i)$, the nonparametric kernel estimator of $E(y_i|x_i)$. In this case, one can no longer concentrate out σ^2. Hence, one has to use a nonlinear optimization algorithm to solve the three F.O.C.'s simultaneously.

Greene (1990) considered the case in which the one-sided error u has a gamma distribution; that is, $u \sim G[\Theta, P]$. In this case, the F.O.C.'s become more complicated, and they involve numerical integrations. It is straightforward to apply the pseudo-MLE method proposed in this article to the case of a gamma-distributed u. One simply replaces $x_i'\beta$ by $E(y_i|x_i) + E(u_i) = E(y_i|x_i) + P/\Theta$ in the F.O.C.'s for Θ, P, and σ^2 as given in equations (35), (38), and (43) of Greene (1990). $E(y_i|x_i)$ is then replaced by $\hat{E}(y_i|x_i)$, the nonparametric kernel estimator of the conditional mean. The remaining steps, requiring some numerical methods to carry out the integrations and a nonlinear optimization, are similar to Greene's (1990). Alternatively, one may use a method-of-moments approach to estimate Θ, P, and σ^2. The advantage of using a method-of-moments approach is that it avoids numerical integrations and nonlinear optimization. The moments estimators using nonparametric regression residuals are $\sqrt{n}$-consistent under quite weak regularity conditions. To illustrate, let $\varepsilon_i = v_i - u_i$, where $v_i \sim N(0, \sigma_v^2)$ and $u_i \sim G[\Theta, P]$. Greene (1990) showed that $\mu \equiv E(u) = P/\Theta, E[(\varepsilon - E(\varepsilon))^2] = \sigma_v^2 + P/\Theta^2, E[(\varepsilon - E(\varepsilon))^3] = -2P/\Theta$, and $E[(\varepsilon - E[\varepsilon])^4] - 3\ \mathrm{var}[(\varepsilon]) = 6P/\Theta$. Hence, the method-of-moments estimators are given by $\hat{\Theta} = -3m_3/m_4^*, \hat{P} = -\hat{\Theta}^3 m_3/2$, and $\hat{\sigma}_v^2 = m_2 - \hat{P}/\hat{\Theta}^2$, where $m_l = n^{-1}\sum_{i=1}^n \hat{e}_i^l\ (l = 2, 3, 4), m_4^* = m_4 - 3m_2^2$, and $\hat{e}_i$ is the nonparametric residual from (1). Only the $\hat{\Theta}$ and $\hat{P}$ are needed to correct the nonzero mean of the error in the nonparametric conditional mean estimation $\hat{\mu} = \hat{P}/\hat{\Theta}$. The frontier function is then estimated by $\hat{g}(x_i) = \hat{E}(y_i|x_i) + \hat{\mu} = \hat{E}(y_i|x_i) + \hat{P}/\hat{\Theta}$.

5. CONCLUSION

In their editor's introduction to the recent special issue of *Journal of Econometrics* on frontier analysis, Lewin and Lovell (1990) pointed out "we seek two prominent directions in which future work might proceed. In the theory/modeling arena, ongoing work in pursuit of a convergence of the two techniques can be expected to expand, both by making DEA stochastic and by relaxing parametric restrictions in econometric models," (p. 5). This article moves in that direction. Specifically, we extend the linear stochastic frontier model proposed by Aigner et al. (1977) to a semiparametric frontier model in which the functional form of the production frontier is unknown. We proposed semiparametric pseudolikelihood estimators of σ^2 and λ based on kernel estimation of $E(y_i|x_i), i = 1, \ldots, n$. Therefore, our estimators are robust to possible misspecifications of the production frontier as opposed to the existing parametric estimators using correctly specified functional form of the frontier. Hence, this article can be viewed as a complement to those that avoid restrictive assumptions on the composed-error terms. Moreover, our estimators converge to their true parameter values at the same rate as their parametric counterparts. Monte Carlo results in Section 2 show that the semiparametric pseudolikelihood estimators of σ^2 and λ proposed in this article perform adequately in finite samples for at least the models considered. Furthermore, the results do not depend on the orthogonality of the extra regressors with the intercept as opposed to MLE and COLS observed by Olson et al. (1980). We presented an empirical application using data from Quebec dairy farms to illustrate the usefulness of the proposed estimation method in practice. We also discussed the extensions to the case in which the frontier function has a partially linear functional form and cases in which the one-sided error has a truncated normal or a gamma distribution.

One drawback of the article is the use of specific distributional assumptions on the composed-error terms. It may be possible to extend some of the tests of error distributions proposed by Lee (1983) and Schmidt and Lin (1984) to our model, the advantage of which is the robustness of the resulting tests to possible misspecifications in the functional form of the frontier. It may also be possible to extend the tests of Fan and Li (1992, in press) to consistent testing of the functional form of the frontier $g(\cdot)$. These topics will be taken up in future research.

ACKNOWLEDGMENTS

We thank George Tauchen, an associate editor and two anonymous referees for their helpful comments. The first author would like to thank the Natural Sciences and Engineering Research Council and Social Sciences and Humanities Research Council (SSHRC) of Canada for their support, the second author would like to acknowledge SSHRC of Canada for its research support, and the third author would like to acknowledge the financial aid of the Ontario Milk Marketing Board.

[Received May 1994. Revised December 1995.]

REFERENCES

Aigner, D. J., Amemiya, T., and Poirier, D. J. (1976), "On the Estimation of Production Frontier: Maximum Likelihood Estimation on the Parameters of a Discontinuous Density Function," *International Economic Review*, 17, 377–396.

Aigner, D. J., Lovell, C. A. K., and Schmidt, P. (1977), "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics*, 6, 21–37.

Battese, G. E., and Corra, G. S. (1977), "Estimation of a Production Frontier Model With Application to the Pastoral Zone of Eastern Australia," *Australian Journal of Agricultural Economics*, 21, 169–179.

Bauer, P. W. (1990), "Recent Developments in the Econometric Estimation of Frontiers," *Journal of Econometrics*, 46, 39–56.

Cornwell, C., Schmidt, P., and Sickles, R. C. (1990), "Production Frontiers With Cross-Sectional and Time-Series Variation in Efficiency Levels," *Journal of Econometrics*, 46, 185–200.

Deprins, D., Simar, L., and Tulkens, H. (1984), "Measuring Labor Inefficiencies in Post Offices," in *The Performance of Public Enterprises: Concepts and Measurements*, eds. M. Marchand, P. Pestieau, and H. Tulkens, Amsterdam: North-Holland, pp. 243–267.

Fan, Y., and Li, Q. (1992), "The Asymptotic Expansion of Kernel Sum of Square Residuals and Its Application in Hypotheses Testing," discussion paper, University of Windsor, Dept. of Economics.

——— (in press), "Consistent Model Specification Tests: Omitted Variables and Semiparametric Functional Forms," *Econometrica*, 64.

Fan, Y., Li, Q., and Stengos, T. (1994), "Root-N-Consistent Semiparametric Estimation With Conditional Heteroskedastic Disturbances," *Journal of Quantitative Economics*.

Farrell, M. J. (1957), "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society*, Ser. A, 120, 253–281.

Greene, W. H. (1990), "A Gamma-Distributed Stochastic Frontier Model," *Journal of Econometrics*, 46, 141–164.

Härdle, W. (1990), *Applied Non-parametric Regression*, Cambridge, U.K.: Cambridge University Press.

Härdle, W., and Stoker, T. M. (1989), "Investigating Smooth Multiple Regression by the Method of Average Derivatives," *Journal of the American Statistical Association*, 84, 986–995.

Hausman, J. A., and Taylor, W. E. (1981), "Panel Data and Unobservable Individual Effects," *Econometrica*, 49, 1377–1398.

Jondrow, J., Lovell, C. A. K., Materov, I. S., and Schmidt, P. (1982), "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics*, 19, 233–238.

Kopp, R. J., and Mullahy, J. (1990), "Moment-Based Estimation and Testing of Stochastic Frontier Models," *Journal of Econometrics*, 46, 165–184.

Lee, L. F. (1983), "A Test for Distributional Assumptions for the Stochastic Frontier Functions," *Journal of Econometrics*, 22, 245–267.

Lewin, A. Y., and Lovell, C. A. K. (1990), "Editors' Introduction," *Journal of Econometrics*, 46, 3–5.

Meeusen, W., and van den Broeck, J. (1977), "Efficiency Estimation From Cobb–Douglas Production Functions With Composed Error," *Interna-*

tional Economic Review, 18, 435–444.

Olson, J. A., Schmidt, P., and Waldman, D. M. (1980), "A Monte Carlo Study of Estimators of the Stochastic Frontier Production Function," *Journal of Econometrics*, 13, 67–82.

Park, B. U., and Simar, L. (1994), "Efficient Semi-parametric Estimation in a Stochastic Frontier Model," *Journal of the American Statistical Association*, 89, 929–936.

Robinson, P. (1988), "Root-N-Consistent Semi-parametric Regression," *Econometrica*, 156, 931–954.

Romain, R., and Lambert, R. (1995), "Economies of Size, Technical Efficiency, and the Cost of Production in Dairy Sectors of Quebec," *Canadian Journal of Agriculture Economics*, 37–55.

Schmidt, P., and Lin, T. (1984), "Simple Tests of Alternative Specifications in Stochastic Frontier Models," *Journal of Econometrics*, 24, 349–361.

Schmidt, P., and Sickles, R. (1984), "Production Frontier and Panel Data," *Journal of Business & Economic Statistics*, 2, 367–374.

Seiford, L. M., and Thrall, F. M. (1990), "Recent Developments in DEA: The Mathematical Programming Approach to Frontier Analysis," *Journal of Econometrics*, 46, 7–38.

Simar, L. (1992), "Estimating Efficiencies From Frontier Models With Panel Data: A Comparison of Parametric, Non-parametric, and Semi-parametric Methods With Bootstrapping," *Journal of Productivity Analysis*, 3, 171–203.

Stevenson, R. E. (1980), "Likelihood Functions for Generalized Stochastic Frontier Estimation," *Journal of Econometrics*, 13, 57–66.

Weinstein, M. A. (1964), "The Sum of Values From a Normal and a Truncated Normal Distribution," *Technometrics*, 6, 104–105 (with some additional material, 469–470).

Wilson, P. W. (1993), "Detecting Outliers in Deterministic Non-parametric Frontier Models With Multiple Outputs," *Journal of Business & Economic Statistics*, 11, 319–323.