THE KOREAN ECONOMIC REVIEW
Volume 23, Number 2, Winter 2007

PRODUCTIVITY GROWTH IN KOREAN AGRICULTURE, 1998-2002: ANALYSIS OF MULTI-OUTPUT PRODUCTION TECHNOLOGY WITH CORRECTION FOR SAMPLE SELECTION BIAS

HYEJUNG KANG *

This article analyzes the productivity growth in Korean agriculture with data on multiple crops, over the five years, from 1998 to 2002. Measurements are obtained from the estimation of a stochastic multi-output distance function. This study corrects for sample selection bias in the context of a translog functional form that appears when some farms produce only a subset of the potential outputs. The results find that technological change has caused a significant productivity change. Larger farms experience the highest rate of productivity growth by the greatest rate of scale effect. More human capital also leads higher productivity growth rates.

JEL Classification: C5, Q1
Keywords: Agriculture Productivity growth, Panel data, Stochastic frontier production function, Multi-output production technology, Sample selection bias

I. INTRODUCTION

Korea has devoted considerable effort in improving productivity, growth, and competitiveness. The agricultural sector has relied heavily on protectionist policies and has not experienced the rapid economic growth

Received for publication: April 7, 2007.　Revision accepted: July 11, 2007.
* Research associate, Korea Rural Economic Institute, 4-102, Hoegi-Dong, Dongdaemoon-Ku, Seoul, 130-710, Korea. Phone: 02-3299-4286. E-mail:kang@krei.re.kr. I wish to thank two anonymous referees for helpful comments and suggestions. Any remaining error, if any, is of course mine.

compared to other parts of the Korean economy. Productivity growth is now an important topic in Korean agriculture, particularly in the context of market opening in international competitiveness. Agricultural imports are expected to increase through further market liberalization, given current trade agreements, and negotiations, such as DDA and FTAs.

Many Korean commentaries have discussed the influence of trade reform on Korean agriculture (e.g., Choi et al., 2000; Kim and Lee, 2004), noting that domestic prices in Korea are far above import prices for many agricultural commodities. Productivity gains have been regarded by some agricultural economists and agricultural policy practitioners as the best way for Korean agriculture to compete with imports. Thus, considerable public efforts have been devoted to improving productivity (e.g., Kim, 1997; Lee, 1998).

In order to better understand the recent performance of these efforts, this paper analyzes the productivity growth in Korean agriculture with panel data on multiple crops over the five-year span from 1998 to 2002. Panel data provide more reliable evidence on Korean farms' performance because the data enables us to track the performance of each producer through a sequence of time periods.

There is no study for agricultural productivity growth in Korea, especially in the context of multiple crops. While panel data are also used in the existing literature, data in these studies come exclusively from livestock farms. For example, Brümmer et al. (2002) consider dairy farms for three European countries (Germany, The Netherlands, and Poland), over the period from 1991 to 1994 and with outputs of milk, meat and other products. Paul et al. (2000) use four outputs from New Zealand farms (wool, lamb, mutton or sheep, and beef or deer). The analysis of farms producing multiple crops from different regions over time yields indications of heterogeneous productivity changes in agricultural production patterns.

This paper applies a panel data production frontier model to measure productivity growth of farms in Korea. Specifically, this study estimates a stochastic multi-output distance function to accommodate multiple outputs and inputs within the frontier framework[1]. Panel data frontier

[1] Since our panel data do not include prices, productivity growth is measured using a primal

models allow the measurement of farm and time specific indices of technical efficiency changes (i.e., technical efficiency is allowed to vary across producers and through time for each producer). Furthermore, stochastic frontier econometric techniques explain deviations from best-practice productivity with a two-part error term including a statistical noise from measurement error and a technical inefficiency arising from farms not reaching the production frontier boundary. This contrasts with conventional econometric approaches that fit a function through the data assuming a normal error distribution and with nonparametric or deterministic econometric frontier approaches that limit statistical inference (Paul et al., 2000).

The methodology used to measure multi-output productivity growth and its components is derived from several sources; Kumbhakar and Lovell (2000) provide an analytical framework for estimation and decomposition of TFP growth using cost and profit frontier functions in the context of multi-output production technology. Brümmer et al. (2002) expand on existing literature by estimating and decomposing traditional TFP growth into technical change, technical efficiency, allocative efficiency (of inputs and outputs), and scale components, all in the context of primal multi-output technology.

This paper also provides evidence about the importance of sources leading overall productivity growth and heterogeneous adjustments of Korean farms based on the production technology available to farms. As well as measurement of productivity growth rates of farms, productivity growth is decomposed as technical change, technical efficiency change, and scale effects. Productivity growth rates are different by categories of farm size, operator human capital, and crops, which result from heterogeneous sources causing productivity growth across farm sizes, farm characteristics, and output composition. These issues are all vitally important to Asian countries, such as China, Japan, and Taiwan, that have experienced rapid growth and a lagging farming sector.

For the empirical implementation of the multi-output distance function, the translog functional form has advantages to allow for a variety of possible production relationships including non-constant returns to scale,

---

parametric method that does not require price data.

non-homothetic production, and non-constant elasticities of outputs and inputs. However, use of the translog functional form raises a problem of zero-valued observations, because taking logarithms of zero values is not possible. If the dropped zero observations from the sample are missing purely by chance, zero observation problems may not cause bias in the estimation of the production frontier. The decision to produce single outputs or multiple outputs, however, is made by individual farms, so those who are not producing multiple outputs constitute a self-selected sample and not a random one. Some studies get around these problems by using time-series or panel data without zero observations or omit the observations with zero values from the sample, while others have intentionally avoided use of translog functional forms and alternatively use the level-based specification of quadratic functions. For example, suppose a generalized linear transformation function has been specified (Diewert, 1973; Felthoven and Paul, 2004). Zero observations are also included in estimation by using a value of one or an arbitrarily small number greater than zero. The results of this method, however, are not independent of the variable's units of measurement. In addition, if the number of zero values is a significant proportion of the total number of sample observations, then the method may result in substantially biased parameter estimates. Accordingly, I develop an approach to deal with the sample selection bias caused by the problems of zero-valued observations in the translog distance function. This paper corrects for sample selection bias in the context of a translog functional form by utilizing an extended version of the Heckman selection model.

II. ANALYTICAL FRAMEWORK

An output-oriented distance function is employed to represent a multi-output production frontier. This approach has been used in the substantial literature on multi-output production frontier analysis (e.g., Lovell et al., 1994; Battese and Coelli, 1995; Paul et al, 2000; Brümmer et al., 2002).

The output distance function $D_o(x, y, t)$, suggested by Shephard (1970), is defined on the output set $P(x, t)$ as:

$$D_o(x, y, t) = \inf_{\delta}\{\delta > 0 : (y/\delta) \in P(x, t)\},$$
$$\text{for all } \ y \in R_+^M \ \text{ and } \ x \in R_+^K, \tag{1}$$

where $P(x,t)$ represents the feasible set of outputs that can be produced with the input vector x, given external production determinant t. Time, t, facilitates the calculation of technical change.

The output distance function indicates the reciprocal of the maximum proportional expansion of the output vector y, given x and t. The function also represents production technology, especially in the case of multiple-outputs. Since the output distance function $D_o(x, y, t)$ is defined in terms of the output set $P(x,t)$, satisfying certain properties, the corresponding output distance function satisfies the following properties: $D_o(x, y, t)$ is non-decreasing, convex, and linearly homogeneous in outputs and non-increasing and quasi-concave in inputs (Färe and Primont, 1995).

The distance function represents distance from the frontier. The distance takes on a value less than one if y is within the function; this also indicates the deviation of the farm from technologically "best-practice" production. The output-oriented measure of technical efficiency thus coincides with the output distance function (Kumbhakar and Lovell, 2000, p.50). The relationship is formulated as,

$$D_0(x, y, t)\exp(-u) = 1 \Leftrightarrow \ln D_0(x, y, t) - u = 0, \tag{2}$$

where u is non-negative, $D_o(x, y, t)$ is a multi-output production frontier and $\exp(-u)$ is the value of the output distance function, which is less than or equal to one. The exponential $-u$ is therefore often represented as the technical efficiency score (i.e., the efficiency of transforming inputs into outputs). Here, u is assumed to be dependent on time if technical efficiency changes over time.

Differentiating the latter expression of (2) with respect to time t and summing up the rate of change in TFP (i.e. $\dot{TFP} = \dfrac{d\ln TFP}{dt} = \sum_{m=1}^{M} R_m \dot{y}_m - \sum_{k=1}^{K} S_k \dot{x}_k$. Here, R_m and S_k denotes the observed revenue share of output y_m and cost share of input x_k, respectively) measured in the context of a multi-output, multi-input setting lead to the decomposition formula of productivity growth for multiple outputs as[2]:

$$\dot{TFP} = (\varepsilon - 1)\sum_{k=1}^{K} \frac{\varepsilon_{x_k}}{\varepsilon} \frac{d \ln x_k}{dt} - \frac{\partial \ln D_0(.)}{\partial t} + \frac{du}{dt}, \tag{3}$$

where $\frac{d \ln x_k}{dt}$ indicates the growth rate of x_k, $\frac{\partial \ln D_0(.)}{\partial \ln x_k}$ denotes as ε_{x_k}, and returns to scale (RTS) denote as $\varepsilon(= -\sum_{k=1}^{K} \varepsilon_{x_k})$. Equation (3) for TFP growth decomposes observable productivity growth into a scale effect (SE= $(\varepsilon - 1)\sum_{k=1}^{K} \frac{\varepsilon_{x_k}}{\varepsilon} \frac{d \ln x_k}{dt}$), a technical change effect (TC=$-\frac{\partial \ln D_0(.)}{\partial t}$), and a technical inefficiency change (TEC=$\frac{du}{dt}$). The sum of these three components is the total change in productivity[3].

To be consistent with the discrete form of the data, the components of the $\dot{TFP}$ in equation (3) are accordingly transformed into (4).

$$\begin{aligned} \ln TFP_{t+1} - \ln TFP_t = (\bar{\varepsilon} - 1)\sum_{k=1}^{K} \frac{\bar{\varepsilon}_{x_k}}{\bar{\varepsilon}} (\ln x_k^{t+1} - \ln x_k^t) \\ - \overline{D_o^t} + (u^{t+1} - u^t), \end{aligned} \tag{4}$$

where $\bar{\varepsilon}_{x_k} = \frac{1}{2}(\varepsilon_{x_k}^t + \varepsilon_{x_k}^{t+1})$, $\bar{\varepsilon} = \sum_{k=1}^{K} \bar{\varepsilon}_{x_k}$, and $\overline{D_o^t} = \frac{1}{2}\{\frac{\partial \ln D_o(x,y,t)}{\partial t} + \frac{\partial \ln D_o(x,y,t+1)}{\partial t}\}$.

If technical efficiency is time invariant, then the third component on the right-hand side of equation (4) drops out, and the productivity change is composed of a scale effect and a technical change. If technical

[2] For the detail, see Brümmer et al., 2002, pp.630-631.

[3] Since price information is unavailable in the panel data used, the allocative inefficiency components cannot be calculated empirically.

efficiency is time invariant and constant returns to scale (i.e., $\varepsilon = 1$) prevail, the first and the third components on the right-hand side of equation (4) both drop out, and productivity change consists solely of technical change (Kumbhakar and Lovell, 2000, p.285).

To decompose *TFP* growth in the context of the primal multi-output production technology as in equation (4), we need to measure the growth rates of inputs from the data (i.e., ($\ln x_k^{t+1} - \ln x_k^t$)). Furthermore, the elasticities of the distance function with respect to inputs and time are required to calculate the scale effect and the technical change (i.e., $\frac{\partial \ln D_0(.)}{\partial \ln x_k} = \varepsilon_k$ and $\frac{\partial \ln D_0(.)}{\partial t}$). Returns to scale (i.e., RTS: $\varepsilon(= -\sum_{k=1}^{K} \varepsilon_{x_k})$) are calculated as the negative sum of distance elasticities with respect to the inputs. The change in technical efficiency is obtained as the difference in the individual technical efficiency estimates from year to year. These calculations are based on the coefficients resulting from the estimation of a specified parametric production model.

III. DATA AND VARIABLES

This study relies primarily upon farm-level data, compiled by the Korean Ministry of Agriculture and Forestry in a national farm survey, for the period 1998 through 2002. The Ministry survey classified and reported statistics for approximately 2,900 randomly selected farm households, spanning nine provinces. The data tracked farm households with the same farm identification number through the five years of observation (1998-2002) to make a balanced panel data set. The resulting panel data set contains statistics for 2,450 farms across eight provinces[4].

For each farm household, data are aggregated into two outputs and four inputs. The outputs are rice and non-rice crops (including vegetables, fruits, and other crops)[5]. The inputs are land, labor, capital, and other inputs.

---

[4] The data used in this article exclude Jeju province (an island off the south coast of the peninsula). Less than one percent of farms in Jeju province - 0.007% - produce rice. This study also excludes livestock farms which tend to be specialized operations in Korea.

[5] Since rice is planted in more than 50 percent of cropland and generates about 50 percent of total crop revenue in the panel data, it is important to focus special attention on rice.

Land and labor are measured by quantities. Land is planted area and includes three types of cropland: paddy, upland, and orchard. Paddy refers to land primarily used for flood-irrigated rice, and upland area is other annual cropland. Labor is hours spent on farm work and includes both family labor and hired labor. Capital and other inputs are measured in value terms. Capital includes the average estimated replacement cost of structures, machinery depreciation, repairs, and leased farm equipment. Other inputs include expenditures on fertilizers, pesticides, fuel, electricity, seeds, and miscellaneous operating expenses. National level output and input-specific deflators were used to rescale those outputs and inputs that are collected in value terms, with 1998 being the base year. In this way, outputs and inputs become implicit quantities.

Descriptive statistics for the two outputs and four inputs are summarized in Table 1, including mean per farm household by year. The data confirm that farms in Korea are small, with an average landholding of 1.06 hectares per farm in the sample. The average farm has a part-time operator with about 1,000 total hours of labor used, including all family and hired labor. Labor use declined over the sample period, while usage of capital and cultivated land per farm increased steadily.

[Table 1] Summary Statistics for Aggregate Outputs and Inputs

	Rice (1,000won)	Non-rice crops (1,000won)	Land (ha)	Labor (hour)	Capital (1,000won)	Other inputs (1,000won)
Mean	7,558	8,964	1.06	1,038	3,616	4,104
	(9,085)	(14,051)	(1.03)	(813)	(3,885)	(4,860)
			Mean value per farm household			
1998	7,034	8,194	1.04	1,050	3,178	3,822
1999	7,490	9,194	1.06	1,068	3,413	4,240
2000	7,469	9,467	1.06	1,045	3,626	4,253
2001	8,264	9,366	1.07	1,037	3,830	4,226
2002	7,533	8,600	1.09	991	4,034	3,981

Note: Standard deviations are in parenthesis.

Non-rice crops denote all crops except rice such as vegetables, fruits, and other crops.

IV. EMPIRICAL IMPLEMETATION

Translog Multi-output Stochastic Distance Function

For empirical implementation, a functional form for the multi-output stochastic distance function first has to be chosen. This study employs the translog functional form that has been adopted widely in frontier studies (Lovell et al., 1994; Grosskopf et al., 1997; Paul et al., 2000; Brümmer et al., 2002). The translog function allows for a variety of possible production relationships including non-constant returns to scale, non-homothetic production, and non-constant elasticities of outputs and inputs.

A translog multi-output stochastic distance function with two outputs, four inputs, and time t is specified as:

$$\begin{aligned}\ln D_{oit}(y,x,t) &= \alpha_0 + \sum_{m=1}^{2}\alpha_m \ln y_{mit} + \sum_{k=1}^{4}\beta_k \ln x_{kit} \\ &+ \frac{1}{2}\sum_{l=1}^{2}\sum_{m=1}^{2}\alpha_{lm}\ln y_{lit}\ln y_{mit} + \frac{1}{2}\sum_{k=1}^{4}\sum_{j=1}^{4}\beta_{kj}\ln x_{kit}\ln x_{jit} \\ &+ \frac{1}{2}\sum_{m=1}^{2}\sum_{k=1}^{4}\gamma_{mk}\ln y_{mit}\ln x_{kit} + \delta_0 t + \frac{1}{2}\delta_{11}t^2 \\ &+ \sum_{m=1}^{2}\delta_{ym}\ln y_{mit}t + \sum_{k=1}^{4}\delta_{xk}\ln x_{kit}t\end{aligned}, \qquad (5)$$

where farms are indexed by subscript, i and time is indexed by subscript, t. D_{oit} denotes the output distance function measure, y_{mit} is a vector of outputs ($m=1$ for non-rice crops; $m=2$ for rice), x_{kit} is a vector of inputs ($k=1$ for land, $k=2$ for labor, $k=3$ for capital, $k=4$ for other inputs). Time t allows for possible shifts of the frontier over time and may reflect technical change or other systematic change over time.

Direct estimation of equation (5) by standard least squares or maximum likelihood techniques is not possible, since the dependent variable, $\ln D_{oit}$, is unobservable. The conversion of equation (5) into an estimable regression model can be accomplished by exploiting the fact that the output distance function is linear homogeneous in outputs. One way of imposing this restriction is to normalize the function by one of the outputs (e.g., Lovell et al., 1994; Paul et al., 2000; Brümmer et al., 2002). Rice, represented as y_{2it}, is chosen for the normalization, which leads to the following expression: $\ln D_{oit}(\frac{y_{it}}{y_{2it}}, x_{it}, t) = \ln \frac{1}{y_{2it}} D_{oit}(y_{it}, x_{it}, t)$.[6] Using this homogeneity restriction, replacing $\ln D_{oit}(y_{it}, x_{it}, t)$ with the technical inefficiency error, $-u_{it}$, and adding a random error term, v_{it}, the estimation of output distance function (5) can be written as:

$$-\ln y_{2it} = \alpha_0 + \alpha_1 \ln \frac{y_{1it}}{y_{2it}} + \sum_{k=1}^{4} \beta_k \ln x_{kit} + \frac{1}{2}\alpha_{11}\left(\ln \frac{y_{1it}}{y_{2it}}\right)^2$$
$$+\frac{1}{2}\sum_{k=1}^{4}\sum_{j=1}^{4} \beta_{kj} \ln x_{kit} \ln x_{jit} + \sum_{k=1}^{4} \gamma_{1k} \ln \frac{y_{1it}}{y_{2it}} \ln x_{ik} + \delta_0 t + \frac{1}{2}\delta_{11} t^2 \text{ , or}$$
$$+\delta_{y1} \ln \frac{y_{1it}}{y_{2it}} t + \sum_{k=1}^{4} \delta_{xk} \ln x_{ik} t + v_{it} - u_{it}$$
$$\ln y_{2it} = TL(x_{it}, y_{1it}/y_{2it}, t, \alpha, \beta, \gamma, \delta) + v_{it} - u_{it}, \tag{6}$$

where v_{it} is a random error term independently and identically distributed as $N(0, \sigma_v^2)$ (intended to capture events beyond the control of farmers), and u_{it} (intended to capture technical inefficiency in outputs) are assumed to vary over both farms and time periods. Battese and Coelli (1992) proposed the following specification of u_{it}.

$$u_{it} = \{\exp[-\eta(t-5)]\}u_i, \tag{7}$$

where the u_{it} are assumed to be independently distributed non-negative truncations of the $N(0, \sigma_u^2)$ distribution suggested by Stevenson (1980). Thus u_i^t decreases, remains constant or increases over time if η>0, η=0, or η<0. If producers improve their level of technical efficiency, then η

[6] Estimation of the ratio form of the output distance function, however, raises a problem since the model examines how an output variable expands, holding output composition constant. This specification imposes perfect complementarities of outputs when representing expansion. If cross terms for ratios are incorporated in the model, this problem can be solved to some extent. Another problem with the ratio form of normalization relates to endogeneity; the dependent variable appears on the right side as the denominator of ratios in the model. Even though the existing literature recognizes these endogeneity problems, they are still not ultimately resolved. For example, Brümmer et al. (2002) argue that these problems in the stochastic distance function are not likely to be more serious than in any production function type of study.

is positive. The model in (7) allows technical efficiency to change over time by adding just one more parameter, η, to be estimated, while it restricts technical efficiency to be monotonically increasing or decreasing over time.

The maximum likelihood estimation of model (6) with the specification in (7) provides estimators for $\alpha, \beta, \gamma, \delta$ and variance parameters, σ_u^2 and σ_v^2 [7]. Within the specification in (6), the production parameters required to measure the components of productivity growth and the distance elasticities with respect to two outputs, four inputs, time, and returns to scale can be obtained using the following equations:

$$\varepsilon_{y_{2it}} = \frac{\partial \ln D_{oit}(.)}{\partial \ln y_{2it}} = 1 - \alpha_1 - \alpha_{11} \ln\left(\frac{y_{1it}}{y_{2it}}\right) - \delta_{y_1} t - \sum_{k=1}^{4} \gamma_{1k} \ln x_{kit} \tag{8a}$$

$$\varepsilon_{y_{1it}} = \frac{\partial \ln D_{oit}(.)}{\partial \ln y_{1it}} = \alpha_1 + \alpha_{11} \ln\left(\frac{y_{1it}}{y_{2it}}\right) + \delta_{y_1} t + \sum_{k=1}^{4} \gamma_{1k} \ln x_{kit} \tag{8b}$$

$$\varepsilon_{x_{kit}} = \frac{\partial \ln D_{oit}(.)}{\partial \ln x_{kit}} = \beta_k + \sum_{j=1}^{4} \beta_{kj} \ln x_{jit} + \delta_{x_k} t + \gamma_{1k} \ln\left(\frac{y_{1it}}{y_{2it}}\right),$$

$$k = 1, 2, 3, \text{ and } 4 \tag{8c}$$

$$\varepsilon_{t_i} = \frac{\partial \ln D_{oit}(.)}{\partial t} = \delta_0 + \delta_{y_1} \ln\left(\frac{y_{1it}}{y_{2it}}\right) + \sum_{k=1}^{4} \delta_{x_k} \ln x_{kit} + \delta_{11} t \tag{8d}$$

$$\varepsilon_{it} = \sum_{k=1}^{K} \varepsilon_{x_{kit}} \tag{8e}$$

Two-Stage Estimation procedures

In the present sample, all farms do not produce both rice and non-rice crops in all years)[8]. When this paper estimates equation (6), use of the translog functional form thus raises a problem of zero-valued observations, because taking logarithms of zero values is not possible. If the dropped zero observations are missing purely by chance, zero observation problems may not cause bias in the estimation of the

---

[7] There are three methods to represent a technical efficiency in the context of panel data, that is the fixed effects model, the random effects model, and maximum likelihood method. For the detail, see Kumbhakar and Lovell, 2000.

[8] When input data are aggregated to the four variables, there are no zero-values in inputs. Thus the zero-observation problems in inputs are not considered in this study.

production frontier. The decision to produce single outputs or multiple outputs, however, is made by individual farms, so those who are not producing multiple outputs constitute a self-selected sample, not a random one. In order to deal with the sample selection bias caused by confining the observations to only those farms that produce positive outputs, this study exploits sample selection method incorporating two selection rules into the system. For the $i\,th$ farm in the entire sample size N, relevant equations are defined as:

$$\ln y_{2it} = TL(x_{it}, y_{1it} / y_{2it}, t, \alpha, \beta, \gamma, \delta) + u_{it} + v_{3it}\,; \tag{9a}$$

$$Rice_{it} = Z_{it}'\gamma_1 + v_{1it}\,; \tag{9b}$$

$$Nrice_{it} = Z_{it}'\gamma_2 + v_{2it}\,; \tag{9c}$$

$$D_{rice_{it}} = \begin{cases} 1 & if \quad Rice_{it} > 0 \\ 0 & if \quad Rice_{it} = 0 \end{cases}; \tag{9d}$$

$$D_{nrice_{it}} = \begin{cases} 1 & if \quad Nrice_{it} > 0 \\ 0 & if \quad Nrice_{it} = 0 \end{cases}; \tag{9e}$$

where (9a) is the multi-output stochastic distance function of primary interest. $Rice_{it}$, is a latent variable with associated indicator function $D_{rice_{it}}$, indicating whether a farm chooses to produce rice, and $Nrice_{it}$ is a latent variable with associated indicator function $D_{nrice_{it}}$, indicating whether a farm chooses to produce non-rice crops. The relationships, between $D_{rice_{it}}$ and $Rice_{it}$ and $D_{nrice_{it}}$ and $Nrice_{it}$, are shown in (9d) and (9e). Equations (9b) and (9c) are the reduced forms for the latent variables capturing two sample selection rules: both latent variables $Rice_{it}$ and $Nrice_{it}$ are assumed to be random functions of observed exogenous variables Z_{it} . γ_1 and γ_2 are vectors of unknown coefficients. Note that the sample selection is based on two indices and two criteria. Due to the log functional form, observations in (9a) are observed if and only of $Rice_{it} > 0$ and $Nrice_{it} > 0$.

The equation (9a) is likely to be influenced by the same unobserved farm characteristics that affect the two sample selection rules (9b) and (9c), leading to a non-zero covariance among the v_{mit}'s. The random error terms v_{mit} (m=1,2,3: 1 is rice equation, 2 is non-rice crop equation, and 3 is the multi-output distance function) thus are multivariate, normally distributed with mean zero and covariance matrix Σ as,

$$Cov(v_{1it}, v_{2it}, v_{3it}) = \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}.$$

Of course, random factors influencing the two selection rules in a farm are also not independent due to unobserved circumstances that could affect both production decisions. That is, the disturbances in the two selection equations (9b) and (9c) have a bivariate normal distribution with mean vector zero and $Cov(v_{1it}, v_{2it}) = \sigma_{12} = \rho$. It is assumed in bivariate probit models that $\sigma_1 = \sigma_2 = 1$, in order to estimate γ_1 and γ_2 (Hausman and Wise, 1978, p.411).

Due to the correlations among the v_{mit} 's operating through the covariance matrix, least squares or maximum likelihood estimation of β over the subsample corresponding to $D_{rice_{it}} = 1$ and $D_{nrice_{it}} = 1$ will generally lead to inconsistent estimates (Maddala, 1983; Heckman, 1979). The conditional expectation for the truncated random error term in (9a) is obtained as follows (Tunali, 1986; Vella, 1998; Khanna, 2001)[9]:

$$\begin{aligned} & E(v_{3it} \mid z_{it}, D_{rice_{it}} = 1, D_{nrice_{it}} = 1) \\ & = E(v_{3it} \mid v_{1it} > -z_{it}'\gamma_1, v_{2it} > -z_{it}'\gamma_2) \\ & = \sigma_{13} \left[\frac{\varphi(z_{it}'\gamma_1)\Phi(z_{it}'(\gamma_2 - \rho\gamma_1)/(1-\rho^2)^{1/2})}{\Phi^b(z_{it}'\gamma_1, z_{it}'\gamma_2; \rho)} \right] , \\ & \quad + \sigma_{23} \left[\frac{\varphi(z_{it}'\gamma_2)\Phi(z_{it}'(\gamma_1 - \rho\gamma_2)/(1-\rho^2)^{1/2})}{\Phi^b(z_{it}'\gamma_1, z_{it}'\gamma_2; \rho)} \right] \\ & = \sigma_{13}\lambda_{1it} + \sigma_{23}\lambda_{2it}, \end{aligned} \tag{10}$$

where ϕ is the probability density function of the standard normal distribution, Φ is the cumulative distribution function of the standard

[9] Tunali (1986) provides a good reference for the double-selection models in various situations. Khanna (2001) analyzes the sequential decision to adopt two site-specific technologies and the impact of adoption on nitrogen productivity by employing a double selectivity model to correct sample selection bias.

normal distribution, and Φ^b is the bivariate normal distribution. The two λ's constitute the double-selection analogs of the inverse Mill's ratio that arises in the context of single-selection.

This study employs two-step procedures to eliminate the nonzero conditional expectation of v_{3it}. The two-step procedure is to first estimate the unknown parameters γ_i ($i=1,2$) and ρ over the entire N observations by bivariate probit model, suggested by Hausman and Wise (1978), and then construct the two terms in brackets in (10). In other words, the two selection models (9b) and (9c) with indicator functions (9d) and (9e) are estimated simultaneously using the bivariate probit procedure to find $\gamma_i (i=1,2)$ and ρ. This approach recognizes that the same unobserved characteristics of a farm could influence the two decisions, and the bivariate probit model is therefore more efficient than univariate probit model analyzing each decision independently.

In the second stage, one can then consistently estimate the parameters by least square or maximum likelihood over the observations producing multiple outputs by including estimates of the above two additional terms, denoted λ_{1it} and λ_{2it}, as additional regressors in (9a). More precisely, the multi-output stochastic distance function over the subsample consisting of farms producing both rice and non-rice crops, with sample size n_1, is constructed as follows:

For subsample of observations with both $D_{rice_{it}} = 1$ and $D_{nrice_{it}} = 1$:

$$\ln y_{2it} = TL(x_{it}, y_{1it}/y_{2it}, t, \alpha, \beta, \gamma, \delta) + \delta_{13}\hat{\lambda}_{1it} + \delta_{23}\hat{\lambda}_{2it} + u_{it} + \xi_{it}, \quad i = 1,.....,n_1, \tag{11}$$

where ξ_{it} ($= v_{3it} - \sigma_{13}\hat{\lambda}_{1it} - \sigma_{23}\hat{\lambda}_{2it}$) are random disturbance terms with zero mean. The t-test on the null hypothesis of $\sigma_{13} = 0$ or $\sigma_{23} = 0$ represent tests of sample selectivity biases, under the maintained distributional assumptions. That is, the t-tests (or Lagrange Multiplier Test) for the coefficients on λ_{1it} and λ_{2it} are based on the correlation between the error in the primary equation and the errors from the two selection equations[10].

[10] The z matrix for the probits includes explanatory variables not found in the distance function itself. However, the z matrices in each of the two probits are identical. Therefore,

V. ESTIMATION RESULTS

The results first consider the preliminary step by presenting the results of the multi-output distance function with selection bias corrections[11]. Then, estimation results of productivity growth and its components are discussed.

Parameter Estimates and Distance Elasticities

Estimates of coefficients of the multi-output distance functions without and with the correction of selection bias are similar, but the value of the log-likelihood in the model with the correction of selection bias is larger than that in the other. In the distance function with selection model, about 70 percent of the parameters in the frontier distance function are statistically significant at the ten percent level or higher (see appendix 2). The Wald-Chi Square Test for significance of the regression rejects the null hypothesis that the coefficients of the explanatory variables are all zero at the one percent level. In Table 2, the two coefficients of λ_{ij} are statistically significant at the ten percent level, which confirms the importance of correcting sample selectivity bias. The variance parameter σ_u^2 which measures the relative importance of inefficiency is statistically significant at the one percent level. The other variance parameter σ_v^2 which indicates inherent randomness in production due to variations in weather and other conditions is statistically significant at the one percent level. The statistical significances of the two variance parameters confirm the importance of using the parametric stochastic approach to estimate the productivity growth of farms.

identification of the λ terms in (11) is achieved through non-linearity.

[11] The results of bivariate probit model are shown in the Appendix 1.

[Table 2] Selection and Variance Parameter Estimates

Variable		Estimate
λ_1	σ_{13}	-0.0734
		(0.0241)*
λ_2	σ_{23}	0.0497
		(0.0227)**
Variance parameters	σ_u	0.0390
		(0.0020)*
	σ_v	0.0532
		(0.0010)*

Note: Standard erros are in parenthesis.
*, **, and *** indicate significance at the 1%, 5% and 10% level.
'Nrice' refers to non-rice crop.

The coefficients of the distance function itself are not useful for interpretation. Table 3 thus presents an overview of the technogical properties of the estimated model based on the average distance elasticities with respects to outputs and inputs using the selection bias corrected model. The distance elasticies are the estimated frontier elasticity or the elasticity of best practice production with respect to the arguments in the function. The distance elasticies with respect to two outputs, four inputs, time, and returns to scale can be obtained using the equation (8a)-(8e).

The high average distance elasticity of rice simply reflects high share of rice production. Because of homogeneity constraint in outputs, the values of these elasticities for the two crops must sum to be one. Note that the distance elasticities for a "well-behaved" input are negative. The very high elasticity of land (around 70%) reflects the large contribution of land to production (high returns to land). The negative elasticity of time implies technical progress. The Korean farms had a high rate of technical progress (nearly 7 percent per year) during the study period. This estimated elasticity is similar to the finding in Kwon and Lee (2004). At the sample mean, relatively small increasing returns to scale are realized (around 1.07).

[Table 3] Average Distance Elasticities across Observations

Variable	Average distance elasticity*	Standard error
Non-rice crops	0.3929	(0.00185)
Rice	0.6071	(0.00185)
Land	-0.7034	(0.00161)
Labor	-0.1393	(0.00071)
Capital	-0.0362	(0.00026)
Other inputs	-0.1897	(0.00090)
Time	-0.0700	(0.00066)
Returns to scale (RTS)	1.0685	(0.00026)

Note: The values are the elasticity of the distance function with respect to the variables.

Productivity growth and its components

Using the estimated coefficients of the multi-output distance function with the correction of selection bias, production parameters needed to measure the components of productivity growth are calculated. Table 4 presents productivity growth rates calculated by estimation of the multi-output stochastic distance function with the correction of sample selection bias. A productivity index of positive or negative value indicates improvement or decline in productivity, respectively.

Productivity growth in the frontier model is decomposed into the three sources of growth; technical efficiency change (TEC) is attributable to improvements in individuals "catching up" with the frontier, technical change (TC) is attributable to a shift in the frontier, and scale effect (SE) reflecting change in scale economies.

Given estimation procedures generate a large subset of productivity growth rates for each of the 2,450 farms, it is necessary to summarize the results to facilitate presentation. To this end, several categories distinguish the average productivity growth rates by time period, farm size, farm operator's age, education, and major crop. Value in each category presents the average productivity growth rate for those farms within that category. T-tests for testing the null hypothesis that the sample mean is identical zero are performed. For technical changes and technical efficiency changes in all categories, the null hypotheses are all rejected at the one percent significance level, implying that the sample means are statistically different zero. However, the sample means for the scale effect in some categories are statistically not different zero.

The average growth rate for all years is about 2.4 percent. The decomposition results provide some insights into net growth; technical change (TC) is larger than technical efficiency change (TEC) or the scale effect (SE). Note that values of technical efficiency change and scale effect in all categories are all smaller than those of technical efficiency. Since this study uses the panel data of identical farms over the five year, the variations of technical efficiency and scale effect over time seem to be small.

At an annual level, productivity growth rates were highest between 1998 and 1999 and lowest from 2001 to 2002. Substantial productivity decline, as seen in the decomposition, between 2001 and 2002. It was due to negative technical change and scale effect overwhelming an increase from improved technical efficiency. The magnitude of such productivity changes, thus resulted in a negative net growth. Interestingly, the fastest and slowest years of productivity growth correspond to the highest and lowest rates of technical change. Note that the term of technical change compounds aggregate productivity shocks such as those due to weather variations with true technological change. Actually, technical change in 2001 to 2002 indicates a negative change, resulting from a poor harvest by bad weather in the period. Overall, productivity growth rates tend to show large variation over years. This tendency may be attributable to the monotonic assumption on the time-variant parameter of the one-side error term in the Frontier Model (Kwon and Lee, 2004).

The average productivity growth rate tends to rise as farm size gets larger. The average rate of growth for farmland operated in the range of 0 to 0.5ha is lowest since this class has the smallest rates of technical change and scale effect. Farms with farmland operated of more than 3.0ha have the highest growth rate, due to this class having the greatest scale effect.

Productivity growth rates across farm operator's age and education differ substantially. The average growth rates show that the productivity growth among younger farmers is greater than that of older operators. This is due to the substantial roles of technical change and scale effect.

Productivity growth for farmers with higher education is greater than that of farmers with lower education. This is due to a higher rate of technical change.

Finally, productivity measures based on farm's major crop indicate a higher rate of average productivity growth in non-rice dominant farms. This is due to this class having the greatest rates of technical change and scale effect. As discussed earlier, farm operators producing non-rice crops averagely have higher human capital, and this higher human capital may lead to technological innovation of non-rice dominant farms. Average productivity growth rate of non-rice dominant farms thus might be higher than that of rice dominant farms.

Note that empirical results are always dictated by the data used. It is important to understand the data in interpreting the results (Kwon and Lee, 2004). The data used tend to fluctuate considerably, beginning and ending with historic low and high productivity years. This implies that our productivity measures are based on a low productivity year, and the results must be interpreted in this context. A five-year period of panel data is relatively short to draw any convincing results on productivity growth. It is unlikely that high productivity growth calculated in this study can be sustained and is rather related to the specific data period.

[Table 4] Productivity Growth Rates and Decompositions

		Decomposition of TFP growth rates		
	TFP growth rates	Technical Efficiency Change(TEC)	Technical Change(TC)	Scale Effect(SE)
Mean for all years	0.0239	0.00165	0.0221	0.00014
	(38.56)*	(252.81)*	(37.39)*	(0.93)
Year mean				
1998-1999	0.0846	0.00153	0.0815	0.00158
	(177.28)*	(356.01)*	(304.94)*	(3.97)*
1999-2000	0.0432	0.00160	0.0417	-0.00016
	(102.14)*	(193.86)*	(153.19)*	(-0.52)
2000-2001	0.0039	0.00171	0.0024	-0.00019
	(9.45)*	(113.21)*	(7.79)*	(-0.69)
2001-2002	-0.0374	0.00178	-0.0385	-0.00065
	(-89.79)*	(94.14)*	(120.00)*	(-2.93)*
Means by farm size				
0-0.5 ha	0.0167	0.00170	0.0177	-0.00272
	(13.95)*	(134.18)*	(15.31)*	(-7.98)*

0.5-1.0 ha	0.0258	0.00162	0.0239	0.00032
	(25.27)*	(165.73)*	(24.36)*	(1.33)
1.0-2.0 ha	0.0268	0.00165	0.0235	0.00161
	(23.01)*	(124.10)*	(21.31)*	(5.95)*
2.0-3.0 ha	0.0276	0.00164	0.0233	0.00261
	(11.62)*	(59.59)*	(10.41)*	(5.70)*
Above 3.0ha	0.0321	0.00172	0.0246	0.00573
	(7.41)*	(34.20)*	(6.25)*	(4.55)*
Means by operator age				
Less 40 years	0.0256	0.00167	0.0227	0.00119
	(8.47)*	(46.63)*	(8.02)*	(1.30)
40-55 years	0.0244	0.00166	0.0222	0.00051
	(22.08)*	(133.15)*	(20.97)*	(1.91)**
55-65 years	0.0240	0.00164	0.0221	0.00028
	(24.56)*	(164.48)*	(23.70)*	(1.23)
Above 65 years	0.0228	0.00166	0.0219	-0.00074
	(18.05)*	(133.91)*	(18.22)*	(-2.17)**
Means by operator education				
0-5 years	0.0214	0.00174	0.0203	-0.00070
	(9.04)*	(55.11)*	(9.01)*	(-1.02)
5-9 years	0.0235	0.00165	0.0215	0.00035
	(26.87)*	(196.59)*	(25.91)*	(1.62)***
9-12 years	0.0243	0.00165	0.0226	0.00004
	(18.15)*	(111.17)*	(17.62)*	(0.11)
Above 12 years	0.0252	0.00165	0.0235	0.00005
	(18.94)*	(110.93)*	(18.42)*	(0.15)
Means by major crop				
Rice dominant farms	0.0225	0.00167	0.0208	0.00004
	(27.96)*	(187.68)*	(27.10)*	(0.21)
Non-rice dominant farms	0.0259	0.00163	0.0240	0.00029
	(26.73)*	(172.19)*	(25.94)*	(1.3)

Note: t-values from t-test of sample mean are in parenthesis

Farm size, measured by hectares of farmland operated, is separated into five categories: 0-0.5ha, 0.5-1.0ha, 1.0-2.0ha, 2.0-3.0ha, and above 3.0ha. Farm operator's age is divided by years into the following groups: less than 40, 40-54, 55-64, and above 65. Farm operator's years of education is divided by years into the following groups: less than 5, 5-9, 9-12, and above 12. Major crop produced is separated into two groups by share of crop receipts: rice dominant farms (i.e., farms with share of rice receipts greater than 50% of total gross farm receipts) and non-rice dominant farms (i.e., farms with share of non-rice receipts greater than 50% of total gross farm).

VI. SUMMARY AND CONCLUSION REMARKS

Despite the caution required in interpreting the results, we can provide some summary and insights. The results indicate that frontier shifts played an important role in productivity growth in Korean agriculture. The fastest and slowest years of productivity growth correspond to the highest and lowest rates of technical change. Larger farms experienced the fastest productivity growth attributed to the greatest rate of scale effect, which suggests farm consolidation is one source of the average productivity growth. The results also indicate that there are higher rates of average productivity growth for younger farmers and farmers with higher education. Based on the empirical analysis, we can conclude that technical innovation was most important for changes in productivity in Korean agriculture between the years of 1998 and 2002. This resulted from larger farm sizes and greater investment in human capital.

Appendix 1. Bivariate Probit Estimates of Parameters in Two Output Choice Equations

	Bivariate Probit model	
	Rice is observed	Non-rice crops are observed
Age	0.0350 (0.0123)*	0.0283 (0.0134)**
$(\text{Age})^2$	-0.00020 (0.00011)***	-0.00033 (0.00012)*
Education	0.0180 (0.0228)	-0.0507 (0.0251)**
$(\text{Education})^2$	-0.0030 (0.0013)**	0.0027 (0.0014)**
Farmland	0.4853 (0.0319)*	-0.6976 (0.0311)*
$(\text{Farmland})^2$	-0.0539 (0.0046)*	0.0516 (0.0045)*
Rent	0.5960 (0.0326)*	0.3007 (0.0360)*
Full-time	-0.1828 (0.0308)*	0.3760 (0.0321)*
Family size	-0.0441 (0.0221)**	0.4343 (0.0246)*
Kangwon	0.0920 (0.0541)***	0.2636 (0.0611)*
N. Chungchong	0.0796 (0.0546)	0.1913 (0.0612)*
S. Chungchong	0.3102 (0.0564)*	-0.2571 (0.0565)*
N. Choolla	0.2947 (0.0561)*	-0.3894 (0.0551)*
S. Choolla	0.4798 (0.0586)*	-0.0022 (0.0583)
N. Kyoungsang	-0.1402 (0.0537)*	0.3061 (0.0642)*
S. Kyoungsang	0.3560 (0.0579)*	-0.1650 (0.0590)*
Year 99	-0.0642 (0.0446)	0.1647 (0.0470)*
Year 00	-0.1360 (0.0440)*	0.0787 (0.0459)***
Year 01	-0.1238 (0.0443)*	-0.0254 (0.0450)
Year 02	-0.1760 (0.0440)*	-0.0118 (0.0450)
Constant	-1.5091 (0.3459)*	-0.3966 (0.3764)
Rho	-0.7262 (0.0223)*	
Log likelihood: -9,552		
Wald chi^2 (42): 2437.69*		

Note: Kyounggi and Year 98 are dropped as the reference province and year.
‘Rent’ is 1 if farm rents farmland, otherwise, 0.
‘Full-time’ is 1 if farm operator is full-time farmer, otherwise, 0.
‘Family size’ is the number of family.
Standard errors in parenthesis.
*, **, and *** indicate significance at the 1%, 5% and 10% level.

Appendix 2. Parameter Estimates of the Multi-Output Stochastic Distance Function with Correction of Sample Selection Bias

Variable		Estimate	Variable		Estimate
(Nrice/Rice)	α_1	-0.7817	(Nrice/Rice) × (Land)	γ_{11}	-0.1563
		(0.0549)*			(0.0083)*
Land	β_1	0.0047	(Nrice/Rice) × (Labor)	γ_{12}	0.0580
		(0.0972)			(0.0082)*
Labor	β_2	0.3906	(Nrice/Rice) × (Capital)	γ_{13}	0.0099
		(0.0980)*			(0.0050)**
Capital	β_3	-0.0492	(Nrice/Rice) × (Other)	γ_{14}	0.0878
		(0.0567)			(0.0067)*
Other	β_4	0.5601	Time	δ_0	0.0416
		(0.0758)*			(0.0284)
$(\text{Nrice/Rice})^2$	α_{11}	-0.1780	$(\text{Time})^2$	δ_{11}	-0.0404
		(0.0073)*			(0.0032)*
$(\text{Land})^2$	β_{11}	-0.2134	(Nrice/Rice) × (Time)	δ_{y1}	-0.0080
		(0.0209)*			(0.0023)*
(Land) × (Labor)	β_{12}	0.1009	(Land) × (Time)	δ_{x1}	-0.0145
		(0.0154)*			(0.0047)*
(Land) × (Capital)	β_{13}	0.0171	(Labor) × (Time)	δ_{x2}	0.0098
		(0.0101)*			(0.0044)**
(Land) × (Other)	β_{14}	0.0951	(Capital) × (Time)	δ_{x3}	-0.0055
		(0.0132)*			(0.0027)**
$(\text{Labor})^2$	β_{22}	-0.0851	(Other) × (Time)	δ_{x4}	0.0174
		(0.0203)*			(0.0039)*
(Labor) × (Capital	β_{23}	0.0219	λ_{1i}	σ_{13}	-0.0734
		(0.0090)**			(0.0241)*
(Labor) × (Other)	β_{24}	-0.0372	λ_{2i}	σ_{23}	0.0497
		(0.0124)*			(0.0227)**
$(\text{Capital})^2$	β_{33}	-0.0057	Variance parameters	σ_u	0.0390
		(0.0071)			(0.0020)*
(Capital) × (Other	β_{34}	-0.0115		σ_v	0.0532
		(0.0077)			(0.0010)*
$(\text{Other})^2$	β_{44}	-0.0651			
		(0.0138)*			
Log likelihood: -952.8					
Wald chi ^{2} (29): 22712.9*					

Note: *, **, and *** indicate significance at the 1%, 5% and 10% level.

References

Battese G. E. and T. J. Coelli. (1992), "Production Frontier Functions, Technical Efficiencies and Panel Data: With Application to Paddy Farmers in India." *Journal of Productivity Analysis*, Vol. 3, 153-169.

Battese G. E. and T. J. Coelli. (1995), "A model for technical inefficiency effects in a stochastic frontier production function for panel data." *Empirical Economics*, Vol. 20, No. 2, 325-32.

Brümmer, B., T. Glauben and G. Thijssen (2002), "Decomposition of productivity growth using distance functions: The case of dairy farms in three European countries." *American Journal of Agricultural Economics*, Vol. 84, No. 3, 628-44.

Choi, S. G., J. O. Lee, M. G. Eor, S. S. Im and J. B. Im (2000), " The impact of market opening of agricultural products on domestic agriculture." Consignment Research Report (C2000-46), Korea Rural Economic Institute.

Diewert, W. E. (1973), "Functional for profit and transformation functions." *Journal of Economic Theory*, Vol. 6, No. 3, 284-316.

Färe, R. and D. Primont (1995), *Multi-output production and duality: Theory and applications*: Kluwer Academic Publication.

Felthoven, R. G. and C. J. M. Paul (2004), "Multi-output, non-frontier primal measures of capacity and capacity utilization." *American Journal of Agricultural Economics*, Vol. 86, No. 3, 619-33.

Grosskopf, S., K. Hayes, L. Taylor and W. Weber (1997),"Budget constrained frontier measures of fiscal equality and efficiency in schooling." *The Review of Economics and Statistics*, Vol. 79, No. 1, 116-124.

Hausman, J. A. and D. A. Wise (1978), "A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences." *Econometrica*, Vol. 46, No. 6, 403-26.

Heckman, J. J. (1979), "Sample selection bias as a specification error." *Econometrica*, Vol. 47, No. 1, 153-161.

Khanna, M. (2001), "sequential adoption of site-specific technologies and its implications for Nitrogen productivity: A double selectivity model." *American Journal of Agricultural Economics*, Vol. 83, No. 1, 35-51.

Kim, H. S. (1997), "Evaluation and policy tasks of large-scale farming." *Korean Rural Economics*, Vol. 20.

Kim, J. H. and B. H. Lee (2004), *Methods of solving Korean agricultural and rural problems in an era of trade reform*. Research material (D189), Korea Rural Economic Institute.

Kumbhakar, S. C. and C. A. K. Lovell (2000), *Stochastic Frontier Analysis*: Cambridge University Press.

Kwon, O. S. and H. O. Lee (2004), "Productivity Improvement in Korean Rice Farming: Parametric and Nonparametric Analysis." *Australian Journal of Agricultural and Resource Economics*, Vol. 48, No. 2, 323-46.

Lee, J. H. (1998), *Transformation of Korean farm structure: From beginning to end*, Research series 21, Korea Rural Economic Institute.

Lovell, C. A. K., S. Richardson, P. Travers and L. L. Wood (1994), "Resources and functions: A new view of inequality in Australia." in W. Eichhorn, eds., *Models and Measurement of Welfare and Inequality*: Berlin: Springer-Verlag Press.

Maddala, G. S. (1983), *Limited dependent and qualitative variables in econometrics*. Econometric Society Monographs 3: Cambridge University Press.

Paul, C. J. M., W. E. Johnston and G. A. G. Frengley (2000), "Efficiency in New Zealand sheep and beef farming: The impacts of regulatory reform." *The Review of Economics and Statistics*, Vol. 82, No. 2, 325-37.

Shephard, R. W. (1970), *Theory of Cost and Production Functions*: Princeton University Press.

Stevenson, R. E. (1980), "Likelihood Functions for Generalized Stochastic Frontier Estimation." *Journal of Econometrics*, Vol. 13, 58-66.

Tunali, I. (1986), "A general structure for models of double-selection and an application to a joint migration/earnings process with remigration." *Research in Labor Economics*, Vol. 8, 235-82.

Vella, F. (1998), "Estimating models with sample selection bias: A survey." *The Journal of Human Resources*, Vol. 33, No. 1, 127-69.