Maximum Likelihood Estimation of Censored Stochastic Frontier Models: An Application to the Three-Stage DEA Method*

Wen-Jen Tsay
Institute of Economics
Academia Sinica

Cliff J. Huang
Department of Economics
Vanderbilt University

Tsu-Tan Fu
Institute of Economics
Academia Sinica

I-Lin Ho
The Institute of Physics
Academia Sinica

January 2009

* The research described in this paper was conducted in part while the second author was a visiting scholar at the Institute of Economics, Academia Sincia. The financial support from the National Research Council of Taiwan is also gratefully acknowledged. The usual disclaimer applies.

Abstract

This paper takes issues with the appropriateness of applying the stochastic frontier analysis (SFA) technique to account for environmental effects and statistical noise in the popular three-stage data envelopment analysis (DEA). A correctly specified SFA model with a censored dependent variable and the associated maximum likelihood estimation (MLE) are proposed. The simulations show that the finite sample performance of the proposed MLE of the censored SFA model is very promising. An empirical example of farmers' credit unions in Taiwan illustrates the comparison between the censored and standard SFA in accounting for environmental effects and statistical noise.

Key words: Three-stage data envelopment analysis, stochastic frontier analysis, censored stochastic frontier model

1. Introduction

The mixed approach of combining data envelopment analysis (DEA) with stochastic frontier analysis (SFA) has been widely used in accounting for environmental effects and statistical noise in the study of production efficiency since its introduction by Fried et al. (1999, 2002). In general, the approach consists of a three-stage analysis that begins with an oriented DEA to obtain input or output slacks. In the second-stage, SFA is used to identify the variation of DEA slacks attributable to environmental impacts and statistical noise. In the third-stage, DEA is applied to obtain pure production efficiency from the adjusted inputs or outputs after "purging" the environmental and statistical noise influence. Some recent empirical applications include Liu and Tone (2008) on the Japanese banking industry, Cordero-Ferrera et al. (2008) on Spanish education, Lee (2008) on global forest and paper companies, Margari et al. (2007) on Italian public transit, Chen et al. (2007) on Taiwan's farmer credit unions, Drake and Simper (2005) on U.K. police force efficiency, and Glass et al. (2006) on U.K. universities.

In this paper we take issue with the appropriateness of applying the standard SFA in the second-stage regressions to estimate the impacts of environmental effects and statistical noise on the first-stage input or output slacks. The second-stage SFA regressions take the general form, $y = f\left(X;\beta\right)+v+u$, where the non-negative $y \geq 0$ represents the input or output slacks obtained from the first-stage DEA, and X represents the observable environmental variables. The two-sided random error v represents statistical noise and $u \geq 0$ reflects managerial inefficiency. Since the dependent variable y in the second-stage SFA represents the input or output slacks obtained from the first-stage DEA, a considerable number of observations are often found to be at value 0[1]. Thus, the stochastic frontier regression is of the censored type rather than the standard (uncensored) SFA model. A censored SFA implies the truncation of the composite error $\left(v+u\right)$, i.e., y = 0 if $v+u \leq -f\left(X;\beta\right)$. Although the standard censored (Tobit) model (Tobin, 1958) without the one-sided error u, i.e., $y = f\left(X;\beta\right)+v$, has often been applied

[1] This result holds especially in situations where the number of decision making units (DMU) is small relative to the sum of the number of inputs and outputs. This situation is an inherent drawback of the linear programming method of DEA.

in efficiency analysis to regress the DEA efficiency scores on some environmental variables, no study exists, as far as we know, that correctly specifies and estimates the censored (Tobit) SFA regression. This issue is not trivial because the standard SFA estimates of the environmental effect $f\left(X;\beta\right)$, the noise *v*, and the inefficiency *u* are inconsistent. The specification error and the inconsistent estimates of the second-stage thus carry over to the third-stage DEA estimates of efficiency scores.

In this paper, we fill this gap in the literature by proposing the maximum likelihood estimation (MLE) of the censored SFA model. Since the model is censored, the MLE requires the computation of the cumulative distribution function (cdf) of the truncated composite error $\left(v+u\right)$ at point $-f\left(X;\beta\right)$ for the observations $y = 0$. To facilitate the computation of MLE, we first provide a closed-form formula for computing the cdf of the composite error. The proposed formula is both easy to implement and accurate. Furthermore, the Monte Carlo results show that the finite sample performance of the proposed MLE of the censored SFA models is very promising under various model specifications considered in this paper.

The remaining parts of this paper are arranged as follows: In section 2 we present a theorem that provides a closed-form formula for computing the cdf of the composite error needed for the maximization of the likelihood function. The accuracy of the proposed formula is examined via an empirical distribution of ten million random drawings of the composite error. In section 3, some Monte Carlo experiments are conducted to evaluate the proposed MLE of the censored SFA and the MLE of the standard (uncensored) SFA. An empirical example of farmers’ credit unions in Taiwan is given in section 4 to illustrate the comparison between the censored and standard SFA in accounting for environmental effects and statistical noise. Section 5 provides a summary and conclusions.

2. Maximum Likelihood Estimation of Censored Stochastic Frontier Regressions

Consider a standard linear stochastic frontier model,

$$y_i = x_i^{\mathrm{T}}\beta + \varepsilon_i, \qquad i = 1, 2, \ldots, n \tag{1}$$

where y_i and ε_i are the i^{th} observation on the dependent variable and the random error, respectively; x_i^{T} is a $1\times k$ vector of the i^{th} observation on the k regressors; and β is a $k\times 1$ vector of unknown parameters to be estimated. The composite error ε_i is specified as:

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

where the random errors v_i are independently and identically distributed (iid) as $N(0,\sigma_v^2)$, and the random errors u_i are the absolute values of the variables that are iid as $N^+(0,\sigma_u^2)$. All v_i's and u_i's are independent of each other, and are also independent of x_i. We follow the reparameterization of Aigner et al. (1977) in setting

$$\sigma^2 = \sigma_u^2 + \sigma_v^2, \text{ and } \lambda = \frac{\sigma_u}{\sigma_v}. \tag{3}$$

The log likelihood function of the model defined in (1)-(3) is shown to be

$$L_0 = \frac{n}{2}\ln\left(\frac{2}{\pi}\right) - n\ln(\sigma) + \sum_{i=1}^{n}\ln\left[\Phi\left(\frac{\lambda}{\sigma}\varepsilon_i\right)\right] - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^2, \tag{4}$$

where $\varepsilon_i = y_i - x_i^{\mathrm{T}}\beta$, and Φ is the cdf of $N(0,1)$. The MLE is obtained by the maximization of (4) with respect to the parameter $\left(\beta^{\mathrm{T}},\lambda,\sigma\right)$.

When the dependent variable is censored, the regression (1) becomes the censored SFA model, i.e.,

$$\begin{cases} y_i^* = x_i^{\mathrm{T}}\beta + \varepsilon_i, & i=1, 2, ..., n; \\ y_i = y_i^*, & \text{if } y_i^* > 0; \\ y_i = 0, & \text{if } y_i^* \le 0. \end{cases} \tag{5}$$

where $\varepsilon_i = v_i + u_i$. It is inappropriate to estimate the parameters of model (5) via the log likelihood function L_0 in (4) because of the presence of the censored dependent variable. Thus, L_0 is called the standard (uncensored) SFA log likelihood function. As in Amemiya (1985), the censored SFA log likelihood function for n independent observations of model (5) is:

$$L_1 = \sum_{y_i>0}\ln\left(f\left(\varepsilon_i\right)\right) + \sum_{y_i=0}\ln\left(F\left(-x_i^{\mathrm{T}}\beta\right)\right), \tag{6}$$

where $f(.)$ and $F(.)$ are the density and distribution function of the composite error $\varepsilon_i = v_i + u_i$, respectively. The first summation is over the observations for which $y_i > 0$ and the second summation is over the observations for which $y_i = 0$.

From the estimation point of view, the uncensored part in (6) is easy to compute because the density $f(\varepsilon_i)$ is well-known,

$$f(\varepsilon_i) = \frac{2}{\sigma}\phi\left(\frac{\varepsilon_i}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma}\varepsilon_i\right), \tag{7}$$

where $\phi(.)$ denotes the density function of $N(0,1)$. However, the difficulty in the maximization of (6) is in computing the censored part,

$$F\left(-x_i^{\mathrm{T}}\beta\right) = \int_{-\infty}^{-x_i^{\mathrm{T}}\beta} f(\varepsilon)d\varepsilon, \qquad \text{for } y_i = 0. \tag{8}$$

Alternatively, the above distribution function $F(.)$ can be expressed as:

$$F(Q) = \frac{2}{\sigma}I(Q), \tag{9}$$

where

$$I(Q) = \int_{-\infty}^{Q}\left(\int_{-\infty}^{a\varepsilon}\phi(\varsigma)d\varsigma\right)\phi(b\varepsilon)d\varepsilon \tag{10}$$

and $a = \dfrac{\lambda}{\sigma}$, $b = \dfrac{1}{\sigma}$, $Q = -x_i^{\mathrm{T}}\beta$.

In the following theorem, we derive an approximated formula $I_{app}(Q)$ for $I(Q)$ in (10). However, we first define two functions, an error function $erf(z)$ and a sign function $sign(z)$:

$$erf(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt = 2\int_0^{\sqrt{2}z}\phi(t)dt, \tag{11}$$

and

$$sign(z) = \begin{cases} 1, & \text{if } z > 0; \\ 0, & \text{if } z = 0; \\ -1, & \text{if } z < 0. \end{cases} \tag{12}$$

Theorem: Under (Q, a, b) $\in$ finite R, $a \geq 0$, and $b > 0$, $I(Q)$ in (10) can be approximated by $I_{app}(Q)$:

$$I_{app}(Q) = exp\left(\frac{a^2 c_1^2}{4b^2 - 4a^2 c_2}\right) \frac{1}{4\sqrt{b^2 - a^2 c_2}} \left[1 - erf\left(\frac{-a\ c_1 + \sqrt{2}Q\left(b^2 - a^2 c_2\right)\ sign(Q)}{2\sqrt{b^2 - a^2 c_2}}\right)\right]$$
$$+ \frac{erf\left(\frac{b\ Q}{\sqrt{2}}\right)}{2b} \frac{1 + sign(Q)}{2},$$

where $c_1 = -1.09500814703333$ and $c_2 = -0.75651138383854$.

The derivation of I_{app} is given in the appendix. Given I_{app}, the cdf $F(Q)$ in (9) can then be approximated by:

$$F_{app}(Q) = \frac{2}{\sigma} I_{app}(Q). \tag{13}$$

Basically, the proposed formula I_{app} involves the error function $erf(z)$, which can easily be computed with the standard statistical package. Accordingly, the computation of F_{app} is extremely straightforward in making the MLE of the censored SFA easy to implement.

Mathematically, the role of the two constants, c_1 and c_2, in the theorem is to ensure that the error function $erf(z)$ can be well approximated by another function, $g(z) = 1 - e^{c_1 z + c_2 z^2}$, for $z \geq 0$. The choice of c_1 and c_2 in the theorem is to make the two functions, $erf(z)$ and $g(z)$, as close to each other as possible. One possible method is to use the first-order Taylor expansion around $z = 1$ to obtain the values of c_1 and c_2. An alternative approach, as adopted in this paper, is to use the nonlinear least squares method to estimate c_1 and c_2 based on 500 equally spaced points within the interval $[0,5]$. The interval is chosen because $erf(0) = 0$ is the lower limit of the error function, and $erf(5) \approx 1 - 0.15 \times 10^{-8}$ is very close to the upper limit, $erf(\infty) = 1$. The estimated c_1 and c_2 are both negative, and are given in the theorem. This choice of interval and function $g(z)$ implies that $g(0) = 0$, and $g(5) \approx 1 - 0.3 \times 10^{-10}$, which is approximately 1.

Moreover, both $g(z)$ and $erf(z)$ are monotonically increasing functions. Thus, the error function $erf(z)$ is well approximated by $g(z)$ for $z \geq 0$.

Table 1 demonstrates that $F_{app}(Q) = \frac{2}{\sigma} I_{app}(Q)$ delivers a very accurate approximation to $F(Q)$ which cannot be exactly known, but can be estimated by the Accept-Reject algorithm based on a large number of independent draws of ε. Ten million random drawings of ε are observed and the cumulative distribution $F(Q)$ is estimated from the empirical distribution of $\varepsilon \leq Q$. Indeed, for various choices of Q and parameter sets of σ_u and σ_v, the absolute difference between $F_{app}(Q)$ and the empirical estimate of $F(Q)$ based on the Accept-Reject algorithm is less than 0.0003 in probability. More importantly, the absolute difference $\left| F_{app}(Q) - F(Q) \right|$ exhibits no apparent pattern either at the truncation point Q, or the values of parameters σ_u and σ_v. The good approximation of $F_{app}(Q)$ to $F(Q)$ also explains the excellent finite sample performance of the MLE of the censored SFA under various model configurations considered in the next section.

[insert Table 1 here]

3. Monte Carlo Experiments and Results

In this section we consider the finite sample performance of the MLE based on the standard SFA likelihood function, L_0, and the MLE based on the censored SFA likelihood function, L_1, when the data-generating processes (DGP) are model (5). Following Olson et al. (1980), we consider a set of experiments with a simple model:

$$y_i^{*l} = \beta_0 + \beta_1 x_i + \varepsilon_i^l, \quad i = 1, 2, \ldots, n, \; l = 1, 2, \ldots, 500, \tag{14}$$

where l denotes the l-th replication of the data, and the regressors x are drawn from $N(0,1)$. The parameters considered in the Monte Carlo experiments are[2]:

$$\xi = \left(\beta_0, \beta_1, \sigma_u, \sigma_v\right)^{\mathrm{T}}.$$

The maximization of the standard SFA likelihood function L_0 in (4) is well established. For the censored SFA likelihood function L_1 in (6), we take the derivatives of the approximated formula with respect to the parameter ξ, i.e.,

$$\frac{\partial L_1}{\partial \xi} \approx \sum_{y_i>0} \frac{\partial \ln\left(f\left(\varepsilon_i\right)\right)}{\partial \xi} + \sum_{y_i=0} \frac{\partial \ln\left(F_{app}\left(-x_i^{\mathrm{T}}\beta\right)\right)}{\partial \xi}. \tag{15}$$

All the programs are written in GAUSS at three sample sizes, 200, 400, and 800. The optimization algorithm used to implement the MLE is the quasi-Newton algorithm of Broyden, Fletcher, Goldfarb, and Shanno (BFGS) contained in the GAUSS MAXLIK library[3]. The maximum number of iterations for each replication is 200. In order to create a more realistic scenario in simulation, the initial value for the MLE procedure is set at the true parameter value plus a random number generated from $N(0,1)$[4].

The experimental design intends to show that a significant bias exists in MLE if the presence of censored dependent variable is not taken into account. Intuitively, the more observations are censored, the more weight the censored part of the likelihood function L_1 in (6) carries in the maximization. Consequently, a larger bias in MLE is expected based on the misspecified standard SFA model[5]. To examine the sensitivity of the MLE with

---

[2] In the experiments, we estimate ξ with the following transformation function: $\xi = \left(\beta_0, \beta_1, \sigma_u, \sigma_v\right)^{\mathrm{T}} = \kappa\left(\tilde{\xi}\right)$, where $\tilde{\xi} = \left(\beta_0, \beta_1, \ln\left(\sigma_u\right), \ln\left(\sigma_v\right)\right)^{\mathrm{T}}$ are the parameters actually estimated when conducting the MLE of the censored and uncensored SFA.

[3] The GAUSS program for the censored SFA is available upon request from the authors.

[4] More precisely, the initial value of ξ is set to be, $\tilde{\xi}_0 = \kappa\left(\xi\right)^{-1} + N\left(0,1\right)$.

[5] Consider the probability limit of the derivative of L_0 with respect to ξ,

$$Plim\ \frac{1}{n}\frac{\partial L_0}{\partial \xi} = \left(1-m\right) Plim\ \frac{1}{n_1}\sum_{y_i>0}\frac{\partial \ln f\left(\varepsilon_i\right)}{\partial \xi} + m\ Plim\ \frac{1}{n_0}\sum_{y_i=0}\frac{\partial \ln f\left(\varepsilon_i\right)}{\partial \xi}$$

where m is the probability and n_0 is the observations that $y_i = 0$, respectively, and $n_1 + n_0 = n$. If the true model is of the censored SFA, we have $Plim\ \frac{1}{n_1}\sum_{y_i>0}\frac{\partial \ln f\left(\varepsilon_i\right)}{\partial \xi} = 0$, but respect to censored observations, various percentages of the latent dependent variable (y_i^*) falling below zero are assumed in the Monte Carlo experiments. To this end, the slope of the censored SFA in (14) is assumed to be zero, $\beta_1 = 0$, for all experiments, and the intercept β_0 is set to an appropriate value to ensure the probability that y_i^* falling below zero is at a specified m. More specifically, in an experiment with a pre-assigned m, the intercept is set to be $\beta_0 = -F^{-1}(m)$, where $F^{-1}(.)$ is the inverse distribution function of ε. Thus the parameters considered for the simulations are:

$$\xi = (\beta_0, \beta_1, \sigma_u, \sigma_v)^{\mathrm{T}} = (-F^{-1}(m), 0, \sigma_u, \sigma_v)^{\mathrm{T}}. \tag{16}$$

All Monte Carlo experiments are conducted with 500 replications. Some easily distinguished patterns of bias in the MLE emerge in Table 2. The results clearly show that in almost all cases, the bias resulting from the standard MLE (L_0) of the parameters is much larger than the bias resulting from the censored MLE (L_1), and as expected, the bias from L_0 seems to increase with the degree of censoring *m*. Furthermore, the bias from the standard MLE (L_0) of the standard SFA is considerable relatively to its true value, especially in β_0, σ_u, and σ_v. This result implies that the empirical results from the three-stage DEA procedure might be problematic if the standard SFA regression is used in the second stage[6]. On the other hand, it appears that the finite sample performance of the censored MLE based on the proposed approximated formula F_{app} in (13) is very promising, as the associated bias is negligible relative to its true value. This

---

$Plim \frac{1}{n_0} \sum_{y_i=0} \frac{\partial \ln f(\varepsilon_i)}{\partial \xi} \neq 0$. Thus, the degree of bias in the standard SFA depends on the probability *m* that $y_i = 0$.

[6] From our experience in conducting Monte Carlo experiments and in studying empirical applications, the MLE of the standard SFA regression (in the second stage of the three-stage DEA) often fails to obtain "normal" convergence in computation when the probability of censoring or the number of censoring observations is, for example, larger than 8%. The failure is due to either the number of iterations in maximization beyond a reasonable limit or, in most cases, the iterative estimates of σ_u (or σ_v) tending to approach to the boundary of the parameter space, i.e., 0. On the other hand, however, the failure of normal convergence almost never occurs in the maximization of L_1.

outcome is direct evidence of the effectiveness of applying F_{app} in the maximum likelihood estimation of censored stochastic frontier models.

[insert Table 2 here]

To further demonstrate the effectiveness of applying F_{app} in the MLE of the censored SFA estimation, we conduct more detailed experiments with various combinations of parameters $\xi = (\beta_0, \beta_1, \sigma_u, \sigma_v)^{\mathrm{T}}$, and sample size. The results in Tables 3-5 show the mean-squared-errors (MSE) of the MLE of ξ. For all 27 combinations of ξ considered in Tables 3-5, the MSE of censored SFA always decreases with increasing sample size. The results assure that the MLE of the censored SFA estimation based on the proposed F_{app} approximation possesses a well-defined asymptotic behavior. This finding is important, but not surprising in that the MLE of censored SFA takes into account the presence of censored dependent variable in estimation. The simulations enhance our understanding and confidence in the use of F_{app} in the MLE of censored SFA.

[insert Tables 3-5 here]

4. An Empirical Application

We illustrate the proposed censored SFA estimation of a sample of 259 Farmers' Credit Unions (FCUs) in the year 2002 in Taiwan. The FCUs have been regarded as the most important financial institution in rural communities. They provide banking services to its members on deposits, loans, handing remittance services and act as a custodian for township treasury. For the purpose of illustration and comparison between the three-stage DEA using the censored and standard SFA regressions in the second stage, we follow the "intermediation approach" in banking to define three inputs: loanable funds (*Funds*), labor (*Labor*), and capital expenses (*Capital*); and two outputs: farm loans (*FLoan*) and non-farm loans (*NFLoan*). Three types of environmental variable are specified: voting member ratio (*VMR*), number of regional financial institutions (*RFI*), and number of branches (*Branch*)[7]. In general, members of FCUs consist of contributed members with voting right and nonvoting associate members. Only the full-time farmers are eligible to be voting members, and the associate members are mostly part-time farmers and rural residents. Thus, the ratio of voting members (*VMR*) indicates the economic and community environment of a FCU. The number of regional financial institutions (*RFI*) serves as a proxy for the degree of banking competition faced by FCUs. The FCUs were highly regulated and not until recently were prohibited from engaging in any vertical (with other financial institutions) or horizontal (with other FCUs) integration. The number of branches (*Branch*) thus serves as a proxy of the environmental constraint on an FCU's scale of operations. All three types of environmental variables are posited to influence the FCUs performance. For more detailed discussion of the FCU and their operation, readers are referred to the studies by Chen, et al. (2007) and Huang, et al. (1999).

The initial FCU performance evaluation is conducted using the input-oriented, variable returns to scale (VRS) DEA with three inputs: *Funds*, *Labor*, *Capital*; and two outputs: *FLoan*, *NFLoan*[8]. The results of the first-stage DEA analysis are summarized in Table 6. These results show that fully 19 FCUs or 7.34% are at the best practice frontier with an efficiency score of one. More importantly, there are 18 FCUs (6.95%) with zero input slack in *Funds* and *Labor*, and 19 FCUs (7.34%) having zero slack in *Capital*. These zero input slacks form the censored observation points in the second-stage SFA regression.

[insert Table 6 here]

---

[7] The variables, *Funds*, *Capital*, and *FLoan* are measured in millions of NT dollars; *Labor* in number of employee. *VMR* is the percentage of voting members to the total; *RFI* and *Branch* are measured in units.

[8] Briefly, the VRS data envelopment analysis involves solving the linear programming problem: $\min_{\theta,\lambda} \theta$ subject to $\left(\theta x \ge X\lambda,\ Y\lambda \ge y,\ \lambda \ge 0,\ e^{\mathrm{T}}\lambda = 1\right)$, where $x = (Funds, Labor, Capital)^{\mathrm{T}}$ is the column input vector, and $y = (FLoan, NFloan)^{\mathrm{T}}$ is the column output vector of the FCU being evaluated; $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$ are the matrices of inputs and outputs of all FCUs; $\lambda = (\lambda_1, \ldots, \lambda_n)^{\mathrm{T}}$ is a column vector of weights, and $e^{\mathrm{T}} = (1, \ldots, 1)$ is a row vector with elements 1. The input-oriented DEA efficiency score is the optimal value $\theta \le 1$. The input slacks are computed as $(x - X\lambda) \ge 0$. The linear programming is solved n times, once for each FCU.

To account for the environmental impact on FCU's managerial inefficiency, three second-stage SFA regressions of the input slacks are modeled on the three environmental variables: *VMR*, *RFI*, and *Branch*. Let the computed nonnegative input slacks on *Funds*, *Labor*, and *Capital* from the first-stage DEA be denoted as $Slack_F$, $Slack_L$, and $Slack_C$. The second-stage censored SFA regressions corresponding to (5) are specified as:

$$Slack_F^* = \beta_{0F} + \beta_{1F} VMR + \beta_{2F} RFI + \beta_{3F} Branch + \varepsilon_F ,$$

$$Slack_L^* = \beta_{0L} + \beta_{1L} VMR + \beta_{2L} RFI + \beta_{3L} Branch + \varepsilon_L ,$$

$$Slack_C^* = \beta_{0C} + \beta_{1C} VMR + \beta_{2C} RFI + \beta_{3C} Branch + \varepsilon_C , \tag{17}$$

where $Slack_j = Slack_j^*$, if $Slack_j^* > 0$; and $Slack_j = 0$, if $Slack_j^* \le 0$, for $j = F, L, C$. The composite error $\varepsilon_j = v_j + u_j$ $(j = F, L, C)$ consists of the random noise component v_j with iid as $N(0, \sigma_{vj}^2)$, and the nonnegative component u_j with iid as $N^+(0, \sigma_{uj}^2)$. Since the input slacks are nonnegative, the above second-stage SFA regressions are censored SFA. Table 7 shows the empirical estimates based on the correctly specified censored SFA log likelihood function L_1 of (6) and the misspecified standard SFA log likelihood function L_0 of (4). Results of the second-stage SFA regressions show a consistent sign on the estimated coefficients β_j across the three slack equations between the censored and standard SFA estimations. Both censored and standard SFA regressions reject the hypothesis of $\sigma_u = 0$ that the one-sided error component u_j makes no contribution to the composite error ε_j. The results suggest that a significant portion of input slacks is attributable to managerial inefficiency. However, the standard SFA estimates of βs seem to consistently underestimate the impacts of the environmental variables when compared to the censored SFA estimates. Furthermore, judging from the ratios of the standard deviations, $\lambda = \frac{\sigma_u}{\sigma_v}$, we see that the standard SFA regression, on the other hand, consistently overestimates the portion of input slacks due to managerial inefficiency, and underestimates the portion due to statistical noise. Let $f_j\left(z_j; \beta_j\right)$ be the deterministic feasible slack frontiers of (17), i.e.,

$\beta_{0j} + \beta_{1j}VMR + \beta_{2j}RFI + \beta_{3j}Branch$. The stochastic feasible slack frontiers (SFSF), $f_j\left(z_j;\beta_j\right)+v_j$, represents the environmental impacts and statistical noise on input utilization to be netted out from evaluating the managerial efficiency of FCUs. Table 8 shows that, combining the underestimation of $f_j\left(z_j;\beta_j\right)$ and v_j, the standard SFA estimates of the stochastic feasible slack frontiers are much smaller than the censored SFA estimates[9]. These underestimations of environmental effects and statistical noise affect the efficiency calculation in the third-stage DEA.

[insert Table 7 here]

[insert Table 8 here]

Instead of directly subtracting the SFSF, i.e., $f_j\left(z_j;\beta_j\right)+v_j$, from inputs to construct the adjusted inputs in the third-stage DEA, Fried et al. (2002) suggest the following procedure to avoid the possibility that some inputs might be adjusted so far downward as to become negative:

$$Funds^{adj} = Funds + Slack^{adj}_{Funds},$$

$$Labor^{adj} = Labor + Slack^{adj}_{Labor},$$

$$Capital^{adj} = Capital + Slack^{adj}_{Capital}$$

where the slack adjustments are defined as

$$Slack^{adj}_j = \left[\max\left(f_j\left(z_j;\beta_j\right)\right) - f_j\left(z_j;\beta_j\right)\right] + \left[\max\left(v_j\right) - v_j\right]$$

---

9 Adopting the Jondrow et al. (1982) procedure to decompose the composite error, the estimate of the i^{th} observation u_{ji} of the j^{th} slack censored SFA regression is based on the conditional mean of u_{ji} given ε_{ji}, or $\hat{u}_{ji} = \sigma_{*j}\left[\frac{\phi\left(\hat{\varepsilon}_{ji}\hat{\lambda}_j/\hat{\sigma}_j\right)}{\Phi\left(\hat{\varepsilon}_{ji}\hat{\lambda}_j/\hat{\sigma}_j\right)} + \frac{\hat{\varepsilon}_{ji}\hat{\lambda}_j}{\hat{\sigma}_j}\right]$, where $\hat{\varepsilon}_{ji} = Slack^*_{ji} - f_j\left(z_{ji},\hat{\beta}_j\right)$ and $\hat{\sigma}_{*j} = \frac{\hat{\sigma}_{uj}\sigma_{vj}}{\hat{\sigma}_j}$.

When $Slack^*_{ji} > 0$, $Slack^*_{ji} = Slack_{ji}$, and hence $\hat{\varepsilon}_{ji} = Slack_{ji} - f_j\left(z_{ji},\hat{\beta}_j\right)$; when $Slack^*_{ji} \le 0$, $Slack^*_{ji}$ is set at the upper bound so that $Slack^*_{ji} = Slack_{ji} = 0$, and hence $\hat{\varepsilon}_{ji} = -f_j\left(z_{ji},\hat{\beta}_j\right)$. The estimate of the noise component v_{ji} is obtained by the subtraction, $\hat{v}_{ji} = \hat{\varepsilon}_{ji} - u_{ji}$.

The above adjustment represents an upward adjustment in inputs of FCUs that have been advantaged by their relatively favorable operating environments or by their relative good luck. Table 9 gives the third-stage DEA efficiency estimates computed from the adjusted inputs, and their respected slacks adjustments $Slack_j^{adj}$. As expected, the efficiency estimate from the correctly specified censored SFA is higher than that from the misspecified standard SFA. However, both estimates are higher than the initial, first-stage efficiency estimate shown in Table 6 where environmental effects and statistical noise are not taken into consideration.

[insert Table 9 here]

5. Summary and Conclusions

This paper takes issues with the appropriateness of applying the standard SFA regression in accounting for environmental effects and statistical noise in DEA. The mixed DEA-SFA approach is based on the popular Fried et al. (1999, 2002) three-stage DEA methodology. In the second-stage, a stochastic frontier regression is estimated separately for each input (output) with the corresponding DEA input (output) slacks calculated from the first-stage as the dependent variable and environmental variables as the independent variables. In practice, however, a considerable number of calculated input (output) slacks are zero. Thus, the stochastic frontier regressions in the second-stage are of the censored type rather than the standard (uncensored) SFA models. The bias estimates of the environmental impacts and noise from the misspecified standard SFA regressions in the second-stage thus carry over to the third-stage DEA estimation of the efficiency score. The Monte Carlo experiments clearly show that the bias is substantial and it increases with the degree of censoring.

As noted in Section 1, none of the published studies on the mixed DEA-SFA approach that we are aware of have correctly specified and estimated the censored SFA regression. In this paper, we propose the MLE of the censored SFA by first providing an easy-to-implement and accurate closed-form formula for computing the cumulative distribution function of the composite error to facilitate the computation of MLE. The

Monte Carlo experiments show that the proposed MLE of the censored SFA models is very promising. We illustrate the proposed censored SFA model and estimation of a sample of farmers' credit unions in Taiwan. The empirical results show that the standard SFA tends to consistently underestimate the impacts of environmental variables and statistical noise than the censored SFA estimates. As a consequence, the third-stage DEA efficiency scores estimated from the misspecified standard SFA are smaller than from the correctly specified censored SFA.

From an application viewpoint, the discrepancy between the standard and censored SFA results clearly depends on the number of censored observations, or the percentage of zero slack values obtained from the first-stage DEA calculation. In our experience, a significant discrepancy is often observed even if the percentage is as small as 6%. Furthermore, the standard SFA is often difficult to estimate by the maximum likelihood technique with the censored observations. The standard computer program routines, such as LIMDEP, STATA, or FRONTIER 4.1, often fail to obtain "normal" convergence in computation. However, the failure of normal convergence almost never occurs in the maximization of the correctly specified censored SFA likelihood function.

Appendix

This appendix shows the derivation of the approximated cumulative distribution function F_{app}. Since $a = \frac{\lambda}{\sigma} \geq 0$, we divide the derivation into two parts: for $(Q \leq 0, a \geq 0)$ and $(Q \geq 0, a \geq 0)$. Furthermore, for ease of exposition, two equations from Abramowitz and Stegun (1970, equations (7.11) and (7.4.32)) are given:

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt = 2\int_0^{\sqrt{2}z} \phi(t)\, dt,$$

$$\int e^{-\left(kx^2+2mx+n\right)} dx = \frac{1}{2}\sqrt{\frac{\pi}{k}}\, e^{\frac{m^2-kn}{k}}\, erf\left(\sqrt{k}x + \frac{m}{\sqrt{k}}\right) + C, \ \ k \neq 0,$$

where C denotes a finite constant.

Given that $(Q, a, b) \in$ finite R, $b > 0$, $erf(-x) = -erf(x)$, and define $\varepsilon = \sqrt{2}v/a$, we have:

$$\begin{aligned}
I_{a\geq 0} &= \frac{\sqrt{2}}{a} \int_{-\infty}^{\frac{a}{\sqrt{2}}Q} \left(\int_{-\infty}^{\sqrt{2}v} \phi(\varsigma)\, d\varsigma \right) \phi\left(\sqrt{2}v \frac{b}{a}\right) dv \\
&= \frac{\sqrt{2}}{2a} \int_{-\infty}^{\frac{a}{\sqrt{2}}Q} \left(1 + erf(v)\right) \phi\left(\sqrt{2}v \frac{b}{a}\right) dv \\
&= \frac{\sqrt{2}}{2a} \int_{-\infty}^{0} \left(1 + erf(v)\right) \phi\left(\sqrt{2}v \frac{b}{a}\right) dv + \frac{\sqrt{2}}{2a} \int_{0}^{\frac{a}{\sqrt{2}}Q} \left(1 + erf(v)\right) \phi\left(\sqrt{2}v \frac{b}{a}\right) dv,
\end{aligned}$$

Note that $erf(z)$ can be well approximated by a function, $g(x) = 1 - e^{c_1 x + c_2 x^2}$ for $x \geq 0$, where c_1 and c_2 are chosen to ensure that $g(x)$ is as close to $erf(x)$ as possible. The choice of c_1 and c_2 is discussed in Section 2.

With the preceding results of $I_{a\geq 0}(Q)$, we then have

$$\begin{aligned}
I_{a\geq 0, Q\geq 0}(Q) &= \frac{\sqrt{2}}{2a} \int_{-\infty}^{0} \left(1 + erf(v)\right) \phi\left(\sqrt{2}v \frac{b}{a}\right) dv + \frac{\sqrt{2}}{2a} \int_{0}^{\frac{a}{\sqrt{2}}Q} \left(1 + erf(v)\right) \phi\left(\sqrt{2}v \frac{b}{a}\right) dv \\
&= \frac{\sqrt{2}}{2a} \int_{0}^{\infty} \left(1 - erf(v)\right) \phi\left(\sqrt{2}v \frac{b}{a}\right) dv + \frac{\sqrt{2}}{2a} \int_{0}^{\frac{a}{\sqrt{2}}Q} \left(1 + erf(v)\right) \phi\left(\sqrt{2}v \frac{b}{a}\right) dv
\end{aligned}$$

$$\approx \frac{\sqrt{2}}{2a}\int_0^{\infty} e^{c_1 v + c_2 v^2}\ \frac{1}{\sqrt{2\pi}}\ e^{-\frac{v^2 b^2}{a^2}} dv + \frac{\sqrt{2}}{a}\int_0^{\frac{a}{\sqrt{2}}Q}\ \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2 b^2}{a^2}} dv$$

$$-\frac{\sqrt{2}}{2a}\int_0^{\frac{a}{\sqrt{2}}Q}\ \left(e^{c_1 v + c_2 v^2}\right)\frac{1}{\sqrt{2\pi}} e^{-\frac{v^2 b^2}{a^2}} dv$$

$$= \frac{1}{2\sqrt{\pi}a}\int_0^{\infty} exp\left(-\left(\tfrac{b^2 - a^2 c_2}{a^2}\right)v^2 + c_1 v\right)dv$$

$$+\frac{1}{\sqrt{\pi}a}\int_0^{\frac{b}{\sqrt{2}}Q}\ exp\left(-v^2\right)d\left(\tfrac{a}{b}v\right) - \frac{1}{2\sqrt{\pi}a}\int_0^{\frac{a}{\sqrt{2}}Q} exp\left(-\left(\tfrac{b^2 - a^2 c_2}{a^2}\right)v^2 + c_1 v\right)dv.$$

When we use (7.4.32) of Abramowitz and Stegun (1970), $I_{a\geq 0, Q\geq 0}$ can be approximated by:

$$I_{a\geq 0, Q\geq 0}(Q) \approx exp\left(\frac{a^2 c_1^2}{4b^2 - 4a^2 c_2}\right)\ \frac{1}{4\sqrt{b^2 - a^2 c_2}}\ \left[1 - erf\ \left(\frac{-a\ c_1 + \sqrt{2}Q\left(b^2 - a^2 c_2\right)}{2\sqrt{b^2 - a^2 c_2}}\right)\right]$$

$$+\frac{1}{2b} erf\left(\frac{b\ Q}{\sqrt{2}}\right).$$

Likewise, we can derive the approximation for $I_{a\geq 0, Q\leq 0}$:

$$I_{a\geq 0, Q\leq 0}(Q) \approx exp\left(\frac{a^2 c_1^2}{4b^2 - 4a^2 c_2}\right)\ \frac{1}{4\sqrt{b^2 - a^2 c_2}}\ \left[1 - erf\ \left(\frac{-a\ c_1 - \sqrt{2}Q\left(b^2 - a^2 c_2\right)}{2\sqrt{b^2 - a^2 c_2}}\right)\right].$$

Combining the preceding results, we prove the Theorem.

Table 1. Accuracy of $F_{app}(Q)$ in Computing $F(Q)$ at Various Q, σ_u, and λ

$\sigma_u = 1$				Q			
Method	–3.0	–2.0	–1.0	0.0	1.0	2.0	3.0
				$\lambda = 1.50\ (\sigma_v = 0.6667)$			
$100 \times F_{app}(Q)$	0.000060	0.022664	1.488730	18.740960	60.943449	90.408314	98.743208
$100 \times F(Q)$	0.000060	0.019500	1.477160	18.725950	60.914390	90.406290	98.746360
$100 \times (AbsD)$	0.000000	0.003164	0.011570	0.015010	0.029059	0.002024	0.003152
				$\lambda = 1.25\ (\sigma_v = 0.8000)$			
$100 \times F_{app}(Q)$	0.001620	0.128156	2.957624	21.493701	59.469658	88.293175	98.086663
$100 \times F(Q)$	0.001390	0.122430	2.954490	21.481860	59.462650	88.285310	98.088210
$100 \times (AbsD)$	0.000230	0.005726	0.003134	0.011841	0.007008	0.007865	0.001547
				$\lambda = 0.85\ (\sigma_v = 1.1765)$			
$100 \times F_{app}(Q)$	0.149700	1.499332	8.425067	27.601621	56.692885	81.965878	94.941287
$100 \times F(Q)$	0.147350	1.499210	8.428980	27.579250	56.708140	81.971100	94.934540
$100 \times (AbsD)$	0.002350	0.000122	0.003913	0.022371	0.015255	0.005222	0.006747

Note: $F_{app}(Q)$ is computed based on $\frac{2}{\sigma} I_{app}(Q)$ in (13) and $F(Q)$ is computed from the Accept-Reject algorithm based on 10 million independent draws of the distribution $f(\varepsilon)$ in (7). *AbsD* denotes the absolute difference $\left|F_{app}(Q) - F(Q)\right|$.

Table 2. Bias of the Censored (L_1) and Standard (L_0) MLE:

$\beta_0 = -F^{-1}(m),\ \beta_1 = 0,\ \sigma_v = 0.6667,\ \sigma_u = 1$

	β_0		β_1		σ_u		σ_v	
L	L_0	L_1	L_0	L_1	L_0	L_1	L_0	L_1
n				$\beta_0 = 0.91\ (m = 0.02)$				
200	–0.0454	0.0664	0.0004	–0.0001	0.0654	–0.0658	–0.0806	0.0094
400	–0.0635	0.0341	0.0019	0.0018	0.0845	–0.0399	–0.0789	–0.0028
800	–0.0702	0.0103	–0.0003	–0.0003	0.0937	–0.0114	–0.0786	–0.0043
n				$\beta_0 = 0.68\ (m = 0.04)$				
200	–0.1243	0.0779	0.0005	0.0003	0.1674	–0.0937	–0.1672	0.0002
400	–0.1353	0.0379	0.0019	0.0020	0.1784	–0.0472	–0.1613	–0.0010
800	–0.1413	0.0129	–0.0004	–0.0002	0.1866	–0.0148	–0.1620	–0.0042
n				$\beta_0 = 0.53\ (m = 0.06)$				
200	–0.2317	0.0890	–0.0001	0.0002	0.2963	–0.1082	–0.3293	0.0044
400	–0.2521	0.0457	0.0008	0.0019	0.3165	–0.0574	–0.3304	0.0024
800	–0.2481	0.0191	–0.0002	–0.0004	0.3142	–0.0229	–0.3100	–0.0024

Note: All results are based on 500 replications. The true parameters in simulation are: $\xi = (\beta_0, \beta_1, \sigma_u, \sigma_v)^{\mathrm{T}} = (-F^{-1}(m),\ 0,\ 1,\ 0.6667)^{\mathrm{T}}$, where β_0 is chosen to ensure the probability that the dependent variable latent y_i^* falls below zero at a specified m under various configurations.

Table 3. Mean Squared Error (MSE) of the Censored (L_1) MLE: $\beta_0 = -F^{-1}(m)$, $\beta_1 = 0$, $\sigma_u = 1.0$

	β_0			β_1			σ_u			σ_v		
						$\lambda = 1.50$ $(\sigma_v \approx 0.6667)$						
n	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06
200	0.0782	0.0753	0.0825	0.0074	0.0045	0.0045	0.1324	0.1145	0.1273	0.0802	0.0155	0.0173
400	0.0363	0.0389	0.0434	0.0019	0.0019	0.0019	0.0506	0.0607	0.0688	0.0077	0.0085	0.0095
800	0.0109	0.0139	0.0192	0.0011	0.0011	0.0011	0.0171	0.0218	0.0303	0.0032	0.0038	0.0047
						$\lambda = 1.25$ $(\sigma_v = 0.80)$						
n	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06
200	0.1398	0.1372	0.1427	0.0090	0.0059	0.0056	0.2180	0.2024	0.2233	0.0520	0.0200	0.0227
400	0.0730	0.0773	0.0887	0.0023	0.0024	0.0024	0.1086	0.1208	0.1403	0.0107	0.0114	0.0133
800	0.0290	0.0392	0.0492	0.0014	0.0014	0.0014	0.0456	0.0620	0.0790	0.0049	0.0061	0.0074
						$\lambda = 0.85$ $(\sigma_v \approx 1.1765)$						
n	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06
200	0.2767	0.2650	0.2669	0.0231	0.0099	0.0096	0.5212	0.3989	0.4140	0.1252	0.0299	0.0323
400	0.2088	0.2180	0.2430	0.0040	0.0041	0.0042	0.3237	0.3465	0.3851	0.0199	0.0207	0.0253
800	0.1454	0.1818	0.2091	0.0025	0.0025	0.0025	0.2327	0.2931	0.3352	0.0110	0.0132	0.0175

Note: All results are based on 500 replications. The true parameters in simulation are: $\xi = (\beta_0, \beta_1, \sigma_u, \sigma_v)^{\mathrm{T}} = \left(-F^{-1}(m),\ 0,\ 1,\ \sigma_v\right)^{\mathrm{T}}$, where β_0 is chosen to ensure the probability that the dependent variable latent y_i^* falls below zero at a specified m under various configurations.

Table 4. Mean Squared Error (MSE) of the Censored (L_1) MLE: $\beta_0 = -F^{-1}(m)$, $\beta_1 = 0$, $\sigma_u = 1.2$

	β_0			β_1			σ_u			σ_v		
						$\lambda = 1.50$ $(\sigma_v = 0.80)$						
n	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06
200	0.1054	0.1119	0.1184	0.0093	0.0065	0.0065	0.1903	0.1704	0.1829	0.1154	0.0226	0.0252
400	0.0462	0.0551	0.0614	0.0027	0.0027	0.0028	0.0718	0.0858	0.0978	0.0106	0.0122	0.0136
800	0.0146	0.0205	0.0263	0.0016	0.0016	0.0016	0.0229	0.0322	0.0416	0.0046	0.0055	0.0066
						$\lambda = 1.25$ $(\sigma_v = 0.96)$						
n	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06
200	0.1938	0.1966	0.2049	0.0114	0.0084	0.0081	0.3139	0.2948	0.3206	0.0748	0.0291	0.0327
400	0.1022	0.1133	0.1289	0.0034	0.0034	0.0034	0.1543	0.1773	0.2040	0.0154	0.0167	0.0192
800	0.0430	0.0566	0.0712	0.0020	0.0020	0.0021	0.0674	0.0893	0.1140	0.0071	0.0088	0.0107
						$\lambda = 0.85$ $(\sigma_v \approx 1.4118)$						
n	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06
200	0.3976	0.3819	0.3800	0.0339	0.0141	0.0137	0.7522	0.5787	0.5987	0.1810	0.0432	0.0465
400	0.2974	0.3185	0.3458	0.0058	0.0059	0.0060	0.4640	0.5064	0.5505	0.0285	0.0297	0.0324
800	0.2102	0.2586	0.2917	0.0036	0.0036	0.0036	0.3368	0.4177	0.4701	0.0159	0.0188	0.0214

Note: All results are based on 500 replications. The true parameters in simulation are: $\xi = (\beta_0, \beta_1, \sigma_u, \sigma_v)^{\mathrm{T}} = (-F^{-1}(m),\ 0,\ 1,\ \sigma_v)^{\mathrm{T}}$, where β_0 is chosen to ensure the probability that the dependent variable latent y_i^* falls below zero at a specified m under various configurations.

Table 5. Mean Squared Error (MSE) of the Censored (L_1) MLE: $\beta_0 = -F^{-1}(m)$, $\beta_1 = 0$, $\sigma_u = 1.4$

	β_0			β_1			σ_u			σ_v		
						$\lambda = 1.50$ $(\sigma_v \approx 0.9333)$						
n	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06
200	0.1361	0.1488	0.1619	0.0116	0.0088	0.0089	0.2568	0.2268	0.2496	0.1569	0.0305	0.0341
400	0.0628	0.0747	0.0800	0.0037	0.0037	0.0038	0.0975	0.1164	0.1267	0.0145	0.0165	0.0180
800	0.0198	0.0252	0.0354	0.0021	0.0021	0.0022	0.0312	0.0399	0.0562	0.0062	0.0072	0.0090
						$\lambda = 1.25$ $(\sigma_v = 1.12)$						
n	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06
200	0.2576	0.2599	0.2782	0.0143	0.0113	0.0110	0.4272	0.3921	0.4350	0.1019	0.0391	0.0442
400	0.1375	0.1514	0.1753	0.0046	0.0047	0.0047	0.2101	0.2366	0.2768	0.0209	0.0225	0.0261
800	0.0602	0.0710	0.0969	0.0027	0.0027	0.0028	0.0941	0.1115	0.1551	0.0098	0.0114	0.0145
						$\lambda = 0.85$ $(\sigma_v \approx 1.6471)$						
n	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06	m=0.02	m=0.04	m=0.06
200	0.5246	0.5163	0.5222	0.0385	0.0190	0.0187	1.0196	0.7870	0.8098	0.2464	0.0587	0.0634
400	0.3996	0.4340	0.4717	0.0079	0.0080	0.0081	0.6272	0.6920	0.7506	0.0387	0.0405	0.0440
800	0.2846	0.3528	0.3966	0.0048	0.0049	0.0050	0.4558	0.5692	0.6387	0.0215	0.0255	0.0289

Note: All results are based on 500 replications. The true parameters in simulation are: $\xi = (\beta_0, \beta_1, \sigma_u, \sigma_v)^{\mathrm{T}} = (-F^{-1}(m),\ 0,\ 1,\ \sigma_v)^{\mathrm{T}}$, where β_0 is chosen to ensure the probability that the dependent variable latent y_i^* falls below zero at a specified m under various configurations.

Table 6. First-Stage Variable Returns to Scale DEA Results

	Mean	Std. Deviation	Minimum	Maximum	% of one/zero
First-Stage Efficiency	0.7473	0.1571	0.2080	1	7.34%
Input Slacks:					
Funds	8.6272	7.8738	0	59.811	6.95%
Labor	46.5180	44.9772	0	326.741	6.95%
Capital	19.9013	9.9740	0	81.557	7.34%

Note: The “% of one/zero” for the first-stage efficiency is the percent of observations (DMUs) with an efficiency score equal to one; for the input slack variables they are the percentage of observations with zero slack value.

Table 7. Second-Stage SFA Results

Independent Variable	Dependent Variable: Input Slack					
	Slack (Funds)		*Slack (Labor)*		*Slack (Capital)*	
	Standard SFA: L_0	Censored SFA: L_1	Standard SFA: L_0	Censored SFA: L_1	Standard SFA: L_0	Censored SFA: L_1
Constant: (β_0)	–1.3254	–2.0339	–7.5637	–8.1212	–2.0352	–2.9683
	(0.5632)	(0.7104)	(3.5505)	(4.7192)	(0.7214)	(0.9041)
VMR: (β_1)	–0.0006	–0.0000	0.0149	0.0162	–0.0016	–0.0017
	(0.0019)	(0.0019)	(0.0102)	(0.0101)	(0.0025)	(0.0022)
RFI : (β_2)	0.1444	0.1982	1.7526	1.9050	0.1743	0.2233
	(0.0579)	(0.0524)	(0.2845)	(0.2811)	(0.0620)	(0.0671)
Branch: (β_3)	0.5085	0.8021	2.9067	3.6681	0.8160	1.1914
	(0.1801)	(0.1919)	(0.9229)	(0.9623)	(0.2474)	(0.2374)
σ_u	9.8638	8.6870	42.7686	37.4492	12.3416	11.1016
	(0.8605)	(0.8609)	(4.8534)	(6.0701)	(1.0479)	(1.0706)
σ_v	2.5189	4.0421	20.2037	25.8043	3.3525	5.1158
	(0.6268)	(0.5483)	(2.7905)	(3.1834)	(0.7396)	(0.6846)
$\lambda = \frac{\sigma_u}{\sigma_v}$	3.9159	2.1491	2.1169	1.4513	3.6813	2.1701

Note: The numbers in parentheses are the asymptotic standard errors.

Table 8. Stochastic Feasible Slack Frontiers (SFSF):
$$f_j\left(z_j;\beta_j\right)+v_j$$

Input	*Funds*	*Labor*	*Capital*
Standard SFA: L_0	1.0995	12.9725	1.4804
Censored SFA: L_1	1.8500	16.7493	2.2454

Table 9. Third-Stage Variable Returns to Scale DEA Results

	Mean	Std. Deviation	Minimum	Maximum
Third-Stage Efficiency:				
Standard SFA (L_0)	0.9514	0.0472	0.7570	1.0000
Censored SFA (L_1)	0.9704	0.0307	0.8250	1.0000
Adjusted Input $Funds^{adj}$:				
Standard SFA (L_0)	43.8714	16.4504	23.0641	146.6080
Censored SFA (L_1)	56.3391	15.3981	37.4209	149.6298
Adjusted Input $Labor^{adj}$:				
Standard SFA (L_0)	428.5208	133.9140	278.6734	1385.6390
Censored SFA (L_1)	471.7183	131.7805	327.7251	1434.4740
Adjusted Input $Capital^{adj}$:				
Standard SFA (L_0)	55.0904	22.3142	27.9083	206.6725
Censored SFA (L_1)	66.5968	21.3230	41.5177	213.7632

REFERENCES

Abramowitz, M. and I. A. Stegun (1970). *Handbook of Mathematical Functions*. Dover, New York.

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

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press, Cambridge.

Chen, P. C., M. M. Yu, C. C. Chang and S. H. Hsu (2007). "Productivity Change in Taiwan's Farmers' Credit Unions: A Nonparametric Risk-Adjusted Malmquist Approach." *Agricultural Economics* 36, 221-231.

Cordero-Ferrera, J. M., F. Pedraja-Chaparro and J. Salinas-Jimenez (2008). "Measuring Efficiency in Education: An Analysis of Different Approaches for Incorporating Non-Discretionary Inputs." *Applied Economics* 40, 1323-1339.

Drake, L. M. and R. Simper (2005). "The Measurement of Police Force Efficiency: An Assessment of U.K. Home Office Policy." *Contemporary Economic Policy* 23, 465-482.

Fried, H. O., S. S. Schmidt and S. Yaisawarng (1999). "Incorporating the Operating Environment into a Nonparametric Measure of Technical Efficiency." *Journal of Productivity Analysis* 12, 249-267.

Fried, H. O., C. A. K. Lovell, S. S. Schmidt and S. Yaisawarng (2002). "Accounting for Environmental Effects and Statistical Noise in Data Envelopment Analysis." *Journal of Productivity Analysis* 17, 157-174.

Fu, T. T., C. J. Huang and M. Y. Huang (1999). "Cost Efficiency in Taiwan Farmers' Credit Unions." in Tsu-Tan Fu, Cliff J. Huang, and C.A. Knox Lovell (eds), Economic Efficiency and Productivity Growth in the Asia Pacific Region, Edward Elgar, 1999.

Glass, J. C., G. McCallion, D. G. McKillop and K. Stringer (2006). "A 'Technically Level Playing-Field' Profit Efficiency Analysis of Enforced Competition between Publicly Funded Institutions." *European Economic Review* 50, 1601-1626.

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

Lee, J. Y. (2008). "Application of the Three-Stage DEA in Measuring Efficiency – An Empirical Evidence." *Applied Economics Letters* 15, 49-52.

Liu, J. and K. Tone (2008). "A Multistage Method to Measure Efficiency and its Application to Japanese Banking Industry." *Socio-Economic Planning Sciences* 42, 75-91.

Mararai, B. B., F. Erbetta, C. Petraglia and M. Piacenza (2007). "Regulatory and Environmental Effects on Public Transit Efficiency: A Mixed DEA-SFA Approach." *Journal of Regulatory Economics* 32, 131-151.

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

Tobin, J. (1958). "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26, 24-36.