Scan to know paper details and author's profile

Stochastic Frontier Model and Factors Influencing Seed Cotton Production Cost

Dr. Eddy Hope Kabasele Bambe Prof. Dr. Roger Ntoto & Dr. Xiao Wen Wei

ABSTRACT

For this analysis, we used survey data from the China Academy of Agricultural Sciences Cotton Research Institute for 1253 China seed cotton producers, employing the framework of a stochastic frontier trans-logarithmic cost function of cross-sectional data. Our findings showed that China seed cotton producers are almost in full efficiency. However, the assumption that farmers minimize cost was not supported by the sample data. Further, both the economic factors related to farm size that might affect the cost of producing seed cotton and the negative externalities that could affect environmental management over short and long durations were captured.

Keywords: economic efficiency, stochastic cost frontier, cost determinants, and cotton-grain farms.

Classification: For Code: 010599

Language: English

London Journals Press

LJP Copyright ID: 925652
Print ISSN: 2631-8490
Online ISSN: 2631-8504

London Journal of Research in Science: Natural and Formal

Volume 19 | Issue 5 | Compilation 1.0

Stochastic Frontier Model and Factors Influencing Seed Cotton Production Cost

Dr. Eddy Hope Kabasele Bambe[α], Prof. Dr. Roger Ntoto[σ] & Dr. Xiao Wen Wei[ρ]

ABSTRACT

For this analysis, we used survey data from the China Academy of Agricultural Sciences Cotton Research Institute for 1253 China seed cotton producers, employing the framework of a stochastic frontier trans-logarithmic cost function of cross-sectional data. Our findings showed that China seed cotton producers are almost in full efficiency. However, the assumption that farmers minimize cost was not supported by the sample data. Further, both the economic factors related to farm size that might affect the cost of producing seed cotton and the negative externalities that could affect environmental management over short and long durations were captured.

Keywords: economic efficiency, stochastic cost frontier, cost determinants, and cotton-grain farms.

Author α: Agricultural Information Institute of Chinese Academic of Agricultural Sciences Department of Agricultural Analysis, Monitoring, and Early Warning.
e-mail: eddiehope61@yahoo.fr
σ: University of Kinshasa Agricultural Economic
ρ: Institute of Cotton Research of CAAS Department of Agricultural Economic and Management.

I. INTRODUCTION

An understanding of Chinese economics is facilitated by an understanding of China's cotton market. Cotton is a strategic product that is China's leading cash crop, providing raw material to factories to make clothing for a large population that currently comprises approximately 20 percent of the world's population. Cotton is both an important economic crop and a raw material for the textile industry, providing income to nearly 100 million cotton farmers and employment to more than 2000 million textile workers in 2012. Furthermore, the cotton industry is responsible for approximately $283.9 billion in textile and apparel exports (USDA Foreign Agricultural Service January 2017). However, the results of current government programs show that Chinese cotton is less competitive in terms of both price and production cost production than either rayon or foreign cotton (ICAA, 2016). To understand the reason for China's new policy, we must evaluate recent developments in its agricultural sector. Two historical events of relevance have characterized the growth in China's cotton market and production sector. First, China's agricultural production increased after the introduction of market-based reforms in 1978 that included the elimination of the collective production system and the relaxation of government direction over certain farmer production and marketing decisions (Lohmar et al. 2009). Second, changes have occurred since China's December 2001 accession to the World Trade Organization (WTO). Consequently, China's sector industries of textile and apparel exports have grown substantially. These positive measures have also been a windfall for producers and other actors involved in China's cotton production and internal market for cotton. It is noteworthy that other factors have caused changes in China's cotton industry that can be explained through both historical and market events.

According to Stephen MacDonald (2015), both China's cotton policy and China's macroeconomic policy in the cotton sector emphasize the main goal of management to create competitive conditions that contribute to farmers' welfare. Example include China's 1995 and 1999 government-initiated policies for the stock acquisition of cotton. In 2011, this policy's goal of

preserving the stock of China's national reserve coupled with the introduction of a formal price-support system for cotton was observed when the price of cotton reached a record level of more than 136.3 percent over 1995, for the first time exceeding China's bar of 1500 dollars per ton. This circumstance contributed (in one way or another) to the increased total production cost of cotton grain incurred to farms. We can understand this cost increase through the observation that the high return led to farmers' neglect of the importance of continuing to use their inputs efficiently. Another argument could be that hired farm workers' improved standing in China allowed hired them to obtain a high wage (Qin zhag, 2015). Perhaps, as markets for goods interact, other producers who create or generate inputs for farmers have seen that the price of cotton has increased and responded by establishing higher input prices[1]. Both China's national economic growth and the global economic fluctuation can explain the increased production cost of the past decade.

Nevertheless, as stressed early in January 2014, the Chinese government's cotton reserve policy, with its functions of two plus one viz., for purchase, holding and selling cotton supplies, had relaxed its policy of regulating the cotton market and supporting producers. Currently, government policy focuses on developing China's cotton industry by producing high-quality cotton and promoting sustained and healthy economic development. These new policies targeting price reform have strongly affected cotton producers. There has been a significant area shift from cotton to other crops that offer higher returns. The overall declining trend in cotton is occurring more rapidly in regions of Eastern China (i.e., the Yellow River and the Yangtze River Valley) and to a lesser extent in Xinjiang Province (Wang and Li, 2006). The area of cotton-grain cultivation has decreased in Eastern China by approximately 34 percent and in Xinjiang Province by 6 percent (USDA., op cit.).

Abundant evidence, however, has shown that cotton producers respond to research on reducing the cost of production, processing, quality improvement and evaluation, and product development (Domingos Sárvio Magalhães Valente et al. 2012, Wang 2011, Wang et al. 2003, Zhao et al. 2003). In the past few years, there has been a gradual increase in the funds available for cotton research to identify hybrids that are resistant to insect pests (Huang, Rozelle, et al., 2002). In 1997, the Cotton Research Institute of the Chinese Academy of Agriculture Sciences (CRICAAS) developed Bt cotton to eradicate the cotton bollworm (Helicopera armigera) (Pray et al., 2001). As a result, cotton yield increased, decreasing farmer production costs by reducing the need for pesticides to combat insect pests (Huang, Rozelle, et al., 2002).

The stochastic frontier to estimate efficiency is a very relevant instrument for productivity growth, especially for a country that yearns to structure and develop its agricultural sector (Ali & Chaudhry 1990). As stressed by Kumbhakar and Lovell (2000), not all producers are technically efficient, not all producers are cost efficient, and not all producers are proficiency efficient. An improved understanding of the level of cost efficiency and its relationship to seed-cotton farmers could greatly aid policy makers both in creating efficient enhancement policies and in judging the efficiency of present and past reforms (Ogundari et al. 2006).

In light of producers' obvious failure, stochastic frontier costs are widely used and support empirical research using both cross-sectional and panel data, which can be used to help answer economic questions (Kumbhakar et al., 2015). Stochastic frontier models (SFMs) have been extensively applied in efficiency analyses of agricultural sectors, and in recent years they have captured the attention of researchers in developing countries (Awal Abdul and Rahaman (2016); Tim Coelli et al. (1999); Bravo-Ureta and Pinheiro (1997); Bravo-Ureta and Evenson (1994); Ali and Flinn (1989); and Hussain (1989); Taylor and Shonkwiler (1986), and Shapiro (1983)).

[1] Although China's farmers have not been forced by the government to grow cotton since 1998, the Provincial Seed Company maintains a strong lobby in favor of, for example, the use of Bt cotton.

Taking these findings into consideration, we estimated the stochastic frontier cost function translogarithm (TL) based on the framework of Wang (2002). This study was conducted to acquire a better understanding both of the cost efficiency of seed-cotton farms in China and of determinants of driver cost inefficiency despite the serious issue of economic theory. The study will help introduce a new dimension to farmers and policy makers concerning how to achieve optimal costs in cotton production by determining the extent to which it is possible to increase the efficiency of cotton-grain farms with the current resources base and the available technology. To address both quantity and quality insufficiency problems in China, the study will also shed light on the empirical literature by examining cost-efficiency measurements of Chinese cotton producers.

II. METHODOLOGY AND LITERATURE ON THE STOCHASTIC FRONTIER COST FUNCTION

The evolution of the concept and history of economic cost and production efficiency in the stochastic frontier is now almost forty years old since the first truly serious work, which was conducted independently in 1977 by Aigner et al. and Meeusen and Broeck (Coelli et al. 2005, Green 2008, Kumbhakar and Lovell 2000). Well before then, however, a great body of literature had been written to explain the existence of the frontier in the production or cost function, including Shapiro (1977), Stigler (1976), Toda (1976), Sidhu (1974), Bardham (1973), Yotopoulos and Lau (1973, 1971), Afriat (1972), Timmer (1971), Bhagwati and Chakravarty (1969), Christensen and Jorgenson (1969), Comanor and Leibenstein (1969), Zellner and Revankar (1969), Aigner and Chu (1968), Schultz (1964) and Sahota (1968), Hildebrand and Liu (1965), Leibenstein (1966), Nerlove (1963), Arrow et al. (1961), Johnston (1959), Farrell (1957), Shephard (1953), Debreu (1951), Koopmans (1951), Dean (1951), Samuelson (1938), Hicks (1935), Cobb and Douglas (1928), and Walras (1834-1910).

However, since the ALS and MB models in 1977, we have observed a great focus on elaborating the methodological and relaxation hypotheses for the computational of technical and cost inefficiencies to identify a more reliable estimation of empirical measurements (Lovell 1993, Greene 2003). The extensive model of Wang (2002) that we will consider in our study, the inefficiency effect model of Battese and Coelli (1992, 1995) with the heteroskedasticity variable of technical inefficiency, and the flexible model of Kumbhakar (1990). Among others, the models of Cornwell et al. (1990) and Lee and Schmidt (1993), along with the pertinent precision model of, inter alia, Stevenson (1980), applied. Consequently, the empirical estimation of production and economic (cost) functions has become a standard exercise in econometrics (Belotti and al. 2012).

The cost function *c(.)* is the function that provides the minimum cost of producing *y* at the factor price w_j. The neoclassical cost frontier model can be written as follows:

$$c_i = c(wj_i, y_i; \beta) \quad i = 1,2,...N; j = 1,..3 \qquad (1)$$

where c_i is the observed cost of farm *i*; w_{ji} is the j-th input price; y_i is the output quantity and β is a vector of technological parameters describing the relationship between input prices, output and the minimum production cost. Numerous properties following from cost minimization must be satisfied: nonnegative, non-decreasing (or non-regressive) in input prices *(w)* and output *(y)*; homogeneous degree of one in *w*; concave in *w*; and continuous in *w* (Kumbhakar et al. 2015, Coelli et al. 2005). The neoclassical cost frontier defined in Eq. (1) is a nonparametric cost function, viz., which overlooks the measurement error and random process (Aigner and Chu 1968, Afriat 1972, Farrell 1957). Therefore, a stochastic frontier (Aigner et al. 1977, Meeusen and van den Broeck 1977) formulation equation can account for the noise, and the inefficiency components can be written by appending the two component error terms into Eq. (1):

$$c_i \geq c(w_{ji}, y_i; \beta) \exp(v_i) \qquad (2),$$

where $v_i = \varepsilon_i - u_i$; therefore, by rearrangement, we can write equation (2) as follows:

$$c^a_i = c(w_{ji}, y_i; \beta) \exp(v_i + u_i) \qquad (3)$$

where c^a_i is the actual cost, which could be greater than the stochastic minimum production cost because of the additive inefficiency term; where ν_i is the idiosyncratic random error reflecting statistical noise that is independently and identically distributed (*iid*) and is usually assumed to follow the standard normal distribution with a mean of zero and a constant variance, i.e. $N(0, \sigma^2_v)$; and where u_i is a nonnegative producer-specific inefficiency error term that follows certain distributional assumptions (Aigner et al. 1977, Meeusen and van den Broeck 1977, Stenvenson 1980, Kumbhakar et al. 1991, Hang et Liu 1994, Greene 1980a,b, 2003). If the farms are completely efficient, the inefficiency error term will be zero, which indicates that the farms are operating on the ray of the stochastic cost frontier. The actual cost function can expressed as follows:

$$c^a{}_i = c(w_{ji}, y_i) \cdot \exp(u_i) \qquad (4),$$

where $u_i \geq 0$ and the remaining terms are explained above. However, if the production function is nonhomogenous, the cost function will assume the following form:

$$c(w_{ji}, y_i \exp(u_i)) \geq c(w_{ji}, y_i) \qquad (5)$$

Therefore, the measurement of CE_i can be defined as the ratio of the stochastic frontier cost (2) to the actual cost (3):

$$CEi = c(w_{ji}, y_i; \beta) \exp(\nu_i) / c(w_{ji}, y_i; \beta) \exp(\nu_i) \exp(u_i) = \exp(-ui), \qquad (6)$$

which by construction will be between 0 and 1. However, in stochastic frontier cost function of the production technology, the inefficiency could be between one and infinite (Belotti and al. 2012, Greene 2005, Kumbhakar and Lovell 2000). Thus, we use the cost frontier to judge producers' performance of producers relative to the optimal effect that can be achieved economically. Because one objective is to examine whether the exogenous variables have any influence on the cost product response to inefficiency, we can include in the exogenous variables of equation (7) below in equation (3),

$$\eta_i = \delta_o + z_{mi}\delta_m \qquad (7)$$

where δ_o is the intercept term, δ_m is the parameter for the *m-th* explanatory variables, and z_{mi}[2] is a vector of covariates of firm-specific characteristics and exogenous variables related to the idiosyncratic error term.

Furthermore, the maximum likelihood estimator (MLE) can be used to disintegrate the two error components and estimate the parameters of the stochastic frontier consistently and regularly (Greene 1980, Schmidt and Sickles 1984). Thus, producer-specific $CE_i = exp(-u_i)$ (8) and $CI_i = u_i$ (9) can be estimated using Battese and Coelli (1988) and Jondrow et al. (1982), respectively.

The stochastic frontier cost function (SFCF) analysis is relatively limited because of a difficult data set from farms limited by survey requirement times, human resources and sufficient money. To our knowledge, very little literature has addressed the issue of economic efficiency in cotton-grain production in China. In addition, no rigorous efficiency study has been conducted in the country's three main raw cotton-growing areas of China. The work of Fan (1991), which was based on an earlier work by Nishimizu and Page, estimated a simplified translog production frontier time series of panel data in Chinese agriculture using aggregate data from 29 provinces, municipalities, and autonomous regions. The study by Jirong Wang et al. (1996) was conducted to assess a shadow price frontier measurement of profit efficiency in Chinese agriculture. Jikun Huang et al. (2002) also used the stochastic frontier to explain the Bt cotton benefit, costs, and impacts in some provinces in Eastern China.

Other works examining cost efficiency address perspectives on EU cotton farming, namely, the technical and scale efficiencies of Greek cotton growers (Christos et al. 2002). The cost frontier was implied in a study of cotton farms in Turkey (Ebru and Nihat 2013, Ferit 2013, Erdal 2009, Ibrahim 2005) and in Nigeria by Odedokun et al. (2015). The technical efficiency of cotton growers has been examined in four counties in Western Texas (Kalyan et al., 2002). Another study

[2] Distributions are the Truncated Normal, $u_i \sim iid N^+(\mu, \sigma^2_u)$. Where $\mu = z_{mi}\delta$, this is the mean parameter of the truncated normal distribution, μ, which is modeled as a linear combination of the set of covariates, z_m.

performed a stochastic frontier analysis of cotton producers in West Africa (Veronique Theriault and Renata Serra 2013).

III. DATA COLLECTION

The dataset was derived from a survey of the China Academy of Agricultural Sciences Cotton Research Institute (CRICAAS) during one season from 2014-2015. Thus, we have carefully selected the variables that match the current concept and phenomena of the cost function during the production process of grain cotton in China.

Administratively, the production of raw cotton in China is divided into three agricultural locations: the Yellow River Basin, the Yangtze River basin, and Northern West Inland area. Taken together, the cotton sown in those areas accounts for approximately 30 percent of the total sown area of all the various cash crops, which together account for 99.5% of the cotton area with a total yield of 99.7% (Dai and Dong January 2014).

Nevertheless, based on their agroclimatic regions, the three locations mentioned above are in agricultural zones. The major zones, in which cotton is a prime crop, are Xinjiang in the Northwest Inland Region, the midstream and downstream basin of the Yangtze River basin, and Hueubei and the Huabei Lowland in the Yellow River Basin. Despite subregional disparity, altogether we had a population of 171 in the subregions. In the Yellow River Basin, the Northwest Inland Region, and the Yangtze River basin there were approximately 88, 50, and 31 sample subregions, respectively. Thus, our dataset was derived from three main regions producing seed cotton in China with a population of 171 subregions. Consequently, we had nearly 156 base counties, which permitted representation of 91.2% of the subregions producing cotton grain in China. Accordingly, major cottonseed farms were selected by consulting the department of agricultural extension in each province, district, township, and village. Because many farms in the Yellow River and Yangtze River basin, for example, were small size because of the agglomeration area and sprawling infrastructure, the selection of one farm for inclusion in the sample was set at a minimum size of one mu (Erling et al. 2017). Therefore, approximately 1253 farms were chosen using a systematic random sampling design in the Yellow River Basin, Northwest Inland Region, and Yangtze River basin of 544, 426, and 283 farms, respectively.

In addition, the data collection was performed with a primary target of evaluating the raw cotton cost through the use of a structured questionnaire similar to the ICAC[3]. The questionnaires included sections on the background characteristics of the farm households, e.g., education level and age[4], among others. Furthermore, there were sections concerning government rent, total land area used for cottonseed production[5], seeding cost, cost of the growth season, cost of the harvest season, loan interest, yield or cottonseed output, indirect costs[6], output prices, fixed costs, seed subsidies, self-agriculture insurance, sum of cotton production and hired workers, among others. We have not found convincing answers to questions about, for example, self-agricultural insurance and loan interest, for which there was too much missing information in the dataset; therefore, these variables were omitted. Furthermore, the prices of the various inputs were either pooled into an aggregate or bundled, clearly by respecting the nature and category of each input (Coelli and al., 2005). Doing so allowed us to obtain not only a limited number of well-manageable variables to examine the source of variability in the economic inefficiency of cottonseed in China's three producing regions but also good results without bias instead of more than one hundred variables that could lead to the loss of some important information but also miss the true value of the actual cost of cotton-grain production in China either because of missing information or because

[3] ICAC, International Cotton Advice Commitment.

[4] Exogenous variables characterize specific farmers' social characteristics that were used as a proxy variable for explaining farmers' experiences.

[5] In China, in many cases, farmers use the same land for either double-cropping (cotton-wheat) or multi-cropping (cotton-wheat-watermelon).

[6] Relative to rent, environment protection and cost, the nature and category of which were not found in the questionnaire. Therefore, to avoid underestimate the total cost we refused to exclude them. However, in the equation model, this variable was not included because some farmers did not pay the correct rent because of the size of their holdings.

the farmers were not using the same category of inputs.

Moreover, we specified that the input prices and total cost were expressed in RMB per mu. Although the output unit is expressed in kilograms per mu band, the output price is expressed in RMBs per kilogram. Therefore, the observations consist of the total costs, C_i, one output Y_i, the unit prices of the three inputs, W_{ni}, and the seven exogenous variables, Z_{mi}, which characterize three farmer-specific social characteristics and four other exogenous or explanatory variables that could explain the particularity of the space or the government policy setting, which is either less or uncontrollable by the farmers. In this class, we preferred to consider farm size as a variable to explain the heterogeneity related to the idiosyncratic noise term instead of as a characteristic of the economic inefficiency of specific farm decisions caused by particular features related to the country's geographical policy (its unique land policy) because farmers are not directly responsible for the choice of the farm dimensions (Erling et al. 2017).[7] The measured variables were as follows:

C_i= total cost of producing cottonseed = sum of the three cost items described below;
Y_i= output of the cottonseed for the season from 2014-2015;

W_{1i} (L_i)= price of labor = wages of the hired workers plus wages of the farmers;
W_{2i} (M_i) = price of material = sum of the prices of fertilizer plus herbicide plus pesticide
plus seed plus plastic mulching plus energy plus irrigation plus plant growth regulator (PGR) plus machine[8];

W_{3i} (F_i) = fixed cost[9];
Z_{1i} (P_i) = output price;
Z_{2i} (S) = seed subsidy received by a farmer;
Z_{3i} (F_Z) = farmer size;
Z_{4i} (AGE)= farmer age;
Z_{5i} (EDU)= farmer level of education EDU;
Z_{6i} (A_E) = interaction age and education;
Z_{7i} (Z)= zone of production is a three-level factor variable (Z_{7i1}= Northwest Inland region, Z_{7i2} =Yangtze River basin, Z_{7i3}= Yellow River Basin).

It is noteworthy that Z_{4i} (AGE), Z_{5i} (EDU), Z_{6i} (A_E), and Z_{7i} (Z) are in real forms, but the remaining variables are in logarithmic forms.

IV. EMPIRICAL MODEL

We start our framework by defining the cost frontier model as when the producer chooses variables as the input prices in the production process and the objective is to minimize costs given the required output. We estimate the effects of inefficiency in stochastic cost frontiers based on the more flexible model of Wang (2002). Thus, the stochastic frontier cost model allows inefficiency effects and traditional random error to be a function of a set of exogenous determinant variables based on a truncated-normal model, the parameters of which are estimated simultaneously in one step with the variates. This study focuses its attention on identifying the determinants of cost functions and assumes a cost minimization function when the input prices indicate economic inefficiency[10]. Therefore, the stochastic frontier is examined in the maximum likelihood estimation and multiple linear regression models used previously. A stochastic frontier model is applied, along with regression equations[11] with cost as a dependent variable and input price factors as endogenous variables along with exogenous variables. The variables in the stochastic frontier are transformed in log form. Thereafter, linear homogeneity is imposed to estimate the simple homogeneity of the input prices; thus, the variables considered are $C_i = \ln(C_i/w_{3i})$, $w_{ji} = \ln(w_{ji}/w_{3i})$, $w_{ki} = \ln(w_{ki}/w_{3i})$, and $y_{mi} = \ln(y_{mi})$. Accordingly, the resulting Cobb-Douglas cost frontier analysis can be written as follows:

[7] Although Chinese farmers have been free since 1978 to decide which type of crop to produce, leader counties have the power to orient the choices of other farmers.
[8] Fee paid by a farmer for activities operated by a machine, such as plowing, harrowing, and harvesting, among others.
[9] Saving for fixed assets, such as storage, tractors, tools for irrigation, devices used for pulverizing, pesticides and herbicides, among others.

[10] Because the input quantity information was unavailable, we assumed that the farms had full allocative efficiency.
[11] It is important to note the regression equations without the exogenous variables.

$$\ln c_i = \beta_0 + \sum_{j=1}^{3} \beta_1 \ln w_{ji} + \beta_y \ln y_i + \sum_{m=1}^{7} \delta_m z_{mi} + v_i + u_i$$

$$\ln c_i = \beta_0 + \sum_{j=1}^{3} \beta_1 \ln w_{ji} + \beta_y \ln y_i + \sum_{m=1}^{7} \delta_m z_{mi} + v_i + u_i$$

,-.,-.,--,-..,-.,-.,-.,--,-..,-.,-.,-.,-.,-.,--,-..,-.,-.,-.,-.,-.

The equation can be used to compare the result with the stochastic frontier trans-logarithmic below. The translog cost frontier model showing the typical symmetry is as follows:

$$ln\ C_i = \beta_0 + \sum_{j=1}^{3} \beta_j\ ln\ w_{ji} + \beta_y\ ln\ y_i + 0.5 \sum_{j=1}^{3} \sum_{k=1}^{3} \beta_{jk} \ln w_{ji} \ln w_{ki}$$
$$+ 0.5\ \beta_{yy} \ln y_i \ln y_i + \sum_{j=1}^{3} \beta_{jy} \ln w_{ji} \ln y_i + \sum_{m=1}^{7} \delta_m z_{mi} + v_i + u_i$$

,-.,-.,--,-..,-.,-.,-.,--,--,-...,-,-..,-,-..,-.,-,-..,-,-..,--,-..,-,-..,-,-..,--,-..,-.,-.,-.,-.,-.,--,-..,-.,-.,-.,--,--,-...,-,-..,-,-..,-.,-,-..,-,-..,--,-..,-,-..,-,-..,--,-..,-.,-.,-.

This equation is similar to equations (3) and (7), with the addition of the log function to both sides. Without including exogenous variables, this equation verifies the presence of inefficiency in the free distribution model[12] when the stochastic form specified in equation (2) is considered.

Following the approach of Wang (2002), which is the extension of both models of CFCFGH[13] motivated by the heteroskedasticity of the random variables v_i, the relationship of the model for KGMHLBC[14] addressed the issue of exogenous determinants of inefficiency u_i. This specification, as demonstrated in equations (3) and (7), enables the distribution of the inefficiency error term to vary between each observation.

The likelihood ratio statistic test (LR) or generalized likelihood ratio statistic can be conducted to test the specification of the models and the presence of the inefficiency term [15] and is given by

$$LR = -2*[L(H_0) - L(H_1)] \qquad (12),$$

where $L(H_0)$ and $L(H_1)$ are the values of the log-likelihood function under the null (H_0) and alternative (H_1) hypotheses. The restrictions form the basis of the null hypothesis (i.e., H_0: $\sigma^2_u = 0$), with the unrestricted model being the alternative hypothesis. The critical values of the mixed chi-square distribution, which are given in Table 1 of Kodde and Palm (1986), were used and compared with the LR statistic that was provided directly using the Stata software package (Stata/MP 10.1, 2010).

Therefore, we can easily estimate the efficiency technology parameters, α, β, δ, u, μ, and ν, in our framework computed using a truncated normal distribution with a maximum likelihood with mean μ_i and variance σ^2_{ui}, along with the variance of the appended idiosyncratic noise term σ^2_{vi}. Free distributions such as the Correct Least Square (COLS) proposed by Winsten (1957) and Correct Mean Absolute Deviation (CMAD) (Forsund and Hjalmarsson 1987) were determined to compute the efficiency. Thus, Schmidt and Lin (1984) proposed an OLS residual of the skewness to check the validity of the stochastic frontier specification of the model, and the skewness must be positive in the stochastic frontier cost function. This test can be strengthened by applying both the test developed by D'Agostino, Belanger, and D'Agostino Jr. (1990) and the M3T statistic suggested by Coelli (1995). Furthermore, thick frontier approach (TFA) and quantile regression were performed for the same purpose (Wang et al.

[12] The free distribution models is itemized below.
[13] Caudill and Ford (1993) and Caudil, Ford, and Gripper (1995) used the whin half-normal model.
[14] Kumbhakar, Ghosh, and McGuckin (1991), Reifschneider and Stevenson (1991), Huang and Lu (1994) and Battese and Coelli (1995) used the whin truncated-normal model.
[15] Examples include the test to compare the functional form of the Cobb-Douglas and Translog model and the estimation models for the distributional form, half-normal and truncated-normal by proposing hypotheses.

2008). The descriptive statistics of the frontier variables and the exogenous variables are depicted in table 1.

V. EMPIRICAL RESULTS

5.1 Commentaries on the Descriptive Statistics of the Dataset

The descriptive statistics of the variables are provided in table 1. Based on the means, medians, and standard deviations, considerable regional differences were observed. The inspections showed that the total cost of all farms in the Northwest Inland Region was greater than the median cost of our sample, whereas 200 farms in the Yangtze River basin had a total cost greater than the sample median. However, entire farms in the Yellow River Basin and their total cost of cottonseed production did not exceed the median cost of the sample. We found that 356 and 264 farms had a labor cost greater than the sample median in the Northwest Inland Region and the Yangtze River basin, respectively, whereas no farms satisfied these criteria in the Yellow River Basin. Further investigation revealed 625 farms with a fixed cost greater than the sample median (426, 98, and 101 farms in the Northwest Inland Region, the Yangtze River basin, and the Yellow River Basin, respectively). Discordance of the material cost was clearly observed in the Northwest Inland Region, the Yellow River Valley and the Yangtze River Valley (426, 191, and 1, respectively). The variability of the variables used in our study and their disagreement showed confidence with a very small p-value of less than 0.001, which showed that producers in the Northwest Inland Region had a higher comparative absolute cost of sowing or planting cotton than producers in the Yellow River Valley and Yangtze River Valley in the second position.

However, this analysis also showed that on the one hand, the Northwest Inland Region had a relatively higher fixed and material cost than the two regions in the east, whereas the Yangtze River and Yellow River regions had a relative labor cost higher than that in the northwest. On the other hand, farmers in the northwest used more materials (fertilizer, seed, deep irrigation, etc.) and invested more to obtain a greater yield by taking advantage of cheaper labor and the region's natural ecological conditions. The northwest region also appeared to be the best location economically to produce raw cotton because of its cheaper labor prices and the quality of the natural conditions, which provide a higher yield (see table 1). Thus, the farms in the northwest are mechanically intensive producers of cotton, whereas the farms in the east and center of the region are labor intensive.

Table 2: Descriptive Statistics of the Variables Used in the Study

	Northwest Inland Region *(N = 426)*			Yangtze River basin *(N = 283)*			Yellow River Basin *(N = 544)*			Whole Country *(N = 1253)*	
Variable	Mean	Median	Standard deviation	Mean	Median	Standard deviation	Mean	Median	Standard deviation	Mean	Median
Total cost	1903.86	1898	77.63	1540.6	1530.5	45.5	1176.6	1181.8	39.2	1506.1	1508
Labor	1077.43	1078	39.9	1079.2	1074	36.21	700.4	705	19.5	914.14	1033
Material	742.74	746	24.31	413.1	411	13.5	430.5	433	19.6	532.72	440
Fixed cost	40.2	40.5	7.2	26	25	4.8	21.07	19.5	6.3	28.7	26.5
Output	287.9	287	57.8	213.04	214	44.05	245.2	240	46.4	252.46	248
Output price	6.3	5.8	1.05	6.15	6	0.96	6.46	6.2	0.97	6.3	6.1
Subside	19.4	15	16.15	30.83	15	37.6	31.81	15	30.82	27.36	15
Farm size	36.7	26.5	34.3	9.76	9	6.2	12.45	8	18.96	20.09	11
Age	47	47	7	56	55	9.6	52	52	7.8	51	51

5.2 Efficiency with the Free Approach Distribution (FAD)

Table 2 shows the results of the free approach distribution of the residuals based on the estimated OLS. The skewness of the residuals has a value equal to 4.079. The positive sign indicates that the distribution of the residuals skews to the right, which is consistent with a cost frontier specification. We have found support for a right-skewed error distribution of the test proposed by Belanger and D'Agostino, Jr. (1990), in which the

skewness is considered statistically significant with a p-value of less than 0.001. The M3T statistic suggested by Coelli (1995) was also analyzed and showed that the skewness was significant with a M3T value greater than 1.96 of the normality distribution for rejection of the null hypothesis of the absence of skewness. Concerning the output statistics for the free approach distribution, the report confirmed an overwhelming rejection of the null hypothesis of no skewness in the OLS residuals. Thus, the existence of cost inefficiency within and between farm producers of cotton grain was confirmed in the different regions of China.

Additionally, analysis using the TFA indicated that the average cost of the inefficiency group farms was approximately 1.7 percent of the efficiency farm group average, which indicates that there was a lower gap value between the farmers in terms of complete efficiency and the farmers who failed to achieve complete efficiency to move to the ray of the frontier. Moreover, quantile regression revealed that for the least efficient farm group, the minimum cost was approximately 82 percent of the actual cost.

However, the estimated efficiency examined in the free approach excluded the existence of the disturbance of random error in the computation of cost efficiency. Consequently, there is a bias in the computation of cost efficiency.

Table 2: Estimation Results of Cost-Free Models

		Schmidt & Lin test	Coelli test	TFA	Quantile approach
Free distribution	Skewness	4.0791***			
	Kurtosis	62.6320***			
	M3T[1]		58.94 81***		
	Mean	2.1			
	Standard. Deviation	0.0031			
Inefficiency	Mean			0.017	0.1822
	Standard. Deviation				0.0317
	Minimum				0.1219
	Maximum				0.2172
	N	1253	1253	1253	313
Significance: * p<0.05, ** p<0.01, *** p<0.001					

The tests described above show the existence of cost inefficiency, but the production costs require some satisfaction under economic conditions of monotonicity of the input prices and output and on the concavity of input prices. Following Kumbhakar et al. (2015) to compute the economic inefficiency of MLE, it is appropriate to check the cost function properties. Thus, the monotonicity condition requires the input prices utilized to be positive for all observations. In the data sample, a violation of the monotonicity condition results in a fixed cost in the output. Our observations showed that only 6 of the 1253 observations violated the monotonicity condition for the output, among which 4 were in the Yangtze River Valley and 2 were in the Yellow River Valley. However, 183 of the 1253 observations did not satisfy the monotonicity condition for the fixed cost, and this violation was identified only in the Northwest Inland Region. This investigation showed that nearly 15 percent of the sample data violated the monotonicity condition and thus were considered to be too low.

In contrast, the concavity condition implied that the conditional input demand functions could not slope upwards, i.e., an increasing input price will not encourage its use. This phenomenon was examined by checking the negative semidefinite of the Hessian matrix for each input price of a farm for all data points. The results of the concavity test

were not consistent with the regularity condition of the cost function because there were 1074 violations in the dataset (approximately 86 percent of the sample population). At this stage, the assumption that producers minimize cost cannot support the theoretical requirements of cost function production.

Following Kumbhakar, we applied Shephard's Lemma to equation (3), which included the shared equations in the estimation model to reduce, in some cases, the number of violations. The implication is that in imposing more structure in the estimation process using the share equations, each input price can make the results align more with the theory.

The actual cost is

$$\ln ca(w, y,\eta) = \ln c^* (w, y) + \eta \qquad (13)$$

There are no changes in the terms of the equation above (13), and all terms in the equations are equivalent to the terms in equation (3). Therefore, the actual cost share S_j of input j is

$$S_j = w_j x_j / c_a = \partial \ln c^* / \partial \ln w_j,\ j=1,..3 \qquad (14)$$

We drop one equation share (input price equivalent to the homogeneity property cost) and then add the jth share equation j=2,..3

Finally, we obtain $Sj = \partial \ln c^* / \partial \ln w_{j} + \zeta_j,\ j=2,..3$ (15)

All items retained an unchanged distribution-free approach, which does not impose any distributional assumption η and ζj. We applied a seemingly unrelated regression equation (SURE) approach and subsequently adjusted ̂η to obtain the inefficiency index in the system equation.[16] The findings showed that the Schmidt and Lin (1984) test for skewness should be positive at approximately 0.31. By applying the noadj test as described by D'Agostino, Belanger, and D'Agostino Jr. (1990), we found that the skewness test for normality was meaningful with a p-value of less than 0.1 percent, as described previously.[17] The same result was obtained for the M3T statistic of Coelli (1995), for which we obtained a value of 4.45 greater than 1.96 of normality. Thus, the results confirmed the rejection of the null hypothesis of no skewness in the SURE residuals.

The average minimum cost efficiency in COLS was approximately 10 percent of the actual cost. In contrast, using the maximum likelihood estimators, the average optimal cost efficiency in the model system equation with a half-normal absence of correlation was 58 percent (Battese and Coelli 1995), and the average cost inefficiency described by Jondrow et al. (1982) was 85 percent. This large difference in the estimated measure of parameters in both approaches remains to be addressed. However, the likelihood ratio supported the existence of inefficiency in the model stochastic cost frontier function test, indicating an absence of economic inefficiency in the model in striking contrast to previously reported outcomes supporting the presence of inefficiency in the dataset (MT3 of Coelli and Schmidt & Lin).[18] The fitted values of the labor share, material share, and fixed share were 0.70, 0.15 and 0.15, respectively. These results align more with the theoretical requirements of the stochastic frontier cost function, demonstrating that on average, the labor cost share was 70 percent of the actual cost, the material cost share was 15 percent, and the fixed cost share was 15 percent. Furthermore, the sample averages of 0.86, 0.58, and 0.4 were reasonably coherent with the above-described input share coefficients.

We computed the returns to scale by first calculating the elasticity cost with respect to output as follows:

$ECY = \partial \ln C / \partial \ln y = \beta ly + \beta ly_2 + \beta_1 ldy + \beta_2 mtdy$ (16),

and thereafter, we observed a return to scale by determining the investment of the cost elasticity. Consequently, we can write the return to scale as follows:

$$RTS = 1/ECY \qquad (17)$$

Thus, we found diseconomies of scale with a return to scale of 0.32 less than 1, which indicates

[16] For those who wish to deepen their knowledge, please see Kumbhakar et al. (2015).

[17] For reasons of space, we do not present the results in the paper.

[18] The likehood ratio was -16856.334, the degree of freedom was one, and the critical values of the mixed chi-square distribution shown in Table 1, Kodde and Palm (1986, Econometrica) were 2.705 at a 5 percent level of confidence.

a decreasing return to scale[19]. All things remaining equal, producers of seed cotton in China experience higher costs with lower returns, and the reduced averages returning to scale are very large in the Yellow River Valley, followed by the Yangtze River Valley and slightly less in the Northwest Inland Region for values of approximately 0.11, 0.41, and 0.52, respectively. The average cost elasticity (EC) is approximately 1.12, indicating 2.12 in the northwest, 0.84 in the Yellow River Basin, and 0.15 in the Yangtze River basin.

It is worth noting that these findings were focused only on the half-normal model, with no correlation between the frontier equation and the shares equations. However, we are not confident that the results of the other models will be equivalent.[20]

Table 3 shows the results of the estimated coefficients using the free approach distribution. We found that all input prices and output were statistically meaningful, with a p-value of less than 0.001. The findings revealed the positive coefficients in input price for labor, material, and fixed costs range from (0.47, 0.38, and 0.0088) to (0.54, 0.40, and 0.08), respectively. However, the coefficient elasticity for the output was negative up to approximately 0.11%. The efficiencies of the free approach using both models (COLS and CMAD) were quite similar but with slight dispersion for the CMAD model in which the distribution of the cost efficiency had a larger tail (see figure 2). The average cost efficiency ranged from 94 percent to 100 percent for the COLS and from 92 to 100 percent for the CMAD.

[19] We are confident of this value in the sense that the estimates are neither too close to one another nor too close to zero.

[20] The empirical illustration for other models (partial and full correlation) was difficult to find because Stata's maximization procedure failed to achieve convergence to a solution; therefore, we encountered the issue about a flat or discontinuous region when Stata did the numerical derivatives. We did not wish to spend time looking for something that could not be found.

Table 3: Estimation Results of the Cost Frontier Models

		Model 1	Model 3	Model 4	Model 5	Model 6
		OLS	CMAD	Half-normal with no heterogeneity	Half-normal with heterogeneity	Truncated-normal
Frontier	llD	0.4706***	0.5391***	0.5014***	0.6113***	0.5897***
		(0.01586)	(0.01035)	(0.01477)	(0.01037)	(0.0082)
	lmtD	0.4052***	0.3816***	0.3917***	0.3565***	0.391***
		(0.01543)	(0.01009)	(0.01418)	(0.00902)	(0.00746)
	ly	-0.1131***	-0.1115***	-0.0768***	-0.1258***	-0.1181***
		(0.01546)	(0.00988)	(0.01565)	(0.00995)	(0.00833)
	llD2	0.2085***	0.2214***	0.2175***	0.1926***	0.20534***
		(0.00593)	(0.00388)	(0.00549)	(0.00373)	(0.00296)
	lmtD2	0.1915***	0.2214***	0.2077***	0.2181***	0.2230***
		(0.00696)	(0.00457)	(0.0064)	(0.00435)	(0.0039)
	ly2	0.02415***	0.0263***	0.018***	0.0329***	0.0313***
		(0.00255)	(0.00163)	(0.00261)	(0.00161)	(0.00132)
	lDmt	-0.1853***	-0.2065***	-0.1996***	-0.1955***	-0.2062***
		(0.00613)	(0.00402)	(0.00583)	(0.00385)	(0.00335)
	lDy	-0.0045***	-0.0103***	-0.0087***	-0.0146***	-0.0132***
		(0.0022)	(0.00142)	(0.00222)	(0.0015)	(0.00128)
	mtDy	0.0019***	0.0072***	0.0084***	0.00619***	0.0047***
		(0.00227)	(0.00148)	(0.00221)	(0.00145)	(0.00121)
	constant	1.2116***	1.1265***	1.0851***	1.0647***	1.0266***
		(0.04856)	(0.03108)	(0.04801)	(0.03299)	(0.02768)

5.3 Efficiency with the Maximum Likelihood Estimation (MLE)

The results of the estimated coefficients of the frontier variables and cost efficiency function using the maximum likelihood estimator are also presented in table 3. Several models were executed based on the type of distribution in the two component errors, along with the functional form that is commonly used in studies of the stochastic production economic function (Greene 2008, Kumbhakar et Lovell 2000). The null hypothesis that there is no technical inefficiency in the model was widely rejected by the test of the likelihood ratio with a p-value less than 0.001, indicating that the coefficients of the frontier cost function were significantly different from the average cost function estimated by OLS.

5.4 Model specification Test

The results in table (4) show that the LR was greater than the critical values for the mixed chi-square distribution of Kodde and Palm (1986) for all tests of the functional form between Cobb-Douglas and Translog; null hypotheses, which were used to select restrictive models in place of unrestrictive models, were rejected with a p-value of less than 0.1percent. In addition to the model in distributional form, the truncated-normal model described by Wang (2002) was preferred in our study because it was optimal for half-normal both with and without heterogeneity variables. However, between the half-normal model with heterogeneity and the half-normal model without heterogeneity, the LR was significant, and the null hypothesis that was not appended to heterogeneity in the model specification was overwhelmingly rejected with a p-value of less than 0.001.

Table 4: Generalized Likelihood Ratio Tests of Hypotheses for Parameters of SFCF for Cottonseed Farms in China

Models	Null hypotheses	LR	Critical value(x^2)	Dof	Decision
Functional form					
Model 4 VS Model 7	Ho: β4-β9=0	2691.6***	21.7	6	Reject Ho
Model 5 VS Model 8	Ho: β4-β9=0	2173.47***	21.7	6	Reject Ho
Model 6 VS Model 9	Ho: β4-β9=0	2081.9***	21.7	6	Reject Ho
Testing for variance parameters and inefficiency effects variables					
Model 4 VS Model 1	Ho: u = δo = 0	314.12***	9.5	1	Reject Ho
Model 5 VS Model 1	Ho: u = δo =...δ7 = 0	1036.7***	25.4	8	Reject Ho
Model 5 VS Model 4	Ho: u = δ1 =...δ7 = 0	722.6***	23.6	7	Reject Ho
Model 6 VS Model 1	Ho: u = δo =...δ7 = 0, Ho: μ = δo =...δ3 = 0	1125.6***	32.2	12	Reject Ho
Model 6 VS Model 4	Ho: u = δ1 =...δ7 = 0, Ho: μ = δo =...δ3 = 0	811.4***	30.5	11	Reject Ho
Model 6 VS Model 5	Ho: μ= δ0 =...δ3 = 0	88.9***	17.6	4	Reject Ho
Model 7 VS Model 1	Ho: u = δo = 0	-2377.5	9.5	1	Do not Reject Ho
Model 8 VS Model 1	Ho: u = δo =...δ7 = 0	-1136.8	25.4	8	Do not Reject Ho
Model 8 VS Model 7	Ho: u = δ1 =...δ7 = 0	1240.7***	23.6	7	Reject Ho
Model 9 VS Model 1	Ho: u = δo =...δ7 = 0, Ho: μ = δo =...δ3 = 0	-956.3	32.2	12	Do not Reject Ho
Model 9 VS Model 7	Ho: u = δ1 =...δ7 = 0, Ho: μ = δo =...δ3 = 0	1421.2***	30.5	11	Reject Ho
Model 9 VS Model 8	Ho: μ= δ0 =...δ3 = 0	-294.5	17.6	4	Do not Reject Ho
Significance: * p<0.05, ** p<0.01, *** p<0.001					

!The critical values of the mixed chi-square distribution are taken from Table 1 of Kodde and Palm (1986, Econometrica), which is obtained directly from Stata software.

5.5 Determinants of the stochastic cost frontier

The findings in table 3 show that the coefficient of the output in the maximum likelihood model was quite similar to the coefficients employed in the free approach model. However, the estimated coefficients of the input prices were different, especially for the half-normal model with heterogeneity, for which the coefficient was 61 percent for labor and 35.6 percent for material[21]. In addition, the elasticity coefficients of the input prices in the half-normal model were greater than the estimated coefficients in the truncated-normal model for labor costs, excluding those for material costs. However, the conventional likelihood ratio (LR) statistical tests were preferred in the truncated-normal model described by Wang (2002). Therefore, statistically, an increase in one unit of the input price could increase the actual cost by approximately 59 percent for labor and 39 percent for material, whereas a fixed cost could increase the actual cost by only 2 percent, ceteris paribus. This finding indicates that the input price for labor is the greatest determinant of the cost of producing seed cotton in China, followed by the material and finally the fixed cost. The coefficient of cost with respect to output was also negative, which was also the case for the free approach distributions. This negative output coefficient is consistent with the current situation in China for the cotton-production sector.

5.6 Index efficiency and its determinants in the stochastic frontier cost function

Reports of the efficiency index of the parametric models are shown in table 5 and indicate that the studied cotton farms were had almost full economic frontier efficiency. The average efficiency index of the stochastic frontier cost models were similar to one another. On average, the minimum cost was approximately 99.6 percent of the actual cost, which was the case in the Northwest Inland Region, the Yangtze River basin, and the Yellow River Basin.

However, following Kumbhakar, we derived the stochastic cost frontier model with a current half-

[21] Kumbhakar et al. (2015) said that it is no surprise that the coefficients in OLS were close to the coefficients in the stochastic frontier because of the consistency of the OLS estimates (except for the intercept).

normal model of output-oriented technical inefficiency.[22] The estimated efficiency index was similar to the input-oriented model in the full translog truncated-normal stochastic economic frontier model that was previously executed. The results indicate that on average, seed-cotton farmers produce approximately 99.6 percent of their maximum output, which means that only 0.4 percent of their potential output is lost because of technical inefficiency. Reports of the efficiency index have been almost the same for all regionally produced cotton grain. The similarity of the two approaches was a surprise to us. However, as mentioned previously, the assumption of cost minimization was not supported by the dataset used in our study, mutatis mutandis, which could also be true for the hypothesis of output maximization without increasing the input quantities.

Tableau 5: Estimation Results of the Cost Frontier Free Distribution and ML Models

	Observation	Mean	Standard Deviation	Minimum	Maximum
Model 2: COLS, eff_cols	1253	0.951824	0.001142	0.9435825	1
Model 3: CMAD, eff_cmad	1253	0.928347	0.0033479	0.9212758	1
Model 4: Half-normal with no heterogeneity, bc_hn	1253	0.996795	0.002855	0.9394708	0.9996913
Model 5: Half-normal with heterogeneity, bc_h	1253	0.998107	0.001142	0.9917311	0.9999183
Model 6: Truncated-normal, bc_t	1253	0.996189	0.0014865	0.9896166	0.9999789

The results for the determinants of cost inefficiency are reported in table 7. As shown in table 7, the estimated coefficients of farmers' specific characteristics related to cost efficiency were 1.30174, 0.0079, and 0.0017 for level of education, age, and the interaction of age and education, respectively. This outcome appears to indicate that the effects of producer-specific characteristics are very small in terms of the inefficiency production cost function, except for education, for which the estimated coefficient was 1.3. However, the critical point for the p-value was not significant. This result implies that the inclusion of education, age, and the interaction of age and education[23] in the models was not sustained by the data. In addition, marginal effects on the unconditional expectation of u, E(u), and the unconditional variance of u, V(u), were calculated, and the output revealed that the mean of the marginal effects on both E(u) and V(u) were positive but not significant.

However, eliminating the test statistic provided insignificant results when the effect of farmer age is considered. On average, both the level of cost inefficiency and the uncertainty of the cost inefficiency were affected. Tabulation of the data revealed that 590 of 1253 observations in the sample had an age greater than 51 years and that only 87 observations in the sample had an age below 40 years. In contrast, with the farmer's lower education level, on average, both the level of inefficiency and the uncertainty of the cost were greater than for farmers with a higher education level, and a significant change was noticeable for farmers with a high school education (see figure), beyond which there was no significant variation. Thus, training can help producers efficiently manage a farm to minimize costs by accessing information using communication tools for the verification and knowledge of input prices in various markets.

[22] The equation model considers the output-oriented cost efficiency to be the same as that employed using input-oriented cost efficiency, with the exception of specifying a restriction of the interaction, which includes all interaction terms between wj,i and yi and the square term of yi in the equation half-normal translogarithmic model (Kumbhakar et al. 2015).

[23] We had found support for including interaction of age and education in the model based on the fact that the link test for model specification performed in the least square linear regression told us that our dependent variable was well specified with the interaction variable. Therefore, the interaction of age and education had explanatory power with a p-value of less than 0.1 percent. Additionally, the cross-validation underpinned on Akaike and Schwarz's information criterion was smaller than the model without interaction.

Furthermore, with respect to the interaction of age and education, our survey demonstrated that the effect was negative for a majority of the farmers; however, at certain levels, the effect of the association between age and education on both cost inefficiency and risk became positive (increased), indicating that there was not only an optimal level of both education and age beyond this minimum level of uncertainty but also an increase in the inefficiency cost. The sorted data obtained in our inspection indicated that the middle of high school was the optimal level for education and 51 years old was the optimal level for age.

Furthermore, table 7 provides the estimated coefficients of the exogenous variables in association with the idiosyncratic noise terms. The results reveal that the estimated parameter values were 3.75, 0.3, -0.2, 0.14, 0.46, and -3.99 for the farm size, output price, seed subsidy, and for the three production zones in the Northwest Inland Region, the Yellow River Basin, and the Yangtze River basin, respectively. The results showed that farm size had a greater positive effect on the idiosyncratic error term, followed by the regional factor variable, which had a negative effect value; both were significant at a p-value of less than 0.001 and 0.001, respectively. This result indicated that the overall area for the production of raw cotton significantly influenced the change in the variance (by approximately 3.8-fold) of the conventional randomness error. In other words, the risk more than tripled when the land area was increased to obtain a cost in the upper region of the seed cotton to scale. Farm location had a negative and significant effect on the magnitude of the idiosyncratic error term (approximately double).[24] In other words, farm location mitigates nearly doubled the uncertainty of achieving an increased total cost of producing seed cotton if all other factors remained unchanged.

In addition, the estimated coefficients of the virtual indicator variables of the zone production above indicated that Z3 (Yellow River Basin) differs from Z1 (Northwest Inland Region) by -4, and this difference is significant at the 0.1% level. In contrast, Z2 (Yangtze River Basin) is not significantly different from the virtual indicator of the production location Z1.[25] However, the joint test to determine whether the coefficients of virtual indicators Z1, Z2, and Z3 were all equal was overwhelmingly rejected.[26]

[24] The overall estimated coefficient of the production location was 1.836749 when the zone of production factor variable was considered a simple categorical variable. Note that in this model, the parameter values of the other variables were unaffected, meaning that they were almost the same as when the production location was considered using three level virtual indicator variables. However, the result of the truncated-normal model when the zone production was considered a simple categorical variable is not reported here

[25] Z1, Z2, and Z3 are the virtual indicator variables that characterized the zone of production factor variable where Z1 is the base level.

[26] The linear regression test of the hypotheses after the estimation revealed that the chi-square value was 31.28 with two degrees of freedom, and the p-value was less than 0.001.

Table 6: Estimation Results of the Cost Frontier Models

		Model 1	Model 4	Model 5	Model 6
		OLS	Half-normal with no heterogeneity	Half-normal with heterogeneity	Truncated - normal
σ_v^2	lfz			1.8215***	3.7571***
				(0.0949)	(0.24389)
	lp			0.3319	0.2925
				(0.6982)	(0.91353)
	lsb			-0.3034**	-0.2223
				(0.116)	(0.20841)
	z_1			0.0093	0.1408
	z_2			0.2056	0.4606
				(0.26756)	(0.82056)
	z_3			-0.5053*	-3.9902***
				(0.23209)	(0.73875)
	constant		13.6885***	-17.8024***	25.9585***
			(0.14078)	(1.30174)	(1.97219)
σ_u^2	edu			-0.1194	-0.0791
				(0.20615)	(0.22542)
	age			0.0079	0.011
				(0.01853)	(0.02078)
	a_edu			0.0017	0.0025
				(0.00386)	(0.00435)
	constant		10.8377***	-12.36562***	-13.5571***
			(0.0565)	(0.97371)	(1.06302)
μ	edu				0.00007
					(0.00025)
	age				0.0000086
					(0.00002)
	a_edu				-0.000003
					(0.000005)
	constant				0.0036
					(0.00109)
marginal effect of edu on E(u)				0.0001	0.000053
marginal effect of edu on V(u)				0.00000025	-0.00000015
marginal effect of age on E(u)				0.0000074	0.00001
marginal effect of age on V(u)				0.000000016	0.00000003
marginal effect of age_edu on E(u)				0.0000016	-0.000002
marginal effect of age_edu on V(u)				0.000000003	0.000000005
log likelihood		5493.4658	5650.5237	6011.8173	6056.2466

Note: Standard errors in parentheses

*Significance: * p<0.05, ** p<0.01, *** p<0.001*

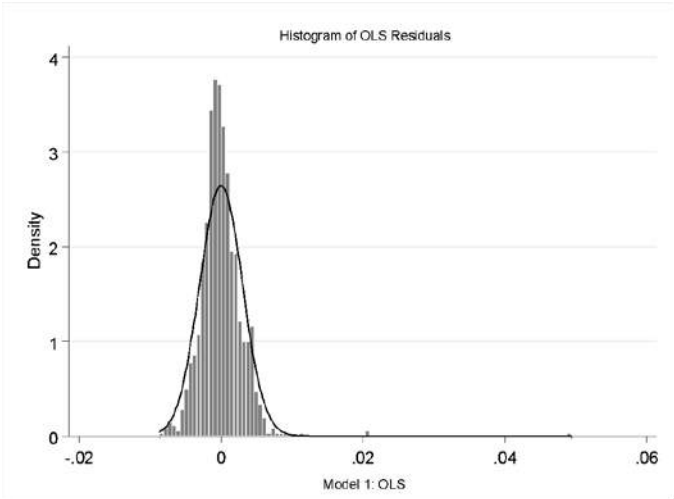

Figure 1

Kendall's Correlation Coefficients

Tables 7 and 8 show Kendall's rank correlation coefficient (τ) between the efficiency index. The null hypothesis of independence between the pair index is rejected at the 0.1 percent level in every case. The findings also showed both that the free approach to distribution efficiency ranking tended to correlate positively and the frontier approach efficiency ranking tended to correlate positively. Additionally, the free approach distribution efficiency ranking tended to negatively correlate with a stochastic frontier model without heterogeneity rather than a stochastic frontier with heterogeneity.

Table 7: Kendall's Correlation Coefficient (τ)

		Model 3: CMAD	Model 4: Half-normal with no	Model 5: Half-normal with	Model 6: Truncated-normal
		eff_cmad	heterogeneity bc_hn	heterogeneity bc_h	bc_t
Model 2: COLS, eff_cols		0.7412***	-0.8217***	-0.4873***	-0.5226***
Model 3: CMAD, eff_cmad			-0.8128***	-0.6327***	-0.6644***
Model 4: Half-normal with no heterogeneity, bc_hn				0.5403***	0.568***
Model 5: Half-normal with heterogeneity, bc_h Significance: * p<0.05, ** p<0.01, *** p<0.001					0.8035***

Table 8: Economic Efficiencies of Cotton Farms in China by Zone Production

Model	Northwest Inland Region *(N = 426)*	Yangtze River basin *(N = 283)*	Yellow River Basin *(N = 544)*	Whole Country *(N = 1253)*
Model 2	0.9513	0.9517	0.9522	0.9517
Model 3	0.9284	0.9282	0.92833	0.92831
Model 4	0.9967	0.9971	0.9966	0.92831
Model 5	0.9982	0.9979	0.998	0.998
Model 6	0.9962	0.9963	0.9961	0.9962
Model 7	0.9832	0.98820	0.9957	0.9897
Model 8	0.9881	0.9989	0.9986	0.9951
Model 9	0.9929	0.9923	0.9941	0.9933

IIV. DISCUSSION AND CONCLUSION

In this paper, we estimated the stochastic cost frontier function to examine the inefficiency of Chinese seed-cotton producers in three main locations. The test results of our study indicate that the assumption of cost minimization based on input prices orienting cost inefficiency may be inappropriate for this sampling of data. Thus, the cost efficiency was viewed as an estimated relationship in the objective to explain the determinants of costs of cotton-grain production in China. The homogeneous estimation of cost inefficiency in different regions indicated consistency of the integration of the cotton-production sector in China. Our study's results, however, suggest significant discordance both in input prices and in different variables. In our empirical stochastic economic frontier, the total cost of the dependent variable was significantly influenced by all the frontier explanatory variables (p-value less than 0.001), along with the overall farm area and farm location with the exception of the farmer's education level and age. Our results were consistent with those of Odedokun et al. (2015), who showed that labor price had a positive impact on production cost.

However, Ebru et al. (2013), in a study examining cotton production in Turkey, did not find that farm size had an effect on technical efficiency. Considering the determinants of cost in our study, with a higher labor cost, the previous study showed evidence of our current findings (Jrong Wang et al., 1996; Jikun Huang et al., 2002). Dai and Dong (Dai and Dong, 2014) demonstrated and argued that the main cause could be explained by urbanization. Jikun et al. (2002) specified that for farmers to produce the cotton required for 500 persons per hectare (approximately 33 persons per mu), approximately the same results were obtained for the estimation of cotton production for approximately 22 workers per mu. The majority of workers were old women and children because intensive cotton farming in China requires a great deal of labor; therefore, the farmers are paid for long hours of work, especially in the Yellow River Valley and the Yangtze River Valley. Hsu and Gale (2001) provided the same evidence, noting that the increase in labor costs was among the main reasons for producers in the Yellow River area to displace land for relative crop returns. Adshead (1997) provided a good reason for this phenomenon, stressing that Chinese cotton was a household crop, not a plantation crop. The shift in work created a labor shortage and reduced efficiency.

Our results showed that under these circumstances, farmers in the Northwest Inland Region are in a good position. Approximately 15 percent of the cotton-growing area is harvested with machinery in the Northwest Inland Region, whereas in the Yellow River Basin and the Yangtze River basin, harvesting is performed manually (Dai and Dong, 2014). Jirong et al. (1996) emphasized that the farmland and higher population density may force farmers to overuse input, increasing cost inefficiencies and thus reducing profit efficiencies. The land constraints offset the losses through the total average cost. Even if our findings indicate that farm size could significantly affect inefficiency, this may result only from agriculture's substantial socioeconomic, ecological contribution at the regional level. In this dataset, the average farm size in the Northwest Inland Region was approximately 37 mu, whereas it was 10 and 12.5 mu in the Yangtze River basin and the Yellow River Basin, respectively. The average cotton yield was 288 kilograms per mu, whereas in the Yangtze River Basin and the Yellow River Basin it was 213 and 245 kilograms per mu, respectively. Although it is difficult to link the yield and farm size to justify this discordance, it is clear that the farmers in the Northwest Inland Region were more competitive. We can affirm our conjecture based on Xufu and Clem's (2009) claim in a comparative study of cotton production in China and Australia, stressing that cotton growth in China depended on variation in yield but not on farm area (size). In contrast, according to Wang and Li (2006), farmers in the Northwest Inland Region are less susceptible to the low price of raw cotton and are less swift to reduce production and therefore cost than farmers in the eastern and southern regions, where they shift the land rapidly to alternative crops. Both authors reached the same

conclusion[27], although Wang and Li were more concrete, and we understand that location plays a key role in the stability of crop production.

This study showed that regional production significantly contributes to the magnitude or shock of the total cost of producing seed cotton. The results showed that the cost of cotton-grain production depends not only on input prices in the market alone but also on farm size. Thus, we cannot assume that cost minimization impacts producer behavior. We encourage an investigation to verify whether producers profit from efficiency by assessing farmers' technical and allocative efficiency. Finally, we claim that the policy must continue to specialize in regional agriculture to decongest the sector without overlooking the responsibility not only to provide employment opportunities to its ever-increasing labor force but also to support environmental sustainability.

REFERENCES

1. Abedullah, Khuda Bakhsh, and Bashir Ahmad. 2006. Technical efficiency and its Determinants in Potato Production, Evidence from Punjab, Pakistan. The Labore Journal of Economics 11: 2 (Winter 2006) pp. 1-22.
2. Adshead, S.A.M. 1997. Material Culture in Europe and China, 14000-1800: The Rise of Consumerism.
3. Aigner, D.J., C.A.K. Lovell and P. Schmidt. 1997. Formulation and Estimation of Stochastic Frontier Production Models. Journal of Econometrics 6, 21-37.
4. Aigner, D.J. and Chu, S.F. 1968. On Estimating the Industry Production Function. American Economic Review, 58(4), 826 39.
5. Afriat, S.N. 1972. Efficiency Estimates of Production Functions. International Economic Review 13, 568-98.
6. Ali, M. and J.C. Flinn. 1989. Profit efficiency in Basmati Rice Production. America Journal of Agriculture Economics 71(1989): 303-310.
7. Ali, M. and M.A Chaudhry. 1990. "Inter-regional Farm Efficiency in Pakistan's Punjab: A Frontier Production Function Models", Journal of Econometrics 6, 21-37.
8. Away Abdul – Rahaman. 2016. Stochastic frontier analysis (SFA) of technical efficiency, insights from smallholder cotton farmers in the Northern Region of Ghana. Global Science Research Journals, Vol. 4(1), pp.361 – 367.
9. Bardham, P.K. 1973. Size, Productivity and Returns to Scale: An Analysis of Farm Level Data in India Agriculture. Journal of Political Economy 18, 1370-86.
10. Belotti, F., Daidone, S., Hardi, G., and Atella, V. 2012. Stochastic frontier analysis using Stata. CEIS Tor Vergata, Volume 10, Issue 12, 251.
11. Bhagwati, J.N., and S Chakravarty. 1969. Contributions to India Economic Analysis: A Survey. American Economic Review 59, 2-67.
12. Bravo-Ureta, B.E., and Pinheiro, A.E. 1997. Technical, Economic and Allocative Efficiency in Peasant: Evidence from the Dominican Republic, The Developing Economies, 35 (1): 48-67.
13. Bravo-Ureta, B., and L. Rieger. 1991. "Dairy Farm Efficiency Measurement Using Stochastic Frontiers and Neoclassical Duality." American Journal of Agricultural Economics, 73, 421-28.
14. Bravo-Ureta, B.E., and Robert E. Evenson. 1994. Efficiency in Agricultural Production: The Case of Peasant in Farmers in Eastern Paraguay. Agricultural Economics 10 (1994): 27-37.
15. Chaudhry, M.R. 2008. Update on Costs of Producing Cotton in the World. Presented at the 29th International Cotton Conference. Bremen, Germany.
16. Chennareddy, V. 1967. Production Efficiency in South Indian Agriculture. Journal of Farm Economics 49, 816-20.
17. Christensen, L.K. and Jorgenson D. W. 1969. The Measurement of US Real Capital Input 1919 – 69. Review of Income and Wealth, 15, n. 4.
18. Cobb, S. and P. Douglas. 1928. A Theory of Production. American Economic Review 18, pp. 139-165.
19. Comanor, W.S., and H. Leibenstein. Allocative Efficiency, X-Efficiency and the Measurement of Welfare Losses. Economica 36, 304-9.
20. Dai, Jianlong, and Dong, Hezhong. 2014."Intensive cotton farming technologies in

[27] The study by Xufu and Clem was performed at the macroeconomic level, whereas the study by Wang and Li was performed at the microeconomic level, facilitating understanding of producers' specific behavior when using the aggregate data.

China: Achievements, challenges, and countermeasures". Science Direct Journal.
21. Debreu, G. 1951. The Coefficient of Resource Utilization. Econometrica 19, pp. 273-292.
22. Domingos Sárvio Magalhães Valente, Daniel Marçal de Queiroz, Francisco de Assis de Carvalho Pinto, Nerilson Terra Santos e Fábio Lúcio Santos. 2012. Revista Ciência Agronômica, v. 43, n. 4, p. 683-690.
23. Erling Li, Ken Coates, Xiaojian Li, Xinyue Ye, and Mark Leipnik. 2017. Analyzing Agricultural Agglomeration in China. MDPI Journal.
24. Farrell, M.J. 1957. The Measurement of Productive Efficiency. Journal of the Royal Statistical Society (A, general), 120, 253 – 281.
25. Forsund, F.R., and Hjalmarsson, L. 1987. Analysis of Industrial Structure: A Putty-Clay Approach. Stockholm: Almqvist & Wiksell.
26. Greene. W.H. 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of Econometrics, 126, 269-303.
27. Green W. 2008. The Measure of Efficiency, chap. The Econometric Approach to Efficiency Analysis. Oxford University Press.
28. Greene, W.H. 1980. On the estimation of a flexible frontier production model. Journal of Econometrics, 13(1), 101-15.
29. Hadri K, Whittaker J. 1999. Efficiency, Environmental Contaminants and Farm Size: Testing for Links using Stochastic Production Frontiers. Journal of Applied Economics, 2: 337-356.
30. Hicks, J. 1935. The Theory of Monopoly: A Survey. Econometrica 3, pp. 1-20.
31. Hsu, H., and Gale, F. (2001). Regional shifts in China's cotton production and use. In United States Department of Agriculture Economic Research Service, Cotton and Wool Situation and Outlook. Washington DC: USAD ERS.
32. Huang, J., Hu. R., Rozelle, S., Pray, C.E., and Wang, Q. 2002. Plant biotechnology in the developing world: the case of China. Science, 295(25), 674-677.
33. Huang, J., Hu. R., Rozelle, S., Qiao, F., and Pray, C.E. (2002). Transgenic varieties and productivity of smallholder cotton farmers in China. Australian Journal of Agricultural and Resource Economics, 46(3), 367-388.
34. Hussain, S.S. 1989. Analysis of Economic Efficiency in Northern Pakistan: Estimation, causes and Policy Implication. Ph.D. Diss., University of Illinois.
35. Jrong Wang, Eric J. Wailes, and Gail, L. Cramer. 1996. A Shadow-Price Frontier Measurement
36. of Profit Efficiency in Chinese Agriculture. American Agriculture Economics Association, 78: 146-156.
37. Koopmans, T.C. 1951. An analysis of Production as Efficient Combination of Activities, in Activity Analysis of Production and Allocation, Koopmans, T.C., eds, Cowles Commission for Research in Economics, Monograph no. 13. New York.
38. Kumbhakar SC. 1997. Efficiency Estimation with Heteroscedasticity in a Panel Data Model. Applied Economics 29: 379-386.
39. Lau, L.J., and P. A. Yotopoulos. 1971. A Test for Relative Efficiency and Applications to Agriculture. American Economic Review 61, 94-109.
40. Laurits, R. Christensen, Dale W. Jorgenson, Lawrence J. Lau. 1973. The Review of Economics and Statistics, Volume 55, Issue 1, 28-45.
41. Leibentein, H. 1966. Allocative Efficiency vs. 'X-Efficiency'. American Economic Review, 56, 392-415.
42. Meeusen, W. and J. van den Broeck. 1977. Efficiency Estimation from Cobb-Douglas
43. production Functions with Composed Error. International Econometric Review, 18, 435-444.
44. Muhammad Sajid Hussain, Tim Coelli, and Phil Simmons. 1999. 43rd Annual AARES.
45. P. Paudel, A. Matsuoka. 2009. Cost efficiency estimates of maize production in Nepal: a case study of the Chitwan district Agric. Econ. Czech, 55.
46. Pray, C.E., Huang, J., Ma, D., Qiao, F. 2001. The impact of Bt cotton in China. World Development 29, 813-825.
47. Qin Zhag. 2015. Precision Agriculture Technology for Crop Farming. CRC Press Taylor & Francis Group Boca Raton London Newyork 232.
48. Sahota, G.S. 1968. Efficiency of Resource Allocation in Indian Agriculture. American

Journal of Agricultural Economics 50, 584-605.
49. Shapiro, K.H. 1983. Efficiency Differences in Peasant Agriculture and their Implications for Development Policies. Journal Development Studies 19 (1983):179-90.
50. Shapiro, K.H., and J. Muller. 1977. Sources of Technical Efficiency: The Roles of Modernization and Information. Economic Development and Cultural Change 25, 293-310.
51. Sidhu, S.S. 1974. Relative Efficiency in Wheat Production in the Indian Punjab. American Economic Review, 64, 742-51.
52. Stephen MacDonald and Fred Gale and James Hansen. 2015. Cotton Policy in China. U.S. Department of Agriculture, Economic Research Service.
53. Stigler, C.J. The Existence of X-Efficiency. 1976. American Economic Review 66, 213-26.
54. Subal C. Kumbhakar and C. A. Knox Lovell. 2000. "Stochastic Frontier Analysis", Cambridge University Press, Cambridge CB22RU, UK
55. Subal C. Kumbhakar, Hung-Jen Wang, and Alan P. Horncastle. 2015. A Practioner's Guide to Stochastic Frontier Analysis Using Stata. Cambridge University Press.
56. Taylor, G.T., and Shonkwiler, J.S. 1986. Alternative Stochastic Specification of the Frontier Production Function in the Analysis of Agricultural Credit Programs and Technical Efficiency, Journal of Development Economics, 21: 149-160.
57. Thath Rido. 2014. Factors Affecting Cost Efficiency of Cambodian Rice Farming Households. Forum of International Development Studies. 45–2.
58. Timmer, C.P. 1971. Using a Probabilistic Frontier Function to Measure Technical Efficiency for Ontario Dairy Farms. Canadian Journal of Agricultural Economics, 38, 439-456.
59. Timothy J. Coelli, D.S. Prasada Rao, Christopher J. O'Donnell, and George E. Battese. 2005. An Introduction to Efficiency and Productivity Analysis Second Edition. Springer Science and Business Media.
60. Toda, Y. 1976. Estimation of a Cost Function When the Cost is not Minimum: The Case of Soviet Manufacturing Industries, 1958-71. Review of Economic and Statistics 58, 259-68.
61. USDA Foreign Agricultural Service: Global Agricultural Information Network, report January 2017, acceded 2017/1/15).
62. Walras, L. 1954 (1874). Elements of Pure Economics. London, George Allen and Unwin.
63. Wang, H.-J. 2002. "Heteroscedasticity and Non-Monotonic Efficiency Effects of a Stochastic Frontier Model," Journal of Productivity Analysis, 18, 241–53.
64. Wang, H.-J., and Schmidt, P. 2002. "One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels," Journal of Productivity Analysis, 18, 129–44.
65. Wang, M.H. 2011. Precision Agriculture. Beijing, China: China Agriculture University Press.
66. Wang, X., C.J. Zhao, Q.J. Meng, L.P Chen, Y.C. Pan, and X.Z. Xue. 2003. Design and experiment of variable rate fertilizer applicator. Transactions of the CSAE, 20(5):114-117.
67. Wang, C.X., and Li, D. (2006). Analysis of Cotton Policy, Price Fluctuation and the Productive Investiment Behavior of Cotton Farmers – Taking Xinjiang Cotton Region as an Example. Price Theory and Practice, (11): 40-41. [In Chinese].
68. Xufu Zhao and Clem Tisdell. 2009. Economic Theory, Applications and Issues. ISSN: 1444-8890.
69. Zhao, C.J., X.Z. Xue, X. Wang, L.P. Chen, Y.C. Pan, and Z.J. Meng. 2003. Advance and prospects of precision agriculture technology system. Transaction of the CSAE, 12(4):7-1.

Table 9: Kendall's Correlation Coefficients (τ) For the Functional Form Analysis

	Model 7	Model 8	Model 9
Model 2	0.0022	0.0377*	-0.2702***
Model 3	-0.1195***	-0.0086	-0.2478***
Model 4	0.0579**	0.233***	0.2166***
Model 5	0.233***	0.074 ***	0.3707***
Model 6	0.2166***	0.0906***	0.3603***
Model 7		0.2848***	0.4439***
Model 8			0.2971***
Significance: * p<0.05, ** p<0.01, *** p<0.001			

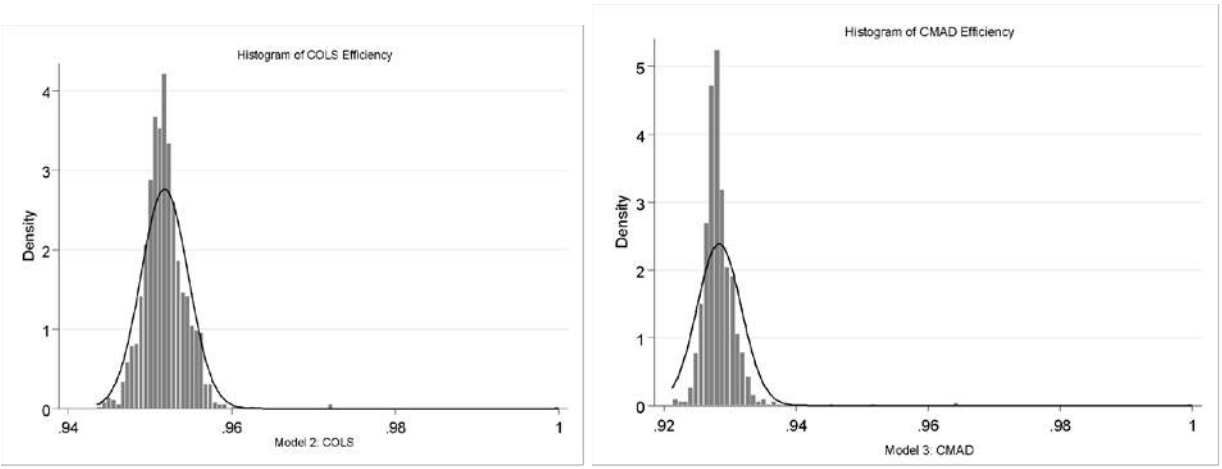

Figure 2: Efficiency of COLS and CMAD

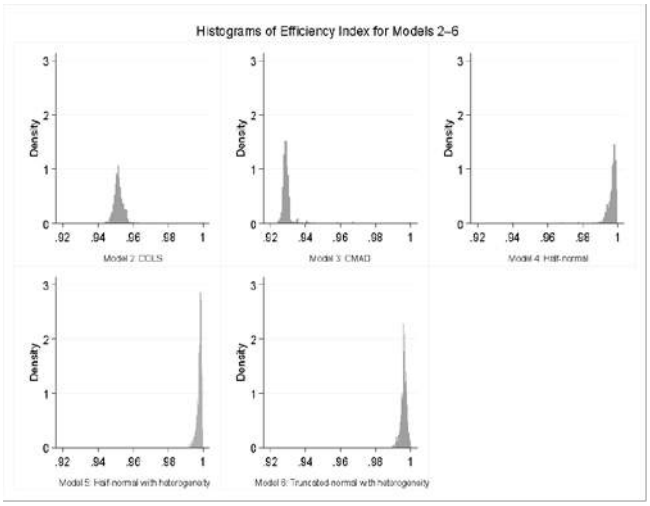

Figure 3: Efficiency of all models

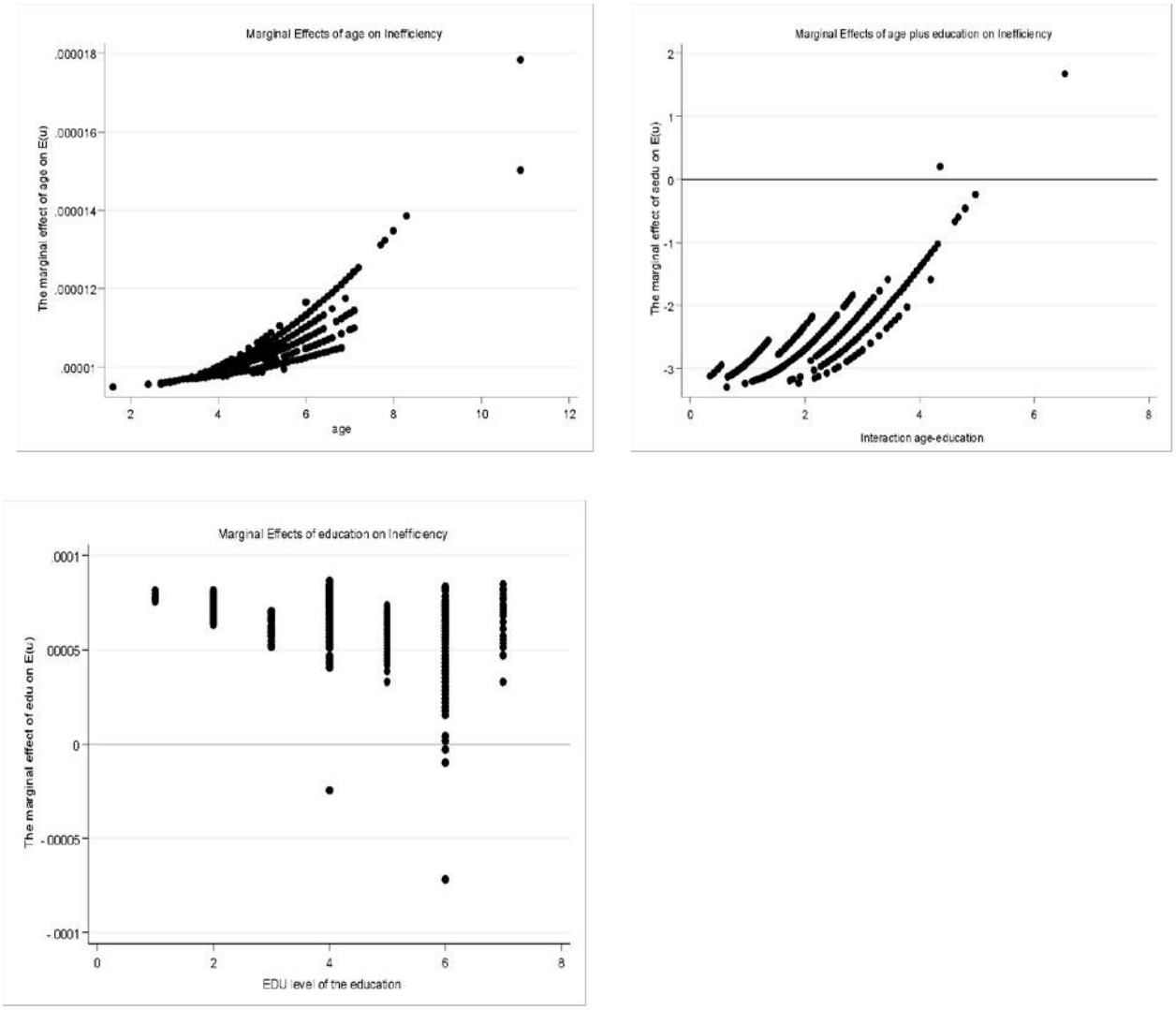

Figure 4: Marginal effects on inefficiency

This page is intentionally left blank