International Journal of Advanced Culture Technology Vol.6 No.4 292-302 (2018)
DOI https://doi.org /10.17703//IJACT2018.6.4.292

IJACT 18-4-37

The Effects of Human Resource Factors on Firm Efficiency: A Bayesian Stochastic Frontier Analysis

Sangwoo Shin[1], Hyejung Chang[2]
[1]*Solbridge International School of Business, Woosong University, Daejun, South Korea,*
[2] *Department of Management, Kyung Hee University, Seoul, South Korea*
sangwoo.shin@solbridge.ac.kr, hjchang@khu.ac.kr

Abstract

This study proposes a Bayesian stochastic frontier model that is well-suited to productivity/efficiency analysis particularly using panel data. A unique feature of our proposal is that both production frontier and efficiency are estimable for each individual firm and their linkage to various firm characteristics enriches our understanding of the source of productivity/efficiency. Empirical application of the proposed analysis to Human Capital Corporate Panel data enables identification and quantification of the effects of Human Resource factors on firm efficiency in tandem with those of firm types on production frontier. A comprehensive description of the Markov Chain Monte Carlo estimation procedure is forwarded to facilitate the use of our proposed stochastic frontier analysis.

Keywords: *Bayesian stochastic frontier analysis; Human resources; Human Capital Corporate Panel data; Markov Chain Monte Carlo*

1. Introduction

The fact that human resources (HR) play a role as a firm's valuable intangible assets has been widely recognized. Beyond a traditional perception of HR as costs to be minimized, a commonly agreed view is that HR serves as a source of value creation via efficiency enhancement and/or revenue growth [1-3]. A more extended view of HR as a strategic lever postulates that HR have economically significant impacts on organizational performance, which in turn can be translated into sustainable competitive advantage [4-6].

Despite this concensual prominence of HR, empirical validation and subsequent quantification of its economic effects have been rather limited in the academic literature. Empirical work in this vein, if conducted anyway, has often been far-fetched from modeling rigor. In addition, the resulting outcomes in many cases are not rich enough to properly measure the size of the effects of HR on organizational performance.

A major hindrance is that researchers have very limited access to HR data of a good quality. First, HR data often contain highly sensitive private information and, therefore, is not publicly available. The HR information observable directly by an outsider has rather small informational value. The HR information hand-collected through survey (by an individual researcher) does not assure information quality either. The generic shortcomings of survey such as high non-response rates and inaccuracy of responses, due to the very nature of

Manuscript Received: November 20, 2018 / Revised: November 28, 2018 / Accepted: November 30, 2018
Corresponding Author: hjchang@khu.ac.kr
Tel:+82-02-961-9432, Fax: +82-02-961-0515
Department of Management, Kyung Hee University, Seoul, Korea

HR information, would be more profound. Second, the nature of empirical work necessitates consolidation of data from multiple sources. For instance, productivity/efficiency analysis requires a measure for production output and inputs, which comes conventionally from finance, accounting or marketing data. Seamless data consolidation, though not impossible, is a daunting task and often goes beyond an individual researcher's capacity. Last but not least, tracking a firm's HR information during a sufficiently long time span is very costly. Despite the importance of panel HR data, its construction is time and effort consuming. In addition, consistency in many aspects of data collection, unless properly maintained over time, raises data quality issues.

Human Capital Corporate Panel (HCCP) data that has been collected and maintained by Korean Research Institute for Vocational Education and Training (KRIVET) opens a new avenue for empirical HR research. The aforementioned challenges are resolved, or at minimum alleviated, in the HCCP data. This unique panel data set contains very detailed HR information about major Korean firms through a total of seven survey waves up to 2018. The fact that the surveys have been conducted by KRIVET, a government-funded independent research institute, assures the data quality issues substantially. KRIVET's endeavor to well organize the data structure makes it easier to consolidate the HCCP data with other secondary data sources.

Given the availability of the HCCP data, the main objective of this study is to conduct an empirical research that sufficiently utilizes the valuable HR information therein. Specifically, this study proposes a flexible model that accommodates both cross-sectional and time-series variation in the HCCP data and, consequently, yields outcomes rich enough to deepen our understanding of the role various HR factors play in influencing organizational performance. To this end, we narrow our focus down to the context of productivity/efficiency analysis. Using a stochastic frontier model as a platform model, we postulate that the level of firm efficiency is moderated by various HR factors and estimate such a model via the use of a popular Bayesian inference method - Markov Chain Monte Carlo (MCMC) sampling.

The remainder of the paper is organized as follows. In the next section, we begin with specifying the Bayesian stochastic frontier model. We then present complete details of its estimation procedure. In section 3, we describe the data used in the empirical application of this study. In so doing, we provide a list of the variables chosen and their descriptive statistics. In section 4, we report our estimation results and discuss the empirical findings and their managerial implications. Lastly, we offer concluding remarks.

2. Model Specification and Estimation

This section lays out our modeling details. We begin with illustrating the model specification for our stochastic frontier analysis (SFA), with an emphasis on how to model the production frontier function and inefficiency, two essential ingredients of SFA. We then brief the hierarchical structure embedded via across-firm heterogeneity distribution of both constructs. Lastly, we forward the details of our MCMC sampling procedure that facilitates Bayesian estimation of the resulting model.

2.1 Production Function Specification

Since the seminal papers of Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) [7,8], SFA has been commonly used in the empirical study of firm productivity and efficiency (see Bauer (1990) for a survey of SFA [9]). The SFA originally for cross-sectional data has been extended later for panel data by Battese and Coelli (1992, 1995) and Bayesian inference for SFA has been introduced by van den Broeck, Koop, Osiewalski and Steel (1994) and Koop, Osiewalski and Steel (1994, 1997) [10-14]. This study draws upon the literature of Bayesian SFA for panel data.

The ideas underlying SFA center upon an economic theory of production where a maximum attainable output of firm i $(i = 1, \cdots, I)$ at time t $(t = 1, \cdots, T_i)$, $\bar{Y}_t$, depends on a combination of K inputs, $X_t = \{X_{t,1}, \cdots, X_{t,K}\}$. Each firm has access to its own best-practice technology for transforming inputs into output.

This technology is assumed to follow a Cobb-Douglas form with a vector of unknown parameters, $\beta_i = \{\beta_{i,1}, \cdots, \beta_{i,K}\}$, as given by

$$\bar{Y}_t = \left(\prod_{k=1,\cdots,K} (X_{t,k})^{\beta_{i,k}}\right) exp\ (\varepsilon_t), (1)$$

Where ε_t is a random error drawn from a Normal distribution with mean 0 and variance σ_ε^2. Inclusion of the random error ε_t is intended to capture measurement or specification error, making the production frontier stochastic. The actual output Y_t then is a function of the frontier output $\bar{Y}_t$ and firm-specific inefficiency δ_i as follows.

$$Y_t = \bar{Y}_t\ exp\ (-\delta_i), (2)$$

Where δ_i is positive (i.e., $\delta_i > 0$). This positivity restriction implies $0 < exp\ (-\delta_i) < 1$ and therefore $Y_t < \bar{Y}_t$. Taking natural logarithms for both sides of equation (2) after plugging equations (1) into (2) leads to

$$\ln(Y_t) = \sum_{k=1,\cdots,K} \beta_{i,k} \ln(X_{t,k}) + \varepsilon_t - \delta_i. (3)$$

The key empirical task is to estimate, with the panel data of output and inputs $\{Y_t, X_t\}_{\forall\, i,t}$, the model parameters of three kinds: (i) firm-specific frontier parameters β_i whose k^{th} element is interpreted as elasticity of frontier output with respect to the k^{th} input (since $\beta_{i,k} = \frac{\partial \ln(Y_t)}{\partial \ln(X_{t,k})}$), (ii) firm-specific inefficiency δ_i, and (iii) stochastic error variance σ_ε^2.

2.2 Hierarchical Structure

Completion of our model setup requires specifying a form of across-firm heterogeneity for two firm-specific parameters, β_i and δ_i. To this end, we keep up with a standard practice in the hierarchical Bayesian modeling literature (see Allenby and Rossi (2006) for reference [15]). We start with assuming β_i to be drawn from the following multivariate Normal distribution:

$$\beta_i \sim MVN(Z_i^\beta \gamma_\beta, V_\beta), (4)$$

where Z_i^β is a vector containing an intercept and firm i's L characteristic variables; γ_β is a $(L+1) \times K$ matrix in which its (*l,k*) element measures contribution of the l^{th} characteristic to the average elasticity of the k^{th} input's production frontier; and V_β is a $K \times K$ unrestricted covariance matrix of β_i. The linkage of Z_i^β to β_i through γ_β allows us to better understand who are more productive with respect to the level of a certain input and, hence, profile firms in terms of their productivity.

We now turn to imposing a hierarchical structure to δ_i. To reserve its positivity constraint, we assume δ_i to be drawn from a log Normal distribution whose mean is a function of firm i's characteristic vector as follows:

$$\ln(\delta_i) \sim N(Z_i^\delta \gamma_\delta, \sigma_\delta^2), (5)$$

where Z_i^δ is a vector containing an intercept and firm i's M characteristic variables (different from those contained in Z_i^β); γ_δ is a response vector of the mean of δ_i to these firm characteristics Z_i^δ; and σ_δ^2 is the variance of δ_i across firms. The linkage of Z_i^δ to δ_i through γ_δ, similar to the previous one of Z_i^β to β_i through γ_β, enables us to shed light on how various firm characteristics affect individual firm's production inefficiency.

Lastly, we make distributional assumptions for the stochastic error variance σ_ε^2 and the second-stage parameters such as γ_β, V_β, γ_δ, and σ_δ^2. Following the standard practice, we assume the conjugate priors for all of them, as given by: $\sigma_\varepsilon^2 \sim InvG\ (\vartheta_\varepsilon, \varphi_\varepsilon)$, $\gamma_\beta \sim MVN(m_\beta, P_\beta)$, $V_\beta \sim InvW\ (n_\beta, Q_\beta)$, $\gamma_\delta \sim MVN(m_\delta, P_\delta)$, and $\sigma_\delta^2 \sim InvG\ (\vartheta_\delta, \varphi_\delta)$, where $InvW$ and $InvG$ stand for inverted Wishart and Gamma distributions, respectively. The hyperparameters, $\{\vartheta_\varepsilon, \varphi_\varepsilon, m_\beta, P_\beta, n_\beta, Q_\beta, m_\delta, P_\delta, \vartheta_\delta, \varphi_\delta\}$, are appropriately chosen to make these priors sufficiently diffuse. Hence, they are not subject to estimation.

2.3 MCMC Estimation

Bayesian inference centers on posterior distribution that contains all relevant information about the model parameters. Analytical derivation of posterior distribution, however, is nontrivial or impossible in many cases, since the target posterior distribution is highly multi-dimensional and, often, of an intractable form. A major breakthrough in Bayesian inference is introduction of MCMC simulation [16]. This numerical method substitutes sequential samplings from a chain of full conditionals for a direct sampling from the target posterior. For the estimation of our SFA, we use an MCMC simulatior that consists of a series of Gibbs samplers and a Metropolis-Hastings (M-H) sampler (see Gelfand (2000) and Chib and Greenberg (1995) for further details of Gibbs and M-H samplings, respectively [17,18]). With that said, all the full conditionals but one, from which an updated value of the model parameters is drawn, are of a known form and, therefore, simulation is fairly straightforward to conduct via Gibbs sampling. For an exceptional case (i.e., sampling of λ in Step 4 below), the corresponding full conditional distribution is easy to evaluate up to its proportionality, facilitating the use of M-H sampling. Specifically, our MCMC sampler cycles through the following 7 steps after a proper initialization of the model parameters $\{\beta_i, \gamma_\beta, V_\beta, \delta_i, \gamma_\delta, \sigma_\delta^2, \sigma_\varepsilon^2\}$.

Step 1: Update β_i, for each firm $i\ (i = 1, \cdots, I)$, with a draw from $MVN(M, V)$ where

$$\mathrm{M} = \left[(V_\beta)^{-1} + \frac{\sum_{t=1}^{T_i}(X_{it})'(X_{it})}{\sigma_\varepsilon^2}\right]^{-1} \text{ and } \mathrm{V} = \mathrm{M}\left[(V_\beta)^{-1} Z_i^\beta \gamma_\beta + \frac{\sum_{t=1}^{T_i}(X_{it})'(Y_{it} + \delta_i)}{\sigma_\varepsilon^2}\right].$$

Step 2: First vectorize γ_β and denote its vectorized version by $vec\ (\gamma_\beta)$. Then update $vec\ (\gamma_\beta)$ with a draw from $MVN(M, V)$ where $\mathrm{M} = \left[(P_\beta)^{-1} + (\sum_{i=1}^{I}(Z_i^\beta)'(Z_i^\beta)) \otimes V_\beta^{-1}\right]^{-1}$ and $\mathrm{V} = \mathrm{M}\left[(V_\beta)^{-1} m_\beta + (\sum_{i=1}^{I}((Z_i^\beta)' \otimes V_\beta^{-1}) vec\ (\gamma_\beta)\right]$.

Note that the $\otimes$ notation indicates a tensor product.

Step 3: Update V_β with a draw from $InvW\ (d, S)$ where $\mathrm{S} = Q_\beta + \sum_{i=1}^{I}(\beta_i - Z_i^\beta \gamma_\beta)'(\beta_i - Z_i^\beta \gamma_\beta)$ and $d = n_\beta + I$.

Step 4: Update δ_i, for each firm $i\ (i = 1, \cdots, I)$, using an M-H algorithm. The full conditional distribution of δ_i is proportional to $\left[\sum_{i=1}^{I}\sum_{t=1}^{T_i}\exp(-\frac{(Y_{it} - X_{it}\beta_i + \delta_i)^2}{2\sigma_\varepsilon^2})\right]\exp(-\frac{\left(\mathrm{h}\ (\delta_i) - Z_i^\delta \gamma_\delta\right)^2}{2\sigma_\delta^2})$.

We generate a proposal value δ_i^{new} using a symmetric random walk M-H process and then replace the old value δ_i^{old} probabilistically. The acceptance probability of δ_i^{new} is $\min(1, \frac{[\delta_i^{new}|rest\ ]}{[\delta_i^{old}|rest\ ]})$ where $[\delta_i^{new}|rest\ ]$ and $[\delta_i^{old}\ |rest\ ]$ denote the value of the aforementioned full conditional distribution evaluated at δ_i^{new} and δ_i^{old}, respectively.

Step 5: Update γ_δ with a draw from $\mathrm{MV}N(M, V)$ where $\mathrm{M} = \left[(P_\delta)^{-1} + \frac{\sum_{i=1}^{I}(Z_i^\delta)'(Z_i^\delta)}{\sigma_\delta^2}\right]^{-1}$ and $\mathrm{V} = \mathrm{M}\left[(P_\delta)^{-1} m_\beta + \frac{\sum_{i=1}^{I}(Z_i^\delta)\,\mathrm{h}\ (\delta_i)}{\sigma_\delta^2}\right]$.

Step 6: Update σ_δ^2 with a draw from $InvG\ (w,s)$ where $w = \frac{\vartheta_\delta \varphi_\delta + \sum_{i=1}^{I} (\mathrm{h}\ (\delta_i) - Z_i^\delta \gamma_\delta)^2}{s}$ and $\mathrm{s} = \varphi_\delta + I$.

Step 7: Update σ_ε^2 with a draw from $InvG\ (w,s)$ where $w = \frac{\vartheta_\varepsilon \varphi_\varepsilon + \sum_{i=1}^{I} \sum_{t=1}^{T_i} (Y_{it} - X_{it}\ \beta_i + \delta_i)^2}{s}$ and $\mathrm{s} = \varphi_\varepsilon + \sum_{i=1}^{I} T_i$.

We run the above sampling cycle a total of 75,000 iterations. The statistical inference of this study is based on 50,000 MCMC draws of $\{\beta_i, \gamma_\beta, V_\beta, \delta_i, \gamma_\delta, \sigma_\delta^2, \sigma_\varepsilon^2\}$ after 25,000 burn-in iterations.

3. Data Description

We have thus far laid out the model specification of our SFA in the general production context and its MCMC estimation procedure. In this section we start with providing a general description of the firm-level panel data used in this study. We then forward a list of variables selected for implementation of our SFA.

3.1 HCCP Balanced Panel Data

The firm-level panel data used in this study is Human Capital Corporate Panel (HCCP) data that has been collected by KRIVET. Since 2005, KRIVET has conducted surveys every other year, resulting a total of seven survey waves. The data set contains various firm-level details regarding Human Resource Development/Management (HRD/HRM) on top of general management information. In addition, KRIVET constructs a secondary data set that consists of various types of financial/accounting information for the firms surveyed. This secondary data set spans 18 years from 2000 to 2017 and its entries are recorded on an annual basis.

For our empirical analysis, we use both the primary HR survey and secondary finance/accounting data. Regarding the former survey data, KRIVET further constructs a balanced panel data set that contains only the firms and survey questions commonly included from the third wave and onward (the recent five waves from 2009 to 2017). This balanced nature of data is essential to implementing panel data analysis of any form, since proper modeling treatment of various "drop-outs" in panel survey data is nontrivial. We therefore use the balanced panel survey data and supplement it with financial/accounting information from the secondary data. Mismatches between the two data sets are resolved in favor of the survey data. Having said that, we extract financial/accounting information not only for the firms surveyed but for the years when surveys were conducted. This results in a total of 317 firms, each of which has five data points. Out of these 317 firms, we further exclude those who have blank observations in the key variables (to be detailed below). Finally, we include 288 firms into our SFA.

3.2 Description of Variables

The data required to estimate our stochastic frontier model are denoted previously by $\{Y_{it}, X_{it}\}_{\forall i,t}$ and $\left\{Z_i^\beta, Z_i^\delta\right\}_{\forall i}$. The former set corresponds to the output and input variables, while the latter represents firm-specific characteristics. We here forward a list of variables constituting both data sets.

Both the output and input variables come from the secondary finance/accounting data. In accordance with the conventional productivity analyses, we select *Sales* for the output variable and postulate that *Labor* and *Capital* are two major inputs. We then operationalize the two input variables by choosing a reasonable proxy for each of them as follows: (i) a total amount of labor costs for *Labor* and (ii) a total amount of assets for *Capital*. The main equation in (iii) is now expressed as

$$\ln(Sales_{t}) = \beta_{i,L}\ln(Labor_{t}) + \beta_{i,C}\ln(Capital_{t}) + \varepsilon_{t} - \delta_i. \quad (6)$$

Table 1 provides the statistics such as mean, standard deviation, and median for the output and input variables. These statistics are computed not only across physicians but over time. The statistics in the top half are computed across physicians after first being averaged over time. The mean therein depicts an "average" firm's typical production performance, which is characterized by a production of 64.41 billion sales with the use of 3.10 billion labor and 265.45 billion capital. Most notable are substantially large standard deviations. They are 3 to 6 times larger than their corresponding means, indicating a significant amount of across-firm heterogeneity in both the output and input variables. The fact that the medians are far less than the means suggests the distributions of these variables across firms are right-skewed.

On the other hand, the statistics in the bottom half, computed over time after first being averaged across physicians, show a different facet of data. Although the change of computation order leaves the means unchanged, the standard deviations are fairly smaller than the previous ones, and the medians now are very close to the means. These findings altogether signify that most variation in data resides in the cross-sectional dimension, not in the time-series dimension. Consequently derived is a strong impetus for a disaggregate analysis, like our SFA, that fully accounts for the firm-level heterogeneity through both observed characteristics and unobserved random components.

Table 1. Descriptive statistics for output and input variables (Unit: Korean billion won)

Classification	Statistics	Sales	Labor	Capital
Across firms	Mean	64.41	3.10	265.45
	SD	202.24	15.39	2,048.71
	Median	16.04	0.59	18.68
Over time	Mean	64.41	3.10	265.45
	SD	3.13	0.33	41.04
	Median	63.59	3.03	260.00

The firm-specific characteristics utilized in our SFA can be classified into two categories. The first category pertains to particular types of the firms in the sample. We include a total of six variables in this vein, all of which are categorical variables: (i) *Industry* (whether a firm belongs to manufacturing, financial service, or non-financial service industry), (ii) *Size* (whether the number of employees is less than 300, between 300 and 999, between 1,000 and 1,999, or more than 2,000), (iii) *Age* (whether the age of a firm is less than 20 years, between 20 and 39 years, or more than 40 years), (iv) *Foreign Investment* (whether foreigners invest a firm) and (v) *Governance* (whether a firm is run fully by an owner, fully by a professional manager, or neither).

Figure 1 presents the shares of these categorical variables. The most typical firm in our sample is a manufacturing company (75.35%) with less than 300 employees (45.83%), 20 to 39 years old (43.40%), without foreign investment (66.32%), and run fully by an owner (42.01%). This case serves as a baseline in the course of estimation so that its impact on the production frontier is captured by an intercept in β_i, The other cases are represented by a series of dummies in β_i, each of which measures their incremental or decremental impact relative to the baseline case.

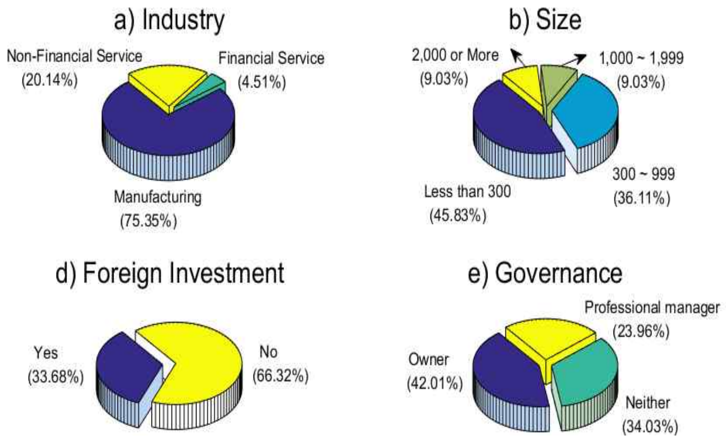

Figure 1. Shares of firm type variables

The second category is related to various HR features of the firms in the sample, which are further classified into three sub-categories: HR Composition (HRC), HR Development (HRD), and HR Management (HRM). For the HRC sub-category, we select three ratio variables: (i) *Full-time* (% of full-time employees), (ii) *Senior* (% of employees who are older than 40), and (iii) *Graduate Education* (% of employees with a master/doctoral degree). Included into the second HRD sub-category are two dummy and one continuous 7 variables: (i) *HRD Department* (whether a firm has a HRD specialized department), (ii) *HRD Planning* (whether a firm sets up HRD planning), and (iii) *HRD Expenses* (a natural logarithm of HRD expenses). Similar to the HRD sub-category, the HRM sub-category embraces two dummy and one continuous variables: (i) *HRM Department* (whether a firm has a HRM specialized department), (ii) *HRM Planning* (whether a firm sets up HRM planning), and (iii) *Employee Benefits* (a natural logarithm of a total amount of employee benefits).

Figure 2 provides general descriptions of these HR factors included in our analysis. A typical HR composition is that 93.68% of their employees are full-time, 39.19% are seniors (i.e., 40 or older), and 5.16% are those who have graduate education (i.e., master/doctoral degree holders). Out of the 288 firms sampled, exactly half (50.00%) have HRD department and 66.32% have HRM department. Cross-tabulation reveals that 132 firms (45.83%) have both, 85 firms (29.51%) have none, and the remaining 71 (24.65%) have either one. Regarding the HR planning, 79.51% and 79.17% of the firms set up development and management planning, respectively. Cross-tabulation discloses that 200 firms (69.44%) do both, 31 firms (10.76%) do none, and the remaining 57 (19.79%) do either one. The firms, on average, spend about 0.5 and 7.4 billion wons for HR development and employee benefits, respectively.

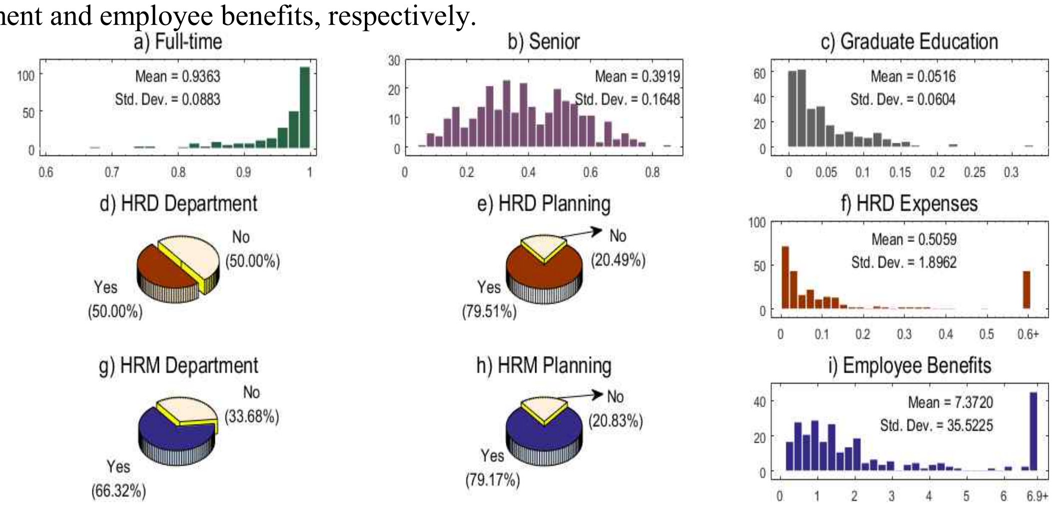

Figure 2. Description of HR factors

4. Results and Discussions

The model estimated in this study is a stochastic frontier model with a two-stage hierarchy. The first stage parameters are all individual firm specific (β_i and δ_i for $i = 1, \cdots, I$) except for the stochastic error variance (σ_ε^2), while the second stage parameters (γ_β, V_β, γ_δ and σ_δ^2) provide an aggregate view of the first stage ones through their mean and covariance. Among them, β_i, δ_i, γ_β, and γ_δ are parameters of primary interest. In this section we report our estimation results and discuss the implied empirical findings.

4.1 First-Stage Parameter Estimates

The primary merit of our SFA is that the production frontier and inefficiency parameters are estimated at the firm level. In our empirical application, we estimate, for each firm, (i) labor elasticity of sales ($\beta_{i,L}$), (ii) capital elasticity of sales ($\beta_{i,C}$), and (iii) inefficiency measured in log sales scale (δ_i).

Figure 3 presents their pairwise scatterplots along with marginal histograms. Regarding the two elasticity measures, their means across firms are 0.514 and 0.966, respectively. This indicates that, for an "average" firm, the percent change of production frontier is affected by that of capital approximately twice more than by that of labor. The standard deviations (0.448 and 0.382) show that a substantial amount of heterogeneity resides in these elasticity measures and labor elasticity is slightly more heterogeneous than capital one. The correlation between labor and capital elasticities is very strongly negative (-0.995), indicating that the two inputs are substitutes, rather than complements, to achieve the maximum production.

On the other hand, the mean and standard deviation of the inefficiency parameter are 0.487 and 0.132, respectively. Since the value of this parameter is measured in a log scale, we convert it into the original ratio scale by exponentiating its negative (i.e., $exp\ (-\delta)$). We then recalculate its mean, which turn out to be 0.615. This value times 100 now can be interpreted as the percentage of actual production relative to the frontier one. Thus, an "average" firm in the sample produces 61.47% of its attainable maximum, so the resulting inefficiency is 38.53%. This suggests there is a potentially large room for productivity improvement for most firms in the sample. A closer look at the distribution of inefficiency across firms reveals that the inefficiency for the best firm is only 4.88%, whereas that for the worst firm is as large as 73.82%.

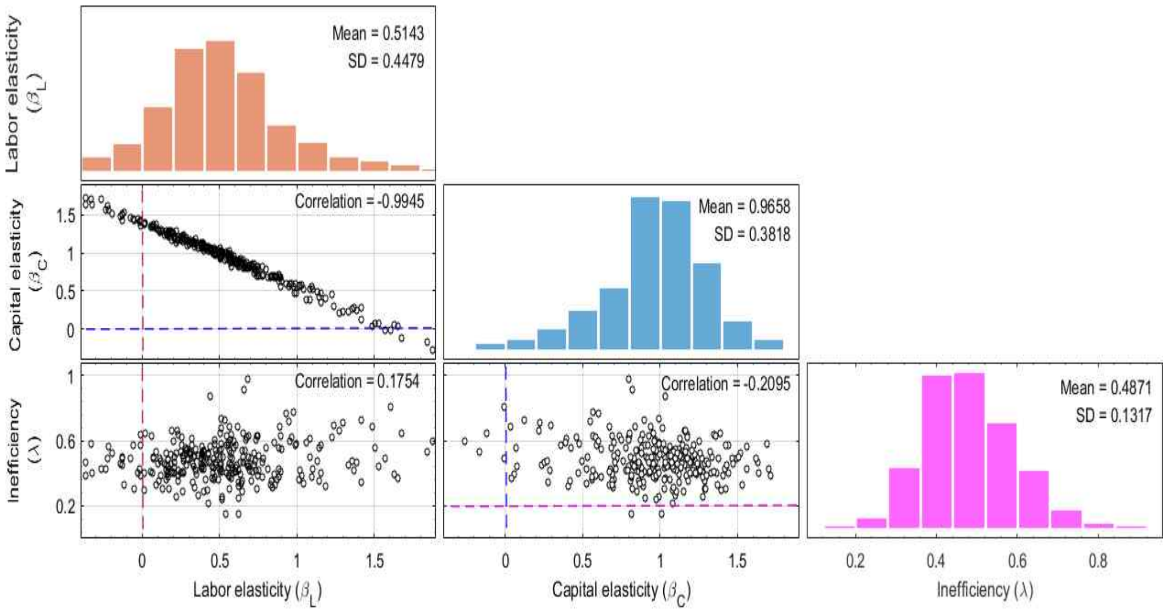

Figure 3. First-stage estimates

4.2 Second-Stage Parameter Estimates

Table 2 presents the estimates of the second-stage production frontier parameters. These second-stage estimates help determine if there are identifiable firm types that covary with the firm-level input elasticities. There are an intercept and ten dummies. The intercept represents the baseline firm type described previously, that is, a manufacturing company with less than 300 employees, 20 to 39 years old, without foreign investment, and run fully by an owner. Our estimation results reveal that input elasticities of this baseline firm are 0.38 for labor and 1.11 for capital. Relative to the "average" firm, the baseline firm has a smaller labor elasticity and a larger capital one.

The coefficient of each dummy variable measures its differential impact on the input elasticities relative to the baseline firm type. Four firm type variables are identified as being related significantly (at the 90% level) to both the labor and capital elasticities, namely, (i) *Industry* = Financial Service, (ii) *Size* = Between 300 and 999, (iii) *Age* = 40 or more, and (iv) *Governance* = Professional manager. Except for the size dummy, the remaining three dummies have an incremental impact on labor elasticity and a decremental impact on capital elasticity. Accordingly, a financial service company with less than 300 employees, 40 or more years old, with or without foreign investment, and run fully by a professional manager has the highest labor elasticity, which is equal to 1.580 (= 0.382 + 0.743 + 0.227 + 0.228) under an assumption that all insignificant coefficients are zero. Likewise, the firm type with the highest capital elasticity (1.293 = 1.112 + 0.181) is the baseline one but with 300 to 999 workers employed.

Table 2. Estimates of second-stage production frontier parameters

		Labor elasticity (β_L)		Capital elasticity (β_C)	
Variable	Parameter	Mean	90% HPD[a)]	Mean	90% HPD
	Intercept[b)]	0.382	[0.201, 0.562]	1.112	[0.962, 1.263]
Industry	Financial service	0.743	[0.320, 1.168]	-0.689	[-1.038, -0.338]
	Non-financial service	0.116	[-0.104, 0.335]	-0.109	[-0.292, 0.074]
Size	300-999	-0.236	[-0.432, -0.041]	0.181	[0.021, 0.343]
	1,000-1,999	-0.119	[-0.436, 0.194]	0.076	[-0.184, 0.339]
	2,000 and more	-0.113	[-0.458, 0.232]	0.042	[-0.248, 0.329]
Age	Less than 20	0.085	[-0.159, 0.326]	-0.078	[-0.278, 0.124]
	40 or more	0.227	[0.027, 0.427]	-0.208	[-0.373, -0.042]
Foreign Investment	Yes	-0.030	[-0.225, 0.165]	0.005	[-0.157, 0.165]
Governance	Neither	0.151	[-0.050, 0.354]	-0.135	[-0.302, 0.031]
	Professional manager	0.228	[0.080, 0.398]	-0.191	[-0.329, -0.066]

a) HPD stands for High Posterior Density interval.

b) Intercept captures the effects of the following baseline case: Industry = Manufacturing; Size = Less than 300; Age = Between 20 and 39; Foreign Investment = No; and Governance = Run fully by an owner.

Table 3 forwards the estimates of the second-stage inefficiency parameters. These estimates shed light on understanding what HR factors are associated with firm efficiency and, if any, quantifying the effects of the HR factors identified. Our results suggest that all the HR factors have a negative sign so that they may play a role in reducing inefficiency (or enhancing efficiency). Not all of them, however, have statistically significant effects. For the HR composition category, only significant is the percentage of full-time employees. For the other HRD and HRM categories, the department variables are only significant for HRM; the planning variables for both HRD and HRM are not significant; and HRD expenses and employee benefits turn out to be both significant. A related issue to the last finding is: given both expenses are efficiency enhancing, which would yield higher bang for the buck? The answer from our analysis is HRD expenses, simply because such a variable has a coefficient approximately twice larger than that of employee benefits.

Table 3. Estimates of second-stage inefficiency parameters

Variable	Parameters	Inefficiency (δ)	
		Mean	90% HPD[a)]
	Intercept	0.5618	[0.2696, 0.8383]
HR Composition	Full-time	-0.0548	[-0.0297, -0.0821]
	Senior	-0.0354	[-0.2311, 0.1371]
	Graduate education	-0.0264	[-0.5635, 0.4321]
HR Development	Department	-0.0151	[-0.0953, 0.0663]
	Planning	-0.0217	[-0.0872, 0.0474]
	Expenses	-0.0017	[-0.0025, -0.0011]
HR Management	Department	-0.0464	[-0.0806, -0.0638]
	Planning	-0.0036	[-0.0671, 0.0676]
	Employee benefits	-0.0009	[-0.0014, -0.0003]

a) HPD stands for High Posterior Density interval.

5. Conclusion

This study proposes a Bayesian stochastic frontier model that allows for estimation of firm-level productivity and efficiency measures and establishment of their linkage to a diverse set of firm characteristics including firm types and HR factors. The resulting rich set of estimates enable us not only to draw detailed micro-level inferences but to identify and quantify the effects of various firm characteristics on organizational performance. We provide complete estimation details of our proposal for the readers who are interested in applying it to other managerial contexts. We hope that this study will spur further interest in modeling of this kind.

Before concluding, we acknowledge that the presence of HCCP data is indispensable to conducting the empirical application of our proposal. The well-organized, multi-sourced, and detailed information contained therein provides an almost ideal playground to estimate our proposed model. At the same time, we also acknowledge that we use only a tiny piece of the HCCP data for this study. We leave a more thorough exploration of the HCCP data for a future research.

References

[1] C.D. Fisher, "Current and recurrent challenges in HRM," *Journal of Management*, Vol. 15, No. 2, pp. 157-180, June 1989. doi: 10.1177/014920638901500203

[2] P.M. Wright and G.C. McMahan, "Theoretical perspectives for strategic human resource management," *Journal of Management*, Vol. 18, No. 2, pp. 295-320, June 1992. doi: 10.1177/014920639201800205

[3] J.B. Arthur, "Effects of human resource systems on manufacturing performance and turnover," *Academy of Management Journal*, Vol. 37, No. 3, pp. 670-687, June 1994. doi: 10.5465/256705

[4] J. Barney, "Firm resources and sustained competitive advantage," *Journal of Management*, Vol. 17, No. 1, 99-120, March 1991. doi: 10.1177/014920639101700108

[5] A.A. Lado and M.C. Wilson, "Human resource systems and sustained competitive advantage: A competency-based perspective," *Academy of Management Review*, Vol. 19, No. 4, pp. 699-727, October 1994. doi: 10.5465/amr.1994.9412190216

[6] C.J. Collins and K.D. Clark, "Strategic human resource practices, top management team social networks, and firm performance: The role of human resource practices in creating organizational competitive advantage," *Academy of Management Journal*, Vol. 46, No. 6, pp. 740-751, December

2003. doi: 10.2307/30040665
[7] D. Aigner, C.A.K. Lovell, and P. Schmidt, "Formulation and estimation of stochastic frontier production function models," *Journal of Econometrics,* Vol. 6, No. 1, pp. 21-37, July 1977. doi: 10.1016/0304-4076(77)90052-5
[8] W. Meeusen, and J. van den Broeck, "Efficiency estimation from Cobb-Douglas production functions with composed errors," *International Economic Review,* Vol. 18, No. 2, pp. 435-444, June 1977. doi:
[9] P. Bauer, "Recent developments in the econometric estimation of frontiers," *Journal of Econometrics,* Vol. 46, No. 1/2, pp. 39-56, October-November 1990. doi: 10.1016/0304-4076(90)90046-V
[10] G.E. Battese and T.J. Coelli, "Frontier production functions, technical efficiency and panel data with application to paddy fanners in India," *Journal of Productivity Analysis,* Vol. 3, No. 1/2, pp. 153-169, June 1992. doi: 10.1007/BF00158774
[11] G.E. Battese and T.J. Coelli, "A model for technical inefficiency effects in a stochastic frontier production function for panel data," *Empirical Economics,* Vol. 20, No. 2, pp. 325-332, June 1995. doi: 10.1007/BF01205442
[12] J. van den Broeck, G. Koop, J. Osiewalski, and M.F.J. Steel, "Stochastic frontier models: a Bayesian perspective," *Journal of Econometrics,* Vol. 61, No. 2, pp. 273-303, March 1994. doi: 10.1016/0304-4076(94)90087-6
[13] G. Koop, J. Osiewalski, and M.F.J. Steel, "Bayesian efficiency analysis with a flexible form: the AIM cost function," *Journal of Business and Economic Statistics,* Vol. *12,* No. 3, pp. 339-346, July 1994. doi: 10.1080/07350015.1994.10524549
[14] G. Koop, J. Osiewalski, and M.F.J. Steel, "Bayesian efficiency analysis through individual effects: hospital cost frontiers," *Journal of Econometrics,* Vol. 76, No. 1/2, pp. 77-105, January-February 1997. doi: 10.1016/0304-4076(95)01783-6
[15] G.M. Allenby and P.E. Rossi, "Hierarchical Bayes model." In R. Grover, and M. Vriens (Eds.), *The handbook of marketing research: uses, misuses, and future advances* (pp. 418-440). Thousand Oaks, CA: Sage, 2006
[16] A.E. Gelfand and A.F.M. Smith, "Sampling-based approaches to calculating marginal densities," *Journal of the American Statistical Association,* Vol. 85, No. 410, pp. 398-409, June 1990. doi:
[17] A.E. Gelfand, "Gibbs sampling." *Journal of the American Statistical Association,* Vol. 95, No. 452, pp. 1300-1304, June 2000. doi: 10.1080/01621459.2000.10474335
[18] S. Chib and E. Greenberg, "Understanding the Metropolis-Hastings algorithm," *American Statistician,* Vol. 49, No. 4, pp. 327-335, November 1995. doi: 10.1080/00031305.1995.10476177