The Efficiency of Firms in India Manufacturing: A Stochastic Frontier Analysis

Qinghe Su[1]

Oklahoma State University

Abstract

This article investigates the efficiency of firms in India manufacturing sector, using a Stochastic Frontier Analysis (SFA) of a data set of 42,772 firms during the period 2014-2015 in India. We then apply a continuous treatment approach Generalized Propensity Score (GPS) methodology to investigate the export intensity and production growth. Our empirical results indicate that larger firms tend to be more efficient; rural-located firms and firms with ISO (International Organization for Standardization) certification are more efficient than firms without the certification. Furthermore, we find that firms' export share has a causal effect on production growth within a specific level of export share. The dose response reveals an inverted-U shape between the export share intensity and production growth.

JEL: O14 D24
Keywords: Stochastic Frontier Analysis, General Propensity Score, Technical Efficiency, Export, India Manufacturing

[1]Email: qinghe.su@okstate.edu

1 Introduction

Whether export-oriented enterprises are more efficient than non-exporters has attracted considerable attention. Fryges and Wagner (2008) survey 54 micro-econometric studies that included data for firms from 34 countries and were published between 1994 and 2006, showing that exporting firms are more productive than non-exporters within a sub-interval of the range of firms' export-sales ratios.[1]

In this paper, we address the question, "Are Indian manufacturing firms that exports are more efficient than firms that cater to domestic markets only?" For this question, we adopt a Stochastic Frontier Analysis to determine whether exporting firms are more efficient than non-exporters. Besides, we also use the generalized propensity score (GPS) that allows continuous treatment to examine whether the impact of exporting firms on efficiency depends on their export-intensity defined as the share of output exported.

Tabrizy and Trofimenko (2010) study the effect of exporting on India firms' productivity and get a similar conclusion that the self-selection of more productive firms into exporting explains the productivity differentials between exporters and non-exporters. However, some studies do find empirical support for learning by exporting. Baldwin and Gu (2003) explore the linkages between export-market participation and productivity performance in Canadian manufacturing plants and assert that export participation is associated with improved productivity, which is interpreted as evidence of learning-by-exporting. Kraay(1999) investigates Chinese enterprises using the data between 1988 and 1992 and concludes that these learning effects are most pronounced among established exporters. Another strand of the literature considers that there is no relation between exporting and efficiency or the relation is unclear. For example, Hung, Salomon, and Sowerby(2004) conclude that exporting activity itself does not promote efficiency in the United States, but due to the increase in manufacturing labor

[1] Bernard and Jensen (1999) and Bernard and Wagner (1997) give two alternative hypotheses for higher efficiency. One is self-selection, where higher productivity firms can afford the cost to serve the international market. The second hypothesis suggests the channel of learning-by-exporting. Globally engaged firms could be more efficient through the spillovers of knowledge in the foreign market.

productivity. Bernard and Jensen (1999) find that efficiency from exporting for American firms are ambiguous. Moreover, some literature attributes the efficiency of firms to trade liberalization generally (Pavcnik2002, Driffield and Kambhampati 2003, Fernandes 2003).

As stated above, most previous literature sheds light on the mechanism of exporter activity to the change of firm efficiency. However, how the export share influences the output growth rate is mostly ignored. Therefore, our paper contributes to the empirical evidence of firm efficiency in India, adding to the previous papers the export and other relevant factors affecting firm efficiency and how these factors are used and combined. Secondly, distinct from previous literature in India (Driffield and Kambhampati 2003, Mukim 2011), we use the continuous treatment effect to identify how export-intensity relate to firm efficiency.

The main findings of the paper are as follows. Regarding firm inefficiency, larger firm size can decrease the inefficiency and firms located in rural areas are associated with efficiency. Second, Firms with ISO certification have higher efficiency than firms without certification. Third, the aged firm increases the inefficiency significantly. However, export has no relation to firms' efficiency in this case based on the Stochastic Frontier Analysis. Then we adopt GPS to test the linkage between export intensity and production growth rate. The result suggests that there is an invert-U shape between the two. When the export share below 46 percent, the growth rate increases along with export intensity. However, it would be another story when the export share above the threshold. When the export share level is too large (above 90 percent), exporters have a lower production growth rate than the non-exporters.

Thus, we can show that exporting improves labor productivity only within a sub-interval of the export intensity. In contrast, it has no or negative effect within another sub-interval, which can partly explain why those studies that confine themselves to firms' export status do not find any impact on firms' export activities on productivity growth (Fryges and Wagner 2008). Investigating the effect of changes in export share on firm efficiency has been a significant issue for economics as export intensity has a direct influence on the output growth rate and the economic growth of the country. This study is also meaningful for policy implication.

The rest of the paper is organized as follows. Section 2 presents the empirical methodology. Section 3 describes the data set and shows the statistics. Section 4 shows the empirical results, and section 5 concludes.

2 Empirical Methodology

2.1 Stochastic Frontier Analysis

First, we use the Stochastic Frontier Analysis (SFA) to estimate a production frontier with inefficiency effects.[2] Technical inefficiency is estimated from the stochastic frontier and simultaneously explained by the firms' characteristics (Battese and Coelli 1995). Following Battese and Coelli (1995) the output for firm i, y_i, can be written as

$$y_i = f(x_i; \beta) \exp(v_i - u_i) \tag{1}$$

where x and β are vectors of inputs and parameters respectively. v_i is random error assumed to be be iid $N(0, \sigma_v^2)$. u_i is a non-negative random variable that captures the inefficiency and it assumed to be distributed independently and obtained by truncation at zero of $N(\mu_i, \sigma_u^2)$. The inefficiency term is

$$u_i = \sum_{k=1}^{m} \delta_k Z_{ik} + W_i \tag{2}$$

where Z_{ik} is a vector of characteristics that may have effects over firm efficiency, δ_k is a parameter to be estimated, and W_i is a random variable defined by the truncation of the normal distribution with zero mean and variance σ^2 (Diaz and Sanchez, 2008). Technical

[2]Several papers have used SFA to examine factors that effect firms' efficiencies. Few examples are Yasar and Paul (2009) for Turkey, Diaz and Sánchez(2008) for Spain, and Lemi and Wright (2020) for Africa.

efficiency (TE) of firm i can be expressed as[3]

$$TE_i = \frac{f(x_i;\beta)\exp(v_i - u_i)}{f(x_i;\beta)\exp(v_i)} = exp(-u_i) \tag{3}$$

TE equals one for an efficient firm and less than one for firms that fail to get the maximum feasible output. In our study, we use the Cobb-Douglas production function, adding a term of inefficiency.[4]

$$\ln Y_i = \beta_0 + \beta_1 \ln(L_i) + \beta_2 \ln(K_i) + \beta_3 \ln(O_i) + \sum_j \varphi_j I_j + v_i - u_i \tag{4}$$

where $\ln Y_i$ is the log value of total output of firm i; $\ln(L_i)$ is the log value of the total wage bill of firm i; $\ln(K_i)$ is the log value of capita; $\ln(O_i)$ is the log value of the other inputs used by firm i; and I_j are indicators for three digit industries.[5]

The pre-truncated mean carries information about inefficiency, and μ_i can be modeled as a function of the critical variables, which are perceived to influence the firms' inefficiency.

$$\begin{aligned}\mu_i =& \varphi_0 + \varphi_1\, export_i + \varphi_2\, (size1)_i + \varphi_3\, (size2)_i + \varphi_4\, (size3)_i + \varphi_5\, (size4)_i + \varphi_6\, rural_i \\ &+ \varphi_7\, quality_i + \varphi_8\, firmAge_i + \varphi_9\, public_i.\end{aligned} \tag{5}$$

where $export_i$ is an indicator for expoting firm (i.e. export share is greater than zero); $size1, size2, size3, size4$ are indicators if firm employ up to 20 workers, 21 to 50 workers, 51 to 100 workers, above 100 workers, respectively; $rural_i$ is an indicator if the firm is located

[3]Technical efficiency is defined as ratio of observed production over the maximum technically possible output (no inefficiency).

[4]In the SFA literature, Cobb-Douglas and translog production functions are commonly used functional forms. Many papers estimate both and choose based on the Hausman test. We have attempted to estimate a translog production function. However, the likelihood function will not converge in the case of the translog production function. Since our objective is to estimate TE, the choice of the production function is not critical as the distribution of TE estimated using different production functions are similar (Ahmad and Bravo-Ureta 1996).

[5]The other inputs include other expenses (work done by others on materials supplied by the industrial undertaking, operating expenses, and purchase value of goods sold in the same condition as purchased) and other inputs, which comprise the gross value of fuel materials.

in rural areas; $quality_i$ takes value of one if firm's product has ISO certification; $firmAge_i$ represents the number of years the firm has been in operation; and $public_i$ is an indicator for public ownership. Equations (4) and (5) are estimated jointly using by maximum likelihood using STATA's *frontier* command.

2.2 The Generalized Propensity Score Methodology

Since exporting firms differ in their export intensity (treatment levels or dose in our case), defined as the share of output exported, the binary export status does not present the complete story. Hence, we use GPS methodology developed by Imbens (2000) and Hirano and Imbens (2004) that allows continuous treatment to examine the relationship between export intensity and efficiency where technical efficiency comes from the SFA detailed in the last subsection. [6]

This bias-removing property of the GPS corresponds to that of the binary propensity score. Conditional on observable characteristics, the level of treatment can be considered random for units belonging to the same GPS strata. Therefore, adjusting for the GPS removes all biases associated with differences in the covariates (Guardabascio and Ventura, 2014).

Following Guardabascio and Ventura (2014) notations, let us define a set of potential outcomes $Y_i(t)$ for $t \in \tau$, where τ represents the continuous set of potential treatments levels defined over the interval $[0, 1]$, and $Y_i(t)$ is referred to as the unit-level dose–response function. For each unit i, we observe the level of the treatment delivered, T_i; and the outcome corresponding to the level of the treatment received, $Y_i = Y_i(T_i)$. We are interested in the average dose–response function $\varphi(t) = E[Y_i(t)]$. If we define $r(t, x) = f_{T|X}(d|x)$ as the conditional density function of the treatment given the covariates, then the GPS is defined as $R = r(T, X)$ (Hirano and Imbens 2004). Hirano and Imbens (2004) show that adjusting for the GPS eliminates any biases associated with differences in the pre-treatment variables.

[6] We also look at labor productivity as an alternative measure of efficiency.

Following Hirano and Imbens (2004), we first estimate the conditional distribution of the treatment variable given the covariates using fractional logit.[7] In the second stage of Hirano and Imbens' GPS methodology, the conditional expectation of outcome Y (technical efficiency in our case) is modeled as a function of the treatment T and the (estimated) generalized propensity score$\hat{R}$. We use a quadratic approximation for the conditional expectation of Y following Hirano and Imbens (2004) and Fryges and Wagner (2008).

$$E[Y_i|T_i, \hat{R}_i] = \alpha_0 + \alpha_1 T_i + \alpha_2 T_i^2 + \alpha_3 \hat{R}_i + \alpha_4 \hat{R}_i^2 + \alpha_5 T_i \hat{R}_i \tag{6}$$

The $\hat{\alpha}$s are estimated using OLS and used to estimate the average expected outcome at the treatment level t. The formula is as follows:

$$\hat{E}[Y(t)] = \frac{1}{N} \sum_{i=1}^{N} (\hat{\alpha_0} + \hat{\alpha_1}\, t + \hat{\alpha_2}\, t^2 + \hat{\alpha_3}\, \hat{r}(t, X_i) + \hat{\alpha_4}\, \hat{r}(t, X_i)^2 + \hat{\alpha_5}\, \hat{r}(t, X_i)) \tag{7}$$

where N is the number of observations. Bootstrapping is used to get confidence intervals of the dose-response function. Instead, the dose-response function we estimate shows the average potential outcome at each dose of the treatment and how average responses vary along with the interval $\tau = [0, 1]$. The paper uses STATA command glmdose (Guardabascio and Ventura, 2014) to estimate the dose-response function.

3 Data and Regression Variables

The data source is from the Annual Survey of Industries (ASI) 2014-2015 conducted by the Central Statistics Office in India. The survey covers all the registered unites whose employment is ten or more workers with power and 20 or more workers without power. The ASI is the principal source of industrial statistics of the registered manufacturing sector

[7] Export intensity (i.e. our treatment) variable is highly skewed and has observations with zero value (non-exporting firm). Hence fractional logit is suitable in our case. Fractional logit is used by others also in similarly skewed treatment distribution (Fryges and Wagner, 2008; Wagner, 2003; and Wagner, 2001).

in India. It includes the information to assess and evaluate the changes in the growth composition and structure of the organized manufacturing sector. The geographical coverage is the entire country. The reference period for the collection of ASI 2014-2015 is the financial year covering from any day between April 1st, 2014 and March 31st, 2015. The actual survey period is from September 2015 to April 2016. The ASI frame is based on the lists of registered factories or units maintained by the Director of Factories in the state or other registration authorities regarding Beedi and Cigar establishments and Electricity undertakings. The sampling design has two parts: the central sector and the state sector. This survey regards the firms that could not initiate production or did not produce anything during the reference period as closed units.

We eliminate all the closed firms and only keep the perennial firms in manufacturing sectors. We exclude the firms whose output value is below zero. Our sample includes $42,772$ firms. Summary statistics of the data are presented in Table 1. The variables used for the estimation of the production frontier are the log value of total output. The input variables include labor, capital, and other input variables (X_i). Furthermore, industrial dummies ($NIC3_i$) are also included. In the following part, we will explain these variables in detail.

4 Results

A firm's performance can be measured by its technical and allocative efficiencies (Charoenrat and Harvie 2014). There are two ways to estimate technical efficiency: Data Envelopment Analysis (DEA) or Stochastic Frontier Analysis. However, SFA has two advantages: one is SFA is a parametric approach where the form of the production function is assumed to be known and estimated statistically; second, SFA can simultaneously estimate a stochastic production model and the technical inefficiency.

4.1 Stochastic Frontier and the Inefficiency Model

4.1.1 Second Stage

The coefficient of the stochastic frontier production function and technical inefficiency effects are estimated by the maximum likelihood methodology (Charoenrat and Harvie 2014). Our estimation result is shown in Table 2.

The value of the estimates in Table 2 allows us to explain the factors that affect the firms' efficiency. From stage one, the coefficient of all the input variables is positive. Moreover, if the other input increases one percent, the output will increase by 71.3 percent. If the average wage rate increases one percent, the output will increase by 22.5 percent. If the value of fixed assets increases one percent, the output will increase by 4.3 percent.

The second stage is the estimation of the inefficiency model. The coefficient of export is negative, which means export increase the efficiency of production. Nevertheless, it is not statistically significant in this case. There is plenty of literature on the positive relation between the exporting and technical efficiency (Granér, and Isaksson 2009; Kim 2003). Furthermore, most of the results show that the positive relation associated with firm efficiency. The reason can explain that export firms can make more innovation through exporting and learning process.

Moreover, exporters are more competitive compared with the non-exporters, therefore more productive. Nevertheless, in our case, there is no relation between the export and inefficiency, which may because the export firm needs to invest more, which increases firms' sometimes cost, which offsets the positive effect.

The influence of firm size on firm efficiency can be divided into two groups. For firms with less than 20 workers, the inefficiency is positively correlated with the size. The decrease in efficiency, maybe because if the firm size is too small, some resources may not be utilized thoroughly. With the increase of firm sizes, the sign of the coefficient becomes positive. Moreover, the absolute values of coefficients are increasingly associated with the firm size.

It may because larger firms can reduce the cost because of scale production. Our result is consistent with the conclusion of previous literature that larger firms are more efficient (Lemi and Wright 2020; Söderbom 2004; Lundvall and Battese 2000).

The coefficient of rural is negative and statistically significant, which means firms located in the rural area will decrease the inefficiency compared with urban firms. At least two reasons can explain this. One is the land rent, and another cost for firms is lower than the urban area. Another reason is probably because the development of rural market can bring benefits to firms located in rural areas.

The coefficient of quality is negative, which suggests that qualified firms are more efficient than non-qualified firms. In our case, we use whether the firm has ISO certification to measure the quality. ISO certified firms mean that the firm has a standardized management system, manufacturing process, or documentation procedure. ISO certification can help firms to improve the efficiency of the business.

The coefficient symbols of firm age are positive and statistically significant, which indicates that longer operating years increase firms' inefficiency. First of all, it may difficult for aged firms to upgrade with new technology, and the management system may be out of date, which leads to inefficiency. From this case, we can conclude that firm longevity and experience are not guaranteed to increase efficiency (Charoenrat and Harvie 2014).

Unlike the previous literature (Dilling-Hansen, Madsen and Smith 2003), the coefficient of variable public firms is positive, but it is not statistically significant. Based on Dilling-Hansen, Madsen and Smith (2003), the public-owned firm is more efficient than other companies, the reason may be that for such companies the commercial risk is limited to the share capital itself. Therefore, they can take higher risks in order to get higher returns.

From Table 2, we can notice that the variance of *gamma* is around 0.003, which lies between 0 and 1, suggests that technical inefficiency is stochastic. It is relevant to obtaining an adequate representation of the data (Diaz and SÃ¡nchez 2008).

4.2 Generalized Propensity Score Methodology

From the SFA analysis, we know that export has no impact on the inefficiency of firms in our case. Some literature gets the negative relation between export and inefficiency, which means that export increases the efficiency of the firms, and exporters are inclined to have more significant production. In order to investigate how production growth responds to the change of export, we adopt a continuous treatment approach: Generalized Propensity Score (GPS) methodology developed by Imbens (2000) and Hirano and Imbens (2004). Different from the traditional propensity score methodology derived by Rosenbaum and Rubin (1983), the GPS allows for continuous treatment.

When identifying the causal effect of a treatment variable T on an outcome variable Y, we need to control the differences in other covariates (pretreatment variables) between the treatment group and non-treatment (control) group (Fryges and Wagner 2008). If not, it will lead to biases associated with differences in the covariates. One of the advantages of GPS is that it can eliminate any biases associated with the differences in pretreatment variables. Based on the Hirano and Imbens (2004) there are two steps for bias removal.[8]

We use a three-step approach to implement GPS, which is suggested by Hirano and Imbens (2004). First, following Fryges and Wagner (2008), we estimate the conditional distribution of the export share intensity (treatment variable) using a fractional logit model developed by Papke and Wooldridge (1996). Second, we use a quadratic approximation for the conditional expectation of Y ($\ln Y$ in our case). Third, we use the regression coefficients from the second stage to estimate the average expected outcome at the treatment level z.

Figure 1 depicts the dose-response function. It shows an inverted U-shaped relation between productivity growth and firms export share. In this graph, we take the production growth rate of non-exporter (where the export share is zero) as the baseline. Therefore, when the export share is zero, the production growth rate is zero instead of 0.793. The maximum productivity growth rate is obtained at an export share of 46 percent. The expected value

[8]The two steps approach is shown in appendix.

of the production growth rate amounts to 2.2 percent compared with the non-exporters (approximately growth rate of 81.56 percent). The treatment effect reveals that, at an export share of 46 percent, the production growth rate is significantly larger than the export intensity of zero. In other words, without considering the firm-specific characteristics in the pre-treatment variables, switch a non-exporter to exporting 46 percent intensity will lead to a 2.2 percent increase in the production growth rate. From Figure 1, we can see that when the level of export share changes from zero to 46 percent, the production growth rate increases but in a diminishing way, as illustrated in Table 3. Therefore, we can conclude that a firm's export activities have a causal effect on its production growth rate when the export share is lower or equal to 46 percent. If the export intensity exceeds this threshold, a firm will exhibit a lower production growth rate. From column (5) to (10) in Table 3, the production response coefficients are negative and statistically significant, which means the expected value of production growth rate is decreasing.

Nevertheless, exporting will positively impact a firm's production growth rate if its export intensity is less than 90 percent. The difference between the expected production growth rate for all firms export intensity above 90 percent and the non-exporters is negative and statistically significant. Considering the tremendous export intensity, firms (export share over 46 percent) need to increase the cost compared mainly with small export share firms for coordination and management to enter a more international market. The increasing international expansion decreases the firm's production growth rate. When the export intensity continues to increase (over 90 percent), the increasing geographic distance and differences in culture and peculiarities of the individual foreign markets raise the costs of exporting and necessitate additional sales personnel dramatically. As a result, it will hurt the production growth rate compared with the non-exporters (Gomes and Ramaswamy 1999). From column (5) to (10) in Table 3, we can see that the production growth rate decreases faster as the export intensity is tremendous.

5 Conclusion

In this study, our objective is to empirically analyze the determinants of the technical efficiency in the Indian manufacturing firms and investigate the relation between export intensity and output growth.

Firm technical efficiency can be determined by the environment or the specific characteristics of firms. In this paper, we analyzed the inefficiency of 42,772 manufacturing firms in India by the SFA methodology. The main foundings are that larger firm size can decrease the inefficiency firms located in a rural area associated with efficiency. Firms with ISO certification have higher efficiency than firms without certification. However, the aged firm increases the inefficiency statistically significant. Furthermore, export has no relation to the efficiency of firms. This founding tells the owner of the firm should increase the size of the firms and select rural areas to lower the cost and increase the market share.

To explore the relation of export and output, we adopt the continuous treatment methodology: GPS with dose-response to analyze the export share intensity and output growth. We show that firms' export share has a causal effect on the production growth rate within some share intensity interval. In our case, there is an invert-U shape between the level of export share and production growth. When the export share is between 0 to 46 percent, the production growth rate is positive. When the export share is above 46 percent, the production will decrease if the export share increases. The growth rate is still positive compared with non-exporters. However, if the export share level is above 90 percent, the difference in production for exporters and non-exporters is negative. Our result indicates that export will contribute to production within a specific threshold. When the level of export share is too large, the cost for firms to connect with distant trade partners or increase the share in the foreign market will be very costly.

References

[1] Ahmad, M., & Bravo-Ureta, B. E. (1996). Technical efficiency measures for dairy farms using panel data: a comparison of alternative model specifications. Journal of Productivity Analysis, 7(4), 399-415.

[2] Baldwin, J. R., & Gu, W. (2003). Export-market participation and productivity performance in Canadian manufacturing. Canadian Journal of Economics/Revue canadienne d'économique, 36(3), 634-657.

[3] Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empirical economics, 20(2), 325-332.

[4] Bernard, A. B., & Jensen, J. B. (1999). Exceptional exporter performance: cause, effect, or both?. Journal of international economics, 47(1), 1-25.

[5] Bernard, A. B., & Wagner, J. (1997). Exports and success in German manufacturing. Weltwirtschaftliches Archiv, 133(1), 134-157.

[6] Charoenrat, T., & Harvie, C. (2014). The efficiency of SMEs in Thai manufacturing: A stochastic frontier analysis. Economic Modelling, 43, 372-393.

[7] Diaz, M. A., & Sánchez, R. (2008). Firm size and productivity in Spain: a stochastic frontier analysis. Small Business Economics, 30(3), 315-323.

[8] Dilling-Hansen, M., Madsen, E. S., & Smith, V. (2003). Efficiency, R&D and ownership–some empirical evidence. International Journal of Production Economics, 83(1), 85-94.

[9] Driffield, N. L., & Kambhampati, U. S. (2003). Trade liberalization and the efficiency of firms in Indian manufacturing. Review of Development Economics, 7(3), 419-430.

[10] Fernandes, A. (2003). Trade policy, trade volumes, and plant-level productivity in Colombian manufacturing industries. The World Bank.

[11] Fryges, H., & Wagner, J. (2008). Exports and productivity growth: First evidence from a continuous treatment approach. Review of World Economics, 144(4), 695-722.

[12] Gomes, L., & Ramaswamy, K. (1999). An empirical examination of the form of the relationship between multinationality and performance. Journal of international business studies, 30(1), 173-187.

[13] Granér, M., & Isaksson, A. (2009). Firm efficiency and the destination of exports: evidence from Kenyan plant-level data. The Developing Economies, 47(3), 279-306.

[14] Guardabascio, B., & Ventura, M. (2014). Estimating the dose–response function through a generalized linear model approach. The Stata Journal, 14(1), 141-158.

[15] Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. Applied Bayesian modeling and causal inference from incomplete-data perspectives, 226164, 73-84.

[16] Hung, J., Salomon, M., & Sowerby, S. (2004). International trade and US productivity. Research in International Business and Finance, 18(1), 1-25.

[17] Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. Biometrika, 87(3), 706-710.

[18] Kim, S. (2003). Identifying and estimating sources of technical inefficiency in Korean manufacturing industries. Contemporary Economic Policy, 21(1), 132-144.

[19] Kraay, A. (1999). Exports and economic performance: Evidence from a panel of Chinese enterprises. Revue d'Economie du Developpement, 1(2), 183-207.

[20] Lemi, A., & Wright, I. (2020). Exports, foreign ownership, and firm-level efficiency in Ethiopia and Kenya: an application of the stochastic frontier model. Empirical Economics, 58(2), 669-698.

[21] Lundvall, K., & Battese, G. E. (2000). Firm size, age and efficiency: evidence from Kenyan manufacturing firms. The journal of development studies, 36(3), 146-163.

[22] Mukim, M. (2011). Does exporting increase productivity? Evidence from India. Manuscript, London school of economics.

[23] Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. Journal of applied econometrics, 11(6), 619-632.

[24] Paul, C. J. M., & Yasar, M. (2009). Outsourcing, productivity, and input composition at the plant level. Canadian Journal of Economics/Revue canadienne d'économique, 42(2), 422-439.

[25] Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants. The Review of Economic Studies, 69(1), 245-276.

[26] Söderbom, M. (2004). Productivity, exports and firm dynamics in Kenya 1999-2002. CSAE paper. Oxford: Centre for the Study of African Economies, Department of Economics, University of Oxford

[27] Tabrizy, S. S., & Trofimenko, N. (2010). Scope for export-led growth in a large emerging economy: Is India learning by exporting? (No. 1633). Kiel working paper.

[28] Wagner, J. (2001). A note on the firm size–export relationship. Small business economics, 17(4), 229-237.

[29] Wagner, J. (2003). Unobserved firm heterogeneity and the size-exports nexus: Evidence from German panel data. Review of World Economics, 139(1), 161-172.

6 Graph and Tables

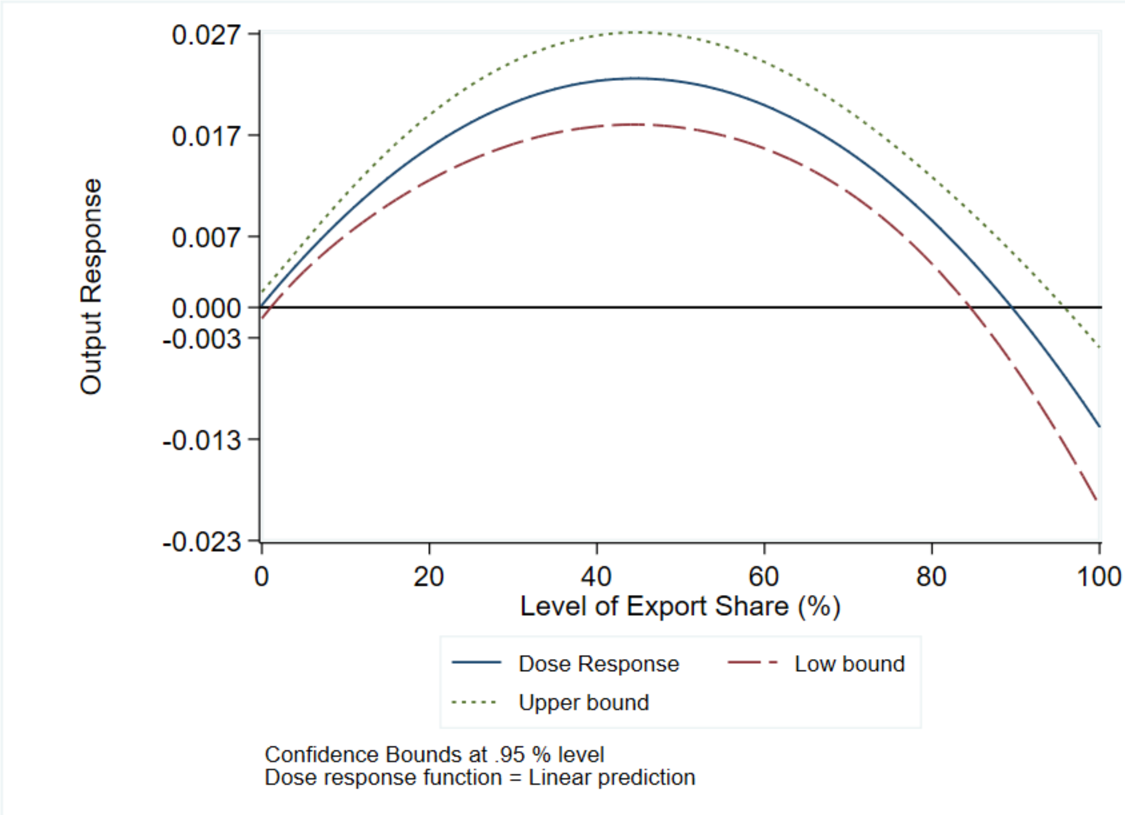

Table 1: Summary Statistics

Variables	Definition	Obs	Mean	Std.Dev	Min	Max
Characteristics of firm						
Rural	If the firm located in rural, rural=1	42772	0.42	0.49	0	1
Export	if the firm is export firm, export=1	42772	0.11	0.32	0	1
Share	percentage of products/by-products directly exported	42772	5.86	20.93	0	100
Quality	If the firm has the ISO certificate, quality=1	42772	0.12	0.32	0	1
Public	If the firm is public owned, public=1	42772	0.40	0.20	0	1
Size of the firm						
Firm size (dummy)	firm size by average person worked	42772	209.80	850.88	1	87427
Size 1	average person worked is from 10 to 20	42772	0.17	0.38	0	1
Size 2	average person worked is from 20 to 50	42772	0.17	0.37	0	1
Size 3	average person worked is from 50 to 100	42772	0.10	0.30	0	1
Size 4	average person worked is above 100	42772	0.37	0.48	0	1
Other important variables						
lnY	log value of total output	42772	18.31	2.26	10.60	28.73
lnL	log value of wage	42772	15.70	2.03	8.34	23.84
lnK	log value of capital (excluding building, land)	42772	15.89	2.66	0.69	26.99
lnO	log value of other input	42772	17.89	2.37	5.99	26.67

Table 2: Stochastic Frontier Analysis Cobb Douglas Production Function Estimates

VARIABLES		coefficient	Standard-error
Constant	β_0	1.619	19.532
lnL	β_1	0.225	0.001
lnK	β_2	0.043	0.001
lnO	β_3	0.713	0.001
Inefficiency model			
Constant	φ_0	0.249	19.532
export	φ_1	-0.002	0.003
Size1: From 10 to 20	φ_2	0.009	0.003
Size 2: From 21 to 50	φ_3	-0.023	0.003
Size 3: From 51 to 100	φ_4	-0.06	0.005
Size 4: Above 100	φ_5	-0.106	0.006
Rural	φ_6	-0.018	0.002
Quality	φ_7	-0.032	0.003
Firm_age	φ_8	0.001	0
public	φ_9	0.010	0.007
NIC3			
Variance parameter			
	σ^2	0.97	0
	γ	0.003	1.038
Observations		42,772	

Table 3: Export Intensity and Production Response

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Export Intensity	0-10%	11%-20%	21%-30%	31%-40%	41%-50%	51%-60%	61%-70%	71%-80%	81%-90%	91%-100%
Production Response	0.09	0.068	0.046	0.024	-0.001	-0.023	-0.046	-0.068	-0.09	-0.114
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.793	0.795	0.8	0.806	0.816	0.827	0.84	0.856	0.874	0.895
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.999	0.998	0.997	0.989	0.058	0.988	0.995	0.999	0.999	0.999

7 Appendix

First, we estimate the conditional expectation of the Y as a function of treatment level T and the GPS R, which is

$$\phi(t, r) = E[Y(t)|r(t, X) = r] = E[Y|T = t, R = r] \tag{8}$$

Second, use the dose-response function at a specific treatment level t,

$$\mu(t) = E[\phi(t, r(t, X))] \tag{9}$$