Production Efficiency versus Ownership: The Case of China

Alice Shiu and Valentin Zelenyuk

Kyiv School of Economics, Kyiv Economics Institute

1. August 2009

Production Efficiency versus Ownership: The Case of China

Alice Shiu* and Valentin Zelenyuk**

March 22, 2010

Abstract

In this study, we explore the pattern of efficiency among enterprises in China's 29 provinces across different ownership types in heavy and light industries and across different regions (coastal, central and western). We do so by performing a bootstrap-based analysis of group efficiencies (weighted and non-weighted), estimating and comparing densities of efficiency distributions, and conducting a bootstrapped truncated regression analysis. We find evidence of interesting differences in efficiency levels among various ownership groups, especially for foreign and local ownership, which have different patterns for light and heavy industries.

JEL classification: C13; C15; O11; O18

Keywords: Efficiency; Data envelopment analysis; Bootstrapping; Ownership; China

Acknowledgement: The authors would like to thank Aliaksandr Abrashkevich, Pavlo Demchuk, Natalya Dushkevych, Daniel Henderson, Christopher Parmeter, Ira Howoritz, Vincent Mok, Anatoly Piliawsky, Harry Wu and Xinpeng Xu for their valuable comments. The research support received from the Belgian Government (Belgian Science Policy) through Phase V (No. P5/24) of the "Interuniversity Attraction Pole" is gratefully acknowledged. We also would like to thank the Editor, Paul Wilson for his useful and insightful comments. The authors are entirely responsible for the views expressed and mistakes made in this work.

* School of Accounting and Finance, Room M805, Li Ka Shing Tower, The Hong Kong Polytechnic University, Hung Hom, Hong Kong; Tel. +852 27664016; fax: +852 23569550; afshiu@inet.polyu.edu.hk
** KEI (Kyiv Economics Institute) and the Kyiv School of Economics, vul., Yakira 13, office 321, Kyiv, 04119, Ukraine; vzelenyuk@eerc.kiev.ua

1. Introduction

Extraordinary changes have taken place in China over the past three decades since the adoption of the open door policy. These changes have been exemplified by those seen in China's industrial structure, especially in the radical moves toward non-state ownership. The corporatization of the state sector, the government's encouragement of merger and acquisition activity among state-owned enterprises (SOEs), and the dramatic development of the non-state sector with enormous foreign investments have dominated both the Chinese economy and political debate for the past decade.

The purpose of this paper is to tackle the timeworn political debate about which type of ownership is more efficient in the Chinese economy and whether it depends on the industry (light or heavy) and/or the region (central, western or coastal). While the literature includes many studies of productivity in China (see the citations below), none have focused on the relative efficiency of various ownership types for both light and heavy industry combined. This is the issue we attempt to address in our study. Our particular focus is on foreign versus local ownership. While there is little doubt that private ownership should outperform state ownership on average, the situation is not so clear for foreign versus local ownership and whether it depends on the type of industry.

To achieve our goal, we use the most recent census data constructed for Chinese enterprises of different ownership types in 1995. Our methodological approach exploits recent developments in the area of efficiency analysis and is implemented in two stages. The first stage involves the estimation of efficiency scores for individual observations (each province in each type of industry) using the data envelopment analysis (DEA) estimator. In the next stage, we analyze the individual efficiency scores obtained in the first stage using three different methods.

The first method is based on the analysis of densities of efficiency distributions for different ownership groups using a kernel density estimator and testing for their equalities using an adaptation of the Li (1996) test. The second method is based on the aggregation method of Färe and Zelenyuk (2003) and investigates group efficiency scores obtained as weighted averages, with the weights representing the economic importance of each observation. Statistical inferences for these group efficiency scores are made via bootstrapping techniques suggested by Simar and Zelenyuk (2007). The third method assumes more of a dependency structure and allows us to analyze the dependency of efficiency scores on hypothetical explanatory variables. Here, we use the truncated regression proposed by Simar and Wilson (2007) in which bootstrapping is used as a means of statistical inference to investigate how the conditional mean of efficiency scores is influenced by explanatory variables such as ownership and regional dummies, as well as by size. These methods yield interesting evidence of performance variations among ownership groups and regions. Remarkably, the pattern of performance for light industry is found to differ from that for heavy industry.

In common with the results of other studies, our results provide robust evidence confirming the expectation that non-state ownership is superior to state ownership in terms of the performance levels achieved. In addition, we confirm our prediction that foreign owned firms in heavy industry perform distinctly better than their counterparts with other ownership types. Somewhat surprisingly, foreign ownership in light industry appears to be associated with *lower* efficiency, on average, than the other non-state ownership types we consider. This unexpected result can nevertheless be explained by the theory of technology diffusion/adoption, which can be traced back at least as far as the studies of Gerschenkron (1962) and Nelson and Phelps (1966).

Among our other findings, we present evidence of agglomeration effects that are pronounced in light industry but are not particularly marked in heavy industry. Interestingly, we find no significant difference in average efficiency between light and heavy industries. Overall, apart from confirming a number of previous findings, our study sheds new light on the pattern of productivity in China that will be of interest to researchers and practitioners.

The remainder of this paper is organized as follows. Section 2 briefly discusses our methodology and Section 3 provides a brief discussion of the data. Section 4 reports the empirical results in detail and Section 5 concludes the paper.

2. Methodology

Estimation of Efficiency (Stage 1)

In the first stage of our analysis, we use the data envelopment analysis (DEA) estimator to obtain efficiency scores for each observation. This approach usually assumes that all *decision-making units* (DMUs) within a sample have *access to* the *same technology* for transforming a vector of N inputs, x, into a vector of M outputs, y.[1] We also assume that technology can be characterized by the *technology set*, T, as[2]

$$T = \{(x, y) \in \mathbb{R}_+^N \times \mathbb{R}_+^M : \ x \in \mathbb{R}_+^N \ can\, produce\ y \in \mathbb{R}_+^M\} \qquad (1)$$

Note that while our approach requires that all DMUs have *access* to the *same technology*, it also allows for any DMU to be either on or away from the *frontier* of such technology. The distance from each DMU in T to the frontier of T is called the

[1] The DEA was originally designed for firm-level analysis, but it has frequently been applied to more aggregated data; see, for example, Färe et al. (1994) and the more recent studies of Kumar and Russell (2002), Henderson and Russell (2005), and Henderson and Zelenyuk (2006).

[2] We assume that the standard regularity conditions of the neo-classical production theory hold (see Färe and Primont (1995) for details).

inefficiency of each DMU caused by endogenous or exogenous factors specific to that DMU. These endogenous factors could include internal economic incentives influenced by motivation systems, ownership structure, management quality, etc. Exogenous factors might include different demographic or geographic environments, regulatory policies, and so on. Our goal is to *estimate* such inefficiency and analyze its dependency on the hypothesized factors.

Technical efficiency for each DMU $j \in \{1, \dots, n\}$ is measured using the Farrell/Debreu-type output-oriented technical efficiency measure

$$TE^j \equiv TE\left(x^j, y^j\right) = \max_\theta\left\{\theta \ : \ \left(x^j, \theta y^j\right) \in T\right\}. \tag{2}$$

Obviously, the true *T* is unobserved, and so we replace it with its *DEA-estimate*, $\hat{T}$, obtained through the following activity analysis model

$$\hat{T} = \{(x, y) \in \mathbb{R}_+^N \times \mathbb{R}_+^M \ : \ \textstyle\sum_{k=1}^n z^k y_m^k \geq y_m\,, \ m = 1, \dots, M,$$
$$\textstyle\sum_{k=1}^n z^k x_i^k \leq x_i\,, \ i = 1, \dots, N, \ z^k \geq 0, \ k = 1, \dots, n\}, \tag{3}$$

where $\{z^k : k = 1, \dots, n\}$ are the intensity variables over which optimization (2) is made. Note that such $\hat{T}$ is the smallest convex free disposal cone (in (x, y)-space) that contains (or 'envelopes') the input-output data.[3] In our discussions, we focus on the constant returns to scale (CRS) model only for several reasons. First, the CRS model (2) has greater discrimination power, making it capable of identifying more inefficiency than non-CRS models. Some of the inefficiency identified under the CRS model will be due to the scale effect (i.e., where a DMU is too small or too large), which will be tested at the second stage by including a proxy for scale. Second, the CRS model compares all DMUs evenly to the same cone, whereas for the non-

[3] Alternatively, if we add $\sum_{k=1}^n z^k \leq 1$ or $\sum_{k=1}^n z^k = 1$ to (3), then we can model the non-increasing returns to scale (NIRS) or the variable returns to scale (VRS), respectively.

CRS DEA estimator, a large proportion of DMUs are often in or near the flat regions of the estimated technology and so obtain high or perfect efficiency scores while being quite inefficient from an economic perspective. Third, the CRS model is a natural choice when aggregate (country- or region-level) data are used.

We choose the Farrell efficiency measure over others for two reasons that make it the most popular in practice. This measure has been shown to satisfy a set of attractive mathematical properties that are desirable in an efficiency measure.[4] Moreover, this estimator is fairly easy in terms of computation and allows for straightforward interpretation.

Note that the true efficiency scores from the Farrell measure are bounded between unity and infinity, where unity represents a perfect (technical or technological) *efficiency* score of 100%. On the other hand, $\left(1/TE^{j}\right)$ would represent the *relative %-level* of the *efficiency* of the j^{th} DMU ($j \in \{1, \ldots, n\}$). By replacing T with $\hat{T}$ in (2), we obtain the DEA estimator of TE^{j} under the assumptions of CRS, additivity, and free disposability. Applying this estimator will give estimates of the true efficiency scores, $\left\{TE^{j}: j = 1, \ldots, n\right\}$, which we denote as $\left\{\widehat{TE}^{j}: j = 1, \ldots, n\right\}$. These estimated efficiency scores have the same range as the true efficiency scores and, as in many other extreme-value type estimates, are subject to small-sample bias, which nevertheless vanishes asymptotically as the estimates are consistent with their true counterparts.[5]

---

[4] These properties include various forms of *continuity*, (weak) monotonicity, commensurability, homogeneity, and (weak) indication for all technologies satisfying certain regularity conditions (see Russell (1990, 1997) for details).

[5] See Korostelev et al. (1995) and Park et al. (2010) for proof of consistency and rates of convergence of the DEA estimator under CRS, and other statistical properties and required assumptions. Also see Kneip et al. (1998, 2008) for related results on VRS.

Analysis of Efficiency Distributions (Method 1 of Stage 2)

The aim of the *second* stage of the analysis is to study the dependency of the efficiency scores obtained in the first stage on DMU-specific factors such as ownership structure, regional location, size, etc.

The starting point of our second stage analysis is to explore the efficiencies within and between groups that might theoretically represent different sub-populations in the population as a whole. For example, state-owned firms have different incentives to other firms which are likely to be reflected in the efficiency distribution of state-owned firms relative to other firms. In particular, we first analyze the *distributions* of efficiency within various groups. Here, we start with estimation and visualization of the densities of corresponding distributions using the *kernel density estimator*. For this we use the Gaussian kernel, Silverman (1986) reflection method (around unity), to take into account the bounded support of efficiency measure, and Sheather and Jones (1991) method for bandwidth selection. We then apply a version of the Li (1996) test (adapted to the DEA context by Simar and Zelenyuk (2006)) to test the equality of efficiency distributions between various groups of interest.

Analysis of Aggregate Efficiency Scores (Method 2 of Stage 2)

We proceed to analyze the various groups by testing the equality of group (aggregate) efficiencies, which is estimated using the weighted and non-weighted averages of the individual efficiency scores for each group. Because the weights used for averaging might be critical here, they must be chosen on the basis of some (more-or-less) objective criterion. We use the weights derived from economic optimization by Färe and Zelenyuk (2003) which were extended to the sub-group case by Simar and

Zelenyuk (2007). In summary, our (weighted) group efficiency score for group l (l = 1, …, L) is estimated as

$$\overline{\widehat{TE}}^{l} = \sum_{j=1}^{n_l} \widehat{TE}^{l,j} S^{l,j} \ , \qquad l = 1, \dots L. \tag{4}$$

where the weights are

$$S^{l,k} = py^{l,j} / p \sum_{j=1}^{n_l} y^{l,j} \ , \qquad j = 1, \dots n_l. \tag{5}$$

in which p is the vector of output prices. For convenience, we would present the *reciprocals* of the estimated group efficiency scores, i.e., $\left(\overline{\widehat{TE}}^{l}\right)^{-1}$, $l = 1, \dots, L$ (and the corresponding confidence intervals) to give them meaning in percentage terms.

To make statistical inferences based on these group efficiency scores, we use the bootstrap-based approach suggested by Simar and Zelenyuk (2007); readers are referred to the same study for further details of this method. In summary, the statistic used for testing the null hypothesis that the aggregate efficiencies for any two groups, e.g., A and Z, are equal (i.e., $H_0{:}\, \overline{TE}^{A} = \overline{TE}^{Z}$) is given by the relative difference (RD) statistic:

$$\widehat{RD}_{A,Z} = \overline{\widehat{TE}}^{A} \ / \ \overline{\widehat{TE}}^{Z} \tag{6}$$

The null hypothesis will be rejected (at certain level of confidence) in favor of $H_1{:}\, \overline{TE}^{A} > \overline{TE}^{Z}$ if $\widehat{RD}_{A,Z} > 1$ (or $H_2{:}\, \overline{TE}^{A} < \overline{TE}^{Z}$ if $\widehat{RD}_{A,Z} < 1$) and the bootstrap-estimated confidence interval of $\widehat{RD}_{A,Z}$ does not overlap with unity.

Regression Analysis of Determinants of Efficiency (Method 3 of Stage 2)

The last method used in our investigation involves the application of regression analysis to study the dependency between efficiency scores and some expected explanatory variables. Here, we assume and test the following specification

$$TE^j \approx a + Z_j\delta + \varepsilon_j, \qquad j = 1, \ldots, n, \tag{7}$$

where a is the constant term, ε_j is statistical noise, and Z_j is a (row) vector of observation-specific variables for DMU j that we expect to influence DMU efficiency score, TE^j, defined in (2), through the vector of parameters δ (common for all *j*) that we need to estimate.

For some time, a practice commonly adopted in the DEA literature was to estimate model (7) using the Tobit-estimator. However, Simar and Wilson (2007) illustrate that this approach would be incorrect here and instead propose an approach based on a bootstrapped *truncated regression*, showing that it performs satisfactorily in Monte Carlo experiments. We follow their approach (specifically, their "Algorithm 2") and instead of using the unobserved regressand in (7), TE^j, use its bias-corrected estimate, $\widehat{TE}^j_{bc}$, which is obtained using the heterogeneous parametric bootstrap they propose. Note that because both sides of (7) are bounded by unity, the distribution of ε_j is restricted by the condition$\varepsilon_j \geq 1 - a - Z_j\delta$. To simplify the estimation process, we follow Simar and Wilson (2007) by assuming that this distribution is a truncated normal distribution with a mean of zero, unknown variance, and a (left) truncation point determined by $\varepsilon_j \geq 1 - a - Z_j\delta$. Formally, our econometric model is given by

$$\widehat{TE}^j_{bc} \approx a + Z_j\delta + \varepsilon_j, \qquad j = 1, \ldots, n, \tag{8}$$

where

$$\varepsilon_j \sim N(0, \sigma_\varepsilon^2), \text{ such that } \qquad \varepsilon_j \geq 1 - a - Z_j\delta, \quad j = 1, \ldots, n. \tag{9}$$

We then use our data to estimate the model shown in (8)-(9) by maximizing the corresponding likelihood function with respect to$(\delta, \sigma_\varepsilon^2)$. To obtain the bootstrap confidence intervals for the estimates of parameters $(\delta, \sigma_\varepsilon^2)$, we use the *parametric* bootstrap for regression that incorporates information on the parametric structure (7) and the distributional assumption (9). For the sake of brevity, we refer readers to Simar and Wilson (2007) for the details of the estimation algorithm.

3. Data

The data used in this paper are drawn from the *Third National Industrial Census of the People's Republic of China* conducted by the State Statistical Bureau in 1995, which is the latest census for which statistics have been put together and published. The data provided in the census are the only industry-level data available that are categorized by type of ownership. Specifically, the census provides cross-sectional data for Chinese enterprises divided into four ownership types that are aggregated at the *province* level (29 provinces) for light and heavy industries in 1995. The four types of ownership are: *(i)* state-owned enterprises (SOEs); *(ii)* foreign-funded enterprises (FFEs); *(iii)* township-owned enterprises (TOEs); and *(iv)* collectively-owned enterprises (COEs). Given these data, we have 8 'representative' DMUs for each of the 29 provinces in China: SOEs, FFEs, TOEs, and COEs in the light and heavy industries, respectively.

A brief explanation of the industry sectors is warranted here. "Light industry" refers to the group of industries that produce consumer goods and hand tools. It consists of two categories distinguished from each other according to the materials used. The first category include industries that use farm products as materials, while the other category includes industries that use non-farm products as materials.[6] "Heavy industry" refers to industries that produce capital goods and provide materials and technical bases required by various sectors of the national economy.[7] The level of

[6] Some examples of the first category of light industries are food and beverage manufacturing, tobacco processing, and textiles and clothing, and some examples of the second category are the manufacturing of chemicals, synthetic fibers, chemical products, and glass products.

[7] Heavy industry consists of three branches distinguished according to the purpose of production or how the products are used. They include (i) the mining, quarrying and logging industry that involves the extraction of natural resources; (ii) the raw materials industry, which provides raw materials, fuel

competition among light industry firms is generally more severe than that among heavy industry participants because there are usually more firms in the former group. Also, because most light industry firms are non-SOEs, they face hard-budget constraints and are fully responsible for their profits and losses. On the other hand, because most heavy industry firms are SOEs which are larger and fewer in number, the level of competition between such firms is usually lower than it is among light industry firms.

To construct the constant returns to scale (CRS) output-oriented activity analysis model for the DEA estimator in the first stage, we use *three inputs* (i.e., total wage, the net value of fixed assets, and the value of intermediate inputs) and *one output* (the gross industrial output of each type of ownership in each province). Some descriptive statistics and a brief discussion of the data are provided in the Appendix. Further details can be found in two studies conducted by Shiu (2000, 2001).

4. Main Results

Analysis of Densities and Means for <u>Light Industry</u>

After obtaining the DEA estimates of efficiency scores, we use the kernel density estimator to approximate the distributions of the individual efficiency scores for the four ownership groups in each of the light and heavy industry sectors. Statistical tests for the equality of distributions suggested by Li (1996) (and adapted to the DEA context by Simar and Zelenyuk (2006)) are used to test for differences in distributions amongst the ownership groups. Figure 1 shows the (estimated) densities of the distributions of the estimated individual efficiency scores for each ownership group in

---

and power to various sectors of the economy; and (iii) the manufacturing industry, which processes raw materials.

the light industry sector. The estimated densities seem to be relatively divergent among groups. Interestingly, the only ownership group that has a density with an estimated mode of unity is the TOEs. Intuitively, this means that for TOEs, the highest frequency at which the level of efficiency is observed is where one would expect it to be for highly competitive firms: at the 100% level of efficiency. Other groups have estimated modes that are not at unity but are instead at some level of inefficiency, which we view as evidence of some degree of 'pathological' inefficiency. The SOE group has the most 'inefficient' mode (around 2, i.e., about 50% efficient), making it radically different from other groups and the least efficient group. Columns 2 and 3 of Table 1 present the results of tests for the equality of distributions between all dyads of ownership groups in the light industry sector; we reject equality for most of the comparisons at the 95% confidence level. The exceptions are the efficiency distributions of foreign-funded enterprises (FFEs) and township-owned enterprises (TOEs) versus collectively- owned enterprises (COEs). We also reject equality for foreign-funded enterprises versus collectively-owned enterprises at the 10% level (est. p-value of 0.06).

Given the evidence of different efficiency distributions among ownership groups, a further issue that arises is whether this divergence is due to differences in group aggregate efficiency scores and whether these differences are statistically significant. The upper part of Table 2 lists the weighted efficiency scores for each light industry ownership group in the 29 provinces in 1995. The aggregate efficiency scores are calculated using Färe-Zelenyuk weights, with bias corrected and confidence intervals estimated on the basis of the Simar-Zelenyuk (2007) group-wise-heterogeneous bootstrap-based approach. The second column indicates the ownership groups. The numbers in the third and fifth columns represent the *reciprocals* of the original DEA efficiency scores and of the bias-corrected efficiency scores, respectively. (Reciprocals are taken for convenience to show the percentage meaning of the efficiency scores.) The last two columns show the lower and upper bounds of the 95% confidence interval.

The results in the upper part of Table 2 indicate that SOE performance is different from non-SOE performance. A relatively large estimated bias in the aggregates of efficiency scores is found among all ownership groups, especially for SOEs (0.55 and 0.43). This indicates that in the light industry sector, the technical efficiency of SOEs varies widely across the provinces.

Furthermore, bootstrap-based tests of the equality of aggregate efficiencies are employed to test for pair-wise comparisons of the aggregate efficiencies of the various sub-groups (see the lower part of Table 2). The relative difference (RD) statistics computed for the DEA and bias-corrected aggregate efficiency scores are shown in the third and fifth columns, respectively. If the RD statistic for group A versus group Z is greater than 1 and the confidence interval does not overlap with 1, then the null hypothesis that the aggregate efficiencies of the two groups are equal is rejected in favor of the alternative hypothesis that the aggregate efficiency of group A is worse than that of group Z.[8] The RD statistics suggest that SOEs are operated in a significantly (at the 1% level) less efficient manner than all the other groups. This finding supports the results obtained in our distributional analysis and can be explained by the fact that SOEs are often ill-equipped to meet their business objectives as they tend to use out-of-date capital equipment and usually have no funding available to them for technological upgrades (for more discussion, see Groves et al., 1994; Weitzman and Xu, 1994; Zheng et al., 1998; Zhang and Zhang, 2001; Dong and Tang, 1995; Lin et al., 1998; Huang et al., 1999; Wu, 1996 and 1998).

---

[8] E.g., the RD-statistic for comparing the weighted average efficiency scores for groups 1 and 2 was estimated as $\overline{\widehat{TE}}^1 \ / \ \overline{\widehat{TE}}^2 = 1.30$, meaning that group 1 is <u>less</u> efficient than group 2, and this difference is significant, since 95% confidence interval is [1.23, 1.62], not overlapping with 1.

Regarding the performance of non-SOEs, it is interesting to find that in the light industry sector, FFEs perform significantly *less* efficiently than COEs and TOEs (at the 5% level for weighted averages and at the 10% level for non-weighted averages). One possible explanation for this result is that the network of bureaucratic restrictions adversely affecting the competitiveness of FFEs offset the benefits gained from the government's preferential policies for foreign investors. Examples include high-profile administrative intervention in the operation of FFEs, the levying of miscellaneous fees of an ambiguous nature, and the imposition of stringent policies. (For more discussion, see ACC, 1998; Melvin, 1998; Weldon and Vanhonacker, 1999; Transparency International, 2001). These issues could lead to higher transaction costs being incurred in FFE operations and thereby cancel out certain competitive advantages enjoyed by FFEs over local firms (see, for example, Yeung and Mok, 2002). Other reasons that may account for the lower level of efficiency in FFE operations include the large initial investment required and the steep learning curve for foreign investors (e.g., see Wei et al., 2002).

Analysis of Densities and Means for <u>Heavy Industry</u>

Figure 2 shows the (estimated) densities of individual efficiency distributions of the four ownership groups for the heavy industry sector. The densities appear to be more tightly grouped in the heavy industry sector than those observed for the light industry sector, other than in the case of FFEs, for which we see a clear difference in the density of efficiency relative to that of the other groups. The SOEs group again has less of its distributional mass close to unity, while the FFEs group has more of its distribution close to unity than the other groups. Columns 4 and 5 of Table 1 formally support these observations via tests for the equality of distributions between the four groups in the heavy industry sector. Note that the overall situation in the heavy industry sector is somewhat different from what we have seen for the light industry sector. The efficiency distributions cannot be statistically distinguished from each other, the sole exception being the FFEs, for which the distribution appears to be significantly different from those of all the other groups. We also observe significance at the 10% level for TOEs versus COEs (which are not significantly different from each other in the light industry sector).

The results reported in the upper part of Table 3 also show that in comparison with the light industry ownership groups, all the ownership groups in the heavy industry sector have relatively small aggregate inefficiency scores and an (absolutely and relatively) lower level of estimated bias. These results suggest that performance varies to a lesser degree among ownership types in the heavy industry sector. This could be explained by the fact that heavy industry operations are more stable than are operations in the light industry sector, which is more dynamic and features larger numbers of firms breaking through and firms lagging behind, thereby causing more variation in efficiency. In addition, because heavy industry is more capital-intensive in nature and light industry is more labor-intensive, greater automation in the production process leads to less human-driven inefficiency (such as human mistakes and shirking on the job) in the heavy industry sector. Firms operating in heavy industries therefore tend to operate in a relatively similar manner and are more similar in terms of performance, both of which contribute to less variation in efficiency estimates and, in turn, less estimated bias.

Although it has long been held that SOEs are less efficient than their non-state owned counterparts, our results from the analysis of densities and aggregate efficiencies do not provide strong support for this view in the case of the *heavy* industry sector. Specifically, a comparison of *weighted* aggregate efficiencies between the heavy industry groups using RD statistics indicates *no* statistical difference between them. This result could be attributed to the high level of automation in production activities in the heavy industry sector, a factor which has been discussed in the previous paragraph.

A similar test for the non-weighted efficiency scores confirms the insignificance of the differences between these group efficiencies, other than for the FFEs, which appear to be more efficient than SOEs and TOEs (at about the 10% significance level) and COEs (at about the 1% significance level). This is consistent with our analysis of the distributions for these groups, but contrasts with the results obtained for the light industry sector, where we find that FFEs perform significantly *less* efficiently than COEs and TOEs, while SOEs perform significantly worse than all of the other groups. We explain this difference between the industry sectors in more detail later in this work.

Truncated Regression Analysis

The regression analysis method we employ is not simply a generalization of the above analysis because it imposes a particular structure on the dependency between the efficiency of a DMU and the hypothesized explanatory variables. Moreover, the dependent variable (i.e., the efficiency score), does not account for the economic weight (e.g., size) of the observations. Nevertheless, this analysis complements the methods used above in a number of very important respects. In particular, it allows for inferences to be drawn about different factors that simultaneously influence efficiency scores by focusing on the (marginal) effect of each variable. One additional advantage of this approach is that it allows for the effects of continuous variables to be investigated.

Our empirical specification shown on the right-hand side of regression equation (8) includes the intercept, dummy variables and one continuous variable.

The first dummy variable is the industry indicator (1 for light industry and 0 for heavy industry). The next three dummies – D2, D3, and D4 – represent the DMU ownership type and take the value of 1 if the observation belongs to a group of FFEs, TOEs and COEs, respectively. Thus, for the sake of convenience in testing, the group of state owned enterprises is taken as the base and so the coefficients on D2, D3 and D4 would estimate the difference in effects between the corresponding group (e.g., FFEs for D2) and the group of SOEs. For example, a negative coefficient on D2 would suggest evidence that FFEs introduce *improvements* relative to SOEs, on average.

The next two dummies – D5 and D6 – represent the regions and are assigned the value of 1 if the observation belongs to the coastal and central regions, respectively.[9] That is, the coefficients on each of these dummies will estimate the difference in effects between their region (e.g., coastal) and the western region, which is taken as the base. The continuous variable on the right-hand side of the regression model is used to control for the size effects (measured as the logarithm of output) of DMUs. The size effect variable is expected to capture at least part of the *agglomeration* effect of the province: the larger the gross output of a particular type of firm in a province in a given industry, the higher we expect the efficiency level to be for this type of ownership. The agglomeration effect is expected to have a positive influence on efficiency for at least two reasons. First, there is a spillover effect derived from the activities of firms that are in the same general industry sector (light or heavy) but are not direct competitors (e.g., shoemakers versus textile producers,

[9] We follow the categorization used by the State Planning Commission of China: (1) the *Coastal region*, which includes Beijing, Tianjin, Heibei, Liaoning, Shandong, Shanghai, Zhejiang, Jiangsu, Fujian, Guangdong, Hainan, and Guangxi; (2) the *Central region*, which includes Shanxi, Inner Mongolia, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hunan, and Hubei; and (3) the *Western region*, which includes Sichuan, Yunnan, Guizhou, Shaanxi, Gansu, Qinghai, Ningxia, Tibet, and Xinjiang.

etc.). Second, there is also a competition effect between firms producing the same products that is expected to encourage firms to strive for greater efficiency. We expect both effects (the spillover and competition effects) to be 'proxied' by this size control variable, but unfortunately cannot decompose it into its two components in our data or results because of the aggregate nature of our data.

The results of our bootstrapped (truncated) regression analysis with DEA are presented in Table 4.[10] We run several specifications to test the robustness of our conclusions. The results confirm our previous findings, but also shed some additional light on the issue under study. We see consistently strong evidence for the argument that at an aggregate level, non-SOEs of all types of ownership have significantly higher efficiency levels than their SOE counterparts. This evidence is robust in that it is confirmed by all the regression specifications we run. While this result is also consistent with those of many studies and is therefore not surprising, we also provide some interesting new results.

Turning to the pooled models (models 1 to 4) in which we consider both industries under the same frontier, the greatest efficiency improvement over that of SOEs comes from TOEs and is followed in turn by FFEs and COEs.

The size effect in all four models is found to be significant such that larger output leads to a better (smaller, i.e., closer to unity) efficiency score, on average. This finding supports the hypothesis of a positive spillover effect on efficiency. That is, the more activities (total output) performed by a particular type of enterprise in a certain province, the higher the efficiency level is expected to be for that type of

[10] The significance tests are based on bootstrapped confidence intervals using Algorithm 2 of Simar and Wilson (2007), with 1000 replications for the bootstrap bias correction of the DEA estimates and 2000 replications for the bootstrapping of the regression coefficients.

enterprise. Notably, the coefficient of the industry dummy is insignificant (and near zero) in Model 1, so we drop it from Model 2 and, as expected, observe almost no change in the estimates. Interestingly, the coefficients on the regional dummies are insignificant in both Model 1 and Model 2, so we drop these dummies from Model 3 and again see almost no change in the coefficients. In Model 4, we drop both the industry dummy and the regional dummies and the coefficients remain almost the same as in the previous three models. These results suggest that, at least on this aggregate level, neither type of industry nor location has a real effect on the level of efficiency. This finding is contrary to the conventional expectation, at least for the coastal region versus the western or even the central region.[11]

More interesting results are revealed when we consider each industry separately. Models 5, 6, and 9 consider light industry alone, while models 7, 8, and 10 consider heavy industry in isolation. There is no qualitative change in most of the results. The region dummies remain insignificant (and almost zero for heavy industry). However, note that the size effect is much more pronounced now for light industry and is much less pronounced in the heavy industry sector relative to what we observed in the pooled models. This suggests that although the agglomeration effect is present in the heavy industry sector, it is much less pronounced than it is in the light industry environment.

Also note that in the heavy industry context, the largest improvement on state ownership comes from FFEs, while the coefficient on the dummy representing the

---

[11] However, Zelenyuk (2009) reports Monte Carlo evidence suggesting that the power of the test of the significance of coefficients on dummy variables in the Simar-Wilson (2006) model is very low, even when the true difference is quite substantial from an economic standpoint. It is therefore likely that in some cases we are simply unable to reject the null hypothesis of equality of efficiencies due to a relatively small sample size, which is clearly not the same as accepting the null hypothesis.

efficiency difference between COEs and SOE is barely significant. (Recall that in the foregoing analysis, we could not confidently reject the differences between the aggregate group efficiency scores for heavy industry.) On the other hand, we find that in light industry, FFEs make the smallest improvement relative to SOEs (smaller than the other types of ownership), so we use Model 9 to test the efficiency difference between FFEs and other types of ownership in the light industry sector. We see that while SOEs are significantly less efficient than FFEs on average (as was also seen in models 5 and 6), the latter are significantly less efficient than the other (non-state local) ownership groups. Although this result might be somewhat unexpected, it is consistent with the results we obtain using other methods and is robust in this sense. Zelenyuk and Zheka (2006) report a similar result for foreign ownership on a disaggregated level in another transitional country (Ukraine).

5. Concluding Remarks

Over the past three decades, the Chinese economy and its industrial structure have experienced remarkable changes which have been rooted in the reform and open door policy initiated by Deng Xiaoping in 1978. Although these changes have continued to gain pace over time, their impact has not been uniform across different types of ownership, industries, and regions in China. Given the continued growth of China's economic power since the turn of the new millennium, it is imperative to gain a better understanding of how China has achieved its economic success and how its economy will evolve in the near future.

In this paper, we investigate efficiency levels and their determinants for different types of ownership, industries and regions in China. The question of the performance of different types of ownership in general, and in China in particular, is a very sensitive issue that often carries political connotations. It goes without saying that great care is required in selecting reliable methods. We employ several recently developed efficiency analysis methods to examine efficiency variations across different cohorts of Chinese industrial firms. In particular, we employ the latest bootstrap-based estimation procedures involving DEA, aggregation, density estimation and truncated regression. The results obtained in this paper provide robust statistical evidence that contributes to the ownership-performance debate. While some results support the earlier work of Shiu (2000, 2001), others shed significant new light on the ownership-performance nexus.

We confirm that in comparison with state ownership, all the other types of ownership we consider result in an improvement in performance. This finding is highly robust, is supported by most of the models and methods employed, and is no great surprise. It confirms the results of many other studies that claim modern China is no exception to the economic laws of the free market and related incentives offered by the 'invisible hand' of Adam Smith.

A somewhat unexpected finding that is nevertheless robust is that foreign–owned firms perform worse on aggregate than non-state local enterprises in the light industry sector, but perform slightly better than firms of all other ownership types in the heavy industry sector. To the best of our knowledge, this finding is new to the productivity literature and therefore warrants a greater degree of attention than our other findings.

We consider that the main explanation for this phenomenon stems from the fact that heavy industry, on average, is more capital-intensive than light industry and that purchasing and adopting new technology requires greater financing. As a result, foreign investors in the heavy industry sector, most of which are huge multinational corporations, are likely to have an advantage over local firms in introducing more advanced capital equipment and expensive technologies, both of which lead to better performance.

In light industries, even when foreigners have initial technological and capital advantages, local private firms should be able to absorb, adopt and disseminate such technology according to local specifications more easily and quickly in light industries than in heavy ones. On the other hand, because light industries tend to be more labor-intensive than heavy ones, the performance of firms active in the former is more likely to be dependent on local content (culture, traditions, habits, etc.). This is likely to give an advantage to local firms and, given a similar level of technology adoption, should enable them to become more efficient than their foreign counterparts — a prediction we confirm in our study.

Our explanation of the foreign versus local ownership question in the heavy versus light industry puzzle is not entirely new or ad hoc. One theoretical foundation for this explanation is closely related to the technology diffusion argument that goes back at least to the work of Gerschenkron (1962) and Nelson and Phelps (1966), as well as the more recent studies of Grossman and Helpman (1991), Parente, Stephen and Prescott (1994), Banks (1994), and Helpman and Rangel (1999), in various areas of economics.

Possible Extensions

It is worth noting that our results are based on cross-sectional data obtained from the most recently available national census and leave to one side the empirical estimation of changes in efficiency and productivity over time which would be possible with a panel data set. This would be a natural extension of our study and we hope that the work presented in this work provides a good foundation for such future research when new census data become available.

Another natural extension to our work would be to use a *non-parametric* truncated regression method, e.g., proposed by Park, Simar and Zelenyuk (2008), which would be possible when more data become available. Yet another interesting extension would be to test for the stochastic dominance of the distributions of efficiency scores of various ownership groups and regions.[12]

Overall, we hope that our study spurs theoretical development of related methodology issues that can improve our work, as well as encourage more of empirical investigations of the current topic using other methods.

---

[12] We thank Paul Wilson for this remark.

References

American Chamber of Commerce (ACC). 1998. Business Outlook Survey.

Banks, Robert B. 1994. Growth and Diffusion Phenomena. Springer Verlag, Berlin.

Dong, F., Tang, Z. 1995. Studies of Institutional Reforms of Chinese State-Owned Enterprises. The People's Publisher, Beijing.

Färe, R., Grosskopf, S., Norris, M. and Zhang, Z. 1994. Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. American Economic Review 84:1, 66-83.

Färe, R., Primont, D. 1995. Multi-Output Production and Duality: Theory and Applications. Kluwer Academic Publishers, Boston.

Färe, R., Zelenyuk, V. 2003. On Aggregate Farrell Efficiencies. European Journal of Operations Research 146:3, 615-621.

Farrell, M.J. 1957. The Measurement of Productive Efficiency. Journal of the Royal Statistical Society 120, 253-281.

Gerschenkron, A. 1962. Economic Backwardness in Historical Perspective. Cambridge, Belknap Press of Harvard University Press.

Grossman, G.M. and Helpman, E. 1991. Trade, Knowledge Spillovers and Growth. European Economic Review 35, 517-26.

Groves, T., Hong, Y., McMillan, J. and Naughton, B. 1994. Autonomy and Incentives in Chinese State Enterprises. Quarterly Journal of Economics 109, 183-209.

Helpman, E. and Rangel, A. 1999. Adjusting to a New Technology: Experience and Training. Journal of Economic Growth 4, 359-383.

Henderson, D.J. and Russell, R.R. 2005. Human Capital and Convergence: A Production-Frontier Approach. International Economic Review 46, 1167-1205.

Henderson, D.J. and Zelenyuk, V. 2006. Testing for (Efficiency) Catching-up. Southern Economic Journal 73:4, 1003–1019.

Huang, H, Cai, F. and Duncan, R. 1999. Reform of State-Owned Enterprises in China: Key Measures and Policy Debate. ANU, Canberra: Asia-Pacific Press.

Korostelev, A., Simar, L. and Tsybakov, A. 1995. Efficient Estimation of Monotone Boundaries. The Annals of Statistics 23:2, 476-489.

Kneip, A., Park, B. and Simar, L. 1998. A Note on the Convergence of Nonparametric DEA Estimators for Production Efficiency Scores. Econometric Theory 14, 783-793.

Kneip A, Simar L, and Wilson P.W. (2008) "Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models." Econometric Theory 24:1663–1697.

Kumar, S. and Russell, R.R. 2002. Technological Change, Technological Catch-up, and Capital Deepening: Relative Contributions to Growth and Convergence. American Economic Review 92:3, 527– 548.

Li, Q. 1996. Nonparametric Testing of Closeness between Two Unknown Distributions. Econometric Reviews 15, 261-274.

Lin, J.Y., Cai, F. and Li, Z. 1998. Competition, Policy Burdens, and State-Owned Enterprise Reform. American Economic Review 88:2, 422-427.

Melvin, S. 1998. Business Group Sizes up Shanghai's Investment Climate. China Business Review 25:1, 4-4.

Nelson, R.R. and Phelps, E.S. 1966. Investment in Humans, Technological Diffusion, and Economic Growth. American Economic Review 56, 69-75.

Parente, S.L. and Prescott, E.C. 1994. Barriers to Technology Adoption and Development. Journal of Political Economy 102, 298-321.

Park, B.U., Jeong, S.-O. and Simar, L. 2010. Asymptotic Distribution of Conical-hull Estimators of Directional Edges. Annals of Statistics 38: 3, 1320-1340.

Park, B. U, Simar, L. and Zelenyuk, V. (2008), Local Likelihood Estimation of Truncated Regression and its Partial Derivatives: Theory and Application. Journal of Econometrics 146, 185-198.

Russell, R.R. 1990. Continuity of Measures of Technical Efficiency. Journal of Economic Theory 52, 255-267.

Russell, R.R. 1997. Distance Functions in Consumer and Producer Theory in Index Number Theory: Essays in Honor of Sten Malmquist. Kluwer Academic Publishers, 7-90.

Sheather S.J. and Jones M.C. (1991), A reliable data based bandwidth selection method for Kernel density estimation. Journal of the Royal Statistical Society 3, 683-690.

Shiu, A. 2000. Productivity in the Chinese Economy and International Comparisons. Ph.D. Dissertation, University of New South Wales, Sydney.

Shiu, A. 2001. Efficiency of Chinese Enterprises. Journal of Productivity Analysis 18, 255-267.

Simar, L. and Wilson, P.W. 2007. Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Processes. Journal of Econometrics 136, 31-64.

Simar, L. and Zelenyuk, V. 2007. Statistical Inference For Aggregates Of Farrell-Type Efficiencies. Journal of Applied Econometrics 22:7, 1367-1394.

Simar, L. and Zelenyuk, V. 2006. On Testing Equality of Two Distribution Functions of Efficiency Scores Estimated via DEA. Econometric Reviews 25:4, 497-522.

State Statistical Bureau, 1995. The Data of the Third National Industrial Census of the People's Republic of China 1995. Beijing: State Statistical Bureau.

Transparency International. 2001. Global Corruption Report (http://www.globalcorruption-report.org/).

Wei, Z., Varela, O. and Hassan, M.K. 2002. Ownership and Performance in Chinese Manufacturing Industry. Journal of Multinational Financial Management 12, 61-78.

Weitzman, M.L. and Xu, C. 1994. Chinese Township-Village Enterprises as Vaguely Defined Cooperatives. Journal of Comparative Economics 18, 121-145.

Weldon, E. and Vanhonacker, W. 1999. Operating a Foreign-Invested Enterprise in China: Challenges for Managers and Management Researchers. Journal of World Business 34, 94-107.

Wu, Y. 1996. Productive Performance in Chinese Enterprises – An Empirical Study. Macmillan Press, London.

Wu, Y. 1998. Redundancy and Firm Characteristics in Chinese State-Owned Enterprises. The Asia Pacific Journal of Economics and Business 2, 33-44.

Yeung, G. and Mok, V. 2002. Government Policy and the Competitive Advantages of Foreign-Financed Firms in Guangdong Province of Southern China. Asian Business and Management 1, 227-247.

Zelenyuk, V. 2009. Power of Significance Test of Dummies in Simar-Wilson Two-Stage Efficiency Analysis Model. Applied Economics Letters 16:15, 1493-1497.

Zelenyuk, V. and Zheka, V. 2006. Corporate Governance and Firm's Efficiency: The Case of a Transitional Country, Ukraine. Journal of Productivity Analysis 25, 143-168.

Zheng, J. and Hu, A. 2004. An Empirical Analysis of Provincial Productivity in China (1979-2001). Paper presented at the 15th Annual Conference of the Chinese Economic Association (UK).

Zheng, J., Liu, X. and Bigsten, A. 1998. Ownership Structure and Determinants of Technical Efficiency: An Application of Data Envelopment Analysis to Chinese Enterprises (1986-1990). Journal of Comparative Economics 26, 465-484.

Zhang, X.G. and Zhang, S. 2001. Technical Efficiency in China's Iron and Steel Industry: Evidence from the New Census Data. International Review of Applied Economics 15, 199-211.

Table 1. Simar-Zenlenyuk-adapted for DEA Li-test for Equality of Efficiency Distributions Across Different Types of Ownership

Null Hypothesis	Test Statistic	p-val.	Test Statistic	p-val.	Test Statistic	p-val.
	Light Industry		Heavy Industry		Both Industries	
$f(eff_{SOE}) = f(eff_{FFE})$	10.61	0.00	5.83	0.00	16.6	0.00
$f(eff_{SOE}) = f(eff_{TOE})$	12.24	0.00	0.47	0.47	18.2	0.00
$f(eff_{SOE}) = f(eff_{COE})$	12.39	0.00	0.21	0.76	13.52	0.00
$f(eff_{FFE}) = f(eff_{TOE})$	2.32	0.01	1.24	0.07	1.01	0.08
$f(eff_{FFE}) = f(eff_{COE})$	1.19	0.06	4.52	0.00	-0.13	0.85
$f(eff_{TOE}) = f(eff_{COE})$	0.55	0.34	1.23	0.06	1.32	0.56

Notes: All calculations are done by authors in Matlab using, after adopting from programs used for the work in Simar and Zelenyuk (2006).

Table 2. The Light Industry: Group-Wise Heterogeneous Sub-Sampling Bootstrap for Aggregate Efficiencies (aggregation into 4 types of ownership)

	Groups	Original	Bootstrap Estimate of		Est. 95% Conf. Int.	
		Reciprocal of DEA Estimates	Standard Deviation	Bias-Corr. Eff. Score (reciprocal)	Lower Bound	Upper Bound
Weighted (output shares) group efficiencies	1	0.55	0.26	0.43	0.39	0.55
	2	0.71	0.13	0.62	0.57	0.74
	3	0.83	0.07	0.75	0.71	0.88
	4	0.79	0.09	0.71	0.67	0.83
	All	0.69	0.14	0.59	0.55	0.70
Non-weighted group efficiencies	1	0.53	0.3	0.41	0.36	0.53
	2	0.66	0.18	0.55	0.51	0.71
	3	0.75	0.12	0.65	0.61	0.80
	4	0.72	0.13	0.62	0.58	0.77
	All	0.65	0.18	0.54	0.50	0.66
RD statistics for comparing groups in terms of weighted average efficiencies	1 vs. 2***	1.30	0.11	1.48	1.23	1.62
	1 vs. 3***	1.50	0.18	1.80	1.38	2.00
	1 vs. 4***	1.43	0.15	1.70	1.35	1.87
	2 vs. 3**	1.16	0.08	1.23	1.02	1.33
	2 vs. 4**	1.10	0.06	1.16	1.01	1.24
	3 vs. 4	0.95	0.05	0.93	0.85	1.06
RD statistics for comparing groups in terms of non-weighted average efficiencies	1 vs. 2***	1.26	0.11	1.40	1.13	1.54
	1 vs. 3***	1.43	0.16	1.67	1.30	1.84
	1 vs. 4***	1.38	0.14	1.60	1.27	1.74
	2 vs. 3*	1.14	0.08	1.21	0.98	1.31
	2 vs. 4*	1.10	0.07	1.15	0.96	1.25
	3 vs. 4	0.96	0.05	0.95	0.84	1.08

Notes: Groups 1, 2, 3 and 4 refer to SOEs, FFEs, TOEs and COEs, respectively. Also, ***, ** and * indicates the rejection of the null hypothesis of $H_0\colon \overline{TE}^A = \overline{TE}^Z$ at 1%, 5% and 10% significance levels, respectively.

For convenience, we present *reciprocals* of estimated efficiency scores, i.e., $1 / \overline{\widehat{TE}}^l$, $l = 1, 2, 3, 4$ (and the corresponding confidence intervals) so that they have percentage meaning. All calculations are done by authors in Matlab using, after adopting from programs used for the work in Simar and Zelenyuk (2007).

Table 3. The Heavy Industry: Group-Wise Heterogeneous Sub-Sampling Bootstrap for Aggregate Efficiencies (aggregation into 4 types of ownership)

	Groups	Original	Bootstrap Estimate of		Est. 95% Conf. Int.	
		DEA Estimates	Standard Deviation	Bias-Corr. Eff. Score (reciprocal)	Lower Bound	Upper Bound
Weighted (output shares) group efficiencies	1	0.81	0.06	0.75	0.70	0.82
	2	0.88	0.04	0.83	0.79	0.87
	3	0.87	0.04	0.81	0.78	0.89
	4	0.83	0.06	0.76	0.72	0.85
	All	0.83	0.04	0.77	0.73	0.83
Non-weighted group efficiencies	1	0.79	0.07	0.74	0.68	0.80
	2	0.86	0.04	0.81	0.78	0.86
	3	0.83	0.05	0.75	0.71	0.82
	4	0.79	0.07	0.71	0.67	0.81
	All	0.82	0.04	0.75	0.72	0.81
RD statistics for comparing groups in terms of weighted average efficiencies	1 vs. 2	1.09	0.06	1.09	0.98	1.21
	1 vs. 3	1.08	0.06	1.08	0.97	1.18
	1 vs. 4	1.03	0.05	1.00	0.90	1.09
	2 vs. 3	0.99	0.04	0.98	0.90	1.07
	2 vs. 4	0.95	0.05	0.91	0.83	1.02
	3 vs. 4	0.95	0.04	0.93	0.85	1.03
RD statistics for comparing groups in terms of non-weighted average efficiencies	1 vs. 2*	1.08	0.06	1.10	0.99	1.21
	1 vs. 3	1.04	0.05	1.01	0.91	1.10
	1 vs. 4	1.00	0.06	0.95	0.84	1.04
	2 vs. 3*	0.96	0.04	0.91	0.83	1.01
	2 vs. 4***	0.92	0.06	0.85	0.76	0.98
	3 vs. 4	0.97	0.04	0.94	0.86	1.03

Notes: Groups 1, 2, 3 and 4 refer to SOEs, FFEs, TOEs and COEs, respectively. Also, ***, ** and * indicates the rejection of the null hypothesis of $H_0\colon \overline{TE}^A = \overline{TE}^Z$ at 1%, 5% and 10% significance levels, respectively.

For convenience, we present *reciprocals* of estimated efficiency scores, i.e., $1 / \overline{\overline{TE}}^l$, $l = 1, 2, 3, 4$ (and the corresponding confidence intervals) so that they have percentage meaning. All calculations are done by authors in Matlab using, after adopting from programs used for the work in Simar and Zelenyuk (2007).

Table 4. Result of Truncated Regression Analysis for Explaining Inefficiency Level

	Interpretation of Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Intercept	Constant	2.96**	2.96**	2.96**	2.94**	3.33**	3.29**	1.61**	1.61**	2.45**	1.36**
Industry	Industry Dummy (1 if Light, 0 o.w.)	0.00	-	-0.04	-	Light	Light	Heavy	Heavy	Light	Heavy
D1	Ownership Dummy (1 if SOE, 0 o.w.)	-	-	-	-	-	-	-	-	0.84**	0.25**
D2	Ownership Dummy (1 if FFE, 0 o.w.)	-0.73**	-0.72**	-0.73**	-0.72**	-0.85**	-0.84**	-0.26**	-0.25**	-	-
D3	Ownership Dummy (1 if TOE, 0 o.w.)	-0.90**	-0.89**	-0.89**	-0.88**	-1.28**	-1.27**	-0.13**	-0.13**	-0.42**	0.12**
D4	Ownership Dummy (1 if COE, 0 o.w.)	-0.69**	-0.69**	-0.68**	-0.68	-1.06**	-1.05**	-0.04*	-0.04*	-0.21**	0.21**
D5	Region Dummy (1 if Coastal, 0 o.w.)	0.05	0.05	-	-	0.09	-	0.00	-	-	-
D6	Region Dummy (1 if Central, 0 o.w.)	-0.04	-0.04	-	-	-0.04	-	0.01	-	-	-
Log(*y*)	Measure of Size	-0.14**	-0.14**	-0.14**	-0.14**	-0.20**	-0.19**	-0.04**	-0.04**	-0.19**	-0.04**
σ^2	Variance of the error term	0.08**	0.08**	0.08**	0.08**	0.10**	0.11**	0.01**	0.01**	0.11**	0.01**

Notes: Dependent variable is "efficiency score" (see eq. (8)-(9) in the text). Note that * and ** indicate *significance* at α being 5% and 10%, respectively. Significance tests are based on bootstrapped confidence intervals, using Algorithm 2 of Simar and Wilson (2006), with 1000 and 2000 bootstrap replications for bias correction and for confidence intervals, respectively. All calculations are done by authors in Matlab using code of Valentin Zelenyuk, which adopted some earlier codes of Leopold Simar.

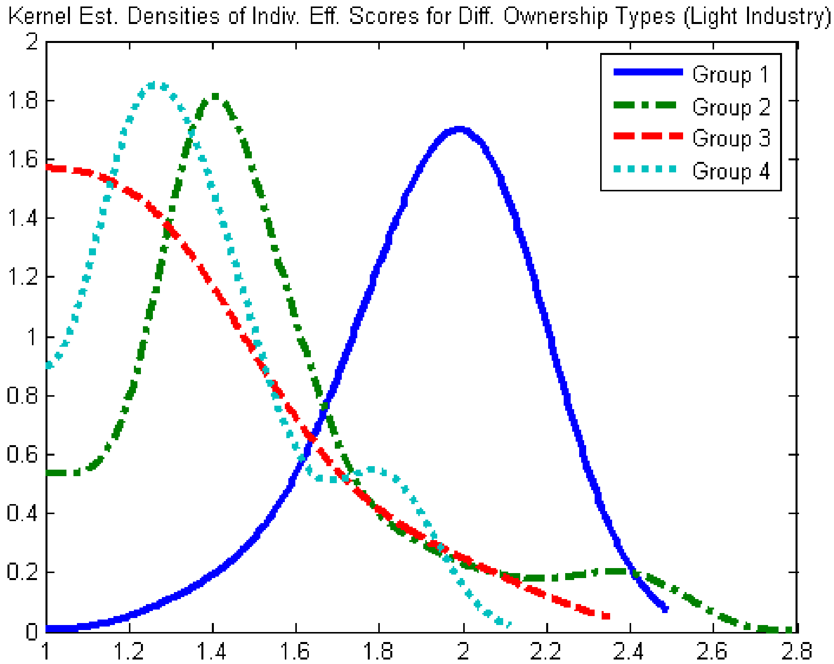

Figure 1. Estimated Densities of Individual Efficiency Scores for Ownership Groups in Light Industry.
Notes: [i] Groups 1, 2, 3 and 4 refer to SOEs, FFEs, TOEs and COEs, respectively.
[ii] Vertical axis refers to (estimated) probability density function of the distribution of the efficiency scores and horizontal axis refers to efficiency scores.

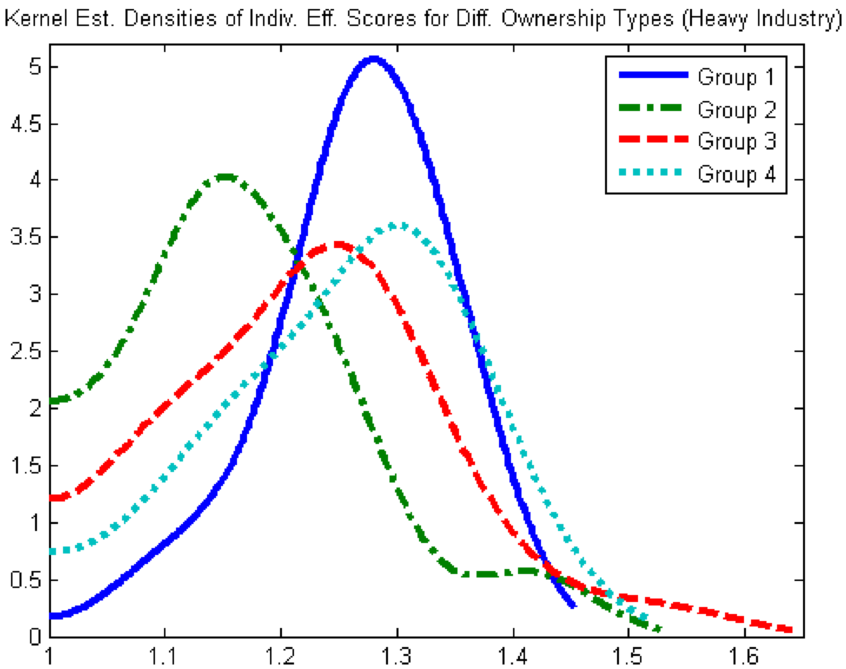

Figure 2. Estimated Densities of Individual Efficiency Scores for Ownership Groups in Heavy Industry.
Notes: [i] Groups 1, 2, 3 and 4 refer to SOEs, FFEs, TOEs and COEs, respectively.
[ii] Vertical axis refers to (estimated) probability density function of the distribution of the efficiency scores and horizontal axis refers to efficiency scores.

APPENDIX

All inputs and outputs used in our activity analysis model for the DEA estimator are measured in units of one hundred million Chinese yuan. Total wage refers to the total remuneration paid to staff and workers during a certain period. This includes wages, salaries and other payments to staff and workers regardless of their source, category and form (in kind or in cash). The net value of fixed assets is calculated as the original value of fixed assets minus depreciation, in which the original value of fixed assets owned by the enterprise is calculated as the price paid at the time the assets were purchased, installed, reconstructed, expanded or subject to technical innovation and transformation. These include expenses incurred in purchasing, packaging, transportation and installation, and so on. The value of intermediate inputs is proxied as the difference between the gross value of industrial output and value added. These are goods that have been processed in one production process and then sold for final processing in another production process. The gross industrial output is the total volume of industrial products sold or available for sale in value terms. It includes the value of finished products and the value of industrial services.

Tables A1 and A2 show the summary statistics for each ownership type in the heavy and light industry sectors, respectively. See Shiu (2000, 2001) for more information and a discussion of the data set.

Table A1. Summary Statistics for Ownership Types (Heavy Industry)

	Gross Industrial Output (hundred million yuan)		Net Value of Fixed Assets (hundred million yuan)		Total Wage (hundred million yuan)		Intermediate Inputs (hundred million yuan)	
Ownership	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
SOEs	609.05	401.22	568.77	322.74	65.89	39.96	396.75	267.34
COEs	152.79	258.26	89.19	181.46	6.71	10.36	111.00	189.47
TOEs	197.71	295.94	58.73	75.45	12.69	15.65	147.86	227.69
FFEs	152.79	258.26	89.19	181.46	6.71	10.36	111.00	189.47

Table A2. Summary Statistics for Ownership Types (Light Industry)

	Panel A:		Panel B:		Panel C:		Panel D:	
	Gross Industrial Output (hundred million yuan)		Net Value of Fixed Assets (hundred million yuan)		Total Wage (hundred million yuan)		Intermediate Inputs (hundred million yuan)	
Ownership	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
SOEs	283.51	206.26	167.44	110.05	22.86	15.94	209.46	157.48
COEs	278.80	354.33	86.06	93.20	22.23	21.97	206.38	273.89
TOEs	204.01	336.27	52.84	88.17	11.36	19.00	158.39	263.53
FFEs	216.65	405.08	82.45	134.91	12.22	23.89	169.26	318.28