

Effects of Ownership and Business Portfolio on Production in the Oil and Gas Industry

Binlei Gong^a

ABSTRACT

The Shale Revolution and the two oil crises have overwhelmingly reshaped the petroleum industry in the last decade. Heterogeneity across companies is also a big concern as many multi-product (oil and gas) and multi-segment (upstream and downstream) firms exist, both state-owned and privately-owned. Therefore, a varying coefficient model is introduced to capture the effects of time, ownership, and business portfolio on both productivity and input elasticities to closely observe the fundamental transition, which is further interpreted using decomposition equations. The shape of the production function is indeed firm- and time-variant, which confirms the transition of the industry and the necessity of using the varying coefficient model. The average productivity achieved tremendous growth after the 2007–2009 financial crisis but lost momentum following the 2014 price crash. Finally, privately-owned, gas production and downstream activities are more productive than state-owned, oil production and upstream activities, respectively. Some policy implications are also discussed.

Keywords: Stochastic frontier analysis, Varying coefficient model, Oil and gas industry, Effects of state ownership, Segment-wide and product-wide business portfolio

<https://doi.org/10.5547/01956574.41.1.bgon>

1. INTRODUCTION

Remarkable growth has been witnessed in the oil and gas industry due to the Shale Revolution over the past two decades (Jacoby, O’Sullivan, and Paltsev, 2012). Technical innovations, such as hydraulic fracturing and horizontal drilling, have boosted the investment and production, especially in the unconventional oil and gas market (Gong, 2018a). During the same period, high volatility in oil prices (especially in 2008 and 2014), also reshaped the petroleum industry as the oil price crashes forced business transformations and a market reconstruction. These events not only significantly affected productivity but also overwhelmingly reshaped the production process of the industry. As a result, the shape of the production function may be time-variant.

The oil and gas industry is a complex sector that involves many steps in the energy supply chain. Some petroleum enterprises are multi-segment, having footprints in both upstream and downstream segments, or multi-product, generating both crude oil and natural gas. The productivity for various segments and products is different, since the technologies and production process are different. As a result, the total factor productivity of a company depends on its business portfolio, which measures the ratio of upstream and downstream activities, as well as the ratio of oil and gas production. Moreover, National Oil Companies (NOCs) and privately owned International Oil Companies

^a Department of Agricultural Economics and Management, School of Public Affairs, China Academy for Rural Development (CARD), Zhejiang University, 310058, Hangzhou, China.

(IOCs) are also found to be significantly different due to various commercial and non-commercial objectives (Hartley and Medlock, 2008), which may cause differences in economic productivity. Therefore, many existing studies have studied the effects of ownership and business portfolio on productivity for oil and gas companies.¹ More importantly, the Shale Revolution over the past two decades first affected gas production and then affected oil production, and this mainly occurred in the upstream segment and IOCs. As a result, the effects of ownership and business portfolio on productivity for oil and gas companies are very likely to vary over time; new investigation of these impacts is needed.

Although the effects of ownership and business portfolio on productivity have been previously evaluated, the effects of these factors on the shape of the production function are overlooked in the literature, as most productivity analyses assume a fixed production function. Different segments and different products involve different technologies; therefore, each has its own production process and input–output relation, which is reflected in the shape change of the production function measured by changes in input elasticities. Furthermore, ownership may also affect input elasticities, since overemployment usually occurred in NOCs for non-commercial reasons (Hartley and Medlock, 2008), which can change the labor elasticity relative to the elasticities of other inputs. Finally, the Shale Revolution may also reshape the production function dramatically. Therefore, the segment-wide and product-wide business portfolio of a company, as well as its ownership and time, may also affect input elasticities (i.e., the shape of the production function), in addition to their impacts on productivity (i.e., the shift of the production function).

To summarize, both ownership and business portfolio, along with time, can influence the production function. The first puzzle is that the classic method assumes the shape of the production function is fixed, which fails to capture the varying input–output relation over time and across firms. To solve this issue, this paper employs a varying coefficient production function to better estimate the true data-generating process, where the production function not only can shift, as in the classic model, but can also be reshaped depending on the aforementioned factors. Although the varying coefficient model captures the heterogeneity in time, ownership, and business portfolio, their effects are modeled nonparametrically and are therefore hard to clarify. This second problem motivates us to further decompose input elasticities (i.e., the shape of the production function) and total factor productivity (i.e., the shift of the production function) parametrically, so that their impacts can be better interpreted.

Econometrically, this paper explores the effects of ownership and business portfolio, along with the time trend, on both the shape and the shift of the production function using a two-step method. In the first step, a varying coefficient stochastic frontier model is established to estimate the production function. The varying input elasticities and productivity can be derived when the heterogeneity in ownership and business portfolio across firms is controlled. In the second step, a parametric regression model is employed to further interpret the impacts of time, ownership, and business portfolio on input elasticities and productivity, respectively.

Three central contributions are made in this paper. First, a varying coefficient method is used to better capture the fundamental transition of the petroleum industry over the past decade. Second, this study extends the comparison between NOCs and IOCs by not only interpreting more accurate effect of ownership on productivity than was estimated with bias in the literature, but also investigating its impacts on the input–output relation that was overlooked. Third, this paper also controls the heterogeneity in segment-wide and product-wide business portfolio and predicts their

1. A detailed introduction of these studies are introduced and discussed in Section 2.

effects on production, which contributes to the studies of multi-product, multi-segment and multidivisional firms.²

Using data of 110 major petroleum enterprises during the time period of 2009–2016, this paper delivers the following empirical findings: 1) the shape of the production function (measured by input elasticities) is indeed firm- and time-variant, which confirms the fundamental transition of the oil and gas industry in the last decade and the necessity of using a varying coefficient model; 2) labor elasticity was increasing, the elasticities of oil and gas reserves were decreasing, and the elasticity of refining capacity was relatively flat over time; 3) average productivity achieved tremendous growth after the 2007–2009 financial crisis but has lost momentum following the oil price crash since the end of 2014; 4) segment-wide and product-wide portfolios, as well as ownership, do significantly affect the input elasticities and productivity; and 5) IOCs, gas production and downstream activities, *ceteris paribus*, are more productive than NOCs, oil production and upstream activities, respectively.

The remainder of the paper is structured as follows. Section 2 reviews literature, Section 3 illustrates the two-step method, and Section 4 describes the data. Section 5 then presents the estimation results, and Section 6 draws a conclusion.

2. LITERATURE REVIEW

Cuñado and de Gracia (2003) and Jiménez-Rodríguez and Sánchez (2005) point out that economies are sensitive to oil supply and price volatility due to their strong dependence on oil-based products. Therefore, it is very important to study the productivity and efficiency of petroleum enterprises, as it directly determines energy supply and affects energy prices. A few studies investigate the productivity and efficiency of petroleum enterprises. The majority of them analyze the difference between NOCs and IOCs and agree that IOCs are more efficient and productive. Al-Obaidan and Scully (1992) introduce a stochastic frontier analysis (SFA) to evaluate the efficiency of 44 oil companies and claim that IOCs are more efficient than NOCs. Wolf (2009) employs a classic production function to study large petroleum enterprises for 1987–2006 and also concludes that NOCs are not as efficient as IOCs. Eller, Hartley, and Medlock (2011) and Hartley and Medlock (2013) predict the firm-level efficiency of oil companies using both SFA and DEA, and they find results consistent with previous studies. Gong (2018b) utilizes two types of SFA to model the production function and asserts that NOCs are significantly less productive. In terms of the reasons that cause the difference, Hartley and Medlock (2008) attribute the inefficiency of NOCs to non-commercial objectives, such as price subsidies and overemployment under political pressure. Cabrales, Bautista, and Benavides (2017) evaluate the effect of domestic fuel subsidies and employment on NOCs' performance, which provides evidence of the trade-offs among non-commercial objectives and commercial ones. The effect of ownership is also a hot topic in other industries and the differences between state-owned and private companies are investigated by many researchers (e.g., Qiang and Tuohan (2016), Zhu (2016), Xiao (2016), Chen and Fu (2017), Moretz (2018), and Rahman (2018)).

Besides the effect of ownership, some economists also discuss the impacts of segment-wide, product-wide, and region-wide heterogeneity across companies. Al-Obaidan and Scully (1992) consider the segment-wide portfolio and use the output share by segment in the production function. This data is used to calculate a vertical integration ratio, which is then treated as an explanatory variable in the production function and can therefore only affect productivity, rather than input elasticities. Ike and Lee (2014) use a DEA model to estimate efficiencies of the 38 oil and gas companies

2. For more information on multidivisional firms, please refer to Gong (2016).

for 2003–2010, and then consider vertical-integration and international operations variables in the second-step regression to evaluate their effects on firm-level efficiency. Sueyoshi and Goto (2012) attribute the difference between domestic and international companies to the more restricted international standard. Ohene-Asare, Turkson, and Afful-Dadzie (2017) compare domestic and international firms by some nonparametric statistical significance tests, after the DEA method is utilized to estimate the efficiencies of 50 petroleum enterprises from 2001–2010. Gong (2018b) investigates the effect of upstream ratio and gas ratio on firm-level efficiency by regression method, where the efficiencies are derived from parametric and semiparametric SFA models. Gong (2018c) evaluates the segment-wide, product-wide, and region-wide spillover effects and competitions among petroleum enterprises.

However, most of these studies assume a fixed production function that only shifts up and down. In other words, the shape of the production function is the same for all companies, regardless of their ownership and business portfolio. This paper has discussed in the introduction that ownership and business portfolio may affect the shape of the production function, as different technologies are used in various segments and products, and the labor elasticity can be affected by ownership due to the overemployment phenomenon among NOCs, which is studied in Hartley and Medlock (2008) and Cabrales, Bautista, and Benavides (2017). In the methodology section, this paper also mathematically illustrates that the fixed production function not only overlooks the effects on input elasticities, but also evaluates the effects on productivity with bias. Gong (2018a) captures segment-wide heterogeneity that affects the shape of the production frontier using a varying coefficient model, but he only studies some oilfield service firms, rather than the petroleum enterprises. Moreover, the effects of ownership and product-wide heterogeneity are overlooked in Gong (2018a). This paper considers both ownership and business portfolio to investigate their effects among petroleum enterprises using a varying coefficient model.

Recent studies suggest using labor, oil reserves, gas reserves and refining capacity as the four inputs, and revenue as the output. Wolf (2009) introduces total resource reserves as a third input to produce oil and gas besides the labor and assets employed by Al-Obaidan and Scully (1992). Eller, Hartley, and Medlock (2011) drop total assets from the input portfolio and further separate oil reserves and gas reserves as two different inputs. Although a severe distortion problem due to inflation can be avoided, the abandonment of total assets also ignores capital of the company other than reserves. Hartley and Medlock (2013) and Gong (2018b) adopt refining capacity as an input on the top of the input portfolio in Eller, Hartley, and Medlock (2011) so that oil and gas reserves measure the major capital in the upstream segment and refining capacity measures the major capital in the downstream segment. In terms of the output, Hartley and Medlock (2008) suggest the use of revenue instead of production as the output for three reasons: 1) quantities of oil and gas overlook price subsidies that are widely observed in NOCs; 2) revenue is a good approach to aggregate multiple products (e.g., oil and gas) by their prices; and 3) data availability is another important concern as revenue data are easier to obtain. Moreover, the strong correlation between physical outputs and revenue in oil and gas companies is found in Wolf (2009), which empirically confirms the robustness of revenue. Due to the aforementioned virtues, revenue is widely accepted to measure the output of petroleum enterprises in recent studies (Eller, Hartley, and Medlock 2011, Hartley and Medlock 2013, Gong 2018b).

To summarize, existing studies reach some agreement on the selection of inputs and outputs. Furthermore, SFA and DEA are the two major methods utilized in the productivity and efficiency analysis of the oil and gas industry. However, the segment-wide and product-wide heterogeneity and ownership characteristics of petroleum enterprises, along with the time trend, are not fully

considered, especially their effects on the shape of the production function. Therefore, this paper uses a two-step method (a varying coefficient model and a parametric decomposition model) to control and then interpret the effects of ownership and business portfolio in the petroleum industry. In terms of input and output selection, this paper is consistent with those choices in recent studies. In terms of the econometric tool, SFA is utilized to model the varying coefficient model in order to better explore the impacts of ownership and business portfolio on input elasticities, which cannot be fulfilled in DEA. Finally, this paper assumes that the production technology in this industry exhibits constant returns to scale (CRS) (Eller, Hartley, and Medlock, 2011; Hartley and Medlock, 2013; Gong, 2018b).

3. METHODOLOGY

Our econometric model includes two parts. First, a varying coefficient stochastic frontier model is employed to estimate the production function of the petroleum enterprises. The predicted input elasticities and total factor productivity are then separately regressed on ownership and business portfolio to further explore their effects on production.

3.1 Stochastic Frontier Analysis

This paper uses stochastic frontier analysis, since it is widely used in energy studies (e.g., Boyd (2008), Kumbhakar and Lien (2017), Llorca et al. (2017), Orea, Álvarez, and Jamasb (2018)). The stochastic frontier model consists of a deterministic frontier production function and a symmetric random error and was first proposed by Aigner, Lovell, and Schmidt (1977) and Meeusen and Van den Broeck (1977) to predict the production frontier in a Cobb-Douglas (C-D) form.

$$y_i = x_i' \beta + v_i - u_i, \tag{1}$$

where y_i is the output and x_i is the vector of inputs of firm i , both in logarithms. v_i is the disturbance, which is the stochastic part and follows a normal distribution. u_i accounts for a non-negative random variable measuring firm-level technical inefficiency, which is assumed to follow a variety of distributions, such as a normal truncated distribution (Stevenson, 1980) or gamma distribution (Greene 1990).

Schmidt and Sickles (1984) extend the utilization of stochastic frontier analysis from the cross-sectional to panel data setting in the form:

$$y_{it} = \alpha + x_{it}' \beta + v_{it} - u_i = \alpha_i + x_{it}' \beta + v_{it}. \tag{2}$$

where α_i measures total factor productivity. Over the next three decades, the focus of stochastic frontier analysis has been how to estimate α_i more properly under different circumstances. According to the characteristics of various industries, different estimation strategies are proposed, including specifications that allow time-variant estimates. Sickles (2005) and Gong and Sickles (2016) introduce a number of competing approaches and compares them under different settings.

It is worth noting that the coefficients β in the classic model are fixed, which implies constant elasticity for each input. In other words, the input–output relation is invariant. However, the production function may be distinct between NOCs and IOCs, considering their differences in both objective and structure. Moreover, the techniques utilized to generate oil products and gas products are not similar and should therefore be product-variant. The same rule applies to upstream activities and downstream activities, where different production processes lead to segment-variant input–out-

put relations. Finally, the Shale Revolution and oil crises in the last decade not only dramatically affected productivity but also overwhelmingly reshaped the production process, which may result in a time-variant production function. To summarize, fixed coefficients β might not be a valid assumption for the oil and gas industry, where companies with different ownerships and business portfolio exist.

The Varying Coefficient Model (VCM) is proposed by Hastie and Tibshirani (1993), where the coefficients are nonparametric functions of some “threshold” variables θ and therefore have the name “varying coefficients” (Chen and Hsu, 2013) or “random coefficients” (Farsi, Fetz, and Filipini 2008). The original model has the form.

$$y = x_1 h_1(\theta_1) + \dots + x_p h_p(\theta_p) + \epsilon, \quad (3)$$

where $\theta_1, \dots, \theta_p$ decide the coefficients of x_1, \dots, x_p through unspecified nonparametric functions $h_1(\cdot), \dots, h_p(\cdot)$.

Time is usually employed as the “threshold” variable to derive time-varying coefficients (Li, Joyeux, and Ripple 2014; Nepal and Foster, 2016). In the field of production analysis, Sun and Kumbhakar (2013) and Zhang et al. (2012) study the Norwegian forest industry and China’s high-tech industry using the varying coefficient model, respectively. As discussed in Kumbhakar and Sun (2013), candidates for “threshold” variables in the production function include tax rate, R&D spending, firm age, and firm size. Gong (2018a) utilizes a varying coefficient stochastic frontier model to explore the efficiency of the global oilfield market using revenue shares by segment as the “threshold” variables, which is an industry that is highly related to the oil and gas industry being studied by the present paper. However, the effect of time is overlooked, let alone the product-variant techniques.

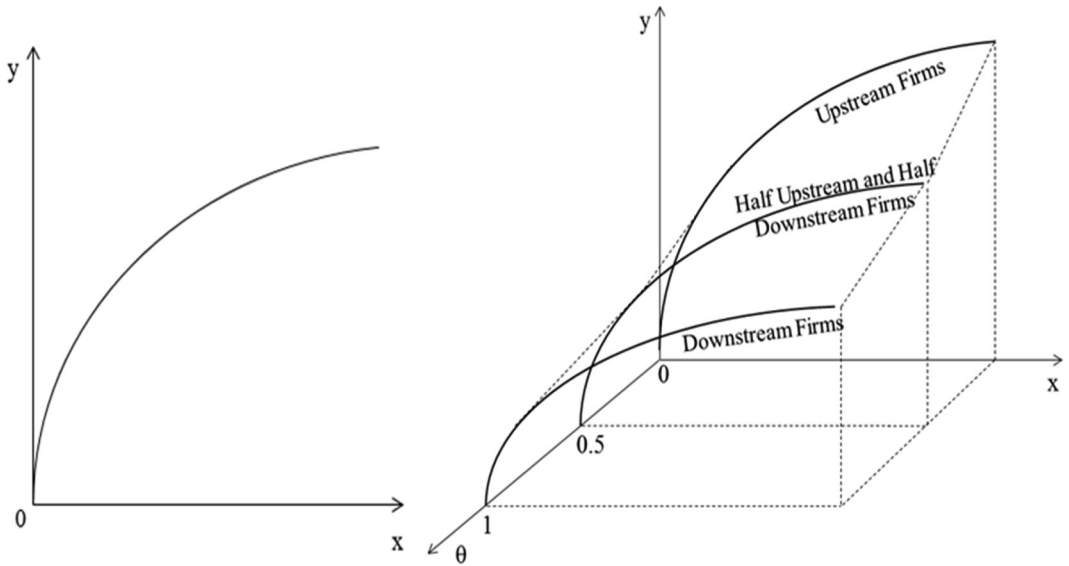
In the oil and gas industry, many companies are multi-product, generating both oil and gas, as well as multi-segment, which has footprints in both upstream and downstream activities (Rogers 2003). Moreover, NOCs and IOCs are quite different, and therefore studying their differences has become the focus of the industry in existing studies. Finally, the heterogeneity over time should not be ignored in the context of the Shale Revolution and oil crises. Therefore, this paper uses time, ownership and business portfolio (measured by revenue shares by product and by segment) as the “threshold” variables that decide the magnitude of the coefficients. As a result, companies share the same production frontier and can be directly comparable only if they have the same level of all “threshold” variables. Using this method, only an “apples to apples” comparison is allowed, whereas comparing apples and oranges is prevented.

A varying coefficient stochastic frontier model is built to model the production function of the petroleum enterprises, where time, ownership and business portfolio are treated as “threshold” variables:

$$y_{it} = f(x_{it}) + v_{it} - u_{it} = \alpha_{it} + \sum_{k=1}^4 h_k(\theta_{it}) x_{it}^k + \tau Z + v_{it} - u_{it}, \quad (4)$$

where y_{it} is the output, x_{it}^k is the k -th inputs, and $h_k(\theta_{it})$ is a nonparametric function to estimate β_{it}^k , the varying elasticity of the k -th input, of firm i at time t . $\theta_{it} = (t, noc_{it}, gas_{it}, down_{it})$ where noc_{it} , gas_{it} and $down_{it}$ measure a dummy variable for NOCs, the output share of gas products, and the output share of downstream activities of firm i at time t , respectively. Z vectors a group of year dummy variables, which controls the shift of the production frontier over time. v_{it} , the disturbance, is assumed to be normally distributed with a mean of zero and a standard deviation of σ_v , and $TE_{it} = \exp(-u_{it})$ accounts for technical efficiency. This study uses the most widely used “Error Components Fron-

Figure 1: Comparison of the Classic Model (Left) and Varying Coefficient Model (Right)



tier” (Battese and Coelli 1992) to model $u_{it} = \exp(-\eta(t-T)) u_i$, where u_i is an i.i.d. non-negative truncation of the $N(\lambda, \sigma^2)$ distribution. Eq. (4) is a semiparametric production frontier model as the coefficients of the Cobb-Douglas production function are nonparametrically estimated. Moreover, the estimate strategy of Eq. (4) is illustrated in Appendix A. Finally, endogeneity problems and robustness checks are discussed in Appendix B.

Figure 1 visualizes the differences between the classic method and the varying coefficient method. Suppose there are three types of oil and gas companies, including specialized upstream firms, specialized downstream firms and integrated firms that have footprints in both segments. The x- and y-axes represent the firm-level input and output, respectively. The left figure shows the spirit of the classic method, where all three types of firms share the same production frontier, which has a fixed shape but can shift over time. Therefore, we are comparing apples (upstream firms) and oranges (downstream firms) directly. The right figure shows that the shape of the production frontier varies according to the share of the output from downstream activities in companies, which is a “threshold” variable θ . This approach considers the heterogeneity in technology and production process across segments. Companies share the same frontier and are directly compared to derive efficiency if and only if they have the same business portfolio (“apples to apples”).

3.2 Decomposition Equations of Input Elasticities and TFP.

Although the varying coefficient model captures the effects of ownership, business portfolio and time, they are modeled nonparametrically and are therefore difficult to clarify. This motivates us to further decompose input elasticities and total factor productivity (TFP) parametrically so that the impacts can be better interpreted. More specifically, this paper further explores the effects of ownership and business portfolio, along with the time trend, on petroleum production from two perspectives, including their impacts on input elasticities and their impacts on productivity through Eqs. (5) and (6), respectively. On the one hand, the decomposition equation for each input elasticity, modeled in Eq. (5), can describe how ownership and business portfolio reshape the production frontier. On the other hand, the TFP determination regression in Eq. (6) can reveal how ownership and

business portfolio shift the production frontier. It is worth noting that the dependent variables in Eqs. (5) and (6), the input elasticities and TFP in logarithms, are predicted from Eq. (3).

$$\log(\beta_{it}^k) = \alpha^k + \delta_1^k \text{noc}_{it} + \delta_2^k \text{gas}_{it} + \delta_3^k \text{down}_{it} + \tau^k Z + \varepsilon^k, \quad \forall k = 1, \dots, 4, \quad (5)$$

$$\log(TFP_{it}) = \alpha + \delta_1 \text{noc}_{it} + \delta_2 \text{gas}_{it} + \delta_3 \text{down}_{it} + \tau Z + \varepsilon, \quad (6)$$

where β_{it}^k is the varying coefficient for the k -th input to measure its elasticity, and TFP_{it} accounts for the total factor productivity. Both of them are derived in Eq. (3) for firm i at time t . noc_{it} is a dummy variable for NOCs. gas_{it} measures the revenue share of gas products, whereas down_{it} accounts for the revenue share of downstream activities. Both of them describe the business portfolio of firm i at time t from different perspectives. It is worth noting that δ_1 indicates the actual effect of ownership on productivity. If a constant coefficient model, rather than a varying coefficient model, is used, the effect of ownership on input elasticities (δ_1^k) will be overlooked and mistakenly included in the TFP part, which leads to the biased effect of ownership on productivity, including both δ_1 and δ_1^k . The same rule applies to the effect of business portfolio. Therefore, the conventional constant coefficient model not only overlooks the effects on input elasticities, but also evaluates the effects on productivity with bias, which confirms the necessity of using the varying coefficient model. Finally, the endogeneity problems of the decomposition regressions are also discussed in Appendix B.

4. DATA

This paper follows the inputs and output selection in traditional literature (e.g., Eller, Hartley, and Medlock, 2011; Hartley and Medlock, 2013; and Gong, 2018b). On the one hand, total revenue, rather than quantities of products, is selected as the output in the stochastic frontier model since it is a classic way to aggregate multiple products (such as oil and gas) to reflect their relative value. Furthermore, the revenue data are easier to collect and verify through multiple sources and can better control the impact of price subsidies for some NOCs (Hartley and Medlock, 2008). On the other hand, four inputs are introduced for oil and gas companies, including the number of employees, oil reserves, gas reserves and refining capacity. Besides inputs and output, the stochastic frontier model also includes oil price realization and gas price realization as explanatory variables, as suggested in Eller, Hartley, and Medlock (2011), Hartley and Medlock (2013) and Gong (2018b). To explore the effects of ownership and business portfolio, this paper also requires ownership information, the product-wide revenue share between oil and gas, and the segment-wide revenue share between upstream and downstream activities.

The primary data source is the Energy Intelligence's "Top 100: Global NOC & IOC Rankings," which includes all the variables introduced above. This dataset provides information from one hundred of the biggest oil companies in the world annually. Since the top 100 ranking changes over time, unbalanced panel data from 110 companies covering the time period of 2009–2016 with a total of 639 observations are collected. In earlier studies of the industry, Al-Obaidan and Scully (1992) use a sample of 44 petroleum enterprises for 1979–1982, Thompson et al. (1996) have data on 14 integrated oil companies for the years 1980–1991, Wolf (2009) builds a sample of the 50 largest oil companies over the period 1987–2006, Eller, Hartley, and Medlock (2011) analyze a sample of 78 oil firms during the period 2002–2004, and Hartley and Medlock (2013) collect a sample of 61 oil companies from 2001–2009. Compared with these existing literatures, this paper has the newest and second-largest dataset, which will help analyze the petroleum industry after the 2007–2009 financial crisis.

Table 1: Summary Statistics

Variable	Explanation	Unit	Mean	S.D.	Min	Max
<i>y</i>	Revenue	billion \$	58.2	92.3	0.7	482.3
<i>Labor</i>	Number of employees	'000	59.1	183.3	0.1	1670
<i>OilRsv</i>	Oil reserves	MMbbl	9,657	38,604	0	300,878
<i>GasRsv</i>	Gas reserves	Bcf	33,252	102,206	0	816,423
<i>RefCap</i>	Refining Capacity	'000 b/d	672	1,146	0	6,271
<i>OilPr</i>	Oil price realization	\$/BOE	67.1	21.8	4.4	112
<i>GasPr</i>	Gas price realization	\$/Mcf	4.24	1.39	1.18	16.36
<i>gas</i>	Share in gas products	%	45	26	0	100
<i>down</i>	Share in downstream	%	34	33	0	100
<i>noc</i>	NOC=1; IOC=0	—	0.34	0.47	0	1

Notes: Sample size is 639; Revenue in billion U.S. dollars is the output; The four inputs include number of employees in thousands, oil reserves in million barrels, gas reserves in billion cubic feet, and refining capacity in thousand barrels per day; Oil prices and gas prices are control variables; Share in gas products and share in downstream measure business portfolio and the dummy variable for NOC is an indicator of ownership.

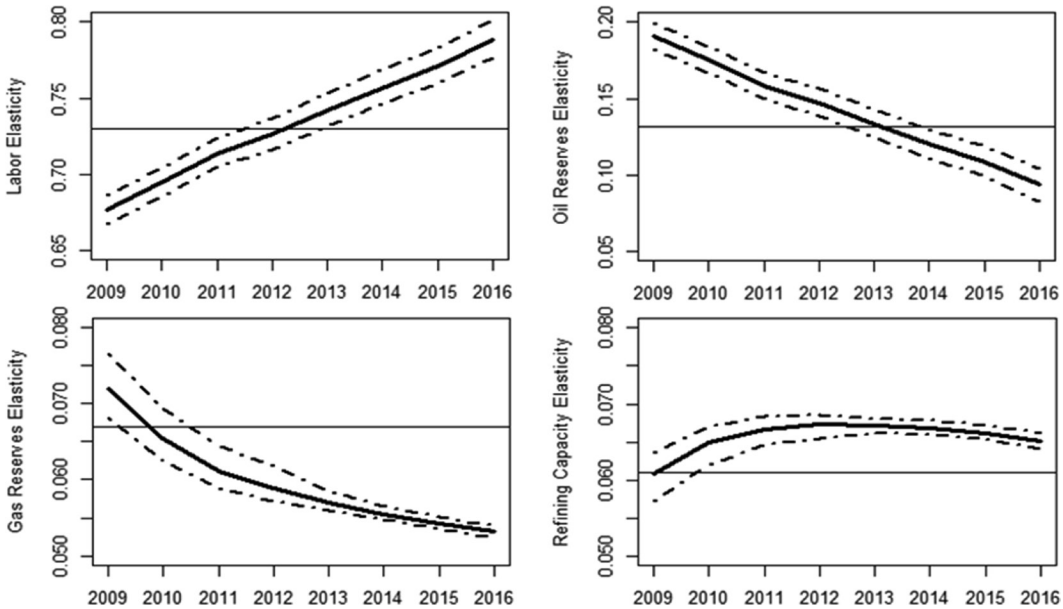
Table 1 provides summary statistics of the variables of interest in the dataset that cover 110 petroleum enterprises during the period of 2009–2016. On average, these companies had 59,100 employees, 9,657 million barrels (MMbbl) of oil reserves, 33,252 billion cubic feet (Bcf) of gas reserves, and a refining capacity of 672,000 barrels per day (b/d) to generate a revenue of 58.2 billion U.S. dollars, on average. During the same period, the average oil price and gas price realizations were 67.1 U.S. dollars per barrel of oil equivalent (\$/BOE) and 4.24 U.S. dollars per thousand cubic feet (\$/Mcf). In terms of the business portfolio product-wide, the revenue share of oil products and gas products is 55% and 45%, respectively. In terms of the business portfolio segment-wide, the revenue share of upstream and downstream activities is 66% and 34%, respectively. Finally, 34% of the companies are NOCs, while 66% of the companies are IOCs. More detailed information of the companies covered in the dataset is provided in Appendix C.

5. ESTIMATION RESULTS

The described model is applied to a panel of 110 petroleum enterprises from 2009 to 2016. This paper estimates the production function using the varying coefficient stochastic frontier model in Eq. (4). The input elasticities and TFP are derived, which are fairly robust when the dynamic stochastic frontier model proposed in Zhang et al. (2015) is adopted. This paper then predicts the effects of ownership and business portfolio on input elasticities and TFP using conventional OLS regressions in Eqs. (5) and (6). These effects are also fairly robust when delay independent variables and propensity score matching are employed to deal with the simultaneity bias. The introduction of these robustness checks are given in Appendix B, while the estimation results of the robustness checks are reported in Appendix D.

5.1 Production Frontiers and Input Elasticity

The stochastic frontier model in Eq. (4) estimates varying elasticities for each of the four inputs (i.e., labor elasticity (β_u^1), oil reserves elasticity (β_u^2), gas reserves elasticity (β_u^3), and refining capacity elasticity (β_u^4)). Figure 2 presents the average elasticities for the four inputs across time. This paper also provides 95% confidence intervals in Figure 2 (the dotted lines) by utilizing Efron’s nonparametric bias-corrected and accelerated (BCa) bootstrap method with 10,000 replications (Briggs, Mooney, and Wonderling, 1999). Moreover, the horizontal lines indicate the values of the fixed input elasticities derived from the classic SFA model.

Figure 2: Average Elasticities of the Four Inputs over Time

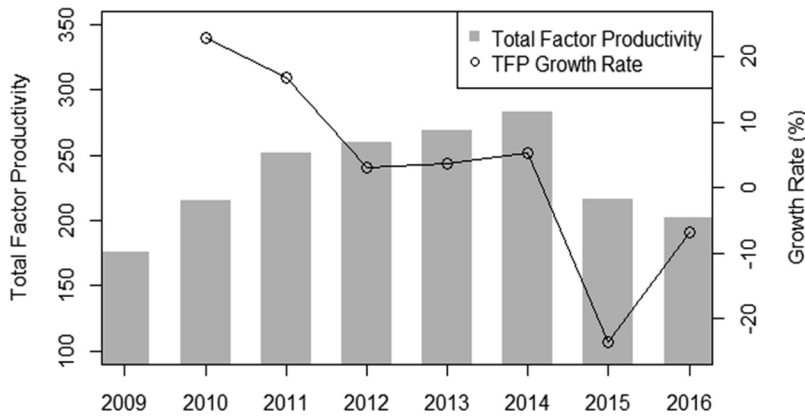
The upper left and right graphs provide the changes in elasticities for labor and oil reserves, respectively, where the former was gradually increasing and the latter was gradually decreasing over time. The lower left graph shows that the elasticity of gas reserves decreased, although at a lower speed in recent years. The lower right graph shows that the elasticity of refining capacity increased in the first period, stayed flat in the second period, and dropped slightly in the third period. A cross-input comparison indicates that labor has the greatest elasticity and therefore contributes the most to revenue growth, followed by oil reserves, while the growths in gas reserves and refining capacity provide the least contribution to revenue growth. Finally, the average levels of the varying elasticities are close to the fixed input elasticities derived by the classic model, as expected.

5.2 Total Factor Productivity

In addition to input elasticities, the varying coefficient stochastic frontier model can also derive total factor productivity (TFP), which ought to be more accurate than the TFP derived from the classic model since the heterogeneous input–output relations are better captured. Figure 3 presents the TFP, as well as its growth rate in the oil and gas industry, over the period of 2009–2016. The petroleum industry achieved tremendous growth after the 2007–2009 financial crisis, when two-digit growth rates in TFP are witnessed in 2010 and 2011. Then the annual growth rates in TFP fell back to around four percent for three years in a row, from 2012 to 2014. Due to the oil price crash started from the fourth quarter of 2014, the market lost momentum dramatically and experienced about a one-quarter loss in TFP in 2015. As the petroleum industry was still experiencing a slowdown, TFP further declined, although at a lower speed in 2016.

To summarize, the changes in TFP over time follow the business cycle as expected: TFP achieved dramatic growth when the oil price recovered quickly after it had bottomed out at \$32 at the end of 2008, climbing to above \$100 per barrel at the end of 2010; the average oil price maintained its rate above \$100 for more than three years until it suddenly crashed below \$60 in the fourth quarter of 2014, which guaranteed positive growth in TFP during the same period. Petroleum TFP

Figure 3: Growth Rate of Total Factor Productivity over Time



lost its momentum in recent years as the average oil price dropped to about \$50 in 2015 and further declined to about \$40 in 2016.

5.3 Effects of Ownership and Business Portfolio

The essential issue this paper seeks to solve regards the effects of ownership and business portfolio on input elasticities and TFP among petroleum enterprises. How did segment-wide and product-wide portfolios, along with ownership of the company, reshape the production frontier and productivity in the global oil and gas industry over time? Table 2 provides results of the correlation analysis, where Pearson’s correlations between each dependent variable and each independent variable are given. Moreover, Table 2 also reports the statistical significance level of these correlation coefficients.

According to Table 2, most of the correlations are significant, indicating that both input elasticities and productivity are correlated with ownership and business portfolio. However, this analysis is a simple bivariate analysis without controlling for other variables. To further explore the actual effect of ownership and business portfolio, multivariate analysis in Eqs. (5) and (6) is employed. Using a likelihood ratio test (LR test), this paper finds that both ownership and business portfolio should be included in all five regressions. Therefore, Eqs. (5) and (6) are appropriate and the regression results are reported in Table 3. The first four columns in Table 3 present the elasticity

Table 2: Correlation Test Results

	$\log(\beta_{it}^1)$ Labor (1)	$\log(\beta_{it}^2)$ Oil Reserves (2)	$\log(\beta_{it}^3)$ Gas Reserves (3)	$\log(\beta_{it}^4)$ Refinery (4)	$\log(TFP_{it})$ TFP (5)
<i>noc</i>	0.219*** (0.000)	-0.210*** (0.000)	-0.161*** (0.000)	0.059 (0.134)	-0.419*** (0.000)
<i>gas</i>	0.396*** (0.000)	-0.351*** (0.000)	-0.329*** (0.000)	0.155*** (0.000)	0.132*** (0.001)
<i>down</i>	0.410*** (0.000)	-0.409*** (0.000)	-0.324*** (0.000)	0.117** (0.003)	-0.030 (0.446)

Notes: The first four columns present the correlations between each of the four input elasticities and the three variables of ownership and business portfolio, whereas the last column reports the correlations between productivity and the three variables of interest.

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 3: Regression Results

	Elasticity Decompositions				TFP Decomposition
	$\log(\beta_{il}^1)$ Labor	$\log(\beta_{il}^2)$ Oil Reserves	$\log(\beta_{il}^3)$ Gas Reserves	$\log(\beta_{il}^4)$ Refinery	$\log(TFP_{it})$ TFP
	(1)	(2)	(3)	(4)	(5)
<i>noc</i>	0.054*** (0.000)	-0.290*** (0.013)	-0.087*** (0.009)	0.038** (0.016)	-0.592*** (0.053)
<i>gas</i>	0.208*** (0.001)	-1.092*** (0.025)	-0.362*** (0.017)	0.179*** (0.031)	0.224** (0.101)
<i>down</i>	0.161*** (0.001)	-0.888*** (0.019)	-0.276*** (0.013)	0.122*** (0.024)	0.140* (0.079)
year effects	controlled	controlled	controlled	controlled	controlled
intercept	-0.564*** (0.001)	-0.764*** (0.023)	-2.372*** (0.016)	-2.991*** (0.028)	5.073*** (0.092)
R^2	0.99	0.89	0.65	0.62	0.61

Note: The first four columns present the elasticity determination for the four inputs in Eq. (5), whereas the last column reports the TFP decomposition in Eq. (6).

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

determination for the four inputs in Eq. (5), whereas the last column in Table 3 reports the TFP decomposition in Eq. (6). Most of the signs and the significances of the correlations reported in Table 2 are confirmed by the results in Table 3.

The first two rows in Table 3 show the impacts of ownership on input elasticities and TFP. Although the comparison between NOCs and IOCs is the focus in previous studies, no literature has analyzed the impact of ownership on the input–output relation since the elasticities were assumed to be fixed in the classic model. On the one hand, the result indicates that NOCs have a significantly higher coefficient of labor than the one for IOCs, which is consistent with the fact that NOCs rely more on their workforce. On the other hand, the coefficients of oil and gas reserves for IOCs are greater than those for NOCs when other things are kept equal, which is compatible with the reality that IOCs have had much more Exploration and Production (E&P) spending and better utilization of their reserves in the last two decades.³ Furthermore, NOCs also have a greater coefficient of refining capacity than IOCs. In terms of the effect on TFP, this paper finds that, *ceteris paribus*, the TFP of NOCs is 59.2% lower than the TFP of IOCs, which is both economically and statistically significant.

The product-wide portfolio for petroleum companies is measured by the revenue share of gas products, *gas*. This paper finds that increasing the ratio of natural gas in a company's portfolio can significantly increase the elasticities of labor and refining capacity but can decrease the elasticity of oil and gas reserves. Moreover, the more negative effect on the coefficient of oil reserves than the one of gas reserves is as expected since the oil reserves are less important when a company chooses to produce more natural gas and less crude oil. Finally, a one percentage point incline in the gas ratio can significantly increase TFP by 0.224%, which implies that natural gas production is more productive than crude oil. This finding is consistent with the fact that gas is less costly than oil production in recent years, as reported in Rystad Energy's UCube database.

The petroleum industry can be divided into upstream activities (exploration and production) and downstream activities (refinery and marketing). The result shows that workforce is more important and productive in downstream activities since an incline in revenue share of downstream activities can increase labor elasticity, which contributes more to output given a fixed number of

3. Since 1999, two-thirds of the newly discovered oil and gas reserves belong to five IOCs, which include British Petroleum (BP), ExxonMobil, Shell, Chevron, and Total.

employees. The findings that companies more reliant on downstream activities have lower elasticity of oil and gas reserves but higher elasticity of refining capacity are reasonable as the formers are the capital used in the upstream, while the latter is the capital employed in the downstream. Furthermore, downstream activities, on average, are more productive than upstream activities. The estimation result shows that a one percentage point incline in downstream ratio in a company's portfolio leads to a 0.14% increase in TFP. However, the difference between upstream and downstream activities is smaller than the difference between oil and gas production.

5.4 Discussions and Policy Implications

This subsection compares our key findings with those in the literature. In terms of ownership, this paper finds that NOCs are about 60% less efficient than IOCs. This efficiency difference between NOCs and IOCs is consistent with the estimates in Al-Obaidan and Scully (1992), Eller, Hartley, and Medlock (2011), and Ike and Lee (2014), but larger than the estimates in Wolf (2009) and Ohene-Asare, Turkson, and Afful-Dadzie (2017). In terms of the effect of gas ratio, this paper finds that gas production is more efficient than oil production, which is contrary to the findings in Gong (2018b) that assumes a simple and constant Cobb-Douglas production frontier. Since TFP is a Solow residual, the effect of gas ratio on input elasticity will be mistakenly included in the effect of gas ratio on productivity when the constant input elasticity method is adopted. In Table 3, the effects of gas ratio on oil reserves and gas reserves are significantly negative (-1.092 and -0.362), whereas the effects of gas ratio on labor and refining capacity are significantly positive (0.208 and 0.179). The magnitudes of the two negative effects are much larger than the magnitudes of the two positive effects, which is very likely to result in downward bias when all four effects are mistakenly included in the effect of gas ratio on productivity. Such a bias can decrease the positive impact of gas ratio on productivity, and may even lead to a negative impact as derived in Gong (2018b). More detailed illustration can be found in Appendix E, which is a good example to highlight the virtue of the varying coefficient model. For the comparison between upstream and downstream activities, this paper finds that the later segment outperforms the former segment. If a conventional fixed coefficient model, rather than the varying coefficient model, is used, this paper expects that the estimated effect of downstream ratio on productivity will be smaller or even negative, for the same reasons provided in Appendix E.

Based on the empirical findings, this paper gives three policy implications. Firstly, the government should be very careful in putting pressure on NOCs to fulfill their non-commercial objectives. Although price subsidies and overemployment are the social responsibility of NOCs, the economic objective is still important as many IOCs and NOCs compete directly in domestic and international markets. This paper finds the TFP of NOCs to be about 60% lower than the TFP of IOCs, which should be considered by the authorities. Moreover, NOCs are also found to be less efficient in converting their oil and gas reserves into production. This paper suggests that NOCs should invest more on exploration and production.

Secondly, privatizing NOCs may be a good option. Ike and Lee (2014) point out that the private sector has advantages in new ideas, good management practices and assuming responsibilities, all of which can help improve firm-level efficiency and revenue. On the contrary, if more resources are controlled by NOCs, we may suffer from a shortage in oil and gas supply and higher energy prices, which is also warned in Eller, Hartley, and Medlock (2011) and Hartley and Medlock III (2013). More importantly, since NOCs have a tradition of providing price subsidies in the

domestic market, the price gap across countries will be further extended. This distortion in energy price will generate a significant impact on global resource allocation.

Thirdly, oil companies can hire more oilfield service companies to explore and produce oil and gas. The upstream segment is a complex process that involves specialized technology in each step of exploration and production. Some of these activities can be more efficiently accomplished by specialized oilfield service companies. The empirical result shows that downstream activities are more efficient than upstream activities for petroleum enterprises. Therefore, these petroleum enterprises can utilize more E&P spending to hire professional oilfield service companies in upstream activities, rather than extract oil and gas by themselves.

6. CONCLUSION

The Shale Revolution that has taken place over the past two decades has overwhelmingly reshaped the oil and gas industry. This paper aims to better capture such structural changes using a varying coefficient production model, where the shape of the production function, along with the magnitude of TFP, can vary according to ownership and business portfolio as well as over time. Compared with the classic model, which assumes a fixed input–output relation, this model is more suitable for the petroleum industry in the context of the Shale Revolution, along with the two waves of oil crises that also reshaped the market dramatically.

Using panel data from 110 of the largest oil and gas companies covering the time period of 2009–2016, the empirical results show that labor elasticity was increasing, elasticities of oil and gas reserves were decreasing, and refinery elasticity was relatively flat. The changes in the productivity of these petroleum giants follows the business cycle, which achieved tremendous growth after the 2007–2009 financial crisis but lost momentum following the oil price crash since the end of 2014. Furthermore, segment-wide and product-wide portfolios, as well as ownership, significantly affect the input elasticities and therefore the shape of the production frontier, which in turn confirms the importance and necessity of employing the varying coefficient model, rather than the classic model. Finally, IOCs, gas production and downstream activities, *ceteris paribus*, are more productive than NOCs, oil production and upstream activities, respectively.

This paper establishes a new model, which enables us to not only capture the effects that ownership and business portfolio have on the shape of the production function, which is overlooked in the literature, but also accurately evaluate their effects on productivity, which is estimated with bias in existing studies. The varying coefficient model, compared with the conventional fixed coefficient model, can better analyze the fundamental transition of the petroleum industry over the past decade. Moreover, the empirical result provides evidence that the fixed coefficient assumption is invalid, which further confirms the necessity to use the new model. These are the major advantages of this paper.

An important limitation of our model is that we ignore the effects of market factors due to a lack of data on the domestic and international market. Future studies can take domestic market endowment factors, such as degree of industrialization and market size, into consideration, since the market-entry/exit decision and productivity of a petroleum enterprise could be constrained by these variables. For international companies with a large proportion of their business abroad, market-entry/exit decision and productivity are also significantly affected by the factors in those overseas markets. When the proportions of domestic and overseas business are available, we can better capture and identify their effects on oil and gas production. Moreover, replacing revenue with physical outputs in the production frontier model is a good robustness check when accurate physical output data is available. Taking environmental and geopolitical concern into consideration is also important

when studying petroleum industry (Barbi, Ferreira, and Guo, 2016; Munro, 2018). Finally, future studies that pay attention to the changes in the industry before and after the 2007–2009 financial crisis are of great importance, if a database that goes back to years before 2008 can be built.

ACKNOWLEDGMENTS

I thank Dr. Robin Sickles, Dr. Peter Hartley, Dr. Kenneth Medlock, Dr. Tang Xun, Dr. Gustavo Grullon, Dr. Steven Chu, Dr. Roger Myerson, Dr. Hulya Eraslan, Dr. Marc Santugini, and participants at presentations at the 12th Asia-Pacific Productivity Conference. I am grateful to Energy Intelligence for providing their data, Weatherford International for giving me the opportunity to closely observe firms in the oilfield market, and Baker Institute for Public Policy and Rice University for supporting my research on energy economics. Whatever shortcomings that remain in the paper are solely my responsibility. The work was supported by the Research Program for Humanities and Social Science Granted by Chinese Ministry of Education [18YJC790034], the National Natural Science Foundation of China [71703149], the Fundamental Research Funds for the Central Universities [526013*172220182], and the Teaching and Research Development Program for Faculties in Arts and Humanities [126000-541903/026] at Zhejiang University.

REFERENCE

- Akerberg, D.A., K. Caves, and G. Frazer. (2015). “Identification Properties of Recent Production Function Estimators.” *Econometrica* 83(6): 2411–2451. <https://doi.org/10.3982/ECTA13408>.
- Ahmad, I., S. Leelahanon, and Q. Li. (2005). “Efficient Estimation of a Semiparametric Partially Linear Varying Coefficient Model.” *Annals of Statistics* 33(1): 258–283. <https://doi.org/10.1214/009053604000000931>.
- Aigner, D., C.A. Lovell, and P. Schmidt. (1977). “Formulation and Estimation of Stochastic Frontier Production Function Models.” *Journal of Econometrics* 6(1): 21–37. [https://doi.org/10.1016/0304-4076\(77\)90052-5](https://doi.org/10.1016/0304-4076(77)90052-5).
- Al-Obaidan, A.M., and G.W. Scully. (1992). “Efficiency Differences between Private and State-Owned Enterprises in the International Petroleum Industry.” *Applied Economics* 24(2): 237–246. <https://doi.org/10.1080/00036849200000122>.
- Amsler, C., A. Prokhorov, and P. Schmidt. (2016). “Endogeneity in Stochastic Frontier Models.” *Journal of Econometrics* 190(2): 280–288. <https://doi.org/10.1016/j.jeconom.2015.06.013>.
- Barbi, F., L.d.C. Ferreira, and S. Guo. (2016). “Climate Change Challenges and China’s Response: Mitigation and Governance.” *Journal of Chinese Governance* 1(2): 324–339. <https://doi.org/10.1080/23812346.2016.1181598>.
- Battese, G.E., and T.J. Coelli. (1992). *Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India*. Springer Netherlands. https://doi.org/10.1007/978-94-017-1923-0_10.
- Boyd, G.A. (2008). “Estimating Plant Level Energy Efficiency with a Stochastic Frontier.” *The Energy Journal* 23–43. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol29-No2-2>.
- Briggs, A.H., C.Z. Mooney, and D.E. Wonderling. (1999). “Constructing Confidence Intervals for Cost-Effectiveness Ratios: An Evaluation of Parametric and Non-Parametric Techniques Using Monte Carlo Simulation.” *Statistics in Medicine* 18(23): 3245–3262. [https://doi.org/10.1002/\(SICI\)1097-0258\(19991215\)18:23<3245::AID-SIM314>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1097-0258(19991215)18:23<3245::AID-SIM314>3.0.CO;2-2).
- Cabrales, S., R. Bautista, and J. Benavides. (2017). “A Model to Assess the Impact of Employment Policy and Subsidized Domestic Fuel Prices on National Oil Companies.” *Energy Economics* 68: 566–578. <https://doi.org/10.1016/j.eneco.2017.10.038>.
- Chen, L., and Y. Fu. (2017). “Dynamics of the Rise of the Privately-Run and Government-Subsidised Think Tank and Its Influence in Talent Policy Making-Case Study of the Zhejiang Institution of Talent Development.” *Journal of Chinese Governance* 2(1): 50–67. <https://doi.org/10.1080/23812346.2017.1286790>.
- Chen, S.-S., and K.-W. Hsu. (2013). “Oil Price Volatility and Bilateral Trade.” *The Energy Journal* 34(1): 207–229. <https://doi.org/10.5547/01956574.34.1.9>.
- Cuñado, J., and F.P. de Gracia. (2003). “Do Oil Price Shocks Matter? Evidence for Some European Countries.” *Energy Economics* 25(2): 137–154. [https://doi.org/10.1016/S0140-9883\(02\)00099-3](https://doi.org/10.1016/S0140-9883(02)00099-3).
- Dimitropoulos, D., and A. Yatchew. (2017). “Is Productivity Growth in Electricity Distribution Negative? An Empirical Analysis Using Ontario Data.” *The Energy Journal* 38(2): 175–200. <https://doi.org/10.5547/01956574.38.2.ddim>.

- Eller, S.L., P.R. Hartley, and K.B. Medlock. (2011). "Empirical Evidence on the Operational Efficiency of National Oil Companies." *Empirical Economics* 40(3): 623–643. <https://doi.org/10.1007/s00181-010-0349-8>.
- Fan, Y., Q. Li, and A. Weersink. (1996). "Semiparametric Estimation of Stochastic Production Frontier Models." *Journal of Business & Economic Statistics* 14(4): 460–468. <https://doi.org/10.1080/07350015.1996.10524675>.
- Farsi, M., A. Fetz, and M. Filippini. (2008). "Economies of Scale and Scope in Multi-Utilities." *The Energy Journal* 38(2): 123–143. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol29-No4-6>.
- Gong, B. 2016. "Efficiency and Productivity Analysis of Multidivisional Firms." Ph.D. Dissertation, Department of Economics, Rice University.
- Gong, B. (2017). "Multi-Dimensional Interactions in the Oilfield Market: A Jackknife Model Averaging Approach of Spatial Productivity Analysis (Forthcoming)." *Energy Economics*. <https://doi.org/10.1016/j.eneco.2017.08.032>.
- Gong, B. (2018a). "The Shale Technical Revolution—Cheer or Fear? Impact Analysis on Efficiency in the Global Oilfield Service Market." *Energy Policy* 112(1): 162–172. <https://doi.org/10.1016/j.enpol.2017.09.054>.
- Gong, B. (2018b). "Different Behaviors in Natural Gas Production between National and Private Oil Companies: Economics-Driven or Environment-Driven?" *Energy Policy* 114(3): 145–152. <https://doi.org/10.1016/j.enpol.2017.12.004>.
- Gong, B. (2018c). "Total-Factor Spillovers, Similarities, and Competitions in the Petroleum Industry." *Energy Economics* 73: 228–238. <https://doi.org/10.1016/j.eneco.2018.04.036>.
- Gong, B. (2018d). "Agricultural Reforms and Production in China Changes in Provincial Production Function and Productivity in 1978–2015." *Journal of Development Economics* 132: 18–31. <https://doi.org/10.1016/j.jdeveco.2017.12.005>. <https://doi.org/10.1016/j.jdeveco.2017.12.005>.
- Gong, B. (2018e). "Like Father Like Son? Revisiting the Role of Parental Education in Estimating Returns to Education in China (Forthcoming)." *Review of Development Economics*. <https://doi.org/10.1111/rode.12538>.
- Gong, B. (2018f). "The Impact of Public Expenditure and International Trade on Agricultural Productivity in China (Forthcoming)." *Emerging Markets Finance and Trade*. <https://doi.org/10.1080/1540496X.2018.1437542>.
- Gong, B., and R.C. Sickles. 2016. *Non-Structural and Structural Models in Productivity Analysis: Study of the British Isles During the 2007–2009 Financial Crisis*. Rice University.
- Gong, B., and R.C. Sickles. 2018. *Resource Allocation in Multi-Divisional Multi-Product Firms*. Rice University, Department of Economics.
- Greene, W.H. (1990). "A Gamma-Distributed Stochastic Frontier Model." *Journal of Econometrics* 46(1): 141–163. [https://doi.org/10.1016/0304-4076\(90\)90052-U](https://doi.org/10.1016/0304-4076(90)90052-U).
- Hartley, P., and K.B. Medlock. (2008). "A Model of the Operation and Development of a National Oil Company." *Energy Economics* 30(5): 2459–2485. <https://doi.org/10.1016/j.eneco.2007.12.017>.
- Hartley, P.R., and K.B. Medlock III. (2013). "Changes in the Operational Efficiency of National Oil Companies." *The Energy Journal* 34(2): 27–57. <https://doi.org/10.5547/01956574.34.2.2>.
- Hastie, T., and R. Tibshirani. (1993). "Varying-Coefficient Models." *Journal of the Royal Statistical Society. Series B (Methodological)* 55(4): 757–796. <https://doi.org/10.1111/j.2517-6161.1993.tb01939.x>.
- Henningsen, A., and S. Kumbhakar. 2009. "Semiparametric Stochastic Frontier Analysis: An Application to Polish Farms During Transition." European Workshop on Efficiency and Productivity Analysis (EWEP) in Pisa, Italy, June.
- Ike, C.B., and H. Lee. (2014). "Measurement of the Efficiency and Productivity of National Oil Companies and Its Determinants." *Geosystem Engineering* 17(1): 1–10. <https://doi.org/10.1080/12269328.2014.887045>.
- Jacoby, H.D., F.M. O'Sullivan, and S. Paltsev. (2012). "The Influence of Shale Gas on Us Energy and Environmental Policy." *Economics of Energy & Environmental Policy* 1(1): 37–52. <https://doi.org/10.5547/2160-5890.1.1.5>.
- Jaunky, V.C., and L. Zhang. (2016). "Convergence of Operational Efficiency in China's Provincial Power Sectors." *The Energy Journal* 37(China Special Issue): 3–27. <https://doi.org/10.5547/01956574.37.S11.vjau>.
- Jiménez-Rodríguez, R., and M. Sánchez. (2005). "Oil Price Shocks and Real Gdp Growth: Empirical Evidence for Some Oecd Countries." *Applied Economics* 37(2): 201–228. <https://doi.org/10.1080/0003684042000281561>.
- Kim, Y.-J. (2013). "A Partial Spline Approach for Semiparametric Estimation of Varying-Coefficient Partially Linear Models." *Computational Statistics & Data Analysis* 62: 181–187. <https://doi.org/10.1016/j.csda.2013.01.006>.
- Kumbhakar, S.C., and G. Lien. (2017). "Yardstick Regulation of Electricity Distribution--Disentangling Short-Run and Long-Run Inefficiencies." *The Energy Journal* 38(5). <https://doi.org/10.5547/01956574.38.5.skum>.
- Kumbhakar, S.C., and K. Sun. 2013. "Estimation of a Flexible Stochastic Cost Frontier Model with Environmental Factors Subject to Economic Constraints." European Economic Association & Econometric Society 2013 Parallel Meetings, Gothenburg, Sweden, 26 - 30 August 2013.
- Lee, K., S. Ni, and R.A. Ratti. (1995). "Oil Shocks and the Macroeconomy: The Role of Price Variability." *The Energy Journal* 16(4): 39–56. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol16-No4-2>.
- Levinsohn, J., and A. Petrin. (2003). "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies* 70(2): 317–341. <https://doi.org/10.1111/1467-937X.00246>.

- Li, R., R. Joyeux, and R.D. Ripple. (2014). "International Natural Gas Market Integration." *The Energy Journal* 35(4): 159–179. <https://doi.org/10.5547/01956574.35.4.7>.
- Llorca, M., J. Baños, J. Somoza, and P. Arbués. (2017). "A Stochastic Frontier Analysis Approach for Estimating Energy Demand and Efficiency in the Transport Sector of Latin America and the Caribbean." *The Energy Journal* 38(5). <https://doi.org/10.5547/01956574.38.5.mllo>.
- Lu, Y., R. Zhang, and L. Zhu. (2008). "Penalized Spline Estimation for Varying-Coefficient Models." *Communications in Statistics-Theory and Methods* 37(14): 2249–2261. <https://doi.org/10.1080/03610920801931887>.
- Meeusen, W., and J. Van den Broeck. (1977). "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error." *International Economic Review* 18(2): 435–444. <https://doi.org/10.2307/2525757>.
- Mirza, F.M., and O. Bergland. (2015). "Market Power in Norwegian Electricity Market: Are the Transmission Bottlenecks Truly Exogenous?" *The Energy Journal* 36(4): 313–330. <https://doi.org/10.5547/01956574.36.4.fmir>.
- Moretz, T. (2018). "An Assessment of China's Ability to Regulate Its Iron and Steel Industries." *Journal of Chinese Governance* 3(1): 101–121. <https://doi.org/10.1080/23812346.2018.1428061>.
- Munro, N. (2018). "Market Justice" in China and Russia." *Journal of Chinese Governance*: 1–24.
- Nepal, R., and J. Foster. (2016). "Testing for Market Integration in the Australian National Electricity Market." *The Energy Journal* 37(4): 215–238. <https://doi.org/10.5547/01956574.37.4.rnep>.
- Ohene-Asare, K., C. Turkson, and A. Aful-Dadzie. (2017). "Multinational Operation, Ownership and Efficiency Differences in the International Oil Industry." *Energy Economics* 68: 303–312. <https://doi.org/10.1016/j.eneco.2017.10.006>.
- Olley, G.S., and A. Pakes. (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64(6): 1263–1297. <https://doi.org/10.2307/2171831>.
- Orea, L., I.C. Álvarez, and T. Jamasb. (2018). "A Spatial Stochastic Frontier Model with Omitted Variables: Electricity Distribution in Norway." *The Energy Journal* 39(3). <https://doi.org/10.5547/01956574.39.3.lore>.
- Qiang, L., and W. Tuohan. (2016). "Social Governance and the Qinghe Experiment." *Journal of Chinese Governance* 1(1): 139–156. <https://doi.org/10.1080/23812346.2016.1138697>.
- Rahman, F. (2018). "Defeating or Delaying the Defaults: Bailout Strategy of the Chinese Government for Its State-Owned Enterprises on Their Bond Payments." *Journal of Chinese Governance* 3(1): 86–100. <https://doi.org/10.1080/23812346.2017.1377806>.
- Rogers, R.P. (2003). "The Effect of the *Energy Policy* and Conservation Act (EPCA) Regulation on Petroleum Product Prices, 1976–1981." *The Energy Journal* 24(2): 63–93. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol24-No2-3>.
- Schmidt, P., and R.C. Sickles. (1984). "Production Frontiers and Panel Data." *Journal of Business and Economic Statistics* 2(4): 367–374. <https://doi.org/10.1080/07350015.1984.10509410>.
- Sickles, R.C. (2005). "Panel Estimators and the Identification of Firm-Specific Efficiency Levels in Parametric, Semiparametric and Nonparametric Settings." *Journal of Econometrics* 126(2): 305–334. <https://doi.org/10.1016/j.jeconom.2004.05.004>.
- Stevenson, R.E. (1980). "Likelihood Functions for Generalized Stochastic Frontier Estimation." *Journal of Econometrics* 13(1): 57–66. [https://doi.org/10.1016/0304-4076\(80\)90042-1](https://doi.org/10.1016/0304-4076(80)90042-1).
- Sueyoshi, T., and M. Goto. (2012). "Data Envelopment Analysis for Environmental Assessment: Comparison between Public and Private Ownership in Petroleum Industry." *European journal of operational research* 216(3): 668–678. <https://doi.org/10.1016/j.ejor.2011.07.046>.
- Sun, K., and S.C. Kumbhakar. (2013). "Semiparametric Smooth-Coefficient Stochastic Frontier Model." *Economics Letters* 120(2): 305–309. <https://doi.org/10.1016/j.econlet.2013.05.001>.
- Thompson, R.G., P. Dharmapala, L.J. Rothenberg, and R.M. Thrall. (1996). "Dea/Ar Efficiency and Profitability of 14 Major Oil Companies in U.S. Exploration and Production." *Computers and Operations Research* 23(4): 357–373. [https://doi.org/10.1016/0305-0548\(95\)00044-5](https://doi.org/10.1016/0305-0548(95)00044-5).
- Wolf, C. (2009). "Does Ownership Matter? The Performance and Efficiency of State Oil Vs. Private Oil (1987–2006)." *Energy Policy* 37(7): 2642–2652. <https://doi.org/10.1016/j.enpol.2009.02.041>.
- Xiao, H. (2016). "Public Financial Management and the Campaign against Extravagant Position-Related Consumption in China." *Journal of Chinese Governance* 1(4): 546–563. <https://doi.org/10.1080/23812346.2016.1241938>.
- Zhang, Q., Z. Xu, T. Feng, and J. Jiao. (2015). "A Dynamic Stochastic Frontier Model to Evaluate Regional Financial Efficiency: Evidence from Chinese County-Level Panel Data." *European Journal of Operational Research* 241(3): 907–916. <https://doi.org/10.1016/j.ejor.2014.09.021>.
- Zhang, R., K. Sun, M.S. Delgado, and S.C. Kumbhakar. (2012). "Productivity in China's High Technology Industry: Regional Heterogeneity and R&D." *Technological Forecasting and Social Change* 79(1): 127–141. <https://doi.org/10.1016/j.techfore.2011.08.005>.
- Zhu, X. (2016). "Dynamics of Central-Local Relations in China's Social Welfare System." *Journal of Chinese Governance* 1(2): 251–268. <https://doi.org/10.1080/23812346.2016.1166903>.

APPENDIX A: ESTIMATION STRATEGY OF THE VARYING COEFFICIENT MODEL

A two-step approach proposed by Fan, Li, and Weersink (1996) is employed to estimate Eq. (4), which is a standard method used to solve semiparametric and nonparametric stochastic frontier models. Suppose μ is the mean of u_{it} so that $\epsilon_{it} = \mu + v_{it} - u_{it}$ has a zero mean. Then Eq. (4) can be transformed to $y_{it} = f'(x_{it}) + \epsilon_{it}$, where $f'(x) = f(x) - \mu$. Firstly, $y_{it} = f'(x_{it}) + \epsilon_{it}$ is regressed using semiparametric or nonparametric methods, which derives the residuals $\hat{\epsilon}_{it}$. Secondly, the residual is then decomposed as $\hat{\epsilon}_{it} = \mu + v_{it} - u_{it}$ using the classic stochastic frontier model. Henningsen and Kumbhakar (2009) employ this approach to study the efficiency levels of Polish farms during transition, and Gong (2018a) adopts this estimation strategy in his applied study on the global oilfield service market.

More specifically, a penalized B-spline approach is used in the first step to estimate $y_{it} = f'(x_{it}) + \epsilon_{it}$ as this method has benign properties, such as strong consistency and asymptotic normality for the varying coefficient model (Lu, Zhang, and Zhu 2008). It is worth noting that a spline-based method, rather than a kernel-based method, is adopted as the latter may suffer from the “curse of dimensionality” when we have four “threshold” variables in $h_k(\theta_{it})$. The spline-based method is utilized in many studies (e.g., Ahmad, Leelahanon, and Li (2005) and Hastie and Tibshirani (1993)) due to its flexibility when multiple smoothing parameters are involved (Kim 2013). However, if we have fewer “threshold” variables, a kernel-based method is an alternative approach (Lee, Ni, and Ratti 1995, Mirza and Bergland 2015, Jaunky and Zhang 2016, Dimitropoulos and Yatchew 2017).

APPENDIX B: ENDOGENEITY PROBLEMS AND ROBUSTNESS CHECKS

The identification in econometric model is challenged by endogeneity problem in many economics studies (Gong 2018e). In production function, this issue exists since some information observed by the petroleum enterprises that helps their production decision-making process turns out to be unobserved by researchers (Akerberg, Caves, and Frazer 2015, Gong and Sickles 2018). One popular approach to deal with this problem is proposed in Olley and Pakes (1996) and then extended in Levinsohn and Petrin (2003), where annual investment or intermediate inputs can be employed to control for the unobserved productivity shocks. However, Akerberg, Caves, and Frazer (2015) warn that the coefficients of the exogenous inputs cannot be identified due to collinearity problems. Another widely used approach with an endogeneity concern is the instrumental variables (IV) estimation. Amsler, Prokhorov, and Schmidt (2016) introduce how to use the IV method in stochastic frontier analysis. Compared with Olley and Pakes (1996) and Levinsohn and Petrin (2003), the IV method can also test the endogeneity of each input prior to the bias correction. This paper adopts the IV approach because of this advantage as well as the lack of data on investment or intermediate inputs. Econometrically, a control function method introduced in Amsler, Prokhorov, and Schmidt (2016) is first used to test the endogeneity of the inputs, where lagged values of inputs are utilized as instruments (Levinsohn and Petrin 2003, Gong 2017). If any input is found to be endogenous, a Corrected Two-Stage Least Square (C2SLS) suggested in Amsler, Prokhorov, and Schmidt (2016) can be utilized to correct the bias.

Zhang et al. (2015) develop a dynamic stochastic frontier model, rather than a static model, to study China’s financial efficiency, as financial institutions may adjust their financial strategies over time, especially during a financial crisis when the environment changes rapidly. For the sample period that is studied in this paper (2009–2016), the petroleum industry also experienced a

significant transformation. Therefore, this paper employs the dynamic stochastic frontier model proposed in Zhang et al. (2015) to check the robustness of our varying coefficient model in Eq. (4). Mathematically, one-year delay outputs and inputs are included in Eq. (4) as independent variables. Finally, this paper follows Gong (2018d) to use correlation and regression to check the robustness of the input elasticities and productivity estimated by the varying coefficient model and the dynamic stochastic model.

In the determination equations of input elasticities and TFP in Eqs. (5) and (6), endogeneity may also be problematic due to simultaneity bias. Causality is a concern since some explanatory variables may conversely be simultaneously affected by the explained variables (Gong 2018f). For instance, companies with greater elasticity of refining capacity (β_{it}^4) are more likely to invest in downstream activities and therefore may increase the share of this segment ($down_{it}$). Analogously, companies with greater elasticity of gas reserves (β_{it}^3) are more likely to produce more gas-related products and thus may increase the share of gas products (gas_{it}). This paper uses lagged values of all independent variables in Eqs. (5) and (6) to correct the potential simultaneity bias, which can also be viewed as a robustness check.

Regarding the potential endogeneity problem of the ownership, this paper uses Propensity Score Matching (PSM) to yield robust results in comparison with the main findings. First, this paper estimates the propensity score by running a Logit regression where the dependent variable is the dummy variable of NOCs. For the matching to provide a causal estimate in the end, this paper includes all four inputs as well as the prices for oil and gas as covariates, since they are related to both the ownership and output. This paper uses nearest neighbor propensity score matching to find pairs of observations that receive very similar propensity scores, but one is an NOC and the other is an IOC. Finally, the treatment effects can be estimated by running the original decomposition model in Eqs. (5) and (6) on the matched sample of NOCs and IOCs.

APPENDIX C: MORE DETAILED INFORMATION ABOUT THE COMPANIES IN THE DATASET

Energy Intelligence's "Top 100: Global NOC & IOC Rankings" publishes data of the largest one hundred petroleum enterprises on an annual basis. The company list varies over time, as some companies entered, whereas others dropped out of the Top 100 rankings. This paper removes observations with incomplete input and output data, resulting in unbalanced panel data from 110 companies covering the time period of 2009–2016 with a total of 639 observations. The distribution of observations by year is given in Table C1. From 2009, there are 87 firms covered in our dataset, 37 of which are NOCs, while 50 are IOCs. From 2016, there are 18 NOCs and 47 IOCs in our dataset. Table C1 shows that the ratio of IOCs in the Top 100 rankings is decreasing over time, indicating the slower development of NOCs compared with IOCs during the sample period. Table C2 lists the names of the 110 petroleum enterprises included in our dataset, along with their ownership information.

Table C1: Distribution of Observations by Year

Year	Number of Companies	Number of NOCs	Number of IOCs
2009	87	37	50
2010	90	37	53
2011	91	39	52
2012	81	26	55
2013	79	24	55
2014	74	17	57
2015	72	19	53
2016	65	18	47

Table C2: Company Name List

40 National Oil Companies:				
CNOOC	CNPC	CPC	Ecopetrol	EGPC
Gazprom	INOC	Kazmunaigas	KNOOC	KPC
Libya NOC	Mubadala	NNPC	Oil India	ONGC
PDO	PDVSA	PEMEX	Pertamina	Petrobras
Petroecuador	PETRONAS	Petrovietnam	PGNiG	PTT
QP	Rosneft	Saudi Aramco	Sinopec	Socar
Sonangol	Sonatrach	SPC	Statoil	Taqa
Turkmengas	Turkmenneft	Uzbekneftegas	YPF	YPFB
70 International Oil Companies:				
Anadarko	Antero Resources	Apache	Bashneft	BG
BHP Billiton	BP	Cabot Oil & Gas Corp.	Cenovus	Centrica
Chesapeake	Chevron	CNR	Concho Resources	ConocoPhillips
Consol Energy	Continental Resources	Cosmo	DEA	Devon Energy
EnCana	Eni	EOG	EP Energy	EQT
ExxonMobil	Freeport McMoRan	Hess	Husky Energy	Idemitsu
Imperial	Impex	IOCL	IPIC	JX Holdings
Linn Energy	Lukoil	Maersk Oil	Marathon Oil	Mitsui
Mol	Murphy Oil	Newfield	Nexen	Noble Energy
Novatek	Occidental	OMV	Pioneer	QEP Resources
Range Resources	Reliance	Repsol	Rice Energy	Russneft
Santos	Sasol	Shell	SK Energy	Southwestern Energy
Suncor	Surgutneftegas	Talisman	Tatneft	Total
Ultra	Whiting Petroleum	Wintershall	Woodside	WPX Energy

APPENDIX D: ROBUSTNESS CHECK OF INPUT ELASTICITY AND TFP DECOMPOSITION

First, this paper denotes the input elasticities and TFP estimated by the main varying coefficient model as β_{vsfa}^1 , β_{vsfa}^2 , β_{vsfa}^3 , β_{vsfa}^4 , and TFP_{vsfa} , whereas the input elasticities and TFP are estimated by the dynamic stochastic frontier model as β_{dsfa}^1 , β_{dsfa}^2 , β_{dsfa}^3 , β_{dsfa}^4 , and TFP_{dsfa} . In order to check the robustness of the production function and productivity, Table D1 presents the estimation results of five regressions, where each of the five estimates from the varying coefficient model is the dependent variable and the corresponding estimate from the dynamic stochastic frontier model is the independent variable. The coefficients of the independent variable in all regressions are all significant, which confirms the robustness of the estimates in the varying coefficient model. Moreover, Table D1 also calculates the correlation between the dependent and independent variables. All the correlation coefficients in the table are above 0.8, which implies a strong uphill (positive) linear relationship across the estimates derived in the two models. To summarize, the robustness of the input elasticities and TFP estimated by the main varying coefficient model is confirmed.

Table D1: Correlations of the Estimates across Models

	β_{vsfa}^1 Labor (1)	β_{vsfa}^2 Oil Reserves (2)	β_{vsfa}^3 Gas Reserves (3)	β_{vsfa}^4 Refinery (4)	TFP_{vsfa} TFP (5)
X_{dsfa}	1.043*** (0.000)	1.239*** (0.000)	0.909*** (0.000)	0.771*** (0.02)	0.886*** (0.019)
intercept	0.222*** (0.000)	0.044*** (0.000)	-0.185*** (0.000)	0.093*** (0.000)	-0.038 (0.028)
correlation	0.995	0.999	0.918	0.826	0.903

Note: Standard error in parentheses; The independent variable, X_{dsfa} , is β_{dsfa}^1 , β_{dsfa}^2 , β_{dsfa}^3 , β_{dsfa}^4 , and TFP_{dsfa} in Columns (1)–(5), respectively.

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

This paper also checks if the effects of ownership and business portfolio are robust with endogeneity concern. On the one hand, Table D2 shows the estimation results of input elasticity and TFP decomposition in Eqs. (5) and (6), where the independent variables are replaced with their lagged values. On the other hand, Table D3 reports the estimation results of the nearest neighbor propensity score matching model. The values in Tables D2 and D3 are comparable and fairly robust with those in Table 3.

Table D2: Regression Results of the Delay Independent Variable Model

	Elasticity Decompositions				TFP Decomposition
	$\log(\beta_{it}^1)$ Labor (1)	$\log(\beta_{it}^2)$ Oil Reserves (2)	$\log(\beta_{it}^3)$ Gas Reserves (3)	$\log(\beta_{it}^4)$ Refinery (4)	$\log(TFP_{it})$ TFP (5)
<i>noc</i>	0.056*** (0.001)	-0.326*** (0.016)	-0.081*** (0.009)	0.014 (0.014)	-0.620*** (0.059)
<i>gas</i>	0.200*** (0.003)	-1.138*** (0.032)	-0.300*** (0.018)	0.070*** (0.027)	0.284** (0.113)
<i>down</i>	0.155*** (0.002)	-0.926*** (0.024)	-0.225*** (0.014)	0.036* (0.021)	0.168* (0.087)
year effects	controlled	controlled	controlled	controlled	controlled
intercept	-0.534*** (0.002)	-0.802*** (0.027)	-2.503*** (0.016)	-2.814*** (0.023)	5.234*** (0.097)
R^2	0.96	0.87	0.57	0.61	0.62

Note: Standard error in parentheses; The first four columns present the elasticity determination for the four inputs in Eq. (5), whereas the last column reports the TFP decomposition in Eq. (6).

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table D3: Regression Results of the Propensity Score Matching Model

	Elasticity Decompositions				TFP Decomposition
	$\log(\beta_{it}^1)$ Labor (1)	$\log(\beta_{it}^2)$ Oil Reserves (2)	$\log(\beta_{it}^3)$ Gas Reserves (3)	$\log(\beta_{it}^4)$ Refinery (4)	$\log(TFP_{it})$ TFP (5)
<i>noc</i>	0.053*** (0.000)	-0.303*** (0.016)	-0.087*** (0.010)	0.034 (0.015)	-0.520*** (0.069)
<i>gas</i>	0.207*** (0.001)	-1.128*** (0.032)	-0.351*** (0.019)	0.123*** (0.030)	0.304** (0.137)
<i>down</i>	0.160*** (0.001)	-0.900*** (0.028)	-0.284*** (0.017)	0.122*** (0.027)	0.248** (0.122)
year effects	controlled	controlled	controlled	controlled	controlled
intercept	-0.562*** (.001)	-0.704*** (0.031)	-2.409*** (0.018)	-2.896*** (0.030)	4.923*** (0.135)
R^2	0.99	0.88	0.64	0.63	0.62

Note: Standard error in parentheses; The first four columns present the elasticity determination for the four inputs in Eq. (5), whereas the last column reports the TFP decomposition in Eq. (6).

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

APPENDIX E: MORE DISCUSSION ON THE EFFECT OF GAS RATIO ON PRODUCTIVITY

Oil production is found to be more efficient than natural gas in Gong (2018b), which is contrary to the findings in this paper, namely that oil production is less productive than natural gas. There are two reasons for the different conclusions.

First, the methodologies used in these two studies are different. Gong (2018b) assumes a simple and constant Cobb-Douglas production frontier. However, in this paper, we have emphasized and illustrated that the input elasticities may be time-variant and firm-variant, which is the motivation to use a varying-coefficient model. Moreover, the empirical result also provides evidence that the input elasticities are affected by the gas ratio. Since TFP is a Solow residual, the effect of gas ratio on input elasticity will be mistakenly included in the effect of gas ratio on productivity when the constant input elasticity method is adopted. In Table 3, the effects of gas ratio on oil reserves and gas reserves are significantly negative (-1.092 and -0.362), whereas the effects of gas ratio on labor and refining capacity are significantly positive (0.208 and 0.179). However, the magnitudes of the two negative effects are much larger than the magnitudes of the two positive effects. Therefore, it is very likely to result in downward bias when all four effects are mistakenly included in the effect of gas ratio on productivity. Such a bias can decrease the positive impact of gas ratio on productivity, and may even lead to a negative impact as derived in Gong (2018b).

Second, the datasets used in these two studies are different. Although both use data from Energy Intelligence's "Top 100: Global NOC & IOC Rankings", Gong (2018b) uses a balanced panel data of 54 companies from 2009–2015, whereas this paper uses an unbalanced panel data from 110 companies from 2009–2016. The sample size is almost doubled in this paper. The MATLAB code provided by the author of CSS and KSS can only be applied on balanced data, which is the major reason to use balanced data in Gong (2018b). Since the production frontier has a simple C-D formation, it does not encounter the problem of degree of freedom to use the smaller sample size in Gong (2018b). In this paper, the R code for the Battese and Coelli estimator allows the unitization of unbalanced data. Moreover, the more complicated semiparametric varying coefficient model has a higher requirement of degree of freedom, which also motivates us to use unbalanced data with a larger sample size. More importantly, many large companies are removed from the balanced data, such as the largest oil company, Aramco, from Saudi Arabia, which makes it less representative of the global petroleum industry. To summarize, the dataset in Gong (2018b) is limited by the model and code, whereas the model and code in this paper allows and requires the utilization of unbalanced data, which makes it a more representative analysis of global oil and gas companies. Finally, adding the data from year 2016 may also expand the advantage of gas production, as oil prices dropped more than gas prices in recent years.