Energy Policy 112 (2018) 162–172

Contents lists available at ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/enpol

# The shale technical revolution – cheer or fear? Impact analysis on efficiency in the global oilfield service market☆

Binlei Gong[1]

*Department of Agricultural Economics and Management, School of Public Affairs, China Academy for Rural Development (CARD), Zhejiang University, 688 Yuhangtang Road, Hangzhou 310058, PR China*

ARTICLE INFO

*JEL Classification:*
D24 L71 C23 C5

*Keywords:*
Stochastic frontier analysis
Semiparametric model
Global oilfield service market
Multidivisional firms
Shale technical revolution

ABSTRACT

The shale technical revolution has reshaped the oil and gas industry dramatically but also controversially as it affects existing energy policies as well. Many related policies, such as the fracking tax in the U.S. and the shale subsidies policy in China, depend heavily on whether or not the innovation is commercially successful. This paper develops a two-step approach to evaluate the effect of the revolution on efficiency in the global oilfield service (OFS) market, which can be divided into five segments. In the first step, a new semiparametric model is introduced to evaluate firm-level technical efficiencies assuming segment-specific production functions for each of the five segments. In the second step, this study tests if companies acquiring directional drilling (DD) and/or hydraulic fracturing (HF) techniques can maintain efficiency. The empirical results show that practicing just one of the techniques will decrease efficiency. However, combining the two can produce significant spillover effects and improve efficiency. Therefore, innovation and integration are both crucial for the OFS market. Some policy implications are also discussed.

## 1. Introduction

The oilfield service (OFS) market, or oil and gas service industry, is a complex process that involves specialized technology at each step of the oil and gas supply chain. Companies in the OFS market provide the infrastructure, equipment, intellectual property, and services needed to explore for and extract crude oil and natural gas. Therefore, this market is the upstream of the petroleum industry. The global OFS market has a total market capitalization of over $4 trillion, generating total revenues over $400 billion in 2014.[2]

The shale revolution, which benefited mainly from new technologies in hydraulic fracturing and directional drilling, has resulted in a 10% compound annual growth rate (CAGR) for the OFS market over the past decade. As conventional oil and gas resources are now being exhausted, oil and gas companies are currently paying more attention to unconventional oil and gas, offshore production, and aging reservoirs to maintain a steady supply. Therefore, the revolution is also called an unconventional revolution.

Hydraulic Fracturing (HF) is a well stimulation technique in which rock is fractured by a pressurized liquid. The process involves the high-pressure injection of "fracking fluid" (primarily water containing sand or other chemical additives) into a wellbore to create cracks in the deep-rock formations through which natural gas, petroleum, and brine will flow more freely. Directional Drilling (DD) is the practice of drilling non-vertical wells, and it includes the popular horizontal drilling. This technology can hit some targets that cannot be reached by vertical drilling and can drain a broad area from a single drilling pad. The combining of two technologies, HF and DD, has led to the shale revolution. Some rock units that were unproductive when drilled vertically can become fantastic producers of oil and/or gas. The magic of converting worthless shales into productive reservoir rocks occurs in many locations, such as the Barnett Shale of Texas, the Fayetteville Shale of Arkansas, the Marcellus Shale of the Appalachian Basin, the Bakken Formation of North Dakota, and the Haynesville Shale of Louisiana and Texas. Fig. 1 illustrates the hydraulic fracturing and directional drilling activities.

The Oilfield Market Report (OMR) by Spears divides the OFS industry into five macro segments: 1) exploration, 2) drilling, 3) completion, 4) production, and 5) capital equipment, downhole tools and offshore services (capital equipment, hereafter). OMR reports segment-level revenue for the 114 public firms in the field, where 68 firms are single-division and 56 firms are multidivisional.[3] These five macro

☆ I am grateful to Spears and Associates, Inc. for providing their data and Weatherford International for giving me the opportunity to closely observe firms in the oilfield market.
*E-mail address:* gongbinlei@zju.edu.cn.
[1] http://person.zju.edu.cn/en/gbl
[2] Data from 2015 Oilfield Market Report (OMR) by Spears.
[3] 28 firms do business in two segments, 10 firms are active in three segments, seven firms have footprints in four segments, and only one firm covers all five segments.

http://dx.doi.org/10.1016/j.enpol.2017.09.054
Received 1 August 2016; Received in revised form 11 August 2017; Accepted 26 September 2017
Available online 18 October 2017
0301-4215/ 

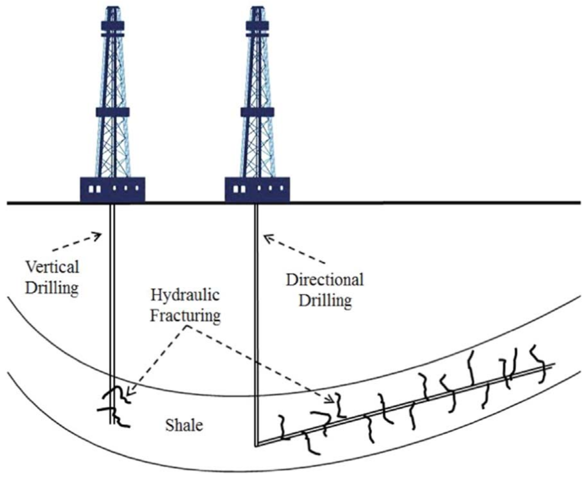


**Fig. 1.** Diagram of hydraulic fracturing and directional drilling.

segments can be further divided into 32 micro-market segments, including Hydraulic Fracturing (under "completion" segment) and Directional Drilling (under "drilling" segment). Based on OMR, the total revenue of the entire OFS market increased by 183% from 2005 to 2015, while the HF segment and the DD segment increased by 395% and 287% respectively during the same period, which implies that these two techniques are leading the development of the entire market.

On the one hand, the shale revolution is generating massive revenues for OFS companies and is producing sufficient energy supplies. Many people cheer the low energy prices and the mitigation of the energy shortage. On the other hand, the new innovations also require huge amounts of investment, such as labor and capital inputs as well as Research and Development (R & D) spending, which is feared for the related financial risk, sustainability, and low input-output ratio.[4] It is difficult to estimate the profitability of the new techniques in practice. Public firms report total inputs and outputs, and possibly segment/division-level outputs, but not segment/division-level inputs. Therefore, it is hard to get cost information for a specific activity or segment to calculate the actual breakeven price for unconventional oil and gas.[5] As a result, whether the innovation is commercially successful is unknown.

But many energy policies depend on whether hydraulic fracturing and directional drilling techniques are earning or losing. For example, what should the tax rate of the fracking tax in the U.S. be? What subsidies should the Chinese government offer to encourage shale resource exploration and extraction? How should the renewable energy policy be adjusted to compete with shale oil and gas?

This paper evaluates whether the innovation has a positive or negative effect on firm-level efficiency using a two-step approach. If firms can maintain or even increase efficiency with hydraulic fracturing and directional drilling programs, it implies that these businesses are at least as competitive and profitable as traditional oil and gas businesses, which will reshape geopolitics and the global energy market.

Managi et al. (2004) and Managi et al. (2006) study the productivity and efficiency of the offshore Gulf of Mexico oil and gas production, using data envelopment analysis (DEA) and stochastic frontier analysis (SFA), respectively. Thompson et al. (1996) analyze the efficiency of 14 major companies in the US oilfield market, using a non-parametric DEA for the period 1980–1991. Non-academic reports on this market are generated by advisory service firms such as Deloitte[6] and Ernst & Young,[7] which predict that the companies will be more efficient in the future. But all the academic and non-academic studies fail to consider the multidivisional structure of the companies and the pure effect of new shale technologies. The oil and gas industry has been better studied (e.g., Wolf, 2009; Eller et al., 2011, and Hartley and Medlock III, 2013) using efficiency analysis. However, their focus is the difference between National Oil Companies and International Oil Companies (i.e., the effect of ownership), rather than the effect of the new shale technologies.

The OFS market is complex and can be divided into multiple segments, each using different technologies and hence following different production functions. In the first step, a semi-varying coefficient stochastic frontier model is introduced to estimate the firm-level efficiency with this multi-segment concern, which standard productivity and efficiency analysis overlooked or chose to ignore. Then, this paper explores whether hydraulic fracturing and directional drilling have a significant effect on a firm's overall technical efficiency.

This study makes three central contributions. Firstly, the semi-parametric production function considers the multi-segment characteristics of a market with multidivisional firms. Secondly, this study focuses on OFS companies, which experience much more volatility than oil and gas companies but are seldom studied.[8] Thirdly, this paper estimates the impact of the shale revolution on efficiency, which provides essential messages to companies for their operational decisions and strategies as well as to governments for their policies and management.

The empirical results show that: 1) the production function is indeed segment-variant, which supports the validation of the multi-segment assumption considered; 2) the output elasticity of labor is consistent, while the output elasticity of capital varies greatly across segments; 3) the average firm-level efficiency for the OFS market is about .4, and the distribution is positive skewed; 4) having a footprint in just a hydraulic fracturing or just a directional drilling business can decrease efficiency, but combining the two generates positive spillover effects; 5) all the findings above are robust when either a Cobb-Douglas or Transcendental Logarithmic production form is adopted.

The remainder of the paper is structured as follows. Section 2 introduces the model. Section 3 provides data descriptions. Empirical results are presented and analyzed in Section 4. Section 5 gives conclusion and policy implications.

## 2. Model

This model includes two steps. Firstly, a stochastic frontier model is used to estimate firm-level aggregated production function as well as efficiency. Secondly, the derived efficiency is regressed on dummy variables of hydraulic fracturing and directional drilling as well as other variables.

### 2.1. Step One: production function and technical efficiency

This subsection develops a partial linear semiparametric varying

[4] The shale revolution is also criticized for climate reasons. The oil and gas from shale is "worse than coal" for the climate since there is greater leakage of methane to the atmosphere in unconventional wells. Moreover, while a high supply of oil and gas decreases energy prices, it discourages the development of renewable energy. However, this paper only focuses on analyzing the economic impact of the revolution from companies' perspectives.

[5] Although some firms report a breakeven price, many of them are wide ranges rather than fixed numbers. The veracity of the reported prices is also suspect since many firms adjust their price ranges frequently and continue to produce when the market price drops far below their reported breakeven prices. Sometimes even the companies themselves find it difficult to calculate the profitability of a certain program/segment because of the joint inputs and spillover effects. Remember, oilfield is a complex process that involves many steps in the energy supply chain.

[6] https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/energy-resources/deloitte-uk-energy-and-resources-outlook-for-oilfield-services.pdf.

[7] http://www.ey.com/Publication/vwLUAssets/EY-review-of-the-UK-oilfield-services-industry-January-2017/$FILE/EY-Review-of-the-UK-oilfield-services-industry-January-2017.pdf.

[8] The productivity and efficiency of oilfield firms is studied much less than oil and gas companies for two reasons: the complex multi-segment characteristics and the lack of segment-level data. This paper uses a very unique dataset to capture the multi-segment characteristics. The empirical result confirms the necessity of doing so.

coefficient stochastic frontier model ("Varying Frontier") to estimate the aggregated production function for multidivisional firms and further predicts firm-level efficiency with multi-segment concern.

#### 2.1.1. Stochastic frontier analysis

Stochastic frontier production function model equals the deterministic frontier production function plus a symmetric random error variable, which is independently and simultaneously proposed by Aigner et al. (1977) and Meeusen and Van den Broeck (1977) in the form

$$\ln Y_i = x_i'\beta + \nu_i - u_i,\ i = 1, \ldots, N,$$

where $Y_i$ is the output of firm $i$, $x_i$ is the vector of inputs typically in logarithms, $\nu_i$ accounts for measurement errors and other sources of non-systematic statistic noise, and $u_i$ is a non-negative random variable representing technical inefficiency (the distance to the frontier).

The stochastic frontier literature in the early 1980s mainly consists of analyses for cross-sectional data. $\nu_i$ is usually assumed to follow a normal distribution that is independent of each $u_i$ while $u_i$ is assumed to follow a variety of distributions including half-normal distribution (Aigner et al., 1977), normal truncated distribution (Stevenson, 1980), and gamma distribution (Greene, 1990). Given panel data, Schmidt and Sickles (1984) proposed panel stochastic frontier model in the form

$$\ln Y_{it} = \alpha + x_{it}'\beta + \nu_{it} - u_i = \alpha_i + x_{it}'\beta + \nu_{it},\ i = 1, \ldots, N,\ t = 1, \ldots, T. \tag{1}$$

Then fixed effects or random effects methods can be used to estimate $\alpha_i$ under different conditions. Other estimators can be found in Cornwell et al. (1990), Kumbhakar (1990), Battese and Coelli (1992), Lee and Schmidt (1993), Kneip et al. (2003), and Sickles (2005).

#### 2.1.2. Weight index and multi-segment concern

One industry may have multiple segments. For example, the global OFS market has five segments including exploration and production. Since the technologies utilized in exploration and production are different, the production function is segment-specific. For a multidivisional firm who has footprints in multiple segments, different production technologies are used to convert inputs to outputs. Therefore, the aggregated production function for this firm is not equal to any of the segment-specific production functions, but a combination of them. This paper attempts to estimate the aggregated production function for multidivisional firms and then derives firm-level efficiency.

Since multidivisional firms use different production technologies, some weight index is needed to estimate the aggregated production function with multi-segment concern. The revenue share by segment/division for firm $i$ at time $t$, $\theta_{it}$, is an eligible weight to capture the heterogeneity in technologies since it indexes the proportion of business using each of the segment-specific techniques. In other words, $\theta_{it}$ measures the frequency of using every segment-specific technology in a multidivisional firm. Imagine a "$M$ inputs – $N$ products/segments – $T$ periods" industry, $\theta_{it} = (\theta_{i1t}, \theta_{i2t}, \ldots, \theta_{iMt})$ where $\theta_{ijt} = \frac{R_{ijt}}{\sum_{J=1}^{N} R_{ijt}}$; $\forall\ j = 1,2,\ldots,N$ and $R_{ijt}$ is the revenue for firm $i$ in segment/division $j$ at time $t$.

As a standard single frontier model, Eq. (1) ignores the heterogeneity in production technologies across segments. The next subsection adds $\theta_{it}$ into the stochastic frontier model to estimate the aggregated production function when firms may have different technologies in their portfolio and use them at different frequencies.

#### 2.1.3. General model

Eq. (1) presents a linear production model and can be generalized to

$$Y_{it} = f(X_{it}; \beta_0) \bullet \exp\ (\tau Z) \bullet \exp\ (\nu_{it}) \bullet \exp\ (-u_i), \tag{2}$$

where $Y_{it}$ is the aggregated output of individual $i$ at time $t$; $X_{it} = (X_{it}^1, X_{it}^2, \ldots X_{it}^M)$ vectors the $M$ types of inputs; $f(X_{it}; \beta_0) \bullet \exp\ (\tau Z)$ is the production frontier over time, where $f(X_{it}; \beta_0)$ is the time-invariant part of the production function, $\beta_0 = (\beta_{01}, \beta_{02}, \ldots \beta_{0M})$ is a vector of technical parameters to be estimated. $Z$ vectors a group of year dummy variables, controls the production frontier change over time and $\tau$ vectors the coefficients of the year dummy variables; $\exp\ (\nu_{it})$ is the stochastic component that describes random shocks affecting the production process, where $\nu_{it}$ is assumed to be normally distributed with a mean of zero and a standard deviation of $\sigma_\nu$, and $TE_i = \exp\ (-u_i)$ denotes the technical efficiency defined as the ratio of observed output to maximum feasible output. $TE_i = 1$ or $u_i = 0$ shows that the $i$-th individual allocates at the production frontier and obtains the maximum feasible output at time $t$, while $TE_i < 1$ or $u_i > 0$ provides a measure of the shortfall of the observed output from the maximum feasible output. This study uses the popular "Error Components Frontier" (Battese and Coelli, 1992) with time-invariant efficiencies to estimate $u_i$ and $TE_i$.

Again, imagine a "$M$ inputs – $N$ products/segments – $T$ periods" industry. The revenue share by segment $\theta_{it}$ is introduced to capture the heterogeneity in total revenue and production technologies. This paper uses $\theta_{it}$ as a weight index and adds it into the aggregated production function:

$$Y_{it} = f(X_{it}; \beta_0, \theta_{it}) \bullet \exp\ (\tau Z) \bullet \exp\ (\nu_{it}) \bullet \exp\ (-u_i) \tag{3}$$

The effect of the business portfolio $\theta_{it}$ can be either dependent or independent with the rest of the production function. If it is independent (i.e., $f(X_{it}; \beta_0, \theta_{it}) = f(X_{it}; \beta_0) \bullet m(\theta_{it})$), then a transfer to the traditional multiproduct stochastic frontier analysis is possible, where one product is a function of all inputs and all other products. Adams et al. (1999) and Liu (2014) use such a canonical regression to check the efficiency of the banking industry with multiple outputs and inputs, where these two papers model $f(X_{it}; \beta_0)$ nonparametrically and parametrically, respectively. This study assumes a Cobb-Douglas form and set a production function where $\theta_{it}$ is independent with $f(X_{it}; \beta_0)$

$$\ln Y_{it} = r(\theta_{it}) + \sum_{k=1}^{M} \beta_k (\ln X_{it}^k) + \tau Z + \nu_{it} - u_i, \tag{4}$$

where $r(\theta_{it})$ is a nonparametric functions of $\theta_{it}$. Although the intercept $r(\theta_{it})$ is a nonparametric functions of $\theta_{it}$ rather than a constant $\alpha$ as in Eq. (1), the core of production function $f(X_{it}; \beta_0)$ is still segment-invariant, which is a strong assumption. Therefore, the frontier estimated by Eq. (4) is called "Single Frontier".

This paper focuses on the other situation, where $\theta_{it}$ can directly affect the production function through their effects on the technical parameters. This model is inspired by the smooth/varying coefficient model (see Hastie and Tibshirani, 1993) and therefore called the varying production frontier, where the coefficients are nonparametric functions of some "threshold" variables ($\theta_{it}$ in this case).

$$Y_{it} = f(X_{it}; \beta_0' = r(\theta_{it})) \bullet \exp\ (\tau Z) \bullet \exp\ (\nu_{it}) \bullet \exp\ (-u_i). \tag{5}$$

Eq. (5) allows the change in aggregated production function when revenue share $\theta$ varies. For example, if a multidivisional firm has major business in segment $A$ and minor business in segment $B$, then the aggregated production function of this firm is likely to be closer to the production function in segment $A$, as this company uses production technology from this segment more frequently. Since using multiple technologies jointly can lead to nonlinear spillover effects caused by shared R & D investment, joint inputs, and so on, we cannot simply take the weighted average of the segment-specific production functions. Hence, a nonparametric function $r(\bullet)$ is used to control the nonlinear combination of technologies.

#### 2.1.4. Semi-varying coefficient model

Productivity and efficiency analysis is dominated by two approaches: the parametric stochastic frontier analysis (SFA) and the nonparametric deterministic data envelopment analysis (DEA). Each method has its own strengths and drawbacks: stochastic frontier analysis is suitable for noisy data, but requires the priori assumption of an explicit functional form; data envelopment analysis does not require specified functional form, but does not allow for statistical noise since no stochastic component is included. In recent years, many new semi-parametric and nonparametric stochastic frontier techniques have been applied to narrow the gap between SFA and DEA. Such development results in new methods to better model the aggregated production function for multidivisional firms who use multiple production technologies.

Fan et al. (1996) propose a semiparametric method that allows for statistical noise and does not need to specify the functional form of the production frontier. Their approach, known as semiparametric frontier analysis, has the form

$$y = f(x) + \epsilon = f(x) + \mu + v - u \tag{6}$$

where $f(x)$ is a semi- or nonparametric production function. Similar to parametric stochastic frontier analysis, $u$ is a non-negative technical inefficiency term and $v$ is a statistical noise term. $\mu$ is a constant that guarantees the expected value of $\epsilon$ equals zero. Therefore, $\epsilon = \mu + v - u$ is the disturbance term with a zero mean.

In practice, the semiparametric model is solved in two steps: in the first step, the semi- or nonparametric regression $y = f(x) + \epsilon$ is run to retrieve the residuals $\hat{\epsilon}$; in the second step, the residual is decomposed as $\hat{\epsilon} = \mu + v - u$ using normal stochastic frontier model where $\hat{\epsilon}$ is the dependent variable and a constant is the only independent variable. Henningsen and Kumbhakar (2009) adopt this approach in their applied study on Polish farms, where they use logarithmic output and input quantities for three reasons: 1) the elasticities are easier to interpret; 2) the observations are more equally distributed when using constant bandwidths; and 3) the usual specification of the production function is easier to adopt.

As Henningsen and Kumbhakar (2009) point out, the nonavailability of software used to prevent applied studies to widely use this approach. This restriction has disappeared in recent years. Take R as an example, the “np” package (Hayfield and Racine, 2008) , the “gam” package (Hastie and Tibshirani, 1990), or the “gamlss” package (Stasinopoulos and Rigby, 2007) can be used in the first step and the “frontier” package (Coelli et al., 2012) can be used in the second step.

This section uses the varying coefficient model (VCM) for the production function $f(x)$ in Eq. (6). Hastie and Tibshirani (1993) first introduce VCM in the form

$$\mathrm{Y} = X_1 r_1(\theta_1) + \ldots + X_p r_p(\theta_p) + \epsilon$$

where $\theta_1, \ldots, \theta_p$ change the coefficients of the $X_1, \ldots, X_p$ through unspecified functions $r_1(\bullet), \ldots, r_p(\bullet)$. The coefficients are nonparametric functions that are not constant, hence the name “varying/smooth coefficient model”. VCM is initially applied to model time-variant coefficient functions in censored data in survival analysis.

In production analysis, environmental factors can only affect the frontier neutrally if treated as independent variables ($X$). Some varying coefficient stochastic frontier analysis treats the environment factors as $\theta_i$ and allows their effect on the frontier to be non-neutrally. R & D Spending is such an environmental factor that is believed to affect the frontier directly. Other examples of such “threshold” variables include tax rate, firm size, firm age, etc. (Kumbhakar and Sun, 2013).

Zhang et al. (2012) develop a varying coefficient production function to study China’s high technology industry, where panel data spanning the period 2000–2007 is used. Sun and Kumbhakar (2013) estimate stochastic production frontier in a Norwegian forest using a cross-section of 3249 active forest owners. Both of these studies use R & D-varying coefficient production functions. However, they use an average production function, not a stochastic frontier model.

This section generates a partial linear semi-varying coefficient stochastic frontier analysis to model the OMR panel data for the OFS market where the revenue distribution by segment $\theta$ can directly affect the technical parameters and the frontier has a Cobb-Douglas (C-D) form.

$$\ln Y_{it} = \alpha + r_1(\theta_{it})\ln L_{it} + r_2(\theta_{it})\ln K_{it} + \tau Z + v_{it} - u_i \tag{7}$$

where $Y_{it}$, $L_{it}$, and $K_{it}$ are the output, number of employees, and capital employed for firm $i$ at time $t$, respectively.

There are two nonparametric approaches to estimate the $r_1(\bullet)$ and $r_2(\bullet)$ in Eq. (7): the kernel-based method (Fan and Huang, 2005; Fan and Li, 2004; Hu, 2014; Su and Ullah, 2006; Sun et al., 2009) and the spline-based method (Ahmad et al., 2005; Hastie and Tibshirani, 1993). Fan and Zhang (2008) think that kernel smoothing methods are more reasonable, as the varying coefficient model is a local linear model, while Kim (2013) argues that spline methods are more attractive for their flexibility to involve multiple smoothing parameters. However, both methods have some disadvantages: the former may suffer from the “curse of dimensionality” and the latter may encounter computational challenges, since the number of spline basis functions can be large.

Since there are five variables in $\theta_{it}$ that will cause a “curse of dimensionality”, this study selects the penalized B-spline approach to estimate the production function. It is assumed that the inefficiency term is time-invariant ($u_i$) so that the Least Square Dummy Variable (LSDV) can be used to derive a fixed effect estimator. Appendix A provides reasons to adopt time-invariant firm-level efficiency in this study. Lu et al. (2008) present results on the strong consistency and asymptotic normality for penalized B-spline estimators of such a varying coefficient model.

This paper uses the two-step approach in Henningsen and Kumbhakar (2009) to estimate Eq. (7): 1) a penalized B-spline method is used to derive consistent coefficients and predict the residuals in the first step; 2) then, a normal stochastic frontier analysis is used where $\hat{\epsilon}$ is the dependent variable and a constant is the only independent variable. This paper also develops a varying coefficient stochastic frontier analysis where the production function has a Transcendental Logarithmic (T-L) form to check the robustness of the varying coefficient model.

#### 2.1.5. Endogeneity problem

Endogeneity is a big issue in production function since input choices are determined by some information that are available by the firms (Ackerberg et al., 2015), which is unavailable by outsiders such as economists. Marschak and Andrews (1944) point out this simultaneity problem is more significant for inputs that adjust rapidly. OFS market is a typical example where the decisions of the companies depend heavily and frequently on exploration and production (E & P) spending from the oil and gas firms and the business cycles. The massive volatility forces companies to divest capital and cut headcount aggressively when the oil price goes down. The potential endogeneity problem in the production function can lead to biased OLS estimates.

One of the solutions to an endogeneity problem is using a set of two-step techniques, advocated by Olley and Pakes (1996). This method uses observed investment to “control” for unobserved productivity shocks (efficiency). Levinsohn and Petrin (2003) extend the idea by using intermediate inputs instead of investment to solve the simultaneity issue as investment is an invalid proxy in many datasets where significant amounts of observations have zero or missing investment. However, as Ackerberg et al. (2015) note, both of the models suffer from the collinearity problems so that the coefficients of the exogenous inputs cannot be identified.

Since the intermediate data is not available in the dataset, this paper uses the most widely used instrumental variables (IV) estimation to solve the endogeneity problem. Recently, Amsler et al. (2015) introduce how to use instrumental variables method in stochastic frontier analysis when the production adopts Cobb-Douglas (C-D) and Transcendental Logarithmic (T-L) form, respectively. On the one hand, they applied a Corrected Two-Stage Least Square (C2SLS) to solve the endogeneity problem in C-D stochastic frontier model.[9] On the other hand, they suggest using the control function method in T-L stochastic frontier model. Moreover, they introduce a method to reduce the number of instrument variables needed.[10]

Following Amsler et al. (2015), this study uses the C2SLS method for the linear C-D production function and the control function method for the nonlinear T-L production. The control function method can also test the exogeneity of the inputs using *t*-tests for the significance of the reduced form residuals (see detail in Amsler et al., 2015). The potential instrument variables include input prices and lagged values of input use (Levinsohn and Petrin, 2003). However, lagged values of inputs are valid instruments only if the lag time is long enough to break the dependence between the input choices and the serially correlated shock. Blundell and Bond (2000) and Guan et al. (2009) both emphasize the input levels lagged at least two periods can be valid instruments. This study uses lag two and lag three input quantities as instruments respectively and get robust results. Therefore, input price and lag two input quantities are selected to be the instruments so that more observations can be pooled into the regression.

### 2.2. Impact of the shale revolution

The shale revolution has mainly occurred in hydraulic fracturing and directional drilling. This paper explores the effect of hydraulic fracturing and directional drilling activities on firm-level efficiency using Eq. (8).

$$TE_i = \beta_0 + \beta_1 \bullet HF_i + \beta_2 \bullet DD_i + \beta_3 \bullet HF_i \bullet DD_i + \beta_4 \bullet R_i + \beta_5 \bullet M_i \tag{8}$$

where $TE_i$ is the technical efficiency for firm $i$, $HF_i$ is the dummy variable of companies who has hydraulic fracturing business, $DD_i$ is the dummy variable of companies who has directional drilling business, $R_i$ refers the revenue for firm $i$ in logarithms to control the size of the company, and $M_i$ is the dummy variable of multidivisional firms who have footprints in multiple segments. Eq. (8) also includes the interaction between $HF_i$ and $DD_i$ to estimate the potential spillover effects of the two technologies.

## 3. Data

This paper applies Eq. (7) in the OFS market using deflated revenue as the output, number of employees[11] as the first input, and capital as the second input. Division-level revenue data from 1997 to 2014 for each of the 114 public firms are collected from the three waves of the OMR (2000, 2011, and 2015) dataset. Appendix B introduces this report, the method employed to combine the three waves of data, and the detailed segmentation of the OFS market.

Data on the annual overall revenue, the number of employees, and total capital for the 114 public firms during same period is collected from Thomson ONE, Bloomberg, and FactSet. The total capital is the accounting capital, which is the sum of equity and long-term debt. This study adjusts the capital data following the unified perpetual inventory method (PIM) in Berlemann and Wesselhöft (2014), which is widely used in productivity analysis. Appendix C explains this data-generating process. Since the labor and capital data are the year-end values, the values at periods $t$ and $t+1$ are averaged to get the average value at period $t+1$.

The overall revenue of a firm is not always equal to the total revenue in the OFS market as reported by the OMR. In some cases, the former may be larger because the company has some business outside the OFS market. On the other hand, the former could be smaller, as the OMR adds the acquired firm's revenue to the mother firm's revenue even in the years before acquisition. The input proportionality assumption suggested by Foster et al. (2008) is used to adjust the labor and capital used in the OFS market. Finally, the Bureau of Labor Statistics publishes the Producer Price Index (PPI) by North American Industry Classification System (NAICS) division. The output price indices deflate the revenue, which can be regarded as output.

Since input prices are selected as instruments to solve endogeneity problem, this paper also collects labor price and capital price: 1) the labor price is the total labor cost divided by the number of employees. Many international firms have compensation cost information, but North American firms have no such information published. This paper sets the labor price of each North American firm to its corresponding NAICS division average. The later information is available in the Labor Productivity and Cost (LPC) Database from the Bureau of Labor Statistics; 2) the capital price is the sum of the depreciation rate and interest rate. i) Thomson ONE, Bloomberg, and FactSet offer depreciation and capital data, which can derive the depreciation rate. ii) the interest rate can be estimated by a capital asset pricing model (CAPM). The needed firm-level beta,[12] the risk-free rate, and the expected market return are all available in Thomson ONE, Bloomberg, and FactSet.

Table 1 summarizes firm-level input and output in the OFS market. Average revenues increased almost three-fold, from \$.97 billion in 1997 to \$2.78 billion in 2014. In the labor market, the average number of employees was around 5750 from 1997 to 2009 and then jumped to 7500 in 2014, which shows a large amount of new employment after the 2007–2009 financial crisis. The average wage almost doubled in the period of 1997–2014. In the capital market, the price of capital is very stable, while the amount of capital in 2014 was more than four times the level it was in 1997. To sum up, both revenues and costs increased dramatically from 1997, which was very likely driven by the shale revolution. Based on the dataset, this paper can build a "2 inputs – 5 products/segments – 16 periods" model[13] for the OFS market.

## 4. Estimation results

This empirical study applies the described models to public firms in the global OFS market. This study estimates production frontiers and firm-level efficiencies in the first step and then predicts the impact of the shale revolution on efficiency in the second step.

### 4.1. Production frontiers

The "Varying Frontier" model in Eq. (7) cannot derive constant coefficients of the production function that are directly comparable with those in the "Single Frontier" model in Eq. (4). Therefore, this study visualizes the varying effects of labor and capital on output in the "Varying Frontier" model. Fig. 2 illustrates the range of such varying

[9] For C2SLS, the first step is to estimate the model by 2SLS and derive the residuals using the instruments. In the second step, these 2SLS residuals are decomposed using the maximum likelihood method, just as in classic stochastic frontier analysis. A somewhat similar two-step procedure is built by Guan et al. (2009).

[10] For example, suppose two inputs, labor and capital, are both endogenous. At least five instruments are needed since all of the two inputs, their square terms, and their interaction are endogenous in the T-L production function. However, under some additional assumptions, consistent estimators can be obtained using only two control functions, not five. This point has been made by some economists, including Blundell and Powell (2004), Terza et al. (2008), and Wooldridge (2010). See detailed discussion in Amsler et al. (2015).

[11] There are contractors (non-fulltime employees) working in the OFS field that are not included in the number of employees. We cannot find the number of these contractors in each firm, let alone their working hours to transfer them into the number of Full Time Equivalent (FTE) employees. Actually, this is a problem happens in many industries, where the companies' financial reports only provide the number of employees rather than the number of FTE employees. The existing studies usually use number of employees to be the proxy of total labor force. In order to reduce the potential bias, this paper uses Cobb-Douglas production function so that the result is not skewed as long as the ratio of employees and non-FTE employees has no large variation across firms.

[12] In finance, the beta of an investment or a company is a measure of the risk arising from exposure to general market movements as opposed to idiosyncratic factors. The market portfolio of all investable assets has a beta of unity.

[13] In regression, the data in 1997 and 1998 are only used as instrument variables to control heterogeneity. Therefore, time $t$ in Eq. (7) refers to 1999–2014 in the empirical study.

**Table 1**
Oilfield market summary statistics.

| Variables | Unit | 1997 | 2001 | 2005 | 2009 | 2014 |
|---|---|---|---|---|---|---|
| Average Revenue | \$1$\cdot10^9$ | .97 | 1.02 | 1.33 | 1.7 | 2.78 |
| Average Number of Employees | $1\cdot10^3$ | 5.78 | 5.63 | 5.64 | 5.86 | 7.5 |
| Average Labor Price | \$1000 | 50.7 | 54.7 | 69.6 | 81.7 | 95.9 |
| Average Capital | \$1$\cdot10^9$ | .61 | .68 | 1.17 | 1.39 | 2.51 |
| Average Capital Price | % | 18.9 | 21.3 | 20.2 | 20.7 | 21.2 |

model are a little less concave than the fixed labor elasticity in the "Single Frontier" model. A similar finding applies to the effect of capital.

Fig. 4 further compares the two models by showing the "Output-Labor-Capital" relations graphically using 3D images and contour graphs. Overall, the comparisons again show that the average effects of the varying coefficients in the "Varying Frontier" model are a little less concave than the constant effects in the "Single Frontier" model, but the difference is not very significant.

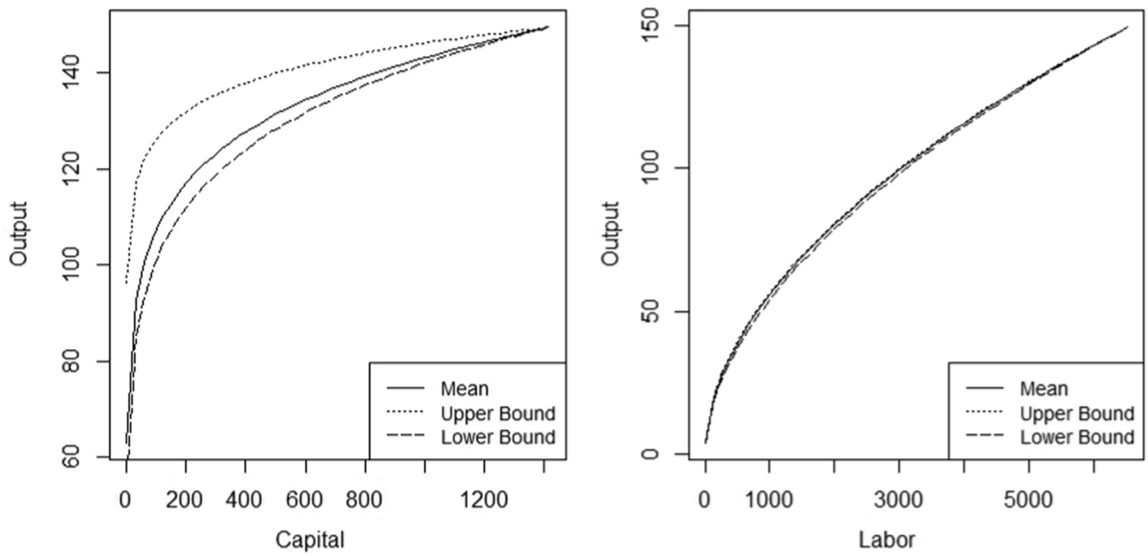


**Fig. 2.** The range of the production frontier in the "Varying Frontier" method.

effects for firms with different business portfolios, which reveals the variation of the aggregated production function under the "Varying Frontier" method. It is clear that the capital effect varies greatly when technologies in different segments are utilized at different frequencies. However, the labor productivities in different segments are very robust. The varying effects support the validation of the segment-specific production assumption.

Fig. 3 calculates the average effects of the varying coefficients in the "Varying Frontier" model and compares them with those in the "Single Frontier" model. The average effects of labor in the "Varying Frontier"

### 4.2. Technical efficiency

Table 2 summarizes the distribution of the efficiency scores in the OFS market. Since the difference is witnessed between the "Single Frontier" model and the "Varying Frontier" model, this paper adds the estimation when the production function takes Transcendental Logarithmic (T-L) form for further comparisons. In practice, this paper drops the top and bottom 2.5% of the estimations to eliminate outliers.

The average efficiency level of the industry is around .3 in the "Single Frontier" model and around .4 in the "Varying Frontier" model.

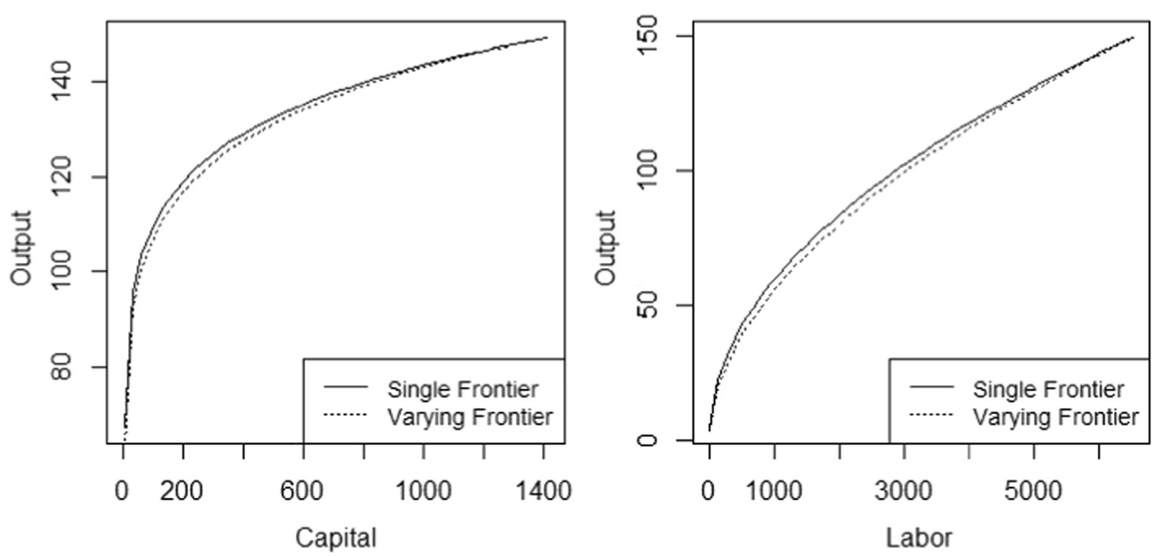


**Fig. 3.** Effect of capital and labor on output in various methods.

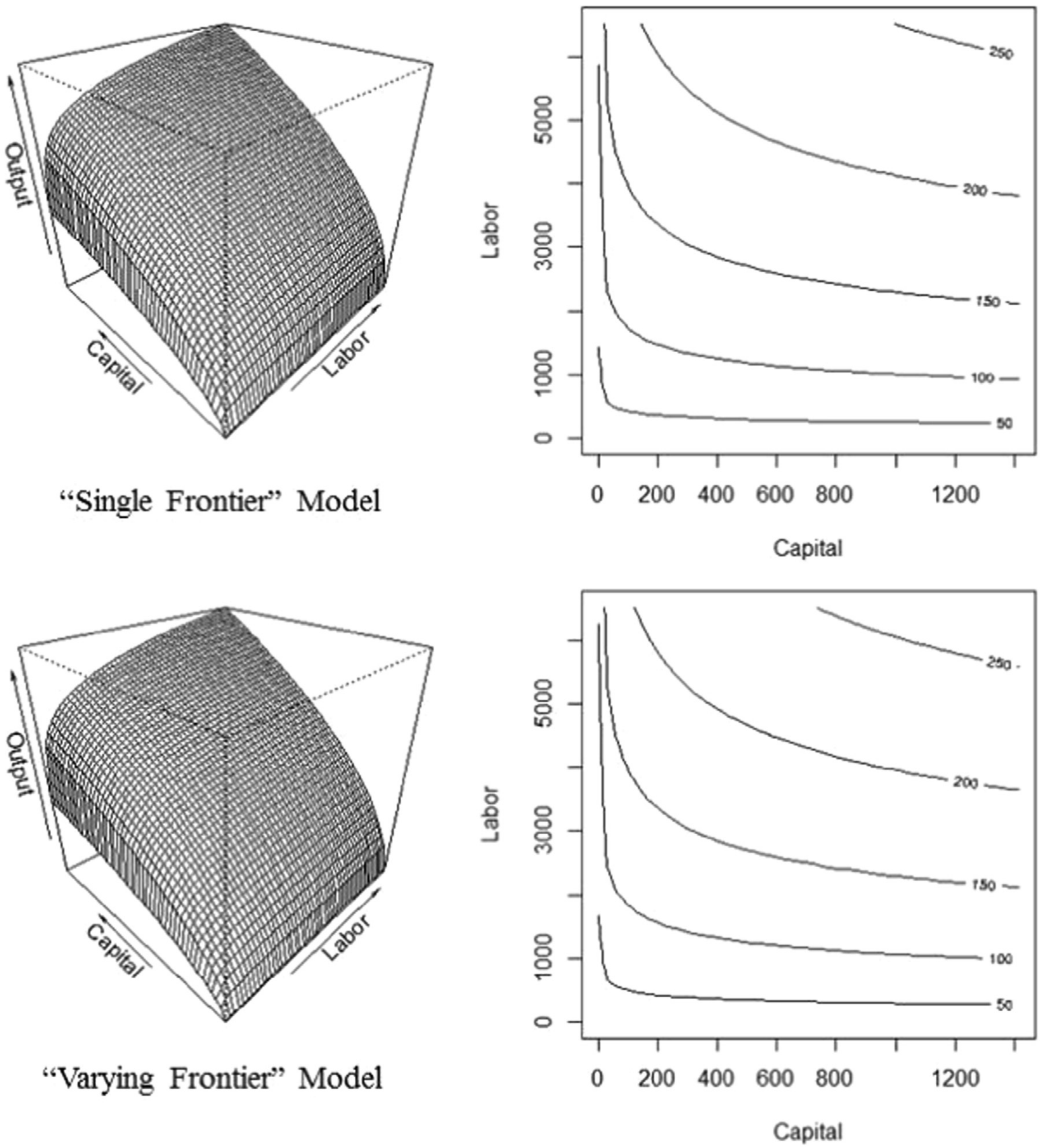


**Fig. 4.** Estimated production frontiers in various methods.

**Table 2**
Technical efficiency statistics.

| | Single frontier | | Varying frontier | |
|---|---|---|---|---|
| | C-D model | T-L model | C-D model | T-L model |
| Mean | .27 | .32 | .43 | .42 |
| Minimum | .06 | .07 | .16 | .16 |
| 25% quantile | .14 | .18 | .29 | .27 |
| 50% quantile | .21 | .26 | .40 | .37 |
| 75% quantile | .34 | .40 | .52 | .49 |
| Maximum | 1.00 | 1.00 | 1.00 | 1.00 |

The "Varying Frontier" model provides more robust estimation than the "Single Frontier" model when assuming different functional forms of the production frontier.

Table 3 presents the distribution of firms by technical efficiency scores, which is classified into four efficiency class intervals. For each of the efficiency class intervals and the overall average estimated technical efficiency scores, the lower and upper bounds of 95% Confidence Interval (CI) are obtained by using a bootstrap technique, called Efron's nonparametric bias-corrected and accelerated (BCa) method, with 10,000 replications (Briggs et al., 1999). It also supports that the efficiency estimated by the "Varying Frontier" model is higher than the one estimated by the "Single Frontier" model.

### 4.3. Impact of the shale revolution

The most important question this paper seeks to answer is whether or not investing in hydraulic fracturing and directional drilling is a good strategy. Especially after the oil price crash in 2014, should we still consider these expensive innovations? In other words, should the shale revolution be cheered or feared, especially in the downturn?

This study uses data in 2013 to estimate Eq. (8) since 113 companies are active this year, which is the largest cross-sectional data in the

**Table 3**
Technical Efficiency Class Interval.

| Eff. range | Single frontier | | | | | | Varying frontier | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | C-D model | | | T-L model | | | C-D model | | | T-L model | | |
| | firm # | Mean (95% CI) | | firm # | Mean (95% CI) | | firm # | Mean (95% CI) | | firm # | Mean (95% CI) | |
| < = .3 | 73 | .17 | (.15–.18) | 65 | .19 | (.17–.21) | 30 | .23 | (.22–24) | 31 | .22 | (.21–.23) |
| .3–.5 | 23 | .38 | (.36–.41) | 24 | .38 | (.36–.41) | 47 | .39 | (.38–.41) | 52 | .39 | (.38–.41) |
| .5–.75 | 8 | .62 | (.58–.67) | 14 | .61 | (.58–.65) | 20 | .58 | (.56–.61) | 14 | .60 | (.57–.62) |
| > .75 | 3 | .89 | (.76–.97) | 4 | .93 | (.88–.98) | 10 | .91 | (.86–.95) | 10 | .88 | (.84–.93) |
| Total | 107 | .27 | (.23–.31) | 107 | .32 | (.28–.36) | 107 | .43 | (.40–.47) | 107 | .42 | (.38–.46) |

panel.[14] There are nine companies that only have hydraulic fracturing techniques, four companies that only have directional drilling techniques, and seven companies that utilize both hydraulic fracturing and directional drilling techniques. Table 4 reports the estimated results, which is the second step regression after estimating the efficiency.

Two things are consistent in all the four columns: 1) larger firms have advantages over smaller firms in terms of efficiency since the coefficient of $R_i$ is always positive and significant[15]; and 2) other things being equal, multidivisional firms on average have neither an advantage nor a disadvantage in efficiency over single-division firms in the OFS market since the coefficient of $M_i$ is statistically and economically insignificant.

The impact of the shale revolution on the "Single Frontier" estimated efficiency is not significant. With the multi-segment concern in the "Varying Frontier" model, however, this impact is both statistically and economically significant. According to the results in the C-D model (column 3 in Table 4), investing in only hydraulic fracturing or only directional drilling will lower a firm's overall efficiency when other things, including firm size, are equal. In other words, if a company has limited resources and cannot expand quickly, it is not a good idea to divest the current business to support innovation in one of hydraulic fracturing and directional drilling. Entering into hydraulic fracturing alone will, on average, have a 12.3 percentage points decrease in efficiency, while entering into directional drilling alone will, on average, have an 8.7 percentage points decrease in efficiency. Since the traditional segments of the OFS are saturated, it is hard to increase market share dramatically in those segments. If a company has sufficient funding in hydraulic fracturing, it has to generate 95% more revenue in order to keep the current efficiency. In other words, a company has to double its revenue in a year so that the benefit from economies of scale can fully compensate the cost of innovating hydraulic fracturing. If investing in directional drilling, this company would also need to increase by two-thirds in revenue to achieve breakeven. This result makes sense because these technologies are innovated to extract oil and gas from the more complex and less productive reservoirs, which involves massive sunk costs and operating costs, and hence it is very difficult to break even.

Does that mean that the shale revolution should be feared and companies need to get rid of their investments? The coefficient of the interaction between hydraulic fracturing and directional drilling is significantly positive, which indicates the existence of spillovers of the combination. For a company that is already engaging in hydraulic fracturing, adding directional drilling to the portfolio can increase efficiency by 9.5 percentage points. For a company that is already engaging in directional drilling, adding hydraulic fracturing to the portfolio can increase efficiency by 5.9 percentage points. Although adding both technologies will still decrease efficiency by 2.8 percentage points for companies that had not previously been engaging in either practice, the breakeven revenue growth is 22%. This growth rate requirement means that each of the two businesses only needs to contribute 11% of the company's sales, which is achievable and much lower than adding hydraulic fracturing alone (95%) and directional drilling alone (67%). This result supports the theory that combining hydraulic fracturing and directional drilling is crucial to the success of the shale revolution. The positive spillover effects of the combination can compensate for the massive investment.

**Table 4**
Efficiency regression result in 2013.

| $\hat{TE}_i$ | Single frontier | | Varying frontier | |
|---|---|---|---|---|
| | C-D model | T-L model | C-D model | T-L model |
| $HF_i$ | −.051 | −.040 | −.123** | −.136** |
| | (.076) | (.080) | (.057) | (.056) |
| $DD_i$ | −.087 | −.092 | −.087** | −.084** |
| | (.057) | (.063) | (.041) | (.040) |
| $HF_i \bullet DD_i$ | .166 | .123 | .182** | .194** |
| | (.110) | (.116) | (.083) | (.082) |
| $R_i$ | .078*** | .092*** | .130*** | .126*** |
| | (.011) | (.012) | (.009) | (.009) |
| $M_i$ | −.007 | −.002 | .0002 | .006 |
| | (.033) | (.035) | (.025) | (.025) |
| Intercept | −.246*** | −.291*** | −.427*** | −.421*** |
| | (.072) | (.075) | (.059) | (.058) |
| $R^2$ | .42 | .46 | .73 | .72 |

Note: Significant at: *10, * *5 and * * * 1%; Standard error in parentheses.

As the first robustness check, this paper repeat the second-step regression using the T-L model derived efficiency as the dependent variable in the fourth column in Table 4, which derives very close numbers, as analyzed above using the C-D model derived efficiency (column 3). As the second robustness check of our results, Table 5 lists the estimated results of the C-D Varying Frontier Model using 2010, 2011, 2012, and 2013, respectively. The results are pretty consistent over time.

### 4.4. More discussions on single and varying frontier models

In the single frontier model, the unique frontier is the highest frontier among all the frontiers in the varying frontier model. Therefore, the efficiency level derived in the single frontier setting must be lower or equal to the one derived in the varying frontier setting, which is verified in Tables 2 and 3. For example, the median efficiency is .21 for single frontier model and .4 for varying frontier model when C-D production function is adopted.

On the one hand, single frontier assumption is invalid logically, as different segments have different production process and use different techniques, which leads to different frontiers. Such invalid assumption

[14] The oil price dropped at 2014, which had significant impact on firms' entry and exit decisions. The oilfield market peaked around 2013 as the stock prices of many oilfield companies hit all-time high, so does the number of firms in the field due to the profitability of the market.

[15] This result is consistent with the opinion of Schlumberger's CEO Paal Kibsgaard, who said in the 4Q2104 Earnings Call that scale is essential to drive performance in the oilfield market. The industry-leading size and integration capabilities are key competitive advantages.

**Table 5**
Efficiency regression results of the C-D varying frontier model over time.

| $\hat{TE}_i$ | 2010 | 2011 | 2012 | 2013 |
|---|---|---|---|---|
| $HF_i$ | −.139** | −.142** | −.138** | −.123** |
| | (.058) | (.056) | (.055) | (.057) |
| $DD_i$ | −.075* | −.110*** | −.103*** | −.087** |
| | (.041) | (.040) | (.037) | (.041) |
| $HF_i \bullet DD_i$ | .192** | .211*** | .195** | .182** |
| | (.080) | (.077) | (.075) | (.083) |
| $R_i$ | .126*** | .131*** | .135*** | .130*** |
| | (.009) | (.009) | (.009) | (.009) |
| $M_i$ | .003 | −.002 | .005 | .0002 |
| | (.026) | (.025) | (.024) | (.025) |
| Intercept | −.348*** | −.398*** | −.453*** | −.427*** |
| | (.054) | (.055) | (.056) | (.059) |
| $R^2$ | .74 | .75 | .75 | .73 |

Note: Significant at: *10, * *5 and * * * 1%; Standard error in parentheses.

of single frontier can lead to biased estimation of efficiency and biased impacts of new techniques sequentially, which reflects on the inaccurate magnitude of the coefficients in the first two columns of Table 4.

On the other hand, the firm-level efficiencies are lack of variations in the single frontier setting, which lead to lower R square and more insignificant coefficients when served as dependent variable. If we look at the coefficients of the new techniques and their interaction term in the first two columns of Table 4, their signs also imply negative effect of either techniques alone and positive spillover effects, as are predicted in the varying frontier model. But due to the lack of variation in efficiency, all the three coefficients are insignificant. Therefore, we cannot provide confident predictions as the one given in varying frontier setting.

To sum, the invalid single frontier assumption can lead to inaccurate magnitude and significance of the coefficients in the second-step regression.

## 5. Conclusion and policy implications

This paper develops a two-step approach to estimate the impact of the shale revolution on firm-level efficiency and investment strategy regarding hydraulic fracturing and directional drilling. Stochastic frontier analysis is applied to derive firm-level efficiency in the first step. The second step regresses the efficiency on indicators of hydraulic fracturing and directional drilling as well as other variables.

The empirical results show that: 1) it is necessary to take multi-segment assumptions into consideration by introducing the revenue share by segment $\theta$; 2) labor elasticity is stable while capital elasticity varies greatly across segments as per the varying coefficient model; 3) investing in only one of the two technologies in the shale revolution is likely to lower the overall efficiency of the company; 4) the combining of hydraulic fracturing and directional drilling produces positive spillover effects, which can compensate for the massive cost and maintain efficiency; 5) more cooperation, alliances, mergers, and acquisitions among experts of the two techniques should be encouraged.

Considering the less endowed reservoirs that the new technologies are working on, the shale revolution has so far been an economic success. OFS companies have the incentive to continue investing, even if there is a fracturing tax. Setting aside the environmental concern,[16] government could encourage the shale revolution since it increases oil and gas recoverable reserves, mitigates the energy shortage, and decreases energy prices. More cooperation, alliances, mergers, and acquisitions among hydraulic fracturing companies and directional drilling companies are also good ideas to share the huge amount of investment and produce spillover benefits. More specifically, there are two policy implications.

Firstly, this article finds that the efficiency enhancement is possible by introducing the new techniques. The empirical results show that these unconventional oil and gas techniques can be either cheer or fear, depending on the investment behavior. The positive spillover effects between hydraulic fracturing and directional drilling is the key to improve efficiency. Therefore, energy policies should be adjusted to encourage more cooperation, mergers and acquisitions among experts of the two techniques in order to achieve the spillover effects, such as the one between Mitchell Energy and Devon Energy, in line with anti-trust laws.

Secondly, either of the two techniques alone leads to lower efficiency. Therefore, the government should help decrease the “entry fee” (e.g., the R & D costs) to encourage more entrants. This is even more important currently, as lower break-even price is required in order to survive in the world with low oil price. Policies to lower the cost can help unconventional shale companies (mainly in the USA) to compete with OPEC members. The government can provide some information sharing platforms to release basic data and knowledge of shale resources, and encourage the cooperation among different companies and institutions.

This study discovers evidence of the competitiveness of the techniques engendered by the shale revolution, which provides information for policy makers and companies. Future studies can focus on the environmental effect of the shale revolution and analyze the social welfare change brought about by these innovations. Moreover, scholars can study how energy policies should be changed to face this competitive entrant in the energy market. Finally, although the varying coefficient model is a semiparametric method and more flexible than the standard parametric method, it still has a rigid functional assumption, such as the C-D and T-L forms. How to relax this assumption would be an interesting field to explore.

## Appendix A. Reasons to adopt time-invariant efficiency

In Eq. (7), the efficiency term, $u_i$, is time-invariant. Then, the dependent variable of Eq. (8) ($TE_i$) is time-invariant as well, since $TE_i = \exp(-u_i)$.

Whether to use time-variant or time-invariant efficiency is a problem that we cannot ignore. Many scholars have studied time-variant technical efficiency. Cornwell et al. (1990) introduced both the within estimator (CSSW) and the generalized least squares estimator (CSSG), where they assumed the firm effects of $\alpha_i$ with $\alpha_{it} = \theta_{i1} + \theta_{i2}t + \theta_{i3}t^2$. Sickles (2005) later examined various specifications of the time-variant firm effect $\alpha_{it}$ modeled in other research, including $\alpha_{it} = \gamma(t)\alpha_i = [1+\exp(bt + ct^2)]^{-1}\alpha_i$ (Kumbhakar, 1990), $\alpha_{it} = \eta_{it}\alpha_i = \exp[-\eta(1 - T)]\alpha_i$ (Battese and Coelli, 1992), $\alpha_{it} = \theta_t\alpha_i$ (Lee and Schmidt, 1993), and the general factor model $\alpha_{it} = c_{i1}g_{1t} + c_{i2}g_{2t} + ... + c_{iL}g_{Lt}$ (Kneip, 1994; Kneip et al., 2003, 2012).

This article tested the difference if time-variant efficiencies are allowed but find no significant difference. Take one of the most popular method (Battese and Coelli: $\alpha_{it} = \exp[-\eta(1 - T)]\alpha_i$) as an example, the estimation of $\eta$ is −.002 with a p-value of .7553, which implies $\eta$ is economically and statistically insignificant. Since $\eta$ is insignificantly different from 0, $\alpha_{it} = \exp[-\eta(1 - T)]\alpha_i \approx \alpha_i$. In other words, the change in efficiency for a company is negligible. Therefore, we can use time-invariant efficiency.

Although the productivity (Total-Factor Productivity, or the frontier of the industry) changes dramatically across time because of the financial crisis, technical revolution, and other events, the firm-specific technical efficiencies are pretty robust according to my own experience after working

[16] Although innovation to make these techniques more environmentally friendly is necessary, it does not fall within the scope of the present article.

in one of the largest OFS companies analyzing competitors in the OFS market. The robustness in efficiency is mainly because companies' reactions are very fast in this industry. OFS market is heavily affected by energy price and economic cycle. When the market is down, companies will sell asset, hold cash, and cut headcount immediately, which guarantees the stability of the efficiency.

To sum up, both the statistical results and my observation of the industry imply that time-invariant efficiency is a valid assumption. The adoption of time-variant efficiencies will lead to negligible difference but cost more degrees of freedom.

## Appendix B. OMR data introduction and adjustment

This study uses data from the Oilfield Market Report (OMR) by Spears & Associates. This report details the global oilfield equipment and service markets associated with five macro-segments: exploration, drilling, completion, production, and capital equipment. Spears & Associates began tracking the OFS market in 1996 and publish its OMR annually. Each year, the report not only releases new data for the current year, but also updates previously published data. Most numbers in the OMR are estimates developed by Spears through five sources: public company reports (about 100 firms), published information, interviews (about 2000 discussions), trade shows, and site visits.

There are several advantages of using the OMR dataset. Firstly, this report brings estimations under the same criteria. Different firms have different segmentations, so direct use of their revenue declarations by product line from their financial reports is not wise. Secondly, this dataset is widely used by most firms and clients in the field. Thirdly, Spears has investigated the numbers through many sources to confirm its estimations in the past twenty years. Lastly, the OMR is updated each year, which alters any incorrect numbers according to the newest information.

In this study, three versions of the OMR (2000, 2011, and 2015) are used to collect firm-level data from 1997 to 2014, which is denoted as OMR1997–2014. OMR2000 includes firm-level revenue by segment from 1997 to 2000, OMR2011 includes firm-level revenue by segment from 1999 to 2011, and OMR2015 includes firm-level revenue by segment from 2005 to 2014. Since different waves of data have different market divisions, this study uses the market segmentation of OMR2015 and adjusts the other two datasets to acquire statistically comparable numbers.

The revision in OMR2000 consist of 1) the "Mud Logging" segment being renamed as the "Surface Data Logging" segment; 2) the "Field Processing Equipment" segment being removed from the market; 3) the "Offshore O & M Services/Contracting" segment being added to the "Offshore Contract Drilling" segment; and 4) the "Production Logging" segment being added to the "Wireline Logging" segment. Moreover, the "Casing & Cementation Products" segment in both OMR2000 and OMR2011 is added to the "Completion Equipment & Services" segment. Finally, the "Pressure Pumping Service" segment in both datasets is divided into the "Cementing" and "Hydraulic Fracturing" segments.

The OMR1997–2014 contains share and size analysis for 32 micro-market segments within the 5 macro-segments from approximately 600 companies working around the world. OMR1997–2014 gives detailed revenue by segment for 275 companies, 114 of which are public firms that publish complete financial information annually. The other 300 smaller companies have been added to "Others" in the report. The detailed segmentation is as follows:

I) Exploration segment includes 1) Geophysical Equipment & Services;
II) Drilling segment includes 2) Cementing, 3) Casing & Tubing Services, 4) Directional Drilling Services, 5) Drill Bits, 6) Drilling & Completion Fluids, 7) Inspection & Coating, 8) Land Contract Drilling, 9) Logging-While-Drilling, 10) Offshore Contract Drilling, 11) Oil Country Tubular Goods, 12) Solids Control & Waste Management, 13) Surface Data Logging;
III) Completion segment includes 14) Completion Eqpt & Services, 15) Coiled Tubing Services, 16) Hydraulic Fracturing, 17) Productions Testing, 18) Rental & Fishing Services, 19) Subsea Equipment, 20) Surface Equipment, 21) Wireline Logging;
IV) Production segment includes 22) Artificial Lift, 23) Contract Compression Services, 24) Floating Production Services, 25) Specialty Chemicals, 26) Well Servicing; and
V) Capital Equipment, Downhole Tools & Offshore Services segment includes 27) Downhole Drilling Tools, 28) Petroleum Aviation, 29) Offshore Construction Services, 30) Rig Equipment, 31) Supply Vessels, and 32) Unit Manufacturing.

## Appendix C. Estimating capital stocks using perpetual inventory method

The perpetual inventory method (PIM) is the most widely employed approach to estimate capital stocks in many statistical offices. Berlemann and Wesselhöft (2014) review the three PIM approaches used most frequently in the literature, consisting of the steady state approach, the disequilibrium approach, and the synthetic time series approach. After comparing the advantages and disadvantages of those three methods, they are able to combine them into a unified approach in order to prevent the drawbacks of the various methods. Their approach follows the procedure proposed by de la Fuente and Doménech (2006).

The PIM interprets a firm's capital stock as an inventory of investments. The aggregate capital stock falls in the depreciation rate per period. Therefore, the capital stock in period $t$ is a weight sum of the history of the capital stock investment:

$$K_t = \sum_{i=0}^{\infty} (1-\delta)^i \bullet I_{t-(i+1)}$$

However, a complete time series of past investments from day one is not available for many companies. Thomson ONE, Bloomberg, and FactSet only cover the recent portion of investment history. Suppose the investment can only be tracked back to period $t_1$, then the current capital stock can be estimated by using

$$K_t = (1-\delta)^{t-t_0} \bullet K_{t_0} + \sum_{i=0}^{t-1} (1-\delta)^i \bullet I_{t-(i+1)} \tag{C-1}$$

Therefore, the information needed to calculate capital stock includes a time series of investment $I_{t-(i+1)}$, the rate of depreciate $\delta$, and the initial capital stock $K_{t_0}$. Firstly, de la Fuente and Doménech (2006) propose smoothing the time-series investment data since the economies are on their adjustment path towards equilibrium rather than staying in a steady state most of the time. Hence, this study smooths the observed capital expenditure (investment) using a regression $I_{it} = \alpha_i + \beta_1 t + \epsilon$ for each firm. Secondly, this study follows the lead of Kamps (2006) and uses time-varying depreciation schemes, which seems to be the most plausible variant. The time-variant smooth depreciation rate can be estimated as the fitted value of the regression $\delta_t = \alpha + \beta_2 t + \epsilon$. This study collects a given firm's annual depreciation and total capital data to calculate the depreciation rate in accounting and use this information to run the regression. Finally, the initial capital stock at time $t_0$ can be calculated from the investment $I_{t_1}$, the long-term investment growth rate $g_I$, and the estimated depreciation rate $\delta$: $K_{t_0} \approx I_{t_1}/(g_I + \delta_{t_1})$, where the growth rate $g_I$ is $\beta_1$ and the investment $I_{t_1}$ is the fitted value in the same regression. Similar to the method used in Berlemann and Wesselhöft (2014), this study assumes all the years before $t_1$ without desegregated data have the same constant depreciation rate as year $t_1$. But for all the recent years that we have investment data, the depreciation rate is time variant. Therefore, Eq. (C-1) becomes:

$$K_t = \prod_{i=t_1}^{t} (1-\delta_i) I_{t_1}/(g_I + \delta_{t_1}) + \sum_{i=0}^{t-1} \prod_{j=t-(i+1)}^{t-1} (1-\delta_j) I_{t-(i+1)}$$

In our empirical study, $t$ is 2014 for most companies that are still active while $t_1$ presents the first year of investment data and varies across firms.

## References


Ackerberg, D.A., Caves, K., Frazer, G., 2015. Identification properties of recent production function estimators. Econometrica 83, 2411–2451.

Adams, R.M., Berger, A.N., Sickles, R.C., 1999. Semiparametric approaches to stochastic panel frontiers with applications in the banking industry. J. Bus. Econ. Stat. 17, 349–358.

Ahmad, I., Leelahanon, S., Li, Q., 2005. Efficient estimation of a semiparametric partially linear varying coefficient model. Ann. Stat. 258–283.

Aigner, D., Lovell, C.A., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. J. Econ. 6, 21–37.

Amsler, C., Prokhorov, A., Schmidt, P., 2015. Endogeneity in stochastic frontier models. J. Econ.

Battese, G.E., Coelli, T.J., 1992. Frontier Production Functions, Technical Efficiency and Panel Data: with Application to Paddy Farmers in India. Springer.

Berlemann, M., Wesselhöft, J.-E., 2014. Estimating aggregate capital stocks using the perpetual inventory method–a survey of previous implementations and new empirical evidence for 103 countries–. Rev. Econ./Jahrb. Wirtsch. 65.

Blundell, R., Bond, S., 2000. GMM estimation with persistent panel data: an application to production functions. Econom. Rev. 19, 321–340.

Blundell, R.W., Powell, J.L., 2004. Endogeneity in semiparametric binary response models. Rev. Econ. Stud. 71, 655–679.

Briggs, A.H., Mooney, C.Z., Wonderling, D.E., 1999. Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using Monte Carlo simulation. Stat. Med. 18, 3245–3262.

Coelli, T., Henningsen, A., Henningsen, M.A., 2012. Package 'frontier'. Technical Report, R.

Cornwell, C., Schmidt, P., Sickles, R.C., 1990. Production frontiers with cross-sectional and time-series variation in efficiency levels. J. Econ. 46, 185–200.

de la Fuente, A., Doménech, R., 2006. Human capital in growth regressions: how much difference does data quality make? J. Eur. Econ. Assoc. 4, 1–36.

Eller, S.L., Hartley, P.R., Medlock, K.B., 2011. Empirical evidence on the operational efficiency of National Oil Companies. Empir. Econ. 40, 623–643.

Fan, J., Huang, T., 2005. Profile likelihood inferences on semiparametric varying-coefficient partially linear models. Bernoulli 11, 1031–1057.

Fan, J., Li, R., 2004. New estimation and model selection procedures for semiparametric modeling in longitudinal data analysis. J. Am. Stat. Assoc. 99, 710–723.

Fan, J., Zhang, W., 2008. Statistical methods with varying coefficient models. Stat. Interface 1, 179.

Fan, Y., Li, Q., Weersink, A., 1996. Semiparametric estimation of stochastic production frontier models. J. Bus. Econ. Stat. 14, 460–468.

Foster, L., Haltiwanger, J., Syverson, C., 2008. Reallocation, firm turnover, and efficiency: selection on productivity or profitability? Am. Econ. Rev. 98, 394–425.

Greene, W.H., 1990. A gamma-distributed stochastic frontier model. J. Econ. 46, 141–163.

Guan, Z., Kumbhakar, S.C., Myers, R.J., Lansink, A.O., 2009. Measuring excess capital capacity in agricultural production. Am. J. Agric. Econ. 91, 765–776.

Hartley, P.R., Medlock III, K.B., 2013. Changes in the operational efficiency of national oil companies. Energy J. 34, 27.

Hastie, T., Tibshirani, R., 1993. Varying-coefficient models. J. R. Stat. Soc. Ser. B (Methodol.) 757–796.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. CRC Press.

Hayfield, T., Racine, J.S., 2008. Nonparametric econometrics: the np package. J. Stat. Softw. 27, 1–32.

Henningsen, A., Kumbhakar, S., 2009. Semiparametric stochastic frontier analysis: An application to Polish farms during transition, In: proceedings of the European Workshop on Efficiency and Productivity Analysis (EWEPA) in Pisa, Italy, June.

Hu, X., 2014. Estimation in a semi-varying coefficient model for panel data with fixed effects. J. Syst. Sci. Complex. 27, 594–604.

Kamps, C., 2006. New estimates of government net capital stocks for 22 OECD countries, 1960–2001. IMF Staff Pap. 53, 120–150.

Kim, Y.-J., 2013. A partial spline approach for semiparametric estimation of varying-coefficient partially linear models. Comput. Stat. Data Anal. 62, 181–187.

Kneip, A., 1994. Nonparametric estimation of common regressors for similar curve data. Ann. Stat. 1386–1427.

Kneip, A., Sickles, R., Song, W., 2003. On Estimating a Mixed Effects Model with Applications to the U.S. Banking Industry. Rice University, Mimeo.

Kneip, A., Sickles, R.C., Song, W., 2012. A new panel data treatment for heterogeneity in time trends. Econom. Theory 28, 590–628.

Kumbhakar, S.C., 1990. Production frontiers, panel data, and time-varying technical inefficiency. J. Econ. 46, 201–211.

Kumbhakar, S.C., Sun, K., 2013. Estimation of a Flexible Stochastic Cost Frontier Model with Environmental Factors Subject to Economic Constraints, In: Proceedings of the European Economic Association & Econometric Society 2013 Parallel Meetings, Gothenburg, Sweden.

Lee, Y.H., Schmidt, P., 1993. A production frontier model with flexible temporal variation in technical efficiency. Meas. Product. Effic.: Tech. Appl. 237–255.

Levinsohn, J., Petrin, A., 2003. Estimating production functions using inputs to control for unobservables. Rev. Econ. Stud. 70, 317–341.

Liu, J., 2014. Essays on Productivity and Panel Data Econometrics. Rice University (Doctoral Thesis). http://hdl.handle.net/1911/77201.

Lu, Q., Yang, C., Li, J., 2008. Rural-urban migration, rural household income and sustainable development in rural areas of China. Chin. J. Popul., Resour. Environ. 06, 70–73.

Managi, S., Opaluch, J.J., Jin, D., Grigalunas, T.A., 2004. Technological change and depletion in offshore oil and gas. J. Environ. Econ. Manag. 47, 388–409.

Managi, S., Opaluch, J.J., Jin, D., Grigalunas, T.A., 2006. Stochastic frontier analysis of total factor productivity in the offshore oil and gas industry. Ecol. Econ. 60, 204–215.

Marschak, J., Andrews, W.H., 1944. Random simultaneous equations and the theory of production. Économ. J. Econom. Soc. 143–205.

Meeusen, W., Van den Broeck, J., 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. Int. Econ. Rev. 435–444.

Olley, G.S., Pakes, A., 1996. The dynamics of productivity in the telecommunications equipment industry. Econometrica 64, 1263–1297.

Schmidt, P., Sickles, R.C., 1984. Production frontiers and panel data. J. Bus. Econ. Stat. 2, 367–374.

Sickles, R.C., 2005. Panel estimators and the identification of firm-specific efficiency levels in parametric, semiparametric and nonparametric settings. J. Econ. 126, 305–334.

Stasinopoulos, D.M., Rigby, R.A., 2007. Generalized additive models for location scale and shape (GAMLSS) in R. J. Stat. Softw. 23, 1–46.

Stevenson, R.E., 1980. Likelihood functions for generalized stochastic frontier estimation. J. Econ. 13, 57–66.

Su, L., Ullah, A., 2006. Profile likelihood estimation of partially linear panel data models with fixed effects. Econ. Lett. 92, 75–81.

Sun, K., Kumbhakar, S.C., 2013. Semiparametric smooth-coefficient stochastic frontier model. Econ. Lett. 120, 305–309.

Sun, Y., Carroll, R.J., Li, D., 2009. Semiparametric estimation of fixed effects panel data varying coefficient models. Adv. Econ. 25, 101–129.

Terza, J.V., Basu, A., Rathouz, P.J., 2008. Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. J. Health Econ. 27, 531–543.

Thompson, R.G., Dharmapala, P., Rothenberg, L.J., Thrall, R.M., 1996. DEA/AR efficiency and profitability of 14 major oil companies in US exploration and production. Comput. Oper. Res. 23, 357–373.

Wolf, C., 2009. Does ownership matter? The performance and efficiency of State Oil vs. Private Oil (1987–2006). Energy Policy 37, 2642–2652.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT press.

Zhang, R., Sun, K., Delgado, M.S., Kumbhakar, S.C., 2012. Productivity in China's high technology industry: regional heterogeneity and R & D. Technol. Forecast. Social. Change 79, 127–141.