



# Efficiency of European oil companies: an empirical analysis

Blanca Sanchez-Robles ·  
Teresa C. Herrador-Alcaide ·  
Montserrat Hernández-Solís

Received: 16 December 2021 / Accepted: 22 October 2022 / Published online: 12 November 2022  
© The Author(s) 2022

**Abstract** We explore the efficiency in the oil industry employing a sample of around 300 companies operating in Europe over 2010–2019. We construct efficiency scores by means of DEA non-parametric techniques. Average efficiency in the sample is modest, 0.27, and decreasing over time. We examine the association of efficiency with economic and financial variables. Results suggest that size is closely associated to efficiency. Large (in the top 10th percentile of income) and very small firms are more efficient, *ceteris paribus*, whereas medium size and small firms exhibit lower levels of efficiency. Firms which have strived more to implement sustainable technologies and cut greenhouse emissions are more efficient too. Increases in employee costs and decreases in financial solvency jeopardize efficiency, which displays a positive correlation with economic activity and oil prices. As a robustness test, we have repeated the exercise by means of the Simar-Wilson methodology, which confirms our main results regarding efficiency levels and trends and correlations with other variables.

Our findings suggest that an industry consolidation is foreseeable in the future.

**Keywords** Efficiency · DEA · Environment · Greenhouse gas emissions · Oil refineries · Simar-Wilson · Bootstrap

**JEL Classification** C44 · C61 · L71 · M21

## Introduction

The oil industry plays a key role in the economy. Fluctuations in the price of oil greatly influence economic variables (Cuñado and Perez de Gracia, 2003). The price of energy is an important component of the cost of many goods and services; as the current scenario shows, rises in energy prices bring about fast growth in inflation rates (International Monetary Fund, 2022; ECB, 2018). The increasing awareness about climate change, greenhouse gas emissions (GHG) and sustainability has also fostered the interest of policymakers and citizens in this industry, whose environmental impact is substantial (Mariano & La Rovere, 2017).

One widespread approach to assess the performance of an industry is to explore its efficiency. There are two main tools to carry out this analysis: stochastic frontier analysis (SFA) and data enveloping analysis (DEA). Both employ as the starting point data on inputs and outputs for a group of decision-making units (DMUs), typically companies. SFA,

---

B. Sanchez-Robles  
Department of Economic Analysis, UNED, Madrid, Spain  
e-mail: bsanchez-robles@cee.uned.es

T. C. Herrador-Alcaide (✉) · M. Hernández-Solís  
Department of Finance and Accounting, UNED, Madrid, Spain  
e-mail: therrador@cee.uned.es

M. Hernández-Solís  
e-mail: montserrat@cee.uned.es

pioneered by Aigner et al. (1977) and Meeusen and Broeck (1977) works from a parametric and statistical approach, estimates a production function and captures inefficiency in the error term. DEA, proposed by Charnes et al. (1978), is instead a non-parametric technique which constructs a best practice frontier by linear programming, computes the relative distance of DMUs to the frontier and generates an efficiency score for every DMU in the sample.

In the last decades DEA has been employed extensively and successfully for the assessment of many industries, including energy (see Emrouznejad & Yang, 2018, for a review). Here the literature has focused primarily on the analysis of electricity and energy saving; the exploration of other aspects, as the oil industry, has been sparser in comparison despite its relevancy (Sueyoshi et al., 2017; Zhou et al., 2008).

This paper contributes to filling this void in the literature by analyzing the efficiency of oil firms operating in Europe over 2010–2019. More in particular, this exploration has the following goals: first, it aims to shed more light on the recent performance of efficiency in this industry. Second, it intends to identify managerial and macroeconomic aspects associated with efficiency. Another goal of our investigation is to provide practical recommendations for firm managers, stakeholders, and policymakers related to this industry.

We pursue our investigation empirically by computing efficiency scores for a large sample of European oil companies using a DEA baseline model. Additionally, we explore the statistical association between efficiency and various managerial and macroeconomic variables as size, activity, environmental policies, financial management, human resource management, oil price, and economic activity, among others.

In accord with part of the literature (Bang et al., 2019; da Silva et al., 2019; Lee et al., 2009; McDonald, 2009; Miao et al., 2020; Sağlam, 2018, among others), we have divided the empirical analysis in two stages: in the first one we compute efficiency scores; in the second stage, we explore the correlation between the efficiency scores estimated in the first stage and different sets of variables.

We have combined several DEA tools for this investigation. We have worked with the baseline DEA model in the first stage and the Tobit specification in the second. Since these methodologies are not devoid of some limitations, we have completed this exercise with another one, grounded upon the Simar

and Wilson methodology (1998; 2000; 2007) and performed by means of bootstrap. This procedure provides a natural robustness test for our findings with baseline DEA and Tobit. By and large the main messages obtained with the first approach carry over when the Simar-Wilson methodology is employed.

One of the aspects of firm management we are interested in is the environmental concern. One potential strategy to explore this issue is to employ DEA models with desirable and undesirable outputs (non-polluting and polluting, respectively). A relevant line of research pioneered by Sueyoshi and Goto (2012a, b), among others, has coined and assessed the distinction between *operational* and *environmental efficiency*.<sup>1</sup>

Our approach in this paper, however, is largely determined by data. We have access to a rich dataset of microdata which enables us to explore and compare the performance of around 300 companies; detailed data on environmental aspects of companies in such a large scale are not available yet, though, either in the dataset we are using or in other compilations, as far as we know. A thorough review of environmental reporting by companies has enabled us to gather environmental data, but only for a reduced subsample of the original sample. Hence, the computation of environmental efficiency along the lines of Sueyoshi and Goto (2012a) is not feasible for our complete sample. For the purpose of this paper, we have opted to exploit as much as possible our primary dataset and focus on operational efficiency. Nonetheless, we complement the information for the large sample with the environmental data we collected and employ this combination in the second stage of our empirical analysis, as it will be detailed below.

This paper contributes to the literature in several ways. First, it provides an exploration of a sample of European oil companies; this is a geographical area which has not been covered enough by the DEA literature so far, despite its large theoretical and practical interest. As “[Related literature and theoretical framework](#)” will discuss in more detail, there are empirical analyses for US, Asian, and Middle East companies

<sup>1</sup> Operational efficiency considers only desirable outputs; environmental efficiency takes also into account the generation of undesirable outputs (pollutants, residuals or waste). See Sueyoshi et al. (2017) for a survey of the related literature.

and for general samples of firms worldwide. There are almost no studies, however, for samples of European companies.<sup>2</sup> We think it is important to fill this void and focus on firms operating in Europe for a number of reasons. From a theoretical approach, Europe is an appealing area for comparative economic analyses because it combines a substantial degree of economic convergence and homogeneity with idiosyncratic features and diversity across regions and countries. These characteristics imply that, while the performance of companies operating in Europe is comparable, it displays significant differences as well which may be exploited by the researcher to identify trends and patterns.

There are other, more practical reasons which justify the investigation of the European oil industry. Europe is an important agent in the worldwide oil sector. In 2019, the EU represented 27% of global imports of crude oil and 30% of global imports of oil products (International Energy Agency, 2022).<sup>3</sup> At the macroeconomic level, this high dependency from other areas poses major risks to the European population if geopolitical tensions arise, as we are currently seeing in regard to the Russian oil. At the microeconomic level, moreover, there is some concern about the performance of the European oil sector in recent years, since policy reports have detected low utilization rates, high operation costs and overcapacity (Lukach et al., 2015). It is true that there has been some progress in the recent past in the correction of these shortcomings, but the improvement has not been sufficient. Increasing competition from other areas (notably Asia) is jeopardizing the survival capacity of European firms (Nivard & Kreijkes, 2017). The energy dependency of Europe from outside (which poses a concern for the European authorities) could be reduced if the European oil industry reached higher levels of efficiency and competitiveness.

Another contribution of our paper is related to the data we use. We have gathered a rich dataset covering around 300 firms over a 10-year period. This allows us to work with a sizeable number of observations

and explore in detail the association of efficiency with size, activity, country, region, and other aspects, obtaining interesting insights for managers, potential investors, and other stakeholders of oil companies. Third, our analysis provides specific policy recommendations which can orientate the industrial policy strategy regarding this sector primarily not only in Europe but also in other geographical areas.

Our investigation is grounded upon the hypothesis that the level of efficiency in European firms is not high, since policy reports have already alerted about various dysfunctions in the European oil sector. Another hypothesis is that there must be some association between efficiency and size, because of the presence of increasing returns to scale linked to the heavy requirements of technology and capital for this industry. We also believe that firms present features which are linked to efficiency and can be approximated by means of variables; thus, the design and estimation of appropriate models may help identify the sign and magnitude of the association of these variables with efficiency.

The structure of this paper is as follows: The “[Related literature and theoretical framework](#)” section summarizes the related literature and our theoretical framework. The “[Methodology](#)” section describes our methodology. The “[Data and variables](#)” section discusses our data and variables. The “[Empirical results](#)” section details our baseline empirical exercise, and “[Sensitivity analysis](#)” carries out a robustness test. The “[Concluding remarks and policy recommendations](#)” section concludes and offers policy recommendations.

## Related literature and theoretical framework

Recent contributions have analyzed the performance of efficiency in oil companies. By and large, the most common tool for this analysis has been DEA, with the number of employees and a proxy for assets as inputs and revenue or production as output.

Table 1 compiles a set of recent papers in this literature. Some of these papers focus on the performance of oil firms from a particular country such as China (Hanrui & Xun, 2011; Song et al., 2015), India (Vikas & Bansal, 2019), the USA (Atris & Goto, 2019; Mekaroonreung & Johnson, 2010; Sueyoshi and Wang, 2014; 2018), and Indonesia (Putra

<sup>2</sup> One exception is Bevilacqua and Braglia (2002). Their analysis, though, covers only seven Italian refineries.

<sup>3</sup> The latest data available are for 2020; we prefer to end the period in 2019 since data for 2020 may be misleading because of the pandemic.

**Table 1** Efficiency in oil companies, selected papers

| Authors                          | Sample   | Period    | Scope     | Finding   |
|----------------------------------|--|-----------|-----------|---|
| Atris and Goto (2019)            | 34 oil and gas companies                           | 2011–2015 | USA       | Integrated firms are more environmentally and operationally efficient than independent  |
| Eller et al. (2011)              | 78 NOCs and SOCs                                   | 2002–2004 | World     | National oil companies (NOCs) less efficient due to government goals  |
| Hanrui and Xun (2011)            | 13 oil fields from one domestic oil company        | 2003–2007 | China     | Comparison of efficiency by DEA and bootstrap-DEA. Bootstrap-DEA allows for a more adjusted calculation of efficiency to the reality of the company |
| Hartley and Medlock (2013)       | 61 NOCs and SOCs                                   | 2001–2009 | World     | Privatizations and mergers increase efficiency. NOCs less efficient than other oil companies  |
| Lim and Lee (2020)               | 30 OECD countries                                  | 2005–2016 | OECD      | Structural decline of oil industry in major OECD countries  |
| Mekaroonreung and Johnson (2010) | 113 refineries                                     | 2006–2007 | USA       | Environmental regulations reduce output but less so for efficient refineries  |
| Ohene-Asare et al. (2017)        | 50 oil and gas companies                           | 2001–2010 | World     | Private firms are more efficient  |
| Putra and Adinugraha (2018)      | 25 oil and gas firms                               | 2005–2014 | Indonesia | Dip in reserves, increases in costs and regulations decrease efficiency   |
| Song et al. (2015)               | 20 listed firms                                    | 2006–2011 | China     | Trade-off between technical and environmental efficiency  |
| Sueyoshi and Goto (2012a)        | 14 NOCs and 5 SOCs                                 | 2005–2009 | world     | Discusses natural and managerial disposability  |
| Sueyoshi and Goto (2015)         | 17 firms   | 2005–2009 | world     | Technological progress is observed, but operational efficiency does not improve   |
| Sueyoshi and Wang (2014)         | 82 independent and 20 integrated oil and gas firms | 2010–2012 | US        | Integrated companies outperform because of scale  |
| Sueyoshi and Wang (2018)         | 30 US oil firms                                    | 2011–2016 | US        | Positive association oil prices and efficiency  |
| Sun et al. (2017)                | 10 top oil firms                                   | 2003–2013 | world     | Not controlling for unobserved heterogeneity biases efficiency in SFA   |
| Vikas and Bansal (2019)          | 22 oil and gas companies                           | 2013–2017 | India     | 59% of firms are efficient  |

Source: own elaboration

& Adinugraha, 2018). Others, instead, pursue an exploration of OECD countries (Lim & Lee, 2020) or the whole world (Eller et al., 2011; Ohene-Asare et al., 2017; Sun et al., 2017). In general, efficiency in emerging countries grows over time (Hanrui & Xun, 2011), whereas it exhibits a decreasing trend in developed countries (Lim & Lee, 2020).

Most of the literature goes beyond the mere discussion of efficiency trends and explores their potential drivers along different dimensions, usually by means of regressions where the dependent variable is efficiency and the regressors are variables capturing the potential drivers. Managerial practices and other variables capturing aspects which

are internal to companies are important candidates in this regard.

Company size may be associated to efficiency when the production function does not exhibit constant decrease to scale. In fact, the oil sector is heavily dependent on fixed assets (plants, machine, equipment) and state-of-the-art technology, which can be best capitalized when the number of units produced is large, entailing in turn increasing returns to scale (Romer, 1990).<sup>4</sup>

<sup>4</sup> A similar phenomenon takes place in the pharmaceutical sector regarding expenditures in research and development, which are also a fixed cost with respect to the number of units produced.

This aspect has also been addressed by the empirical literature on this industry. Sueyoshi and Wang (2014) suggest that integrated companies operating along the entire supply chain outperform independent companies because integration facilitates economies of scale. Atris and Goto (2019) argue that large companies outperform small ones in terms of efficiency. Vikas and Bansal (2019) detect inefficiencies of scale in a significant number of firms included in a sample of Indian oil and gas listed companies.

It is also possible that there are competitive advantages associated to niches, thus benefiting small firms. This kind of pattern has also been detected in other industries, where the efficiency of very big and extremely small firms dominates the rest, due in turn to the coexistence of increasing returns to scale and niche specialization (Díaz & Sanchez-Robles, 2020, 2022). Some papers suggest this possibility. For example, Ismail et al. (2013) document for their sample that large and small oil firms outperform medium size companies in terms of efficiency. Mekaroonreung and Johnson (2010) find that very specialized, small firms outperform their counterparts in a sample of US oil companies.

The literature has also documented the impact of labor costs and human resource management in the performance of the firms in this sector. According to Lukach et al. (2015), operating costs are higher in companies of Europe than in other regions and exhibit an increasing trend, partly because of personnel expenses. Al-Najjar and Al-Jaybajy (2012) find that an excessive number of employees damages efficiency in a sample of oil firms from Iraq.

Financial stability has also been found to be a factor associated with efficiency (Díaz & Sanchez-Robles, 2020, 2022). The financial structure is usually crucial for capital-intensive companies and may have an important effect on its performance. On a priori grounds, it can be argued that more leveraged balance sheets have a positive impact on the firm, since external finance is normally cheaper than the return requested by shareholders. The opposite link may also exist if information asymmetries prevail in financial markets, since under this hypothesis, banks usually charge a higher interest rate premium to more indebted firms, harder to monitor, and riskier (Bernanke & Gertler, 1995). It is, thus, an empirical issue which depends on the degree of competitiveness of

financial markets and must also be elucidated for each sector.

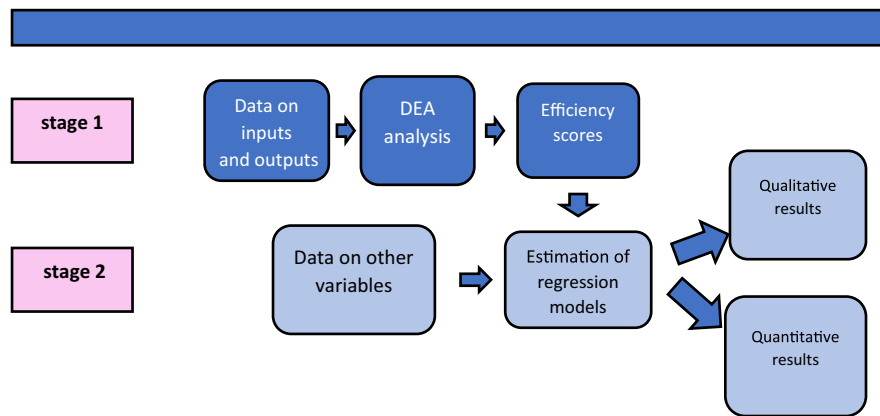
Another managerial aspect we are interested in is the relationship between efficiency and environmental conscience. A priori, it is not clear whether there is a positive or negative correlation between efficiency and environmental concern. Both scenarios are feasible from a theoretical point of view. If firms must incur in large outlays to operate sustainably, a trade-off may exist between efficiency and environmental concern.<sup>5</sup> If state-of-art, more advanced technologies allow simultaneously a better utilization of inputs and a reduction of undesirable outputs, then efficiency and environmental concern may be positively associated. Empirical evidence in this regard is mixed. Boyd and McClelland (1999) and Berman and Bui (2001) document a positive relationship between efficiency and environmental awareness for US firms. Ismail et al. (2013) arrive at the same conclusion using a sample made up of firms worldwide.

Instead, Song et al. (2015) report a negative association for 20 listed Chinese firms. Mekaroonreung and Johnson (2010) analyzed the efficiency of 113 US oil refineries in 2006 and 2007. They show that environmental regulations have damaged company efficiency, although this impact is smaller for efficient firms. Sueyoshi and Wang (2018) find that the environmental pressure exerted by regulations and stakeholders has damaged efficiency in a sample of 30 US firms.<sup>6</sup>

The literature has also explored the association of external variables with efficiency. Key variables for the oil industry are the oil price, which proxies for supply, and the economic activity, which captures the demand side of the market. The sign of the impact of oil price on efficiency is not straightforward. Growing oil prices should positively affect oil production (Friedman, 1992); low prices, however, could also increase output if companies try to offset this factor by expanding production (Mohaddes & Pesaran, 2016). In this regard, Sueyoshi and Wang (2018)

<sup>5</sup> This is also the outcome under weak disposability of output, where undesirable outputs may only be cut if desirable outputs diminish (Mekaroonreung and Johnson, 2010).

<sup>6</sup> The issue of pressure from stakeholders has also been detected by Atris and Goto (2019), who stress the positive impact of this aspect on the environmental efficiency in large companies.



**Fig. 1** The methodology in this paper. A summary. Note: The Figure illustrates our methodology in this paper. It is divided in two stages. In the first stage, we use data on inputs and outputs to get efficiency scores per company and year in the sample. The tool to perform this computation is DEA. In the second stage, we estimate regression models using the efficiency

scores from stage 1 as the dependent variable and data on various variables which capture internal and external aspects of companies as regressors. Regression models are estimated according to different procedures (Tobit, Simar-Wilson) and provide quantitative and qualitative results

detect a positive association between oil prices and efficiency while Putra and Adinugraha (2018) do not.

It is also possible that efficiency be affected by factors associated to the country of origin of companies (through features like government policies, the institutional environment, and the education of the labor force, among others). Putra and Adinugraha (2018), for example, document a negative correlation between government intervention and efficiency for their sample.

The discussion of the literature suggests, therefore, an association between efficiency on the one hand and aspects such as size, human resource management, financial management, environmental policies, country characteristics, and oil price on the other. We explore in more detail these associations in the second stage of our empirical exercise by means of the estimation of several models relating efficiency to variables capturing these aspects.

## Methodology

Our methodology in this paper is empirical and carried out in two stages. Figure 1 provides a visual summary of the methodology. In the first stage, we employ DEA to compute efficiency scores from data for each year and company in the sample. In the second stage, we

construct a panel; then, we use regression analysis and the efficiency scores estimated in the first stage to examine the performance of several variables potentially associated with efficiency, qualitatively, and quantitatively. This section discusses this methodology in more detail, distinguishing between the first and second stages.

### First stage

In the first stage, we employ data on inputs and outputs for the firms in our sample. We get an efficiency score per company and year. For this computation, we use DEA, a popular tool in operations research, whose main features we describe next.

Essentially, DEA constructs a technological frontier from data about inputs and outputs of companies. This frontier can be understood as the geometrical locus of production plans in an input/output space.<sup>7</sup> The best performers in the group lie on the frontier and register the maximum efficiency score by definition. DEA computes efficiency scores for the rest of the companies (placed inside the frontier) according to their distance to the best performers.

In terms of the solution strategy, DEA solves a linear programming problem, searching for a vector

<sup>7</sup> Hence, the frontier determines the maximum attainable output with a given level of input (in output oriented models) or,

in the feasible set satisfying the optimality criteria. Both the objective function and the constraints are linear equations.<sup>8</sup> Constraints specify technical and economic conditions, typically in the form of total available quantities of inputs (in the output-oriented model), desired levels of output (in the input-oriented model), and non-negativity of variables representing quantities. The search for the solution requires the use of algorithms when the number of outputs and inputs (and hence the dimensionality of the problem) gets large.<sup>9</sup>

The theoretical setup for DEA is grounded on the following, very general assumptions (Simar & Wilson, 1998)<sup>10</sup>:

1. There are  $n$  DMUs indexed by  $j$  ( $j = 1, \dots, n$ ), and a technology  $\Gamma$  transforming inputs into outputs:

$$\Gamma = \{(x, y) : y \text{ can be produced by } x\} \tag{1}$$

where  $x_j$  is a vector of  $m$  inputs and  $y_j$  is a vector of  $s$  outputs for DMU  $j$ .

$$\begin{aligned} x_j &= (x_{1j}, \dots, x_{mj}) \in R^m \\ y_j &= (y_{1j}, \dots, y_{sj}) \in R^s \end{aligned} \tag{2}$$

The feasible set in this problem is the production possibility frontier  $P(x)$  or, alternatively, the input requirement set  $L(y)$ :

$$P(x) \equiv \{y : (x, y) \in \Gamma\} \tag{3}$$

$$L(y) \equiv \{x : (x, y) \in \Gamma\} \tag{4}$$

where  $P(x)$  is closed, convex, and bounded for all  $x \in R^m$  and  $L(y)$  is closed and convex for all  $y \in R^s$ .

2. Inputs and outputs are freely disposable, i.e., technology is monotonic.
3. Given a vector of control variables  $z$ , which are potentially associated with efficiency, there exists a density function  $f(x, y|z)$ , strictly positive and continuous.<sup>11</sup>
4. Within the production possibility frontier, technology is differentiable.

Since DEA is a non-parametric technique, there is no need to assume a specific functional form (as Cobb–Douglas or Translog) for the technology mapping inputs to outputs. The only crucial technological feature to be specified is whether returns to scale are constant or variable.

In this paper, we work with the input-oriented setup of the problem because it is, in our view, intuitively more appealing and closer to the actual practice in firms and other institutions.<sup>12</sup> For the specific case of oil companies, the assumption of variable (increasing) returns to scale seems more realistic: the installation, maintenance and upgrading of the necessary technology, infrastructure, and equipment entail larger fixed costs which bring about decreasing average costs in production.

Define  $\theta_j$  as the input-oriented efficiency score for DMU  $j$ , with  $0 < \theta_j \leq 1$ . Intuitively,  $\theta_j$  informs of the reduction in inputs which DMU <sub>$j$</sub>  should carry out in order to become efficient, relative to the best performers in the sample. In this setting, the best performers attain an efficiency score of 1 by construction. To compute an efficiency score for every DMU in the sample means solving Eq. (5) for each unit:

$$\min\left\{\theta > 0 : \theta x \geq \sum_{j=1}^n \lambda_j x_j; y \leq \sum_{j=1}^n \lambda_j y_j, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n\right\} \tag{5}$$

where  $\lambda$  stands for the set of multipliers in the linear combinations of the DMUs' inputs and outputs, i.e., the weight of each DMU within the peer group of

Footnote 7 (continued)

alternatively, the minimum necessary input level for a given amount of output (in the input oriented framework).

<sup>8</sup> Charnes and Cooper (1962) designed DEA as a fractional programming problem; later on, they redefined it in terms of linear programming (Charnes et al., 1978). See Raith et al. (2019) and Førsund (2018).

<sup>9</sup> One of the most common algorithms is simplex. Simplex looks for and evaluates an extreme point of a feasible solution; if it is not optimal, the algorithm looks for an adjacent solution and repeats the process until convergence.

<sup>10</sup> In this presentation of DEA, we omit the explicit mention of the time subscript to alleviate notation. Note, however, that we work with yearly data on inputs and outputs and obtain efficiency scores for each company and year.

<sup>11</sup> Typically called environmental variables in the literature. We prefer to name them control variables here to avoid confusion with the term environmental in the sense of ecological.

<sup>12</sup> Strategies oriented to increase efficiency in firms tend to act reducing over-dimensioned resources, rather than expanding output.

DMUs.<sup>13</sup> The constraint  $\sum_{j=1}^n \lambda_j = 1$  is the convexity condition associated to the variable returns to scale assumption.

In the particular case of this paper, this first stage provides a set of efficiency scores for each company and year in our sample, which plays an important part in the second stage of the analysis.

## Second stage

In the second stage, we design a statistical model to explore the correlation between efficiency scores from the first stage, on the one hand, and different aspects of firm management and the macroeconomic and institutional setup where companies operate, on the other. The literature has already employed successfully this method (Bang et al., 2019; da Silva et al., 2019; Lee et al., 2009; McDonald, 2009; Miao et al., 2020; Sağlam, 2018, among others).

More formally, we estimate Eq. (6).

$$\hat{\theta}_{it} = z_{it}\beta + \epsilon_{it} \quad (6)$$

The dependent variable  $\hat{\theta}$  in Eq. (1) is made up by the efficiency scores constructed in stage 1 for each firm and year within the framework of baseline DEA and discussed in the previous subsection. The controls in  $z$  are internal and external variables, as described in “Data and variables.” Since the dependent variable and most of the regressors exhibit time and cross-section dimensions, we organize our data in a panel, with  $i$  indexing firms and  $t$  indexing time.<sup>14</sup> This practice allows to exploit heterogeneity across firms and over time and to increase the degrees of freedom in our estimation.  $\beta$  is a vector of parameters to be estimated, and  $\epsilon_{it}$  is the error term.

We have chosen a Tobit model for the functional form of Eq. (6); it seems appropriate since efficiency scores are censored by construction at the maximum value of efficiency, 1. Nonetheless, this assumption will be relaxed in “Second stage results” below. Parameters are estimated by maximum likelihood, as in previous research on DEA for oil companies (Dalei

& Joshi, 2020; Putra & Adinugraha, 2018). The model has been specified with random effects because fixed effects yield inconsistent estimates in non-linear models (Greene, 2004). Moreover, in fixed effects models, covariates with little or no variability over time, as many of our categorical variables, are not identified.

The model may be characterized more precisely by Eq. (7).

$$\begin{aligned} \hat{\theta}_{it}^* &= z_{it}\beta + u_i + e_{it} \\ \hat{\theta}_{it} &= 1 \text{ if } \hat{\theta}_{it}^* \geq 1; \hat{\theta}_{it} = \hat{\theta}_{it}^* \text{ if } 0 \leq \hat{\theta}_{it}^* \leq 1; \hat{\theta}_{it} = 0 \text{ if } \hat{\theta}_{it}^* \leq 0 \\ u_i &\sim N(0, \sigma_u^2); e_{it} \sim N(0, \sigma_e^2) \\ i &= 1, 2, \dots, n; t = 2010, 2011, \dots, 2019 \end{aligned} \quad (7)$$

where  $\hat{\theta}_{it}^*$  is the latent or unobservable efficiency,  $\hat{\theta}_{it}$  is the observable efficiency,  $u_i$  is the panel level variance component, and  $e_{it}$  is the overall variance component. In general, the unobservable and the observable efficiency coincide except in the case of the most efficient DMUs, where efficiency does not surpass the threshold of 1 by construction.

Heteroskedasticity is a frequent problem in panel data estimations. In order to handle this issue, we have performed the estimations with observed information matrix (OIM) corrected standard errors. Regressors have been included in the equation sequentially to reduce the risk of multicollinearity.

## Data and variables

### Data

Our sample is composed by around 300 companies operating in Europe. Most of them are European firms; there are also affiliates of non-European multinationals working on European soil. Our time horizon is 2010–2019.

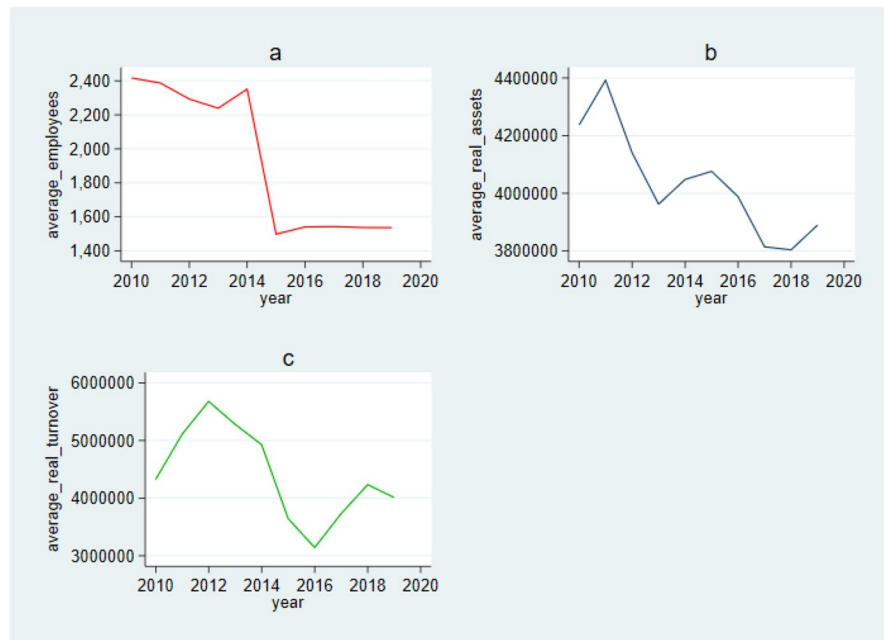
We have constructed a rich and detailed dataset with yearly information at the firm level. We have organized data in variables, which can be classified in internal and external (see Appendix 1 Table 14 for details of all the variables defined). Internal variables refer to key aspects of the business which can be controlled by the company (number of employees, total assets, turnover, size, activity, financial management, human resources management, environmental policies). External variables proxy for relevant features of the macroeconomic and institutional environment where firms operate. Internal variables vary over two dimensions, company and time (we

<sup>13</sup> The peer group or references for DMU $_j$  is made up by firms which are more efficient than  $j$  while presenting similar combinations of outputs and inputs.

<sup>14</sup> There are regressors common for all companies, varying over the time dimension  $t$  but not over the firm dimension  $i$ , as GDP growth. Country variables vary across firms but are invariant timewise.



**Fig. 2** **a** Average employees, **b** real assets, and **c** real turnover, 2010–2019. Note: Average real turnover and average real assets in thousand euros. They have been computed averaging over the companies in our sample. Source: Amadeus



get one observation per company and year). External variables are the same for all (or a subset of) companies; some change on a year-by-year basis (as the oil price); others are constant over time (as country of origin).

#### Internal variables

Internal economic variables have been constructed from the Amadeus database (Van Dijk, 2021), a rich and detailed collection of microdata disaggregated at the company level on a yearly basis. We use a set of internal variables in stage 1 and another set of internal variables in stage 2.

In the first stage of our empirical exercise, and in order to compute efficiency scores, we need to employ internal variables which proxy for inputs and output for each firm and year. We approximate the input labor with the total number of employees per company and year. The input capital is approximated by the quantity of total assets, in euros, also per company and year. In turn, output is proxied by turnover (i.e., operational revenues) in euros.<sup>15</sup> The use of these variables is in accord

with the literature in the area (Ismail et al., 2013; Song et al., 2015; Sueyoshi & Wang, 2014, among others).

Figure 2 provides a first synthetic approximation to the evolution of these variables over time for the firms in our sample (the figure is constructed with averages over companies). As Fig. 2a shows, the average number of employees diminishes abruptly until 2015 and is fairly stable ever since. Average real assets (Fig. 2b) exhibits as well a decreasing profile over most of the time horizon but grows slightly in 2018. Real turnover (Fig. 2c) increases at the beginning of the period, decreases between 2012 and 2016, and partially recovers at the end. The figure points to a mixed performance of the industry over the time horizon considered, with large oscillations in resources and production.

In the second stage of the empirical exercise, we work with other internal variables which capture different aspects of firm performance and may be potentially associated with efficiency. Some of them have been already explored by the literature, as discussed in “[Related literature and theoretical framework](#)”. Most of these variables have been compiled or constructed from the data available in Amadeus.

One important aspect is firm size. To construct size indicators, first we have distributed the oil companies in our sample in five size categories: *very big*, *big*, *medium*, *small*, and *very small*. These categories are determined

<sup>15</sup> Since DEA computes efficiency scores in relative terms, i.e., as a proportion of the scores of the best performers in the sample, computations in nominal and real terms yield the same result. Therefore, it is not necessary to deflate nominal variables for this first stage.

by the 90th, 75th, 50th and 25th percentiles of real turnover (defined as nominal turnover over the Harmonized Index of Consumer Prices, HICP, for the EU).

The size categories are as follows:

- *Very big*: for those companies whose real turnover is higher than the 90th percentile of real turnover in the sample.
- *Big*: if real turnover is less or equal than the 90th percentile and higher than the 75th percentile of real turnover in the sample.
- *Medium*: if real turnover is less or equal than the 75th percentile and higher than the 50th percentile of real turnover in the sample.
- *Small*: if real turnover is less or equal than the 50th percentile and higher than the 25th percentile of real turnover in the sample.
- *Very small*: if real turnover is less or equal than the 25th percentile of real turnover in the sample.

Next, we have created dummy (1,0) variables for each of those size categories. The dummy *very big* takes the value 1 for company  $i$  in time  $t$  if the real turnover of company  $i$  in time  $t$  is higher than the 90th percentile of average real turnover in the sample; otherwise, it takes the value 0. We have proceeded similarly for the *big*, *medium*, *small*, and *very small* categories.

Another dimension we want to explore is activity. The firms in the sample belong to a sector with NACE code 19, *Manufacture of coke and refined petroleum products*; this sector is classified in two subsectors: 19.1 (*Manufacture of coke oven products*) and 19.2 (*Manufacture of refined petroleum products*).

Ninety-four percent of the firms in our sample are refineries, and 6% are coke plants. We have captured the main activity of firms with another dummy variable, *refineries*, defined as 1 if a particular firm  $i$  is a refinery in time  $t$  and 0 otherwise.

We have defined two more internal economic variables. The variable *employee cost* captures the total cost of the payroll over the number of employees. It is a proxy of the unit cost of labor. The *solvency ratio* is defined as current assets/current liabilities. It measures the firm's capacity to confront short-term financial obligations. These two variables inform about the human resource and financial management of the firm, respectively.

### *Environmental aspects*

The reduction and control of emissions and residuals have become an important element of the production processes of firms operating with fossil energies. In recent years, many companies in the sector have started to define and monitor environmental goals and to adopt less polluting practices. To explore the connection between environmental awareness and efficiency, we have gathered additional data and constructed two indicators.

Unfortunately, and although sustainability reporting has improved very much over the last decade, it is still largely in its infancy. Unlike what happens with economic and financial variables, there is not yet a unified source of detailed, quantitative information about environmental aspects, disaggregated by company. To circumvent this problem, we have collected and compiled a data set on environmental aspects manually, checking the annual sustainability reports of groups and companies individually for each year in the sample. Ismail et al. (2013) and Bang et al. (2019) also construct indicators from sustainability reports. Furthermore, most groups do not publish environmental information disaggregated at the plant or affiliate level. Instead, they elaborate reports at the group level and disclose figures of greenhouse gases for the whole group. In the case of affiliates, we assume that affiliates and individual firms follow the directives and policies stated for the whole group, and thus, the paths in emission reductions for the whole group and for the individual affiliates exhibit the same trend over time.

Detailed and timely environmental reporting may be considered as a proxy of sound managerial green practices. We have defined a variable, *sustainability reporting*, reflecting the commitment of a company with the pursuit of sustainable goals; it is a categorical variable equaling 1 for company  $i$  in year  $t$  if its group has published a sustainability report in  $t$  and zero otherwise.<sup>16</sup>

<sup>16</sup> Companies prepare sustainability reports in accordance with Directive 2014/95/EU of the European Parliament and of the Council of October 22, 2014, amending Directive 2013/34/EU on the disclosure of non-financial and diversity information by certain large undertakings and groups (European Parliament and the Council of EU, 2014). It regulates social, environmental, and governance aspects. The directive is particularized in each country through national transpositions which may be more or less restrictive in terms of the reporting format. The EU has considered various frameworks as references for the directive and provided guidelines rather than a fully standardized format. Nonetheless, since the information standards of GRI or CDP (Carbon Disclosure Project) have wide acceptance in companies, ultimately there is some homogeneity in sustainability reports.

Many companies follow the guidelines of the EU and the GRI<sup>17</sup> and control the emissions of CO<sub>2</sub> and other substances related to the GHG effect. The path of reduction in GHG emissions is a result of the measures implemented to achieve environmental efficiency. More commitment with the environment results in more investment aimed toward less polluting technologies. We have proxied the degree of commitment with clean technologies using the information conveyed by the data of scope 1 GHG emissions, taken again manually from yearly sustainability reports.<sup>18</sup> We focus on scope 1 because those are the emissions that the firm can manage more directly. In addition, they are the lion's share of total emissions of the firms in our sample. Usually, CO<sub>2</sub> is the main component. Sueyoshi and Goto (2012b) use CO<sub>2</sub> as the undesirable input in their analysis of 19 oil firms.

We have defined a categorical variable, GHG reduction, equaling 1 for company  $i$  in year  $t$  if the decrease in GHG scope 1 emissions with respect to the previous year is higher than or equal to 1.5% (the average yearly change in GHG in our sample), and 0 otherwise. More formally, if we define "GHG emissions <sub>$i,t$</sub> " as the variable capturing the scope 1 emissions for firm  $i$  and year  $t$ , then

$$\text{GHG reduction}_{i,t} = \begin{cases} 1 & \text{if } \frac{\text{GHG emissions}_{i,t} - \text{GHG emissions}_{i,t-1}}{\text{GHG emissions}_{i,t-1}} \leq -0.015 \\ 0 & \text{otherwise} \end{cases}$$

where GHG emissions for each firm and year are proxied by GHG emissions for its group and year.

Among the most successful firms in our sample in terms of GHG reductions, there are not only large firms as ENI or BP but also smaller ones like Neste, INA, or Petrogal.

#### External variables

Several external variables help us assess the impact of the external economic and institutional

environment. The first external variable is oil price, defined as the amount of dollars per barrel of the Brent-Europe. This information comes from Federal Reserve Bank of St. Louis (2021). To express this variable in real terms, nominal prices have been deflated with the Harmonized Index of Consumer Prices (HICP) for EU 27, from Eurostat. Economic activity is proxied by the growth rate of real GDP for EU 27. It has been deflated also with the HICP. Data on GDP come from Eurostat.

Figure 3a shows the performance of the real oil price, suggesting that real oil prices have indeed fluctuated substantially over the time horizon considered. Figure 3b illustrates the trajectory of economic activity over our time horizon.

Moreover, efficiency might also differ by country and by larger geographical areas (as Western versus Eastern Europe) because of different degrees of technological diffusion, institutional environments, education of the labor force, and economic policy, among others. In order to test for these potential influences, we have defined other dummy variables which control for the country and geographical area of origin of companies. For example, the dummy Germany takes the value 1 for the companies operating in Germany and 0 for the rest. Likewise, the dummy Western Europe is 1 for companies in the area and 0 elsewhere.

## Empirical results

As already mentioned, we have organized our empirical analysis in two stages. In the first stage, we compute efficiency scores for each firm and year in our sample. In the second stage, we explore the association between the efficiency scores obtained in the first stage and several candidate variables potentially correlated with them.

First stage: efficiency, levels, and evolution

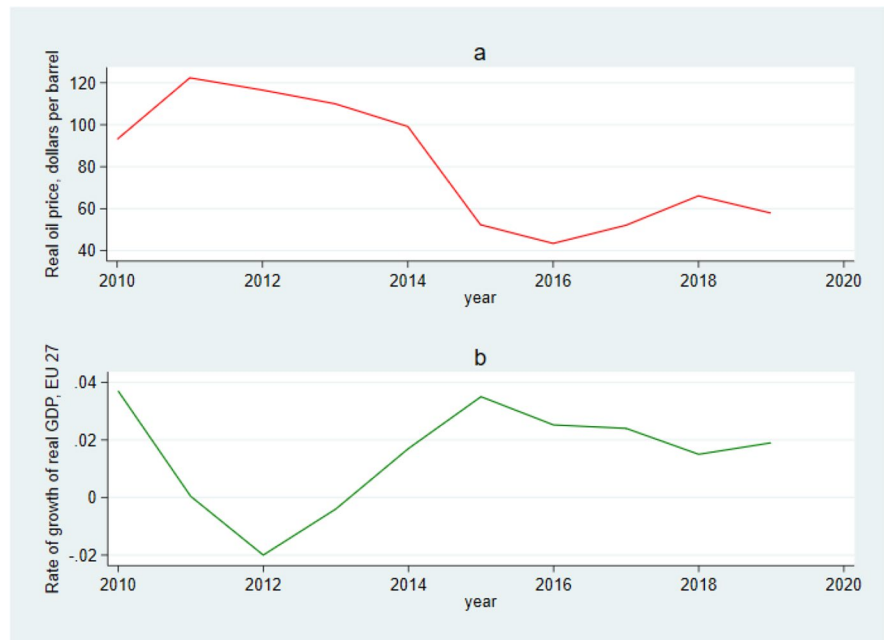
#### *Average efficiency: levels and trends*

Banker (1993) showed that, under fairly general conditions (deviations from the frontier being independent and identically distributed), the expected value of the baseline DEA estimator in the single-output case converges to the true value in large samples. Korostel'ev et al. (1995) relaxed the conditions for convergence for the

<sup>17</sup> Global Reporting Initiative: promotes standards in sustainability reporting and elaborates guidelines.

<sup>18</sup> Established in 2001, the Greenhouse Gas Protocol (GHGP) is the most used international standard for measuring GHG. GHGP classifies greenhouse emissions in three categories (GHGP, 2021): (i) Scope 1: caused directly by the activity of the organization. (ii) Scope 2: indirect emissions from energy consumption associated with the purchase of electricity, heating or cooling. (iii) Scope 3: all other indirect emissions caused along the entire value chain of an organization.

**Fig. 3** Real oil price and real GDP growth, 2010–2019. Notes: **a** Real oil price: dollars per barrel of Brent Europe, deflated with HIPC for EU 27. **b** Real GDP growth: rate of growth of real GDP for EU 27, deflated with HIPC for EU 27



single-output case and showed that, if pairs  $(x,y)$  have a strictly positive density, the baseline DEA estimator is the maximum likelihood estimator, although the rate of convergence is slow.

Our sample is relatively large (almost 300 firms and 2500 observations), and thus, it is reasonable to assume that baseline DEA provides consistent efficiency scores. Nonetheless, in the next section, we check the robustness of these results by computing as well bootstrap estimators (Daraio & Simar, 2007; Simar & Wilson, 1998, 2000, 2007).

We denote the efficiency score obtained with the baseline radial DEA procedure by  $\hat{\theta}$ . Table 2 provides summary statistics for  $\hat{\theta}$ . Its mean across firms in our sample over the period 2010–2019 is 0.27, with a standard deviation of 0.24 (Table 2). The median is 0.19. This suggests a modest level of mean efficiency in the sample: Firms could reduce their input consumption 73% on average with respect to the best performers. This result is in accord with the excess capacity and low utilization rate documented by Lukach et al. (2015).

Figure 4 displays the evolution of average efficiency over time. Average efficiency exhibits a decreasing trend until 2013 and grows thereafter, albeit it shrinks again in 2018.

#### Top performers according to efficiency

Table 3 details the top performers for the entire period, defined as companies with an average efficiency for 2010–2019 of 0.9 or more. There are three companies achieving the maximum efficiency, 1, for the whole period: Eni, Total Belgium, and Total Raffinage France. Other firms with excellent results are Tamoil, Waxoil, and Gilops, which register an average efficiency over the period larger than 0.98. Gunvor Deutschland and Repsol achieve a mean efficiency exceeding 0.95. Cepsa closes the list of top performers. There is some country heterogeneity among the top performers: Eni, Tamoil, and Waxoil are Italian, whereas Gilops is Belgian and Repsol and Cepsa are Spanish.

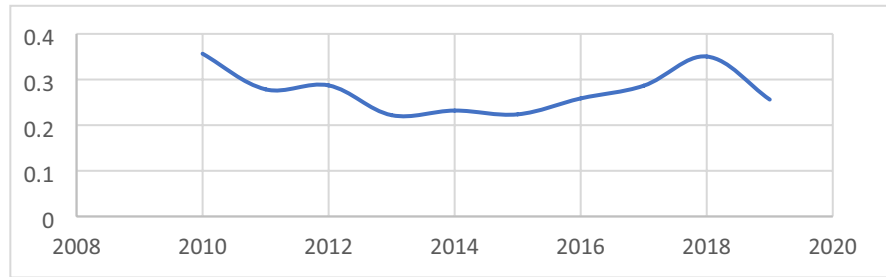
The best performing companies in terms of efficiency are rather stable over time; there is little

**Table 2** Summary statistics, baseline DEA score, 2010–2019

| Variable       | <i>N</i> | Mean | St. dev | Minimum | Maximum |
|----------------|----------|------|---------|---------|---------|
| $\hat{\theta}$ | 2484     | 0.27 | 0.24    | 0.01    | 1       |

*N* is number of observations. St. dev. is standard deviation

**Fig. 4** Average efficiency over time, 2010–2019



**Table 3** Top performers, baseline DEA efficiency, 2010–2019

| Rank | Company                 | Average efficiency, 2010–2019 |
|------|-------------------------|-------------------------------|
| 1    | Eni                     | 1                             |
| 1    | Total Belgium           | 1                             |
| 1    | Total Raffinage France  | 1                             |
| 4    | Tamoil Italy            | 0.988                         |
| 5    | Waxoil S.R.L            | 0.981                         |
| 6    | Gilops                  | 0.981                         |
| 7    | Gunvor Deutschland GmbH | 0.964                         |
| 8    | Repsol                  | 0.955                         |
| 9    | Cepsa                   | 0.914                         |

**Table 4** Efficiency by size, 2010–2019

| Size category | Mean | St. dev | N    | p value test equality of means |
|---------------|------|---------|------|--------------------------------|
| Very big      | 0.6  | 0.34    | 248  | 0***                           |
| Big           | 0.23 | 0.22    | 383  | 0***                           |
| Medium        | 0.19 | 0.14    | 641  | 0***                           |
| Small         | 0.21 | 0.15    | 610  | 0***                           |
| Very small    | 0.32 | 0.25    | 602  | 0***                           |
| Whole sample  | 0.27 | 0.24    | 2478 |                                |

\*\*\* Significant at 99%

internal mobility in the sample as far as efficiency is concerned.

*Average efficiency and internal variables*

It may be useful to conduct a first explanatory analysis of the connection of efficiency scores with various variables. Moreover, this exploration provides grounds for the second stage, where the evaluation will be carried out more thoroughly.

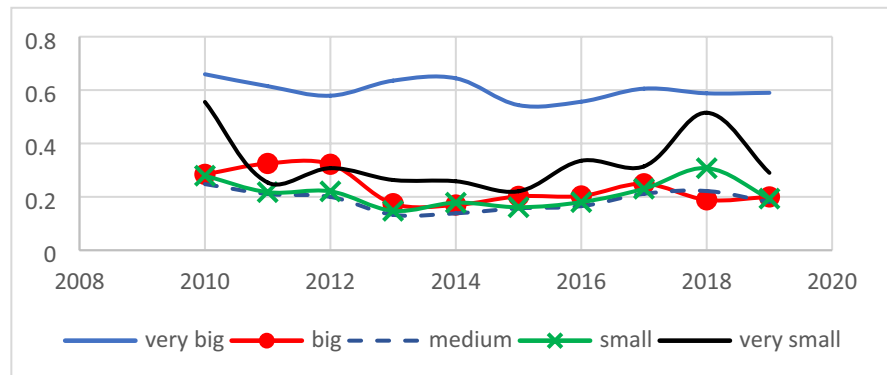
**Company size** Size may be a key determinant of the performance of a company, especially if the production function does not exhibit constant returns to scale, as it is the case in this kind of industry, capital-intensive. In order to explore the potential connection between efficiency and size, we have distributed the oil companies in our sample in five categories. These categories are determined by the 90th, 75th, 50th, and 25th percentiles of real turnover, as detailed above.

Table 4 displays summary statistics of efficiency for each category of size. The last column reports the p value of the Krishnamoorthy and Yu (2004) test of equality of means for each category and the rest of the sample.<sup>19</sup> This test should be considered with caution and just as an orientation, though, since it assumes normality.

According to Table 4, efficiency is not homogeneous across different sizes. The highest average efficiency is reported by very big companies, whose efficiency (0.6) greatly exceeds the global average efficiency (0.27). Very small companies register an average efficiency of 0.32, larger than the global average. Medium firms exhibit the lowest value of average efficiency; the average for big and small firms is also small and below the global average efficiency. The null hypothesis of equality of means is rejected for all sizes.

These results are similar to those of Ismail et al. (2013), who argue that in their sample efficiency is higher in very large and very small firms. They also agree with Mekaroonreung and Johnson (2010), who find in a sample of US oil companies

<sup>19</sup> The null hypothesis is the equality of means.

**Fig. 5** Efficiency by size, 2010–2019**Table 5** Efficiency by main activity, 2010–2019

| Activity    | Mean | St. dev | <i>N</i> | <i>p</i> value test equality of means |
|-------------|------|---------|----------|---------------------------------------|
| Refineries  | 0.29 | 0.3     | 2324     | 0***                                  |
| Coke plants | 0.19 | 0.15    | 163      |                                       |

\*\*\*Significant at 99%

that very specialized, small firms perform better than the rest of the sample. Other industries, as the biopharmaceutical, display this behavior as well (Díaz & Sanchez-Robles, 2020, 2022). This pattern is consistent with the coexistence of increasing returns to scale and niche advantages due to specialization in the production technology. The standard deviation of efficiency by size category is larger for the very big and very small companies, suggesting a higher level of heterogeneity in those categories.

Figure 5 shows the evolution of efficiency over time by size categories. Efficiency in very big firms has been rather consistently larger and steady over the period, although with a dip in 2015. Efficiency in big, medium, and small firms has also been quite stable. In very small firms, instead, efficiency has been more volatile over time. This last point also agrees with Mekaroonreung and Johnson (2010), who state that small firms are more vulnerable to oscillations in oil prices.

**Activity** Refineries exhibit a higher value of average efficiency, ten points above coke plants (Table 5), although this result should be taken with caution due to the asymmetric number of firms in either category in our sample. As above, the last

**Table 6** Efficiency and environmental commitment, 2010–2019

| Variable                   | Mean | St. dev | <i>N</i> | <i>p</i> value test |
|----------------------------|------|---------|----------|---------------------|
| Environmental reporting    | 0.47 | 0.37    | 236      | 0***                |
| No environmental reporting | 0.25 | 0.22    | 2251     |                     |
| Large GHG reduction        | 0.48 | 0.37    | 96       |                     |
| No GHG reduction           | 0.27 | 0.23    | 2391     | 0***                |

*p* value corresponds to test of equality of means

\*\*\*Significant at 99%

column of the Table displays the *p* value of the Krishnamoorthy and Yu (2004) test of equality of means for both subsectors, which can be rejected at the 99% significance level.

**Environmental commitment** We have compared the average efficiency for the companies exhibiting a higher level of environmental commitment, as captured by the two variables discussed above, and for the rest. Table 6 shows the main results. Average efficiency is 0.47 for the firms regularly elaborating and publishing their sustainability report, while it is 0.25 for those who do not. The difference is significant at conventional values. The situation is very similar for the firms reducing scope 1 emissions in a proportion equal to or larger than 1.5% per year (0.48 versus 0.27). This suggests tentatively that greater environmental awareness can be associated with larger efficiency scores in oil companies. The standard deviations of efficiency in the environmentally conscious groups (measured by GHG reduction and environmental reporting) are greater than those of the non-environmentally conscious group.

**Table 7** Efficiency by geographical area and country, 2010–2019

|                  | Mean | St. dev | <i>N</i> | <i>p</i> value test equality of means |
|------------------|------|---------|----------|---------------------------------------|
| Western Europe   | 0.28 | 0.25    | 1843     | 0.26                                  |
| Eastern Europe   | 0.27 | 0.22    | 644      |                                       |
| Belgium          | 0.40 | 0.39    | 137      | 0.0001***                             |
| Nordic countries | 0.35 | 0.27    | 98       | 0.0061***                             |
| Italy            | 0.26 | 0.22    | 1176     | 0.0007***                             |
| Spain            | 0.47 | 0.37    | 60       | 0.0001***                             |
| Romania          | 0.31 | 0.25    | 142      | 0.07**                                |
| Ukraine          | 0.24 | 0.24    | 154      | 0.06**                                |
| Whole sample     | 0.27 | 0.24    | 2478     |                                       |

Statistics for Romania and Ukraine computed with respect to the subsample of Eastern European firms

\*Significant at 90%; \*\*significant at 95%; \*\*\*significant at 99%

#### Average efficiency and external variables

**Geographical variables** We have explored the differences between firms from Western Europe (74% of the sample) and Eastern Europe (26%).<sup>20</sup> Average efficiency is one point higher in Western European countries (0.28) when compared with Eastern Europe (0.27). The *p* value of the Krishna-moorthy and Yu (2004) test of equality of means is 0.26, non-significant. Therefore, we cannot reject the hypothesis that the average efficiency in both areas is the same. Dispersion is larger for Western Europe countries. These figures suggest that the gap in efficiency between Western and Eastern Europe in this sector seems to be very small or negligible, implying that technological diffusion in Europe is rather advanced (Table 7).

Italy has the largest number of observations in the sample (1176 versus 2484), 47.34% of the total. Efficiency is slightly lower in Italy than in the whole sample (0.26 versus 0.27). The *p* value of

the test of equality of means is significant at 1% level, suggesting that the averages are indeed different. Within the Western countries, the best performers in terms of average efficiency are Spanish (0.47) and Belgian (0.40) firms.<sup>21</sup> This fact can be traced back to the existence of two very efficient multinational plants in Belgium, Total, and Esso. In turn, there are two solid oil companies (Repsol and Cepsa) operating in Spain.

Regarding Eastern Europe, average efficiency is slightly higher than the global mean in Romania (0.31), which historically has been an important oil producer since the beginning of this industry, in the mid-nineteenth century. Average efficiency is somewhat lower than the global mean for Ukraine (0.24). Although there are differences, by and large efficiency is rather homogeneous across countries in Europe. This is reasonable because of the traditional internationalization of oil companies, entailing FDI flows among countries, which has favored convergence in technology and processes. The gradual accession of many countries in the sample to the EU has sped up this trend.<sup>22</sup>

#### Second stage: econometric analysis of variables potentially correlated with efficiency

In this section, we discuss the main results from the estimation of several econometric models designed to explore the correlation between efficiency, on the one hand, and different aspects of firm management and the macroeconomic and institutional setup where companies operate, on the other.

The basis for this exploration is the estimation of Eq. (8).

$$\hat{\theta}_{it} = z_{it}\beta + \epsilon_{it} \quad (8)$$

where  $\hat{\theta}$  represents the efficiency scores constructed in stage 1 for each firm and year within the framework of baseline DEA and discussed in the previous subsection. The regressors in  $z$  are internal and external variables, as described above.

<sup>20</sup> Western Europe: Austria, Belgium, Denmark, Finland, France, Germany, Italy, Norway, Portugal, Spain, Sweden.

Eastern Europe: Bosnia, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, North Macedonia, Poland, Romania, Serbia, Slovakia, and Ukraine.

<sup>21</sup> The null hypothesis of equality of means between average efficiency of firms from these countries and from the total sample is rejected in the three cases.

<sup>22</sup> The Eastern European countries in the sample belong to EU, except Bosnia, North Macedonia, Serbia, and Ukraine.

**Table 8** Efficiency and internal variables

|                          | Model 1             | Model 2             | Model 3             | Model 4             | Model 5             | Model 6             | Model 7             |
|--------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Refineries               | 0.106*<br>(1.83)    |                     |                     |                     |                     |                     |                     |
| Very big                 |                     | 0.22***<br>(8.48)   |                     |                     |                     |                     |                     |
| Big                      |                     |                     | 0.005<br>(0.33)     |                     |                     |                     |                     |
| Medium                   |                     |                     |                     | -0.021<br>(1.63)    |                     |                     |                     |
| Small                    |                     |                     |                     |                     | -0.037***<br>(3.76) |                     |                     |
| Very small               |                     |                     |                     |                     |                     | 0.026**<br>(2.23)   |                     |
| Sustainability reporting |                     |                     |                     |                     |                     |                     | 0.031<br>(1.60)     |
| GHG reduct               |                     | 0.032*<br>(1.9)     | 0.039**<br>(2.27)   | 0.039**<br>(2.26)   | 0.039**<br>(2.28)   | 0.04**<br>(2.31)    |                     |
| Year10                   | 0.107***<br>(11.36) | 0.110***<br>(11.65) | 0.108***<br>(11.48) | 0.108***<br>(11.46) | 0.109***<br>(11.61) | 0.109***<br>(11.56) | 0.108***<br>(11.43) |
| Year13                   | -0.049***<br>(5.80) | -0.047***<br>(5.58) | -0.048***<br>(5.67) | -0.048***<br>(5.64) | -0.047***<br>(5.62) | -0.047***<br>(5.63) | -0.048***<br>(5.67) |
| Intercept                | 0.186***<br>(3.31)  | 0.262***<br>(20.06) | 0.283***<br>(19.8)  | 0.289***<br>(20.20) | 0.293***<br>(20.65) | 0.278***<br>(19.39) | 0.282***<br>(20.00) |
| <i>N</i>                 | 2487                | 2487                | 2487                | 2487                | 2487                | 2487                | 2487                |

Dependent variable is efficiency  $\hat{\theta}$  computed within the baseline DEA framework. Specification: Tobit. Estimations performed with observed information matrix (OIM) corrected standard errors to circumvent potential heteroskedasticity

*N* number observations

\*Significant at 90%; \*\*significant at 95%; \*\*\*significant at 99%

In order to correct for heteroskedasticity, we have performed the estimations with observed information matrix (OIM) corrected standard errors. Regressors have been included in the equation sequentially to reduce the risk of multicollinearity. For example, the first column of Table 8 displays the results from the estimation of the following model:

$$\text{Efficiency scores} = \beta_0 \times \text{intercept} + \beta_1 \times \text{refineries} \\ + \beta_2 \times \text{year 2010} + \beta_3 \times \text{year 2013} + \text{error term}$$

### Internal variables

Table 8 summarizes the results from estimating different versions of the baseline Eq. (8). Covariates capture internal company features. One potentially important

internal feature is activity. The dummy variable refineries are positive and significant at the 90% level in model 1, backing up the initial evidence described in “Average efficiency and internal variables,” whereby average efficiency is larger for refineries than for coke plants.

As discussed, size appears to be associated to efficiency in our sample. To test this issue further, we have included in the estimation the dummy variables for each of the five size categories defined in “Methodology.” The results in models 2–6 provide interesting insights in this regard. The dummy very big in model 2 is positive and significant at 99% level, showing that firms registering real turnover larger than the 90 percentile are more efficient, ceteris paribus. The dummy big is also positive (model 3) but not significant. The dummy capturing the medium size is negative but not significant (model 4). As conveyed by model 5, small



firms exhibit lower levels of efficiency, *ceteris paribus*, according to the negative sign of the corresponding dummy, significant at 99%. Finally, the dummy very small displays a positive and significant association with efficiency (model 6). These results confirm those of “Average efficiency and internal variables” above and are consistent with other contributions, as Ismail et al. (2013) and Mekaroonreung and Johnson (2010). They are also reasonable from an economic point of view. Large firms can exploit scale economies — associated to large investments in technology, machinery, and equipment — and reach higher levels of productivity, whereas medium and small firms cannot (Lim & Lee, 2020). Very small firms may profit from their size and enjoy niche advantages associated to specialization (Mekaroonreung & Johnson, 2010).

As suggested by models 2–6, companies achieving a reduction of greenhouse emissions of 1.5% or more in the period analyzed are more efficient. The point estimate is fairly stable in all estimations and significant at 95% (except in model 2). Companies with a more transparent reporting policy about sustainable goals and achievements do not seem to register higher levels of efficiency, *ceteris paribus*: The variable sustainability reporting is positive but not significant. Finally, the dummies for years 2010 and 2013 are positive and negative, respectively.

Human resource and financial management are key internal aspects of firms since they handle the inputs labor and capital, respectively. They have proved to be associated with efficiency in other sectors (Díaz & Sanchez-Robles, 2020, 2022).

The financial structure of companies has been captured with the first difference of the solvency ratio, defined as current assets/current liabilities. The human resource management of companies has been proxied by the unit cost of employees, defined as total costs in employees over number of employees. In order to circumvent potential endogeneity issues, we have worked with the first lag of this variable.<sup>23</sup>

Table 9 displays some results from estimations including these variables. Increases in the level of solvency, as captured by mounting ratios of current assets to current liabilities, are positively associated

with efficiency. The point estimate is stable across estimations and significant at 99% level. The variable is robust to the introduction of most regressors capturing size, which maintains their sign with respect to Table 8. Now, however, the dummies for big firms and medium size firms are positive and significant and negative and significant, respectively (models 9 and 10).

The variable capturing unit employee costs is not significant when considered for the whole sample. If the sample is divided by size, it is positive and non-significant for the very big and big firms (model 13). It is negatively and significantly associated to efficiency for medium, small, and very small firms, however, suggesting that an increase in the employee costs reduces efficiency for these companies (model 14). This is in accord with Al-Najjar and Al-Jaybajy (2012), who report that one of the reasons of inefficiency in their sample of Iraqi oil companies is the excess of workforce and its underutilization. Lukach et al. (2015) attribute as well to rising human resource costs part of the loss of competitiveness of oil companies operating in Europe.

#### *External variables*

The external economic environment where firms operate influences their performance. In order to control for this aspect, we have included in the estimations an indicator of the dynamism of economic activity, the rate of growth of real GDP for the EU 27. Some papers have uncovered a positive association between oil prices and efficiency, and thus, this is other potential external variable (Sueyoshi & Wang, 2018). Finally, we have included several country and area dummies as covariates in our baseline equations.

Table 10 summarizes the main results in this regard. Real GDP growth displays a positive and significant association with efficiency (models 15–21), suggesting a procyclical behavior of efficiency.<sup>24</sup> This effect, which has also been detected in other industries, is reasonable, since a strong level of activity reduces slackness in the use of resources (Díaz & Sanchez-Robles, 2020, 2022).

<sup>23</sup> Anyhow, we work with unit labor costs, which may be beyond the control of many firms operating in European labor markets, and thus, the risk of endogeneity is not severe.

<sup>24</sup> Time dummies have been removed in the estimations including GDP growth because of collinearity.

**Table 9** Efficiency, human resources, and financial structure

|           | Model 8             | Model 9             | Model 10            | Model 11            | Model 12            | Model 13            | Model 14            |
|-----------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Solvency  | 0.016***<br>(3.06)  | 0.015***<br>(2.96)  | 0.016***<br>(3.09)  | 0.015***<br>(2.97)  | 0.015***<br>(3.05)  |                     |                     |
| Very big  | 0.208***<br>(6.41)  |                     |                     |                     |                     |                     |                     |
| Big       |                     | 0.038**<br>(2.03)   |                     |                     |                     |                     |                     |
| Medium    |                     |                     | -0.042***<br>(2.99) |                     |                     |                     |                     |
| Small     |                     |                     |                     | -0.029**<br>(2.57)  |                     |                     |                     |
| V. small  |                     |                     |                     |                     | 0.024*<br>(1.79)    |                     |                     |
| Empl cost |                     |                     |                     |                     |                     | -0.030<br>(1.24)    | -0.149***<br>(4.54) |
| Year10    |                     |                     |                     |                     |                     | 0.063***<br>(3.04)  | 0.14***<br>(11.82)  |
| Year13    | -0.048***<br>(5.07) | -0.049***<br>(5.24) | -0.048***<br>(5.28) | -0.047***<br>(5.18) | -0.047***<br>(5.20) | -0.054***<br>(2.85) | -0.045***<br>(4.52) |
| Intercept | 0.262***<br>(20.04) | 0.274***<br>(18.82) | 0.286***<br>(21.42) | 0.281***<br>(21.29) | 0.269***<br>(20.18) | 0.406***<br>(8.96)  | 0.297***<br>(17.79) |
| <i>N</i>  | 1658                | 1658                | 1658                | 1658                | 1658                | 521                 | 1480                |

Dependent variable is efficiency  $\hat{\theta}$  computed within the baseline DEA framework. Specification: Tobit. Estimations with observed information matrix (OIM) corrected standard errors to circumvent potential heteroskedasticity

*N* number of observations

\*Significant at 90%; \*\*significant at 95%; \*\*\*significant at 99%

Models 15–21 include the real price of Brent oil as a regressor. The point estimate is positive and significant, suggesting that periods of escalating oil price are correlated with higher levels of efficiency. Two macroeconomic effects may be intertwined here. On the one hand, both GDP growth and oil prices display positive partial correlations with efficiency. The partial correlation of growth and real oil prices, however, is  $-0.5289$ , significant at 99%. It can be argued, then, that high oil prices foster efficiency directly but indirectly reduce growth and hence ultimately jeopardize efficiency. The direct effect of oil prices on efficiency partially offsets their negative impact through the indirect channel of slower growth.

We have explored the link between efficiency and the social and cultural features of the country where the affiliate or firm is located by including area and country dummies. We have constructed

an additional dummy, Western Europe, equaling 1 if the firm is located in a country belonging in this area, 0 otherwise. This dummy may shed some light on the degree of technological diffusion in the industry. For example, if firms in Western Europe are systematically more efficient than companies in the East, this will suggest that Eastern European countries have not totally converged in technology with their western counterparts and that technological diffusion has been incomplete. The dummy for Western Europe is positive but not significantly correlated with efficiency (model 20), suggesting that Western European firms do not systematically allocate their resources more efficiently than those in Eastern Europe and thus that technological catch up is almost total.

The dummy for Italy is negative and significant (model 20); the categorical variable for Belgium, instead, is positively and significantly correlated with

**Table 10** Efficiency and external variables

|           | Model 15            | Model 16            | Model 17            | Model 18            | Model 19            | Model 20           | Model 21            |
|-----------|---------------------|---------------------|---------------------|---------------------|---------------------|--------------------|---------------------|
| Growth    | 0.729***<br>(3.2)   | 0.694***<br>(3.04)  | 0.697***<br>(3.05)  | 0.688***<br>(3.02)  | 0.693***<br>(3.04)  | 0.696***<br>(3.04) | 0.696***<br>(3.05)  |
| Oil price | 0.046***<br>(3.43)  | 0.045***<br>(3.34)  | 0.045***<br>(3.38)  | 0.045***<br>(3.38)  | 0.046***<br>(3.43)  | 0.045***<br>(3.35) | 0.045***<br>(3.34)  |
| Very big  | 0.227***<br>(8.5)   |                     |                     |                     |                     |                    |                     |
| Big       |                     | 0.010<br>(0.62)     |                     |                     |                     |                    |                     |
| Medium    |                     |                     | -0.026**<br>(2.01)  |                     |                     |                    |                     |
| Small     |                     |                     |                     | -0.036***<br>(3.50) |                     |                    |                     |
| V. small  |                     |                     |                     |                     | 0.025**<br>(2.06)   |                    |                     |
| Italy     |                     |                     |                     |                     |                     | -0.057*<br>(1.66)  |                     |
| W. Europe |                     |                     |                     |                     |                     | 0.045<br>(1.19)    |                     |
| Belgium   |                     |                     |                     |                     |                     |                    | 0.174***<br>(2.65)  |
| Intercept | 0.219***<br>(11.89) | 0.242***<br>(12.51) | 0.250***<br>(12.89) | 0.253***<br>(13.08) | 0.237***<br>(12.16) | 0.237***<br>(8.03) | 0.236***<br>(12.15) |
| N         | 2487                | 2487                | 2487                | 2487                | 2487                | 2487               | 2487                |

Dependent variable is efficiency  $\hat{\theta}$  computed within the baseline DEA framework. Specification: Tobit. Estimations with observed information matrix (OIM) corrected s.e. to account for potential heteroskedasticity

\*Significant at 90%;  
\*\*significant at 95%;  
\*\*\*significant at 99%

efficiency, implying that firms in that country are ceteris paribus more efficient (model 21). The variables capturing size exhibit similar signs and point estimates than before.

**Sensitivity analysis**

Efficiency scores

We have performed a sensitivity analysis of our results in two main dimensions. First, we have checked the adequacy of the efficiency scores computed in the framework of baseline DEA; moreover, we have explored the robustness of the correlation between the efficiency scores and the control variables employed in the second stage.

The baseline DEA framework does not consider explicitly the possibility of measurement errors or sample bias in the data. Typically, the efficient frontier and the underlying data generating process (DGP) of the efficiency scores are unknown. In a series of

influential papers, Simar and Wilson (1998, 2000, 2007) and Daraio and Simar (2007) design some bootstrapping tools which, by means of repeated sampling, provide approximations to the unknown distribution of the DGP of efficiency and enable the computation of bias-corrected scores. This methodology can also be used to compute standard errors and confidence intervals of the efficiency scores at a specific significance level. The size of our sample suggests that the scores computed within the baseline DEA framework may be regarded as consistent, according to Korostelëv et al. (1995). Nonetheless, we have compared them with a new set of efficiency scores,  $\hat{\theta}_b$ , estimated by means of the Simar and Wilson (1998) bootstrap methodology (henceforth SW). We have used the same data and variables as in our exercise employing baseline DEA and set up the computation under variable returns to scale and input orientation as well.<sup>25</sup>

<sup>25</sup> The software employed for this exercise is the DEA toolbox for Matlab (Álvarez et al., 2020).

**Table 11** Descriptive statistics, DEA and bootstrap efficiency

| Variable           | Mean | Median | St. Dev | Min  | Max  |
|--------------------|------|--------|---------|------|------|
| $\hat{\theta}$     | 0.27 | 0.191  | 0.24    | 0.01 | 1    |
| $\hat{\theta}_b$   | 0.22 | 0.16   | 0.17    | 0.01 | 0.92 |
| $\hat{\theta}_b^U$ | 0.26 | 0.193  | 0.22    | 0.01 | 0.99 |
| $\hat{\theta}_b^L$ | 0.19 | 0.15   | 0.15    | 0.01 | 0.83 |

$\hat{\theta}$  DEA efficiency,  $\hat{\theta}_b$  bootstrap efficiency,  $\hat{\theta}_b^U$  upper bound 95% confidence interval,  $\hat{\theta}_b^L$  lower bound 95% confidence interval

Table 11 compares some descriptive statistics of the efficiency scores obtained by the baseline DEA and the bootstrap models,  $\hat{\theta}$  and  $\hat{\theta}_b$ , respectively.  $\hat{\theta}_b^U$  and  $\hat{\theta}_b^L$  are the upper and lower bounds of the 95% confidence interval. Means and standard deviations computed by bootstrap methods are slightly lower. This is reasonable for two main reasons: first, because bootstrap repeated sampling acts as a smoothing procedure which tends to give less preponderance to observations with extreme values and more to those which are close to others (and hence are more likely to be included several times in different resamplings); second, because the baseline DEA tends to compute upward biased efficiency scores. Anyway, the means of  $\hat{\theta}$  and  $\hat{\theta}_b$  are rather close. The difference between them, 0.05, is below the figure reported in other studies, as López-Penabad et al. (2020). Hanrui and Xun (2011) find a difference of 0.02 between the original DEA and the Bootstrap DEA scores, which is not far from our result.

The mean of the DEA estimator is above the upper bound of the 95% confidence interval for just one point (0.27 versus 0.26). The Mann–Whitney test cannot reject the null hypothesis that the distributions of the DEA and the upper bound of the 95% confidence level are equal ( $p$  value 0.43).

Since the distributions for the efficiency scores are typically skewed, it seems reasonable to compare the medians as well. According to Table 11, the medians for  $\hat{\theta}$  and  $\hat{\theta}_b^U$  are quite similar. In fact, the median efficiency computed by DEA lies within the 95% confidence interval of the median of the bootstrap efficiency.

We can conclude that, while there is not total convergence between the DEA and the bootstrap estimators, the DEA estimator does converge in distribution to the upper bound of the 95%

**Table 12** Partial correlations

| Variable           | $\hat{\theta}$ | $\hat{\theta}_b$ | $\hat{\theta}_b^U$ | $\hat{\theta}_b^L$ |
|--------------------|----------------|------------------|--------------------|--------------------|
| $\hat{\theta}$     | 1***           |                  |                    |                    |
| $\hat{\theta}_b$   | 0.95***        | 1***             |                    |                    |
| $\hat{\theta}_b^U$ | 0.98***        | 0.99***          | 1***               |                    |
| $\hat{\theta}_b^L$ | 0.95***        | 0.99***          | 0.99***            | 1***               |

$\hat{\theta}$  DEA efficiency,  $\hat{\theta}_b$  bootstrap efficiency,  $\hat{\theta}_b^U$  upper bound 95% confidence interval,  $\hat{\theta}_b^L$  lower bound 95% confidence interval

\*\*\*Significant at 99% level

confidence interval estimator. Moreover, medians do converge. This suggests that the set of baseline DEA efficiency scores is a reasonable approximation to the true scores for our sample.

The last column informs about maximum values. As the Table conveys, the baseline model places some DMUs on the frontier, with efficiency of 1, while the bootstrap method does not: in the SW framework DMUs may theoretically approach an efficiency score of 1, but this occurs with 0 probability.

Table 12 informs about partial correlations between the estimators. The correlation coefficient between the mean efficiencies of the DEA and bootstrap models is high, 0.95. The correlation between the DEA scores and the upper bound of the confidence interval is 0.97, larger than the correlation between  $\hat{\theta}$  and  $\hat{\theta}_b^L$ . The higher correlation between  $\hat{\theta}$  and  $\hat{\theta}_b^U$  is to be expected due to the upward bias of baseline DEA scores. Again, we attribute the small discrepancy among them to the different treatment of efficient DMUs in either methodology. It is well known that the bootstrap estimator does not perform equally well in the proximity of the efficient frontier (Simar & Wilson, 1998). Nonetheless, we intend to explore this issue more thoroughly in future research.

Figure 6 in Appendix 2 compares the two distributions of efficiency scores as computed by the DEA and the bootstrap methodology. The interquartile range is somehow larger for the baseline model. This is also reasonable since the SW estimator does not assign scores of 1 in practice, while the baseline does. Figure 7 in Appendix 3 displays the evolution of the means of both sets of efficiency scores over time and shows that the time pattern is almost identical.

By and large, and although we have not found convergence in means, we may conclude that there is a remarkable level of similarity between the results from the baseline DEA and the bootstrap estimators, which backs up the results discussed in “[Empirical results.](#)”

## Second stage results

Another drawback of baseline DEA is the potential presence of serial correlation in the efficiency scores. Simar and Wilson (2007) design a strategy in order to handle this issue. They propose a double bootstrap procedure which provides consistent results. Basically, the first bootstrap computes the bias-corrected efficiency scores, as discussed in “[Efficiency scores.](#)” The second bootstrap obtains a set of estimates of the first and second moments of the parameters of interest in a truncated regression of efficiency scores on the control variables of the following form, in matrix notation:

$$\hat{\theta}_b = z\delta + \xi \quad (9)$$

where the error term  $\xi$  corresponds to a normal distribution with left truncation. Maximum likelihood estimation in this setting provides consistent estimators of  $\delta$ . A common assumption about the DGP in Eq. (7) or (9) is that the error term follows a censored distribution, since efficiency scores cannot exceed 1 by construction, and hence, the appropriate specification for the second stage is a Tobit model. This was our hypothesis in “[Second stage: econometric analysis of variables potentially correlated with efficiency.](#)” Simar and Wilson (1998, 2007) argue, however, that the correct DGP is not censored but truncated because the upper limit of 1 is a true feature of the distribution and not an artifact of the computation procedure.

This is a controversial point. Some researchers still prefer the Tobit model (Greene, 2003). McDonald (2009) argues that the dilemma between the censored and the truncated models is primarily methodological and conceptual, and that in practical applications the Tobit model provides more robust results. The clarification of this issue is beyond the scope of this paper. Anyway, we choose to check the robustness of the second stage results to the assumption of a Tobit versus a truncated distribution by using the SW approach as well.

Notice that there are three main differences between the baseline DEA (Eq. 7) and the SW (Eq. 9) approaches: (i) the dependent variable, constructed by the DEA baseline estimation in the first case and the bootstrap replications in the second; (ii) the distribution of the error term, censored in Eq. (7) and truncated in Eq. (9); and (iii) the correction for serial correlation, implemented in Eq. (9) but not in Eq. (7).

Table 13 displays the main results obtained with the SW procedure. Qualitative results are roughly the same in the Tobit and in the SW models, although the SW methodology seem to provide more efficient estimations (Simar & Wilson, 2007). By and large, signs and order of magnitudes are similar in both cases.

There are a few differences, though. The variable *big* is positive in the Tobit estimations, non-significant in two cases (models 3 and 16; Tables 8 and 10) and significant in another (model 9; Table 9); it is negative and significant at the 95% significance level in the SW specification (model 23). Since the mean efficiency for *big* firms is 0.23, below the global mean (Table 4), we conclude that in this case, the SW result is more plausible. Another difference has to do with the external variables capturing general conditions for oil firms, growth, and oil prices. They are positive but non-significant in the SW specification (model 20), in line with Putra and Adinugraha (2018).

By and large, the Simar-Wilson methodology provides similar results to those obtained with the baseline DEA and the Tobit estimation and presented in “[Empirical results.](#)” This, in our view, provides robustness both to the efficiency scores computed in the first stage and to the association between efficiency and other variables found in the second stage.

## Concluding remarks and policy recommendations

This paper analyzes the level and evolution of efficiency in oil companies operating in Europe. It also explores variables potentially associated with efficiency, both internal and external. Our sample is encompassed by almost 300 firms over the period 2010–2019. The main insights from our empirical exercise are as follows:

1. The average level of efficiency in the sample is relatively low, 0.27, and has exhibited a decreasing trend in the period 2010–2019. This is in

**Table 13** Second stage results, Simar and Wilson model

|               | Model 22           | Model 23           | Model 24            | Model 25            | Model 26           | Model 27           | Model 28            | Model 29            | Model 30           |
|---------------|--------------------|--------------------|---------------------|---------------------|--------------------|--------------------|---------------------|---------------------|--------------------|
| GHG reduction | 0.032<br>(0.81)    | 0.248***<br>(5.77) | 0.214***<br>(5.61)  | 0.219***<br>(5.28)  | 0.278***<br>(6.72) |                    |                     |                     |                    |
| Year10        | 0.155***<br>(5.42) | 0.186***<br>(5.62) | 0.182***<br>(5.64)  | 0.182***<br>(5.51)  | 0.193***<br>(5.90) | 0.183***<br>(5.79) |                     | 0.180***<br>(5.98)  |                    |
| Year13        | 0.017<br>(0.57)    | 0.022<br>(0.64)    | 0.028<br>(0.85)     | 0.025<br>(0.75)     | 0.029<br>(0.85)    | 0.029<br>(0.91)    |                     | -0.013<br>(0.42)    |                    |
| Very big      | 0.32***<br>(11.48) |                    |                     |                     |                    |                    | 0.444***<br>(12.12) |                     |                    |
| Big           |                    | -0.12**<br>(4.13)  |                     |                     |                    |                    |                     |                     |                    |
| Medium        |                    |                    | -0.192***<br>(7.64) |                     |                    |                    |                     |                     |                    |
| Small         |                    |                    |                     | -0.102***<br>(4.15) |                    |                    |                     |                     |                    |
| V. small      |                    |                    |                     |                     | 0.146***<br>(6.68) | 0.163***<br>(7.41) |                     |                     |                    |
| Sust. report  |                    |                    |                     |                     |                    | 0.251***<br>(8.15) |                     |                     |                    |
| Refineries    |                    |                    |                     |                     |                    |                    | 0.083**<br>(2.03)   |                     |                    |
| Solvency      |                    |                    |                     |                     |                    |                    | 0.035**<br>(2.17)   |                     |                    |
| Empl. cost    |                    |                    |                     |                     |                    |                    |                     | -0.291***<br>(7.11) |                    |
| Growth        |                    |                    |                     |                     |                    |                    |                     |                     | 0.05<br>(0.05)     |
| Oil price     |                    |                    |                     |                     |                    |                    |                     |                     | 0.046<br>(0.90)    |
| Italy         |                    |                    |                     |                     |                    |                    |                     |                     | -0.07***<br>(2.87) |
| W. Europe     |                    |                    |                     |                     |                    |                    |                     |                     | 0.05*<br>(1.84)    |
| Intercept     | -0.018<br>(0.75)   | -0.020<br>(0.68)   | 0.023<br>(0.86)     | -0.013<br>(0.44)    | -0.069**<br>(2.14) | -0.077**<br>(2.51) | -0.043<br>(0.90)    | 0.152***<br>(6.95)  | -0.08<br>(1.17)    |
| <i>N</i>      | 2482               | 2482               | 2482                | 2482                | 2482               | 2482               | 1653                | 1465                | 2482               |

Specification: Simar-Wilson second stage. Dependent variable is efficiency obtained by bootstrap. Models 28 and 30 exclude time dummies to avoid multicollinearity

\*Significant at 90%; \*\*significant at 95%; \*\*\*significant at 99%

accord with our first hypothesis. Top performers (Guvnor, Neste, Total, ENI) are quite stable over time.

- Efficiency is positively associated with size, activity, financial stability, controlled labor costs, and environmental commitment. Very large firms

(with turnover higher than the 90% of the distribution) exhibit higher levels of efficiency, *ceteris paribus*, suggesting the presence of scale economies in the industry. The existence of increasing returns is reasonable due to the requirements in terms of technology, machines and equipment

which entail considerable volumes of assets and fixed costs for companies. Very small firms, however, also perform well in term of efficiency, implying the presence of niche competitive advantages associated to specialization. While we hypothesized beforehand the presence of increasing returns, the sound performance of very small firm suggesting niche advantages was unexpected. Mounting employee costs jeopardize efficiency, but only for medium, small, and very small firms, and not for the firms in the top quartile.

3. From the point of view of environmental variables, large reductions in greenhouse emissions (i.e., above the sample mean) are positively correlated with efficiency. These results suggest that there is not a trade-off between environmental and operational efficiency. Instead, they mutually enhance each other. Our evidence, however, does not suggest that the mere disclosure of information about sustainability goals and performance (i.e., the publication of a sustainability report by the company) is enough to impact efficiency significantly.
4. In terms of macroeconomic variables, results suggest that efficiency in oil companies is procyclical; it is positively correlated with oil prices as well, although these two findings are not robust. According to our empirical analysis, the technological catch up of Western European firms by their Eastern European counterparts seems to be almost complete. There are not substantial differences in efficiency across the majority of countries.
5. We have complemented the results from the DEA baseline with those obtained with the Simar-Wilson methodology in both stages. There is not a total convergence in the means of both sets of efficiency scores, but the medians fall in the same 95% confidence interval. Both sets of efficiency scores display a high correlation (95–97%) and a similar time pattern. The basic messages of the baseline second stage carry over to the Simar-Wilson methodology, which nonetheless provide more efficient estimates.

These findings have practical implications. Companies with low levels of efficiency are not sustainable in the long run and require careful analysis so that the best strategy for the future is designed and implemented. This strategy will not be the same for all companies. Some of them may still be able

to recover through investment, rationalization, and modernization; in other cases, survival will not be possible. Moreover, stakeholders and firm managers should keep in mind that, according to our results, the consolidation of the industry by means of mergers and acquisitions is feasible in the future. This process would bring companies closer to their optimal size, where increasing returns can be exploited (Lim & Lee, 2020). Medium size and small firms with poor performances are particularly at risk, especially if their managers do not strive to reduce inefficiencies and boost productivity.

Our analysis provides some insights which may be useful to orientate the industrial policy for this sector; this discussion is especially relevant now in Europe because of the launching of the Next Generation Plan, which provides funds for the transformation of the economy.

Our exercise has shown that average efficiency in the European oil sector is modest and declining; low levels of efficiency and overcapacity, in turn, are damaging the competitiveness of this industry. One avenue for the correction of these shortcomings is to allow the gradual reallocation of resources from low efficiency to high efficiency companies. Barriers to exit (especially from local authorities) which hamper the closure of inefficient plants because of political reasons should be gradually suppressed (Nivard & Kreijkes, 2017). Consolidation of this industry through mergers and acquisitions is also desirable and should not be prevented by policymakers. Legislation which entails higher bureaucratic, fiscal, and labor costs for large firms and hinder company growth should be discouraged as well (Hsieh & Klenow, 2014). These processes could increase global productivity in this sector and the rest of the economy and foster economic growth (Acemoglu et al., 2018; Aghion & Howitt, 1992; Hsieh & Klenow, 2009; Lentz & Mortensen, 2008). Instead, measures like indiscriminate subsidies for firms in the sector could maintain underperformers artificially active and perpetuate inefficiencies.

In parallel, the modest levels of efficiency in the industry suggest that the growth in oil prices should not be automatically passed through to fuels and other consumer products. They can be absorbed by the producers by means of increases in efficiency through rationalization, better resource reallocation, and process innovation. This implies that policies intended to subsidize the price of oil products to final

consumers may not be the best solution and should be carefully considered before implementation.

Finally, policymakers should assess carefully the upsides and downsides of new and existing regulations for this industry, especially about the labor market, since excessive employee cost seems to be one reason underlying poor efficiency. In particular, European authorities should avoid measures which introduce further rigidity in labor markets.

The industrial policy for this sector, in any case, must be very prudently designed and implemented. Measures should be conceived prioritizing the long run over the short run and economic goals over political goals.

Because of the crucial role of the oil sector in the supply chain of most products and services, an increase in the efficiency of this industry will spillover to the rest of the economy, favoring its smooth operation. The restructuring and consolidation of the sector may also be convenient from the macroeconomic point of view. A more solid and productive oil industry may help decrease the

dependence of Europe from oil imports, something particularly desirable in a scenario of geopolitical risks.

The absence of environmental data for a part of our sample is the main limitation of our paper. In future research, we intend to circumvent this issue and further explore the association between efficiency, environmental commitment, and other indicators of firm management. One promising avenue is to compare European and non-European firms, assessing the impact of environmental regulations in each case.

**Acknowledgements** Sanchez-Robles gratefully acknowledges project ECO2017-83183-P, Spanish Ministry of Science and Innovation.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

**Declarations**

**Conflict of interest** The authors declare no competing interests.

## Appendix 1

**Table 14** Definition of variables

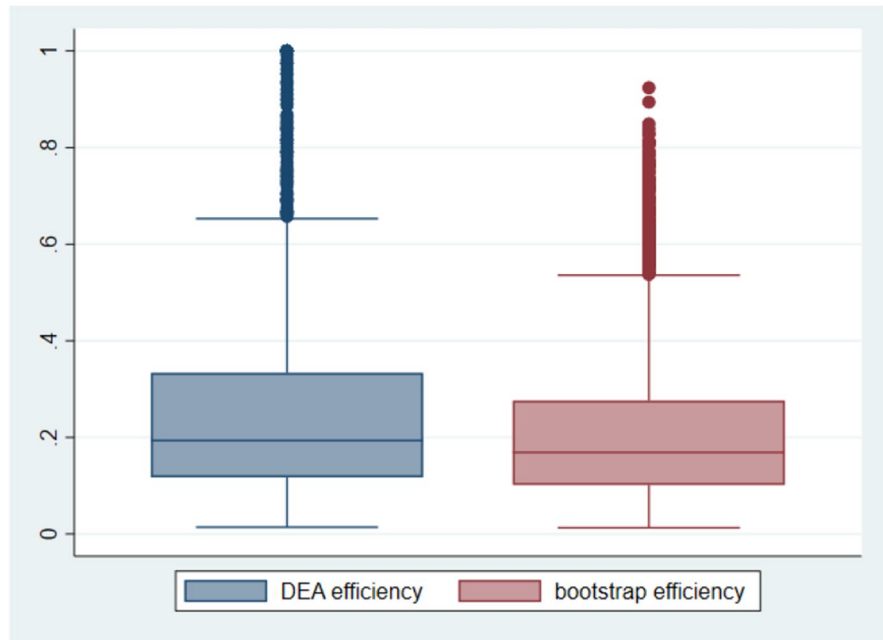
| Variables                | Definition  |
|--------------------------|---|
| Turnover                 | Operating revenue = net sales + other operating revenues + stock variations   |
| Number of employees      | Total number of employees included in the company's payroll   |
| Total assets             | Total assets (fixed assets + current assets)  |
| Deflator                 | HIPC, harmonized index of consumer prices for EU 27 (source: Eurostat)  |
| Employee cost            | Total cost of employees / number of employees   |
| Solvency ratio           | Current assets / current liabilities  |
| Very big                 | Dummy equal to 1 if real turnover > 90 percentile of its distribution, 0 otherwise  |
| Big                      | Dummy equal to 1 if 75 percentile < real turnover ≤ 90 percentile, 0 otherwise  |
| Medium                   | Dummy equal to 1 if 50 percentile < real turnover ≤ 75 percentile, 0 otherwise  |
| Small                    | Dummy equal to 1 if 25 percentile < real turnover ≤ 50 percentile, 0 otherwise  |
| Very small               | Dummy equal to 1 if real turnover ≤ 25 percentile, 0 otherwise  |
| Oil prices               | Price of barrel of Brent-Europe oil in US dollars. Deflated with HIPC for EU 27   |
| Growth                   | Rate of change of real GDP of EU 27, deflated with HIPC for EU 27   |
| Sustainability reporting | Dummy equal to 1 in $t$ if the firms published a sustainability report in $t$ , 0 otherwise. (Source: company reports)                        |
| GHG reduction            | Dummy equal to 1 if the reduction in scope 1 gas emissions has been equal or greater than 1.5% (source: own elaboration from company reports) |
| Refineries               | Dummy equal to 1 if NACE activity of company is 1920, 0 otherwise   |
| Western Europe           | Dummy equal to 1 if country belongs to Western Europe, 0 otherwise  |
| Belgium                  | Dummy equal to 1 if country is Belgium, 0 otherwise   |
| Spain                    | Dummy equal to 1 if country is Spain, 0 otherwise   |
| Italy                    | Dummy equal to 1 if country is Italy, 0 otherwise   |

Source is Amadeus if it is not detailed otherwise



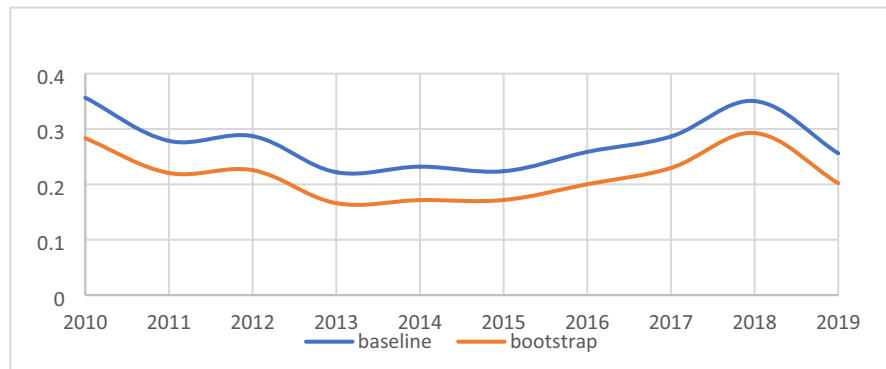
### Appendix 2

**Fig. 6** Distributions of efficiency scores, baseline, and bootstrap models



### Appendix 3

**Fig. 7** Evolution over time of efficiency scores, baseline, and bootstrap models



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Acemoglu, D., Acikcigit, U., Alp, H., Bloom, N., & Kerr, W. (2018). Innovation, reallocation, and growth. *American Economic Review*, 108(11), 3450–3491.
- Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2), 323–351.
- Aigner, D., Lovell, C., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, 21–37.
- Al-Najjar, S. M., & Al-Jaybaj, M. A. (2012). Application of data envelopment analysis to measure the technical efficiency of oil refineries: A case study. *International Journal of Business Administration*, 3(5), 64.
- Álvarez, I., Barbero, J., & Zofío Prieto, J. (2020). A DEA toolbox for MATLAB. *Journal of Statistical Software (online)*, 95, 3.
- Atris, A. M., & Goto, M. (2019). Vertical structure and efficiency assessment of US oil and gas companies. *Resources Policy*, 63, 101437.
- Bang, Y. Y., Lee, D. S., & Lim, S. R. (2019). Analysis of corporate CO<sub>2</sub> and energy cost efficiency: The role of performance indicators and effective environmental reporting. *Energy Policy*, 133, 110897.
- Banker, R. D. (1993). Maximum likelihood, consistency and DEA analysis: A statistical foundation. *Management Science*, 39, 1265–1273.
- Bevilacqua, M., & Braglia, M. (2002). Environmental efficiency analysis for ENI oil refineries. *Journal of Cleaner Production*, 10(1), 85–92.
- Berman, E., & Bui, L. T. M. (2001). Environmental regulation and productivity: Evidence from oil refineries. *The Review of Economics and Statistics*, 83, 498–510.
- Bernanke, B. S., & Gertler, M. (1995). Inside the black box: The credit channel of monetary policy transmission. *Journal of Economic Perspectives*, 9(4), 27–48.
- Boyd, G. A., & McClelland, J. D. (1999). The impact of environmental constraints on productivity improvement in integrated paper plants. *Journal of Environmental Economics and Management*, 38, 121–142.
- Charnes, A., & Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, 9, 181–186.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of DMUs. *European Journal of Operational Research*, 2, 429–444.
- Cuñado, J., & Pérez de Gracia, F. (2003). Do oil price shocks matter? Evidence for some European countries. *Energy Economics*, 25(2), 137–154.
- Dalei, N. N., & Joshi, J. M. (2020). Estimating technical efficiency of petroleum refineries using DEA and Tobit model: An India perspective. *Computers & Chemical Engineering*, 142, 107047.
- Daraio, C., & Simar, L. (2007). Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach. *Journal of Productivity Analysis*, 28(1), 13–32.
- Da Silva, A. V., Costa, M. A., Lopes, A. L. M., & Do Carmo, G. M. (2019). A close look at second stage data envelopment analysis using compound error models and the Tobit model. *Socio-Economic Planning Sciences*, 65, 111–126.
- Díaz, R. F., & Sanchez-Robles, B. (2020). Non-parametric analysis of efficiency: An application to the pharmaceutical industry. *Mathematics*, 8(9), 1522.
- Díaz, R. F., & Sanchez-Robles, B. (2022). Efficiency in the CRO industry, 2010–2019. *Economic Research-Ekonomska Istraživanja*, forthcoming.
- Eller, S. L., Hartley, P. R., & Medlock, K. B. (2011). Empirical evidence on the operational efficiency of National Oil Companies. *Empirical Economics*, 40(3), 623–643.
- Emrouznejad, A., & Yang, G. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, 61, 4–8.
- European Central Bank. (2018). Economic Bulletin Issue 7.
- European Parliament and Council of EU. (2014). Directive 2014/95 / EU of the EP and of the Council of October 22, 2014, which modifies Directive 2013/34/EU on the disclosure of non-financial information and on diversity by certain large companies and groups.
- Eurostat. (2022). <https://ec.europa.eu/eurostat>.
- Federal Reserve Bank of St. Louis. (2021). Federal Reserve Economic Data. <https://fred.stlouisfed.org>.
- Friedman, B. L. (1992). Productivity in crude oil and natural gas production. *Monthly Labor Review*, 115, 9.
- Førsund, F. R. (2018). Economic interpretations of DEA. *Socio-Economic Planning Sciences*, 61, 9–15.
- Greene, W. H. (2003). *Econometric analysis*. Prentice-Hall.
- Greene, W. (2004). The behavior of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Economic Journal*, 7, 98–119.
- Greenhouse Gas Protocol (GHG). (2021). <https://ghgprotocol.org>.
- Hanrui, B. A. O., & Xun, A. N. (2011). Reliability test on oil field efficiency with DEA. *Energy Procedia*, 5, 1473–1477.
- Hartley, P. R., & Medlock, K. B., III. (2013). Changes in the operational efficiency of national oil companies. *The Energy Journal*, 34, 2.

- Hsieh, C. T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, *124*(4), 1403–1448.
- Hsieh, C. T., & Klenow, P. J. (2014). The life cycle of plants in India and Mexico. *The Quarterly Journal of Economics*, *129*(3), 1035–1084.
- International Energy Agency. (2022). <https://www.iea.org/>
- International Monetary Fund. (2022). World Economic Outlook update, July.
- Ismail, Z., Tai, J. C., Kong, K. K., Law, K. H., Shirazi, S. M., & Karim, R. (2013). Using DEA in comparing the environmental performance and technical efficiency of selected companies in their global petroleum operations. *Measurement*, *46*, 3401–3413.
- Korostelev, A. P., Simar, L., & Tsybakov, A. B. (1995). Efficient estimation of monotone boundaries. *The Annals of Statistics* 476–489.
- Krishnamoorthy, K., & Yu, J. (2004). Modified Nel and Van der Merwe test for the multivariate Behrens-Fisher problem. *Statistics and Probability Letters*, *66*(2), 161–169.
- Lee, C., Lee, J., & Kim, T. (2009). Innovation policy for defense acquisition and dynamics of productive efficiency: A DEA application to the Korean defense industry. *Asian Journal of Technology Innovation*, *17*(2), 151–171.
- Lentz, R., & Mortensen, D. T. (2008). An empirical model of growth through product innovation. *Econometrica*, *76*(6), 1317–1373.
- Lim, C., & Lee, J. (2020). An analysis of the efficiency of the oil refining industry in the OECD countries. *Energy Policy*, *142*, 111491.
- López-Penabad, M. C., Maside-Sanfiz, J. M., TorrellesManent, J., & Iglesias-Casal, A. (2020). Application of the DEA double bootstrap to analyze efficiency in Galician sheltered workshops. *Sustainability*, *12*(16), 6625.
- Lukach, R., Marschinski, R., Bakhtieva, D., Mraz, M., Temurshoev, U., Eder, P., & Sancho, L. D. (2015). EU petroleum refining fitness check: Impact of EU legislation on sectoral economic performance. JRC Science for Policy Report, European Commission.
- Mariano, J., & La Rovere, E. (2017). *Environmental impacts of the oil industry*. LAP Lambert Academic Publishing.
- McDonald, J. (2009). Using LS and Tobit in second stage DEA efficiency analyses. *European Journal of Operational Research*, *197*(2), 792–798.
- Meeusen, W., & Broeck, J. V. D. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, *18*, 435–444.
- Mekaroonreung, M., & Johnson, A. L. (2010). Estimating the efficiency of American petroleum refineries under varying assumptions of the disposability of bad outputs. *International Journal of Energy Sector Management*, *4*(3), 356–398.
- Miao, C. L., Duan, M. M., Sun, X. X., & Wu, X. Y. (2020). Safety management efficiency of China's coal enterprises and its influencing factors—based on the DEA-Tobit two-stage model. *Process Safety and Environmental Protection*, *140*, 79–85.
- Mohaddes, K., & Pesaran, M. H. (2016). Country-specific oil supply shocks and the global economy: A counterfactual analysis. *Energy Economics*, *59*, 382–399.
- Nivard, M., & Kreijkes, M. (2017). The European Refining sector: A diversity of markets. Clingendael International Energy Programme (CIEP). Available online at [https://www.clingendaelenergy.com/inc/upload/files/CIEP\\_paper\\_](https://www.clingendaelenergy.com/inc/upload/files/CIEP_paper_).
- Ohene-Asare, K., Turkson, C., & Afful-Dadzie, A. (2017). Multinational operation, ownership and efficiency differences in the international oil industry. *Energy Economics*, *68*, 303–312.
- Putra, I. S., & Adinugraha, D. (2018). Factors affecting efficiency of oil and gas companies in Indonesia two stage DEA analysis. Data Envelopment Analysis and Performance Measurement: Recent 93.
- Raith, A., Rouse, P., & Seiford, L.M. (2019). *Benchmarking using data envelopment analysis: Application to stores of a post and banking business*. Multiple Criteria Decision Making and Aiding. (pp. 1–39).Springer.
- Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*, 5, Part 2, S71-S102.
- Sağlam, Ü. (2018). A two-stage performance assessment of utility-scale wind farms in Texas using data envelopment analysis and Tobit models. *Journal of Cleaner Production*, *201*, 580–598.
- Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, *44*(1), 49–61.
- Simar, L., & Wilson, P. W. (2000). A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics*, *27*(6), 779–802.
- Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, *136*(1), 31–64.
- Song, M., Zhang, J., & Wang, S. (2015). Review of the network environmental efficiencies of listed petroleum enterprises in China. *Renewable and Sustainable Energy Reviews*, *43*, 65–71.
- Sueyoshi, T., & Goto, M. (2012a). Data envelopment analysis for environmental assessment: Comparison between public and private ownership in petroleum industry. *European Journal of Operational Research*, *216*(3), 668–678.
- Sueyoshi, T., & Goto, M. (2012b). Returns to scale and damages to scale under natural and managerial disposability: Strategy, efficiency and competitiveness of petroleum firms. *Energy Economics*, *34*(3), 645–662.
- Sueyoshi, T., & Goto, M. (2015). DEA environmental assessment in time horizon: Radial approach for Malmquist index measurement on petroleum companies. *Energy Economics*, *51*, 329–345.
- Sueyoshi, T., & Wang, D. (2014). Sustainability development for supply chain management in US petroleum industry by DEA environmental assessment. *Energy Economics*, *46*, 360–374.
- Sueyoshi, T., & Wang, D. (2018). DEA environmental assessment on US petroleum industry: Non-radial approach with translation invariance in time horizon. *Energy Economics*, *72*, 276–289.

- Sueyoshi, T., Yuan, Y., & Goto, M. (2017). Literature study for DEA applied to energy and environment. *Energy Economics*, *62*, 104–124.
- Sun, C., Luo, Y., Huang, Y., & Ouyang, X. (2017). A comparative study on the production efficiencies of China's oil companies: A true fixed effect model considering the unobserved heterogeneity. *Journal of Cleaner Production*, *154*, 341–352.
- Van Dijk, B. (2021). Amadeus database. Bureau van Dijk Electronic Publishing.
- Vikas, V., & Bansal, R. (2019). Efficiency evaluation of Indian oil and gas sector: Data envelopment analysis. *International Journal of Emerging Markets*, *14*, 2. <https://doi.org/10.1108/IJoEM-01-2018-0016>
- Zhou, P., Ang, B. W., & Poh, K. L. (2008). A survey of DEA analysis in energy and environmental studies. *European Journal of Operational Research*, *189*(1), 1–18.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.