

# Environmental performance of the Chinese cement enterprise: An empirical analysis using a text-based directional vector

Yi Zhou, Wenji Zhou, Chu Wei<sup>\*</sup>

School of Applied Economics, Renmin University of China, PR China

## ARTICLE INFO

### Keywords:

Environmental performance  
Cement industry  
Text analysis  
Directional vector  
Directional distance function

## ABSTRACT

Cement industry is both energy-intensive and pollution-intensive, imposing significant negative externalities on sustainability. However, previous studies have not sufficiently focused on China's cement industry; thus, there is a lack of information on efficiency and cost, limiting policy effectiveness. Using a directional-distance-function-based model and firm-level data, this study measures the technical efficiency and shadow price of SO<sub>2</sub> and PM in China's cement industry from 1998 to 2010. In response to the debates on the selection of directional vector, this is the first study that proposes a new directional vector determined via text analysis. We use the five-year plans as a breakthrough point to construct a corpus, establish a lexicon, and introduce a text analysis in order to reveal the policy preference, and finally, solve this directional vector. Through this improvement, we highlight the inherent meanings of the directional vector and provide a new insight into the directional-distance-function-based method. We solve this new directional vector to be (0.98, -0.13, -0.13), implying that the preference for economic growth is significantly greater than that for environmental protection. The conditional median marginal abatement costs of SO<sub>2</sub> and PM are estimated to be 29,070 and 1374 yuan/ton, respectively. The annual trends of technical efficiency and shadow price are depicted, and the heterogeneities of regions, provinces, cities, and pollutants are revealed.

## 1. Introduction

Cement has long been an essential material supporting human civilization development for >2000 years (Amato, 2013; Schneider et al., 2011), contributing to infrastructure construction, economic growth, and social development. Around 2020, the global human-made mass may surpass all global living biomass for the first time, among which buildings and infrastructure mainly composed of cement may reach 1100 Gt, accounting for the largest mass proportion (Elhacham et al., 2020). However, it is also argued that the cement industry is energy-intensive and pollution-intensive that imposes significant negative externalities on sustainability. The cement industry accounts for 12–15% of the world's industrial energy consumption. It emits a large number of greenhouse gases (GHGs) and harmful air pollutants, including particulate matter (PM), sulfur dioxide (SO<sub>2</sub>), nitrogen oxides, and carbon monoxide (Hua et al., 2016; Madlool et al., 2011). It has become the largest emitter of PM and the third-largest source of SO<sub>2</sub> emissions in China (Zhang et al., 2019).

As the largest cement producer and user since 1986, China's

expanding cement industry contributed significantly to industrialization, urbanization, and modernization (Lin and Zhang, 2016). However, it necessitates increasingly intense energy conservation and emission reduction. It is estimated that the cement industry used 6% of China's coal and 4% of electricity, and contributed 4% of China's SO<sub>2</sub> emissions and 9% of PM emissions in 2015 (China Cement Association, 2001–2014; Liu et al., 2021; National Bureau of Statistics, 1983–2021, 2021; Shan et al., 2019). Although the government implemented numerous command-and-control and market-based measures, some of these policies failed to reach the desired goal (Price et al., 2011; Yang et al., 2017). This practical gap is partially due to the absence of firm information, such as the abatement potential and abatement cost (Teng et al., 2014). This information asymmetry between the regulated enterprises and policy makers has affected policy effectiveness.

To solve these practical challenges, the academia has made considerable efforts for assessing the environmental performance of the firms when accounting for the undesirable by-product output associated with the production process. Among the various modelling strategies, the directional distance function (DDF) has gained considerable attention

<sup>\*</sup> Corresponding author at: No.59, Zhongguancun Street, Haidian District, Beijing, PR China.

E-mail address: [xiaochu@ruc.edu.cn](mailto:xiaochu@ruc.edu.cn) (C. Wei).

and has been applied in this field (Bostian et al., 2022; Färe et al., 2005; Färe et al., 2012; Färe et al., 2006; Ma et al., 2019; Wang et al., 2020; Zhou et al., 2014). The major merit of DDF is that it can achieve both the expansion of desirable outputs and the reduction of undesirable outputs simultaneously. It can estimate the technical efficiency (TE) and shadow price that can serve as a reference for target setting, allowance allocation, and trading mechanism. However, because the estimation result is sensitive to the directional vector, there is no consensus on the selection of the directional vector yet. Consequently, most empirical studies disregard the intrinsic practical meaning of the directional vector and set it arbitrarily, thereby weakening the credibility of the results significantly.

The empirical analysis to assess the environmental performance of Chinese cement firms is relatively limited. Existing empirical studies that applied the advanced DDF model considered mainly the Chinese provinces (Du et al., 2015; Kaneko et al., 2010; Wang et al., 2019; Wei et al., 2021; Xie et al., 2016; Xie et al., 2021; Zeng et al., 2018), cities (Ji and Zhou, 2020; Wei, 2021; Yang et al., 2021; Yu and Zhang, 2021), thermal power plants (Nakaishi, 2021; Wei et al., 2013; Wei and Zhang, 2020; Zhang et al., 2021b), iron and steel enterprises (Wang et al., 2017), and others. Contrarily, the existing studies analyzing China's cement industry mostly rely on bottom-up (Dinga and Wen, 2022; Hasanbeigi et al., 2013; Zhang et al., 2021a) and top-down models (Pang et al., 2014). To the best of our knowledge, the only related study is Yang et al. (2017), wherein a non-parametric DDF approach was used to estimate the production and environmental performance of 54 cement plants in the Sichuan Province and compare the changes before and after the Great Sichuan Earthquake in 2008. However, such results do not offer a generalized implication for the industry.

To address these practical needs and research gaps, we apply the directional-distance-function-based model with a text-based directional vector to the Chinese cement firms. We first establish a new directional vector by applying text analysis to the government's five-year plans (FYP) for 288 cities during 1998–2010. This text-based directional vector of (0.98, -0.13, -0.13) is obtained, implying that the policy makers give higher priority to economic growth than to environmental protection. We apply this directional vector in the DDF model with two primary undesirable outputs during cement production (that is, SO<sub>2</sub> and PM). The spatiotemporal pattern of technical efficiency and marginal abatement cost (MAC) for 2646 cement firms are also presented.

Compared with previous studies, our study affords two major contributions. First, we develop the DDF model by deriving a new text-based directional vector. Second, we apply this new DDF model with Chinese cement firm-level data to quantify the performance. These efficiency and cost information are vital because they can guide policy makers in the identification of the benchmark and formulation of the targeting policy at the micro level.

The remaining part of this paper is structured as follows: Section 2 introduces China's cement industry. Section 3 presents the methodology and data. In Section 4, we derive the text-based directional vector. Section 5 discusses the results. The last section presents the conclusion.

## 2. Development of China's cement industry

In 1949, there were only 35 cement plants in China, with a limited output of 660 thousand tons, or equivalently 1.5 kg per capita. Thereafter, China's cement industry has undergone tremendous development. Fig. 1 shows the cement production from 1949 to 2020. In 1986, China's cement output reached 166 million tons and accounted for 17% of the world's production. By 1986, China's cement production capacity and consumption have consistently ranked first globally. The cement output in 2020 increased to 2395 million tons, approximately 3628 times higher than that in 1949, and contributed to 57% of the world's production (National Bureau of Statistics, 1983-2021). At present, there are >3300 cement firms in China, 95 times that in 1949.

The expansion of the cement industry also leads to negative

consequences (Shen et al., 2017).

**The first challenge** is overcapacity. Until 1995, China's cement output caught up to the market demand and ended the 45-years cement shortage. Since then, the exceeded production capacity necessitates restructuring and quality improvement. In 1999, the central and local governments reinforced efforts to eliminate outdated production processes, develop new technologies, and promote the merger and restructuring of enterprises. However, in response to the financial crisis in 2008, the government launched the 4-trillion Yuan (\$586-billion) Economic Stimulus Plan, of which infrastructure investment was the highest. This stimulus package enabled the recovery of the cement industry, with a large number of new cement production lines added to meet the policy-driven new demand (Zheng et al., 2021).

**Second**, the cement industry is one of the most energy-intensive sectors. Coal and electricity are the primary energy sources in cement production, accounting for approximately 85% and 10%, respectively. From 1996 to 2016, the coal and electricity consumption of China's cement production show an average annual growth rate of 7% and 8%, respectively. Considering 2016 as an example, the cement production consumed 255.87 million tons and accounted for 7% of industrial coal consumption and nearly 7% of China's total consumption. The electricity consumption was 248.70 billion kWh, accounting for 6% of industrial consumption and 4% of national consumption (China Cement Association, 2001-2014; National Bureau of Statistics, 1983-2021; Shan et al., 2019).

**The third challenge** is production-associated environmental pollution. The cement production process is usually associated with a large amount of atmospheric pollutants, with PM emissions topping the list of industrial sectors (Yang et al., 2013). As shown in Fig. 1, China's cement industry emitted 2.11 million tons of PM<sub>2.5</sub>, 3.32 million tons of PM<sub>10</sub>, and 0.43 million tons of SO<sub>2</sub> in 1990. In 2015, the general PM emissions from the cement industry declined to 1.38 million tons, but still accounted for 11% of the industrial PM emissions and 9% of China's total PM emissions. The SO<sub>2</sub> emissions reached 0.66 million tons, accounting for 4% of industrial emissions and approximately 4% of China's total emissions (Liu et al., 2021; National Bureau of Statistics, 2021).

In response to these problems, the central government implemented intensive policies. The timeline is as follows:

- In 1985, the Ministry of Ecology and Environment<sup>1</sup> first issued the “Cement Industry Air Pollutant Emission Standard”, stipulating the air pollutant emission limit and supervision requirements of cement enterprises. Since then, it has been revised three times in 1996, 2004, and 2013. In 2013, the “Cement Industry Pollution Prevention and Control Technology Policy” was issued, proposing the technical routes, principles, and methods that can be adopted for cement pollution abatement.
- In 2006, eight ministries including the National Development and Reform Commission clarified the development direction of cement industry during the period of the 11th FYP (2006–2010), leading the cement industry to accelerate industrial restructuring. Thereafter, 60 large-scale cement enterprises were also identified for government support to vigorously develop the novel dry process cement production.
- In 2009, the State Council issued “Opinions on Restraining Overcapacity and Duplicate Construction in Some Industries and Guiding the Healthy Development of Industries,” strictly controlling the new cement production capacity and eliminating the outdated capacity by the same amount. Since then, in 2010 and 2013, the State Council issued policies again, urging the elimination of outdated cement capacity. In 2018, it was strictly prohibited to add production capacity in cement industry.

<sup>1</sup> It was formerly known as the National Environmental Protection Agency.

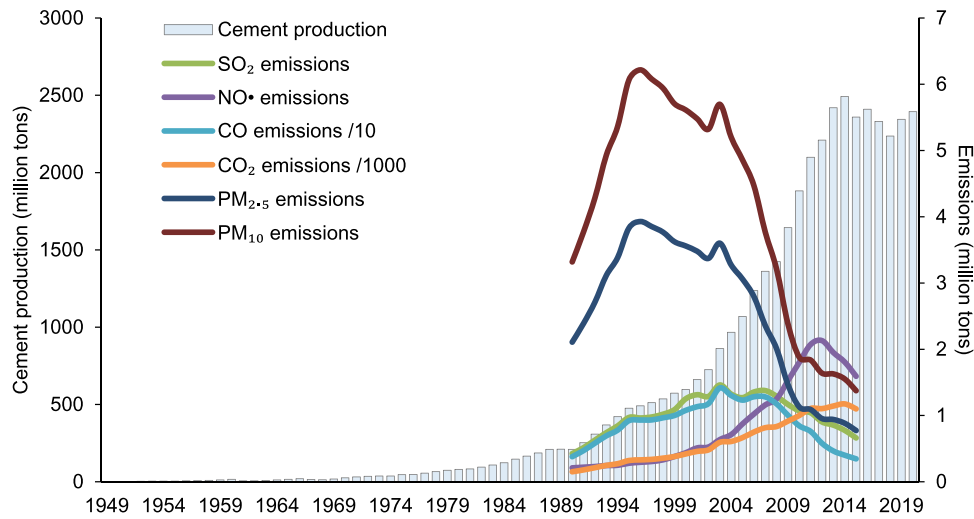


Fig. 1. Cement production and associated emissions.

Sources: The cement production data are collected from National Bureau of Statistics (1983–2021). The emission data are sourced from Liu et al. (2021).

- From the 6th FYP (1981–1985) to the 9th FYP (1996–2000), the statements concerning cement still focused on increasing production capacity and improving technology. Since the 10th FYP (2001–2005), cement-related expressions have gradually shifted to promoting new technologies and mergers and reorganizations, reflecting the adjustment of government's priorities. Furthermore, starting from the 10th FYP, there have been increasing expressions on energy conservation and emission reduction, with the related goals becoming more comprehensive.

In summary, the policies for the cement industry are mostly command-and-control measures, lacking the necessary market-based measures (Chen et al., 2022; Li et al., 2023; Zhang et al., 2020; Zhang et al., 2022). The lack of information on efficiency and cost may be the potential obstacles. In terms of the market-based strategy, relevant economic theories have proved that the theoretical price of emission permit trading can be estimated through the cost information, providing an important reference for the establishment and adjustment of the emissions trading system (Wei et al., 2013). Therefore, this study affords practical significance for the performance measurement of cement enterprises.

### 3. Methodology and data

Modelling the multiply-input multiple-output production process, particularly with the by-product undesirable outputs such as pollutant, has become a research hotspot and frontier. Benefitting from the solid production theory in microeconomics and flexibility in empirical analysis based on the historical data (Ma et al., 2019), the distance-function-based method gained wide application. It includes two branches: first is the non-parametric method mainly referring to data envelopment analysis (DEA) (He et al., 2021; Li et al., 2021). It does not require pre-setting the specification, but has shortcomings in terms of non-differentiability and sensitivity to outliers (Färe et al., 2005; Ma et al., 2019; Xie et al., 2016). The second approach is the parametric method mainly including parametric linear programming (LP) and stochastic frontier analysis (SFA).

In this case, we use the parametric method because it has the advantage of differentiability and easier estimation and calculation. Regarding the distance function form, we use the DDF because it is a generalized form of the conventional Shephard distance function. More importantly, the DDF can achieve both the expansion of desirable

outputs and the reduction of undesirable outputs simultaneously, fitting appropriately with the multiple-dimension decision-making in reality.

#### 3.1. Theoretical model

We model the environmental production technology and the DDF as follows (Färe et al., 2005; Färe et al., 2006). Suppose that a production unit can produce a desirable output vector  $y = (y_1, \dots, y_M) \in \mathbb{R}_+^M$  and an undesirable output vector  $b = (b_1, \dots, b_J) \in \mathbb{R}_+^J$  from an input vector  $x = (x_1, \dots, x_N) \in \mathbb{R}_+^N$ . The environmental production technology is described by the output possibility set,  $P(x) = \{(y, b) : x \text{ can produce } (y, b)\}$ . This production possibility set satisfies two standard assumptions. First,  $P(x)$  is a compact set and  $P(0) = \{0, 0\}$ . Second, the inputs are free disposable, that is, if  $x' \geq x$ , then  $P(x') \supseteq P(x)$ . To accommodate the pollutant production, we impose three specific assumptions. First, the free disposability of desirable outputs; that is, if  $(y, b) \in P(x)$  and  $y' \leq y$ , then  $(y', b) \in P(x)$ . Second, the weak disposability of desirable outputs and undesirable outputs together; that is, if  $(y, b) \in P(x)$  and  $0 \leq \theta \leq 1$ , then  $(\theta y, \theta b) \in P(x)$ . Third is the null-jointness assumption; that is, if  $(y, b) \in P(x)$  and  $b = 0$ , then  $y = 0$ . It implies that undesirable outputs are by-products.

Accordingly, the environmental production technology can be represented by the output DDF, that is,  $\vec{D}_o(x, y, b; g_y, -g_b) = \max\{\beta : (y + \beta g_y, b - \beta g_b) \in P(x)\}$ , of which  $g = (g_y, -g_b) \in \mathbb{R}_+^M \times \mathbb{R}_-^J$  is a directional vector (DV), making it feasible for a decision-making unit to hit the boundary of  $P(x)$  along a trajectory, where the desirable outputs are expanded, while undesirable outputs are reduced. The DDF represents the degree of “inefficiency,” namely, it measures the maximum degree of efficiency improvement for an output combination  $(y, b)$  in  $P(x)$ . Therefore, the greater the value of  $\vec{D}_o(x, y, b; g_y, -g_b)$ , the greater the degree of inefficiency of the point in  $P(x)$ , that is, the lower the efficiency. If a point is exactly located on the boundary of  $P(x)$ , then we obtain  $\vec{D}_o(x, y, b; g_y, -g_b) = 0$ .

Some properties that DDF needs to satisfy are inherited from the assumptions imposed on  $P(x)$ , such as the non-negativity, the monotonicity on inputs and outputs, and the weak disposability of desirable and undesirable outputs. In addition, it should also satisfy the translation property:

$$\vec{D}_o(x, y + \alpha g_y, b - \alpha g_b; g_y, -g_b) = \vec{D}_o(x, y, b; g_y, -g_b) - \alpha, \alpha \in \mathbb{R} \quad (1)$$

where if we substitute  $\alpha = \beta^* = \bar{D}_o(x, y, b; g_y, -g_b)$  into this translation property, we can obtain  $\bar{D}_o(x, y + \beta^* g_y, b - \beta^* g_b; g_y, -g_b) = 0$ .

The partial technical efficiency of desirable and undesirable outputs can be computed as follows:

$$TE_m = \frac{y_m}{y_m^*} = \frac{y_m}{y_m + \beta^* g_{y,m}} = \frac{y_m}{y_m + \bar{D}_o(x, y, b; g_y, -g_b) g_{y,m}} \quad (2)$$

$$TE_j = \frac{b_j^*}{b_j} = \frac{b_j - \beta^* g_{b,j}}{b_j} = \frac{b_j - \bar{D}_o(x, y, b; g_y, -g_b) g_{b,j}}{b_j}$$

where all the values of partial technical efficiency range from 0 to 1, and the production unit with higher efficiency has a higher value. If the TE value is displayed as 1, then it indicates that the output combination is located on the frontier.

Pertaining to the shadow price, we need to utilize the relationship between the revenue function and the DDF. We begin with the following revenue maximization problem:  $R(x, p, q) = \max_{y,b} \{py - qb : \bar{D}_o(x, y, b; g_y, -g_b) \geq 0\}$ , where  $p = (p_1, \dots, p_M) \in \mathbb{R}_+^M$  denotes the observable price vector of desirable outputs, and  $q = (q_1, \dots, q_J) \in \mathbb{R}_+^J$  represents the shadow price vector of undesirable outputs. To solve this maximization problem, we construct the Lagrangian function and solve the first-order conditions with respect to every  $y$  and  $b$ . Finally, we can derive the shadow price of the  $j$ -th undesirable output via the price of the  $m$ -th desirable output:

$$q_j = -p_m \frac{\partial \bar{D}_o(x, y, b; g_y, -g_b) / \partial b_j}{\partial \bar{D}_o(x, y, b; g_y, -g_b) / \partial y_m} \quad (3)$$

This expression implies that the shadow price can be regarded as the value of desirable output that must be dropped in order to reduce an additional unit of undesirable output. Therefore, the shadow price of an undesirable output can be used as a proxy of the marginal abatement cost (Coggins and Swinton, 1996).

### 3.2. Empirical specification

The translog and quadratic form are two commonly used functional forms used to represent the parametric DDF (Ma et al., 2019; Zhou et al., 2014). Here, we use the quadratic specification because it is easier to impose constraints to meet the translation property (Chambers, 1998), and it exhibits improved approximation properties (Färe et al., 2010) over the translog function. It is specified as follows:

$$\begin{aligned} \bar{D}_o(x_k, y_k, b_k; g_y, -g_b) = & \alpha_0 + \sum_{s=1}^S \psi_s SD_{sk} + \sum_{t=1}^T \varphi_t TD_{tk} \\ & + \sum_{n=1}^N \alpha_n x_{nk} + \sum_{m=1}^M \beta_m y_{mk} + \sum_{j=1}^J \gamma_j b_{jk} + \frac{1}{2} \sum_{n=1}^N \sum_{n'=1}^N \alpha_{nn'} x_{nk} x_{n'k} \\ & + \frac{1}{2} \sum_{m=1}^M \sum_{m'=1}^M \beta_{mm'} y_{mk} y_{m'k} + \frac{1}{2} \sum_{j=1}^J \sum_{j'=1}^J \gamma_{jj'} b_{jk} b_{j'k} \\ & + \sum_{n=1}^N \sum_{m=1}^M \delta_{nm} x_{nk} y_{mk} + \sum_{n=1}^N \sum_{j=1}^J \nu_{nj} x_{nk} b_{jk} + \sum_{m=1}^M \sum_{j=1}^J \mu_{mj} y_{mk} b_{jk} \end{aligned} \quad (4)$$

where subscript  $k$  denotes the number of observations. In addition,  $SD$  and  $TD$  represent the spatial and time dummy variables, respectively, in order to measure the fixed effects in the panel data. In this regard, the subscript  $s$  and  $t$  denote the division of space and time, respectively. Accordingly, the parameters in Eq. (4) can be estimated by a deterministic linear programming algorithm (Aigner and Chu, 1968):

$$\begin{aligned} \min & \sum_{k=1}^K [\bar{D}_o(x_k, y_k, b_k; g_y, -g_b) - 0] \\ \text{s.t.} & \quad (i) \bar{D}_o(x_k, y_k, b_k; g_y, -g_b) \geq 0, k = 1, \dots, K \\ & \quad (ii) \frac{\partial \bar{D}_o(x_k, y_k, b_k; g_y, -g_b)}{\partial x_{nk}} \geq 0, n = 1, \dots, N; k = 1, \dots, K \\ & \quad (iii) \frac{\partial \bar{D}_o(x_k, y_k, b_k; g_y, -g_b)}{\partial y_{mk}} \leq 0, m = 1, \dots, M; k = 1, \dots, K \\ & \quad (iv) \frac{\partial \bar{D}_o(x_k, y_k, b_k; g_y, -g_b)}{\partial b_{jk}} \geq 0, j = 1, \dots, J; k = 1, \dots, K \\ & \quad (v) \begin{cases} \sum_{m=1}^M g_{y,m} \beta_m - \sum_{j=1}^J g_{b,j} \gamma_j = -1 \\ \sum_{m=1}^M g_{y,m} \beta_{m'n'} - \sum_{j=1}^J g_{b,j} \mu_{nj} = 0, m = 1, \dots, M \\ \sum_{m=1}^M g_{y,m} \mu_{mj} - \sum_{j=1}^J g_{b,j} \gamma_{jj} = 0, j = 1, \dots, J \\ \sum_{m=1}^M g_{y,m} \delta_{nm} - \sum_{j=1}^J g_{b,j} \nu_{nj} = 0, n = 1, \dots, N \end{cases} \\ & \quad (vi) \begin{cases} \alpha_{nn'} = \alpha_{n'n}, n, n' = 1, \dots, N; n \neq n' \\ \beta_{mm'} = \beta_{m'm}, m, m' = 1, \dots, M; m \neq m' \\ \gamma_{jj} = \gamma_{jj}, j, j' = 1, \dots, J; j \neq j' \end{cases} \end{aligned} \quad (5)$$

where the optimization problem is to minimize the sum of deviations of all the observations from the frontier. The constraint conditions in Eq. (5) inherit the properties of DDF. The first constraint ensures that all the observations are feasible. Constraint (ii)–(iv) are the monotonicity conditions. Constraint (v) reflects the translation property. The symmetry conditions are imposed by Constraint (vi).

### 3.3. Data

We collect our firm-level data from *China Industrial Enterprise Database* and *China Manufacturing Industry Emissions Integrated Database*. After matching the two databases and addressing the missing values and outliers,<sup>2</sup> we obtain an unbalanced panel data set of China's cement industry from 1998 to 2010, containing a total of 8334 observations. Our data covers 2646 cement firms in mainland China that have >5 million Yuan sales and account for >85% of the total local emissions. Therefore, our sampled enterprises are representative. Our data also cover many business types, including state-owned enterprises, collective enterprises, private enterprises, Hong Kong, Macao and Taiwan-funded enterprises, and foreign-funded enterprises. In terms of the geographical distribution, our data cover 29 provincial-level administrative regions and 284 prefecture-level regions, accounting for 85% of the total number of prefecture-level regions in China. Furthermore, our sample has a wide range of spatial distribution, as shown in Fig. 2.

Table 1 shows the descriptive statistics of these variables. Our data contains three inputs (capital, labor, and energy), one desirable output

<sup>2</sup> After matching, the raw data contain 51,490 observations, with a large number of missing values. Processing steps: First, we combine the linear interpolation and approximate variable substitution to fill in the missing values. Second, we drop observations that still have missing values after filling, resulting in 13,435 observations. Third, for outliers, we trim all variables by 5% by year group and subsequently 5% in the full sample. Finally, we drop observations with trimmed variables, retaining 8334 observations.

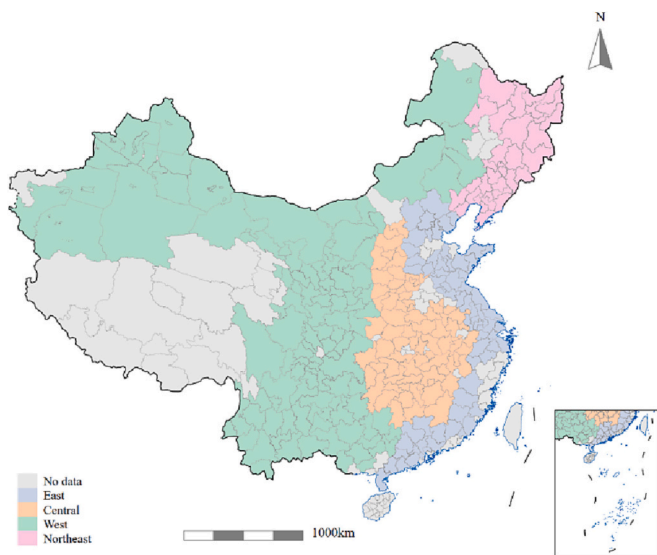


Fig. 2. City-level distribution of sample firms.

(gross value of industrial output), and two undesirable outputs (SO<sub>2</sub> and PM). We use the total fixed assets to measure the capital input that is uniformly converted to 1998 constant Chinese Yuan. The fixed assets investment price indices are sourced from China's National Bureau of Statistics. The labor input is measured by the average annual number of employees at the end of the year. The energy input, dominated by the feed coal and fuel coal, is aggregated into the standard coal equivalent. The conversion factors are sourced from the *China Energy Statistical Yearbook*. Regarding the output, we use the gross industrial output value in 1998's price to measure the desirable output. The undesirable outputs include SO<sub>2</sub> emissions and PM emissions, where the PM emissions are the sum of smoke and dust emissions.

4. Establishment of a new directional vector

The selection of the DV could be critiqued from two aspects (Ma et al., 2019). In the first case, the setting of DV is arbitrary. For example, previous studies usually set the DV as  $(g_y, -g_b) = (1, -1)$ . However, some works have shown that the results are significantly sensitive to the DV (Murty et al., 2012; Vardanyan and Noh, 2006; Wang et al., 2017), imposing worries on the credibility and comparability of the results. The second case is that the DV fails to depict the actual abatement policies or measures in reality (Ma et al., 2019). In other words, the DV is not without practical meanings and cannot reflect the decision maker's trade-offs among desirable and undesirable outputs. Therefore, the directional vector selected arbitrarily ignores its inherent meanings, significantly weakening the guiding significance of the empirical results in reality.

There are several efforts aimed to solve these problems. The first attempt is to enumerate many DVs and select the DV associated with a comparable result; for example, comparison of the estimated marginal abatement cost with the pollutant's actual market price (Ji and Zhou,

2020; Ji et al., 2021). However, this strategy may reverse the decision-making order. The second approach attempts to endogenize the DV (Färe et al., 2017). It adds each element of the DV into the original parametric linear programming model as a decision variable. In this manner, the DV value is combined with the coefficient of the DDF to be optimized simultaneously, transforming linear programming into non-linear programming. Although this new idea endogenizes the DV within the framework of optimal decision-making modelling, it lacks a realistic basis. Moreover, it usually leads to no solution or multiple solutions because it is transformed as non-linear programming. The third strategy is to set the DV equal to the observed outputs or inputs, actually forming a radial model (Aparicio et al., 2016; Cheng and Zervopoulos, 2014; D'Inverno and De Witte, 2020; Foltz et al., 2012; Pastor et al., 2020; Picazo-Tadeo et al., 2012; Singbo et al., 2014; Yu et al., 2021; Zhou et al., 2012). However, such vectors are generally combined with DEA, rather than parametric methods and are usually used in theoretical models. More importantly, this approach still evades the practical meanings of the directional vector.

4.1. Defining a new directional vector

We propose a new DV, namely a text-based DV determined via text mining and analysis. The logic is, we apply the text analysis with China's five-year plans to reveal the underlying policy preference between economy and environment. We subsequently quantify this policy preference as our new DV. Fig. 3 presents the analytical theme. The specific step is as follows.

We first argue that policy preference can generate a directional vector with inherent meanings. Recalling the most widely used DV,  $(g_y, -g_b) = (1, -1)$ , it represents a direction of efficiency improvement where the decision-maker expands one unit of desirable output and

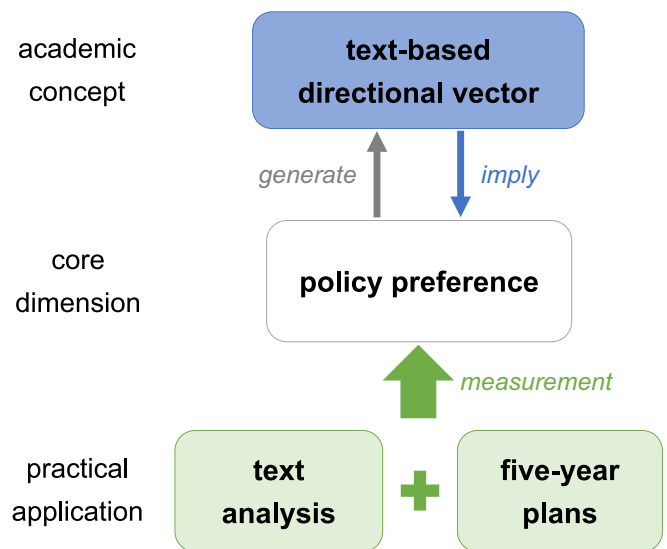


Fig. 3. Analytical theme of the new directional vector.

Table 1  
Descriptive statistics of input–output variables.

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Capital (10 <sup>4</sup> yuan)	8334	2327.33	2465.44	176.64	16,006.38
Labor (employee)	8334	312.44	193.44	95	1083
Energy (tce)	8334	18,859.90	14,296.87	3316.10	87,474.29
Output value (10 <sup>4</sup> yuan)	8334	3076.91	2675.12	598.26	17,186.63
SO <sub>2</sub> (ton)	8334	186.31	185.44	10.85	932.87
PM (ton)	8334	949.17	1056.45	20.14	5154.50

Notes: All the price unit “yuan” is the 1998 constant Chinese Yuan. The unit “tce” means ton of standard coal equivalent.

simultaneously reduces one unit of undesirable output. A different DV will induce a different path for efficiency improvement or a different strategy for emission abatement (Wang et al., 2017). For example, the vector  $(g_y, g_b) = (1, 0)$  reflects a strategy of only expanding desirable outputs without reducing undesirable outputs. Contrarily, the vector  $(g_y, g_b) = (0, -1)$  describes a strategy of only mitigating undesirable outputs without increasing the desirable outputs. Although the two DVs depict the paths for efficiency improvement, the strategy preference is rather different. Moreover, the relative size of  $g_y$  to  $g_b$  implies the relative preference for the economy to the environment when the decision-maker is aiming for efficiency improvement (Agee et al., 2012; Ji and Zhou, 2020; Ji et al., 2021). Therefore, the inherent meaning of a DV is exactly the preference relationship of decision-making units for all the outputs that should be sourced from the actual abatement measures or policies (Ma et al., 2019).

Second, China's FYP is a suitable reference sample, which can be quantified by text analysis for the measurement of policy preference. The national FYP mainly clarifies the national strategic intentions and the government's priorities and guides market participants in their activities, marking it as one of the most important policy tools in China (Hu, 2016). Therefore, the FYP has become a suitable reference for capturing the national strategic intentions, policy preference, and government priorities across various periods. In recent years, thanks to the rapid development of the Internet and the openness of government information, quantitative text analysis has been widely used. In addition to the information and library science, this method has also been introduced into the fields of management or economics to analyze China's policy attention and tool preference (Tan et al., 2021; Wang and Li, 2017). This paper further develops its application scenarios by using it to quantify policy preference.

#### 4.2. Quantifying the policy preference

Here, we apply text analysis to analyze the FYP and characterize the policy preference. The specific steps are as follows. **First**, we construct a corpus of the semantic expression of national policies from the 6th FYP (1981–1985) to the latest 14th FYP (2021–2025).<sup>3</sup> **Second**, we establish a lexicon that can reflect the policy preference. The Chinese word segmentation module “jieba” in Python is used, generating a total of 22,559 words after removing the single characters. Subsequently, we organize five doctoral students and five master students to screen out a total of 1293 words whose word frequency is not <20.<sup>4</sup> Finally, combined with the re-judgment and discussion, this procedure produces total 108 (= 88 + 20) words that will be added into the user-defined dictionary. Fig. 4 illustrates the words in this lexicon by a word cloud, while a detailed word list is reported in Appendix A.

**Third**, we conduct a text analysis of the policy documents during the sample period from 1998 to 2010. This research period spans three FYPs: the 9th FYP Plan (1996–2000), the 10th FYP (2001–2005) and the 11th FYP (2006–2010). Therefore, we use the texts of these three FYPs for text analysis. We use the 108 words above as a user-defined dictionary to expand the vocabulary in “jieba”, and perform the word

<sup>3</sup> It should be noted that this step is a preparatory work to establish a lexicon that can reflect the policy preference, so the period of FYPs is not limited to the sample period (1998–2010) of this paper.

<sup>4</sup> The scoring process is as follows: the word is assigned a value of 1, if it is considered to reflect the decision-maker's preference for “economic growth”, and it is assigned a value of -1, if it is considered to reflect the preference for “environmental protection”. If the word is deemed to reflect neither preference, then it is assigned a value of 0. Subsequently, we sum up the scoring results of 10 postgraduate students, and select words whose result is not lower than 6 or not higher than -6, obtaining 100 words reflecting the preference for “economic growth” and 24 words reflecting the preference for “environmental protection”. A robust analysis for this step is provided in Appendix B.

segmentation and word frequency analysis of the target documents. Based on the equal weight method (Hoberg and Phillips, 2016), we sum up the word frequencies of words reflecting the preference for “economic growth” and “environmental protection” and obtain the results of 5736 and 749, respectively. Accordingly, the “economy-environment” preference ratio is equal to 7.66 (= 5736/749).

#### 4.3. Constructing the text-based directional vector

We denote the preference ratio between economy and environment as  $\rho$ ; accordingly, the text-based DV  $(g_{y,1}, \dots, g_{y,M}, -g_{b,1}, \dots, -g_{b,J})$  can be solved by the following system of equations:

$$\begin{cases} \text{(i)} & \frac{g_{y,m}}{g_{b,j}} = \rho, m = 1, \dots, M; j = 1, \dots, J \\ \text{(ii)} & g_{y,m} = g_{y,m'} \geq 0, m, m' = 1, \dots, M; m \neq m' \\ \text{(iii)} & g_{b,j} = g_{b,j'} \geq 0, j, j' = 1, \dots, J; j \neq j' \\ \text{(iv)} & \sqrt{\sum_{m=1}^M (g_{y,m})^2 + \sum_{j=1}^J (g_{b,j})^2} = 1 \end{cases} \quad (6)$$

where Eq. (i) is the definition of the preference ratio. Eq. (ii) and (iii) assume the indifference in preference among the expansion of desirable outputs and the reduction of undesirable outputs, respectively.<sup>5</sup> The last equation constrains the norm of the DV to be 1.<sup>6</sup>

We substitute  $\rho = 7.66$  into Eq. (6), solving our text-based DV to be (0.98, -0.13, -0.13). These three coordinates represent the expansion of the gross industrial output value, the reduction of SO<sub>2</sub> emissions, and the reduction of PM emissions, respectively. This new DV is significantly different from the conventional vector, that is, (1, -1, -1). It shows that, in China's context, the preference for economic growth is considerably higher than that for the environmental protection. This is also consistent with the national economic characteristics and priorities during the sample period of this paper.

### 5. Results and discussions

Table 2 shows the parameters estimated by CPLEX solver in GAMS. It should be noted that to overcome the convergence problem in linear programming, we normalize all inputs and outputs by their mean value (Färe et al., 2005).

<sup>5</sup> Limited by the availability of policy text, this paper is unable to further identify the preference between SO<sub>2</sub> and PM, which is to be addressed by subsequent studies.

<sup>6</sup> This constraint is to ensure the comparability of the results under different DV. To explain this intention, recalling the DDF's definition, we can calculate the Euclidean distance from the output combination (y,b) to the frontier as

$$\begin{aligned} \text{Euclidean distance} &= \sqrt{\sum_{m=1}^M (y_m^* - y_m)^2 + \sum_{j=1}^J (b_j^* - b_j)^2} \\ &= \sqrt{\sum_{m=1}^M (y_m + \beta^* g_{y,m} - y_m)^2 + \sum_{j=1}^J (b_j - \beta^* g_{b,j} - b_j)^2} \\ &= \beta^* \sqrt{\sum_{m=1}^M (g_{y,m})^2 + \sum_{j=1}^J (-g_{b,j})^2} = \vec{D}_o(x, y, b; g_y, -g_b) \cdot |\vec{DV}| \end{aligned}$$

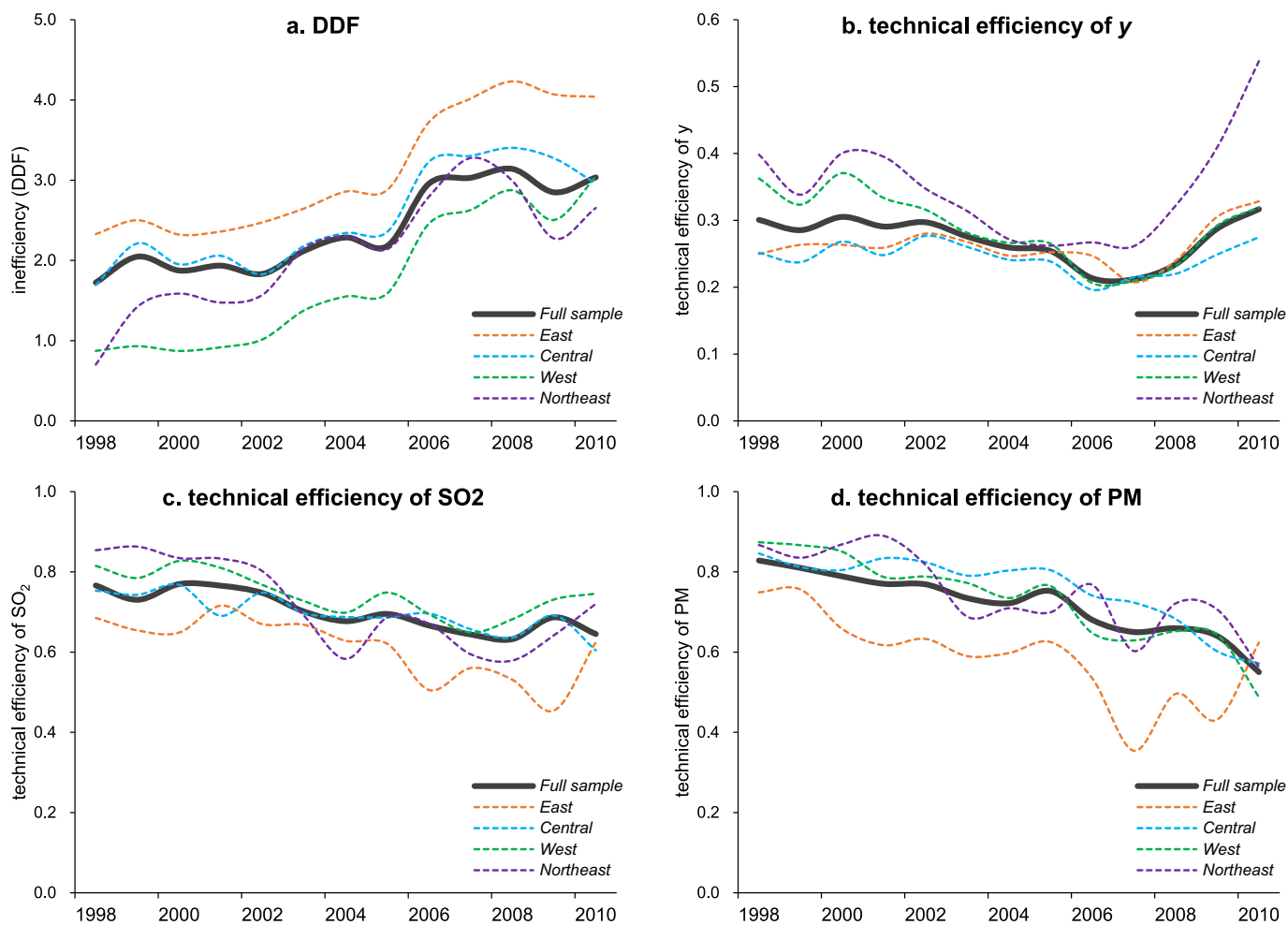
where DV denotes the directional vector. The Euclidean distance from point (y,b) to the frontier is equal to the product of DDF and the norm of DV. Therefore, in order to compare and analyze the influence of different DV on the results, we restrict all the DVs to be unit vectors, causing the value of DDF to be equal to the Euclidean distance regardless of the DV.



**Table 3**  
Results under the text-based directional vector.

Variable	Observation	Mean	Median	Standard Deviation	Minimum	Maximum
DDF	8298	2.30	2.32	1.20	0	7.53
TE <sub>y</sub>	8298	0.31	0.27	0.18	0.03	1
TE <sub>SO<sub>2</sub></sub>	6630	0.65	0.71	0.25	0.00	1
TE <sub>PM</sub>	5744	0.67	0.75	0.25	0.00	1
MAC <sub>SO<sub>2</sub></sub> (10 <sup>4</sup> yuan/ton)	8294	2.96	2.91	0.82	0.03	6.52
MAC <sub>PM</sub> (10 <sup>4</sup> yuan/ton)	8297	0.18	0.14	0.14	0	0.88

Notes: DDF denotes the directional distance function. TE represents the partial technical efficiency. MAC is the marginal abatement cost, that is, the shadow price. All the price unit “yuan” is the 1998 constant Chinese Yuan. We exclude 36 observations with negative DDF. Due to the negative  $b^*$ , there are 1668 observations with negative partial TE of SO<sub>2</sub>, and 2554 observations with negative partial TE of PM. There are also four observations with a negative shadow price of SO<sub>2</sub>, and one observation with a negative shadow price of PM.



**Fig. 5.** Annual median inefficiency and partial TE in the different regions.

inefficiency. The Northeast region exhibits a close pattern with nation's trend, but with considerable variation.

Fig. 5b shows the technical efficiency of the gross industrial output value. It shows a slight decline before 2006, but exhibits subsequent increase. Combined with the reverse trend in Fig. 5a, the rapid development of the cement industry since 2006 suggests efficiency improvement from the economic output side, but the overall efficiency reduces when environmental pollution is considered. This conclusion is confirmed with the trend shown in Fig. 5c and d, of which the technical efficiency of SO<sub>2</sub> and PM experience continuous or even accelerating decline.

Moreover, Fig. 5b–d also reveal substantial regional disparity. Prior to 2006, the technical efficiency of output value of the four regions presents a convergence trend, but the Northeast region shows a different

and remarkable rebound pattern since 2006. Regarding the performance of pollutant emission, the East region is significantly lower than the nation's level and exhibits considerable variation. These findings reflect the importance of regional difference when establishing nation-wide policy.

Fig. 6 shows the city-level distribution of the inefficiencies of cement firms. The most efficient cities appear in provinces such as Anhui, Heilongjiang, and Jiangxi. Furthermore, Yunnan, Jilin, Ningxia, and Shaanxi also feature cities with high efficiency. This finding is consistent with the conclusion in Fig. 5a.

5.2. Spatial-temporal pattern of abatement cost

The shadow price of the pollutant is a suitable proxy of the marginal

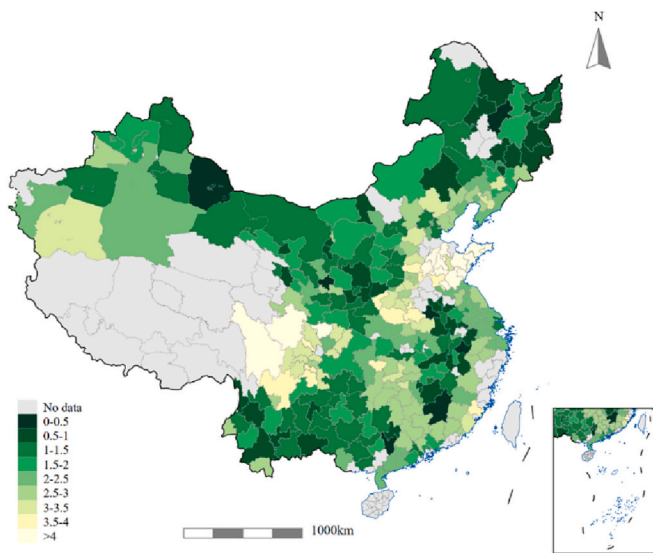


Fig. 6. City-level distribution of median inefficiency.

abatement cost. This information forms the foundation and benchmark for different market-based instruments such as emission permit trading and environmental tax. Given intensive previous estimates on abatement cost of SO<sub>2</sub> in the literature, Table 4 compares our results with those of the other relevant studies. The previous estimates on abatement cost for one ton of SO<sub>2</sub> reduction usually range from tens of thousands to hundreds of thousands of yuan, and in few studies, even as high as millions of yuan. Our estimate of 29,574 yuan/ton is within a reasonable range. A few reasons may be attributed to this disparity. First, the data-level and sample period are considerably different. For example, among these studies, the data covers the province-level, city-level, or firm-level, limiting the comparability of the results. Second, the unconditional marginal abatement cost estimated under a single pollutant is often overestimated owing to the ancillary benefit (Burtraw et al., 2003; Kesicki and Ekins, 2012; Ma and Hailu, 2016). Moreover, the choice of DV may remarkably alter the results.

Fig. 7a first suggests a downward trend. The abatement cost for the reduction of one ton of SO<sub>2</sub> was 31,600 yuan in 1998 that gradually decreased to 25,100 yuan in 2010. In contrast, Fig. 7b shows that the cost to curb one ton of PM fluctuates, ranging from 1200 to 1700 yuan/ton. The literature suggests that the cleaner, the more expensive is the further reduction of the pollutants (Murty et al., 2007). Accordingly, our case suggests a worsening environmental aspect of the Chinese cement industry during our research period. This is also consistent with the technical efficiency pattern. In terms of the regional gaps, Fig. 7a suggests the abatement cost of SO<sub>2</sub> in the central region that is higher than that at the national level, while that in the east region is relatively lower. During the sample period, the abatement cost of SO<sub>2</sub> in the west region

Table 4  
Comparison of SO<sub>2</sub> shadow price in different studies.

Study	Pollutant	DMU	Sample Period	Model/Method	MAC of SO <sub>2</sub>
Kaneko et al. (2010)	SO <sub>2</sub>	Thermal power sector, China	2003, 2006	DDF/DEA	20,720–47,298 yuan/ton
Xie et al. (2016)	SO <sub>2</sub>	30 provinces, China	1998–2011	QDDF/LP	419,000–1,493,000 yuan/ton
Zeng et al. (2018)	SO <sub>2</sub>	30 provinces, China	2001–2013	QDDF/LP	13,765–14,711 yuan/ton
Wei and Zhang (2020)	CO <sub>2</sub> , SO <sub>2</sub>	93 coal-fired power plants, China	2005–2010	QDDF/LP	16,705 yuan/ton
Ji and Zhou (2020)	CO <sub>2</sub> , SO <sub>2</sub> , NO <sub>x</sub>	105 cities, China	2006–2014	QDDF/LP	132,519–171,412 yuan/ton
Nakaishi (2021)	CO <sub>2</sub> , SO <sub>2</sub>	316 coal-fired power plants, China	2014	QDDF/LP	223,401 yuan/ton
Ji et al. (2021)	CO <sub>2</sub> , SO <sub>2</sub> , NO <sub>x</sub> , PM	30 major cities, China	2006–2015	QDDF/LP	99,330 yuan/ton (median) <sup>a</sup>
This study	SO <sub>2</sub> , PM	Cement industry, firm-level, China	1998–2010	QDDF/LP	29,574 yuan/ton

Notes: DMU denotes the decision-making unit. QDDF represents the quadratic directional distance function. DEA indicates the data envelopment analysis. LP indicates the linear programming. MAC is the marginal abatement cost, that is, the shadow price.

<sup>a</sup> The authors only report the median instead of the mean owing to the highly skewed distribution.

exhibits the largest decrease of approximately 25%. As shown in Fig. 7b, the regional differences of the PM abatement cost are mixed and vary significantly.

Fig. 8 depicts the relationship between the inefficiency and abatement cost by provinces. The bubble size indicates the inefficiency level. The horizon axis in Fig. 8 shows that Chongqing has the lowest SO<sub>2</sub> abatement cost (24,381 yuan/ton), followed by Shanghai (26,873 yuan/ton) and Yunnan (27,055 yuan/ton). Contrarily, Henan holds the highest cost (31,717 yuan/ton), followed by Hubei (31,104 yuan/ton) and Anhui (31,032 yuan/ton). The vertical axis shows that Guizhou registers the lowest cost for PM reduction (792 yuan/ton), followed by Fujian (926 yuan/ton) and Jiangxi (987 yuan/ton). The most expensive province to control PM is Shanghai (3346 yuan/ton) that is four times that of Guizhou, followed by Tianjin (2513 yuan/ton) and Chongqing (1939 yuan/ton). Fig. 8 also suggests that the abatement cost of PM varies significantly among the provinces, while the cost of SO<sub>2</sub> is relatively concentrated. Moreover, Shanghai, Tianjin, and Chongqing show different patterns with the other provinces.

### 5.3. Comparison of the directional-vector-based results

Here, we compare the existing results with those under some other directional vectors. We establish the following alternative DVs ( $g_y, -g_{b,1}, -g_{b,2}$ ). As Fig. 9 illustrates, we increase the angle between the DV and the vertical axis (representing the expansion of the desirable output) from 0° to 90° in increments of 5°, constructing 19 alternative vectors, denoted in black. The vertical axis  $g_y$  represents the expansion of the desirable output, and the two horizontal axes  $-g_{b,1}$  and  $-g_{b,2}$  represent the reduction of the two undesirable outputs. The orange curved surface denotes the constraint on the vector norm, and the light blue planes denote the preference ratio; the red vector denotes our text-based DV.

In particular, for  $\zeta \in \{0^\circ, 5^\circ, 10^\circ, \dots, 90^\circ\}$ , these alternative DVs are solved by the following system of equations:

$$\begin{cases} \text{(i)} & g_y = \cos\zeta \\ \text{(ii)} & g_y \geq 0 \\ \text{(iii)} & g_{b,1} = g_{b,2} \geq 0 \\ \text{(iv)} & \sqrt{g_y^2 + g_{b,1}^2 + g_{b,2}^2} = 1 \end{cases} \quad (7)$$

where Eq. (i) implies the preference ratio between economic growth and environmental protection. The smaller the angle  $\zeta$ , the stronger the preference for economic growth. Eq. (iii) assumes the indifference in preference between the abatement of two pollutants. Eq. (iv) constrains the directional vector to be a unit vector, to ensure that the results under different directional vectors are comparable.

Moreover, we adopt the optimized directional vector proposed by Färe et al. (2017) and solve the model by using the non-linear programming solver CONOPT. This optimized directional vector via non-linear programming is solved to be (0.01, -0.71, -0.71), with an angle of 89.35° from the vertical axis.

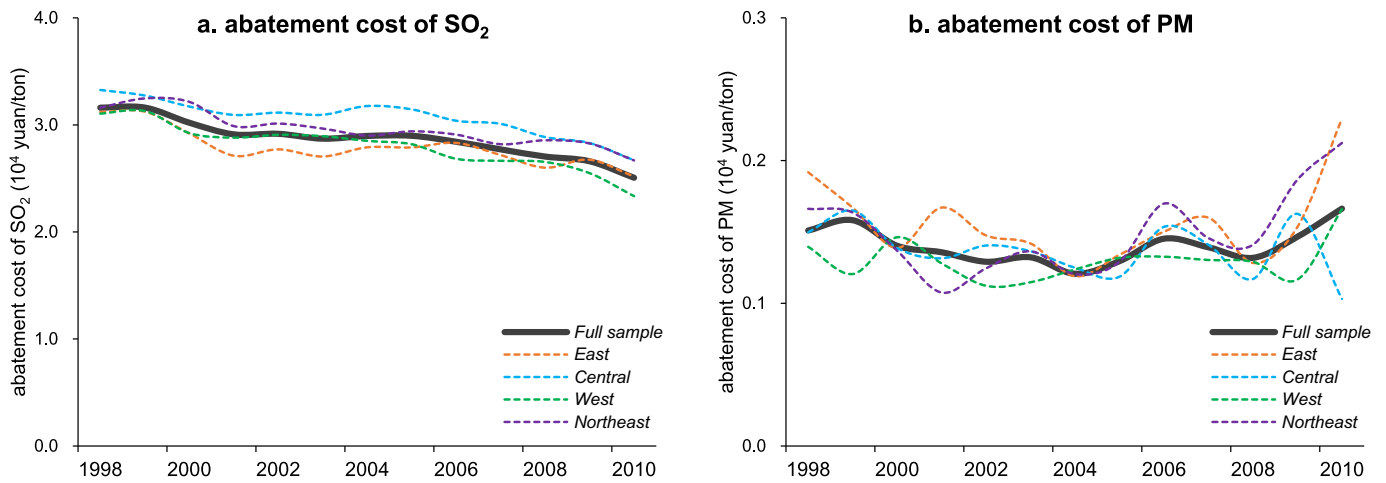


Fig. 7. Annual median shadow price in different regions.

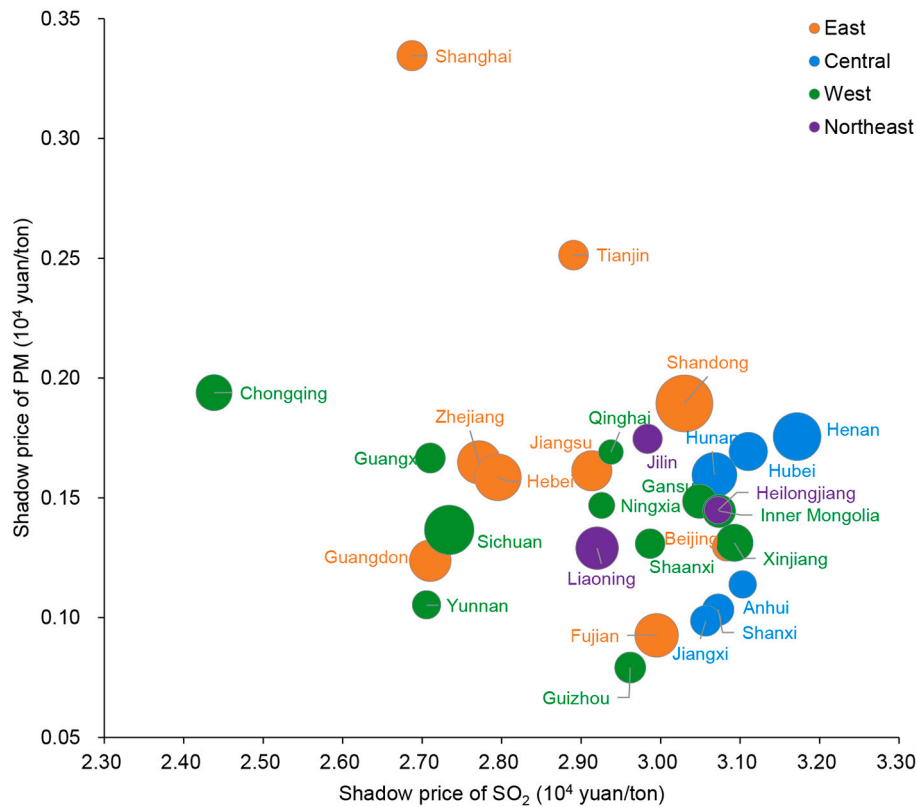


Fig. 8. Median shadow price and inefficiency in the different provinces and regions.

Notes: The size of the bubble represents the degree of inefficiency. The larger the bubble, the higher the inefficiency.

To compare the performance of these different DV-based results, we define the “goodness-of-fit” of the model as follows (Ji and Zhou, 2020):

$$Goodness-of-fit = \frac{K - \sum_{k=1}^K F_k}{\sum_{k=1}^K [\vec{D}_o(x_k, y_k, b_k; g_y, -g_b) - 0]} \quad (8)$$

where  $K$  denotes the total number of observations. In addition,  $F_k$  is a binary variable,  $F_k = 1$ , if  $\vec{D}_o(x_k, y_k, b_k; g_y, -g_b) < 0$  or  $b_k - \vec{D}_o(x_k, y_k, b_k; g_y, -g_b)g_b < 0$ , or  $F_k = 0$  otherwise. The larger the numerator, the better the model fits the data. The denominator

represents the sum of inefficiency that is the objective function that needs to be minimized in Eq. (5). The smaller the denominator, the better the model fits the data. Therefore, the fraction of Eq. (8) can comprehensively reflect how appropriately the model fits the data.

Fig. 10 depicts the results of the above-mentioned 20 alternative DVs as well as our text-based DV. It shows that the shadow prices under different DVs vary significantly. For the case of DV ranging from  $0^\circ$  to  $90^\circ$ , the marginal abatement cost of  $SO_2$  increases 1877 times, that is, rises from 4950 yuan/ton to 9,298,027 yuan/ton. Simultaneously, the marginal abatement cost of PM also rises from 987 yuan/ton to 2,014,717 yuan/ton, which is an increase of 2040 times. This indicates that the marginal abatement cost estimate is extremely sensitive to the

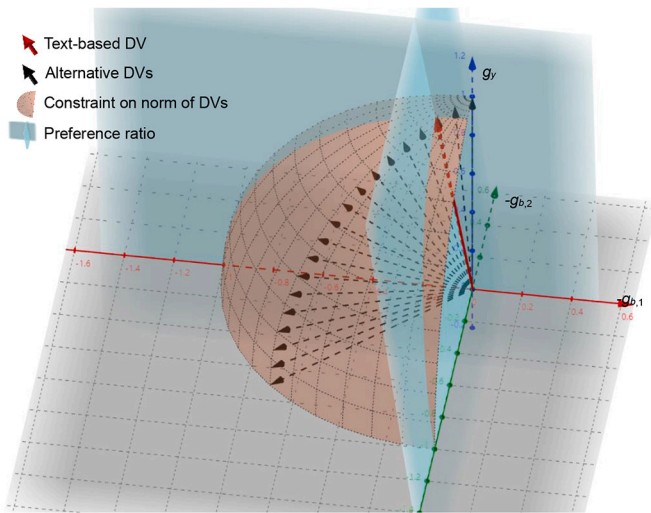


Fig. 9. Illustration of different directional vectors.

DV and suggests that the DV should be set with caution.

In comparison, for the case of vector (0.01, -0.71, -0.71), the corresponding marginal abatement cost of SO<sub>2</sub> and PM is 9,322,842 and 2,020,181 yuan/ton, respectively. This is significantly different from the range of the results presented in previous studies and is also inconsistent in terms of practicality.

The goodness-of-fit shows a downward trend as the angle increases. However, it tends to stabilize when the angle exceeds 40°. Overall, the goodness-of-fit under our text-based DV is at a high level, indicating that our model is a suitable fit for the data. It should be noted that, this goodness-of-fit is purely constructed from a mathematical perspective and does not endogenously originate from the inherent meanings of the directional vector; it can only be used as a reference for comparison instead of a basis for selecting a directional vector.

Comparing the results of various alternative DVs as well as our DV, we once again reveal and confirm that the marginal abatement cost is extremely sensitive to the DV. This highlights the contribution of our

text-based DV that originates at the inherent meanings and is determined with extreme caution.

### 6. Conclusion

As one of the industries with the highest energy consumption and heaviest pollutant emissions, China's cement industry poses considerable pressure on energy conservation and emission reduction. Information on efficiency and cost play a key role in policy regulation, which is an urgent question that the academia needs to address. However, studies on environmental performance have not paid sufficient attention to China's cement industry; this has not been conducive to the high-quality development of the cement industry. This study is the first to use the directional distance function to measure the technical efficiency and shadow price in China's cement industry, bridging the gap in this research field and providing a valuable reference for policy formulation at the micro level. This is the first contribution of this study.

In response to the debates on the directional vector in the DDF-based method, this study is the first to propose a new text-based DV. We first constructed a corpus of the semantic expression of national policies, thereafter established a lexicon reflecting the preference, conducted a text analysis of the target documents to reveal the policy preference, and finally solved it. For the first time, we highlight and start at the inherent meanings to determine the DV. Through this method, we provide a new insight into the DDF-based method.

Applying this model to the firm-level data of China's cement industry from 1998 to 2010, we solve the text-based DV to be (0.98, -0.13, -0.13). The conditional median marginal abatement costs of SO<sub>2</sub> and PM are estimated to be 29,070 yuan/ton and 1374 yuan/ton, respectively that are more practical compared with those in the previous studies. Through the temporal and spatial decomposition analysis, it is found that the overall inefficiency of China's cement industry during the sample period shows an upward trend. It is further found that the increase in the overall inefficiency is actually caused by the decline in the partial technical efficiency of pollutant emissions. The marginal abatement cost of SO<sub>2</sub> shows a significant downward trend during the sample period, implying a worsening environment performance during our sample period. Moreover, the heterogeneities of regions, provinces,

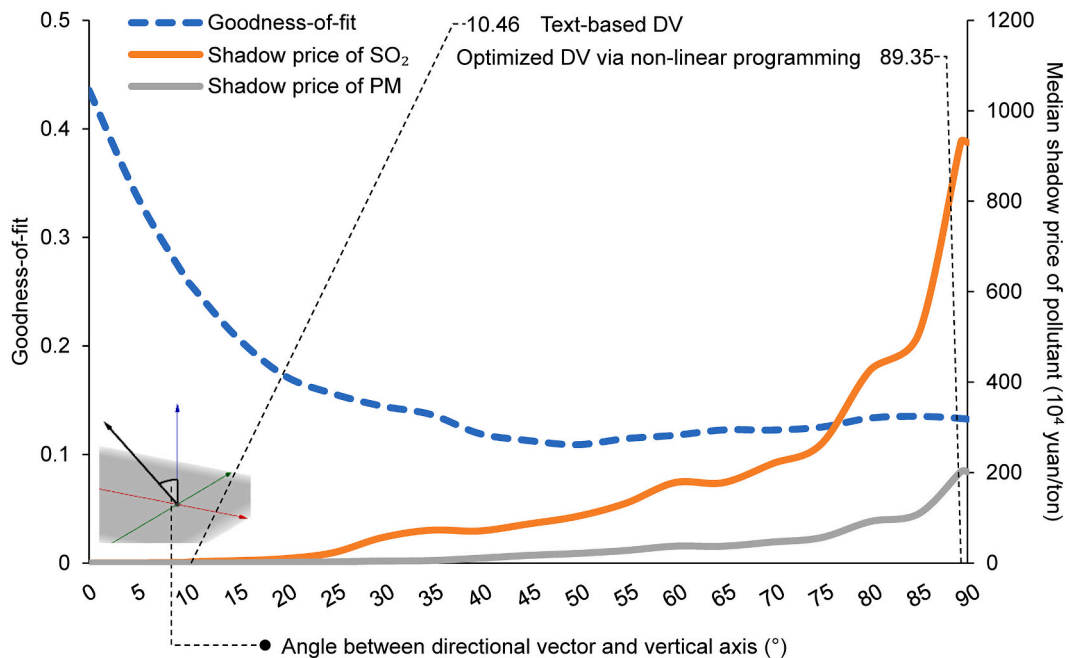


Fig. 10. Goodness-of-fit and shadow price under different directional vectors. Notes: The definition of the goodness-of-fit is shown in Eq. (8).

cities, and pollutants are revealed.

There is a broad application prospect for our new DV. Because any research involving the DDF will inevitably require the selection of DV, our new DV can be widely applied. Therefore, it can be combined with many methods in performance analysis or externality research, serving as an effective supplement. For example, in the Meta-frontier analysis considering heterogeneous production technology, our text-based DV can be applied separately in each group. In this way the inter-group heterogeneities can be better reflected. Another example is a combination with the non-linear programming by Färe et al. (2017). The measure of preference by our text-based DV can be added as a constraint, providing a realistic basis for this method.

The potential limitation of this DV lies in the availability of policy text. Because different countries or regions may have different political and economic environments, it is not certain whether there is any available policy text that can be used to reveal the preference. In addition, due to the limitation of data availability, the data presented in this

paper is not up-to-date. Based on richer text information and improved models, subsequent studies can experiment with “dynamic” directional vectors to portray the changes of preference in more detail.

**CRedit authorship contribution statement**

**Yi Zhou:** Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Wenji Zhou:** Methodology. **Chu Wei:** Conceptualization, Methodology, Writing – review & editing, Project administration, Funding acquisition.

**Acknowledgements**

This research is supported by the National Natural Science Foundation of China (72261147760), and the Fundamental Research Funds for the Central Universities and the Research Funds of Renmin University of China (23XNH158).

**Appendix A. Word list of the lexicon reflecting policy preference**

Preference for economic growth (88 words)					
Chinese	Chinese Pinyin	English	Chinese	Chinese Pinyin	English
发展	Fazhan	Development	商品	Shangpin	Commodity
建设	Jianshe	Construction	经济体制	Jingji tizhi	Economic system
经济	Jingji	Economy	对外开放	Duiwai kaifang	Opening up
推进	Tuijin	Advance	经济社会	Jingji shehui	Economy and society
加快	Jiakuai	Speed up	生产总值	Shengchan zongzhi	Total output value
技术	Jishu	Technology	经济效益	Jingji xiaoyi	Economic benefit
生产	Shengchan	Production	效益	Xiaoyi	Benefit
改革	Gaige	Reform	进口	Jinkou	Import
能力	Nengli	Capacity	供应	Gongying	Supply
促进	Cujin	Promote	资本	Ziben	Capital
创新	Chuangxin	Innovation	新兴产业	Xinxing chanye	Emerging industry
市场	Shichang	Market	商业	Shangye	Commerce
推动	Tuidong	Promote	产业化	Chanyehua	Industrialization
增长	Zengzhang	Growth	消费品	Xiaofeipin	Consumer goods
产业	Chanye	Industry	创新能力	Chuangxin nengli	Innovation ability
产品	Chanpin	Product	产业结构	Chanye jie gou	Industrial structure
投资	Touzi	Investment	融资	Rongzi	Financing
工业	Gongye	Industry	资产	Zichan	Asset
提升	Tisheng	Elevate	振兴	Zhenxing	Revitalization
就业	Jiuye	Employment	利用外资	Liyong waizi	Use of foreign investment
基础设施	Jichu sheshi	Infrastructure	进出口	Jinchukou	Import and export
开放	Kaifang	Open	市场化	Shichanghua	Marketization
现代化	Xiandaihua	Modernization	固定资产	Guding zichan	Fixed asset
规模	Guimo	Scale	产业链	Chanyelian	Industry chain
服务业	Fuwuye	Service industry	改革开放	Gaige kaifang	Reform and opening up
国民经济	Guomin jingji	National economy	国有资产	Guoyou zichan	State-owned assets
大型	Daxing	Large-scale	增产	Zengchan	Increase production
经营	Jingying	Operate	脱贫	Tuopin	Poverty elimination
收入	Shouru	Income	生产力	Shengchanli	Productivity
消费	Xiaofei	Consumption	市场经济	Shichang jingji	Market economy
制造	Zhizao	Manufacture	外贸	Waimao	Foreign trade
资金	Zijin	Fund	产业政策	Chanye zhengce	Industrial policy
金融	Jinrong	Finance	市场主体	Shichang zhuti	Market entity
大力发展	Dali fazhan	Vigorously develop	财政收入	Caizheng shouru	Fiscal revenue
出口	Chukou	Export	国民收入	Guomin shouru	National income
贸易	Maoyi	Trade	外资	Waizi	Foreign investment
价格	Jiage	Price	经济区	Jingji qu	Economic zone
财政	Caizheng	Public finance	经济特区	Jingji tequ	Special economic zone
产量	Chanliang	Yield	总产值	Zongchanzhi	Total output value
劳动	Laodong	Labor	第三产业	Disan chanye	Tertiary industry
生产能力	Shengchan nengli	Production capacity	劳动生产率	Laodong shengchanlyu	Labor productivity
供给	Gongji	Supply	产能	Channeng	Production capacity
新增	Xinzeng	Increase	集体经济	Jiti jingji	Collective economy
制造业	Zhizaoye	Manufacturing industry	优势产业	Youshi chanye	Dominant industry
Preference for environmental protection (20 words)					
Chinese	Chinese Pinyin	English	Chinese	Chinese Pinyin	English
资源	Ziyuan	Resource	排放	Paifang	Emission
保护	Baohu	Protection	生态环境	Shengtai huanjing	Ecological environment

(continued on next page)

(continued)

Preference for environmental protection (20 words)					
Chinese	Chinese Pinyin	English	Chinese	Chinese Pinyin	English
生态	Shengtai	Ecology	节水	Jieshui	Water-saving
环境	Huanjing	Environment	污染物	Wuranwu	Pollutant
能源	Nengyuan	Energy	自然	Ziran	Nature
节能	Jieneng	Energy-saving	气候变化	Qihou bianhua	Climate change
绿色	Lyuse	Green	新能源	Xinnengyuan	New energy
污染	Wuran	Pollution	自然资源	Ziran ziyuan	Natural resources
环境保护	Huanjing baohu	Environmental protection	自然保护区	Ziran baohuqu	Natural reserve
环保	Huanbao	Environmental protection	环境治理	Huanjing zhili	Environmental governance

Notes: From top to bottom and left to right, the words are sorted in descending order according to their word frequencies in the corpus.

### Appendix B. Robust analysis for word selection

Since the word selection for the lexicon may have an impact on the results, here we conduct a robust analysis.

The words in the lexicon are selected based on their scoring results. Specifically, in a value range of -10 to 10, the words are selected with the absolute value of score not lower than 6. In this regard, we define a percentage index called “word intensity”, which is equal to the absolute value of score divided by 10, ranging from 0 to 100%. This index can indicate the “intensity” of the word in reflecting the preference for “economic growth” or “environmental protection”. The higher the value, the more obvious the preference attitude of the word.

Based on the concept of this index, the existing results of this paper can be expressed as the benchmark results with word intensity not lower than 60%. Therefore, here we increase the minimum limit of the word intensity to 70%, 80%, and 90%, respectively, to minimize the uncertainty caused by “neutral” words. Then we repeat the entire solving process separately. Table B.1 presents the new results of text analysis and text-based DV after changing the word intensity. Table B.2 compares a series of estimates for efficiency and cost with different word intensity.

**Table B.1**  
Results of text analysis and text-based DV with different word intensity.

Text analysis and text-based DV	Benchmark results (word intensity $\geq 60\%$ )	Results after changing the word intensity		
		$\geq 70\%$	$\geq 80\%$	$\geq 90\%$
Number of words with preference for				
Economic growth	88	63	35	22
Environmental protection	20	17	13	7
Frequency of words with preference for				
Economic growth	5736	5099	4013	3183
Environmental protection	749	731	639	246
Preference ratio $\rho$	7.66	6.98	6.28	12.94
Text-based DV ( $g_y, -g_{b,1}, -g_{b,2}$ )				
$g_y$	0.98	0.98	0.98	0.99
$-g_{b,1}$	-0.13	-0.14	-0.16	-0.08
$-g_{b,2}$	-0.13	-0.14	-0.16	-0.08

Notes: Word intensity is a percentage index, which is equal to the absolute value of score divided by 10. For the words with intensity  $\geq 60\%$ , we also conduct a re-judgment and discussion.

Table B.2 shows that the estimates of efficiency and cost do not change significantly after changing the word intensity. With the new intensity limit of 70%, 80% and 90%, the changes in the DDF and technical efficiency are almost all within 10%. The change of shadow price is slightly larger in comparison, with the largest variation of about 60%. However, the absolute difference is small compared to the benchmark results. In terms of the values of tens of thousands, hundreds of thousands, or even millions of yuan/ton in other related studies (see Table 4), this difference is almost negligible. Therefore, it can be concluded that the existing results of this paper are robust.

**Table B.2**  
Median estimates of efficiency and cost with different word intensity.

Variable	Benchmark results (word intensity $\geq 60\%$ )	Results and their relative changes after changing the word intensity		
		$\geq 70\%$	$\geq 80\%$	$\geq 90\%$
DDF	2.32	2.32 (-0.19%)	2.29 (-1.25%)	2.30 (-0.72%)
TE <sub>y</sub>	0.27	0.27 (0.23%)	0.27 (0.71%)	0.27 (0.86%)
TE <sub>SO<sub>2</sub></sub>	0.71	0.70 (-2.24%)	0.68 (-5.55%)	0.81 (12.69%)
TE <sub>PM</sub>	0.75	0.73 (-1.79%)	0.72 (-4.33%)	0.83 (10.41%)
MAC <sub>SO<sub>2</sub></sub> (10 <sup>4</sup> yuan/ton)	2.91	2.90 (-0.14%)	4.64 (59.69%)	1.08 (-62.77%)
MAC <sub>PM</sub> (10 <sup>4</sup> yuan/ton)	0.14	0.15 (8.95%)	0.22 (59.53%)	0.13 (-4.80%)

Notes: DDF denotes the directional distance function. TE represents the partial technical efficiency. MAC is the marginal abatement cost, that is, the shadow price. All the price unit “yuan” is the 1998 constant Chinese Yuan.

### Appendix C. Parameters for province and year dummies in DDF

Parameter $\psi$ for province dummies			
Province	Estimate	Province	Estimate
Shanghai	0.25	Jiangxi	-0.58
Yunnan	-0.48	Hebei	1.43
Inner Mongolia	-0.29	Henan	1.76
Jilin	-0.33	Zhejiang	1.71
Sichuan	1.90	Hubei	0.55
Tianjin	0.89	Hunan	1.11
Ningxia	-0.77	Gansu	-0.15
Anhui	-0.33	Fujian	1.37
Shandong	3.21	Guizhou	-0.29
Shanxi	-0.58	Liaoning	0.85
Guangdong	1.09	Chongqing	-0.21
Guangxi	-0.52	Shaanxi	-0.46
Xinjiang	-0.04	Qinghai	-1.01
Jiangsu	1.30	Heilongjiang	-0.57

Parameter $\varphi$ for year dummies			
Year	Estimate	Year	Estimate
1999	0.23	2005	1.05
2000	0.16	2006	2.07
2001	0.23	2007	2.32
2002	0.40	2008	2.66
2003	0.75	2009	2.60
2004	0.96	2010	3.24

## References

- Agee, M.D., Atkinson, S.E., Crocker, T.D., 2012. Child maturation, time-invariant, and time-varying inputs: their interaction in the production of child human capital. *J. Prod. Anal.* 38, 29–44.
- Aigner, D.J., Chu, S.F., 1968. On estimating the industry production function. *Am. Econ. Rev.* 58, 826–839.
- Amato, I., 2013. Green cement: concrete solutions. *Nature* 494, 300–301.
- Aparicio, J., Pastor, J.T., Vidal, F., 2016. The directional distance function and the translation invariance property. *Omega* 58, 1–3.
- Bostian, M., Färe, R., Grosskopf, S., Lundgren, T., 2022. Prevention or cure? optimal abatement mix. *Environ. Econ. Policy Stud.* 24, 503–531.
- Burtraw, D., Krupnick, A., Palmer, K., Paul, A., Toman, M., Bloyd, C., 2003. Ancillary benefits of reduced air pollution in the US from moderate greenhouse gas mitigation policies in the electricity sector. *J. Environ. Econ. Manag.* 45, 650–673.
- Chambers, R.G., 1998. Input and output indicators. In: Färe, R., Grosskopf, S., Russell, R. R. (Eds.), *Index Numbers in Honour of Sten Malmquist*, 1998. Kluwer Academic Publishers, Boston, pp. 241–272.
- Chen, S., Xue, M., Wang, Z., Tian, X., Zhang, B., 2022. Exploring pathways of phasing out clean heating subsidies for rural residential buildings in China. *Energy Econ.* 116, 106411.
- Cheng, G., Zervopoulos, P.D., 2014. Estimating the technical efficiency of health care systems: a cross-country comparison using the directional distance function. *Eur. J. Oper. Res.* 238, 899–910.
- China Cement Association, 2001–2014. *China Cement Almanac, 2001–2014 ed.* China Building Materials Industry Press, Beijing (in Chinese).
- Coggins, J.S., Swinton, J.R., 1996. The price of pollution: a dual approach to valuing SO<sub>2</sub> allowances. *J. Environ. Econ. Manag.* 30, 58–72.
- Dinga, C.D., Wen, Z., 2022. China's green deal: can China's cement industry achieve carbon neutral emissions by 2060? *Renew. Sust. Energ. Rev.* 155, 111931.
- D'Inverno, G., De Witte, K., 2020. Service level provision in municipalities: a flexible directional distance composite indicator. *Eur. J. Oper. Res.* 286, 1129–1141.
- Du, L., Hanley, A., Wei, C., 2015. Estimating the marginal abatement cost curve of CO<sub>2</sub> emissions in China: provincial panel data analysis. *Energy Econ.* 48, 217–229.
- Elhacham, E., Ben-Uri, L., Grozovski, J., Bar-On, Y.M., Milo, R., 2020. Global human-made mass exceeds all living biomass. *Nature* 588, 442–444.
- Färe, R., Grosskopf, S., Noh, D.-W., Weber, W., 2005. Characteristics of a polluting technology: theory and practice. *J. Econ.* 126, 469–492.
- Färe, R., Grosskopf, S., Weber, W.L., 2006. Shadow prices and pollution costs in U.S. agriculture. *Ecol. Econ.* 56, 89–103.
- Färe, R., Martins-Filho, C., Vardanyan, M., 2010. On functional form representation of multi-output production technologies. *J. Prod. Anal.* 33, 81–96.
- Färe, R., Grosskopf, S., Pasurka, C.A., Weber, W.L., 2012. Substitutability among undesirable outputs. *Appl. Econ.* 44, 39–47.
- Färe, R., Pasurka, C., Vardanyan, M., 2017. On endogenizing direction vectors in parametric directional distance function-based models. *Eur. J. Oper. Res.* 262, 361–369.
- Foltz, J.D., Barham, B.L., Chavas, J.-P., Kim, K., 2012. Efficiency and technological change at US research universities. *J. Prod. Anal.* 37, 171–186.
- Hasanbeigi, A., Morrow, W., Masanet, E., Sathaye, J., Xu, T., 2013. Energy efficiency improvement and CO<sub>2</sub> emission reduction opportunities in the cement industry in China. *Energy Policy* 57, 287–297.
- He, W., Zhang, B., Li, Y., Chen, H., 2021. A performance analysis framework for carbon emission quota allocation schemes in China: perspectives from economics and energy conservation. *J. Environ. Manag.* 296, 113165.
- Hoberg, G., Phillips, G., 2016. Text-based network industries and endogenous product differentiation. *J. Polit. Econ.* 124, 1423–1465.
- Hu, A., 2016. The five-year plan: a new tool for energy saving and emissions reduction in China. *Adv. Clim. Chang. Res.* 7, 222–228.
- Hua, S., Tian, H., Wang, K., Zhu, C., Gao, J., Ma, Y., Xue, Y., Wang, Y., Duan, S., Zhou, J., 2016. Atmospheric emission inventory of hazardous air pollutants from China's cement plants: temporal trends, spatial variation characteristics and scenario projections. *Atmos. Environ.* 128, 1–9.
- Ji, D., Zhou, P., 2020. Marginal abatement cost, air pollution and economic growth: evidence from Chinese cities. *Energy Econ.* 86, 104658.
- Ji, D., Zhou, P., Wu, F., 2021. Do marginal abatement costs matter for improving air quality? Evidence from China's major cities. *J. Environ. Manag.* 286, 112123.
- Kaneko, S., Fujii, H., Sawazu, N., Fujikura, R., 2010. Financial allocation strategy for the regional pollution abatement cost of reducing sulfur dioxide emissions in the thermal power sector in China. *Energy Policy* 38, 2131–2141.
- Kesicki, F., Ekins, P., 2012. Marginal abatement cost curves: a call for caution. *Clim. Pol.* 12, 219–236.
- Li, Y., Wang, Z., Wang, K., Zhang, B., 2021. Fuel economy of Chinese light-duty car manufacturers: an efficiency analysis perspective. *Energy* 220, 119622.
- Li, H., Wang, Z., Zhang, B., 2023. How social interaction induce energy-saving behaviors in buildings: interpersonal & passive interactions v.s. public & active interactions. *Energy Econ.* 118, 106515.
- Lin, B., Zhang, Z., 2016. Carbon emissions in China's cement industry: a sector and policy analysis. *Renew. Sust. Energ. Rev.* 58, 1387–1394.
- Liu, J., Tong, D., Zheng, Y., Cheng, J., Qin, X., Shi, Q., Yan, L., Lei, Y., Zhang, Q., 2021. Carbon and air pollutant emissions from China's cement industry 1990–2015: trends, evolution of technologies, and drivers. *Atmos. Chem. Phys.* 21, 1627–1647.
- Ma, C., Hailu, A., 2016. The marginal abatement cost of carbon emissions in China. *Energy J.* 37, 111–127.
- Ma, C., Hailu, A., You, C., 2019. A critical review of distance function based economic research on China's marginal abatement cost of carbon dioxide emissions. *Energy Econ.* 84, 104533.
- Madloul, N.A., Saidur, R., Hossain, M.S., Rahim, N.A., 2011. A critical review on energy use and savings in the cement industries. *Renew. Sust. Energ. Rev.* 15, 2042–2060.
- Murty, M.N., Kumar, S., Dhavala, K.K., 2007. Measuring environmental efficiency of industry: a case study of thermal power generation in India. *Environ. Resour. Econ.* 38, 31–50.
- Murty, S., Robert Russell, R., Levkoff, S.B., 2012. On modeling pollution-generating technologies. *J. Environ. Econ. Manag.* 64, 117–135.
- Nakaishi, T., 2021. Developing effective CO<sub>2</sub> and SO<sub>2</sub> mitigation strategy based on marginal abatement costs of coal-fired power plants in China. *Appl. Energy* 294, 116978.
- National Bureau of Statistics, 1983–2021. *China Statistical Yearbook, 1983–2021 Ed.* China Statistics Press, Beijing (in Chinese).

- National Bureau of Statistics, 2021. China Statistical Yearbook on Environment, 2021 ed. China Statistics Press, Beijing (in Chinese).
- Pang, J., Shi, Y., Feng, X., Wu, S., Sun, W., 2014. Analysis on impacts and co-abatement effects of implementing the low carbon cement standard. *Adv. Clim. Chang. Res.* 5, 41–50.
- Pastor, J.T., Lovell, C.A.K., Aparicio, J., 2020. Defining a new graph inefficiency measure for the proportional directional distance function and introducing a new Malmquist productivity index. *Eur. J. Oper. Res.* 281, 222–230.
- Picazo-Tadeo, A.J., Beltrán-Estevé, M., Gómez-Limón, J.A., 2012. Assessing eco-efficiency with directional distance functions. *Eur. J. Oper. Res.* 220, 798–809.
- Price, L., Levine, M.D., Zhou, N., Fridley, D., Aden, N., Lu, H., McNeil, M., Zheng, N., Qin, Y., Yowargana, P., 2011. Assessment of China's energy-saving and emission-reduction accomplishments and opportunities during the 11th five year plan. *Energy Policy* 39, 2165–2178.
- Schneider, M., Romer, M., Tschudin, M., Bolio, H., 2011. Sustainable cement production - present and future. *Cem. Concr. Res.* 41, 642–650.
- Shan, Y., Zhou, Y., Meng, J., Mi, Z., Liu, J., Guan, D., 2019. Peak cement-related CO<sub>2</sub> emissions and the changes in drivers in China. *J. Ind. Ecol.* 23, 959–971.
- Shen, W., Liu, Y., Yan, B., Wang, J., He, P., Zhou, C., Huo, X., Zhang, W., Xu, G., Ding, Q., 2017. Cement industry of China: driving force, environment impact and sustainable development. *Renew. Sust. Energ. Rev.* 75, 618–628.
- Singbo, A.G., Oude Lansink, A., Emvalomatis, G., 2014. Estimating farmers' productive and marketing inefficiency: an application to vegetable producers in Benin. *J. Prod. Anal.* 42, 157–169.
- Tan, H., Zheng, Q., Wang, H., 2021. Local government big data industry policy: tool preferences and compatibility analysis based on the text of Guizhou policy. *Chin. Public Adm.* 52–58 (in Chinese).
- Teng, F., Wang, X., Pan, X., Yang, X., 2014. Understanding marginal abatement cost curves in energy-intensive industries in China: insights from comparison of different models. *Energy Procedia* 61, 318–322.
- Vardanyan, M., Noh, D.-W., 2006. Approximating pollution abatement costs via alternative specifications of a multi-output production technology: a case of the US electric utility industry. *J. Environ. Manag.* 80, 177–190.
- Wang, Y., Li, M., 2017. Study on local government attention of ecological environment governance: based on the text analysis of government work report in 30 provinces and cities (2006-2015). *China Population, Resources and Environment* 27, 28–35 (in Chinese).
- Wang, K., Che, L., Ma, C., Wei, Y., 2017. The shadow price of CO<sub>2</sub> emissions in China's iron and steel industry. *Sci. Total Environ.* 598, 272–281.
- Wang, Q., Wang, Y., Hang, Y., Zhou, P., 2019. An improved production-theoretical approach to decomposing carbon dioxide emissions. *J. Environ. Manag.* 252, 109577.
- Wang, Z., Chen, H., Huo, R., Wang, B., Zhang, B., 2020. Marginal abatement cost under the constraint of carbon emission reduction targets: an empirical analysis for different regions in China. *J. Clean. Prod.* 249, 119362.
- Wei, X., Zhang, N., 2020. The shadow prices of CO<sub>2</sub> and SO<sub>2</sub> for Chinese coal-fired power plants: a partial frontier approach. *Energy Econ.* 85, 104576.
- Wei, C., Löschel, A., Liu, B., 2013. An empirical analysis of the CO<sub>2</sub> shadow price in Chinese thermal power enterprises. *Energy Econ.* 40, 22–31.
- Wei, C., 2021. Economic loss and environmental gain from regulation: examining the two-fold effect using data from Chinese cities. *Lett. Spat. Resour. Sci.* 15, 423–434.
- Wei, X., Chang, Y.-T., Kwon, O.-K., Zhang, N., 2021. Potential gains of trading CO<sub>2</sub> emissions in the Chinese transportation sector. *Transp. Res. Part D: Transp. Environ.* 90, 102639.
- Xie, H., Shen, M., Wei, C., 2016. Technical efficiency, shadow price and substitutability of Chinese industrial SO<sub>2</sub> emissions: a parametric approach. *J. Clean. Prod.* 112, 1386–1394.
- Yang, X., Teng, F., Wang, G., 2013. Incorporating environmental co-benefits into climate policies: a regional study of the cement industry in China. *Appl. Energy* 112, 1446–1453.
- Yang, Q., Kaneko, S., Fujii, H., Yoshida, Y., 2017. Do exogenous shocks better leverage the benefits of technological change in the staged elimination of differential environmental regulations? Evidence from China's cement industry before and after the 2008 great Sichuan earthquake. *J. Clean. Prod.* 164, 1167–1179.
- Xie, H., Wang, X., Shen, M., Wei, C., 2021. Abatement costs of combatting industrial water pollution: convergence across Chinese provinces. *Environ. Dev. Sustain.* 24, 10752–10767.
- Yang, J., Cheng, J., Zou, R., Geng, Z., 2021. Industrial SO<sub>2</sub> technical efficiency, reduction potential and technology heterogeneities of China's prefecture-level cities: a multi-hierarchy meta-frontier parametric approach. *Energy Econ.* 104, 105626.
- Yu, Y., Zhang, N., 2021. Low-carbon city pilot and carbon emission efficiency: quasi-experimental evidence from China. *Energy Econ.* 96, 105125.
- Yu, M.M., See, K.F., Hsiao, B., 2021. Integrating group frontier and metafrontier directional distance functions to evaluate the efficiency of production units. *Eur. J. Oper. Res.* 301, 254–276.
- Zeng, S., Jiang, X., Su, B., Nan, X., 2018. China's SO<sub>2</sub> shadow prices and environmental technical efficiency at the province level. *Int. Rev. Econ. Financ.* 57, 86–102.
- Zhang, T., Wu, C., Li, B., Wang, J., Ravat, R., Chen, X., Wei, J., Yu, Q., 2019. Linking the SO<sub>2</sub> emission of cement plants to the sulfur characteristics of their limestones: a study of 80 NSP cement lines in China. *J. Clean. Prod.* 220, 200–211.
- Zhang, B., Chen, H., Du, Z., Wang, Z., 2020. Does license plate rule induce low-carbon choices in residents' daily travels: motivation and impacts. *Renew. Sust. Energ. Rev.* 124, 109780.
- Zhang, C., Yu, B., Chen, J., Wei, Y., 2021a. Green transition pathways for cement industry in China. *Resour. Conserv. Recycl.* 166, 105355.
- Zhang, N., Huang, X., Liu, Y., 2021b. The cost of low-carbon transition for China's coal-fired power plants: a quantile frontier approach. *Technol. Forecast. Soc. Chang.* 169, 120809.
- Zhang, B., Xin, Q., Tang, M., Niu, N., Du, H., Chang, X., Wang, Z., 2022. Revenue allocation for interfirm collaboration on carbon emission reduction: complete information in a big data context. *Ann. Oper. Res.* 316, 93–116.
- Zheng, C., Zhang, H., Cai, X., Chen, L., Liu, M., Lin, H., Wang, X., 2021. Characteristics of CO<sub>2</sub> and atmospheric pollutant emissions from China's cement industry: a life-cycle perspective. *J. Clean. Prod.* 282, 124533.
- Zhou, P., Ang, B.W., Wang, H., 2012. Energy and CO<sub>2</sub> emission performance in electricity generation: a non-radial directional distance function approach. *Eur. J. Oper. Res.* 221, 625–635.
- Zhou, P., Zhou, X., Fan, L., 2014. On estimating shadow prices of undesirable outputs with efficiency models: a literature review. *Appl. Energy* 130, 799–806.