Estimating the shadow prices of SO$_2$ and NO$_x$ for U.S. coal power plants: A convex nonparametric least squares approach

Maethee Mekaroonreung, Andrew L. Johnson

Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX 77843-3131, USA

**A R T I C L E   I N F O**

Article history:
Received 28 February 2011
Received in revised form 30 December 2011
Accepted 22 January 2012
Available online 2 February 2012

**JEL classification:**
C14
D24
Q53

**Keywords:**
Frontier estimation
Nonparametric regression
Parametric programming
Shadow pricing

**A B S T R A C T**

Weak disposability between outputs and pollutants, defined as a simultaneous proportional reduction of both outputs and pollutants, assumes that pollutants are byproducts of the output generation process and that a firm can “freely dispose” of both by scaling down production levels, leaving some inputs idle. Based on the production axioms of monotonicity, convexity and weak disposability, we formulate a convex nonparametric least squares (CNLS) quadratic optimization problem to estimate a frontier production function assuming either a deterministic disturbance term consisting only of inefficiency, or a composite disturbance term composed of both inefficiency and noise. The suggested methodology extends the stochastic seminonparametric envelopment of data (StoNED) described in Kuosmanen and Kortelainen (2011). Applying the method to estimate the shadow prices of SO$_2$ and NO$_x$ generated by U.S. coal power plants, we conclude that the weak disposability StoNED method provides more consistent estimates of market prices.

© 2012 Elsevier B.V. All rights reserved.

**1. Introduction**

Coal power plants generate 47–56% of the electricity consumed in the U.S since 1989 (EIA, 2010). However, burning coal produces several byproduct pollutants, notably sulfur oxide (SO$_2$) and nitrogen oxide (NO$_x$), the major cause of acid rain. To address this problem, the Clean Air Act Amendments of 1990 (CAAA) set goals to reduce annual SO$_2$ emissions by 10 million tons and NO$_x$ by 2 million tons from 1980 levels via a two-phase tightening of the restrictions placed primarily on coal plants (EPA, 2007). Phase I (1995–1999) regulated 445 boiler units at mostly coal plants and Phase II (2000–present) regulated over 2000 boiler units with a capacity greater than 25 MW at all fossil fuel plants. In 2011, the U.S. Environmental Protection Agency (EPA) released new environmental regulations requiring coal power plants to lower emissions of 84 toxic chemical levels within four years (EPA, 2011a).

An analysis of the effect of these regulations is helpful in understanding the impacts in terms of reductions in pollution and the associated costs for continued reductions. For this purpose we estimate a frontier production function, as first proposed by Farrell (1957). Data

---

This paper has benefited from comments and suggestions on earlier drafts from an anonymous referee and the editorial services of Ann Stewart, bondcliffs@gmail.com.

Corresponding author.
E-mail addresses: pokkynz_77@tamu.edu (M. Mekaroonreung), ajohnson@tamu.edu (A.L. Johnson).

to estimate the technical efficiency of the U.S. oil refineries. See Zhou et al. (2008) for a summary of other DEA weak disposability applications in energy and environmental studies.

Recently, Sueyoshi and Goto (2011) proposed the concept of natural and managerial disposability and applied the concepts to a DEA frontier. A non-radial efficiency measure compared environmental performances and computed the returns to scale and damages to scale of national oil companies in several countries and international oil companies. This paper will focus on the more standard weak disposability assumption as the frontier for undesirable outputs implied by managerial disposability violates free disposability of inputs.

The implementation of the weak disposability assumption relative to a variable returns to scale (VRS) frontier has been subject to considerable debate. For instance, Färe and Grosskopf (2003) proposed a new model to construct a VRS weakly disposable production possibility set by introducing a single abatement factor across all firms whereas Kuosmanen (2005) used a non-uniform abatement factor across firms. In demonstrating that a production possibility set constructed by a single abatement factor model does not satisfy convexity, Kuosmanen and Podinovski (2008), proved that using non-uniform abatement factors allows the estimation of a VRS weakly disposable production possibility set that satisfies standard production axioms and the minimum extrapolation principle.

Many previous studies have estimated the shadow prices of undesirable outputs using distance functions. A ratio of the derivative of the distance function with respect to desirable output and the derivative of the distance function with respect to undesirable output characterizes the relative shadow price of the undesirable output, and parametric or nonparametric approaches can be used to estimate the distance function. The parametric approach is more widely used, because functions are everywhere differentiable. Färe et al. (1993) used an output distance function with the translog functional form to estimate a shadow price of four undesirable pollutants for 1976 data describing pulp and paper mills in Michigan and Wisconsin. Coggins and Swinton (1996) took the same approach to estimate the shadow price of SO2 for Wisconsin coal plants in 1990–1992. Färe et al. (2005) used a quadratic directional output distance function to estimate both technical efficiency and a shadow price of SO2 for the U.S. electric utilities in 1993 and 1997. Despite its common usage, the parametric approach can be biased if the functional form is misspecified. Alternatively, a nonparametric approach, specifically DEA, can estimate a production frontier and the shadow prices of pollutants. Boyd et al. (1996) used a DEA production function to estimate the shadow price of SO2 for coal plants. Lee et al. (2002) used DEA when accounting for technical inefficiency to derive the shadow prices of SO2, NOx and total suspend particulates (TSP) for Korean coal- and oil-burning plants in 1990–1995. Researchers also acknowledge some major limitations of the alternative approach: greater sensitivity to outliers, and the use of only a few observations to construct the production frontier. Moreover, DEA as a deterministic method does not incorporate statistical noise, and thus the observations of the production units must be observed without error and the production model specified without omitting any inputs or outputs.

Such drawbacks motivated the development of other nonparametric methods such as Convex Nonparametric Least Squares (CNLS), Kuosmanen (2006, 2008), which uses all available data to estimate a piecewise linear production function satisfying production axioms such as continuity, monotonicity and concavity. Kuosmanen and Johnson (2010) have shown that DEA is a special case of CNLS with sign constraints on error terms. To decompose statistical noise and inefficiency for cross-sectional data in a semi-parametric fashion, Kuosmanen and Kortelainen (2011) have proposed a two-stage method called Stochastic Non-parametric Envelopment of Data (StoNED). It applies CNLS in the first stage to estimate an average production function and estimates the conditional expectation of inefficiency based on the CNLS residuals in the second stage.

The advantages of CNLS and StoNED over DEA motivated us to apply them to estimate a weak disposability production frontier. While DEA with weak disposability is well studied, to the best of our knowledge we are unaware of research that incorporates weak disposability with CNLS and StoNED. We describe our proposed model and apply it to measure the technical efficiency and to jointly estimate the shadow prices of SO2 and NOx using the U.S. coal power plants during Phase II of CAAA. To our knowledge there are no studies on the productive performance and shadow prices of SO2 and NOx using the U.S. coal power plants during Phase II of CAAA. The paper is organized as follows: the next section describes a nonparametric method of estimating a production function under weak disposability and the associated technical efficiency and shadow prices of SO2 and NOx. Section 3 describes the data set of 336 boilers of the U.S. bituminous coal power plants in operation from 2000 to 2008. Section 4 presents the analysis and discusses the results and Section 5 summarizes the conclusions.

2. Model

2.1. A production possibility set assuming weak disposability

For each firm \( i \in 1, \ldots, n \) let \( x \in \mathbb{R}^M_+ \) be a vector of inputs, \( y \in \mathbb{R}^T_+ \) be a vector of good outputs and \( b \in \mathbb{R}^R_+ \) be a vector of bad outputs. The production possibility set is defined as \( T=\{(x,y,b):x \text{ can produce } (y,b)\} \). The assumptions defining the production possibility set are: 1. \( T \) is convex 2. There are variable returns to scale

Originally proposed by Shephard (1970), the following axioms regarding production are restated when undesirable outputs are also produced:

3. Free disposability of inputs
   \[
   \text{If } (x,y,b) \in T \text{ and } x \geq x^*, \text{ then } (x^*, y, b) \in T.
   \]

4. Free disposability of outputs
   \[
   \text{If } (x,y,b) \in T \text{ and } y^* \leq y, \text{ then } (x,y^*, b) \in T.
   \]

5. Weak disposability between outputs and pollutants
   \[
   \text{If } (x,y,b) \in T \text{ and } 0 \leq \varphi \leq 1, \text{ then } (x,\varphi y, cb) \in T.
   \]

Based on the production possibility axioms stated above, the variable returns to scale weakly disposable production possibility set \( T \) can be written as:

\[
T = \left\{ (x,y,b) \in \mathbb{R}^M_+ \times \mathbb{R}^T_+ \times \mathbb{R}^R_+ \mid
\begin{array}{c}
\sum_{i=1}^n \lambda_i = 1,
\lambda_i, \mu_i \geq 0,
\sum_{i=1}^n \mu_i y_i \geq \sum_{i=1}^n \lambda_i y_i;
\sum_{i=1}^n \lambda_i b_i \geq \sum_{i=1}^n (\lambda_i + \mu_i) x_i,
\lambda_i, \mu_i \geq 0
\end{array}
\right\}
\]

(1)

where \( \lambda, \mu \) allows the convex combination of observed firms and \( \mu \) allows firms to scale down both outputs and pollutants while maintaining the same level of inputs.

Formulation (1) differs from the Kuosmanen (2005) formulation in that the inequality sign in the pollutant constraints implies a negative shadow price on additional pollution and satisfies the economic intuition that pollutants incur costs to firms.
Using the weak disposable production possibility $T$ in (1), the variable returns to scale output-oriented weak disposability DEA estimator can be written as:

$$\begin{align*}
\max_{\theta_0, \Lambda, \mu} & 
\sum_{i=1}^{n} \sum_{t=1}^{S} y_{it}^\theta \
\text{s.t.} & 
\sum_{i=1}^{n} \sum_{t=1}^{S} x_{it} \geq \sum_{i=1}^{n} \sum_{t=1}^{S} y_{it} \theta_0, \
& \sum_{i=1}^{n} \sum_{t=1}^{S} b_{ij} \leq b_{ij} \quad \forall j = 1, ..., J \
& \sum_{i=1}^{n} (\lambda_i + \mu_i) x_{im} \leq x_{om} \quad \forall m = 1, ..., M \
& \sum_{i=1}^{n} \lambda_i = 1 \
& \lambda_i, \mu_i \geq 0 \quad \forall i = 1, ..., n.
\end{align*}$$

(2)

where $y_{om}, b_{om}, x_{om}$, and $\theta_0$ are outputs, bad outputs, inputs and the technical efficiency for specific firm $o$.

The DEA problem (2) constructs the weak disposability production frontier and estimates technical efficiency as the radial expansion of outputs.

2.2. Production frontier estimation

Consider a single output production function with a multiplicative disturbance term

$$y_i = f(x_i, b_i) \exp(e_i) \forall i = 1, ..., n$$

(3)

where $f(x_i, b_i)$ is the production function satisfying continuity, monotonicity, concavity and weak disposability and $e_i$ is the disturbance term. Note that the production function in (3) treats pollutants as independent variables following Cropper and Oates (1992), who defined this treatment as the standard approach to including pollutants within the environmental economics literature. Treating pollutants as independent variables has been used in several papers such as Pittman (1981) and Considine and Larson (2006).

Our motivations to employ a multiplicative disturbance model are twofold. First, as suggested in Kuosmanen and Kortelainen (2011), the multiplicative model allows the direct imposition of the assumptions of Constant Returns to Scale (CRS), Non-Increasing Returns to Scale (NIRS) or Non-decreasing Returns to Scale (NDRS). Specifically, CRS, NIRS and NDRS do not hold after an additive shift of the estimated frontier production function. Since the assumption of weak disposability between an output and pollutants requires the origin to be part of the convex production possibility set similar to the NIRS, the multiplicative disturbance term model is appealing. Second, the multiplicative disturbance term model helps to control for heteroskedasticity resulting from increased variability in output levels for production units operating at larger scale sizes.

Applying the log transformation to (3) gives:

$$e_i = \ln(y_i) - \ln(f(x_i, b_i)).$$

(4)

To estimate the weak disposability production frontier, we apply the CNLS technique to minimize the sum of the above multiplicative disturbances squared and assume the composition of $e_i$ to be

1. Deterministic (all deviations are attributed to inefficiency) or
2. Composite (mixture of inefficiency and random noise) or
3. Random (all deviations are random noise).

The results of estimating (4) by minimizing the sum of squared deviations is the production function under the assumption of random disturbances. If the disturbances are assumed to be deterministic, we could apply a one-stage method by solving the CNLS problem with sign constraints on the disturbances. The results of applying a one-stage method define an estimated production frontier and technical efficiencies. If the disturbance terms are assumed to be a mixture, we could apply the two-stage StoNED method, solving the CNLS problem with no sign constraint on the disturbances and then decomposing the CNLS residuals into statistical noises and technical efficiencies using Jondrow et al. (1982) see also Kuosmanen and Kortelainen (2011). The estimated averaged CNLS production function is shifted by the average technical efficiency level to obtain a production frontier. Below, we elaborate on these deterministic, composite and random disturbance term assumptions.

2.2.1. Deterministic disturbance term

We assume that there is no statistical noise in the data; thus, any deviations from the estimated frontier are due to technical efficiency. Specifically:

$$e_i = -u_i \forall i = 1, ..., n$$

(5)

where $u_i \geq 0$ is the firm-specific technical inefficiency.

While noting that the CNLS objective function is to minimize the sum of square disturbances, when all of the disturbances are less than or equal to zero in the deterministic case, we can replace the sum of square disturbances by the sum of disturbances, see Kuosmanen and Johnson (2010). The CNLS problem is then formulated as:

$$\begin{align*}
\min_{\theta_0, \Lambda, \mu} & -\sum_{i=1}^{n} e_i^2 \\
\text{s.t.} & \quad y_i = \ln(y_i) - \ln(f(x_i, b_i)) \quad \forall i = 1, ..., n \\
& \quad \alpha_h + w_i x_i + c_i b_i \leq \epsilon_h + w_i x_i + c_i b_i \quad \forall i, h = 1, ..., n \\
& \quad \epsilon_h \geq 0, c_i \leq 0 \quad \forall i = 1, ..., n.
\end{align*}$$

(6)

The objective function maximizes the sum of the disturbance terms. Intuitively, the CNLS problem (6) estimates a production frontier that makes all firms look as efficient as possible using the minimum extrapolation principle of Banker et al. (1984), also referred to as the benefit of the doubt principle by Moesen and Cherchye (1998). The first equality constraints define the disturbance term. The second inequality constraints comprise a system of Afriat inequalities, Afriat (1972), imposing the underlying production function to be continuous and concave. The third inequality constraints impose the weak disposability between desirable and undesirable outputs. The last constraints enforce monotonicity of both inputs and the costs associated with additional undesirable outputs.

Solving the CNLS problem (6) obtains the production frontier. Technical efficiency is obtained from the estimated CNLS residual, $e_i$, $\forall i$:

$$TE_i = \exp(e_i) \forall i = 1, ..., n.$$ 

(7)

Proposition 1. In a single output case, the deterministic CNLS production function (6) is equivalent to the output-oriented weak disposability DEA production function (2).

Proof. See Appendix.

Proposition 2. The technical efficiency estimates from (7) equal the reciprocal of the technical efficiency estimates from DEA (2).

Proof. See Appendix.

DEA or CNLS with a deterministic disturbance as described above can be used to nonparametricly estimate a weak disposability production function; however, it is not appropriate if the production model is imperfectly specified or the data set contains noise. Again,
CNLS is more advantageous because it can be extended to estimate weak disposability production functions including a model of statistical noise.

2.2.2. Composite disturbance term

Similar to Stochastic Frontier Analysis (SFA) by Aigner et al. (1977), we assume that there is statistical noise in the data; thus any disturbance terms can be written as:

$$
\epsilon_i = v_i - u_i \forall i \in 1, \ldots, n
$$

(8)

where $v_i$ is a random noise component.

As Kuosmanen and Kortelainen (2011) have pointed out, the composite disturbance term in (8) violates the Gauss–Markov properties that $E(\epsilon_i) = E(-u_i) = -\mu_0 \leq 0$ where $\mu$ is the expected technical inefficiency. Therefore, we modify the composite disturbance term in (8). The multiplicative disturbance production function $y_i = f(x_i, b_i)$ is written as:

$$
\ln(y_i) = \ln(f(x_i, b_i)) - \mu_i + |v_i| = \ln(g(x_i, b_i)) + \theta_i \forall i \in 1, \ldots, n
$$

(9)

where $\theta_i = \epsilon_i + \mu$, the modified composite disturbance term. Note that $E(\theta_i) = E(\epsilon_i + \mu) = 0$. The CNLS problem is formulated as:

$$
\min_{\alpha, \omega, \varepsilon} \sum_{i=1}^{n} \theta_i^2
$$

s.t. $\theta_i = \ln(y_i) - \ln\left(\alpha_i + w_i x_i + c_i b_i\right)$ $\forall i \in 1, \ldots, n$

(10)

$$
\alpha_i + w_i x_i + c_i b_i \leq \alpha_k + w_k x_k + c_k b_k
$$

$\forall i, k \in 1, \ldots, n$

$$
\alpha_i + w_i x_i + c_i b_i \geq 0
$$

$\forall i \in 1, \ldots, n$

$$
\theta_i, \epsilon_i \geq 0
$$

$\forall i \in 1, \ldots, n.$

Because CNLS identifies a production function that minimizes the sum of squared disturbances among all production functions that are continuous, monotonic increasing, concave and satisfy the weak disposability assumptions, it is important to check the following condition for the objective function:

**Proposition 3.** The objective function in the CNLS problem (10) is a convex function if and only if $\frac{y_i}{\alpha_i + w_i x_i + c_i b_i} \geq \frac{1}{e} \forall i \in 1, \ldots, n.$

**Proof.** See Appendix.

If the CNLS problem (10) has a convex objective function, then a local optimum to (10) is also a global optimum simplifying the optimizing algorithms needed to find the global optimal solution to (10).

The second stage of CNLS separates the technical efficiency and statistical noise components using the estimated modified CNLS residuals $\delta_i \forall i$ from (10). Assuming that technical efficiency is independent and identically distributed (i.i.d) and has a half normal distribution and that the statistical noise is i.i.d. and normally distributed, $u_i \sim N(0, \sigma_u^2)$ and $v_i \sim N(0, \sigma_v^2)$, the method of moments can be applied (Aigner et al. 1977) as in Kuosmanen and Kortelainen (2011). Specifically,

$$
\sigma_u = \sqrt{\frac{M_3}{\mu^2/(1-\mu^2)}} \quad \text{and} \quad \sigma_v = \sqrt{\frac{M_3-\frac{1}{2}}{\mu^2}} \quad \text{with}
$$

(11)

$$
M_3 = \frac{1}{n} \sum_{i=1}^{n} \left( \hat{\theta}_i - \bar{E}(\hat{\theta}_i) \right)^2 \quad \text{and} \quad M_3 = \frac{1}{n} \sum_{i=1}^{n} \left( \hat{\theta}_i - \bar{E}(\hat{\theta}_i) \right)^2.
$$

Unlike the deterministic disturbance case, after solving the CNLS problem (10), the average production function $g(x_i, b_i)$ is obtained instead of the production frontier. Next, the average production function is multiplied by the expected technical efficiency to estimate the production frontier. Specifically,

$$
\ln(\tilde{g}(x_i, b_i)) = \left[ \ln\left( \hat{f}(x_i, b_i) \right) - \hat{\mu} \right] = \ln\left( \hat{f}(x_i, b_i) \exp(-\hat{\mu}) \right)
$$

(12)

$$
\hat{f}(x_i, b_i) = \hat{g}(x_i, b_i) \exp(\hat{\mu})
$$

where $\hat{\mu} = \sigma_u \sqrt{\frac{\mu}{2}}$. Given $\sigma_u$ and $\sigma_v$, the method introduced in Jondrow et al. (1982) can be used to estimate firm-specific inefficiency. Specifically,

$$
\bar{E}(u_i|\epsilon_i) = -\frac{\epsilon_i \sigma_u^2}{\sigma_u^2 + \sigma_v^2} \quad \text{and} \quad \bar{E}(\epsilon_i|\sigma_v^2) = \frac{\phi\left(\epsilon_i/\sigma_v\right)}{1-\Phi\left(\epsilon_i/(\sigma_v^2)^{1/2}\right)}
$$

(13)

where $\epsilon_i = \delta_i - \mu$ is the standard normal density function and $\phi$ is the standard normal cumulative distribution.

2.2.2. Random disturbance term

Here, the CNLS estimator in (4) is used to obtain the residual directly. Assuming statistical noise is i.i.d. and normally distributed, $v_i \sim N(0, \sigma_v^2)$, the method of moments is applied in the second stage of CNLS to estimate $\sigma_v$. Specifically, $\sigma_v = M_4$ where $M_4 = \frac{1}{n} \sum_{i=1}^{n} \left( \delta_i - \bar{E}(\delta_i) \right)^2$.

We note that composite disturbances and random disturbances differ in that the former is skewed by inefficiency. This skewness can be used to determine if the neoclassical production function (no inefficiency) or the frontier production function (with inefficiency) is the proper model. In Section 4, we use the test proposed by Kuosmanen and Fosgerau (2009) to select between a neoclassical or frontier production function based on the skewness of the residuals. The results provide evidence for the presence of inefficiency.

2.3. Estimating shadow prices of pollutants

Assuming profit-maximizing behavior for each firm, the profit maximization problem for a production process with outputs and pollutants is

$$
\pi(p_y, p_b, p_x) = \max_{y, b, x} p_y y - p_b b - p_x x
$$

s.t. $F(x, b, y) = 0$

(14)

where $p_y = (p_{y_1}, p_{y_2})$, $p_b = (p_{b_1}, p_{b_2})$, and $p_x = (p_{x_1}, p_{x_2})$ represent the price vectors of outputs, pollutants and inputs, respectively. $F(x, b, y)$ is the transformation function corresponding to a multi-output production function. Since we are interested in the shadow prices of pollutants, we impose the constraint $F(x, b, y) = 0$ so that only the frontier of the production possibility set is considered. Problem (15) applies the method of Lagrangian multipliers to (14)

$$
\max_{y, b, x} p_y y - p_b b - p_x x + \zeta F(x, b, y)
$$

(15)

where $\zeta$ is a Lagrangian multiplier of the constraint.

The first-order conditions (FOCs) of the problem (15) are:

$$
p_{yx} + \zeta \frac{\partial F(x, b, y)}{\partial y_x} = 0
$$

$$
p_{by} + \zeta \frac{\partial F(x, b, y)}{\partial y_b} = 0
$$

$$
p_{bx} + \zeta \frac{\partial F(x, b, y)}{\partial y_b} = 0
$$

$$
0 = F(x, b, y).
$$

(16)
The shadow prices can be written as:

\[
P_{bi} = p_{yi} \left( \frac{\partial f(x_i, b_i)}{\partial b_i} \right),
\]

(17)

In the case of a single output production function (\(S = 1\)), the first equality of the FOC (16) can be written as \(p_{y} \cdot \zeta = 0\), that is, the Lagrangian multiplier is equal to the price of the output. Thus, if the price of the output is known, the shadow prices of each pollutant can be estimated using the second equality in (16). The relative shadow prices of pollutants for firm \(i\) are estimated as:

\[
P_{bij} = \frac{\partial f(x_i, b_i)}{\partial b_}\]

(18)

where \(p_{yi}\) is the price of an output for firm \(i\).

Assuming a deterministic disturbance term, solving the CNLS problem (6) estimates a weak disposability production frontier \(\hat{f}(x_i, b_i)\) directly. We obtain the variable \(\frac{\partial f(x_i, b_i)}{\partial b}\) for each firm from the estimated variable \(c_i = (c_{i1}, c_{i2}, ..., c_{ij})\) in (6) directly.2

Under the composite disturbance term assumption, in the first stage we estimate the average weak disposability production function \(\bar{g}(x_i, b_i)\) by solving the CNLS problem (10). In the second stage, we decompose the estimated composite CNLS residuals and obtain the relative shadow prices of pollutants for each firm as \(\frac{\partial g(x_i, b_i) \exp(\mu)}{\partial b}\)

where the variable \(c_{ij} = (c_{i1}, c_{i2}, ..., c_{ij})\) is obtained from solving (10).

3. Data set

Our balance panel boiler-level data characterizes 336 units of the U.S. bituminous coal burning electricity plants in operation from 2000 to 2008. Bituminous coal plants are mostly located in the eastern states and these power plants produce about 50% of the total electricity generated from coal. This form of coal has very high sulfur content. All boilers in the sample are either wall or tangential fired boilers, sub-groups of pulverized coal-fired boilers, which are regulated by the Acid Rain Program.

The output is the annual amount of electricity generated (in Megawatt-hours, MWh). The pollutants are the annual amount of \(\text{SO}_2\) (tons) and \(\text{NO}_x\) (tons). The two inputs are capital and heat. The heat input (mmBtu), calculated by multiplying the quantity of fuel with the fuel’s heat content, is the measure of fuel utilization. Information on electricity generated, amount of pollutants and heat input quantities are reported by the EPA database (EPA, 2011b).

The boiler size (MW), the maximum rated output of a generator under specific conditions, is used as an instrumental variable for capital. The U.S. EPA’s database reported the maximum heat input capacity (mmBtu/h), a unit’s maximum designed hourly heat input rate observed in the past five years, for each boiler unit. We convert the maximum heat input capacity to estimate the boilers’ sizes. The boilers’ sizes in our sample range between 100 and 1426 MW.

Electricity prices ($/MWh) of each utility are reported in EIA861 (EIA, U.S. Energy Information Administration, 2011b). Some of the utilities do not generate electricity; thus, we match our power plants in the sample to those utilities in which they have electricity production data and assume that electricity price in those utilities are the same as in power plants. Following Färe et al. (2005) approach, we assume that all generating boilers in the same power plants have the same electricity prices. We derived electricity prices for each boiler by the average price of electricity sales for customers and for resale of each corresponding utility. For some utilities without electric price information, we use the state average retail electricity price reported in EIA (2011a).

From the original 491 bituminous coal power plant boilers data we are able to collect, we construct the 9 years balance panel data set based on the input output information described above. We drop 97 boilers for which their size are less than 100 MW, 55 boilers for which they are not pulverized coal-fired boilers and 3 boilers for which there are missing data on electricity and pollutants, leaving 336 boilers units in the sample. There was no entry or exit of coal power plants observed in the data gathered over this time horizon. The summary statistics are presented in Table 1.

<table>
<thead>
<tr>
<th>Year</th>
<th>Variable</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>Electricity</td>
<td>2141</td>
<td>1640</td>
<td>257</td>
<td>8315</td>
</tr>
<tr>
<td></td>
<td>SO2</td>
<td>12.57</td>
<td>11.45</td>
<td>0.30</td>
<td>76.28</td>
</tr>
<tr>
<td></td>
<td>NOx</td>
<td>4.69</td>
<td>3.55</td>
<td>0.78</td>
<td>18.68</td>
</tr>
<tr>
<td></td>
<td>Heat input</td>
<td>20,855</td>
<td>15,672</td>
<td>3201</td>
<td>79,135</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>52.26</td>
<td>14.51</td>
<td>17.26</td>
<td>113.80</td>
</tr>
<tr>
<td>2001</td>
<td>Electricity</td>
<td>2029</td>
<td>1618</td>
<td>236</td>
<td>10,378</td>
</tr>
<tr>
<td></td>
<td>SO2</td>
<td>11.64</td>
<td>10.79</td>
<td>0.25</td>
<td>63.57</td>
</tr>
<tr>
<td></td>
<td>NOx</td>
<td>4.31</td>
<td>3.34</td>
<td>0.52</td>
<td>20.89</td>
</tr>
<tr>
<td></td>
<td>Heat input</td>
<td>19,741</td>
<td>15,306</td>
<td>2283</td>
<td>86,749</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>54.13</td>
<td>17.45</td>
<td>20.71</td>
<td>115.50</td>
</tr>
<tr>
<td>2002</td>
<td>Electricity</td>
<td>2062</td>
<td>1689</td>
<td>261</td>
<td>10,474</td>
</tr>
<tr>
<td></td>
<td>SO2</td>
<td>11.33</td>
<td>10.73</td>
<td>0.24</td>
<td>87.59</td>
</tr>
<tr>
<td></td>
<td>NOx</td>
<td>4.22</td>
<td>3.42</td>
<td>0.36</td>
<td>20.97</td>
</tr>
<tr>
<td></td>
<td>Heat input</td>
<td>19,945</td>
<td>15,762</td>
<td>2737</td>
<td>88,046</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>50.44</td>
<td>17.06</td>
<td>21.44</td>
<td>111.60</td>
</tr>
<tr>
<td>2003</td>
<td>Electricity</td>
<td>2127</td>
<td>1721</td>
<td>242</td>
<td>10,210</td>
</tr>
<tr>
<td></td>
<td>SO2</td>
<td>12.02</td>
<td>12.07</td>
<td>0.26</td>
<td>83.56</td>
</tr>
<tr>
<td></td>
<td>NOx</td>
<td>3.99</td>
<td>3.19</td>
<td>0.69</td>
<td>20.17</td>
</tr>
<tr>
<td></td>
<td>Heat input</td>
<td>20,412</td>
<td>15,854</td>
<td>2291</td>
<td>92,378</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>52.83</td>
<td>16.60</td>
<td>21.54</td>
<td>124.40</td>
</tr>
<tr>
<td>2004</td>
<td>Electricity</td>
<td>2097</td>
<td>1700</td>
<td>251</td>
<td>9940</td>
</tr>
<tr>
<td></td>
<td>SO2</td>
<td>11.67</td>
<td>11.38</td>
<td>0.22</td>
<td>75.75</td>
</tr>
<tr>
<td></td>
<td>NOx</td>
<td>3.57</td>
<td>2.80</td>
<td>0.40</td>
<td>15.17</td>
</tr>
<tr>
<td></td>
<td>Heat input</td>
<td>20,027</td>
<td>15,674</td>
<td>2526</td>
<td>83,167</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>55.13</td>
<td>17.46</td>
<td>22.42</td>
<td>125.50</td>
</tr>
<tr>
<td>2005</td>
<td>Electricity</td>
<td>2166</td>
<td>1785</td>
<td>266</td>
<td>11,155</td>
</tr>
<tr>
<td></td>
<td>SO2</td>
<td>11.81</td>
<td>11.85</td>
<td>0.19</td>
<td>80.98</td>
</tr>
<tr>
<td></td>
<td>NOx</td>
<td>3.48</td>
<td>2.73</td>
<td>0.42</td>
<td>15.15</td>
</tr>
<tr>
<td></td>
<td>Heat input</td>
<td>20,605</td>
<td>16,299</td>
<td>2803</td>
<td>92,853</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>59.65</td>
<td>18.67</td>
<td>24.71</td>
<td>139.50</td>
</tr>
<tr>
<td>2006</td>
<td>Electricity</td>
<td>2180</td>
<td>1749</td>
<td>200</td>
<td>10,363</td>
</tr>
<tr>
<td></td>
<td>SO2</td>
<td>11.25</td>
<td>11.40</td>
<td>0.20</td>
<td>71.92</td>
</tr>
<tr>
<td></td>
<td>NOx</td>
<td>3.42</td>
<td>2.74</td>
<td>0.38</td>
<td>16.59</td>
</tr>
<tr>
<td></td>
<td>Heat input</td>
<td>20,401</td>
<td>15,939</td>
<td>2191</td>
<td>83,026</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>65.43</td>
<td>19.36</td>
<td>28.58</td>
<td>154.50</td>
</tr>
<tr>
<td>2007</td>
<td>Electricity</td>
<td>2187</td>
<td>1765</td>
<td>83</td>
<td>10,094</td>
</tr>
<tr>
<td></td>
<td>SO2</td>
<td>10.34</td>
<td>11.40</td>
<td>0.13</td>
<td>92.63</td>
</tr>
<tr>
<td></td>
<td>NOx</td>
<td>3.30</td>
<td>2.75</td>
<td>0.31</td>
<td>14.78</td>
</tr>
<tr>
<td></td>
<td>Heat input</td>
<td>20,789</td>
<td>16,388</td>
<td>886</td>
<td>95,973</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>67.53</td>
<td>20.85</td>
<td>23.57</td>
<td>152.20</td>
</tr>
<tr>
<td>2008</td>
<td>Electricity</td>
<td>2180</td>
<td>1763</td>
<td>83</td>
<td>10,094</td>
</tr>
<tr>
<td></td>
<td>SO2</td>
<td>10.31</td>
<td>11.41</td>
<td>0.13</td>
<td>92.63</td>
</tr>
<tr>
<td></td>
<td>NOx</td>
<td>3.30</td>
<td>2.75</td>
<td>0.31</td>
<td>14.78</td>
</tr>
<tr>
<td></td>
<td>Heat input</td>
<td>20,740</td>
<td>16,388</td>
<td>886</td>
<td>95,973</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>73.98</td>
<td>22.04</td>
<td>36.06</td>
<td>165.70</td>
</tr>
<tr>
<td></td>
<td>Boiler size</td>
<td>336</td>
<td>240</td>
<td>100</td>
<td>1426</td>
</tr>
</tbody>
</table>

2 The CNLS production frontier is piecewise linear. Infrequently observations lie on a kink of the CNLS production frontier and do not have unique shadow prices. Using the method proposed by Kuosmanen and Kortelainen (2011), we use the minimum marginal product of pollutants to estimate shadow prices.
4. Empirical results

To test whether the assumption of the frontier production function is more appropriate than the neoclassical production function, we apply the skewness and kurtosis tests proposed in Kuosmanen and Fosgerau (2009). The null hypothesis $H_0$: disturbances that are normally distributed are tested against an alternative hypothesis $H_1$: disturbances are negatively skewed. Table 2 reports the $\sqrt{b_1}$ and $b_2$ test statistics and the relevant p-values of the normality tests. As expected, the $\sqrt{b_1}$ statistics are negatively signed. At the 5% significance level, normality is rejected in favor of skewness in 2001–2004 and 2006–2008, which supports the frontier model, and cannot be rejected in 2000 and 2005, which supports the neoclassical assumption. Thus, in these two years we use the neoclassical production model in which the disturbances contain only noise.

Table 3 reports the estimated shadow prices of SO$_2$ and NO$_x$ using StoNED and related statistics, assuming a deterministic disturbance. The estimated average prices of SO$_2$ over the 9-year time horizon, range between 201 and 343$/ton and the estimated average prices of NO$_x$ between 409 and 1352$/ton. The estimated average technical inefficiencies range between 0.883 and 0.902.

Table 4 shows the estimated shadow prices of SO$_2$ and NO$_x$ using StoNED and related statistics, assuming a deterministic disturbance term case. The values in parentheses represent the statistics excluding zero shadow price.

4. Table 2 Results of the skewness and kurtosis tests.

<table>
<thead>
<tr>
<th>Year</th>
<th>$\frac{\sqrt{b_1}}{b_1}$</th>
<th>$b_2$</th>
<th>P-values</th>
<th>$\frac{\sqrt{b_1}}{b_1}$</th>
<th>$b_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>-0.084</td>
<td>3.987</td>
<td>0.263</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>-0.271</td>
<td>4.022</td>
<td>0.020</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>-0.382</td>
<td>4.069</td>
<td>0.002</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>-0.546</td>
<td>4.004</td>
<td>0.000</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>-0.496</td>
<td>4.738</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>-0.168</td>
<td>3.334</td>
<td>0.102</td>
<td>0.095</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>-0.618</td>
<td>4.785</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>-0.560</td>
<td>5.334</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>-0.488</td>
<td>5.318</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

5. Table 3 Statistics of the estimated shadow prices of SO$_2$ and NO$_x$ ($/ton) and technical inefficiency for the deterministic disturbance term case.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean$^a$</th>
<th>Std. dev.$^a$</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>Price$_{SO2}$</td>
<td>743</td>
<td>11,140</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Price$_{NOx}$</td>
<td>7536</td>
<td>9053</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.884</td>
<td>0.077</td>
<td>0.674</td>
</tr>
<tr>
<td>2001</td>
<td>Price$_{SO2}$</td>
<td>1291</td>
<td>16,816</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Price$_{NOx}$</td>
<td>9234</td>
<td>11,713</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.879</td>
<td>0.077</td>
<td>0.665</td>
</tr>
<tr>
<td>2002</td>
<td>Price$_{SO2}$</td>
<td>507</td>
<td>1041</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Price$_{NOx}$</td>
<td>4597</td>
<td>6425</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.888</td>
<td>0.072</td>
<td>0.685</td>
</tr>
<tr>
<td>2003</td>
<td>Price$_{SO2}$</td>
<td>691</td>
<td>2712</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Price$_{NOx}$</td>
<td>6187</td>
<td>11,262</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.886</td>
<td>0.073</td>
<td>0.637</td>
</tr>
<tr>
<td>2004</td>
<td>Price$_{SO2}$</td>
<td>734</td>
<td>4296</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Price$_{NOx}$</td>
<td>7799</td>
<td>11,272</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.883</td>
<td>0.073</td>
<td>0.612</td>
</tr>
<tr>
<td>2005</td>
<td>Price$_{SO2}$</td>
<td>1033</td>
<td>11,274</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Price$_{NOx}$</td>
<td>11,679</td>
<td>13,692</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.892</td>
<td>0.066</td>
<td>0.697</td>
</tr>
<tr>
<td>2006</td>
<td>Price$_{SO2}$</td>
<td>954</td>
<td>9457</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Price$_{NOx}$</td>
<td>9994</td>
<td>13,356</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.902</td>
<td>0.063</td>
<td>0.700</td>
</tr>
<tr>
<td>2007</td>
<td>Price$_{SO2}$</td>
<td>780</td>
<td>8074</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Price$_{NOx}$</td>
<td>4044</td>
<td>7367</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.875</td>
<td>0.075</td>
<td>0.684</td>
</tr>
<tr>
<td>2008</td>
<td>Price$_{SO2}$</td>
<td>2020</td>
<td>20,215</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Price$_{NOx}$</td>
<td>3671</td>
<td>5583</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>0.869</td>
<td>0.077</td>
<td>0.618</td>
</tr>
</tbody>
</table>

$^a$ Weighted average by the amount of pollutants.

6. Table 4 Statistics of the estimated shadow prices of SO$_2$ and NO$_x$ ($/ton) and technical inefficiency for the composite disturbance term case.

<table>
<thead>
<tr>
<th>Year</th>
<th>Price$_{SO2}$</th>
<th>Price$_{NOx}$</th>
<th>TE</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>201</td>
<td>1540</td>
<td>N/A</td>
<td>0</td>
<td>2573</td>
</tr>
<tr>
<td></td>
<td>255</td>
<td>8035</td>
<td>N/A</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>2001</td>
<td>293</td>
<td>3218</td>
<td>0.043</td>
<td>0.036</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>388</td>
<td>11,955</td>
<td>0.818</td>
<td>0.818</td>
<td>1</td>
</tr>
<tr>
<td>2002</td>
<td>318</td>
<td>4899</td>
<td>0.038</td>
<td>0.042</td>
<td>0.801</td>
</tr>
<tr>
<td></td>
<td>415</td>
<td>16,542</td>
<td>0.743</td>
<td>0.743</td>
<td>1</td>
</tr>
<tr>
<td>2003</td>
<td>230</td>
<td>3338</td>
<td>0.927</td>
<td>0.049</td>
<td>0.743</td>
</tr>
<tr>
<td></td>
<td>418</td>
<td>17,648</td>
<td>0.683</td>
<td>0.683</td>
<td>1</td>
</tr>
<tr>
<td>2004</td>
<td>219</td>
<td>3215</td>
<td>0.934</td>
<td>0.044</td>
<td>0.740</td>
</tr>
<tr>
<td></td>
<td>410</td>
<td>30,712</td>
<td>0.683</td>
<td>0.683</td>
<td>1</td>
</tr>
<tr>
<td>2005</td>
<td>246</td>
<td>37,648</td>
<td>0.935</td>
<td>0.043</td>
<td>0.757</td>
</tr>
<tr>
<td></td>
<td>1467</td>
<td>19,071</td>
<td>0.683</td>
<td>0.683</td>
<td>1</td>
</tr>
<tr>
<td>2006</td>
<td>343</td>
<td>108,436</td>
<td>0.937</td>
<td>0.042</td>
<td>0.777</td>
</tr>
<tr>
<td></td>
<td>3097</td>
<td>32,640</td>
<td>0.683</td>
<td>0.683</td>
<td>1</td>
</tr>
<tr>
<td>2007</td>
<td>409</td>
<td>4907</td>
<td>0.937</td>
<td>0.042</td>
<td>0.777</td>
</tr>
<tr>
<td></td>
<td>1456</td>
<td>16,775</td>
<td>0.683</td>
<td>0.683</td>
<td>1</td>
</tr>
<tr>
<td>2008</td>
<td>239</td>
<td>11,572</td>
<td>0.931</td>
<td>0.046</td>
<td>0.745</td>
</tr>
<tr>
<td></td>
<td>666</td>
<td>28,770</td>
<td>0.683</td>
<td>0.683</td>
<td>1</td>
</tr>
</tbody>
</table>

$^a$ Weighted average by the amount of pollutants.

$H_1$: disturbances are negative skewed. Table 2 reports the $\sqrt{b_1}$ statistics and the kurtosis test statistics for the composite disturbance term. The simulated distribution of the skewness test statistic, $\sqrt{b_1}$, and the kurtosis test statistic, $b_2$, are constructed by a simple Monte Carlo simulation using $M=10,000$ pseudo-samples of $n=336$ observations from $N(0,1)$. 
The estimated average technical inefficiencies range between 0.927 and 0.943. For all data sets, the estimated second and third moments of the residual, $M_2$ and $M_3$, have the correct signs; thus, the expected inefficiency terms can be calculated and used to estimate the shadow prices of both pollutants. The convexity condition, $\sum y_i \leq 1$, in proposition 3 is satisfied for each year of data, which indicates that the objective function is globally convex. Therefore, a global optimal solution can always be found using standard nonlinear programming methods.

We find that applying the deterministic methods results in higher shadow price estimates than when assuming a composite disturbance term. Moreover, the estimated shadow prices in the deterministic case have a wider range. Weak disposability production frontiers constructed in the deterministic case are more sensitive to variation. If outliers are present in the data set, the estimated frontier tends to have larger steep regions, thus $\frac{\partial f(x_i, b_i)}{\partial b_i}$ is large and the estimated shadow prices are higher. In general, when only a few extreme observations are used to construct a frontier, the result is more variation in the estimated shadow price.

Fig. 1 shows the average shadow price estimates compared with previous studies; note that every study uses different data sets and estimation methods as summarized in Table 5. From Fig. 1, three conclusions can be drawn. First, average shadow price estimates for SO$_2$ from previous studies ranges between 76 to 3107 $/\text{ton}$, is consistent with the previous literature and EPA auction prices. For the composite disturbance model, our average shadow price estimates for SO$_2$, ranging from 201 to 343 $/\text{ton}$, contain the estimates of Coggins and Swinton (1996) and are close to Färe et al. (2005). More importantly, our estimates from the composite disturbance model are in the range of EPA’s SO$_2$ allowance auction prices.

The results also confirm that the shadow price estimates from composite disturbance models are generally lower than those from deterministic models, and are likely better estimates of the prices from the EPA’s allowance markets. Excluding Coggins and Swinton (1996), the average shadow price for SO$_2$ from deterministic models (including DEA) are 509–3107 $/\text{ton}$ compared to 76–343 $/\text{ton}$ for the composite models. Table 6 shows that the SO$_2$ market prices are 130–1550 $/\text{ton}$ and the allowance auction price is 126–860 $/\text{ton}$.

We conclude that the weak disposability StoNED method provides more consistent estimates of market prices compared to weak disposability DEA. Third, the shadow prices of NO$_x$ are higher than the shadow prices of SO$_2$. Using a composite disturbance term, the NO$_x$ average shadow prices of 409–1352 $/\text{ton}$ are higher than the SO$_2$ shadow prices of 201–343 $/\text{ton}$. This conclusion coincides with the observed prices in the SO$_2$ and NO$_x$ allowance markets.

Fig. 1. Comparison of the estimated average shadow prices of SO$_2$ and NO$_x$.

Table 5
Data set comparisons of the electricity price used to estimate shadow prices of the pollutants.

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Year</th>
<th>Sample size</th>
<th>Price of electricity ($/\text{MWh}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boyd et al. (1996)</td>
<td>U.S.</td>
<td>1989</td>
<td>62</td>
<td>50.00</td>
</tr>
<tr>
<td>Present study</td>
<td>U.S.</td>
<td>2000–2008</td>
<td>336</td>
<td>17.26–165.70</td>
</tr>
</tbody>
</table>
spikes during 2004–2005 and declines after 2005. Our average NOx prices are within the range or close to NOx market prices except in 2003, 2004 and 2005. During this time, our estimates are lower than the market price because the NOx market price increases sharply. However, our NOx price estimates have a similar trend of rising prices in 2004, 2005 and 2006 and dropping prices in 2007 and 2008. Figs. 2 and 3 illustrate the average shadow price estimates and the market prices.

Recall that the shadow prices of pollutants are estimates of the marginal abatement costs which should reflect the market prices for EPA’s pollutant allowances. Our shadow price estimates are derived based solely on the plants’ production data; however, several other factors can affect the market allowance price. By allowing plants to buy, sell and bank allowances, the allowance prices reflect the cost of compliance with future regulation. The sharp increase in SO2 and NOx prices resulting from EPA’s Clean Air Interstate Rule (CAIR) which requires further SO2 and NOx reduction from coal boilers beginning in 2010, caused an increase in the expected pollutant control costs in the future and provided incentives for plants to buy allowances and bank them for future use. Thus, allowance prices rose due to increased demand for allowances. After 2005, emission levels fell due to the increased use of gas-fired boilers and pollution control equipment. Thus, a sufficient supply of allowances in the market caused allowance market prices to fall.

5. Conclusions

This paper proposed a nonparametric methodology to estimate a production frontier when pollutants are a result of the production process. We assumed that the traditional production axioms such as continuity, monotonicity and concavity with weak disposability between an output and the pollutants characterized the shape of an underlying production frontier. To estimate the production frontier empirically, we extended the two-stage CNLS method to incorporate the weak disposability assumption. In deterministic cases assuming no noise in the data and an exact model specification, we modified CNLS to minimize the sum of firms’ one-sided deviations. In composite disturbance cases where noise was explicitly modeled, we extended the StoNED method to include the weak disposability axiom. The composite CNLS residuals were decomposed into noise and technical inefficiency terms and the estimated expected inefficiency was used to multiplicatively shift an average CNLS production function to obtain the weak disposability production frontier. The proposed methodology was applied to derive the technical efficiencies of 336 boilers for the U.S. coal power plants and the shadow prices of SO2 and NOx.

The main finding of this study is that, applying the StoNED method with a composite disturbance term, average shadow prices estimates of SO2 are between 201 and 343$/ton and average shadow prices of NOx are between 409 and 1352$/ton. Both SO2 and NOx shadow price estimates are in reasonable ranges comparing to allowance market prices. The proposed method can be applied to estimate shadow prices of other pollutants which can be used as references for marginal abatement costs for the industry. This marginal abatement cost is solely derived from production data so that it is not affected by market complexity.

From the results in this study, we recommend the use of weak disposability StoNED method over weak disposability DEA which is likely to overestimate shadow prices due to extreme observations. Further cost analysis tools, such as the ones proposed in this paper, will allow the EPA to investigate the outcomes of their on-going pollution control policies.

Acknowledgement

This paper has benefited from comments and suggestions on earlier drafts from an anonymous referee and the editorial services of Ann Stewart, bondcliffs@gmail.com.

---

Table 6

<table>
<thead>
<tr>
<th>Year</th>
<th>SO2 prices</th>
<th>NOx prices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Market</td>
<td>Present study</td>
</tr>
<tr>
<td>2000</td>
<td>130–155</td>
<td>201</td>
</tr>
<tr>
<td>2001</td>
<td>135–210</td>
<td>293</td>
</tr>
<tr>
<td>2002</td>
<td>130–170</td>
<td>318</td>
</tr>
<tr>
<td>2003</td>
<td>150–220</td>
<td>230</td>
</tr>
<tr>
<td>2004</td>
<td>215–700</td>
<td>219</td>
</tr>
<tr>
<td>2005</td>
<td>700–1550</td>
<td>246</td>
</tr>
<tr>
<td>2006</td>
<td>430–740</td>
<td>343</td>
</tr>
<tr>
<td>2007</td>
<td>500–600</td>
<td>237</td>
</tr>
<tr>
<td>2008</td>
<td>179–509</td>
<td>239</td>
</tr>
</tbody>
</table>

*a* Information on SO2 allowance market prices is reported in EPA Acid Rain Program 2000–2008 Progress Reports, information on NOx allowance market prices is reported in EPA OTC NOx Budget Trading Program 1999–2002 Progress Reports and EPA NOx Budget Trading Program 2003–2008 Progress Reports. All reports are available at: http://www.epa.gov/airmarkt/progress/progress-reports.html.

*b* For 2003, 2004 and 2005, the range of NOx market prices are approximated from graphs; for the other years, the range of NOx market prices are explicitly stated in the EPA reports.

*c* For 2000, 2001 and 2002, EPA published three years of progress in a single OTC NOx budget program report.
Appendix A

Proof of Proposition 1. In a single output case, we can transform the problem (2) into an additive form:

\[
\begin{align*}
\max_{\theta_b, \lambda, \mu} & \quad \sum_{i=1}^{n} \lambda_i y_i \geq y_0 + \Theta_0 \\
& \quad \sum_{i=1}^{n} \lambda_i b_{ij} \leq b_{ij} \quad \forall j = 1, ..., J \\
& \quad \sum_{i=1}^{n} (\lambda_i + \mu_j) x_{im} \leq x_{sm} \quad \forall m = 1, ..., M \\
& \quad \sum_{i=1}^{n} (\lambda_i + \mu_j) = 1 \\
& \quad \lambda_i, \mu_j \geq 0 \quad \forall i = 1, ..., n
\end{align*}
\]

(19)

where \( \theta_b = 1 + \frac{\Theta_0}{\sum_{i=1}^{n} \lambda_i} \).

Applying duality theory of linear programming, the LP problem (19) has a dual problem:

\[
\begin{align*}
\min_{\alpha, w, \beta, \gamma} & \quad (\alpha + w x_0 + c \beta) - y_0 \\
& \quad \alpha + w x_i \geq 0 \quad \forall i = 1, ..., n \\
& \quad w, c \geq 0 \\
& \quad \alpha + w x_i \geq 0 \quad \forall i = 1, ..., n
\end{align*}
\]

(20)

We remove \( y_0 \) from the objective function since it is a constant. Then we take the logarithm of the objective function and the first set of constraints, because the logarithm is a monotonic transformation for values greater than or equal to 1. Next we add the negative of \( \ln y_0 \) to the objective function since it is a constant. Problem (20) becomes:

\[
\begin{align*}
\min_{\alpha, w, \beta, \gamma} & \quad \ln(\alpha + w x_0 + c \beta) - \ln y_0 \\
& \quad \ln(\alpha + w x_i) \geq 0 \quad \forall i = 1, ..., n \\
& \quad w, c \geq 0 \\
& \quad \alpha + w x_i \geq 0 \quad \forall i = 1, ..., n
\end{align*}
\]

(21)

Introducing a new variable \( e_\alpha = \ln y_0 - \ln(\alpha + w x_0 + c \beta) \) and adding an additional constraint, problem (21) can be equivalently written as:

\[
\begin{align*}
\min_{\alpha, w, \beta, \gamma} & \quad -e_\alpha \quad \ln(\alpha + w x_i + c \beta) \geq \ln y_0 \quad \forall i = 1, ..., n \\
& \quad \alpha + w x_i \geq 0 \quad \forall i = 1, ..., n \\
& \quad w, c \geq 0 \\
& \quad \alpha + w x_i \geq 0 \quad \forall i = 1, ..., n
\end{align*}
\]

(22)

Instead of solving (22) separately for each firm, we combine \( n \) optimization formulations and solve simultaneously for all firms. Since \( e_\alpha, \alpha, w, \) and \( c \) are estimated independently for each firm, we minimize the sum of \( e_\alpha \) as:

\[
\begin{align*}
\min_{\alpha, w, \beta, \gamma} & \quad -\sum_{i=1}^{n} e_i \quad \ln(\alpha + w x_i + c \beta) \geq \ln y_0 \quad \forall i = 1, ..., n \\
& \quad \alpha + w x_i \geq 0 \quad \forall i = 1, ..., n \\
& \quad w, c \geq 0 \\
& \quad \alpha + w x_i \geq 0 \quad \forall i = 1, ..., n
\end{align*}
\]

(23)

By construction, \( e_i \leq 0 \); we add this constraint to the problem. We add the inefficiency term \( e_i \) to the right side of the second set of constraints because of the monotonicity assumption. Note that the constraints are binding if \( i = h \), and inequality otherwise:

\[
\begin{align*}
\min_{\alpha, w, \beta, \gamma} & \quad -\sum_{i=1}^{n} e_i \quad \ln(\alpha + w x_i + c \beta) \geq \ln y_0 \quad \forall i = 1, ..., n \\
& \quad \ln(\alpha + w x_i + c \beta) \geq \ln y_0 \quad \forall i = 1, ..., n \\
& \quad \alpha + w x_i \geq 0 \quad \forall i = 1, ..., n \\
& \quad w, c \geq 0 \\
& \quad \alpha + w x_i \geq 0 \quad \forall i = 1, ..., n
\end{align*}
\]

(24)

Since \( \ln y_1 - e_i = \ln(\alpha + w_i x_i + c_i b_i) \), the second set of constraints can be written as \( \ln(\alpha + w_i x_i + c_i b_i) \geq \ln(\alpha + w_i x_i + c_i b_i) \forall i \), \( h = 1, ..., n \). Removing the logarithm from this second set of constraints allows the problem (24) to be equivalently written as:

\[
\begin{align*}
\min_{\alpha, w, \beta, \gamma} & \quad -\sum_{i=1}^{n} e_i \quad \alpha + w_i x_i + c_i b_i \geq \alpha + w_i x_i + c_i b_i \quad \forall i = 1, ..., n \\
& \quad \alpha + w_i x_i + c_i b_i \geq 0 \quad \forall i = 1, ..., n \\
& \quad w, c \geq 0 \\
& \quad \alpha + w_i x_i + c_i b_i \geq 0 \quad \forall i = 1, ..., n
\end{align*}
\]

(25)

which is the formula (6).

Proof of proposition 2. By construction, \( \theta_b y_i = y_i + \Theta_b \), thus \( \theta_b = 1 + \Theta_b / y_i \). By duality between the problem (19) and (20), \( \Theta_b := (\alpha + w_i x_i + c_i b_i) / y_i \). This gives \( \theta_b = (\alpha + w_i x_i + c_i b_i) / y_i \). By construction the variable \( \psi_i = \ln y_i - \ln(\alpha + w_i x_i + c_i b_i) \), thus \( \psi_i = \psi(y_i / (\alpha + w_i x_i + c_i b_i)) = 1 / \theta_b \).

Proof of proposition 3. Let the function \( \Omega(\phi_1, ..., \phi_n) = \sum_{i=1}^{n} (\ln y_i - \ln y_i)^2 \). Since \( \frac{\partial^2 \Omega}{\partial \hat{d}_i^2} = 2 (1 - \ln \theta_b + \ln y_i) / \theta_b \) and \( \frac{\partial^2 \Omega}{\partial \hat{d}_j^2} = 0 \forall i, j = 1, ..., n, \) all non-diagonal elements in the Hessian matrix of the function \( \Omega \) are equal to zero. Thus, the function \( \Omega \) is convex if and only if \( \frac{\partial^2 \Omega}{\partial \hat{d}_i^2} = 2 (1 - \ln \theta_b + \ln y_i) / \theta_b \) is nonnegative. This condition is equivalent to \( \theta_b \leq \exp y_i / \psi_i \). Since the objective function of the CNSL problem (10) is a composition with an affine function \( \phi_i = \alpha_i + w_i x_i + c_i b_i \forall i \), it is convex if the function \( \Omega(\phi) \) is convex if and only if \( \alpha_i + w_i x_i + c_i b_i \leq \exp y_i / \psi_i \).

References
