Discussion Papers in Economics

Measuring output-based technical efficiency of Indian coal-based thermal power plants: A by-production approach.

Sushama Murty & Resham Nagpal

Discussion Paper 18-07

Centre for International Trade and Development

School of International Studies

Jawaharlal Nehru University

India

*

Measuring output-based technical efficiency of Indian coal-based thermal power plants: A by-production approach.

June 25, 2018

Abstract

The by-production approach is employed in conjunction with data from the Central Electricity Authority (CEA) of India to compute the output-based Färe, Grosskopf, Lovell (FGL) efficiency index and its decomposition into productive and environmental efficiency indexes for the Indian coal-based thermal power plants (ITPPs). We show that given (i) the aggregated nature of data on coal reported by CEA (ii) CEA's computation of CO_2 emissions through a deterministic linear formula that does not distinguish between different coal-types and (iii) the tiny share of oil in coal-based power plants, the computed output-based environmental efficiency indexes are no longer informative. Meaningful measurement of environmental efficiency using CEA data is possible only along the dimension of the coal input. Productive efficiency is positively associated with the engineering concept of thermodynamic/energy efficiency and is also high for power plants with high operating availabilities reflecting better management and O&M practices. Both these factors are high for private and centrally-owned as opposed to state-owned power-generating companies. The example of Sipat demonstrates the importance of (ultra)supercritical technologies in increasing productive and thermodynamic efficiencies of the ITPPs, while also reducing CO_2 emitted per-unit of the net electricity generated.

JEL classification codes: Q50, Q40, D24,

Keywords: emission-generating technologies; by-production technologies; output-based measurement of productive, environmental, and technical efficiency indexes; coal-based Indian thermal power plants.

1 Introduction.

The National Climate Change Action Plan (NCCAP) was launched in India amidst increasing national and global concerns for the environment. As part of India's Intended Nationally Determined Contributions under UN Framework Convention on Climate Change, India intends to reduce the emissions intensity of its GDP by 33-35 percent by 2030 over 2005 levels. While the country strives to increase the share of renewables in its energy-mix to achieve this target, the energy sector's dependence on coal is expected to remain high. More than a third of carbon emissions in India are directly linked to coal. In India, roughly 67 percent of electricity is generated by coal-fired thermal power plants, making this sector one of the largest emitters in the country. Against this backdrop, National Mission for Enhanced Energy Efficiency (NMEEE) was launched as one of the eight missions of the NCCAP to promote energy efficient economic development.

Economists have long been concerned with measuring technical efficiency of producing units. Technical (in)efficiency is a broader concept than energy (in)efficiency as it measures the extent to which firms are (not) realising the full potential of some or all of their inputs (including energy inputs) in maximising the production of their intended/economic/good/marketed outputs and minimising production of the bad outputs. If the technical inefficiency of the Indian thermal energy sector is high, then it can be hypothesised that technical efficiency improvements in this sector can play a big role in controlling CO_2 emission generation without compromising on power generation in India.

The first step in developing a methodology for measuring technical efficiency that discounts the performance of producing units for excessive generation of bad outputs is the specification of a technology relative to whose efficient frontier technical efficiency will be measured. Here, though the literature was in agreement that a model of emission-generating technology should capture the empirically observed positive relation between emission generation and intended production, there was no consensus about how the standard neoclassical model of a production technology should be extended to include the generation of bad outputs as byproducts of intended production. Two rival approaches, namely, the input approach and the weak-disposability based output approach to modelling emission-generating technologies were popularly employed in many applied works.'[1] However, the appropriateness of the technological specifications of these two approaches were questioned in a series of articles beginning with Førsund

[1]The input approach was advocated in the classic works of Baumol and Oates (1988) and Cropper and Oats (1992) and was employed in several works such as Reinhard et al. (1999, 2000), Lee et al (2002), and Hailu and Veeman (2001). The weak-disposability based output approach can be traced back to Färe et al (1986). Since then, it gained immense popularity. It has been employed in many works such as Coggins and Swinton (1996), Murty and Kumar (2002, 2003) and Sahoo et al. (2017). See Zhou and Poh (2008) for a comprehensive survey of a number of papers employing this approach.

(1998); Murty and Russell (2002); Førsund (2009); and Murty, Russell and Levkoff (2012) (henceforth, MRL).[2] Some anomalies and counterintuitive implications of the assumptions underlying these models were brought to light and alternative specifications of emission-generating technologies were proposed. We focus on the by-production approach, which decomposes the overall emission-generating technology into (i) a standard neo-classical sub-technology, denoted by T_1 and (ii) a sub-technology, denoted by T_2, that relates usage of emission-causing inputs to emission generation based on considerations such as the thermodynamic laws.

In the context of all the above technological specifications, several output-based technical efficiency indexes such as the hyperbolic and directional distance function-based indexes have been employed in many applied works.[3] These are defined and computed with the help of mathematical programmes that measure the maximum feasible extents to which good outputs can be increased and bad outputs can be decreased when all inputs are held fixed. Most earlier works on Indian coal-based thermal power plants such as Khanna et al. (1999) measure technical efficiency relative to conventional neo-classical specifications of technologies that exclude emission generation. Later works such as Murty et al. (2007) and Sahoo et al. (2017), which extend the traditional technical efficiency measures to discount the production of CO_2 as the bad output produced by these plants, tend to adopt the input or the weak-disposability based methodologies.

The objective of this work is to employ, for the first time, the by-production approach to compute the output-based FGL efficiency index that is defined in MRL for the Indian coal-based thermal power sector. This index decomposes the technical efficiency index into (i) a productive efficiency component, which is computed relative to the sub-technology T_1 and (ii) an environmental efficiency component, which is computed relative to the sub-technology T_2. During the computation of this index using data reported by the Central Electricity Authority of India (CEA), we reached two conclusions of conceptual nature.

Firstly, while the CEA provides an aggregate measure of the physical quantity of coal employed by thermal power plants, there are significant differences in coal types employed by the thermal power plants with respect to gross calorific values, emission factors, *etc.* We argue that employing aggregated (as opposed to disaggregated) data on the coal input in applied work is consistent with the properties of inputs and outputs in the theoretical models of emission-generating technologies only if the aggregate coal input is measured in heat rather than in mass units. This is in contrast to some works in the literature such as Sueyoshi and Goto (2010, 2011, 2012) where the aggregate coal input is measured in mass units. The intuition has to do with the fact that, for a given mass of the aggregate coal input, the mix of coal types that

[2]See also Murty (2015) and Murty and Russell (2016, 2017); and Førsund (2017) for later works.

[3]See the references in footnote 1.

maximises the intended output, namely, electricity, could be quite different from the mix that minimises CO_2 emission level. In contrast, the maximum electricity that a given amount of heat input can produce is independent of the mix of coal types that produce that level of heat, including the mix that minimises emission generation.[4]

Secondly, data on CO_2 emissions reported by the CEA is not obtained by direct observations. Rather, it is computed using a linear formula based on the heat inputs and emission factors of coal and oil.[5] We argue that, since the share of oil in the total (aggregate) heat input employed by coal-based power plants is insignificant and because CEA employs a single average emission factor across all coal types and across all plants for measuring emissions, the emission levels generated and the amounts of the aggregate heat input employed by thermal power plants are almost perfectly correlated. Thus, we find that the production points of all thermal power plants are located very close to the linear frontier of sub-technology T_2, which is obtained by minimising the level of emission for any given level of the fossil fuel (here coal) measured in aggregate heat units. Hence, our results show that the output-based environmental efficiency index of Indian thermal power plants computed using the CEA data and the by-production specification of the technology is very high and shows little variability. Thus, apparently, there is very little scope for achieving environmental efficiency improvements (in the form of reduction of CO_2 emissions), when the heat input levels of these plants are held fixed.[6]

The second finding should not be interpreted as a weakness of the by-production methodology. Rather, it is a result of the way data on CO_2 emissions is computed by the CEA. This finding does not, however, exclude the possibility of existence of environmental efficiency improvements in the direction of the aggregate heat input in the Indian coal-based thermal power sector.[7] This is because the FGL efficiency index in MRL that we have computed for the Indian thermal power sector in this paper is output-based and holds levels of all inputs, including the heat input, fixed. We note also in Section 2 that the weak-disposability and input based approaches may not be suitable to compute technical efficiency when data on emissions are generated by (linear) formulae involving emission factors and gross calorific values of fossil fuels.

Given that output-based environmental efficiency of Indian thermal power plants is not so

[4]For a detailed explanation, see Section 3.2.

[5]Note that oil is a secondary fuel in coal-based thermal power plants, whose use is mainly limited to meeting start-up fuel requirements and flame stabilisation.

[6]See Section for a detailed exposition.

[7]In an on-going work, we are computing technical (in)efficiency of the Indian thermal power sector in the dimensions of all inputs and outputs using the by-production approach. This is being done using the FGL index of graph efficiency developed in Murty and Russell (2017). Employing the weak-disposability based output approach, Sahoo et al. (2017) have also computed technical efficiency in the input dimension for the coal-based Indian thermal power sector.

informative due to the way data on emission is generated by the CEA, we turn our attention to the second component of MRL's output-based FGL technical efficiency, namely, the productive efficiency index. This is a conventional neo-classical efficiency index. Here we find that the proportion of high performing plants has increased over the years and that the distribution of mean productive efficiency of thermal power plants has a fat right tail. The factors that affect productive efficiency of power plants the most are the station heat rate (SHR), total capacity of the plant, and operational availability. The first, is the inverse of energy (thermodynamic) efficiency as defined by engineers. Thus, there seems to be a strong positive relation between energy efficiency as defined by engineers and productive efficiency as defined by economists. The second captures the extent of economies of scale tapped by power plants, while the last is indicative of efficient managerial practices that minimise temporary plant shutdowns. At the level of the companies that operate these power plants, privately owned Reliance followed by centrally owned NTPC are the top performers. Both are associated with low SHR, high operational availability, and high average unit capacities indicating better management practices and high thermodynamic efficiency in plants run by these companies. Most state-owned companies have low productive efficiency as they perform poorly on these counts. We also highlight the important need for adopting supercritical and ultra-supercritical technologies in the Indian thermal power sector to improve both productive and environmental efficiencies. Sipat, the only supercritical power plant in our data set, is among the top performers whose SHR is lowest not only compared to the average of all plants but also compared to the average of the most efficient plants. It also has the lowest CO_2 emission per unit of net electricity generated.

Section 2 of this paper provides an intuitive informal review of the literature on modelling emission-generating technologies. It compares and contrasts the input and weak-disposability based approaches to the by-production approach. Section 3 distinguishes between aggregated and the disaggregated measures of coal input and argues that measuring aggregate input of coal in heat units is consistent with the theoretical models of emission-generating technologies. Section 4 describes the Data Envelopment Analysis (DEA) methodology for computing the output-based FGL efficiency index of MRL, while Section 5 discusses the data employed in this study. Section 6 presents the key results and findings of this study, while Section 7 concludes.

2 Specifying the technology for conducting efficiency measurement.

The first step for measurement of efficiency is the specification of the underlying technology and its efficient frontier. For pollution-generating technologies, there are contrasting approaches for specifying the technology based on different assumptions. We can broadly classify these approaches into the input, output, and by-production-based approaches. We lay out below a modelling framework for reviewing these approaches.

2.1 A suitable modelling framework and disposability properties of conventional outputs and inputs.

We consider a simple model with one intended output whose quantity is denoted by $y \in \mathbf{R}_+$, one unintended bad output (by-product) whose quantity is denoted by $z \in \mathbf{R}_+$, and n inputs. A vector of inputs is denoted by $x \in \mathbf{R}_+^n$. A production vector is denoted by $\langle x, y, z \rangle \in \mathbf{R}_+^{n+2}$. The technology set is denoted by $T \subset \mathbf{R}_+^{n+2}$. A production vector $\langle x, y, z \rangle$ is feasible under technology T if it is an element of set T. Technical efficiency of producing units is measured relative to the efficient frontier of the technology. A production vector in T is a (weakly) efficient point of technology T if there exists no other production vector in T having smaller amounts of the inputs and emission and bigger amount of the intended output. The set of all efficient points of technology T is called its efficient frontier.

Conventionally, technology T is assumed to satisfy free disposability of all inputs and the intended output. Free input disposability implies that if an input vector can produce a given amount of the intended output, then so can an input vector with even larger amounts of inputs. Thus, it is possible that inputs are wastefully used under technology T – more than than the minimum amounts of inputs required may have been used to produce a given level of the intended output. Free output disposability implies that, ceteris-paribus, if a particular level of intended output is produced with a given input vector, then any smaller amount of the intended output can also be produced with the same input vector. This can also be interpreted as saying that any arbitrary reduction of the intended output does not require additional inputs, *i.e.,* its disposal is free. Under these assumptions, it can be shown that, along the efficient frontier of the technology, there is non-negative relation between any input and the intended output when all other inputs are held fixed (see, *e.g.,* Murty and Russell (2017)). Hence, it is these disposability assumptions that lead us to models of technology where the economic output is non-decreasing in inputs when an economic unit operates efficiently. Such models conform to our general intuition about the relationship between the economic output produced and the levels of inputs employed.

2.2 A brief review of alternative specifications of emission-generating technologies.

It is intuitive that emission, which is a by-product of intended production, is not a freely disposable output: its reduction is not costless in terms of resources. The input, output, and the by-production approaches to modelling emission-generating technologies make distinctly different assumptions regarding its disposability properties. But the common feature that all these approaches wish to capture is the positive association between emission generation and intended output production that is generally observed in real life – as intended production increases, more amount of the emission tends to be generated.

2.2.1 The input and weak-disposability based output approaches.

To capture this feature the input approach assumes that emission is a freely disposable input, while the output approach assumes that the technology satisfies "null-jointness" and "weak disposability of the emission and the intended output."[8] Null-jointness implies that when no emission is generated then there is also no production of the intended output (equivalently, a positive production of the intended output must necessarily imply generation of some positive amount of the emission). Weak disposability of the emission and the intended output implies that although the technology does not permit free output disposability of emission, it does permit proportionate (simultaneous) reductions in emission and the intended output when all inputs are held fixed. These alternative set of assumptions made by the input and output approaches ensure that, *when all inputs are held fixed*, then there is a non-negative relation between emission generation and production of the intended output along the efficient frontier of the technology. For instance, as discussed above in Section 2.1, free input and output disposability imply a non-negative relation between an input and the intended output when all other inputs are held fixed. Hence, treating emission as a conventional input will also ensure a non-negative relation between it and the intended output.

For any input vector $x \in \mathbf{R}^{n+2}_{+}$, the set of combinations of intended output and emission that are feasible under technology T with input vector x is denoted as $P(x)$. Lets call the efficient frontier of $P(x)$ as the production possibility frontier (PPF) of $P(x)$. It is the set of combination of intended output and emission levels in $P(x)$ for which there exist no other intended output and emission combinations in $P(x)$ with bigger amounts of the intended output and smaller amounts of the emission. Under the assumptions made by the input and the weak-disposability based output approaches to modelling emission-generating technologies, the slope of the PPF of $P(x)$ is non-negative, indicating a non-negative relation between emission and

[8] For references on these two approaches, see footnote 1 in the introductory section.

intended output levels along this frontier. This also implies that the PPF gives a rich menu of emission and intended output combinations that can be produced with input vector x. The left panel of Figure 1 is an example of the set $P(x)$ under the input approach, while the right panel illustrates a case under the weak-disposability-based output approach. The efficient frontier of $P(x)$ or the PPF in each case is the bold part of the boundary of $P(x)$. It has many emission-intended output combinations and is non-negatively sloped.

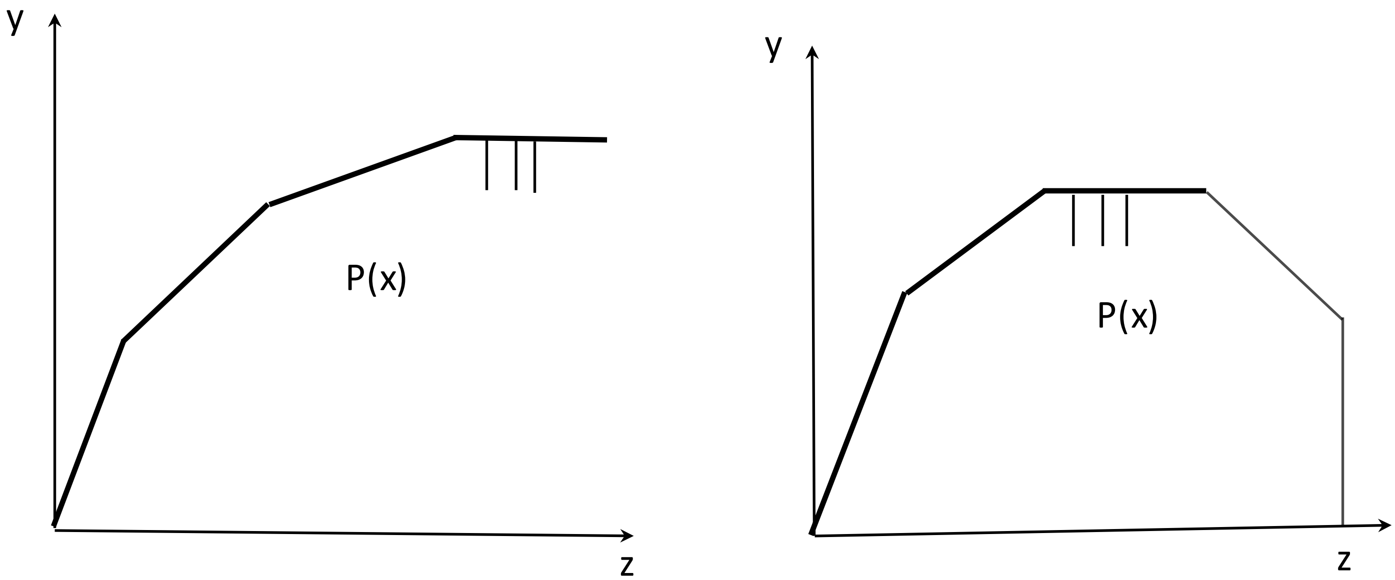

Figure 1

2.2.2 Some critiques of the input and output approaches.

Concerns with these two approaches have been discussed in detail in MRL, Murty (2015), and Murty and Russell (2016, 2017). Here we review two of these concerns. Firstly, it is well known in production theory that standard input disposability implies that iso-quants are non-positively sloped, *i.e.,* when the producing unit operates efficiently, then there is a non-positive relation between any two inputs when the intended output level and amounts of all other inputs are held fixed. Thus, the input approach to modelling emission-generating technologies, which treats emission as an input, implies a non-positive relation between emission and any other emission-causing input along the efficient frontier of the technology, which is counterintuitive.[9] For example, it would imply a negative relation between CO_2 emission and the input of coal.

[9]See Murty and Russell (2017) for a proof.

The fundamental laws of thermodynamics imply that there is a positive relation between emission and emission-causing inputs. Emission generation increases if and only if the use of emission-causing inputs increases. This is in contradiction with the rich menu of emission-intended output combinations on the PPF derived under the the input and weak-disposability based output approaches (see Figure 1). This is because the PPF in these two approaches is constructed holding all inputs, *including emission-causing inputs*, fixed. Thus, the second concern with the input and weak-disposability based output approached is that the rich variations in the levels of emission that are seen along the PPF of $P(x)$ derived under these approaches are not scientifically consistent with a fixed level of the emission-causing inputs.

2.2.3 The by-production approach.

Technology as an intersection of an intended production technology and an emission-generation set.

This paper adopts the by-production approach, which has been developed in MRL, Murty (2015), and Murty and Russell (2016, 2017), to model the emission-generating technology of a thermal power plant. We provide below a brief intuitive review of this approach to modelling emission-generating technologies.[10] This approach distinguishes between emission-causing and non-emission causing inputs and defines the emission-generating technology set T as an intersection of two sub-technologies $T_1 \subset \mathbf{R}_+^{n+2}$ and $T_2 \subset \mathbf{R}_+^{n+2}$:

$$T = T_1 \cap T_2. \tag{1}$$

Sub-technology T_1 is a standard neo-classical technology that defines the engineering rules governing the transformation of all inputs (including emission-causing inputs) into the intended output. In this work we assume that this transformation is not affected by (or is independent of) changes in the emission level. Sub-technology T_2, on the other hand, describes the transformation of emission-causing inputs into the emission in accordance with the laws of thermodynamics, *e.g.,* the rules governing the transformation of coal into CO_2. It is assumed in this work that emission-generation is unaffected by the usage of the non-emission causing inputs and the levels of the intended output produced, *i.e.,* changes in the levels of these goods do not affect the level of emission generated by sub-technology T_2.[11]

[10]For more detailed and formal exposition, the reader is referred to MRL, Murty (2015), and Murty and Russell (2016, 2017). See also Serra et al. (2016) and Ray et al. (2017) for further extensions and empirical applications of this approach.

[11]In the context of thermal power plants, these assumptions imply that the emission of the plant does not directly impact (*e.g.,* detrimentally) its generation of electricity, and that the level of CO_2 emission generated by it is not impacted by its usage of inputs such as labour and capital.

The upper frontier of sub-technology T_1 reflects the maximum amount of the intended output that can be produced by any given level of inputs. On the other hand, the economist and environmentalist will be concerned with the lower frontier of sub-technology T_2, *i.e.,* the set of points in T_2 where emission generation is minimum for any given vector of the emission-causing inputs. As in the case of a standard neo-classical technology, points in T_1 that lie below its upper frontier are technically inefficient in the production of the intended output. In contrast, points in T_2 that lie above its lower frontier are technically inefficient in emission generation: at these points, the given levels of the emission-causing inputs are generating more than the minimum feasible levels of the emission.

It is assumed that T_1 satisfies standard free input disposability with respect to all inputs and output free disposability with respect to the intended output. Hence, as explained above in Section 2.1, inputs and the intended output are non-negatively related along the upper frontier of T_1. This is exactly the case of a standard neo-classical technology where marginal products of inputs are non-negative.

Figure 2 assumes that there is only one input, which is emission-causing. An example of sub-technology T_1 (or rather its projection into the space of the input and intended output) is shown in the top panel. It includes its upper frontier and all points that lie below it. An example of sub-technology T_2 (or rather its projection into the space of of the input and emission) is shown in the bottom panel.[12] In this example, T_2 is a straight line through the origin, *i.e.,* all points in T_2 lie on this straight line. This will be true, for instance, when data on emission is generated by a linear formula, which relates the emission level linearly to fossil fuel usage through use of constant emission factors.[13] In real life data on CO_2 is generally generated in this manner.

The point $\langle x', y', z' \rangle$ is feasible under technology $T = T_1 \cap T_2$. It lies below the upper frontier of T_1, as y' is less than the maximum feasible output that can be produced under sub-technology T_1 with x' amount of the input. Points on or below the upper frontier of T_1 are feasible under sub-technology T_1, which satisfies standard disposability with respect to the input and the intended output. On the other hand, point $\langle x', y', z' \rangle$ lies on the lower frontier of sub-technology T_2, which in this example is sub-technology T_2 itself.

Disposability properties of the emission-generating set T_2 and the slope of its efficient frontier.

To restrict the scope of sub-technology T_2 to valid sets whose lower frontiers exhibit the realistic non-negative relation between emissions and and emission-causing inputs, it is assumed

[12]The emission-dimension is ignored in the top panel of Figure 2, as it is assumed not to affect the production of the intended output; while the intended output-dimension is ignored in the bottom panel of this figure as it is assumed not to affect the generation of the emission.

[13]The emission factor of a fuel is the amount of emission generated per unit combustion of the fuel.

that the lower frontier of set T_2 is the same as the lower frontier of its costly disposable hull $\bar{T}_2$, where the set $\bar{T}_2$ is defined as containing all production vectors in T_2 as well as production vectors with arbitrarily higher levels of emissions and lower levels of emission-causing inputs than those permitted by T_2. Figures 3 and 4, which assume a single input that is emission-causing, illustrate this point. In both figures, the upper panels depict sub-technology T_2, while the lower panels depict its costly disposal hull $\bar{T}_2$, whose lower frontier is depicted by the bold curves in black. Figure 3 is an example of a case that satisfies our assumption about set T_2, namely, that both sets T_2 and $\bar{T}_2$ share a common lower frontier, which is linear. Figure 4 is an example of a case where the two sets have different non-linear lower frontiers. As seen in the figures, set T_2 is a subset of set $\bar{T}_2$, which also contains points which have arbitrarily higher amounts of emission and lower amounts of the input than in set T_2.

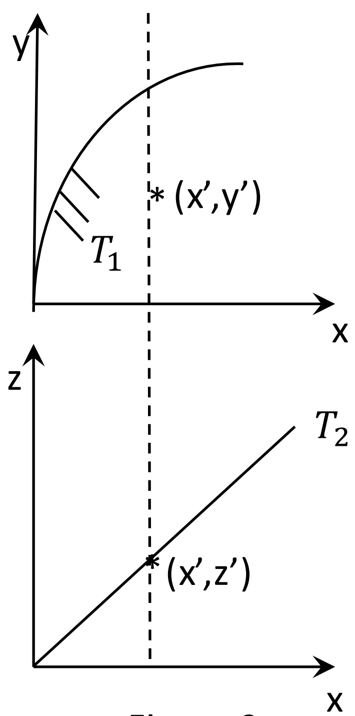

Figure 2

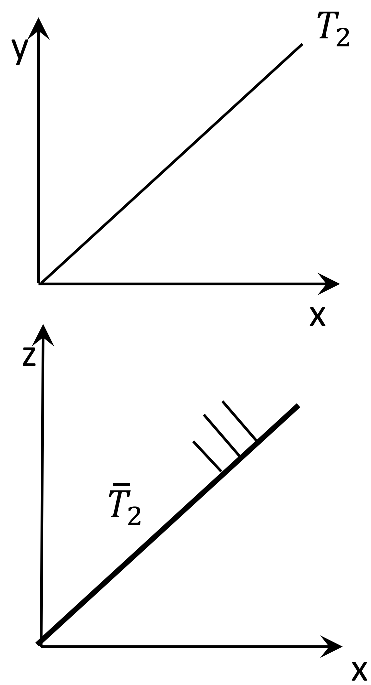

Figure 3

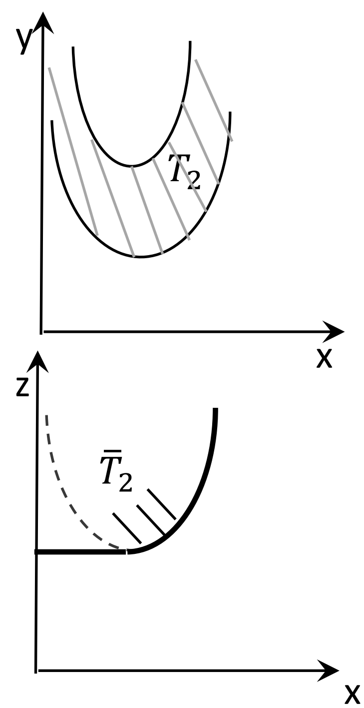

Figure 4

The construct $\bar{T}_2$ satisfies costly disposability of emission and emission-causing inputs. These disposability properties are defined in MRL and together imply that, if a given production vector lies in the constructed sub-technology $\bar{T}_2$, then so does any production vector with arbitrarily greater amount of emission and arbitrarily smaller levels of the emission-causing inputs, holding all other goods fixed at the initial levels. Thus, these properties are polar opposites of the properties of free input and output disposability. Note that set $\bar{T}_2$ in Figures 3 and 4 satisfies these properties. It is shown in Murty (2015) and Murty and Russell (2016, 2017) that these costly disposability properties of set $\bar{T}_2$ imply that the relation between emission and emission-causing inputs along its lower frontier is non-negative. This is apparent in the lower panels of both Figures 3 and 4. If we assume that the original sub-technology T_2 and the construct $\bar{T}_2$ have the same lower frontier, then the relation between emission and emission-causing inputs along the lower frontier of sub-technology T_2 is also non-negative. This is true in Figure 3, where our assumption is satisfied; but not in Figure 4 where our assumption fails. As seen in the upper panel of Figure 4, the lower frontier of sub-technology T_2 has a negatively sloped region.

To summarise, the assumption that sub-technology T_2 and its costly disposal hull $\bar{T}_2$ have a common lower frontier implies that, along the efficient frontier of sub-technology T_2, the relation between emission and any emission-causing input is non-negative: As the amount of the emission-causing input increases, the emission generated either increases or remains the same.

The costly disposal hull of the emission-generating technology T.

Define the set $\bar{T}$ as the intersection of sub-technology T_1 and the costly disposal hull of sub-technology T_2:

$$\bar{T} = T_1 \cap \bar{T}_2. \tag{2}$$

It follows from (1) and (2) that the original technology T is a subset of $\bar{T}$. All data points in T will also be contained in $\bar{T}$. The two sets differ only with respect to the emission-generating sets T_2 and $\bar{T}_2$, where $T_2 \subseteq \bar{T}_2$. Under our maintained assumption that the two sets have the same lower frontier, the distance of a data point from the lower frontier of sub-technology T_2 will be the same as its distance from the lower frontier of set $\bar{T}_2$. Hence, for the purpose of measuring efficiency, we need focus only on set $\bar{T}_2$ rather than the original sub-technology T_2. This is helpful because, as we shall see below in Section 4, the former set has a convenient DEA representation. We will call $\bar{T}$ a costly disposal hull of the original technology T. The technology defined and adopted in MRL can be interpreted as the costly disposal hull $\bar{T}$ of the actual technology.

2.2.4 Some comparisons of by-production approach with input and weak-disposability based output approaches.

Under the by-production approach, the emission-causing inputs are a part of both sub-technologies T_1 and $\bar{T}_2$. Changes in these inputs have both an effect on intended output production (in accordance with the engineering rules defining sub-technology T_1) and on emission generation through relations such as the thermodynamic laws that define sub-technology T_2 and hence set $\bar{T}_2$. The two sub-technologies are connected through emission-causing inputs. Hence, under the maintained assumptions on sets T_1 and $\bar{T}_2$, changes in the levels of these inputs imply increases both in intended-output production and in emission generation for an efficient production unit when all other inputs are held fixed. Thus, under the by-production approach, along the efficient frontier of the over-all technology set T (or its costly disposal hull $\bar{T}$), emission generation and production of the intended output are correlated – in particular, the co-movement in these bad and good outputs occurs due to changes in the usage of emission-causing inputs by the producing unit. Clearly, this is in contrast to the input and weak-disposability based output approaches to modelling emission-generating technologies where, as discussed in Section 2.2.1 and illustrated in Figure 1, the positive relation relation between emission generation and intended production along the efficient frontier is demonstrated holding all (including emission-causing) inputs fixed.

If data on emission was generated by some formula involving data on all emission-causing inputs and their emission factors, then the lower frontier of T_2 will simply be the graph of this formula; *e.g.,* suppose data for emission was computed using the formula: $z = \alpha_c x_{z_c}$, where α_c is the amount of the emission generated per unit of coal input and x_{z_c} is the input of coal used by the producing unit. If α_c is constant across all producing units, then the lower frontier of T_2 will be the set of all production vectors that satisfy this formula, and will hence be linear like in the top panel of Figure 3. Observed production vectors of all producing units will lie on it. Hence, because of the way the data on emission is generated, no unit can be deemed inefficient in emission-generation. The input and weak-disposability based output approaches will not be able to capture this case. As, discussed above, such a positive relation between the emission and the emission-causing input cannot be obtained in these approaches. Rather, in these approaches, the positive association between emission-generation and intended production along the frontier of the technology was shown for given levels of all inputs (*i.e.,* holding levels of all (including emission-causing) inputs fixed). But in this case, the efficient (lowest) emission level is fixed once the coal usage level is fixed.

It is also clear that, since T_1 satisfies free-disposability with respect to emission-causing inputs and $\bar{T}_2$ satisfies the polar opposite property of costly-disposability with respect to emission-causing inputs, the costly disposal hull $\bar{T}$ of technology T cannot be freely disposable with respect to these inputs. Again, this is in contrast to the input and weak-disposability based output approaches, where the technologies are freely disposable in the emission-causing inputs.

3 Aggregated and disaggregated measures of coal input: heat as the appropriate unit of measurement.

3.1 Aggregated versus disaggregated measures of coal.

We will be concerned with measuring technical efficiency of Indian coal-based thermal power plants, which use different types of coal varying with respect to their composition, gross calorific values (GCVs), emission-generating potentials, *etc.*[14] For example, a broad categorisation of different types of coal based on the GCV includes varieties such as lignite, sub-bituminous coal, bituminous coal, and anthracite. In applied works, we can run into a degrees of freedom or a curse of dimensionality problem if we employ a highly disaggregated model of a technology that treats each of the different grades of a given fossil fuel such as coal as a unique input. Data may also often not be available at such a disaggregated level. Hence, a single aggregated amount (or an index) of coal input is generally used in most applied works.[15] Usually, this aggregated amount is expressed either in mass units obtained by summing up over the masses of all the different types of coal used by the power plant (see *e.g.,* Sueyoshi and Goto (2010, 2011, 2012)) or in heat units obtained by summing up over the heat contents of the different types of coal used. The heat contents of the different types of coal can be computed either by multiplying their GCVs with their respective physical amounts (see *e.g.,* Färe et al (1989), Pasurka (2006), MRL, Sahoo et al (2017)) or by employing their GCVs to express their amounts in equivalent units of a reference fossil-fuel, *e.g.,* equivalent units of oil or any particular type of coal (see *e.g.,* Khanna et al (1999)).

3.2 Consistency of measuring aggregate input of coal in heat units with theoretical models of emission-generating technology.

It is important to verify whether measuring the coal input in aggregate terms is consistent with the disposability properties of inputs and outputs that are proposed in the theoretical models of emission-generating technologies. In this regards, we argue below that it is better to measure the aggregate value of coal in heat units rather than in mass units. We discuss below that it is the heat value of aggregate coal and not its physical mass that is positively related to *both* production of electricity and the generation of emission along the efficient frontier of an emission-generating technology. Recall that the by-production approach discussed in Section 2.2.3 incorporated these relations between emission-causing inputs, emissions, and the intended

[14] The GCV of a fuel is the heat content of a physical unit of the fuel.

[15] This is also true with respect to other fossil-fuels such as oil and gas, each of which too may have several types. The arguments made below hence hold also for these fossil fuels.

output. Precisely, along the efficient frontier of the technology, an increase in usage of coal input measured in heat units implies increases in both electricity and emission generation. This may not hold if we measure the aggregate amount of coal (or the coal index) in mass units.[16]

From the point of view of production of electricity, it is intuitive that an increase in the heat input will, if efficiently used, have a direct increasing effect on the electricity generated immaterial of the mix of coal-types used. In coal-fired power plants, it is the heat generated by a pulverised coal-fired boiler that is used to boil water to create steam at high pressure, which then expands through and turns a steam turbine. The turbine hence converts steam energy to work, which the alternator converts into electrical power. Hence, a given increase in the heat input, if efficiently used, can be expected to result in an increase in the electricity generated. This is immaterial of the change in the mix of coal types that is burned in the boiler as long as the change in the mix achieves the given increase in heat. Engineers who design thermal power plants are concerned with maximising energy efficiency, which they define as the ratio of electricity (output) generated to the heat input employed.[17]

It is to be noted that the CO_2 emission level from burning a particular type of coal depends on its carbon content, while its GCV (and hence its heat content) depends also on other factors in addition to its carbon content, *e.g.,* GCVs of fuels with more hydrogen and less moisture content is higher.[18] A given level of heat input can be generated by several combinations of coal types varying with respect to emission factors and GCVs. The combination that leads to the least amount of emission given the fixed level of heat would correspond to a point that lies on the lower frontier of sub-technology T_2, where coal, the emission-causing input is measured in heat units. To see this, suppose there are L types of coal and the amount of the l^{th} type measured in mass units is denoted by X_l. Suppose α_l denotes GCV of the l^{th} type of coal and β_l denotes the amount of CO_2 emission generated from burning one mass unit of this type of coal. The various combinations/mixes of coal types that can generate x_z amount of heat input is given by the following set:

$$\Omega(x_z) := \left\{ \langle X_1, \ldots, X_L \rangle \in \mathbf{R}_+^L \;\middle|\; \sum_{l=1}^{L} \alpha_l X_l = x_z \right\}.$$

The emission generated by using a vector $\langle X_1, \ldots, X_L \rangle \in \mathbf{R}_+^L$ of coal types is given by

$$z = \sum_{l=1}^{L} \beta_l X_l.$$

[16]Thus, measuring the aggregate coal input in mass units may imply violations of the desired disposability properties of an emission-generating technology discussed in Section 2.2.

[17]See Appendix I of the study by Coal Industry Advisory Board (CIAB) (2010) for a thorough justification.

[18]See e.g., Schweinfurth (2009).

The minimum amount of emission that can be generated when x_z amount of heat input is used by a power plant is given by function $g(x_z)$, which is obtained by choosing a combination of coal types that solves the following problem:

$$g(x_z) := \min_{X_1,\ldots,X_L} \left\{ \sum_{l=1}^{L} \beta_l X_l \middle| \ \langle X_1, \ldots, X_L \rangle \in \Omega(x_z) \right\} \tag{3}$$

In the problem above, emission generation is minimised holding the level of heat input fixed. Suppose the solution vector to the above problem is denoted by the function $\psi(x_z) = \langle X_1, \ldots, X_L \rangle = X$. Clearly, if x_z amount of heat input is employed, then the levels of emission generated by all the different coal combinations in $\Omega(x_z)$ are feasible under technology T_2. Of these emission levels, $g(x_z) = \sum_{l=1}^{L} \beta_l X_l$ (where $X = \psi(x_z)$) is the minimum level of emission generated. Thus, $\langle x_z, g(x_z) \rangle$ lies on the lower frontier of sub-technology T_2.[19] The proposition below states that function g is increasing in x_z, *i.e.*, the relation between the heat input and emission along the lower frontier of sub-technology T_2 is positive.

Proposition 1 *If $\bar{x}_z = \lambda x_z$ and $\psi(x_z) = X$, then $\psi(\bar{x}_z) = \lambda\psi(x_z) = \lambda X$ and $g(\bar{x}_z) = \lambda g(x_z)$, where $\lambda > 0$.* [20]

From the arguments above it follows that if the heat input increases from x_z to $\bar{x}_z$ then both the minimum amount of emission and the maximum amount of electricity that can be generated increase. This is achieved when mix of coal types underlying the increase in heat changes from $\psi(x_z)$ to $\psi(\bar{x}_z)$.

3.3 Measuring aggregate input of coal in mass units.

The positive correlation between emission and electricity along the efficient frontier of the technology discussed above may not hold when coal is measured in aggregated mass units. This is because, in this case, the mix of the coal types forming the aggregate input of coal affects both the generation of electricity and emission.[21] Consider a unit increase in mass of the aggregated coal input that increases the maximum amount of electricity that can be produced.

[19] To keep notation simple, we are ignoring the intended output and non-emission-generating inputs components of this production vector, as emission-generation is assumed to be unaffected by these goods.

[20] **Proof.** Suppose $\psi(\bar{x}_z) \neq \lambda X$. Then there exists $\bar{X} \neq \lambda X$ such that $\bar{X} = \psi(\bar{x}_z)$. Since $\sum_{l=1}^{L} \alpha_l X_l = x_z$, we have $\sum_{l=1}^{L} \alpha_l \lambda X_l = \lambda x_z = \bar{x}_z$. Hence, $\lambda X \in \Omega(\bar{x}_z)$. Since $\bar{X} = \psi(\bar{x}_z)$, $\lambda X \in \Omega(\bar{x}_z)$, and $\lambda X \neq \psi(\bar{x}_z)$, we have $g(\bar{x}_z) = \sum_{l=1}^{L} \beta_l \bar{X}_l < \sum_{l=1}^{L} \beta_l \lambda X_l$. This implies $\sum_{l=1}^{L} \beta_l \frac{\bar{X}_l}{\lambda} < \sum_{l=1}^{L} \beta_l X_l$. Also, $\sum_{l=1}^{L} \alpha_l \bar{X}_l = \bar{x}_z$ implies $\sum_{l=1}^{L} \alpha_l \frac{\bar{X}_l}{\lambda} = \frac{\bar{x}_z}{\lambda} = x_z$. Hence, $\frac{\bar{X}}{\lambda} \in \Omega(x_z)$. This contradicts $\psi(x_z) = X$ as $\sum_{l=1}^{L} \beta_l \frac{\bar{X}_l}{\lambda} < \sum_{l=1}^{L} \beta_l X_l$. ■

[21] Recall that, when the aggregate coal input is measured in heat units, the changes in the mix of coal types had no affect on the generation of electricity employed, provided all these mixes achieve the same level of heat input. However, such changes in the mix of coal types impacted the amount of CO_2 generated.

This increase in mass of the aggregated coal input must involve an increase in the heat input. Suppose this happens with an increase in the usage (in terms of physical mass) of high GCV types of coal that outweighs the decrease (if any) in the usage of low GCV varieties of coal. However, if the carbon contents of the high GCV varieties of coal are lower than those of the relatively lower GCV varieties of coal, then this increase in aggregate coal input would imply a reduction in the level of emission generated.[22] Thus, when the aggregated coal input is measured in mass units, an increase in the usage of aggregated coal input brought about by a *given* change in the mix of coal types used may not imply increases in both electricity and emission generation, which is contrary to what is exhibited by the models of emission-generating technologies discussed in Section 2.2.3. Given a fixed amount of aggregate input of coal measured in mass units, the mix of coal types corresponding to the maximum generation of electricity could be quite different from the mix associated with the minimum level of emission.[23]

4 Methodology for measuring efficiency of Indian coal-based thermal power plants.

In this paper, the by-production approach will be adopted to specify the technology, and efficiency of Indian thermal power plants will be computed using the Färe, Grosskopf, Lovell (FGL) output-based efficiency index discussed in MRL. This index is motivated by the works of Färe and Lovell (1978) and Färe, Grosskopf, and Lovell (1985). This index measures efficiency in the directions of the intended output and emission holding input levels fixed. We will employ DEA techniques to measure technical efficiency of Indian thermal power plants. As discussed in Section 2.2.3, we will construct and work with the costly disposal hull $\bar{T}$.

We first distinguish between emission-causing and non-emission causing inputs. We assume that the first n_o (where $0 < n_o < n$) inputs are non-emission causing whose quantity vector is denoted by x_o. The remaining $n - n_o \equiv n_z$ inputs are emission-causing whose quantity vector is denoted by x_z. Thus an input vector can be partitioned as $x = \langle x_o, x_z \rangle$.

Suppose there are U decision making units (DMUs), which in our case are thermal power

[22] For example, such an increase in physical units of coal can be achieved by shifting from using anthracite coal to bituminous coal, where the GCV of the former is lower than that of the latter variety (see, *e.g.,* Schweinfurth (2009)).

[23] Output-based efficiency measurement in this scenario would more often than not imply that, given fixed amounts of all inputs including the aggregate coal input measured in mass units, if a DMU employing a particular mix of coal types operates on the upper frontier of sub-technology T_1 then it will not operate on the lower frontier of sub-technology T_2, and vice-versa.

plants. These are indexed by u. The $U \times n$-dimensional data matrix of inputs is denoted by

$$X = \begin{bmatrix} X_o & X_z \end{bmatrix},$$

where X_o and X_z are, respectively, the $U \times n_o$ and $U \times n_z$-dimensional data matrices of non-emission causing and emission-causing inputs. The $U \times 1$-dimensional data matrices of intended output and emission are denoted by Y and Z, respectively.

DEA construction of set $\bar{T}$ is along the lines discussed in MRL. The DEA specification of sub-technology technology T_1 is[24]

$$T_1 = \left\{ \langle x, y, z \rangle \in \mathbf{R}_+^{n+2} \;\middle|\; \lambda^\top X \leq x,\ \lambda^\top Y \geq y,\ \lambda \geq 0_U \right\} \tag{4}$$

This specification implies that sub-technology technology T_1 exhibits constant returns to scale, is convex, and satisfies free input and output disposability with respect to all inputs and the intended output, respectively. The DEA specification of the sub-technology $\bar{T}_2$ is[25]

$$\bar{T}_2 = \left\{ \langle x_o, x_z, y, z \rangle \in \mathbf{R}_+^{n+2} \;\middle|\; \mu^\top X_z \geq x_z,\ \mu^\top Z \leq z,\ \mu \geq 0_U \right\} \tag{5}$$

Sub technology $\bar{T}_2$ also exhibits constant returns to scale and is convex. It is to be noted that this set satisfies costly disposal of emission and the emission-causing inputs. The costly disposal hull of the overall technology is then obtained as the intersection of sub-technologies T_1 and $\bar{T}_2$:

$$\bar{T} = \left\{ \langle x_o, x_z, y, z \rangle \in \mathbf{R}_+^{n+2} \;\middle|\; \begin{array}{l} \lambda^\top X \leq x,\ \lambda^\top Y \geq y, \\ \mu^\top X_z \geq x_z,\ \mu^\top Z \leq z, \\ \lambda \geq 0_U,\ \mu \geq 0_U \end{array} \right\}$$

This overall technology also exhibits constant returns to scale, is convex, satisfies free input disposability of the non-emission causing inputs and free output disposability of the intended output, but is not freely disposable in the emission-causing inputs. It satisfies costly disposability with respect to the emission.

Computation of the FGL output-based efficiency index as defined in MRL for any DMU u amounts to (i) computing its efficiency in the production of the intended-output relative to sub-technology T_1. This is denoted by β_y^u; (ii) computing its efficiency in the generation of the emission relative to sub-technology $\bar{T}_2$. This is denoted by β_z^u; and (iii) computing its total efficiency by taking an average of the efficiency in intended-output production and the efficiency

[24]The vector $\lambda^\top$ denotes the transpose of a U dimensional vector λ. 0_U is a U-dimensional zero vector.

[25]The vector $\mu^\top$ denotes the transpose of a U dimensional vector μ.

in emission generation. This is denoted by $\beta^u = \frac{\beta_o^u + \beta_z^u}{2}$.

The efficiency in intended-output production for DMU u whose production vector is denoted by $\langle x^u, y^u, z^u \rangle$ is obtained by first computing θ^u that solves the problem

$$\max \left\{ \theta \geq 0 \;\middle|\; \langle x^u, \theta y^u, z^u \rangle \in T_1 \right\},$$

where T_1 is defined in (4). Since $\langle x^u, y^u, z^u \rangle \in T_1$, we have $\theta^u \geq 1$. The efficiency in intended output production for DMU u is then obtained as $\beta_y^u = \frac{1}{\theta^u}$. For a DMU that lies on the upper frontier of set T_1, we have $\theta^u = 1$ and hence, $\beta_y^u = 1$.

The efficiency in emission-generation for DMU u is given by β_z^u that solves the problem

$$\min \left\{ \beta_z \geq 0 \;\middle|\; \langle x_o^u, x_z^u, y^u, \beta_z z^u \rangle \in \bar{T}_2 \right\},$$

where $\bar{T}_2$ is defined in (5). Since $\langle x_o^u, x_z^u, y^u, z^u \rangle \in \bar{T}_2$, we have $0 \leq \beta_z^u \leq 1$. For a DMU that lies on the lower frontier of set $\bar{T}_2$, we have $\beta_z^u = 1$.

5 Data.

The study uses the annual data of 48 coal-fired thermal power plants for 11 years starting from 2003 to 2015, excluding years 2007 and 2012. Data on most variables was collected from the annual publication of the Central Electricity Authority (CEA) "Review of performance of thermal power plants in India (RPTPPI)."[26] The 48 plants studied are run by 16 major power generating companies operating in various states of India.[27] Of these, NTPC is the central government undertaking, two are private undertakings, while the remaining are state electricity companies. The power plants in our sample constitute 50 to 70 percent of the total installed capacity of thermal power plants in India over the years of our study. Table A in the appendix and Table 11 give the details of power generating companies, the names and numbers of plants of each of these companies included in our sample.

The data set constitutes an unbalanced panel of 495 observations on electricity, CO_2 emissions, plant capacity, fossil-fuel measured in heat units, and plant operating availability. Electricity generated is considered as the intended output, while CO_2 emission is taken as the bad

[26] RPTPPI was not published in 2012. Also, since some crucial data was missing for the year 2007, this year was excluded from the study.

[27] A thermal power plant/station comprises of multiple units which could be of same or different capacity. Each unit of the thermal power plant comprises of main plant equipment while some other plant infrastructure like coal handling facility, chimney etc are common across units. These units could be commissioned together or at different times. Since unit level data is not available for several variables, we use the plant-level data for our analysis.

output in the estimation. The remaining variables are inputs used by power plants, with fossil-fuel being the emission-causing input. Table 1 provides the descriptive statistics of the variables used in the estimation of the efficiency indexes for the first and the last years of our study.

TABLE 1: Descriptive statistics of data on inputs and outputs

	mean		maximum		minimum	
	2003	2015	2003	2015	2003	2015
Net generation	6091.35	8576.66	15747.8328	27594	738.974	485.136
CO_2 emission	6616878.65	8782090.23	16437781.79	26548438	1050098.97	966999.495
Capacity	977.90	1471.30	2340	4260	135	240
Aggregate heat	17282919.46	23099027.02	41488624.13	70859304.46	3046659	2397060
Operating availability	85.22	78.23	95.14	97.47	60.46	34.61

Units:

Net generation-GWh, CO2 emission-Mtons, Capacity-MWh, Aggregate heat-mill of Kcal, Operating availability-MWh

Net electricity produced by the plant during a year is measured in gigawatt hours (GWh). The data for the variable is collected from the CEA website.[28] It is computed as the difference between the gross electricity generation of a power plant and its auxiliary power consumption.[29] A power plant may be generating high gross electricity but low net electricity due to high auxiliary power consumption. Using net electricity generation helps us capture plant's inefficiency due to high auxiliary power consumption.

Systematic data on labour and capital employed in Indian power plants was not available. As in other recent papers studying Indian thermal power plants (see *e.g.,* Sahoo et al. (2017) and Behera et al., (2010)) plant capacity is used as a proxy for capital used by power plants. It is the installed capacity expressed in megawatt (MW). Labour is excluded as an input of thermal power plants in this study. It has been argued that the contribution of labour cost to total operating costs of these power plants is very small (see *e.g.,* Kumar et al (2015)).

The primary fuel input of coal based thermal power plants is coal, while oil is used as a secondary fuel. Oil is used only to cover the start-up fuel requirements and for flame stabilisation and does not contribute significantly to electricity generation in coal-fired plants. At the same time, as will be discussed below, CEA reports data on total emission from combined coal and oil combustion by each thermal power plant.[30] However, comprehensive data on oil consumption by coal-based thermal power plants is not released by CEA for all the years of our study. Hence, in this work, aggregate heat from coal and oil consumption by the coal-based thermal power

[28]Link: *http : //cea.nic.in/tpeandce.html*

[29]Auxiliary power consumption by thermal power stations comprises the power consumption by all the unit auxiliaries as well as the common station/unit requirements such as station lighting, air conditioning etc.

[30]See (8).

plants measured in millions of kilocalories (mill of Kcal) is taken as the single fossil-fuel input. Data on total heat content of the two fuels is calculated by multiplying the station heat rate (SHR) with the amount of gross electricity generated, where the SHR measures the heat input used per unit of gross electricity generation. CEA provides plant-level data on SHR, which it calculates using the following formula:

$$SHR = (SCC \times GCV_c) + (SOC \times GCV_o), \tag{6}$$

where SCC and SOC denote, respectively, the specific coal consumption and specific oil consumption; while GCV_c and GCV_o denote, respectively, the GCVs of coal and oil.[31] If Y denotes the amount of gross electricity generated by the thermal power plant, then total amount of heat (the emission-generating input) used by it, denoted by x_z, is given by

$$\begin{aligned} x_z &= Y \times SHR \\ &= (Y \times SCC \times GCV_c) + (Y \times SOC \times GCV_o) \quad = \quad x_{z_c} + x_{z_o}, \end{aligned} \tag{7}$$

where $Y \times SCC \times GCV_c = x_{z_c}$ is the heat obtained from coal and $Y \times SOC \times GCV_o = x_{z_o}$ is the heat obtained from oil.

It is to be noted that, although the power plants in our data set use a range of coal-types varying with respect to GCVs, for every power plant, CEA reports SHR and SCC data for only an aggregate measure of coal used by it that is measured in mass units.

Data on emissions is not generated by direct observation. Rather, the following formula is used by CEA to compute the emission level (measured in metric tons (Mtons) from fossil-fuel consumption by plant u in year t:

$$z^{u,t} = \left[X_c^{u,t} \times GCV_c^{u,t} \times EF_c \times Oxid_c\right] + \left[X_o^{u,t} \times GCV_o \times EF_o \times Oxid_o\right], \tag{8}$$

where for fuel type $i = c, o$, EF_i and $Oxid_i$ are the emission and oxidation factors of the fossil-fuel i, respctively. EF_i is the amount of emission per unit of heat generated by burning fossil-fuel of type i. EF_c and $Oxid_c$ are assumed to be constant across all plants and years taking values 92.5 grams per megajoules (g/MJ) and .98, respectively; while $Oxid_o$ takes a value equal to one for all years and plants and EF_o takes a value 73.7 g/MJ prior to year 2008 and a value 71.9 g/MJ in the post 2008 period for all plants.[32] We hence interpret these emission and oxidation factors as averages across all coal and oil types employed by all power

[31] SCC measures the physical units of coal required to generate one unit of electricity and GCV_c is the heat content of a physical unit of coal. SOC and GCV_o are similarly defined.

[32] See Appendix B of the user guide of CO_2 baseline database for the Indian power sector, published by CEA.

plants. The GCV of coal varies across plants and years and is not reported, while the GCV of oil is taken to be a constant equal to 10100 Kcal per litre by the CEA. $X_c^{u,t}$ and $X_o^{u,t}$ denote the physical units of coal and oil consumed by power plant u in year t, respectively. These can be computed as follows:

$$X_{u,t} = \left[SCC^{u,t} \times Y^{u,t}\right] \qquad \text{and} \qquad X_o^{u,t} = \left[SOC^{u,t} \times Y^{u,t}\right], \tag{9}$$

where $Y^{u,t}$ is the gross generation of electricity by power plant u in year t.

Substituting (9) into (8) and recalling (7), the total emission generated is given by

$$\begin{aligned} z^{u,t} &= \left[(\text{heat from coal})^{u,t} \times EF_c \times Oxid_c\right] + \left[(\text{heat from oil})^{u,t} \times EF_o \times Oxid_o\right] \\ &= \left[x_{z_c}^{u,t} \times EF_c \times Oxid_c\right] \quad + \quad \left[x_{z_o}^{u,t} \times EF_o \times Oxid_o\right] \end{aligned} \tag{10}$$

Following authors such as Sueyoshi and Goto (2010, 2011, 2012) and Sahoo et al. (2017), our model also includes a managerial input, namely the plant operating availability. It is the percentage of total capacity (measured in MWh) that is unutilized by the plant. It is defined as

$$\text{Operating availability} = 100 - \text{forced outage} - \text{planned maintenance},$$

where forced outage is power plant shutdown due to unexpected breakdown and planned maintenance is the planned shutdown for scheduled maintenance of a power plant. Both of these are measured as percentages of total capacity.

6 Results and interpretations.

Efficiency indexes of coal-based thermal power plants in India are computed for each year, by constructing a separate DEA by-production technology for every year. Production, environmental, and technical efficiency indexes are derived for each plant in each year relative to subtechnologies T_1 and $\bar{T}_2$ and the overall by-production technology $\bar{T}$, respectively, constructed for that year.

6.1 Environmental efficiency of Indian coal-based thermal power plants.

Since we assume technological constant returns to scale and only one emission-causing input, namely, aggregate heat from coal and oil used by coal-based thermal power plants, the lower frontier of the technology T_2 (or its costly disposal hull $\bar{T}_2$) constructed using DEA methods will be linear, *i.e.,* in the space of the heat input (plotted on the X-axis) and the CO_2 emission (plotted on the Y axis), it will be an upward sloping ray through the origin. An example is shown in Figure 3. Suppose the slope of this ray is $\gamma \geq 0$, so that the equation of the lower frontier in the heat-CO_2 space is

$$z \quad = \quad x_z \times \gamma \quad = \quad [x_{z_c} + x_{z_o}] \times \gamma \tag{11}$$

However, as discussed in Section 5, the emission data reported by CEA are computed using formula (10) as

$$z \quad = \quad x_{z_c} \times EF_c \times Oxid_c \; + \; x_{z_o} \times EF_o \times Oxid_o. \tag{12}$$

It is clear from (12) that data on CO_2 emission is, in general, not linear in the *aggregate* heat input employed, unless $EF_c \times Oxid_c$ is equal to $EF_o \times Oxid_o$; in which case, (11) will imply that $\gamma = EF_c \times Oxid_c = EF_o \times Oxid_o$, so that sub-technology T_2 and its lower frontier will coincide. This case is illustrated in the upper panel of Figure 3, where *all* points of T_2 are environmentally efficient. This case will, however, generally not arise in reality. Plants employing same amount of heat input can generate different amounts of CO_2 emissions because the shares of coal and oil in the total heat input employed could differ across the plants. Since values employed by CEA imply that $EF_o \times Oxid_o < EF_c \times Oxid_c$ (see Section 5), the plant with a higher (respectively, lower) share of oil in the total heat input will have a lower (respectively, higher) level of emission. In this case, set T_2 will also have points that lie above its lower frontier. The emission level of a plant that lies on the lower frontier of sub-technology T_2 (defined in (11)) are minimal among all plants who employ the same heat input as itself. From the arguments made in the context of problem (3) in Section 3.2, such a plant employs the emission-minimizing mix of coal and oil given its heat input. In particular, the share of oil in the total heat consumed must be highest for this plant among all plants consuming the same amount of heat as itself.

Despite the discussion above, our empirical analysis reveals that the correlation between data on CO_2 emission and heat input for Indian thermal plants is near perfect for every year (see Table 2). Thus, the heat-CO_2 plots are nearly linear for each year of our study in the Indian case. The descriptive statistics in Table 3 show that the output-based environmental efficiency index β_z takes high values for all thermal power plants and shows very little variability. For example, in 2015, for 45 out of the 46 plants in our data set, β_z ranged between one to 0.83, with 76 percent of the the plants scoring between one to 0.89.

Table 2: Correlation between CO2 emission and heat in our data

2003	2004	2005	2006	2008	2009	2010	2011	2013	2014	2015
0.9945	0.9912	0.9918	0.9985	0.9985	0.9982	0.9986	0.9992	0.9996	0.9993	0.9985

TABLE 3: Descriptive statistics of environmental efficiency index.

Year	Mean	Std Dev
2003	0.859	0.064
2004	0.802	0.070
2005	0.916	0.049
2006	0.951	0.038
2008	0.826	0.038
2009	0.797	0.046
2010	0.735	0.053
2011	0.895	0.039
2013	0.909	0.027
2014	0.900	0.027
2015	0.890	0.040

Table 4: Share of oil in total heat in plants with environmental efficiency equal to one.

		Share of oil			
Plant Name	**Year of T2 efficiency**	**2008**	**2009**	**2010**	**2011**
BANDEL TPS	2003	5.96E-04	6.71E-04	2.04E-03	NA
BHUSAWAL TPS	2014	3.03E-03	1.90E-03	1.23E-03	3.32E-03
BUDGE BUDGE TPS	2008	3.97E-05	NA	6.99E-04	3.38E-04
GH TPS (LEH.MOH)	2009, 2015	8.59E-05	2.22E-03	1.16E-04	2.19E-04
GND TPS (BHATINDA)	2004, 2006, 2013	4.44E-04	8.04E-04	7.70E-04	1.30E-03
METUR TPS	2005	1.80E-04	2.75E-04	1.42E-04	3.96E-04
PARLI TPS	2011	9.88E-04	2.35E-03	1.14E-03	2.38E-03
SIKKA REP. TPS	2010	9.64E-04	1.74E-03	1.37E-03	1.17E-03

TPS: Thermal power plants

The near perfect correlation between data on CO_2 emission and heat input and the small variability of the environmental efficiency index in every year of our study can be attributed to the insignificant share of oil in the total heat input. Table 4 lists the plants that lie on the lower frontier of sub-technology T_2 (for which $\beta_z = 1$) and the share of oil in the total heat employed by them.[33] As seen in this table, the share of oil in total heat input employed by coal-based thermal power plants is tiny. This implies that, for coal-based thermal power plants, the total heat input can be approximated by its coal component, *i.e.*, $x_z \approx x_{z_c}$. Hence, we can ignore the contribution of oil to emission generation in coal-based thermal power plants and rewrite

[33] These figures are reported for years for which some limited data on oil is available from the CEA.

the formula (12) that computes emission generated by a plant as

$$z \quad \approx \quad x_z \times EF_c \times Oxid_c. \tag{13}$$

Table 5 compares the slope of the lower frontier of sub-technology T_2 (given by γ in (11)) with the product $EF_c \times Oxid_c$, which is the slope of the line in (13) that approximates the amounts of emission generated by different power plants. The value of γ is given by the ratio of CO_2 emission and heat input for the plant on the lower frontier of sub-technology T_2 constructed using DEA methods. The product $EF_c \times Oxid_c$ is assumed to take a constant value of 0.3795 Mtons/mill of Kcal by CEA.[34] As seen in the table, the values of γ for each year is not very different from $EF_c \times Oxid_c$. This implies that the lower frontier of sub-technology T_2 represented by (11) and the actual data on CO_2 and heat, which solve equation (13), are not very far apart.

Table 5: Slope of T2 frontier (ratio of emission and heat for plant on T2 frontier)

2003	2004	2005	2006	2008	2009	2010	2011	2013	2014	2015
0.326	0.307	0.352	0.368	0.310	0.300	0.275	0.342	0.345	0.339	0.339

$EF_c \times Oxid_c$ = .3795 Mtons/mill Kcal

Thus, given that (i) oil contributes insignificantly to the heat input in coal-based power plants (ii) CEA provides only aggregate data on the physical units of coal employed by power plants and (iii) CEA employs a constant value of $EF_c \times Oxid_c$ across all plants and years, the data on emission of CO_2 and heat provided by CEA will be in a strong (but generally not perfect) linear relationship in each year, where the linear relationship can be approximated by either (11) or (13). Thus, given the way CEA generates data on emission in India, the entire sub-technology T_2 corresponds approximately to a straight line that represents also its lower frontier.[35]

[34] See Section 5. We convert the unit g/MJ into Mtons/mill of Kcals for comparability with our analysis.

[35] Suppose the emission level was computed using different emission factors for different types of coal. If there were L coal types indexed by l, then the emission level would be computed as

$$z = \sum_{l=1}^{L} \left[x_{z_{c_l}} \times EF_{c_l} \times Oxid_{c_l} \right] + \left[x_{z_o} \times EF_o \times Oxid_o \right].$$

Since share of oil is small, we have

$$z \approx \sum_{l=1}^{L} \left[x_{z_{c_l}} \times EF_{c_l} \times Oxid_{c_l} \right]. \tag{14}$$

On the other hand, the lower frontier of sub-technology T_2 in the space of aggregate heat input and emission continues to be given by (11), as $z = \gamma x_z$, where $x_z = \sum_{l=1}^{L} x_{z_{c_l}}$. From (14) and (11) it follows that, as compared to the case where a single emission factor is employed, in this case, there will be considerable deviations of the data from the linear lower frontier of sub-technology T_2. Thus, as in MRL, considerable variations in

The output-based environmental efficiency index β_z computed for a by-production technology provides a measures of the extent to which the level of emission generated by a plant is above the minimum possible level corresponding to the heat input employed by it. Thus, (as discussed also in Section 4), this index holds the level of the emission-causing input (here heat) fixed. Since the CEA data set implies a sub-technology T_2, where all plants operate at points that are very close to its lower frontier, MRL's output-based environmental efficiency index β_z will take values very close to one for almost all power plants in the CEA data set. This implies that no significant improvements in environmental efficiency can be achieved by lowering the emission level of any plant holding its consumption of heat input fixed. Thus, an output-based efficiency index such as β_z loses its importance as a measure of environmental (in)efficiency in the context of data sets such as the CEA's.

6.2 Productive efficiency of Indian coal-based thermal power plants.

Productive efficiency of the thermal power plant, which is derived relative to sub-technology T_1, depends upon its efficiency in generating the intended output (electricity) from inputs. The factors that influence productive efficiency include basic plant design (technology embodied in the equipment), age and size of power plant, coal type used, operations and maintenance (O&M) practices, expenditure on modernization, and management practices of the generating company. Some non-controllable factors that affect plant's efficiency are the environmental conditions around power plant (especially ambient temperature, humidity etc.)

6.2.1 Annual trends.

Table 6 shows the mean values of the productive efficiency indexes for each year of our sample. The mean productive efficiency ranges from 0.83 to 0.89.

We categorise power plants on the basis of their performance as high, moderate and low performers. Plants with efficiency score in the top tertile of the range of efficiency scores obtained in a given year are called high performers in that year, plants with score in mid-tertile are moderate performers and those with score in the bottom tertile are low performers in that year. Table 7 gives the number of plants in the three categories for the earliest (2003), middle (2009), and last year (2015) of our sample. While the overall average productive efficiency score over the years has remained nearly the same (see Table 6), we see improvements in the efficiency distribution of plants.

environmental efficiency indexes of DMUs can be expected.

Table 6: Mean productive efficiency index

Year	Productive Efficiency	Year	Productive Efficiency
2003	0.84	2010	0.86
2004	0.84	2011	0.88
2005	0.88	2013	0.89
2006	0.87	2014	0.83
2008	0.88	2015	0.87
2009	0.87		

Table 7: Number of plants in low, moderate, and high efficiency categories.

Productive efficiency	year		
Efficiency category	2015	2009	2003
High	35	24	22
Moderate	10	12	16
Low	1	9	6

The high performing plants have increased from 50 percent in 2003 to 75 percent in 2015 whereas the low performing plants have decreased from 14 percent to 2 percent during the same period. This hints towards improvement in productive efficiency across these years.

6.2.2 Plant-level analysis.

Table 8 gives the ranking of power plants on the basis of their average productive efficiency scores over the years of study. The scores lie between 0.54 and one with an average efficiency score of 0.86 and standard deviation of 0.11. Roughly two third of the 48 plants that operated in all the years of our study are above average performers indicating that the distribution of mean productive efficiency of thermal power plants has a fat right tail.

All the top ranking plants – Ramagundem, Korba, Dahanu, Vindhyanchal, Simhadri – lie on the efficient frontier of sub-technology T_1 for most of the years. These plants are also performing consistently well throughout the period of study. On the other hand, the worst performing plants – Ennore, Rajghat, Sikka, Satpura – show decline in their performance. Incidentally, eight out of 10 top ranking plants – Ramagundem, Korba, Rihand, Singrauli, Vindhyanchal, Simhari, Talcher and Sipat are run by the public sector undertaking NTPC.[36] The other two plants are Dahanu, which is run by private company REL, and Kakatiya, which is owned by state-run APGENCO.

[36]See Table A in the appendix.

Table 8: Ranking of plants based on their average productive efficiency scores.

Plant name	Average efficiency score	Rank	Plant name	Average efficiency score	Rank	Plant name	Average efficiency score	Rank
Ramagundem	1	1	Kahalgaon	0.93	17	Chandrapura (DVC)	0.82	33
Korba	1	2	Tuticorin	0.91	18	Gandhi Nagar	0.82	34
Dahanu	0.99	3	Suratgarh	0.91	19	Chhabra	0.82	35
Vindhyanchal	0.99	4	Wanakbori	0.9	20	Nasik	0.81	36
Simhadri	0.99	5	Mettur	0.9	21	Panipat	0.79	37
Rihand	0.98	6	Kothagudem	0.9	22	Ukai	0.79	38
Singrauli	0.98	7	North Chennai	0.89	23	Bhusawal	0.79	39
Kakatiya	0.98	8	IB Valley	0.89	24	Sanjay Gandhi	0.75	40
Talcher	0.98	9	Kota	0.88	25	GNDTPS (leh Moh.)	0.71	41
Sipat	0.97	10	Chandrapur	0.88	26	Koradi	0.71	42
Rayalseema	0.97	11	Ropar	0.88	27	Bokaro B	0.71	43
Dr N Tata Rao	0.96	12	GHTPS (Bhatinda)	0.88	28	Satpura	0.7	44
Dadri	0.96	13	Raichur	0.86	29	Bandel	0.7	45
Unchahar	0.95	14	Khaparkheda	0.86	30	Sikka	0.68	46
Budge Budge	0.95	15	Parli	0.85	31	Rajghat	0.63	47
Farakka	0.93	16	Korba West	0.85	32	Ennore	0.54	48

On the other hand, all 10 bottom ranked plants are run by state-owned power generating companies. Besides Dahanu, the only other privately owned plant in our study, Budge Budge, is amongst the best performing plants.

6.2.3 Factors affecting productive efficiency.

In this section, we try to identify variables that would help in explaining the determinants of productive efficiency of Indian coal-based thermal power plants. For this, we run an ordinary least square regression with plant's efficiency scores as the dependent variable. In the past, several studies estimating efficiency of power plant (See and Coelli (2012); Khanna et al. (1999); Kumar and Managi (2009)) have used a number of explanatory variables, such as plant age, plant's capacity, ownership, fuel type, capacity utilisation, labor input and O&M expenditure.

In our regression model, we include plant availability, plant age, SHR, capacity, average unit capacity, and ownership dummies as explanatory variables. Table 9 gives the descriptive statistics of SHR, average unit capacity, and weighted average age.[37]

We report the regression results for two representative years, namely, an early year in our

[37] Weighted average age of a power plant is calculated by taking the capacity weighted average of the age of all units comprising the plant. For descriptive statistics of the remaining explanatory variables see Table 1 in Section 5.

sample such as 2005 and a later year in our sample such as 2013. There are 42 plants in our study in the year 2005 and 46 in 2013. The results of the regression are reported in Table 10.

Table 9: Descriptive statistics of new variables used in regression

Variables	Unit	Mean	Maximum	Minimum	Std Dev
2005					
SHR	Kcal/KWh	2619.44089	3746	2288	286.633653
Average unit capacity	MWh	228.72	500	90	97.93
weighted average age	Years	14.98	31.69	2.5	6.89
2013					
SHR	Kcal/KWh	2631.97427	4329	2283	382.807341
Average unit capacity	MWh	261.09	745	67.5	123.13
weighted average age	Years	18.54	41	2.34	8.71

Table 10: Regression results.

	Coefficients	Standard	t Stat	P-value
2005				
Intercept	1.296231	0.12901347	10.04725	1.44E-11
capacity	2.69E-05	1.05E-05	2.555735	0.015388
operating availability	0.002658	0.000733131	3.62535	0.000961
SHR	-0.00025	2.51E-05	-9.78363	2.79E-11
Centre	-0.00517	0.031746083	-0.1629	0.871594
state	-0.03072	0.030189497	-1.01766	0.316244
avg unit cap	-1.20E-05	7.92E-05	-0.15474	0.87797
age	-4.10E-05	0.000888978	-0.04628	0.963367
2013				
Intercept	1.283002	0.076596	16.75035	2.08E-19
capacity	1.11E-05	9.08E-06	1.219098	0.230129
operating availability	0.001861	0.000444	4.193459	0.000153
SHR	-0.00022	1.61E-05	-13.39	3.62E-16
centre	-0.00792	0.026294	-0.30116	0.764894
state	-0.01957	0.023861	-0.82003	0.41718
avg unit cap	4.94E-05	6.23E-05	0.793575	0.432245
age	9.05E-05	0.000734	0.123304	0.9025

The signs of the coefficients of variables are consistent with theory.[38] We discuss below the regression results with respect to each explanatory variable:

The SHR is a measure of the combined performance of all systems of a power plant: turbine cycle, boiler and other associated auxiliaries. A reduction in SHR imply improvement in these

[38]However, there is some variation in the significance of some variables across years.

systems and thus a lower heat requirement to generate one unit of electricity. This results in fuel saving and thus improvement in productive efficiency and lower emissions. The coefficient of SHR is significant and negative validating our proposition. Engineers define energy (thermodynamic) efficiency to be the inverse of SHR. Thus, our regression results show that there is a strong positive relation between energy efficiency as defined by engineers and productive efficiency as defined by economists.

Several studies show capacity utilisation as an important determinant of plant's efficiency (Hiebert, 2002; Khanna et al., 1999). In this study, we capture capacity utilization through operating availability. This is a managerial input, which captures management practices related to repair and maintenance and upkeep of the plant. However, lower operating availability can also result from fuel shortages. Its coefficient is positive and significant for both the years under consideration. Higher operating availability on account of less plant shutdown leads to high efficiency.

Plants in our dataset belong to the centre, state, and private power-generating companies. This is captured in the regression by including dummy variables for centre and state. The coefficient of centre and state is not significant hinting that ownership of plant does not impact its efficiency. However, this seems to be in contrast to our observation in Section 6.2.2 where we found that the majority of plants run by NTPC, a centre-owned power-generating company, are highly efficient, while plants run by state-owned utilities are mostly inefficient. One plausible reason could be that the ownership impacts the efficiency of a power plant through SHR and operating availability. Thus, plants of centre-owned power-generating companies are more productive efficient when they also have low SHR and high operating availability. This is analysed further in Section 6.2.5.

A-priori, there is evidence of age impacting both positively due to better adaptation to the production conditions and negatively due to older technology and greater wear and tear (Khanna et al, 1999). In our study, we find that age is not a significant determinant of the productive efficiency in any of the given years.

A plant comprises of several units of different capacities. In our sample, the unit size of plants vary between 62.5 and 660 MW. Moreover, the average unit capacity of plants is steadily increasing with time suggesting that new units added are of higher capacity. Bigger units are efficient because specific coal consumption reduces up to some point as the size of boiler increases (see Chan et al. (2014)). Larger units tend to have lower design heat rate and they also optimize on auxiliary power consumption (see Narayana and Bhatt (2004)). Moreover, plants with higher unit sizes also have large total capacity. There is also a reason to believe that a plant will experience economies of scale if it has larger capacity since there would be economies in usage of auxiliary plant facilities. We include both the total capacity and average unit capacity of a plant in the regression. The coefficient of capacity is significant for the year 2005 but not significant for 2014, while the coefficient of the average unit capacity of a power plant is not significant.

6.2.4 Performance of supercritical plants.

The term supercritical refers to main steam operating conditions in the boiler of a power plant being above the critical pressure of water (225.56 Kg/cm^2) while if the steam operating conditions are below this pressure, it is subcritical. In India, usually the power plants in the capacity of 100-600 MW are sub-critical while those with capacity greater than or equal to 660 MW are supercritical. Supercritical technology is primarily beneficial as it is more efficient and essentially requires less fuel per unit of electricity generation (low SHR) and thus leads to less emissions. Moreover, it offers more operational flexibility in running of the plant. Going forward, India needs to adopt supercritical technology to achieve large improvements in efficiency.

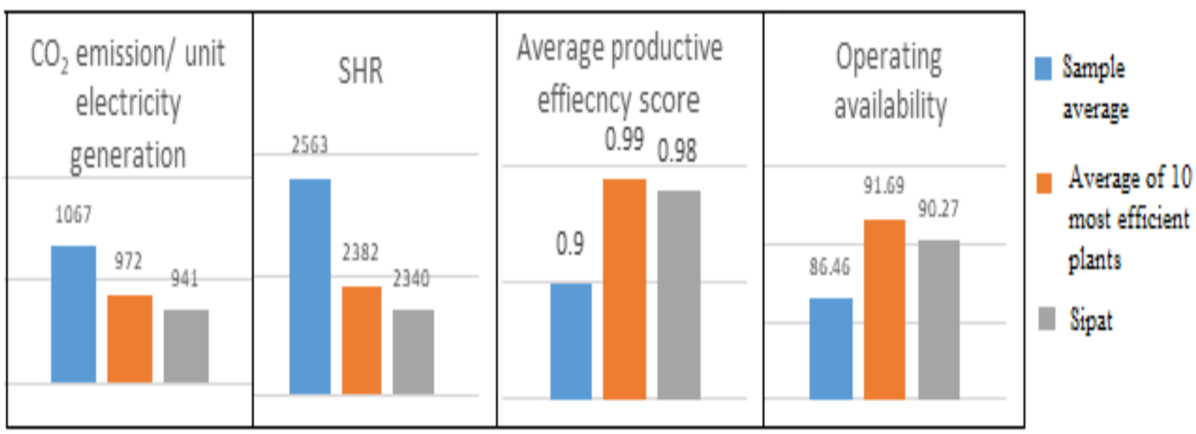

Figure 5

The only plant in our dataset which contains supercritical units is Sipat. Three out of five units of Sipat operate with supercritical technology. Sipat is one of the best performers with an average productive efficiency score equal to 0.98. Figure 5 maps the performance of Sipat against the average performance of all plants in our sample as well as average of 10 most efficient plants. As anticipated, Sipat performs better in terms of SHR not only compared to the average of all plants but also compared to the average of the most efficient plants. Besides SHR, Sipat also has the lowest CO_2 emissions per unit of net electricity generated in the same comparison. It fares slightly behind the 10 most efficient plants in terms of operating availability, a managerial input.

Adopting supercritical technologies is important for improving the environmentally efficiency of the thermal power sector. A supercritical plant generates 13 percent less emissions than a subcritical plant.[39] Currently, supercritical units account for 11 percent of the total installed capacity. In comparison, China has 90 percent of the top 100 most efficient plants running on ultra-supercritical technology.[40]

6.2.5 Company-level analysis.

The 48 power plants in our sample belong to 16 power generating companies. Many of the state-owned power generating companies are loss making and are facing severe liquidity problems. Focus on efficient power generation can partially help alleviate this problem. A firm level assessment of the efficiency of the power generation can provide a mechanism to benchmark and compare across various companies in the thermal power sector. Table 11 gives the company-wise installed capacity of the coal-fired thermal power plants considered in the sample and their average efficiency scores.[41]

Table 11: Company-wise average productive efficiency scores.

Company	Sector	Number of plants	Total capacity (MWh)	Average productive efficiency
REL	Private	1	500	0.99
NTPC	Centre	12	29750	0.98
APGENCO	State	4	5030	0.94
CESC	Private	1	750	0.94
RRVUNL	State	3	3740	0.89
OPGC	State	1	420	0.89
KPCL	State	1	1720	0.86
TNEB	State	4	4770	0.85
PSEB	State	3	2620	0.85
CSEB	State	1	1340	0.85
MAHAGENCO	State	6	7800	0.83
GSECL	State	4	3690	0.83
DVC	Centre	3	2880	0.78
MPGPCL	State	2	2670	0.72
WBPDC	State	1	450	0.69
IPGPCL	State	1	135	0.63

[39] Coal in India, 2015, Report published by Office of Chief Economist, Department of Industry and Science, Australian Government.

[40] http://thediplomat.com/2017/07/chinas-clean-coal-power-viable-model-or-cautionary-tale/ Accessed on 15 July 2017.

[41] Company level efficiency score is the weighted average of individual plant level efficiency scores (weighted with plant?s installed capacity).

REL turns out to be the most efficient power generating company with an average efficiency of 0.99. It is closely followed by NTPC with a score of 0.98. APFENCO and CESC also perform better with scores of 0.94. IPGPCL, WBPDC and MPGPCL are the least efficient firms with scores ranging between 0.63 and 0.72. Table 12 shows that the high mean productive efficiency of REL is driven by its better performance in terms of SHR (high thermodynamic efficiency) and plant availability (better management). NTPC closely follows REL. The worst performing company, namely, IPGCL, has the highest SHR, low average unit capacity, as well as low operating availability.

Table 12: Input usage of power generating companies

Company	Sector	SHR (Kcal/KWh)	CO_2 emissions / net electricity generated Tonnes/GWh	Operating Availability (%)	Average unit capacity (MWh)
NTPC	Centre	2389.47	976.86	91	382.093565
MAHAGEN CO	State	2726.34	1155.29	82	246.245619
APGENCO	State	2422.48	1006.48	91	189.265909
TNEB	State	2687	1140	83	206.09923
RRVUNL	State	2547.5	1076.17	87	215.468673
GSECL	State	2684.15	1135.49	78	184.233555
DVC	Centre	2905.45	1240.44	70	177.454959
MPGPCL	State	3154.88	1325.86	84	201.681654
PSEB	State	2624.44	1109.61	88	195.284738
KPCL	State	2604.9	1089.6	85	212.028302
CSEB	State	2689.17	1097.39	87	225.64926
CESC	Private	2481.79	1025.84	95	250
REL	Private	2302.09	949.36	96	250
WBPDC	State	3028.42	1310.75	67	100.429799
OPGC	State	2432.42	1032.74	89	210
IPGPCL	State	3365.7	1444.89	77	67.5

With respect to ownership, both the privately-owned companies perform high on the efficiency. However, the performance of two centre-owned companies differs starkly. While NTPC is amongst the best performing utilities, DVC is amongst the worst ones. There is a considerable difference in the shareholding structure of the two companies. NTPC is a listed company with 70 percent shareholding with the government of India and the remaining with the foreign and domestic institutional investors, while DVC is owned by the government of India, government of West Bengal, and government of Jharkhand. One managerial practice that distinguishes NTPC from several other generating companies is the plant related performance pay which increases accountability and hence motivation to run a plant better. Better management practices directly feed into lower SHR and improved plant availability, which makes it one of the most efficient companies. Overall, most of the state owned companies fare poorly on efficiency. APGENCO is the only state-owned company that performs well. Looking at the SHR and operating availability under different ownership structures, we find that the SHR for private companies is lowest closely matched by that of centre-owned NTPC, while it is maximum for the state-owned plants. The distribution is similar for operating availability. This validates our assumption that the efficiency under any given ownership structure is driven by SHR and operating availability. Lower SHR also has one to one relation with CO_2 emissions per unit of net generation. Thus, private and central utilities emit substantially less CO_2 emissions per unit of net generation compared to most of the state utilities.

7 Conclusions.

This is a first study that adopts the by-production approach for measuring the productive and environment efficiency of Indian thermal power plants. The study captures key input and output data for 48 power plants from the year 2002-03 to 2014-15 and identifies the variables driving the efficiency of these plant. The results show that during this period, the percentage of plants that performed high on the productive efficiency frontier went up from 50 percent to 75 percent. The study highlights that efficiency of Indian thermal power plants depends predominantly on the SHR and operating availability. Other factors like plant's ownership, installed capacity and age impact the above two factors and through them the plant's efficiency. This reaffirms the importance of schemes like Perform Achieve Trade (PAT) which provide an incentive structure for plants to reduce their SHR. Moreover, SHR is the inverse of energy (thermodynamic) efficiency. Our study indicates that, for Indian coal-based thermal power plants, productive efficiency, as measured by economists, is strongly positively associated with the engineering concept of thermodynamic efficiency.

A company-level analysis suggests that most state owned power generating companies lag behind in SHR and operating availability compared to centre-owned and private-owned power generating companies. This calls for an increased focus on improving managerial and O&M practices of the state run plants. For example, even simple practices like soot blowing of boiler over time helps improve SHR, but the practice is usually not followed at most state run power generating companies.[42]

Going forward, the sector's efficiency is slated to increase with more number of supercritical power plants being installed. Further efficiency gains are possible through measures like shift

[42] http://www.cea.nic.in/reports/others/thermal/tpece/report_85_pul_coal.pdf.

towards ultra-supercritical technology, adoption of the international managerial and O&M best practices and better coal management to ensure its timely availability. Encouraging plants to adhere to the targets under PAT in its second phase would also play a key role in not only improving the productive efficiency but also lowering emissions. This is especially important given India's commitment to lower its CO_2 emissions.

While we could analyse productive efficiency of power plants using the by-production methodology, the deterministic linear nature of CO_2 calculation restricts us from analyzing the environment efficiency in two ways. First, since most of the plants operate very close to the environment frontier, MRL's output-based environmental efficiency index will take values very close to one for almost all power plants in the CEA data set. This implies that no significant improvements in environmental efficiency can be achieved by lowering the emission level of any plant holding its consumption of heat input fixed. Second, the only way to improve environment efficiency by reducing the emission level holding the heat input fixed would be to increase the proportion of oil in the fuel mix to obtain the same heat. This is not desirable in a coal-based thermal power plant. Thus, an output-based environmental efficiency index loses its importance as a measure of environmental (in)efficiency in the context of CEA-type datasets.

However, there is another index of efficiency which continues to remain relevant for measuring environmental efficiency even in the context of CEA-type datasets. The graph efficiency index (see e.g., Färe, Grosskopf, and Lovell (1985) and Murty and Russell (2017)) measures efficiency of a producing unit in both output and input dimensions. So, while there may not be much scope for reducing emission levels when the heat input level is held fixed, it is possible that a power plant may be using its heat input wastefully. Thus, it may be possible for it to cut back on its excessive usage of its heat input without decreasing electricity production and increasing emission generation. Scientists and engineers call the phenomenon of decreasing usage of heat input without decreasing electricity generation as improvements in thermodynamic efficiency of a power plant. In fact, while doing so, it may be possible for it to also reduce its emission level. This is because cutting down on usage of heat input would imply burning less physical units of the carbon containing fossil fuel. One of our ongoing research projects aims to compute such a measure of efficiency for the Indian thermal power plants using a by-production specification of the technology.

REFERENCES

Australian Government (2015), "Coal in India," *Office of Chief Economist, Department of Industry and Science.*

Baumgärtner, S. (2012), *Ambivalent Joint Production and the Natural Environment: An Economic and thermodynamic Analysis*, Springer Science & Business Media.

Baumgärtner, S., and J. de Swaan Arons (2003), "Necessity and Inefficiency in the Generation of Waste," *Journal of Industrial Ecology*, 7: 113–123.

Baumol, W. J., and W. E. Oates (1975, 1988), *The Theory of Environmental Policy*, 1st and 2nd editions, Cambridge University Press.

Behera, S.K., A.P. Dash and J. A. Farooquie (2010), "Performance Analysis of Coal Fired Power Plants in India," *Proceedings of the 2010 International Conference on Industrial Engineering and Operations Management Dhaka*, Bangladesh.

Chan, H. S. R., M. L. Cropper, and K. Malik (2014), "Why are power plants in India less efficient than power plants in the United States?,"*American Economic Review*,104(5), 586–90

Coal Industry Advisory Board (CIAB) (2010) "Power Generation from Coal Measuring and Reporting Efficiency Performance and CO2 Emissions," *International Energy Agency (IEA*

Coggins, J. S., and J. R. Swinton (1996), "The Price of Pollution: A Dual Approach to Valuing SO_2 Allowances," *Journal of Environmental Economics and Management*, 30: 58–72.

Cropper, M. L., and W. E. Oates (1992), "Environmental Economics: A Survey," *Journal of Economic Literature*, 30: 675–740.

R. Färe, R., S. Grosskopf, and C.A. Knox Lovell (1985) *The Measurement of Efficiency of Production*, Kluwer-Nijhoff, Boston.

Färe, R., S. Grosskopf, C. A. K. Lovell, and C. Pasurka (1989), "Multilateral Productivity Comparisons When Some Outputs are Undesirable: A Nonparametric Approach," *The Review of Economics and Statistics*, 71: 90–98.

Färe, R., and C.A.K. Lovell (1978) "Measuring the technical efficiency of production," *Journal of Economic Theory* 19, 150–162.

Førsund, F. (1998), "Pollution Modeling and Multiple-Output Production Theory," Discussion Paper #D-37/1998, Department of Economics and Social Sciences, Agricultural University of Norway.

Førsund, F. (2009), "Good Modelling of Bad Outputs: Pollution and Multiple-Output Production," *International Review of Environmental and Resource Economics*, 3, 1–38.

Førsund, F. (2017), "Multi-equation Modeling of Desirable and Undesirable Outputs Satisfying the Material Balance," *Empirical Economics*, online.

Hailu, A., and T. S. Veeman (2000), "Environmentally Sensitive Productive Analysis of the Canadian Pulp and Paper Industry, 1959–1994: an input distance function approach," *Journal of Environmental Economics and Management*, 40, 251–274

Hiebert, L. D. (2002), "The determinants of the cost efficiency of electric generating plants: a stochastic frontier approach,"*Southern Economic Journal*, 68:4, 935-946.

Khanna, M., K. Mundra, and A. Ullah (1999) "Parametric and Semi-parametric Estimation of the Effect of Firm Attributes on Efficiency: the Electricity Generating Industry in India," *Journal of International Trade & Economic Development*, 8, 419–436

Kumar, S., and S. Managi (2009), *The economics of sustainable development: The case of India*,Volume 32, Springer Science & Business Media.

Kumar, R., A. K. Sharma, and P. C. Tewari (2015), "Cost analysis of a coal-fired power plant using the NPV method," *Journal of Industrial Engineering International*, DOI 10.1007/s40092-015-0116-8

Lee, J. D., J. B. Park, T. Y. Kim (2002), "Estimation of the shadow prices of pollutants with production/environment inefficiency taken into account: A nonparametric directional distance function approach," *Journal of Environmental Management*, 63, 365-375

Ministry of Power (2010), Government of India, Ministry of Power, Central Electricity Authority, "Mapping of 85 pulverized coal fired thermal power generating units in different states," under Indo-German Energy Program (IGEN).

Ministry of Power (2002-03, 2003-04, 2004-05, 2005-06, 2006-07, 2007-08, 2008-09, 2009-10, 2010-11, 2012-13, 2013-14, 2014-15), *Review of Performance of Thermal Power Plants in India*, Central Electricity Authority, Ministry of Power, Government of India.

Murty, M. N., and S. Kumar (2002), "Measuring Cost of Environmentally Sustainable Industrial Development in India: A Distance Function Approach," *Environment and Development Economics*, 7: 467–86.

Murty, M. N., and S. Kumar (2003), "Win-Win Opportunities and Environmental Regulation: Testing of Porter Hypothesis for Indian Manufacturing Industries," *Journal of Environmental Management*, 67: 139–44.

Murty, M. N., S. Kumar, K. Dhawala (2007), "Measuring environmental efficiency of industry: A case study of thermal power generation in India," *Environmental and Resource Economics*, 38: 31–50.

Murty, S (2015), "On the Properties of an Emission-Generating Technology and its Parametric Representation," *Economic Theory*. 60, 243–282.

Murty, S, and R. R. Russell (2002), "On Modeling Pollution-Generating Technologies," Department of Economics, University of California, Riverside, Discussion Papers Series, No. 02-14.

Murty, S., R. R. Russell, and S. B. Levkoff (2012), "On Modeling Pollution-Generating Technologies," *Journal of Environmental Economics and Management*, 64, 117–135.

Murty, S., and R. R. Russell (2016), "Modeling Emission-Generating Technologies: Reconciliation of Axiomatic and By-Production Approaches," *Empirical Economics*, online: DOI 10.1007/s00181-016-1183-4.

Murty, S., and R. R. Russell (2017) "Bad outputs," *Discussion Papers in Economics* No. 17-06, Centre for International Trade and Development, School of International Studies, Jawaharlal Nehru University, India.

Narayana, B. H. and M. S. Bhatt (2004), "Minimisation of auxiliary power consumption in coal powered thermal power stations," Paper presented at the National Power Systems Conference 2004, Indian Institute of Technology, Kanpur.

Pasurka, C. A. (2006) "Decomposing electric power plant emissions within a joint production framework," *Energy Economics* 28 26–43.

Ray, S. C., K. Mukherjee, and A. Venkatesh (2017), "Nonparametric Measures of Efficiency in the Presence of Undesirable Outputs: A By-production Approach with Weak Disposability," *Empirical Economics*, online.

Reinhard, S, C. A. K. Lovell, and G. J. Thijssen, (1999), "Econometric Estimation of Technical and Environmental Efficiency: An Application to Dutch Dairy Farms," *American Journal of Agricultural Economics*, 81: 44–60.

Reinhard, Stijn, C. A. Knox Lovell, and Geert J. Thijssen (2000) "Environmental Efficiency with Multiple Environmentally Detrimental Variables; Estimated with SFA and DEA." *European Journal of Operational Research*, 121: 287–303.

Sahoo, N. R., Pratap K.J. Mohapatra, B. K. Sahoo, and B. Mahanty (2017) "Rationality of Energy Efficiency Improvement Targets under the PAT Scheme in India – A Case of Thermal Power Plants," *Energy Economics* doi:10.1016/j.eneco.2017.06.004

Serra, T., R. G. Chambers, and A. O. Lansink (2016), "Measuring Technical and Environmental Efficiency in a State-Contingent Techology," *European Journal of Operational Research*, 236: 706–717.

See, K. F. and T. Coelli (2012), "An analysis of factors that influence the technical efficiency of Malaysian thermal power plants,"*Energy Economics*,34.3: 677-685.

Schweinfurth, S. P., (2009) "An Introduction to Coal Quality," Chapter C, *The National Coal Resource Assessment Overview*, Eds: Brenda S. Pierce and Kristin O. Dennen, U.S. Geological Survey Professional Paper 1625?F, U.S. Department of the Interior, U.S. Geological Survey, Reston, Virginia

Sueyoshi, T., and M. Goto (2010), "Should the US Clean Air Act Include CO_2 Emission Control?: Examination by Data Envelopment Analysis," *Energy Policy*, 38, 5902–5911

Sueyoshi, T., and M. Goto (2011), "DEA approach for unified efficiency measurement: Assessment of Japanese fossil fuel power generation," *Energy Economics*, 33, 292–303

Sueyoshi, T., and M. Goto (2012), "DEA environmental assessment of coal fired power plants: Methodological comparison between radial and non-radial models," *EnergyEconomics*, 34, 1854–1863

Wenyuan, Wu (2017), "China's Clean Coal Power: Viable Model or Cautionary Tale?," *The Diplomat*, online: $http://thediplomat.com/2017/07/chinas-clean-coal-power-viable-model-or-cautionary-tale/$

Zhou P., B. Ang, and K-L Poh (2008), "A Survey of Data Envelopment Analysis in Energy and Environmental Studies," *European Journal of Operational Research*, 189: 1–18.

APPENDIX

TABLE A

PLANT NAME	STATE	ORGANISATION	SECTOR
BANDEL TPS	WEST BENGAL	STATE	WBPDC
BHUSAWAL TPS	MAHARASHTRA	STATE	MAHAGENCO
BOKARO B TPS	JHARKHAND	CENTER	DVC
BUDGE BUDGE TPS	WEST BENGAL	PVT	CESC
CHANDRAPUR (DVC)TPS	JHARKHAND	CENTER	DVC
CHANDRAPUR(MAHARASHTRA) STPS	MAHARASHTRA	STATE	MAHAGENCO
CHHABRA TPS	RAJASTHAN	STATE	RRVUNL
DADRI (NCTPP)	UTTAR PRADESH	CENTER	NTPC
DAHANU TPS	MAHARASHTRA	PVT	REL
DR.N.TATA RAO TPS	ANDHRA PRADESH	STATE	APGENCO
ENNORE TPS	TAMIL NADU	STATE	TNEB
FARAKKA STPS	WEST BENGAL	CENTER	NTPC
GANDHI NAGAR TPS	GUJARAT	STATE	GSECL
GH TPS (LEH.MOH)	PUNJAB	STATE	PSEB
GND TPS (BHATINDA)	PUNJAB	STATE	PSEB
IB VALLEY TPS	ORISSA	STATE	OPGC
KAHALGAON TPS	BIHAR	CENTER	NTPC
KAKATIYA TPS	ANDHRA PRADESH	STATE	APGENCO
KHAPARKHEDA TPS	MAHARASHTRA	STATE	MAHAGENCO
KORADI TPS	MAHARASHTRA	STATE	MAHAGENCO
KORBA STPS	CHATTISGARH	CENTER	NTPC
KORBA WEST TSP	CHATTISGARH	STATE	CSEB
KOTA TPS	RAJASTHAN	STATE	RRVUNL
KOTHAGUDEM TPS	ANDHRA PRADESH	STATE	APGENCO
METUR TPS	TAMIL NADU	STATE	TNEB
NASIK TPS	MAHARASHTRA	STATE	MAHAGENCO
NORTH CHENNAI TPS	TAMIL NADU	STATE	TNEB
PANITPAT TIPS	BIHAR	CENTER	DVC
PARLI TPS	MAHARASHTRA	STATE	MAHAGENCO
RAICHUR TPS	KARNATAKA	STATE	KPCL
RAJGHAT TPS	DELHI	STATE	IPGPCL
RAMAGUNDEM STPS	ANDHRA PRADESH	CENTER	NTPC
RAYALASEEMA TPS	ANDHRA PRADESH	STATE	APGENCO
RIHAND STPS	UTTAR PRADESH	CENTER	NTPC
ROPAR TPS	PUNJAB	STATE	PSEB
SANJAY GANDHI TPS	MADHYA PRADESH	STATE	MPGPCL
SATPURA TPS	MADHYA PRADESH	STATE	MPGPCL
SIKKA REP. TPS	GUJARAT	STATE	GSECL
SIMHADRI	ANDHRA PRADESH	CENTER	NTPC
SINGRAULI STPS	UTTAR PRADESH	CENTER	NTPC
SIPAT STPS	CHATTISGARH	CENTER	NTPC
SURATGARH TPS	RAJASTHAN	STATE	RRVUNL
TALCHER STPS	ORISSA	CENTER	NTPC
TUTICORIN TPS	TAMIL NADU	STATE	TNEB
UKAI TPS	GUJARAT	STATE	GSECL
UNCHAHAR TPS	UTTAR PRADESH	CENTER	NTPC
VINDHAYACHAL STPS	MADHYA PRADESH	CENTER	NTPC
WANAKBORI TPS	GUJARAT	STATE	GSECL