

Puggioni, Daniela; Stefanou, Spiro E.

Working Paper

The value of being socially responsible: A primal-dual approach

Working Papers, No. 2018-12

Provided in Cooperation with:

Bank of Mexico, Mexico City

Suggested Citation: Puggioni, Daniela; Stefanou, Spiro E. (2018) : The value of being socially responsible: A primal-dual approach, Working Papers, No. 2018-12, Banco de México, Ciudad de México

This Version is available at:

<https://hdl.handle.net/10419/202478>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Banco de México
Documentos de Investigación

Banco de México
Working Papers

N° 2018-12

The Value of Being Socially Responsible. A Primal-
Dual Approach

Daniela Puggioni
Banco de México

Spiro E. Stefanou
University of Florida, Wageningen University

August 2018

La serie de Documentos de Investigación del Banco de México divulga resultados preliminares de trabajos de investigación económica realizados en el Banco de México con la finalidad de propiciar el intercambio y debate de ideas. El contenido de los Documentos de Investigación, así como las conclusiones que de ellos se derivan, son responsabilidad exclusiva de los autores y no reflejan necesariamente las del Banco de México.

The Working Papers series of Banco de México disseminates preliminary results of economic research conducted at Banco de México in order to promote the exchange and debate of ideas. The views and conclusions presented in the Working Papers are exclusively the responsibility of the authors and do not necessarily reflect those of Banco de México.

The Value of Being Socially Responsible. A Primal-Dual Approach*

Daniela Puggioni[†]
Banco de México

Spiro E. Stefanou[‡]
University of Florida, Wageningen University

Abstract: This paper models Corporate Social Responsibility (CSR) as one of the outputs that results from a firm's decisions regarding what and how to produce. The framework developed allows for studying technical efficiency and deriving a system of internal shadow prices to quantify the value of implementing socially responsible activities. The empirical application focuses on the food and beverage manufacturing sector worldwide. The results indicate high levels of technical efficiency in this sector and document a positive average shadow price of CSR activities, implying that the net value of implementing this kind of activities is positive to the firm as their benefit exceeds the cost. In particular, it is shown that increasing the CSR engagement at the margin positively contributes to the creation of firm value, while reducing it has a negative marginal impact.

Keywords: Productivity and competitiveness; Decision processes; Non-parametric technology; Shadow prices; Corporate Social Responsibility.

JEL Classification: C14, D24, M14, M20.

Resumen: Este documento modela la Responsabilidad Social Empresarial (RSE) como uno de los productos que resultan de las decisiones de la empresa acerca de qué y cómo producir. El marco desarrollado permite estudiar la eficiencia técnica y derivar un sistema de precios sombra internos para cuantificar el valor de la implementación de actividades socialmente responsables. La aplicación empírica se centra en el sector de fabricación de alimentos y bebidas a nivel mundial. Los resultados indican altos niveles de eficiencia técnica en ese sector y documentan un precio sombra promedio positivo de las actividades de RSE, lo que implica que el valor neto de la implementación de este tipo de actividades es positivo para la empresa, toda vez que su beneficio supera su costo. En particular, se muestra que aumentar el compromiso a la RSE en el margen contribuye positivamente a la creación de valor de las empresas, mientras que disminuirlo tiene un impacto marginal negativo.

Palabras Clave: Productividad y competitividad; Procesos de decisión; Tecnología no paramétrica; Precios sombra; Responsabilidad Social Empresarial.

*We are particularly thankful to Bob Chambers for his precious suggestions and very constructive comments. We are also grateful to the analysts at Sustainalytics in New York City for assisting with obtaining the data and providing helpful insights on the details of their data collection process and methodology. Financial support for this research was partially provided through a post-doctoral fellowship by the Food and Resource Economics Department at the University of Florida.

[†] Dirección General de Investigación Económica, Banco de México. Email: dpuggionih@banxico.org.mx.

[‡] Food and Resource Economics Department, University of Florida, Florida, USA. Business Economics Group, Wageningen University, Wageningen, The Netherlands. Email: sstefanou@ufl.edu.

1 Introduction

Corporate social responsibility (CSR) has experienced a rather unique and difficult path toward becoming a mainstream subject of academic interest¹ likely because there is a certain degree of ambiguity and disagreement with respect to its definition and the nature of the core principles that identify CSR. According to the World Business Council for Sustainable Development,² “*Corporate Social Responsibility is the continuing commitment by business to contribute to economic development while improving the quality of life of the workforce and their families as well as of the community and society at large*”. While CSR seeks to create long-term economic value, as any conventional business activity, this definition highlights the fact that its scope is much broader as it calls for a more comprehensive commitment of the firm to society. The advancement of CSR to a core management or board-level function together with the growing attention toward CSR among civil society and government has fostered a dramatic increase in CSR implementation and reporting, thus providing access to CSR data and allowing for a shift of the research focus from a normative and theoretical framework to a more applied one. Nonetheless, the literature on CSR, so far, has been mainly restricted to the business management and financial fields and only in the 2000s has this literature started to build a direct connection between CSR and the economic concept of profit maximization.

Baron (2001) and McWilliams and Siegel (2001) were the first to explicitly model ‘strategic’ and ‘profit-maximizing’ CSR suggesting that firms undertake CSR activities expecting a net benefit from them.³ Both contributions emphasize that CSR is a way for firms to compete for socially responsible consumers by either linking their social contribution to product sales or adding social attributes and features to their products. A key implication of this perspective on CSR is that it represents a product differentiation strategy to gain competitive advantage as argued by Bagnoli and Watts (2003) and Siegel and Vitaliano (2007). In addition, there is empirical evidence in the literature supporting the conjecture that CSR practices have an influence on consumers’ purchasing intentions and willingness to pay and that increasing customer awareness of these practices can have a positive impact on firm value (see Creyer (1997), Auger, Burke, Devinney, and Louviere (2003), De Pelsmacker, Driesen, and Rayp (2005), Ailawadi, Neslin, Luan, and Taylor (2014), or Servaes and Tamayo (2013)).

CSR has been analyzed also in the theoretical finance literature with models assuming the

¹For a historical overview and a review of the academic research that has both prompted, and responded to, issues of corporate social responsibility see Crane, McWilliams, Matten, Moon, and Siegel (2008).

²WBCSD (1998) Stakeholder Dialogue on CSR, The Netherlands, Sept 6-8, 1998.

³For a collection of empirical contributions on the relationship between CSR and economic performance see Paul et al. (2006).

existence of a class of investors who prefer to invest in CSR stocks, but the impact of their preferences on the value of the stocks is not always clear. Empirical evidence, on the other hand, seems to identify a negative association between CSR and systemic risk and cost of equity capital, and a positive association between CSR and firm value and shareholder wealth. Thus, higher levels of CSR (or higher CSR scores) imply a lower systemic risk, a lower cost of capital, a higher firm value, and higher stocks valuations (see Sharfman and Fernando (2008), Oikonomou, Brooks, and Pavelin (2012), Albuquerque, Durnev, and Koskinen (2015), Margolis, Elfenbein, and Walsh (2009), Galema, Plantinga, and Scholtens (2008), Dimson, Karakaş, and Li (2015), or Flammer (2015)). Eccles, Ioannou, and Serafeim (2014) further document that the positive impact of a higher engagement in corporate sustainability is not limited to the stock market and accounting performance but extends to organizational processes and governance.

All these contributions, in different ways and with different emphases, try to identify and quantify to what extent the pursuit of a more advanced and comprehensive social agenda, that may go beyond short-run profits or mandated minimum standards, impacts the economic value of the firm. In fact, they provide formal and rigorous support to a conviction that has already reached a wide consensus among firms, consumers, and policy makers; i.e. that CSR should be a prominent business practice. In spite of the efforts spent in corroborating the strategic role of CSR, practically no work has been done to formalize and explain the process through which CSR is created at the firm level. Simply put, the existing literature justifies CSR practices with economic, managerial, or financial motivations trying to provide some insight into whether CSR is good for business, in general, and for what it might be good, in particular. However, considerations regarding the way CSR is produced and what implications CSR production has on a firm's dynamics, such as the technology underlying CSR production, the impact of CSR on the other outputs produced, or the kind of inputs needed to produce CSR are completely ignored. This may be due to CSR being a relatively new and still unfolding concept for both firms and scholars that is not easy to precisely define, model, measure, and quantify.

As the existing research has been almost exclusively focused on *why* CSR is done, this study takes a very different perspective and attempts to shed light on *how* CSR is done and *how much* it is worthy for the firm by incorporating it into a formal production framework. This seems especially appropriate as a considerable share of CSR activities is directly related to production. Our approach is based on the fundamental economic notion of opportunity cost. Since firms have to operate under scarcity, it is crucial to recognize that the decision of allocating resources to CSR activities necessarily implies taking away resources from other

activities. Hence, the value of being socially responsible needs to be weighed against its cost. The framework we develop takes into account these trade-offs in a rigorous yet flexible way consistent with economic theory.

We adopt a joint production model for characterizing the technology and representing the transformation process of multiple inputs into multiple outputs. The types of different inputs and outputs involved in the production process are a distinctive feature of the model. In particular, each firm is assumed to produce a desirable, marketed output but, because the production of this desirable output may require the use of some undesirable input, an undesirable output can be generated along with the desirable one. Thus, the firm needs to engage in socially responsible activities, namely CSR, which is an additional output produced to mitigate the unwanted output. Common inputs as well as ‘socially responsible’ inputs are used to produce CSR. The joint technology is obtained as a composition of two separate sub-technologies: one describing the production of the desirable output and the other describing the generation of the undesirable output. CSR is the link between these two technologies as it simultaneously represents the opportunity cost of producing socially responsible activities in terms of desirable output and its mitigating effect with respect to undesirable output.

The empirical application focuses on the food and beverage manufacturing sector that is particularly interesting since it faces specific CSR challenges, such as food safety controversies, demand for healthier food products, responsible sourcing of raw materials, along with more common CSR issues such as water and energy efficiency, supply chain management, labor standards and safety in the workplace. This rich and diversified array of activities allows for observing and exploring multiple dimensions of CSR engagement making the food sector an insightful example of how our theoretical and empirical frameworks can be employed to analyze the impact and the value of CSR. The implementation of the analysis relies on a parsimonious non-parametric approach known as Data Envelopment Analysis (DEA)⁴ which allows for constructing the joint technology as the intersection of the desirable production and the undesirable production sub-technologies. Once the technology is fully characterized, a measure of technical inefficiency as well as a set of internal shadow values for inputs and outputs can be obtained for each firm. Moreover, because DEA technologies are not amenable to standard differential calculus arguments, at least for some firms, recent developments by Chambers and Färe (2008) are used to derive the shadow value of CSR as a measure of willingness to gain for producing one more unit of CSR and willingness to lose for relinquishing the production of one unit of CSR.

⁴For an in-depth review of DEA theory and empirical applications see Charnes, Cooper, Lewin, and Seiford (1995) and Cooper, Seiford, and Zhu (2011).

Our contribution offers several interesting managerial insights. Specifically, given that CSR is a non-marketed activity there is not a market-based measure of competitiveness in CSR. With the framework proposed in this study firms are able to benchmark competitors within the boundaries of the technology. Also, even though several contributions in the literature have tried to motivate the role of CSR as a strategic tool to attract investors/consumers/workers, there is no consensus on whether CSR is indeed effective in doing so. That is, it is challenging to determine the value of CSR. Our methodology delivers a specific measure of the internal value of implementing CSR activities. Finally, it is important for managers to understand the benefit/cost of any activity at the margin to evaluate the worthiness and the risk of increasing/decreasing the effort and resources allocated to this activity. We are able to quantify the marginal benefit/cost of CSR activities thus providing valuable information to guide firms' decisions on the extent of their CSR engagement.

The results of the analysis provide valuable indications regarding the impact of CSR on firm's performance and value. In particular, our findings suggest that it is important to account for the complexity of the production process and that ranking firms in terms of their performance and identifying leaders can be quite challenging when they engage in very differentiated activities. Nonetheless, a diversified portfolio that includes activities such as CSR can also represent an opportunity for firms to compete and excel on new grounds. In addition we find that, on average, being good is worthy, i.e. socially responsible engagement positively contributes to adding (internal) value to the firm. The positive value of CSR is emphasized also at the margin where we observe firms willing to do more CSR for a positive price but willing to be compensated for doing less.

The remainder of the paper is organized as follows. Section 2 presents a multiple inputs and outputs theoretical framework where CSR is explicitly modeled as an output of production. Section 3 illustrates the empirical approach and explains how DEA methods can be used to characterize the technology, analyze technical efficiency and derive a set of internal values of inputs and outputs to quantify the value of socially responsible activities overall and at the margin. Section 4 describes the data used to carry out the empirical analysis. In Section 5 the results of the analysis are presented and discussed. Section 6 concludes.

2 A Multi-Output Model of Corporate Social Responsibility

Multiple outputs are the rule rather than the exception at the micro level of production. This is because the same input, or set of inputs, can be employed to produce different outputs

and because there are many instances of jointness in production that can reach the extreme form of different outputs needed to be produced in fixed proportions. To acknowledge these instances, we model CSR incorporating it into a multiple input multiple output technology framework and assuming that CSR is an additional output in the production process. Our choice is motivated by the fact that CSR activities are not freely available nor easily acquired on the market like standard inputs, therefore their implementation requires firms to allocate resources (such as capital, labor and materials) to the production of CSR effectively diverting them from the production of other outputs. The fact that CSR needs to be actually generated at an opportunity cost makes it more characteristic of an output rather than an input. In addition, because of the nature of socially responsible activities usually carried out by firms (i.e. environmental programs, sustainability programs, community programs), CSR embodies the notion of mitigation particularly well. Much like abatement is implemented to clean up pollution, CSR can be implemented to improve a dirty production processes, to support the sourcing and use of sustainable inputs, or to establish a good reputation among consumers and community. Hence, building on Murty, Russell, and Levkoff (2012), our model regards CSR as an output produced to mitigate the negative effects of another undesirable output, i.e. an output that is unwanted but inevitably generated within the production process and can potentially be detrimental to the firm.⁵

Consider a joint production technology in which N^6 inputs x_1 , x_2 and x_3 are utilized to produce M^7 outputs that can be categorized as desirable output(s) y_D , undesirable output(s) y_U , and the socially responsible output(s) y_R . More specifically, y_D is the primary, marketed output for which the production process is set up, y_U is a by-product generated during the production of the desirable output, and y_R consists of socially responsible activities implemented to reduce the undesirable output. The production process of the desirable output y_D requires inputs x_1 , x_2 , and x_3 , where x_1 is a standard input used in the production of y_D , x_2 is an undesirable input that leads to the generation of the by-product, i.e. the undesirable output y_U , and x_3 is a socially responsible input that, while used to produce y_D , specifically contributes to the production of the socially responsible output y_R as well. Because of the nature of this joint technology, a firm that aims at being socially responsible needs to engage in the production of y_R to mitigate the undesirable output. Nonetheless, producing y_R is costly, meaning that a firm has to divert resources away from the desirable output production

⁵For a comprehensive literature review on modeling technologies in the presence of undesirable (bad) inputs and outputs see Førsund (2009).

⁶Where N is a $(N_1 + N_2 + N_3)$ -dimensional vector of inputs.

⁷Where M is a $(M_D + M_U + M_R)$ -dimensional vector of outputs.

to generate socially responsible activities.

To fix ideas consider the example of a confectionery firm producing chocolate bars as its main, desirable, marketed output y_D . To produce chocolate bars, this firm utilizes conventional inputs of production x_1 : capital, labor, raw materials; undesirable inputs x_2 : polluting energy sources, unsafe machinery, unethically sourced (e.g. through child labor) cocoa beans; as well as socially responsible inputs x_3 : certified fair trade and sustainable cocoa beans, renewable energy sources. Since the production of chocolate bars requires the use of some undesirable inputs, undesirable/bad outputs y_U such as pollution, waste, worker injuries, negative publicity in the media are also generated. To mitigate and limit the negative impact of these undesirable outputs the firm implements socially responsible activities y_R in the form of programs to reduce energy and water use, pollution, waste, programs to improve the safety on the workplace, programs to source inputs ethically.

The *by-production*⁸, or joint technology T can be characterized as the intersection of two different technologies T_1 and T_2 , so that

$$T = T_1 \cap T_2 \quad (2.1)$$

$$T_1 = \{ \langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in \mathbb{R}_+^{M+N} \mid f(y_D, y_R, x_1, x_2, x_3) \leq 0 \} \quad (2.2)$$

$$T_2 = \{ \langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in \mathbb{R}_+^{M+N} \mid y_U \geq q(y_R, x_2) \} \quad (2.3)$$

with f and q being continuously differentiable functions. The set T_1 is a conventional convex technology set representing the transformation process of the inputs into the desirable output and the socially responsible output. Assuming that f satisfies

$$f_{y_D}(y_D, y_R, x_1, x_2, x_3) \geq 0 \quad (2.4)$$

$$f_{y_R}(y_D, y_R, x_1, x_2, x_3) \geq 0 \quad (2.5)$$

$$f_{x_n}(y_D, y_R, x_1, x_2, x_3) \leq 0 \quad \text{for } n = 1, 2, 3 \quad (2.6)$$

then the technology T_1 displays the standard free disposability properties in desirable output,

⁸By-production is a term specifically coined by Murty (2010) to describe a technology in which the desirable production process generates unwanted residuals, or by-products.

socially responsible output, and inputs, respectively, i.e.

$$\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T_1, \bar{y}_D \leq y_D \Rightarrow \langle \bar{y}_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T_1 \quad (2.7)$$

$$\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T_1, \bar{y}_R \leq y_R \Rightarrow \langle y_D, y_U, \bar{y}_R, x_1, x_2, x_3 \rangle \in T_1 \quad (2.8)$$

$$\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T_1, \bar{x}_1 \geq x_1 \Rightarrow \langle y_D, y_U, y_R, \bar{x}_1, x_2, x_3 \rangle \in T_1 \quad (2.9)$$

$$\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T_1, \bar{x}_2 \geq x_2 \Rightarrow \langle y_D, y_U, y_R, x_1, \bar{x}_2, x_3 \rangle \in T_1 \quad (2.10)$$

$$\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T_1, \bar{x}_3 \geq x_3 \Rightarrow \langle y_D, y_U, y_R, x_1, x_2, \bar{x}_3 \rangle \in T_1 \quad (2.11)$$

For simplicity it is further assumed that the technology T_1 is independent of y_U , which implies that the production of the undesirable output does not have any direct effect on the production of the desirable output y_D .

The set T_2 is also convex and represents the undesirable output generating process. Assuming that q satisfies

$$q_{y_R}(y_R, x_2) < 0 \quad (2.12)$$

$$q_{x_2}(y_R, x_2) > 0 \quad (2.13)$$

and given the definition of T_2 in (2.3), the following properties hold

$$\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T_2, \bar{y}_U \geq y_U \Rightarrow \langle y_D, \bar{y}_U, y_R, x_1, x_2, x_3 \rangle \in T_2 \quad (2.14)$$

$$\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T_2, \bar{y}_R \geq y_R \Rightarrow \langle y_D, y_U, \bar{y}_R, x_1, x_2, x_3 \rangle \in T_2 \quad (2.15)$$

$$\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T_2, \bar{x}_2 \leq x_2 \Rightarrow \langle y_D, y_U, y_R, x_1, \bar{x}_2, x_3 \rangle \in T_2 \quad (2.16)$$

These properties are sometimes referred to costly disposability of undesirable output, socially responsible output, and by-product generating input, respectively. The relations expressed in (2.14), (2.15), and (2.16) describe the fact that the undesirable output y_U is a by-product of the production process whose disposability is not free. As shown in (2.12) and (2.15), the trade-off between the undesirable output and the socially responsible output y_R is negative, capturing the mitigating effect that socially responsible activities have on the production of the undesirable output. On the other hand, as shown in (2.13) and (2.16), the trade-off between the undesirable output y_U and the input x_2 , which is responsible for the generation of y_U , is non-negative, capturing the fact that the optimal level of undesirable output is increasing in the use of the by-product generating input x_2 .

Given the properties of T_1 and T_2 derived above, it is possible to rationalize the properties

of the joint technology T . Specifically, T is convex and satisfies free disposability of the desirable output and the first and third input because T_1 satisfies the same condition with respect to y_D , x_1 and x_3 and T_2 does not impose any restrictions on them. However, T violates free disposability of the socially responsible output and of the second input because, while T_1 satisfies the free disposability condition with respect to y_R and x_2 , T_2 violates free disposability with respect to y_R and imposes a restriction on x_2 that is, in fact, the exact opposite of free disposability. Finally, T displays costly disposability with respect to the undesirable output because T_1 does not impose any restriction on y_U while T_2 implies that y_U can be mitigated only through a decrease in the use of by-product generating input x_2 , or the costly implementation of the socially responsible activities y_R .

The weakly efficient points of T are defined as the quantity vectors $\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T$ satisfying $f(y_D, y_R, x_1, x_2, x_3) = 0$ and $y_U = q(y_R, x_2)$. This is because, if a quantity vector $\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T$ satisfies $f(y_D, y_R, x_1, x_2, x_3) < 0$, it is technologically feasible to decrease the usage of the inputs x_1 and x_3 without affecting the production levels of the desirable output y_D and the usage of x_2 . Therefore such a vector cannot be efficient. Similarly, a quantity vector $\langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in T$ satisfying $y_U > q(y_R, x_2)$ cannot be efficient because it is technologically feasible to decrease the production level of the undesirable output y_U without modifying the usage of the inputs and the production level of the desirable output y_D .

Consider the quantity vector $\langle \hat{y}_D, \hat{y}_U, \hat{y}_R, \hat{x}_1, \hat{x}_2, \hat{x}_3 \rangle$ which is a weakly efficient point of T since it satisfies $f(\hat{y}_D, \hat{y}_R, \hat{x}_1, \hat{x}_2, \hat{x}_3) = 0$ and $y_U - q(\hat{y}_R, \hat{x}_2) = 0$. Let $f_{y_D}(\hat{y}_D, \hat{y}_R, \hat{x}_1, \hat{x}_2, \hat{x}_3) \neq 0$ and $q_{y_R}(\hat{y}_R, \hat{x}_2) \neq 0$, then the matrix

$$\begin{bmatrix} f_{y_D}(\hat{y}_D, \hat{y}_R, \hat{x}_1, \hat{x}_2, \hat{x}_3) & f_{y_R}(\hat{y}_D, \hat{y}_R, \hat{x}_1, \hat{x}_2, \hat{x}_3) \\ 0 & -q_{y_R}(\hat{y}_R, \hat{x}_2) \end{bmatrix} \quad (2.17)$$

has full rank and, by the implicit function theorem, there exists a neighborhood V around $\langle \hat{y}_U, \hat{x}_1, \hat{x}_2, \hat{x}_3 \rangle \in \mathbb{R}_+^{N+M_U}$, a neighborhood W around $\langle \hat{y}_D, \hat{y}_R \rangle \in \mathbb{R}_+^{M_D+M_R}$ and continuously differentiable mappings $\psi : V \rightarrow \psi(V)$ and $h : W \rightarrow h(W)$ with images $y_D = \psi(y_U, x_1, x_2, x_3)$ and $y_R = h(y_U, x_2) = q^{-1}(y_U, x_2)$ such that $\langle \psi(y_U, x_1, x_2, x_3), h(y_U, x_2) \rangle \in W$ and

$$\begin{aligned} f(\psi(\cdot), h(\cdot), x_1, x_2, x_3) &= 0 \\ y_U - q(h(\cdot), x_2) &= 0 \end{aligned} \quad (2.18)$$

Then the trade-off between the desirable and undesirable output at the weakly efficient point

$\langle \hat{y}_D, \hat{y}_U, \hat{y}_R, \hat{x}_1, \hat{x}_2, \hat{x}_3 \rangle$ is given by:

$$\frac{\partial \psi(\hat{y}_U, \hat{x}_1, \hat{x}_2, \hat{x}_3)}{\partial y_U} = -\frac{f_{y_R}(\hat{y}_D, \hat{y}_R, \hat{x}_1, \hat{x}_2, \hat{x}_3)h_{y_U}(\hat{y}_U, \hat{x}_2)}{f_{y_D}(\hat{y}_D, \hat{y}_R, \hat{x}_1, \hat{x}_2, \hat{x}_3)} \geq 0 \quad (2.19)$$

This is because (2.4) and (2.5) establish that $f_{y_D} \geq 0$ and $f_{y_R} \geq 0$, and (2.12) imposes that $q_{y_R} < 0$, thus $h_{y_U} < 0$ given that $q^{-1}(y_U, x_2) = h(y_U, x_2)$. Intuitively, the trade-off between the desirable and undesirable output is non-negative because in a local neighborhood of the weakly efficient point $\langle \hat{y}_D, \hat{y}_U, \hat{y}_R, \hat{x}_1, \hat{x}_2, \hat{x}_3 \rangle$ of the technology T , holding the levels of all the inputs fixed, an increase in y_U can be generated only by a reduction in socially responsible activities and hence, as the input usage is constant, the resources must be diverted from the production of the socially responsible output to the production of the desirable output.

Defining $\check{f}(y_D, y_U, x_1, x_2, x_3) = f(y_D, h(y_U, x_2), x_1, x_2, x_3)$, where $h(y_U, x_2) = y_R$, the technology T can be reformulated as

$$T = \{ \langle y_D, y_U, y_R, x_1, x_2, x_3 \rangle \in \mathbb{R}_+^{M+N} \mid \check{f}(y_D, y_U, x_1, x_2, x_3) \leq 0 \wedge y_R \geq h(y_U, x_2) \} \quad (2.20)$$

and the function $\check{f}(\cdot)$ can be used to analyze the trade-off between the desirable output and the by-product generating input, i.e. the input that is responsible for the generation of the undesirable output. This trade-off can be expressed as

$$-\frac{\check{f}_{x_2}(y_D, y_U, x_1, x_2, x_3)}{\check{f}_{y_D}(y_D, y_U, x_1, x_2, x_3)} = -\frac{f_{x_2}(y_D, y_R, x_1, x_2, x_3) + f_{y_R}(y_D, y_R, x_1, x_2, x_3)h_{x_2}(y_U, x_2)}{f_{y_D}(y_D, y_R, x_1, x_2, x_3)} \quad (2.21)$$

Because the technology T violates free disposability of x_2 by exhibiting the opposite costly disposability property, the sign of the trade-off in (2.21) is ambiguous. This is because an increase in x_2 has a composite effect on the desirable output for fixed levels of inputs x_1 and x_3 and undesirable output y_U . On the one hand, an increase in x_2 generates the standard non-negative effect on y_D given by $-\frac{f_{x_2}(y_D, y_R, x_1, x_2, x_3)}{f_{y_D}(y_D, y_R, x_1, x_2, x_3)} \geq 0$, which directly depends on the conventional free disposability properties of the technology T_1 expressed in (2.4) and (2.6). On the other hand, an increase in x_2 generates a non-positive effect given by $-\frac{f_{y_R}(y_D, y_R, x_1, x_2, x_3)h_{x_2}(y_U, x_2)}{f_{y_D}(y_D, y_R, x_1, x_2, x_3)} \leq 0$, which depends on the technology T_2 displaying costly disposability in x_2 , as verified by the condition in (2.13). Intuitively, this second effect is non-positive because an increase in x_2 needs to be compensated by an increase in socially responsible activities in order to keep the level of undesirable output constant. However, since the levels of the other inputs x_1 and x_3 are also constant, implementing more socially responsible activities requires resources to be diverted from the production of the desirable output to the production of the socially responsible out-

put. Thus, the term $-\frac{f_{y_R}(y_D, y_R, x_1, x_2, x_3)h_{x_2}(y_U, x_2)}{f_{y_D}(y_D, y_R, x_1, x_2, x_3)}$ reflects the fact that the costly disposability of the input x_2 has negative repercussions on the production of desirable output. Nonetheless, not knowing the relative magnitudes of these two opposite effects, the sign of the trade-off between the costly disposable input x_2 and the desirable output cannot be determined.

In addition, the function $\check{f}(\cdot)$ allows for analyzing the trade-off between the socially responsible input and the undesirable output. Recall that input x_3 is identified as socially responsible because, while contributing to the production of the desirable output as any other conventional input, it simultaneously contributes to the production of socially responsible activities as well. Differentiating $\check{f}(\cdot)$ with respect to x_3 and y_U yields

$$\begin{aligned} -\frac{\check{f}_{x_3}(y_D, y_U, x_1, x_2, x_3)}{\check{f}_{y_U}(y_D, y_U, x_1, x_2, x_3)} &= -\frac{f_{x_3}(y_D, y_R, x_1, x_2, x_3)}{f_{y_U}(y_D, y_R, x_1, x_2, x_3) + f_{y_R}(y_D, y_R, x_1, x_2, x_3)h_{y_U}(y_U, x_2)} \\ &= -\frac{f_{x_3}(y_D, y_R, x_1, x_2, x_3)}{f_{y_R}(y_D, y_R, x_1, x_2, x_3)h_{y_U}(y_U, x_2)} \leq 0 \end{aligned} \quad (2.22)$$

The non-positive trade-off between the socially responsible input and the undesirable output captures the fact that an increase in x_3 decreases the undesirable output through the mitigating effect of socially responsible activities, given fixed levels of inputs x_1 and x_2 and desirable output.

In sum, this model provides a rigorous yet flexible framework that describes the transformation process of multiple inputs into multiple outputs with CSR being one of these multiple outputs of production. The properties of the joint technology reflect the fundamental economic notions of trade-off and opportunity cost and highlight two crucial issues that firms face when organizing and managing their production. First, firms engage in a variety of activities - some of them are desirable, marketable and directly profitable, some of them can be undesirable and bad, some of them are not marketable, and hence not directly profitable, but may be valuable in different ways. Second, as firms operate under scarcity, it is crucial to recognize that the decision of allocating resources to CSR activities necessarily implies taking away resources from other production activities.

3 Empirical Methodology

In this section we develop an empirical framework that has several practical implications for evaluating the impact of CSR on firms' production structure and value. We start by introducing the output directional distance function, a parsimonious and computationally accessible way

of describing the joint technology accommodating for desirable, undesirable, and mitigating outputs developed in our model. Given this data-friendly characterization of the technology, we then show how to generate a measure of inefficiency for each firm that provides an implicit ranking and permits the identification of the leaders in the industry, i.e. the firms that are able to produce the product mix (of desirable, undesirable, and CSR outputs) in the most efficient way. In addition, we demonstrate how to obtain an actual measure of the (internal/shadow) value of CSR for each firm even if CSR is a non-marketed output. Lastly, we derive the marginal shadow value of CSR as a measure of willingness to gain and willingness to lose to emphasize that firms face both a benefit of enhancing their socially responsible effort as well as a cost of diminishing it and that benefits and costs may be significantly different.

3.1 Primal Problem: Measuring (In)Efficiency

The set representation of the technology illustrated in section 2 is conceptually useful in characterizing the properties of the transformation process and the relationships between inputs and outputs, but it is not very helpful from an empirical perspective. To this end it is useful to turn to a function representation of the technology that is computationally accessible while maintaining the same assumptions of convexity, feasibility, and disposability discussed in the model set-up. The function representation chosen here is the directional output distance function, a more general and flexible variation of Luenberger's shortage function,⁹ defined as

$$\begin{aligned} \overrightarrow{D}_O(y_D, y_U, y_R, x_1, x_2, x_3; g_{y_D}, -g_{y_U}, g_{y_R}) = \\ \max \{ \beta \mid \langle y_D + \beta g_{y_D}, y_R + \beta g_{y_R}, x_1, x_2, x_3 \rangle \in T_1, \langle y_U - \beta g_{y_U}, y_R + \beta g_{y_R}, x_2 \rangle \in T_2 \} \end{aligned} \quad (3.1)$$

where $g_y = (g_{y_D}, -g_{y_U}, g_{y_R})$ is a vector that determines the direction in which \overrightarrow{D}_O is defined.

This function seeks to simultaneously expand the good outputs (desirable and socially responsible output) while contracting the undesirable output. The fact that the directional vector is preassigned allows for expanding or contracting any output in different directions making the directional output distance function particularly suitable in the presence of bad outputs. The same suitability does not apply to radial output distance functions¹⁰ because they only allow for expanding every output proportionally and at the same rate as much as it is feasible, which is certainly not appropriate when undesirable outputs are produced along

⁹The directional output distance function was first introduced by Chambers, Chung, and Färe (1998). It is derived from the shortage function (Luenberger (1992, 1995)) and the translation function (Blackorby and Donaldson (1980)).

¹⁰For radial output distance functions see Shepherd (1970).

with desirable ones.

The directional output distance function measures the distance, in the preassigned direction g_y , to the boundary of the technology T . Therefore, it can be interpreted as a measure of inefficiency, i.e. by how much desirable and socially responsible outputs can be expanded and undesirable output contracted within the feasibility limits imposed by the technology. In other words, a firm with output bundle (y_D, y_U, y_R) producing inside T operates efficiently if, given the directional vector g_y , it is able to expand the desirable and socially responsible outputs and contract the undesirable output to the boundary of T at the point $\langle y_D + \beta^* g_{y_D}, y_U - \beta^* g_{y_U}, y_R + \beta^* g_{y_R} \rangle$, where $\beta^* = \overrightarrow{D_O}(y_D, y_U, y_R, x_1, x_2, x_3; g_{y_D}, -g_{y_U}, g_{y_R})$.

For each firm i the problem of maximizing efficiency (or minimize inefficiency) in the primal (quantity) space consists of finding β^{i*} , which is the directional output distance function, and can be formalized as

$$\max_{\beta^i, \lambda^j} \beta^i \quad (3.2)$$

$$\text{s.t. } y_D^i + \beta^i g_{y_D} \leq \sum_{j=1}^I \lambda^j y_D^j \quad (3.3)$$

$$y_U^i - \beta^i g_{y_U} \geq \sum_{j=1}^I \lambda^j y_U^j \quad (3.4)$$

$$y_R^i + \beta^i g_{y_R} = \sum_{j=1}^I \lambda^j y_R^j \quad (3.5)$$

$$x_n^i \geq \sum_{j=1}^I \lambda^j x_n^j \quad n = 1, 3 \quad (3.6)$$

$$x_2^i = \sum_{j=1}^I \lambda^j x_2^j \quad (3.7)$$

$$\lambda^j \geq 0 \quad \forall j = 1, \dots, I \quad (3.8)$$

This problem is perfectly consistent with the axiomatic approach developed in section 2 as the constraints reflect the properties of the joint technology. In particular, (3.3) and (3.6) represent free disposability of desirable output and inputs x_1 and x_3 ; (3.4) represents costly disposability of undesirable output; (3.5) and (3.7) represent the fact that free disposability of socially responsible output and input x_2 is violated in the joint technology T because the sub-technology T_1 satisfies free disposability with respect to y_R and x_2 while the sub-technology

T_2 satisfies the opposite condition, i.e. costly disposability, with respect to these elements of the production process.

The maximization problem above is equivalent to the following

$$\max_{\beta^i, \lambda^j} \beta^i \quad (3.9)$$

$$\text{s.t.} \quad - \sum_{j=1}^I \lambda^j y_D^j \leq -y_D^i - \beta^i g_{y_D} \quad (3.10)$$

$$\sum_{j=1}^I \lambda^j y_U^j \leq y_U^i - \beta^i g_{y_U} \quad (3.11)$$

$$\sum_{j=1}^I \lambda^j y_R^j \leq y_R^i + \beta^i g_{y_R} \quad (3.12)$$

$$- \sum_{j=1}^I \lambda^j y_R^j \leq -y_R^i - \beta^i g_{y_R} \quad (3.13)$$

$$\sum_{j=1}^I \lambda^j x_n^j \leq x_n^i \quad n = 1, 3 \quad (3.14)$$

$$\sum_{j=1}^I \lambda^j x_2^j \leq x_2^i \quad (3.15)$$

$$- \sum_{j=1}^I \lambda^j x_2^j \leq -x_2^i \quad (3.16)$$

$$-\lambda^j \leq 0 \quad \forall j = 1, \dots, I \quad (3.17)$$

since any equality constraint is mathematically equivalent to a pair of inequality constraints of opposite sign. The reason why it is useful to express the primal problem as in (3.9)-(3.17) will become particularly evident as we move on to the characterization of the dual problem.

3.2 Dual Problem: Internal Values

The primal problem occurs in the quantity space and requires each firm i to maximize β , which translates to minimizing inefficiency by projecting its input-output mix as close to the technological frontier as possible. The mirror image to this problem is called the dual problem

because it is derived from the primal using duality theorems¹¹ and can be formalized as

$$\max_{p^i, w^i} p_D^i y_D^i - p_U^i y_U^i + (\hat{p}_R^i - p_R^i) y_R^i - w_1^i x_1^i - (w_2^i - \hat{w}_2^i) x_2^i - w_3^i x_3^i \quad (3.18)$$

$$\text{s.t. } p_D^i g_{y_D} + p_U^i g_{y_U} + (\hat{p}_R^i - p_R^i) g_{y_R} \geq 1 \quad (3.19)$$

$$p_D^i y_D^1 - p_U^i y_U^1 + (\hat{p}_R^i - p_R^i) y_R^1 - w_1^i x_1^1 - (w_2^i - \hat{w}_2^i) x_2^1 - w_3^i x_3^1 \leq 0 \quad (3.20)$$

⋮

$$p_D^i y_D^I - p_U^i y_U^I + (\hat{p}_R^i - p_R^i) y_R^I - w_1^i x_1^I - (w_2^i - \hat{w}_2^i) x_2^I - w_3^i x_3^I \leq 0 \quad (3.21)$$

$$p_D^i, p_U^i, p_R^i, \hat{p}_R^i, w_1^i, w_2^i, \hat{w}_2^i, w_3^i \geq 0 \quad (3.22)$$

The interpretation of the dual problem is insightful and quite straightforward. For each firm i minimizing inefficiency in the primal (quantity) space is equivalent to finding a system of optimal, relative (to the numeraire bundle), internal/shadow values that rationalize profit maximization in the dual (price) space. The shadow prices p^i and w^i that solve the dual problem are different for each firm as they are not market prices but internal valuations that each firm assigns to its outputs and inputs, consistently with profit maximization, representing the contribution of each output and input in creating value for the firm.

Because the technology is characterized by the presence of an undesirable output, y_U , the internal value of this output is negative since disposing of y_U represents actually a cost for the firm. It is also of interest to analyze the internal values associated with the socially responsible output y_R and the by-product generating input x_2 , i.e. $(\hat{p}_R - p_R)$ and $-(w_2 - \hat{w}_2)$, respectively. Putting in place socially responsible activities is costly for the firm and this is represented by the negative value p_R . At the same time, the production of CSR positively contributes to the mitigation of the undesirable output, as reflected in the positive value \hat{p}_R . As a result, the overall value of CSR for the firm depends on the relative magnitude of these opposite effects. A similar argument applies for the input x_2 . In fact, this input represents a cost to the firm when it is used in the production of the desirable output, as reflected by the negative sign of w_2 . Nonetheless, the same input has also a beneficial effect, as shown by the positive value \hat{w}_2 , because reducing this input use also reduces the amount of undesirable output produced. Once again, the total contribution of the by-product generating input in terms of profits depends on the relative magnitude of these contrasting effects. This simultaneously costly and beneficial nature of the internal values associated with y_R and x_2 directly depends on the fact that these elements are part of the technology with different and possibly contrasting roles and

¹¹For the exposition of the primal and dual problems in matrix form and a full derivation of the dual from the primal see the Appendix.

properties.¹²

Finally, note that the constraints of the dual problem provide an alternative characterization of the technology. Specifically, the constraint in (3.19) is a normalization implying that all the internal values p^i and w^i derived in the dual problem are expressed in terms of the numeraire bundle g_y . This is because the directional distance function in the primal problem can be also interpreted as a collection of outputs and inputs, thus it can be thought as a numeraire bundle whose price in terms of itself is always one. In addition, the constraints (3.20)-(3.21) reflect the fact that, since the system of internal/shadow prices that solves the dual problem for firm i is optimal only for firm i , this set of constraints holds at equality for firm i only if evaluated at firm's i optimal prices p^i and w^i . For every other firm, p^i and w^i cause this set of constraints to hold with inequality because at firm's i optimal internal prices every other firm is inefficient in the sense that it is not able to match the internal cost of the input bundle with the internal value of the output bundle. That is, p^i and w^i are necessarily inconsistent with profit maximization for any firm other than firm i .

3.3 Shadow Value and Marginal Impact of CSR

In economics the concept of marginal value refers to the change in a value associated with a specific change in some controlled variable, or the measure of the worthiness of a good in terms of other goods. In many instances marginal values are more insightful than overall values as they allow to isolate the effects of single variables variations and quantify trade-offs. Marginal values are usually derived by differentiating smooth functions that characterize the environment of interest (e.g. production, profits, costs, utility, expenditure).

Data Envelopment Analysis (DEA) technologies are conservative approximations derived as convex hulls of observed data points and present, by construction, kinks. The kinks are determined by the extreme efficient firms - the firms that are not simply efficient as they are on the technological frontier, but actually determine its shape by identifying its vertexes. This lack of smoothness renders DEA models not amenable to conventional differential arguments, at least for the extreme efficient firms. More specifically, the kinks associated with extreme efficient units in the primal (quantity) space map into flat portions in the dual (price/internal value) space. Thus, non-differentiability at the kinks in the primal problem translates to non-uniqueness of the internal/shadow values in the dual. Simply put, the dual problem described

¹²Expressing the primal problem solely with inequalities as in (3.9)-(3.17) yields a dual problem that clearly shows the ambivalent nature of y_R and x_2 as represented by their beneficial (\hat{p}_R and \hat{w}_2) and costly (p_R and w_2) contributions to firm value.

in section 3.2 has multiple optimal solutions (p, w) for extreme efficient firms. To overcome this issue Chambers and Färe (2008) apply generalized differential arguments (i.e. directional derivatives and superdifferentials) to DEA representations of technologies to infer marginal shadow values of inputs based on the concept of willingness to pay and willingness to accept.

In the context of this study, we make use of their approach and extend it to the directional output distance function to derive the shadow prices that each (extreme efficient) firm attaches to increasing or forgiving the production of one marginal unit of CSR.¹³ While extreme efficient firms potentially present infinitely many (normalized) shadow prices for the socially responsible output y_R , we identify the only two prices that are economically relevant: the gaining shadow price and the losing shadow price. The first represents *willingness to gain* - a measure of what an extreme efficient firm would be willing to receive for engaging in the production of one extra unit of the socially responsible output. The second represents *willingness to lose* - a measure of what an extreme efficient firm would be willing to give up to forgive the production of one unit of CSR. At the kinks these two prices diverges but are still uniquely identified.

The non-smoothness of the technology generates a gap between willingness to gain and willingness to lose with the gaining shadow price being greater than or equal to the losing shadow price. The fact that for the extreme efficient firms willingness to gain and willingness to lose differ is crucial to recognize that these firms face both a value of being more and a cost of being less socially responsible. The value and the cost are asymmetrical since the benefit of doing more CSR should exceed the damage of doing less.

4 Data

4.1 General and Application-Specific Data Issues

The empirical analysis investigates the food and beverage manufacturing sector because it presents peculiar and interesting production and CSR characteristics. The sector is populated by firms producing very differentiated products, thus competing on different grounds in terms of desirable outputs. The high differentiation in marketed products is not equally prominent with respect to CSR activities as all food manufacturers face similar challenges concerning food safety controversies, food sustainability and security, demand for healthier food products, ethical sourcing of raw materials. Along with these specific issues, firms in

¹³The full derivation of the marginal shadow prices associated with socially responsible activities in the context of the directional output distance function is provided in Appendix.

the food manufacturing sector also face CSR issues common to every manufacturing sector such as responsible usage of water and energy, as well as supply chain and labor standards management. This rich and diversified array of activities allows for observing and exploring multiple dimensions of CSR engagement making the food sector an insightful example of how our theoretical and empirical frameworks can be employed to analyze the impact and the value of CSR at the firm level.

Collecting data to conduct an empirical analysis based on a multiple input, multiple output model presents difficulties as detailed disaggregated measures of inputs and outputs at the micro level are usually not available. With respect to inputs the task is less demanding since the popularity of the KLEM (capital, labor, energy, materials) model in production economics has established the practice of collecting input data, or more often data on input expenditures, at least for the general input categories of capital and assets, labor, energy and materials use. With respect to outputs the task is more complicated as aggregate sales at the firm level are the most commonly available measure of output. This last consideration emphasizes the challenge of finding good measures of non-marketed outputs, namely outputs that are produced but are not sold in a market, such as the undesirable output and the CSR output in the joint production model proposed here.

Data on CSR implementation are particularly difficult to acquire for the following reasons. First, even if firms seem to agree on the fact that CSR activities are essential for their business and increasingly engage in their production, they have yet to develop a consistent and precise way of recording the resources they actually devote to CSR. Second, the need for bringing CSR to the core of the business strategy is clearly accompanied by the need for transparent communication. However, in the absence of mandatory criteria and strict guidelines, firms' reporting on CSR activities is not homogeneous and not easily comparable across firms. Lastly, the strategic importance that CSR has achieved has triggered the proliferation of consulting firms and institutions working on providing scores and rankings that summarize in one final number the CSR performance of each firm. Unfortunately, since this final number is usually obtained as some weighted combination of inputs and outputs involved in the generation of CSR, data on scores and rankings are normally not appropriate in a multiple input/output framework. Even so, this kind of data are practically the only available information on CSR performance at the firm level.

Sustainalytics¹⁴ is a global responsible investment research firm dedicated to support

¹⁴<http://www.sustainalytics.com/> "Sustainalytics is an award-winning global responsible investment research firm specialized in environmental, social and governance (ESG) research and analysis. The firm offers global perspectives and solutions that are underpinned by local expertise, serving both values-based and mainstream

investors with the development and implementation of responsible investment strategies. Sustainalytics' research focuses on developing a reliable and structured scoring system for firms with respect to their ESG/CSR performance and it is based on a methodology that identifies specific issues for each industry, scores every issue for each firm belonging to the same industry, and provides a CSR ranking that evaluates the relative performance of each firm with respect to their peers in the industry. Even if consisting of scores, the dataset provided by Sustainalytics is particularly suitable for the analysis developed in this study because it consists of detailed scores for different CSR indicators along with a final ranking. These detailed CSR scores are available for each firm included in the sample, so that each firm presents several disaggregated data points. Moreover, firms belonging to the same industry are scored on the same issues, so that firms can be consistently evaluated and compared and the occurrence of missing values is minimized. In addition, their methodology focuses on identifying strengths and weaknesses for every CSR category (environment, social, governance) in which the single indicators are organized. The fact that scores for detailed indicators are available together with the distinction of these indicators between favorable/positive and controversial/negative aspects of CSR is extremely helpful in identifying measures of mitigating CSR outputs and socially responsible inputs (scores for the positive indicators) and measures of undesirable outputs and undesirable inputs (scores for the negative indicators).

To carry out the analysis we construct a dataset matching information on CSR performance from Sustainalytics with information on desirable/marketed output and conventional production inputs based on the information included in companies' annual financial reports. Specifically, sales, fixed assets, and cost of goods sold are obtained from Orbis (Bureau van Dijk) and number of employees from Orbis, ThompsonOne (Thomson Reuters) or firms' on-line accessible reports depending on where this information was available. Sales are used as a measure of the desirable, marketed output y_D , while fixed assets, number of employees and cost of goods sold are used as measures of capital, labor and variable inputs, respectively, and constitute the elements of the conventional inputs of production vector x_1 . The construction of the remaining part of the dataset, i.e. measures of the undesirable output y_U , the undesirable input x_2 and the socially responsible input x_3 is one of the innovations and contributions of this study and deserves a thorough illustration.

Recall that the information provided by Sustainalytics is in the form of scores. For each industry a certain (usually quite large) number of indicators across the three Environment, Social and Governance dimensions of CSR performance are chosen and assigned a raw

investors that integrate ESG information and assessments into their investment decisions."

score from 0 to 100 where 0 denotes a very poor performance and 100 denotes an excellent performance. Along with the raw scores, Sustainalytics provides also a system of industry-specific weights for each indicator that reflect its importance for characterizing the overall ESG performance in each sector. The weights are sector-specific to capture the idea that different indicators might matter more or less for CSR depending on the industry. For example, while managing emissions and toxic waste could be very important in the chemical sector, it certainly is not as important in the banking sector. To construct the measure of socially responsible output y_R we select 9 indicators among those available and aggregate them into one weighted score using a re-scaled system of weights that reflects the relative importance given to these indicators in the original Sustainalytics dataset. Similarly, we construct measures of undesirable output y_U , undesirable input x_2 and socially responsible input x_3 aggregating 10, 3 and 2 indicators, respectively. Note that, as the scores assigned by Sustainalytics are increasing in the performance, i.e. the better the performance the higher the score, for the undesirable output and undesirable input the inverse of the original score ($100 - \text{original score}$) is used to be consistent with the theoretical framework. Table 1 provides further details on the variables constituting the dataset used in the empirical analysis. The choice of the indicators used to construct the measures of socially responsible and undesirable inputs and outputs is the result of a careful analysis of the proprietary ESG report compiled by Sustainalytics for each firm that describes each indicator in detail to insure that only indicators that can be clearly and meaningfully identified as inputs or outputs measures and are consistently scored for all the firms are included in our analysis.

Because of the difficulty of matching data from several different sources, the presence of missing values and the necessity of being cautious in selecting indicators from the Sustainalytics dataset that can appropriately represent inputs and outputs, the final dataset consists of a cross-section of 175 publicly traded firms. The sample includes the major players in the food and beverage manufacturing industry worldwide. For each firm we have data on a total of 8 variables - one desirable output, one undesirable output, one socially responsible output, three conventional inputs, one undesirable input, and one socially responsible input. All the data refer to 2014, the latest available year in our data sources with the most complete information. Table 3 reports descriptive statistics of the variables used in the analysis specifying in what units they are expressed¹⁵ while table 2 displays the weights assigned to each component used to construct the variables y_U , y_R , x_2 and x_3 .

¹⁵Understanding the units in which the variables are expressed is important to understand the results and the insights of the empirical analysis

4.2 Summary Statistics and Characteristics of the Dataset

The information provided in table 3 and displayed in figures 1 and 2 describes several insightful patterns in the data. Desirable output (sales) and conventional production inputs (capital, labor, variable inputs) have very high variability and a considerable number of outliers while undesirable output, undesirable input and socially responsible input have a less disperse distribution with almost no outliers¹⁶. Also, desirable output and conventional production inputs seem to be log-normally distributed, as shown in figure 1, where a normal curve is imposed over the histogram of each variable, and confirmed statistically by the result of a test for normality which fails to reject the null hypothesis of normality for all these conventional production variables.¹⁷ Given the truncated nature of the data (scores from 0 to 100) used for the socially responsible and undesirable inputs and outputs these variables are obviously not normally distributed and their distributions are not symmetric with the undesirable output, socially responsible output and socially responsible input being clearly skewed to the right.

Figures 3 - 5 present scatter plot matrices of desirable output with the other outputs (undesirable and socially responsible), desirable output with conventional production inputs (capital, labor and variable inputs), and desirable output with the other production inputs (undesirable and socially responsible), respectively. The matrix diagonal contains the kernel density of each variable. Figure 3 suggests the existence of a positive correlation between the desirable and undesirable output as predicted in the theoretical model where y_D and y_U are positively correlated because y_U is a by-product of y_D . The correlation between the desirable output and the socially responsible output, on the other hand, is not decisively positive confirming the mechanism that in the theoretical model makes socially responsible efforts beneficial in terms of their mitigating effect but costly in terms of production resources. The scatter plot matrix in figure 4 shows a positive correlation between conventional inputs and desirable output, as expected. However, the nature of the correlation between the other inputs of production and the desirable output in figure 5 is varied, showing no clear pattern.

Figures 6-7 characterize the geographical distribution of the firms. Asia is the most represented area with almost 40 percent of the firms included in the analysis followed by Europe and North America (USA and Canada). USA is by far the most represented country with almost 20 percent of the entire sample being located in the USA. The second most

¹⁶This is partially due to the fact that, even if the scale of the raw scored assigned by Sustainalytics is from 0 to 100, the scores are usually assigned in quintiles, i.e. 0, 25, 50, 75, 100.

¹⁷The test for normality is the one proposed by D'agostino, Belanger, and D'Agostino Jr (1990), with the empirical correction developed by Royston (1991). The null hypothesis cannot be rejected with a 5 percent confidence level for the desirable output and with a 1 percent confidence level for capital, labor and variable inputs.

represented country is Japan with 13 percent of the firms. UK and China follow with about 6 percent of the firms being located in each of these countries.

5 Results

5.1 Efficiency Measure

We start our empirical exploration of the role of socially responsible activities within the firm's production dynamics by reviewing the results of our benchmarking exercise, as defined by the primal problem, that involves characterizing the joint production technology with CSR as an additional output and comparing the overall technical performance of the firms included in the sample. Since we include three outputs and five inputs in our empirical application, it is reasonable to expect a technology that is flexible in encompassing this complex production process yielding high levels of efficiency.

Technically, the solution to the primal problem¹⁸ in (3.2) - (3.8) provides a measure of inefficiency that quantifies how much desirable and socially responsible outputs can be expanded and undesirable output contracted within the feasibility constraint imposed by the technology T . The expansion of y_D and y_R and the contraction of y_U are in the pre-assigned direction of $g_y = (g_{y_D}, -g_{y_U}, g_{y_R})$. In this case g_y has been arbitrarily chosen to be $g_y = (1, -1, 1)$ which simply means that all the outputs are considered equally important when moving toward the frontier. Note that, when the directional vector is chosen such that it enters the constraints of the primal problem additively, as in this case, the inefficiency score β has a lower bound at 0 and an upper bound that depends on the scale and magnitude of the data. Therefore, $\beta = 0$ indicates efficiency while $\beta > 0$ indicates a margin for technical improvement where the higher the value of β the higher the inefficiency.

In practice, this measure of inefficiency, β , quantifies the distance between each firm's observed input-output bundle and the technological frontier. The frontier represents what is feasible in terms of production possibilities, hence the higher the distance from the frontier the larger the gap between what a firm is able to produce and what is possible to produce given the technology. Solving the primal problem constitutes a benchmarking exercise as it generates a ranking of firms depending on their inefficiency score. Broadly, we can separate between inefficient firms, which are inside the technology set and have a β bigger than zero, and efficient firms, which are on the technological frontier and have a β equal to zero. Among

¹⁸We assume variable returns to scale and solve the primal problem for each firm imposing the additional constraint that $\sum_{j=1}^J \lambda^j = 1$.

the efficient firms we can further separate between just efficient and extreme efficient firms. Just efficient firms are geometrically located on a flat portion of the frontier while extreme efficient firms are located on the vertexes of the convex hull characterizing the technology. This means that the extreme efficient firms are those shaping the technology and serving as the benchmark for all the other firms.

In our sample of food and beverage manufacturing firms efficiency levels are very high. Approximately 75 percent of the firms are technically efficient, among them only 30 are just efficient while the remaining 102 are extreme efficient. Less than 25 percent of the firms (43 firms) are found to be inefficient. In figure 8 the distribution of the efficiency scores is illustrated through a histogram and a nonparametric kernel density. The distribution is clearly concentrated around zero since $\frac{3}{4}$ of the firms are technically efficient.

This result is quite reasonable for several reasons. First, the empirical analysis is carried out with three outputs and five inputs but each of these outputs and inputs have a specific and peculiar role in the production process. Thus, while firms have more freedom in articulating the scope of their production along different dimensions, their decisions are also necessarily more complex as these different dimensions can be conflicting. For example, more desirable outputs generates more undesirable output, which then needs to be mitigated. Similarly, socially responsible activities can add value to the firm but are costly in terms of resources that need to be allocated to their production. These trade-offs translate to the constraints defining the technology 'pulling' the boundaries of the feasibility set in different, sometimes opposite, directions. Therefore, the more freedom and more choices available to firms generate a very peculiar technology set that can accommodate different production 'recipes' and makes it easier for firms to be efficient. Second, while the variability in conventional inputs and output is quite large in the data this is not the case for undesirable output, socially responsible output and undesirable and responsible inputs. This is in part due to data limitations, but also to the fact that the CSR performance of firms seems to be much more homogeneous - that is, there seems to be a minimum standard that every firm strives to achieve. Hence, even firms that are not extremely competitive in terms of sales or conventional productive resources (capital, labor, variable inputs) are, on the contrary, very competitive in terms of socially responsible efforts. Because of the linkages between the different elements of the production structure, this also generates higher levels of efficiency.

These findings suggest that it is important to account for the complexity of the production process and that ranking firms in terms of their performance and identifying leaders can be quite difficult when they engage in very differentiated activities. Nonetheless, a diversified portfolio that includes activities such as CSR can also represent an opportunity for firms to

compete and excel on new grounds.

5.2 Shadow Values

To understand and quantify whether engaging in socially responsible activities is valuable for the firm and how their value compares with the value of the other production activities we turn to the shadow prices derived as the solution to the dual problem in (3.18)-(3.22). Recall that these prices are not market prices but are determined optimally given the technology and consistently with profit maximization. That is, even if we do not have information on input and output market prices, because we observe the choices of each firm on the level of inputs and outputs, we assume that these choices are made in accordance with a profit maximizing behavior. However, because our framework allows for inefficiency in production, the value that each firm assigns to its input and outputs may not coincide with their market price. Also, because of duality, there is a correspondence between the primal and the dual problem. In particular, firms that are technically efficient will have zero ‘shadow profits’, i.e. the objective in (3.18) will be equal to zero. This means that being on the technological frontier is equivalent to making production decisions in terms of inputs and outputs such that the internal value of the output bundle is equivalent to the internal value of the input bundle. Thus, being efficient means to allocate resources without any ‘waste’, perfectly balancing their value and their cost.

For inefficient and just efficient firms the solution to the dual problem is unique and provides a measure of the internal value that each firm assigns to the elements of its input-output bundle. These values are firm-specific and represents the contribution of each input and output in creating value for the firm. Table 4 reports summary statistics for the output shadow prices¹⁹ while figure 9 depicts their distribution.

The average shadow price of desirable output p_D is positive²⁰ implying that, as expected, producing the desirable output positively contributes to the creation of value for the firm. Not surprisingly, the average shadow price of undesirable output p_U is negative²¹ and remarkably higher (in absolute value) than the other output shadow prices suggesting that the production of the undesirable output represents a considerable cost for the firm. The average shadow

¹⁹Note that these outputs are expressed in different units, i.e. million of USD for the desirable output and scores for the undesirable and socially responsible output, and their shadow prices are relative to the numeraire bundle g_y which makes the interpretation of the results not immediate.

²⁰By construction p_D is always non-negative, but the solution to the primal problem delivers an average p_D that is actually greater than zero.

²¹By construction p_U is always non-negative. However, since p_U enters the objective function and some of the constraints with a negative sign, the contribution of the undesirable output to the firm’s internal value, measured by p_U , is negative.

price of the socially responsible output is positive, meaning that the cost of producing CSR is compensated by the benefit of its mitigating effect. This result is remarkable in the context of our analysis because it is not imposed by the assumptions of the model where CSR has both a beneficial and a costly effect and can, therefore, have a positive or negative value depending on which effect is prevailing. Given that we find a high negative value attached to the undesirable output, it is necessary for firms to engage in socially responsible activities. However, it is relevant to register that overall firms are able to extract value from these activities even if their implementation is costly. It is interesting to point out that the shadow price of CSR is exactly zero for all the 30 just efficient firms, which indicates that these firms are capable of balancing the costs associated with socially responsible efforts with their mitigating nature in a perfectly efficient way.

This set of results is an important piece of the puzzle that our study attempts to solve - namely, is it worth it to be good? Our answer is yes. On average, CSR activities positively contribute to adding (internal) value to the firm. This answer is particularly meaningful because it is derived in a context that fully accounts for the opportunity costs and trade-offs that engaging in CSR entail and, even once these costs have been considered, we still find that being socially responsible is valuable.

5.3 Marginal Value of Socially Responsible Activities

In this section we focus on extreme efficient firms, those on the vertexes of the technology hull. The shadow value of CSR for these firms delivers the most interesting insights for understanding the value of being socially responsible at the margin. After establishing that for the firms included in our analysis it is worthy to engage in CSR, it is interesting to investigate whether further value can be added, or possibly lost, for a firm by changing its socially responsible commitment.²² Our analysis can deliver useful guidance to managers in terms of how much there is to gain (lose) for improving (diminishing) the socially responsible effort of the firm.

Recall that the dual problem in (3.18)-(3.21) does not have a unique solution for extreme efficient firms. However, focusing on the highest and lowest shadow price for the socially responsible output allows for obtaining the only two prices that are economically relevant, i.e. the gaining shadow price (the price an extreme efficient firm is willing to receive to produce one more unit of CSR) and the losing shadow price (the price an extreme efficient firm is

²²The effects of a marginal change in CSR performance are analyzed in the neighborhood of the current CSR performance.

willing to ‘pay’/give up to forgive the production of one unit of CSR).²³ These prices are normalized and expressed in the units of the numeraire bundle g_y .

Because the measure of socially responsible output used here is a score from 0 to 100, identifying the units and understanding the meaning of producing one more or one less unit of CRS is not straightforward. Technically, one unit of CSR corresponds to one score point but, since the raw scores are mostly given in quintiles (0, 25, 50, 75, 100) and the raw scores are weighted and aggregated into one single measure of socially responsible output, establishing the precise magnitude of one score point is potentially complicated. Nonetheless, to understand the results it is sufficient to loosely interpret the marginal shadow prices of CSR discussed here as a measure of how much an extreme efficient firm is willing to gain for improving its CSR performance (thus getting a higher score) or lose for worsening its CSR performance (thus getting a lower score). Table 5 provides some descriptive statistics for the upper and lower bound of p_R and their distribution is shown in figure 10.

The average upper bound for p_R , that can be interpreted as the average gaining shadow price²⁴ of engaging in the production of one additional unit of CSR, is positive and equal to 0.666. Even if this number might be complicated to interpret in terms of units and magnitude, its sign is indicative of the fact that, on average, extreme efficient firms attach a positive value to CSR activities and are willing to increase their socially responsible efforts for a positive price. Note that, since CSR is a mitigating yet costly activity, its price does not need to be necessarily positive. The fact that the average gaining shadow price is positive signals that extreme efficient firm consider a higher socially responsible commitment to be beneficial for adding value to their business. Conversely, the average lower bound for p_R , that can be interpreted as the average losing shadow price²⁵ of relinquishing the production of one unit of CSR, is negative and equal to -0.071. This result is both somewhat unanticipated but also very insightful. A negative losing price of CSR implies that, on average, extreme efficient firms are

²³See Mills (1956) for the mathematical foundations of finding marginal values in linear programs.

²⁴This upper bound for p_R is calculated averaging over the maximum shadow price of CSR obtained by solving a modified version of the dual problem for the 102 extreme efficient firms present in the sample. The modified dual problem calls for maximizing p_R^i under the same constraints of the standard dual problem presented in (3.18)-(3.22) and the additional constraint that $p_D^i y_D^i - p_U^i y_U^i + (\hat{p}_R^i - p_R^i) y_R^i - w_1^i x_1^i - (w_2^i - \hat{w}_2^i) x_2^i - w_3^i x_3^i = 0$.

²⁵This lower bound for p_R is calculated averaging over the minimum shadow price of CSR obtained by solving a modified version of the dual problem for 55 extreme efficient firms. In this case, the modified version of the dual problem consists of minimizing p_R^i under the same constraints of the standard dual problem presented in (3.18)-(3.22) and the additional constraint that $p_D^i y_D^i - p_U^i y_U^i + (\hat{p}_R^i - p_R^i) y_R^i - w_1^i x_1^i - (w_2^i - \hat{w}_2^i) x_2^i - w_3^i x_3^i = 0$. Note that the modified dual problem that generates the lower bound of p_R can be optimally solved only for 55 out of the 102 extreme efficient firms present in the sample. For the remaining 47 firms this problem is unbounded. The lack of a lower bound for p_R for some of the extreme efficient firms does not change the meaning of the results as it simply signifies that these 47 firms are willing to ‘pay’ an infinitely negative price for decreasing their socially responsible engagement.

not willing to ‘pay’ any price for reducing their socially responsible effort, instead they want to realize a gain. Hence, our findings suggest that engaging in less CSR is considered so costly and damaging that firms want to be compensated for doing so. This is particularly interesting and it is in line with the increasingly pervasive evidence that CSR has become an activity that firms feel compelled to do and, more importantly, that certain minimum standards/levels of CSR are perceived as necessary.

This second set of results on the marginal shadow prices of CSR provides a clearer picture about the value that firms can attach to socially responsible activities, highlighting two crucial features. First, there is a sharp asymmetry in terms of gaining and losing marginal values which reflects an asymmetry in willingness to engage in more and willingness to engage in less CSR. If we think of CSR as an array of risk management activities and we assume that firms have very high incentives to minimize the risk connected to undesirable outputs because they are very costly, it is perfectly reasonable to observe this asymmetry at the margin. Second, whether firms want to engage in CSR because they consider it beneficial or because they are afraid of bearing the cost of not doing so, our analysis clearly indicates that there is a positive value for a business in devoting resources and effort to being socially responsible.

6 Concluding Remarks

The conviction that CSR should be a prominent business practice has reached a wide consensus among firms, consumers, investors, and policy makers. While the academic literature in recent years has provided formal and rigorous support to this conviction, the existing research has almost exclusively focused on *why* CSR is done. This study takes a very different perspective and attempts to shed light on *how* CSR is done incorporating it into a formal production framework.

To this extent, we develop a joint production model for characterizing the technology and representing the transformation process of multiple inputs into multiple outputs. Specifically, each firm is assumed to produce a desirable output but, because the production of this desirable output may require the use of some undesirable input, an unwanted output can be generated along with the desirable one. Thus, the firm needs to engage in socially responsible activities to mitigate the unwanted output. The overall technology supporting this joint production is obtained as a composition of two distinct technologies: one describing the desirable-output production and the other describing the generation of the undesirable output. CSR is the link between these two technologies as it simultaneously represents the opportunity cost of

producing socially responsible activities in terms of desirable output and its mitigating effect with respect to undesirable output.

Empirically, the implementation of the analysis is carried out using a sample of 175 global firms in the food and beverage manufacturing sector and is based on a non-parametric DEA approach. DEA techniques allow for constructing the joint technology as the intersection of the desirable production and the undesirable production technologies. Once the technology is fully characterized it is possible to derive a measure of efficiency for benchmarking firms based on their technical performance as well as a set of internal/shadow values for inputs and outputs that reveals how much the production of CSR is worth to the firm in terms of the other outputs produced. For the most technically efficient firms (extreme efficient firms), the Chambers-Färe *calculus* method for DEA technologies is used to identify unique shadow values for CSR as measures of willingness to gain for producing one more unit of CSR and willingness to lose for giving up the production of one more unit of CSR.

The rigorous yet flexible theoretical and empirical frameworks adopted in this study are powerful tools for analyzing the impact of CSR on firm's performance and value and, together with the results of the analysis deliver insightful indications to managers. In particular, we find that in the sample of firms included in the analysis efficiency levels are very high as approximately 75 percent of the firms are found to be technically efficient. This suggests that it is important to account for the complexity of the production process and that ranking firms in terms of their performance and identifying leaders can be quite difficult when they engage in very differentiated activities. In addition we find that, on average, socially responsible activities positively contribute to adding (internal) value to the firm implying that the cost of implementing these activities is compensated by their beneficial mitigating effect. When focusing on quantifying the value of CSR at the margin we also find that the average marginal value of intensifying the socially responsible commitment is positive, indicating that more CSR is considered beneficial for adding value to the firm. Conversely, the average marginal value of decreasing the CSR effort is negative, indicating that lower levels of CSR are perceived so costly and damaging that firms want to be compensated for reducing their socially responsible engagement. Overall, these results reveal that for a business being good is worthy because a diversified portfolio that includes activities such as CSR can represent an opportunity for firms to compete and excel on new grounds and CSR activities support the creation of firm's value. It is important to emphasize that, while the results are specific to the food and beverage industry application, the model and empirical methodology are generalizable to any industry.

The road to fully understanding the impact and the value of CSR, a firm activity that is

so pervasive and yet so heterogeneous and not systematically recorded and documented, is still long and arduous for several reasons. First, CSR is a very popular but still relatively new concept and firms appear to be still in the process of learning what to do with CSR and how to do it in the most effective way. Second, even if a considerable number of CSR issues are common to many firms, the voluntary nature of CSR and the freedom of choosing the scope and the intensity that comes with it generates a very broad spectrum of CSR activities that are difficult to compare across firms. Further, the lack of a uniform and standardized *recipe* for CSR is necessarily reflected into the lack of straightforward measures for the *ingredients* of CSR. The fact that information on CSR is so diversified and fragmented and just plain difficult to acquire imposes a very hard constraint on the possibility of studying CSR empirically. All these considerations point to an obvious requirement for developing a better applied research agenda on CSR: data availability. Having access to more complete and more easily comparable data on a large number of firms would allow for answering interesting and deeper questions on the effect and the value of CSR. For example, with more and better data a production model like the one presented in this paper could be estimated parametrically and the decision of a firm to invest/engage in CSR could be modeled and investigated in a more structural and sophisticated way, similarly to what has been done for other firm-level decisions, such as investment, exporting, or R&D. Having enough data to estimate structural parameters would also open the door to analyzing interesting counterfactual scenarios (e.g. what if the level of CSR is imposed by the government and not chosen by the firm; what if having firms doing CSR is welfare improving as opposed to having consumers, or NGOs, or the government taking care of social responsibility).

References

- Ailawadi, K. L., Neslin, S. A., Luan, Y. J., & Taylor, G. A. (2014). Does retailer csr enhance behavioral loyalty? A case for benefit segmentation. *International Journal of Research in Marketing*, *31*(2), 156–167.
- Albuquerque, R., Durnev, A., & Koskinen, Y. (2015). *Corporate social responsibility and firm risk: Theory and empirical evidence*. (ECGI – Finance working paper No. 359, available at <http://dx.doi.org/10.2139/ssrn.1961971>)
- Auger, P., Burke, P., Devinney, T. M., & Louviere, J. J. (2003). What will consumers pay for social product features? *Journal of Business Ethics*, *42*(3), 281–304.
- Bagnoli, M., & Watts, S. G. (2003). Selling to socially responsible consumers: Competition and the private provision of public goods. *Journal of Economics & Management Strategy*, *12*(3), 419–445.
- Baron, D. P. (2001). Private politics, corporate social responsibility, and integrated strategy. *Journal of Economics & Management Strategy*, *10*(1), 7–45.
- Blackorby, C., & Donaldson, D. (1980). A theoretical treatment of indices of absolute inequality. *International Economic Review*, 107–136.
- Chambers, R. G., Chung, Y., & Färe, R. (1998). Profit, directional distance functions, and nerlovian efficiency. *Journal of Optimization Theory & Applications*, *98*(2), 351–364.
- Chambers, R. G., & Färe, R. (2008). A “calculus” for data envelopment analysis. *Journal of Productivity Analysis*, *30*(3), 169–175.
- Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. (1995). *Data envelopment analysis: Theory, methodology, and application*. Kluwer Academic Publishers.
- Cooper, W. W., Seiford, L. M., & Zhu, J. (2011). *Handbook on data envelopment analysis* (Vol. 164). Springer Science & Business Media.
- Crane, A., McWilliams, A., Matten, D., Moon, J., & Siegel, D. S. (2008). *The oxford handbook of corporate social responsibility*. Oxford University Press on Demand.
- Creyer, E. H. (1997). The influence of firm behavior on purchase intention: Do consumers really care about business ethics? *Journal of Consumer Marketing*, *14*(6), 421–432.
- D’agostino, R. B., Belanger, A., & D’Agostino Jr, R. B. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, *44*(4), 316–321.
- De Pelsmacker, P., Driesen, L., & Rayp, G. (2005). Do consumers care about ethics? Willingness to pay for fair-trade coffee. *Journal of Consumer Affairs*, *39*(2), 363–385.
- Dimson, E., Karakaş, O., & Li, X. (2015). Active ownership. *Review of Financial Studies*,

28(12), 3225–3268.

- Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, 60(11), 2835–2857.
- Flammer, C. (2015). Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach. *Management Science*, 61(11), 2549–2568.
- Førsund, F. R. (2009). Good modelling of bad outputs: Pollution and multiple-output production. *International Review of Environmental & Resource Economics*, 3(1), 1–38.
- Galema, R., Plantinga, A., & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking & Finance*, 32(12), 2646–2654.
- Luenberger, D. G. (1992). New optimality principles for economic efficiency and equilibrium. *Journal of Optimization Theory & Applications*, 75(2), 221–264.
- Luenberger, D. G. (1995). *Microeconomic theory*. McGraw-Hill College.
- Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). *Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance*. (Harvard Business School working paper, available at <http://dx.doi.org/10.2139/ssrn.1866371>)
- McWilliams, A., & Siegel, D. S. (2001). Corporate social responsibility: A theory of the firm perspective. *Academy of Management Review*, 26(1), 117–127.
- Mills, H. D. (1956). Marginal values of matrix games and linear programs. In H. Kuhn & A. Tucker (Eds.), *Linear inequalities and related systems* (Vol. 38, pp. 183–193). Princeton University Press, Princeton, NJ.
- Murty, S. (2010). *On the theory of a firm: the case of by-production of emissions*. (University of Warwick working paper, available at <http://wrap.warwick.ac.uk/3525/>)
- Murty, S., Russell, R. R., & Levkoff, S. B. (2012). On modeling pollution-generating technologies. *Journal of Environmental Economics & Management*, 64(1), 117–135.
- Oikonomou, I., Brooks, C., & Pavelin, S. (2012). The impact of corporate social performance on financial risk and utility: A longitudinal analysis. *Financial Management*, 41(2), 483–515.
- Paul, C. J., Siegel, D. S., Shadbegian, R. J., Gray, W. B., Vitaliano, D. F., Stella, G. P., . . . Granderson, G. (2006). Special issue on corporate social responsibility (CSR) and economic performance. *Journal of Productivity Analysis*, 26(3), 207–287. (Siegel D, Morrison Paul, C guest Eds.)
- Royston, P. (1991). Estimating departure from normality. *Statistics in Medicine*, 10(8),

1283–1293.

- Servaes, H., & Tamayo, A. (2013). The impact of corporate social responsibility on firm value: The role of customer awareness. *Management Science*, *59*(5), 1045–1061.
- Sharfman, M. P., & Fernando, C. S. (2008). Environmental risk management and the cost of capital. *Strategic Management Journal*, *29*(6), 569–592.
- Shepherd, R. W. (1970). *Theory of cost and production functions*. Princeton University Press.
- Siegel, D. S., & Vitaliano, D. F. (2007). An empirical analysis of the strategic use of corporate social responsibility. *Journal of Economics & Management Strategy*, *16*(3), 773–792.
- WBCSD. (1998). *Meeting changing expectations*. First report on corporate social responsibility, available at: <http://old.wbcds.org/pages/edocument/edocumentdetails.aspx?id=82&nosearchcontextkey=true>.

A Mathematical Appendix

A.1 Primal and dual problem in matrix form

The primal problem in equations (3.9)-(3.17) is formulated in its canonical matrix form as

$$\max_z c'z \quad (\text{A.1})$$

$$\text{s.t. } Az \leq b \quad (\text{A.2})$$

$$z \geq 0 \quad (\text{A.3})$$

which is

$$\max_z \left[\overbrace{\begin{matrix} c' \\ 1 & 0 & \cdots & 0 \\ \hline & 1 \times (1+I) & & \end{matrix}}^{c'} \right] \begin{matrix} \overbrace{\begin{matrix} \beta^i \\ \lambda^1 \\ \vdots \\ \lambda^I \end{matrix}}^z \\ \hline \\ \hline \\ (1+I) \times 1 \end{matrix} \quad (\text{A.4})$$

s.t.

$$\left[\overbrace{\begin{matrix} A \\ \hline g_{yD} & -y_D^1 & \cdots & -y_D^I \\ g_{yU} & y_U^1 & \cdots & y_U^I \\ -g_{yR} & y_R^1 & \cdots & y_R^I \\ g_{yR} & -y_R^1 & \cdots & -y_R^I \\ 0 & x_1^1 & \cdots & x_1^I \\ 0 & x_2^1 & \cdots & x_2^I \\ 0 & -x_2^1 & \cdots & -x_2^I \\ 0 & x_3^1 & \cdots & x_3^I \end{matrix}}^{A} \right] \begin{matrix} \overbrace{\begin{matrix} \beta^i \\ \lambda^1 \\ \vdots \\ \lambda^I \end{matrix}}^z \\ \hline \\ \hline \\ (1+I) \times 1 \end{matrix} \leq \begin{matrix} \overbrace{\begin{matrix} b \\ -y_D^i \\ y_U^i \\ y_R^i \\ -y_R^i \\ x_1^i \\ x_2^i \\ -x_2^i \\ x_3^i \end{matrix}}^b \\ \hline \\ \hline \\ (M_D+M_U+2M_R+N_1+2N_2+N_3) \times 1 \end{matrix} \quad (\text{A.5})$$

Duality theorems allow for deriving the dual problem from the primal that, in matrix form,

is given by

$$\min_{\gamma} \gamma' b \quad (\text{A.6})$$

$$\text{s.t. } \gamma' A \geq c' \quad (\text{A.7})$$

$$\gamma \geq 0 \quad (\text{A.8})$$

which is

$$\min_{\gamma} \left[\overbrace{p_D^i \ p_U^i \ p_R^i \ \hat{p}_R^i \ w_1^i \ w_2^i \ \hat{w}_2^i \ w_3^i}_{1 \times (M_D + M_U + 2M_R + N_1 + 2N_2 + N_3)} \right] \begin{matrix} b \\ \left[\begin{array}{c} -y_D^i \\ y_U^i \\ y_R^i \\ -y_R^i \\ x_1^i \\ x_2^i \\ -x_2^i \\ x_3^i \end{array} \right] \\ (M_D + M_U + 2M_R + N_1 + 2N_2 + N_3) \times 1 \end{matrix} \quad (\text{A.9})$$

s.t.

$$\left[\overbrace{p_D^i \ p_U^i \ p_R^i \ \hat{p}_R^i \ w_1^i \ w_2^i \ \hat{w}_2^i \ w_3^i}_{1 \times (M_D + M_U + 2M_R + N_1 + 2N_2 + N_3)} \right] \begin{matrix} A \\ \left[\begin{array}{cccc} g_{yD} & -y_D^1 & \cdots & -y_D^I \\ g_{yU} & y_U^1 & \cdots & y_U^I \\ -g_{yR} & y_R^1 & \cdots & y_R^I \\ g_{yR} & -y_R^1 & \cdots & -y_R^I \\ 0 & x_1^1 & \cdots & x_1^I \\ 0 & x_2^1 & \cdots & x_2^I \\ 0 & -x_2^1 & \cdots & -x_2^I \\ 0 & x_3^1 & \cdots & x_3^I \end{array} \right] \\ (M_D + M_U + 2M_R + N_1 + 2N_2 + N_3) \times (1+I) \end{matrix} \geq \begin{matrix} c' \\ \left[\begin{array}{cccc} 1 & 0 & \cdots & 0 \end{array} \right] \\ 1 \times (1+I) \end{matrix} \quad (\text{A.10})$$

With some manipulation,²⁶ the problem of each firm i in the dual (price) space is formalized as in equations (3.18)-(3.22).

A.2 Derivation of the marginal shadow prices for extreme efficient firms

In what follows we illustrate the *calculus* for DEA proposed by Chambers and Färe (2008) using the directional output distance function presented in (3.1). To facilitate the exposition this function is now simply redefined as $\overrightarrow{D}_O(y, x, g_y)$. Recall that \overrightarrow{D}_O is a function representation of the technology thus the assumptions on the joint technology T determine the properties of \overrightarrow{D}_O . Specifically, since T is convex, $\overrightarrow{D}_O(y, x, g_y)$ is concave in (y) and satisfies the translation property

$$\overrightarrow{D}_O(y + \delta g_y, x, g_y) = \overrightarrow{D}_O(y, x, g_y) - \delta, \quad \delta \in \mathbb{R} \quad (\text{A.11})$$

and the representation property

$$\overrightarrow{D}_O(y, x, g_y) \geq 0 \Leftrightarrow y \in T \quad (\text{A.12})$$

Because $\overrightarrow{D}_O(y, x, g_y)$ is concave in y , its directional derivative²⁷

$$\overrightarrow{D}_O'(y, x, g_y; y^0) = \lim_{\delta \rightarrow 0^-} \left\{ \frac{\overrightarrow{D}_O(y + \delta y^0, x, g_y) - \overrightarrow{D}_O(y, x, g_y)}{\delta} \right\} \quad (\text{A.13})$$

is a superlinear function of y^0 satisfying $\overrightarrow{D}_O'(y, x, g_y; 0) = 0$ and $-\overrightarrow{D}_O'(y, x, g_y; -y^0) \geq \overrightarrow{D}_O'(y, x, g_y; y^0)$. By the translation property in (A.11) and the definition of directional derivative in (A.13)

$$\begin{aligned} \overrightarrow{D}_O'(y, x, g_y; g_y) &= \lim_{\delta \rightarrow 0^+} \left\{ \frac{\overrightarrow{D}_O(y + \delta g_y, x, g_y) - \overrightarrow{D}_O(y, x, g_y)}{\delta} \right\} \\ &= \lim_{\delta \rightarrow 0^+} \left\{ \frac{\overrightarrow{D}_O(y, x, g_y) - \delta - \overrightarrow{D}_O(y, x, g_y)}{\delta} \right\} \\ &= -1 \end{aligned} \quad (\text{A.14})$$

²⁶The dual problem is originally a minimization problem. For expositional purposes, i.e. to present it in a more familiar (shadow) profit-maximization context, the dual has been rendered into a maximization problem by reversing the sign of the objective function and some of the constraints.

²⁷See Rockafellar (1970), Theorem 23.1.

Similarly,

$$\begin{aligned}
\overrightarrow{D}_O'(y + \beta g_y, x, g_y; y^0) &= \lim_{\delta \rightarrow 0^+} \left\{ \frac{\overrightarrow{D}_O(y + \beta g_y + \delta y^0, x, y^0) - \overrightarrow{D}_O(y + \beta g_y, x, g_y)}{\delta} \right\} \\
&= \lim_{\delta \rightarrow 0^+} \left\{ \frac{\overrightarrow{D}_O(y + \delta y^0, x, g_y) - \beta - \overrightarrow{D}_O(y, x, g_y) + \beta}{\delta} \right\} \\
&= \overrightarrow{D}_O'(y, x, g_y; y^0) \quad \forall \beta \in \mathbb{R}
\end{aligned} \tag{A.15}$$

The derivation in (A.15) simply implies that directional derivatives for directional distance functions are translation invariant with respect to the direction defining the directional distance function.

The superdifferential of \overrightarrow{D}_O in y , denoted as $\partial \overrightarrow{D}_O(y, x, g_y)$, is given by

$$\partial \overrightarrow{D}_O(y, x, g_y) = \left\{ v \in \mathbb{R}^N \mid \overrightarrow{D}_O(y, x, g_y) + v'(y^0 - y) \geq \overrightarrow{D}_O(y^0, x, g_y) \quad \forall y^0 \in \mathbb{R}^N \right\} \tag{A.16}$$

which can be also expressed as²⁸

$$\partial \overrightarrow{D}_O(y, x, g_y) = \left\{ v \mid v'y^0 \geq \overrightarrow{D}_O'(y, x, g_y; y^0) \quad \forall y^0 \right\} \tag{A.17}$$

or equivalently

$$\overrightarrow{D}_O'(y, x, g_y; y^0) = \inf \left\{ v'y^0 \mid v \in \partial \overrightarrow{D}_O(y, x, g_y) \right\} \tag{A.18}$$

Denoting $\nabla \overrightarrow{D}_O(y, x, g_y)$ as the gradient of \overrightarrow{D}_O in y and considering that when $\overrightarrow{D}_O(y, x, g_y)$ is differentiable in y $\overrightarrow{D}_O'(y, x, g_y; y^0)$ is the inner product of the gradient and y^0 , i.e. $\overrightarrow{D}_O'(y, x, g_y; y^0) = \nabla \overrightarrow{D}_O(y, x, g_y)'y^0$, it can be proven²⁹ that if $v \in \partial \overrightarrow{D}_O(y, x, g_y)$

$$v'g = 1 \tag{A.19}$$

$$v \in \partial \overrightarrow{D}_O(y + \beta g_y, x, g_y) \quad \forall \beta \in \mathbb{R} \tag{A.20}$$

These two mathematical results have important economic implications. First, the fact that the inner product of any element of the superdifferential $\partial \overrightarrow{D}_O(y, x, g_y)$ and g_y must be equal to one reflects the fact that $\partial \overrightarrow{D}_O(y, x, g_y)$ contains the shadow prices of the output bundle normalized

²⁸See Rockafellar (1970), Theorems 23.3 and 23.4.

²⁹See Lemma 1 and its proof in Chambers and Färe (2008).

by the shadow value of the numeraire bundle g . Second, not only directional derivatives but also superdifferentials of directional distance functions are translation invariant in the direction of g .

To see how the concepts of directional derivatives and superdifferentials allow for deriving shadow prices for extreme efficient units consider the revenue function associated with T for given output prices $p \in \mathbb{R}_+^N$

$$R(x, p) = \max\{p'y \mid y \in T\} \quad (\text{A.21})$$

As long as there exists a y such that $y + \beta g_y \in T$ for some β , by the representation property

$$\begin{aligned} R(x, p) &= \max \left\{ p'(y + \overrightarrow{D}_O(y, x, g_y)g_y) \right\} \\ &= \max \left\{ p'y + \overrightarrow{D}_O(y, x, g_y)p'g_y \right\} \end{aligned} \quad (\text{A.22})$$

Now take any solution to (A.22) and denote it as y^* which is the efficient level of output maximizing revenue. The directional derivative of (A.22) in an arbitrary direction y^0 away from y^* is given by

$$\begin{aligned} &\lim_{\delta \rightarrow 0^+} \left\{ \frac{p'(y^* + \delta y^0) + \overrightarrow{D}_O(y^* + \delta y^0, x, g_y)p'g_y - p'(y^* + \overrightarrow{D}_O(y^*, x, g_y)g_y)}{\delta} \right\} \\ &= p'y^0 + \overrightarrow{D}_O'(y^*, x, g_y; y^0)p'g_y \end{aligned} \quad (\text{A.23})$$

If y^* is optimal, the directional derivative $p'y^0 + \overrightarrow{D}_O'(y^*, x, g_y; y^0)p'g_y$ is non-positive in every possible direction so that

$$\frac{p'y^0}{p'g_y} \leq -\overrightarrow{D}_O'(y^*, x, g_y; y^0) \quad (\text{A.24})$$

which implies that $\frac{p}{p'g_y} \in \partial \overrightarrow{D}_O(y^*, x, g_y)$ for every y^0 . As mentioned before, this means that for efficient firms (those efficiently selecting y^*) $\partial \overrightarrow{D}_O(y^*, x, g_y)$ contains all the possible normalized shadow prices for \overrightarrow{D}_O at y^* .

Recall that $\overrightarrow{D}_O'(y^*, x, g_y; g_y) = -1$ by (A.14), then for $y^0 = g_y$ (A.23) becomes

$$p'y^0 + \overrightarrow{D}_O'(y^*, x, g_y; g_y)p'g_y = p'g_y - p'g_y = 0 \quad (\text{A.25})$$

Hence, translations of y^* in the direction of g_y do not have any impact on the objective

$R(x, p)$, thus if y^* solves the revenue maximization problem so does any translation of y^* in the direction of g_y . That is, for extreme efficient firms on the primal kinks of the technology there are multiple optimal solutions (i.e. any y^* and any translation of it in the direction of g_y) to the dual revenue maximization problem. This solution indeterminacy is simply solved by setting $D_O(y^*, x, g_y)$ to ensure that y^* is on the frontier of T .

Denote e_m as the m th element of the standard orthonormal basis and consider an increase in the production of y_m^* ³⁰ by one unit, which implies a movement from the efficient point y_m^* in the direction of e_m , then (A.24) becomes

$$\begin{aligned} \frac{p_m}{p'g_y} &\leq -\overrightarrow{D_O}'(y^*, x, g_y; e_m) \\ &= -\inf \left\{ v'(e_m) \mid v \in \partial \overrightarrow{D_O}(y^*, x, g_y) \right\} \\ &= -\inf \left\{ v_m \mid v \in \partial \overrightarrow{D_O}(y^*, x, g_y) \right\} \end{aligned} \quad (\text{A.26})$$

Therefore, any normalized price $\frac{p_m}{p'g_y}$ at which y^* is efficient is a lower bound for $-\overrightarrow{D_O}'(y^*, x, g_y; -e_m)$ implying that $-\overrightarrow{D_O}'(y^*, x, g_y; e_m)$ represents *willingness to gain*, i.e. a measure of what an extreme efficient firm would be willing to receive for engaging in the production of one extra unit of y_m . In the same fashion, considering a movement in the direction of $-e_m$, which is associated with holding off the production of one unit of y_m and forfeit the revenue from that unit, yields

$$\begin{aligned} \frac{-p_m}{p'g_y} &\leq -\overrightarrow{D_O}'(y^*, x, g_y; -e_m) \\ \frac{p_m}{p'g_y} &\geq \overrightarrow{D_O}'(y^*, x, g_y; -e_m) \\ &= \inf \left\{ v'(-e_m) \mid v \in \partial \overrightarrow{D_O}(y^*, x, g_y) \right\} \\ &= -\sup \left\{ v_m \mid v \in \partial \overrightarrow{D_O}(y^*, x, g_y) \right\} \end{aligned} \quad (\text{A.27})$$

which establishes that $\frac{p_m}{p'g_y}$ is an upper bound for $\overrightarrow{D_O}'(y^*, x, g_y; -e_m)$. Thus, $\overrightarrow{D_O}'(y^*, x, g_y; -e_m)$ can be interpreted as *willingness to lose*, i.e. a measure of what an extreme efficient firm would be willing to give up to forgo the production of one unit of y_m .

Since directional derivatives are positively linearly homogeneous and concave functions

³⁰In the empirical analysis I will focus specifically on CSR so $y_m = y_R$.

of y

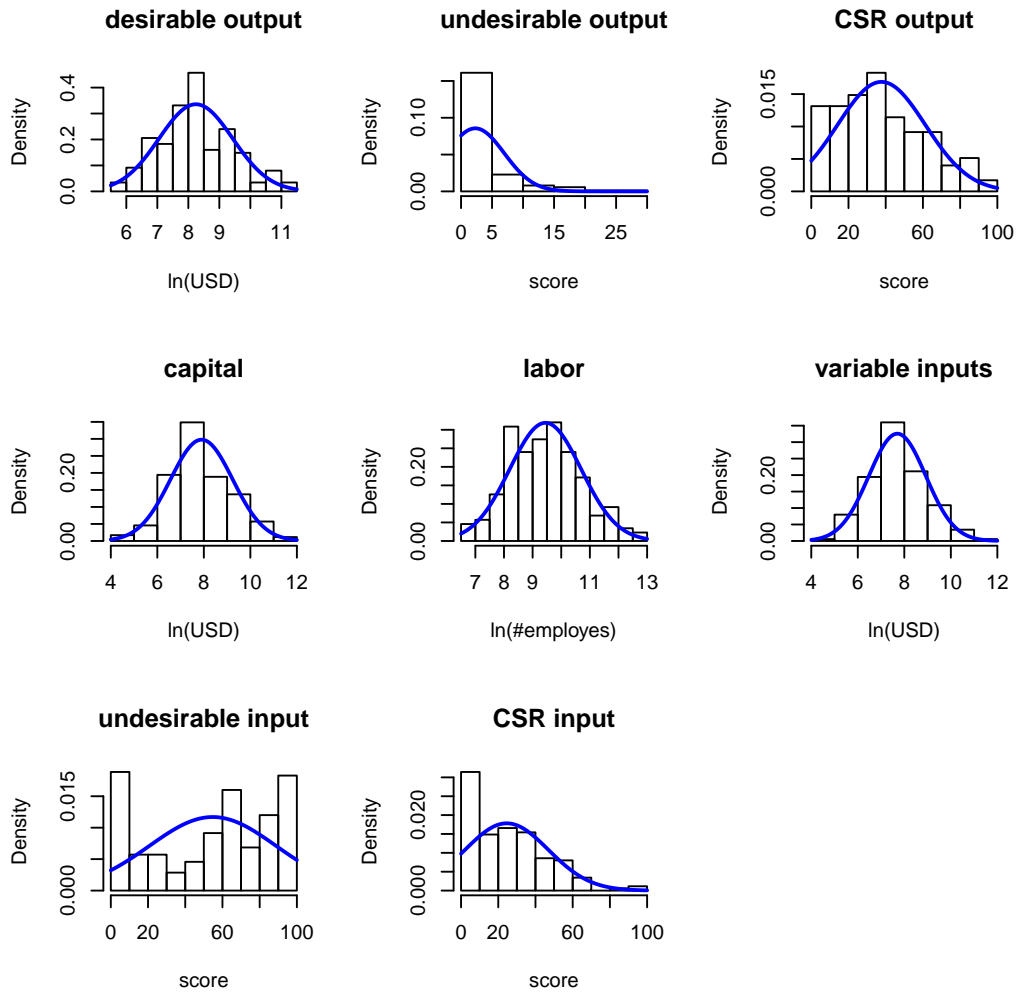
$$\overrightarrow{D}_O'(y^*, x, g_y; -e_m) \leq -\overrightarrow{D}_O'(y^*, x, g_y; e_m) \quad (\text{A.28})$$

which formally represents the gap between willingness to gain and willingness to lose generated by the non-smoothness of the technology. Intuitively, for a firm operating efficiently the marginal gain of producing one additional unit of y_m should be higher than the marginal loss of relinquishing one unit of it. Note that, even if there are potentially infinitely many (normalized) shadow prices for y_m , this approach allows for identifying the only two prices that are economically relevant: the gaining shadow price and the losing shadow price. At the kinks these two prices diverge but are still uniquely identified by $-\overrightarrow{D}_O'(y^*, x, g_y; e_m)$ and $\overrightarrow{D}_O'(y^*, x, g_y; -e_m)$, respectively.

It is important to remark that the interpretation of the directional derivatives $-\overrightarrow{D}_O'(\cdot; e_m)$ and $\overrightarrow{D}_O'(\cdot; -e_m)$ as willingness to gain and willingness to lose, respectively, applies only to extreme efficient units that are at the kinks of the technological frontier. For efficient firms that are on the technological frontier, but not at the kinks, willingness to gain and willingness to lose coincide as $-\overrightarrow{D}_O'(y^*, x, g_y; e_m) = \overrightarrow{D}_O'(y^*, x, g_y; -e_m)$. For inefficient firms the interpretation of $-\overrightarrow{D}_O'(y^*, x, g_y; e_m)$ is still insightful but different. Specifically, if y^* is not efficient, $-\overrightarrow{D}_O'(y^*, x, g_y; e_m)$ simply measures the change in the directional distance function resulting from a small move in the direction of e_m .

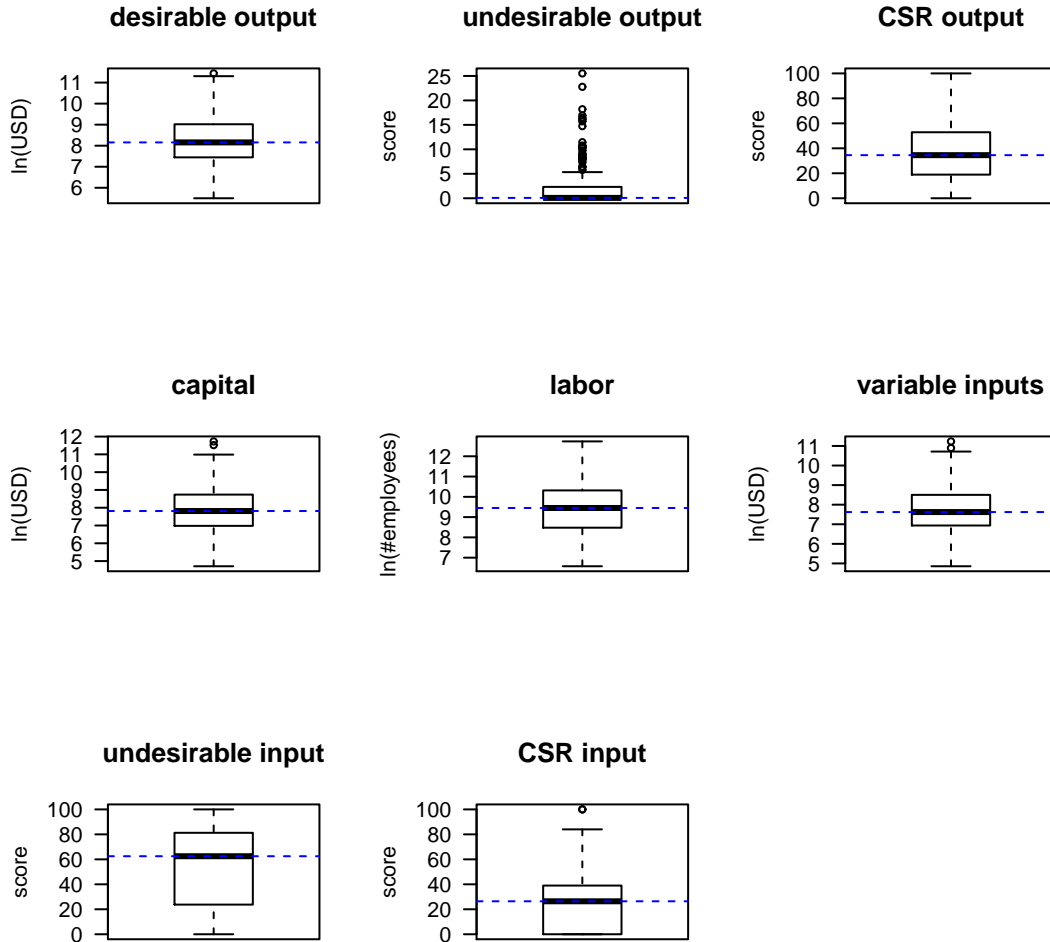
B Figures and Tables

Figure 1: Histogram of the variables used in the empirical analysis



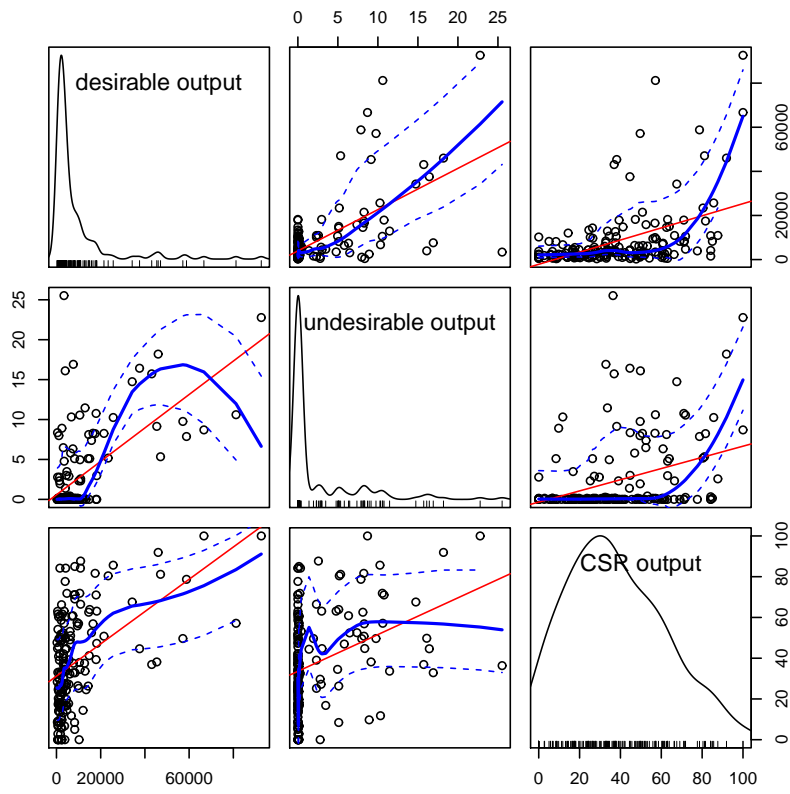
Sources: Orbis - Bureau van Dijk (desirable output, capital, labor, variable inputs), ThomsonOne - Thomson Reuters (labor), and Sustainalytics (undesirable output, CSR output, undesirable input, CSR input).

Figure 2: Box plots of the variables used in the empirical analysis



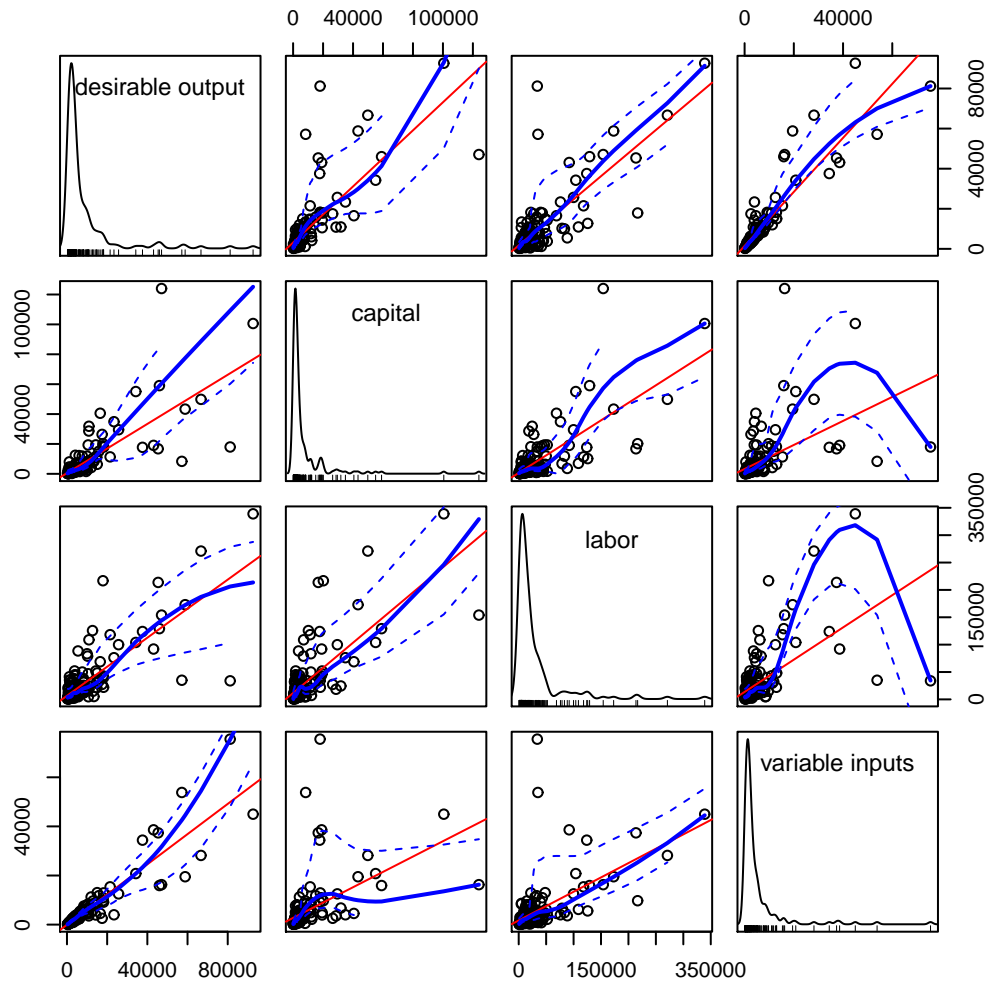
Sources: Orbis - Bureau van Dijk (desirable output, capital, labor, variable inputs), ThomsonOne - Thomson Reuters (labor), and Sustainalytics (undesirable output, CSR output, undesirable input, CSR input).

Figure 3: Scatter plot matrices of the desirable output with the other outputs



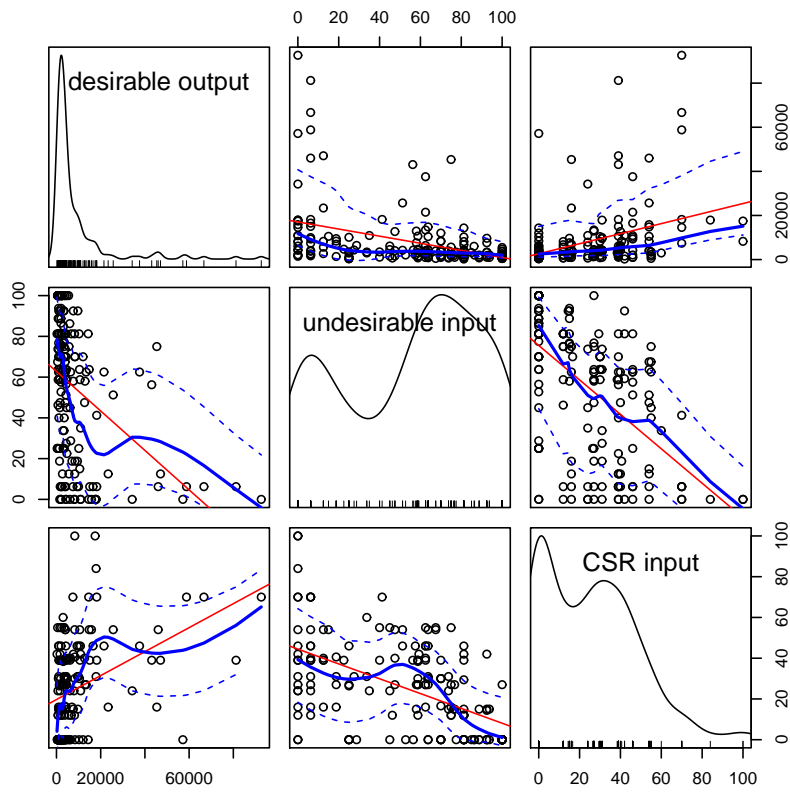
Sources: Orbis - Bureau van Dijk (desirable output), and Sustainalytics (undesirable output, CSR output, undesirable input, CSR input).

Figure 4: Scatter plot matrices of the desirable output with the conventional production inputs



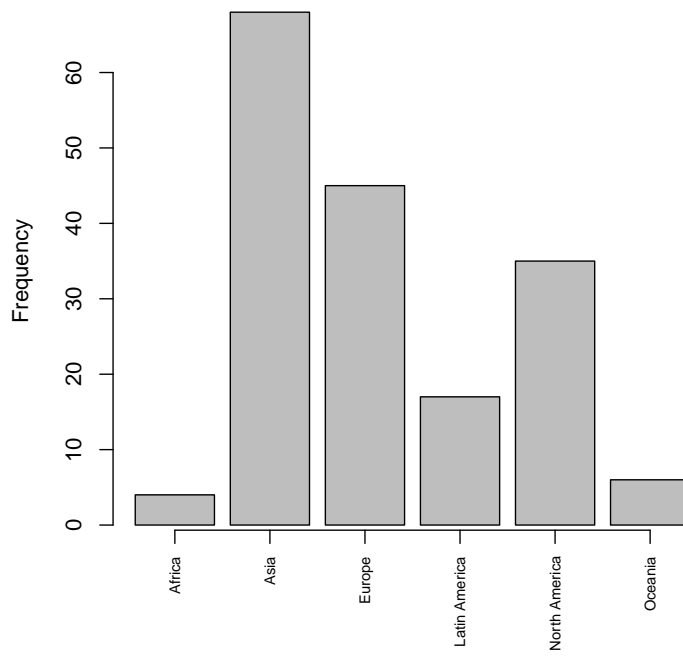
Sources: Orbis - Bureau van Dijk (desirable output, capital, labor, variable inputs), ThomsonOne - Thomson Reuters (labor).

Figure 5: Scatter plot matrices of the desirable output with the other production inputs



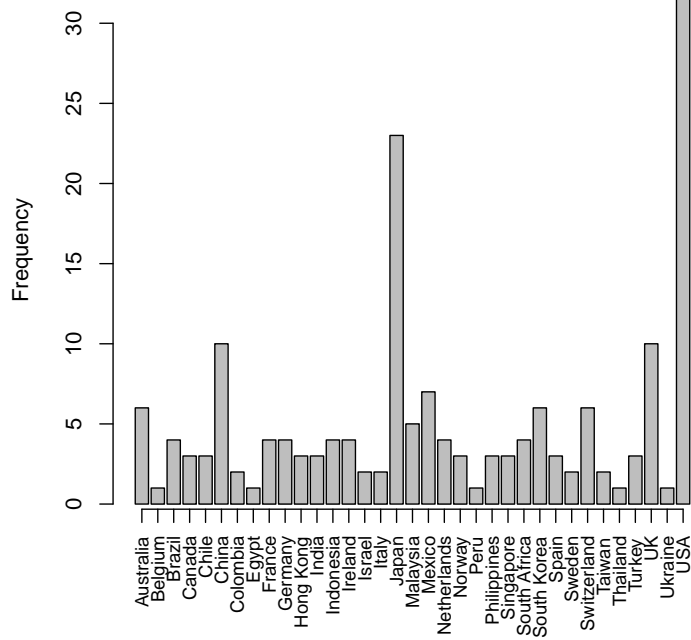
Sources: Orbis - Bureau van Dijk (desirable output), and Sustainalytics (undesirable input, CSR input, undesirable input, CSR input).

Figure 6: Distribution of firms by geographical area



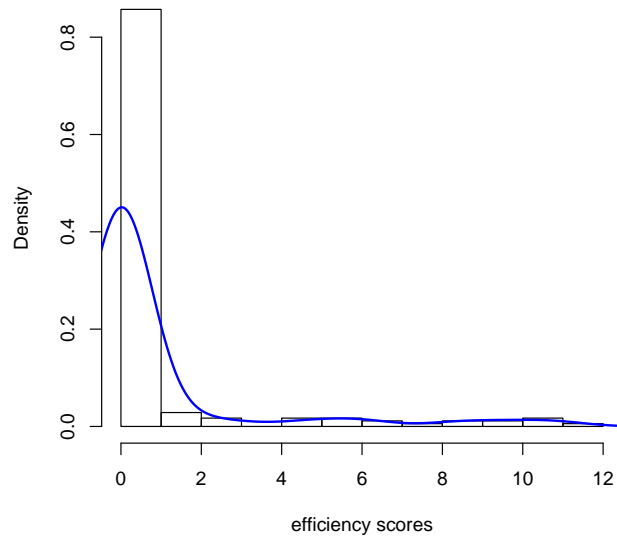
Source: Based on the geographical location of the firms included in the analysis.

Figure 7: Distribution of firms by country



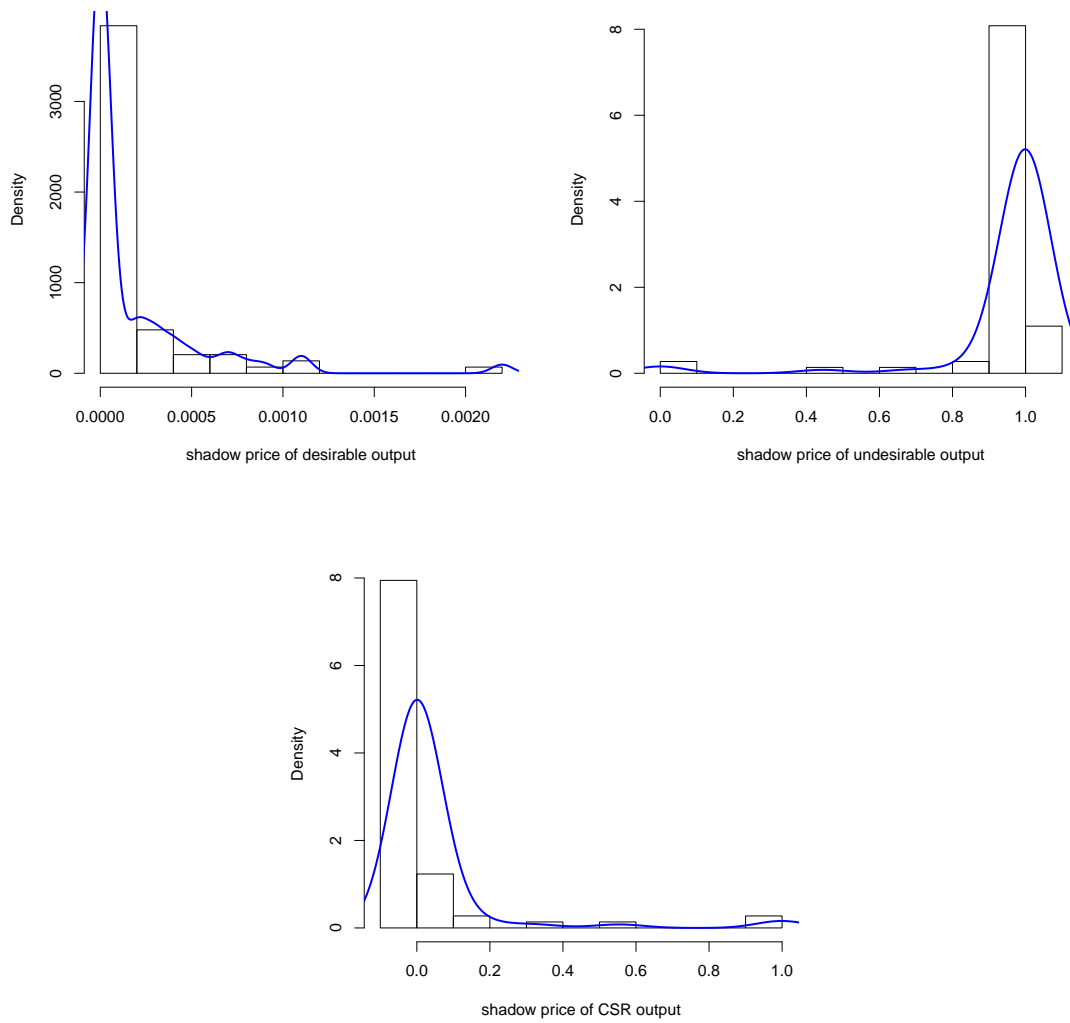
Souce: Based on the geographical location of the firms included in the analysis.

Figure 8: Distribution of efficiency scores



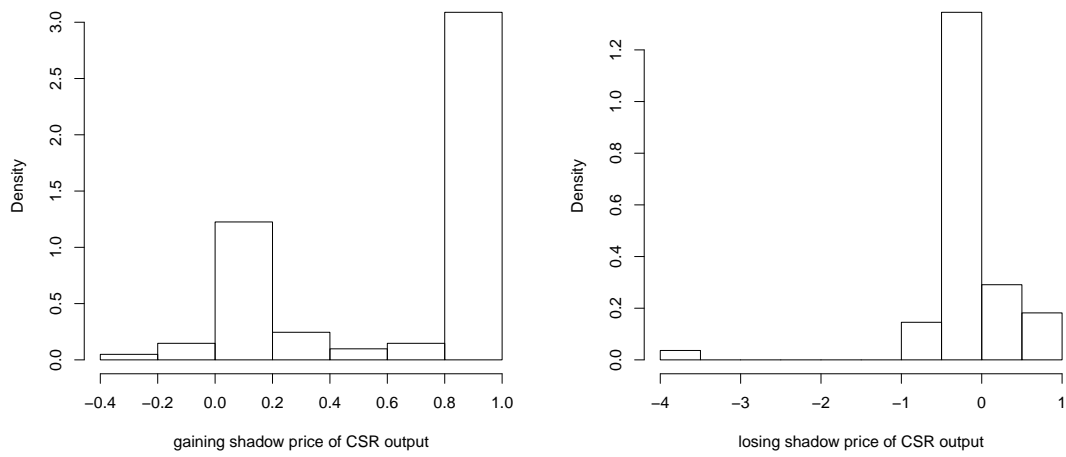
Source: Estimation results of the primal problem.

Figure 9: Distribution of outputs shadow prices for inefficient and just efficient firms



Souce: Estimation results of the dual problem.

Figure 10: Distribution of gaining and losing shadow prices of CSR for extreme efficient firms



Source: Estimation results of the modified dual problem for extreme efficient firms.

Table 1: List of variables included in the empirical analysis

| Variable | Description | Indicator | Source |
|-----------|-----------------------------|--|--------------------|
| y_D | desirable output | Sales | Orbis |
| y_U | undesirable output | Operations Incidents | Sustainalytics - E |
| | | Environmental Supply Chain Incidents | Sustainalytics - E |
| | | Product and Services Incidents | Sustainalytics - E |
| | | Employee Incidents | Sustainalytics - S |
| | | Social Supply Chain Incidents | Sustainalytics - S |
| | | Product and Services Incidents | Sustainalytics - S |
| | | Society and Community Incidents | Sustainalytics - S |
| | | Business Ethics Incidents | Sustainalytics - G |
| y_R | socially responsible output | Governance Incidents | Sustainalytics - G |
| | | Public Policy Incidents | Sustainalytics - G |
| | | Environmental Policy | Sustainalytics - E |
| | | Environmental Management System | Sustainalytics - E |
| | | Sustainable Agriculture Programs | Sustainalytics - E |
| | | Freedom of Association Policy | Sustainalytics - S |
| | | Discriminatory Policy | Sustainalytics - S |
| | | Supply Chain Monitoring | Sustainalytics - S |
| x_{1_k} | conventional input | Bribery and Corruption Policy | Sustainalytics - G |
| | | Global Compact Signatory | Sustainalytics - G |
| | | Board Independence | Sustainalytics - G |
| x_{1_l} | conventional input | Labor - Number of Employees | Orbis/ThomsonOne |
| x_{1_v} | conventional input | Variable Inputs - Cost of Goods Sold | Orbis |
| x_2 | undesirable input | Water Management Programs | Sustainalytics - E |
| | | GHG Reduction Programs | Sustainalytics - E |
| | | Scope of Social Supply Chain Standards | Sustainalytics - S |
| x_3 | socially responsible input | Green Procurement Policy | Sustainalytics - E |
| | | Diversity Programs | Sustainalytics - S |

Note: The scores for the indicators used to construct measures of undesirable output and undesirable input have been transformed as 100 minus the original score to be consistent with the theoretical framework. Sustainalytics - E, S or G signifies that the indicator comes from either the E (Environment), S (Social) or G (Governance) category, in which the indicators in the Sustainalytics database are organized.

Table 2: Weights assigned to each component used to construct the variables y_U , y_R , x_2 , x_3

| Variable | Description | Indicator | Weight |
|----------|-----------------------------|--|--------|
| y_U | undesirable output | Operations Incidents | 16.71 |
| | | Environmental Supply Chain Incidents | 6.96 |
| | | Product and Services Incidents | 4.64 |
| | | Employee Incidents | 11.60 |
| | | Social Supply Chain Incidents | 11.60 |
| | | Product and Services Incidents | 13.92 |
| | | Society and Community Incidents | 6.69 |
| | | Business Ethics Incidents | 10.90 |
| | | Governance Incidents | 9.74 |
| | | Public Policy Incidents | 6.69 |
| | | | 100.00 |
| y_R | socially responsible output | Environmental Policy | 10.76 |
| | | Environmental Management System | 14.35 |
| | | Sustainable Agriculture Programs | 13.45 |
| | | Freedom of Association Policy | 13.45 |
| | | Discriminatory Policy | 13.45 |
| | | Supply Chain Monitoring | 13.45 |
| | | Bribery and Corruption Policy | 4.48 |
| | | Global Compact Signatory | 4.48 |
| | | Board Independence | 12.11 |
| | | | 100.00 |
| x_2 | undesirable input | Water Management Programs | 45.00 |
| | | GHG Reduction Programs | 30.00 |
| | | Scope of Social Supply Chain Standards | 25.00 |
| | | | 100.00 |
| x_3 | socially responsible input | Green Procurement Policy | 40.00 |
| | | Diversity Programs | 60.00 |
| | | | 100.00 |

Note: The weights have been re-scaled to reflect the relative importance given to these indicators in the original Sustainability dataset.

Table 3: Descriptive statistics of the variables used in the empirical analysis

| Variable | Description | Mean | Median | Trimmed Mean 10% | Standard deviation | MAD |
|----------|--|----------|----------|------------------|--------------------|---------|
| y_D | desirable output - sales (USD) | 8233.86 | 3486.88 | 3557.63 | 13921.03 | 2318.59 |
| y_U | undesirable output (score) | 2.30 | 0.07 | 0.22 | 4.66 | 0.07 |
| y_R | socially responsible output (score) | 37.71 | 34.51 | 35.41 | 23.63 | 17.12 |
| x_{1k} | capital - fixed assets (USD) | 7278.93 | 2474.88 | 2497.23 | 15091.47 | 1605.00 |
| x_{1l} | labor - number of employees | 28394.50 | 12700.00 | 12433.58 | 47236.23 | 8490.00 |
| x_{1v} | variable inputs - cost of goods sold (USD) | 5010.73 | 2046.09 | 2224.33 | 9406.09 | 1366.61 |
| x_2 | undesirable input (score) | 54.78 | 62.50 | 56.43 | 34.12 | 27.50 |
| x_3 | socially responsible input (score) | 24.54 | 26.40 | 20.05 | 22.38 | 19.60 |

Note: The variables expressed in USD are in million of USD; whenever the values were expressed in other currencies they have been converted into USD using the Dec 31 2014 exchange rate provided by the IMF. The variables expressed in scores are expressed on a potential scale from 0 to 100.

Table 4: Descriptive statistics of the outputs shadow values

| Variable | Mean | Median | St. dev | Min | Max |
|----------|----------|----------|---------|----------|---------|
| p_D | 0.00017 | 0.00000 | 0.00036 | 0.00000 | 0.00220 |
| p_U | -0.95530 | -1.00000 | 0.17996 | -1.05970 | 0.00000 |
| p_R | 0.04456 | 0.00000 | 0.17971 | -0.06070 | 0.99930 |

Note: The statistics are calculated for the subsample of 43 inefficient and 30 just efficient firms. The statistics relative to the shadow price of the undesirable output p_U are presented accounting for the fact that this shadow price enters the objective function and some of the constraints with a negative sign.

Table 5: Descriptive statistics of the marginal value of CSR

| Variable | Mean | Median | St. dev | Min | Max |
|-------------------|----------|----------|---------|----------|---------|
| p_R upper bound | 0.66563 | 0.99550 | 0.43619 | -0.21440 | 1.00000 |
| p_R lower bound | -0.07141 | -0.00380 | 0.55956 | -3.52720 | 0.78010 |

Note: The statistics for the upper bound of p_R are calculated for all the 102 extreme efficient firms present in the sample while the statistics for the lower bound p_R are calculated for a subsample of 55 extreme efficient firms.