Assessing Corporate Social Responsibility Efficiency for the International Food and Beverage Manufacturing Industry

Magdalena Kapelko

Wroclaw University of Economics and Business, Department of Logistics, Poland
e-mail: magdalena.kapelko@ue.wroc.pl

ABSTRACT

One of the major challenges in the research on corporate social responsibility (CSR) is the aggregation of the CSR metrics into overall measures of CSR practices by firms. This paper computes composite indicators of CSR from an efficiency perspective using data envelopment analysis (DEA) for a sample of international food and beverage manufacturing firms over the period 2011-2018. The study's contributions to the literature are twofold. First, this paper contributes by being the first to compare efficiency in CSR practices of food and beverage companies across regions of Europe, the United States and Canada, Latin America and Caribbean and Asia-Pacific. Second, methodologically we extend the composite indicators within DEA, allowing for non-convexities of the production set. The study finds a considerable potential for improvement in CSR practices as revealed by the values of CSR composite indicators. The study also shows the differences in CSR efficiency between food and beverage firms in the regions considered, with the most CSR efficient region being Latin America and Caribbean, and the least CSR efficient being firms in Asia-Pacific region. The CSR composite scores fluctuate over the analyzed period, with an increase in efficiency in 2018 experienced by all regions.

Keywords: Corporate social responsibility; data envelopment analysis; benefit of the doubt; non-convexity; food and beverage manufacturing

1 Introduction

Corporate social responsibility (CSR) relates to firms’ actions that promote social good beyond economic interest of the firm or shareholders and beyond legal requirements (McWilliams and Siegel, 2001). In the food and beverage manufacturing industry, CSR engagement is increasing in importance, because of this sector’s strong impact on the economy, the environment, and the society (Hartmann, 2011). In particular, food and beverage industry faces CSR challenges related with, for example, food safety scares which made consumers more concerned with food safety, consumer awareness of the links between food and health and of the responsible consumption (impact on climate change, animal welfare or social and economic inequality) (European Commission, 2016).

CSR encompasses a broad range of activities, including its main dimensions of social, environmental and governance CSR engagement, therefore multiple metrics are required in order to reflect its full scope. One of the major challenges in the research on CSR is the aggregation of these metrics into overall measures of CSR practices by firms. The aggregation of CSR metrics into overall measures of CSR performance can be undertaken using different methods, including Data Envelopment Analysis (DEA) (Banker et al., 1984; Charnes et al., 1978). The use of DEA in CSR research was pioneered by Bendheim et al. (1998), who applied the DEA model without inputs (Lovell and Pastor, 1999, 1997) to assess the efficiency of CSR with respect to five key stakeholders domains using US data in light manufacturing, consumer products, primary industry, service, heavy industry and transportation. Later, Chen and Delmas (2011) computed CSR indicators based on the benefit of the doubt (BoD) DEA model (Cherchye et al., 2007a,b; Cherchye et
DEA is a methodology based on mathematical programming for the measurement of efficiency of decision making units (DMUs) that convert multiple inputs into multiple outputs. It constructs a piece-wise frontier enveloping the data and determines the distance of each DMU to this frontier. In the BoD DEA model, the aggregated subindicators are all treated as outputs to generate an overall and objective aggregated indicator for each analyzed unit through the determination of its efficiency. In the literature many different methodological extensions of the BoD model have been developed. The examples include, the non-radial model with slacks (Sahoo and Acharya, 2012), the directional model that recognizes the preference structure among indicators (Fusco, 2015), the directional model that accounts for undesirable output indicators (Zanella et al., 2015), the robust and non-compensatory composite indicator based on the directional model (Vidoli et al., 2015), the robust BoD model that considers external factors directly (Fusco et al., 2018), a translation invariant directional distance function model (Aparicio and Kapelko, 2019), and a model assuming a least distance to the frontier (Aparicio et al., 2020). All of these studies assumed convexity which means that points on the frontier used to evaluate observations are constructed based on linear combinations of actual observations, and not on actual observations themselves. However, such assumption is often unrealistic since some of such observations can never actually be realized, and recent evidence shows that production set is often non-convex (e.g., Kerstens et al., 2019; Wilson, 2021). The direct empirical implication is non-convexity and the free disposal hull (FDH) framework as the approach to data envelopment. FDH approach was introduced by Deprins et al. (1984) with an application to post offices. Tulkens (1993) presented an early overview with applications to several sectors, while Cherchye et al. (2000; 2001) presented additional theoretical refinements with an illustration to banking. However, to the best of our knowledge, FDH approach has so far never been considered in the context of composite indicators created by BoD.

In this study, we aim to fill in the gap in the literature outlined above and we construct composite measures of CSR from an efficiency perspective using DEA and BoD model with non-convex structure for a sample of international food and beverage manufacturing firms. In this way we focus on actual data points in deriving composite indicators measures. The study contributes to the literature in two ways. First, this paper is the first to analyze and compare efficiency in CSR practices by food and beverage companies across regions of Europe, the United States and Canada, Latin America and Caribbean and Asia-Pacific. Second, methodologically we extend the measurement of composite indicators within DEA, allowing for non-convexities of the production set.

The rest of this paper is organized as follows. The next section describes the methodology to compute composite indicators using DEA. The section to follow provides information on the dataset, followed by the description and interpretation of the results. The final section concludes.

2 Methods

To aggregate CSR scores across three dimensions of environmental, social and governance CSR, and measure efficiency in CSR we used BoD model or the model without inputs (Cherchye et al., 2004; Lovell and Pastor, 1999, 1997) within DEA framework and FDH approach.

The initial BoD model is a DEA model without inputs, where subindicators are all treated as outputs to generate an overall and objective aggregated indicator for each assessed firm. The usage of DEA provides

---

an advantage in that it avoids any possible controversies related to the selection of weights for each specific indicators since weighting is objective. The linear programming model to construct composite indicator is given below (Cherchye et al., 2004):

\[
\begin{align*}
\text{Cl}_j &= \max \sum_{r=1}^{s} w_r y_{rj} \\
\text{s.t.} & \sum_{r=1}^{s} w_r y_{rj} \leq 1 \quad j = 1,\ldots,n, \\
& w_r \geq 0 \quad r = 1,\ldots,s.
\end{align*}
\] (1)

In model (1), \( y_{rj} \) corresponds to the value of the output indicator \( r \) \((r = 1,\ldots,s)\) in company \( j \) \((j = 1,\ldots,n)\), in our case these are environmental, social and governance CSR scores, while \( w_r \) \((r = 1,\ldots,s)\) indicates weights.

Van Puyenbroeck (2018) shows that BoD model given by (1) is a reciprocal of the output-oriented variable returns to scale (VRS) model without inputs of Lovell and Pastor (1999, 1997), given by:

\[
\begin{align*}
\text{I}_j &= \max \theta \\
\text{s.t.} & \sum_{j=1}^{n} \lambda_j y_{rj} \geq \theta y_{rj} \quad r = 1,\ldots,s, \\
& \sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0 \quad j = 1,\ldots,n.
\end{align*}
\] (2)

Therefore:

\[
\text{CI}_j = \frac{1}{\text{I}_j}.
\]

In order to relax the convexity assumption in (2) and develop output-oriented FDH model without inputs, the additional constraint needs to be added to model (2) that restricts the weights to be bivalent:

\[
\begin{align*}
\text{IF}_j &= \max \theta \\
\text{s.t.} & \sum_{j=1}^{n} \lambda_j y_{rj} \geq \theta y_{rj} \quad r = 1,\ldots,s, \\
& \sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \in \{0,1\}.
\end{align*}
\] (3)

Hence, the FDH BoD model is a reciprocal of the output-oriented FDH model without inputs:

\[
\text{CIF}_j = \frac{1}{\text{IF}_j}.
\]
To solve the model (3), the mixed-integer programming is used.

Kneip et al. (2016) developed a statistical test to assess the convexity in DEA based on the single sample split. Simar and Wilson (2020) provided further improvement in this test in the form of multiple sample splits, which remove ambiguity surrounding the choice of a single split. In both cases the null hypothesis of convexity (that is, that both FDH and DEA estimators are consistent) is tested against an alternative one (only FDH estimator is consistent). In the paper we applied the test by Simar and Wilson (2020) in order to check if for our data only proposed FDH estimator is consistent.

The composite indicators can be estimated with regard to all firms in the sample, regardless of the group they belong to, measuring so called metafrontier efficiency (Battese et al., 2004; O'Donnell et al., 2008). But they can also be calculated with reference to this specific group, leading to the measurement of group-specific or managerial efficiency (Battese et al., 2004; O'Donnell et al., 2008; Charnes et al., 1981). Metafrontier indicator can be written as the product of group-specific indicator and a residual measure of metatechnology ratio or program efficiency (Battese et al., 2004; O'Donnell et al., 2008; Charnes et al., 1981). Given that in our empirical application the group participation is established based on the region where each firm has headquarters, and given that the composite indicator concerns CSR, the decomposition used in this paper can be written as follows:

\[
\text{Metafrontier CSR indicator} = \text{Region-specific CSR indicator} \cdot \text{Regional CSR gap}
\] (4)

Metafrontier CSR indicator allows for the correct comparison of CSR efficiency between regions. Region-specific CSR indicator measures how close a firm is operating to region-specific frontier, and concerns inefficiency due to shortcomings in managerial practices with regard to CSR. Regional gap assesses how close is region-specific frontier to the metafrontier with regard to CSR.

3 Dataset

Our study is based on CSR data obtained from Sustainalytics, a global leader in environmental, social and governance research and ratings (see www.sustainalytics.com). In contrast to most other CSR data providers, Sustainalytics does not restrict itself to a specific geographic region, hence it is suitable for the purpose of this study to compare firms in different regions. Sustainalytics dataset has largely been used in recent research (for example, Engida et al., 2018; Auer, 2016; Kim et al., 2016). Firms in the food and beverage manufacturing industry were distinguished using the industry classification of Sustainalytics, which is based on Global Industry Classification Standard (GICS), in which this industry belongs to the wider sector of consumer staples (in which remaining industry groups are food and staples retailing, and household and personal products). The information on CSR indicators in Sustainalytics comes from multiple sources such as annual reports, CSR reports, CSR websites, press releases, local newspapers or relevant websites (Auer, 2016).

The Sustainalytics dataset includes scores on different indicators that capture three dimensions of CSR: environmental (for example, involvement in recycling, waste reduction and renewable energies), social (for example, employee profit sharing, product safety, employment of minorities and charitable donations) and governance (for example, board independence and shareholder rights). The data availability in Sustainalytics allows us to analyze the 2011-2018 period covering the food and beverage companies from Europe, the United States and Canada, Latin American and Caribbean and the Asia-Pacific region.

Table 1 reports basic descriptive statistics of CSR scores (averages and standard deviations) for our sample sub-divided by geographic regions.

---

2 Due to the data limitation of Sustainalytics it was impossible to undertake analysis by country. Europe includes Austria, Belgium, Cyprus, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Slovenia, Spain, Sweden, Switzerland, Ukraine and UK. Asia-Pacific contains Australia, China, Hong Kong, India, Indonesia, Japan, Malaysia, New Zealand, Papua New Guinea, Philippines, Singapore, South Korea, Sri Lanka, Taiwan, Thailand, and Vietnam. Latin America and Caribbean includes Bahamas, Bermuda, Brazil, Cayman Islands, Chile, Colombia, Guatemala, Mexico and Peru. Such regional classification is consistent with Sustainalytics and previous research (e.g., Auer, 2016).
### Table 1.
Arithmetic means and standard deviations of CSR variables, 2011-2018

<table>
<thead>
<tr>
<th>Region</th>
<th>Social</th>
<th>Environmental</th>
<th>Governance</th>
<th>No of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>60.270</td>
<td>60.488</td>
<td>65.907</td>
<td>416</td>
</tr>
<tr>
<td></td>
<td>(10.404)</td>
<td>(12.653)</td>
<td>(9.736)</td>
<td></td>
</tr>
<tr>
<td>United States and Canada</td>
<td>56.593</td>
<td>60.594</td>
<td>68.287</td>
<td>345</td>
</tr>
<tr>
<td></td>
<td>(9.562)</td>
<td>(12.278)</td>
<td>(8.224)</td>
<td></td>
</tr>
<tr>
<td>Latin America and Caribbean</td>
<td>62.166</td>
<td>61.429</td>
<td>69.896</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td>(9.135)</td>
<td>(10.566)</td>
<td>(9.752)</td>
<td></td>
</tr>
<tr>
<td>Asia-Pacific</td>
<td>51.326</td>
<td>47.445</td>
<td>58.424</td>
<td>704</td>
</tr>
<tr>
<td></td>
<td>(7.164)</td>
<td>(11.884)</td>
<td>(8.492)</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>55.774</td>
<td>54.924</td>
<td>63.536</td>
<td>1618</td>
</tr>
<tr>
<td></td>
<td>(9.758)</td>
<td>(13.720)</td>
<td>(10.024)</td>
<td></td>
</tr>
</tbody>
</table>

### 4 Results and Discussion

First of all, we tested for convexity using the test of Simar and Wilson (2020) with 10 sample splits and 1000 bootstrap replications. Because Kerstens et al. (2019) questioned convexity assumption in metafrontier analysis, hence, in Table 2 we summarize the results of convexity test for metafrontier CSR indicator per year.

<table>
<thead>
<tr>
<th>Year</th>
<th>Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>-1.066</td>
<td>0.552</td>
</tr>
<tr>
<td>2012</td>
<td>0.085</td>
<td>0.078</td>
</tr>
<tr>
<td>2013</td>
<td>0.268</td>
<td>0.020</td>
</tr>
<tr>
<td>2014</td>
<td>-0.098</td>
<td>0.258</td>
</tr>
<tr>
<td>2015</td>
<td>-0.944</td>
<td>0.576</td>
</tr>
<tr>
<td>2016</td>
<td>0.669</td>
<td>0.010</td>
</tr>
<tr>
<td>2017</td>
<td>-1.417</td>
<td>0.608</td>
</tr>
<tr>
<td>2018</td>
<td>0.941</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 2 reveals that the test does not reject convexity in 2011, 2014, 2015 and 2017, but it rejects convexity in 2012, 2013, 2016 and 2018. Hence, results provide substantial evidence of non-convexity. Moreover, not rejecting null hypothesis does not mean that null hypothesis is true. Also FDH estimator is consistent (Wilson, 2021). Therefore, our dataset seems to provide evidence of some non-convexity, hence it is suitable for the estimations of non-convex CSR indicator to reach the purpose of this paper.

Table 3 presents the geometric average values of metafrontier CSR composite indicators over the 2011-2018 period and for each of the years of this period, for food and beverage manufacturing firms in each of the regions considered (Europe, the United States and Canada, Latin America and Caribbean, and the Asia-Pacific). The use of geometric averages assures that the multiplicative decomposition shown by (4) holds exactly (Kerstens et al., 2019).
Table 3.
CSR composite indicators, firms’ geometric means across years and regions, metafrontier

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>0.820</td>
<td>0.847</td>
<td>0.827</td>
<td>0.846</td>
<td>0.815</td>
<td>0.808</td>
<td>0.800</td>
<td>0.875</td>
<td>0.827</td>
</tr>
<tr>
<td>United States and Canada</td>
<td>0.843</td>
<td>0.869</td>
<td>0.840</td>
<td>0.848</td>
<td>0.840</td>
<td>0.808</td>
<td>0.781</td>
<td>0.827</td>
<td>0.830</td>
</tr>
<tr>
<td>Latin America and Caribbean</td>
<td>0.861</td>
<td>0.854</td>
<td>0.852</td>
<td>0.869</td>
<td>0.855</td>
<td>0.860</td>
<td>0.815</td>
<td>0.889</td>
<td>0.855</td>
</tr>
<tr>
<td>Asia-Pacific</td>
<td>0.739</td>
<td>0.741</td>
<td>0.720</td>
<td>0.727</td>
<td>0.697</td>
<td>0.691</td>
<td>0.692</td>
<td>0.739</td>
<td>0.710</td>
</tr>
<tr>
<td>Overall</td>
<td>0.800</td>
<td>0.813</td>
<td>0.789</td>
<td>0.798</td>
<td>0.755</td>
<td>0.750</td>
<td>0.754</td>
<td>0.811</td>
<td>0.777</td>
</tr>
</tbody>
</table>

The results in table show that across all years and regions the value of CSR metafrontier composite indicator, on average, reaches the level of 0.777, which indicates that firms could, on average, increase all of their CSR outputs by approximately 22%. Furthermore, the differences in CSR efficiency are observed between regions. In particular, the results show that over the period 2011-2018 the most CSR efficient region is Latin America and Caribbean (score of 0.855), while the least CSR efficient are food and beverage firms in Asia-Pacific region (score of 0.710). Finally, the CSR scores fluctuate over the analyzed period, with an increase in efficiency in 2018 experienced by all regions.

Tables 4 and 5 present the decomposition components of metafrontier CSR indicator that is region-specific CSR indicator and regional gap, for each year and region, and in overall terms.

Table 4.
CSR composite indicators, firms’ geometric means across years and regions, region-specific frontier

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>0.824</td>
<td>0.857</td>
<td>0.873</td>
<td>0.854</td>
<td>0.819</td>
<td>0.810</td>
<td>0.800</td>
<td>0.882</td>
<td>0.835</td>
</tr>
<tr>
<td>United States and Canada</td>
<td>0.884</td>
<td>0.890</td>
<td>0.862</td>
<td>0.883</td>
<td>0.884</td>
<td>0.873</td>
<td>0.837</td>
<td>0.827</td>
<td>0.866</td>
</tr>
<tr>
<td>Latin America and Caribbean</td>
<td>0.912</td>
<td>0.880</td>
<td>0.852</td>
<td>0.878</td>
<td>0.867</td>
<td>0.884</td>
<td>0.885</td>
<td>0.932</td>
<td>0.881</td>
</tr>
<tr>
<td>Asia-Pacific</td>
<td>0.785</td>
<td>0.789</td>
<td>0.822</td>
<td>0.830</td>
<td>0.808</td>
<td>0.824</td>
<td>0.825</td>
<td>0.834</td>
<td>0.817</td>
</tr>
<tr>
<td>Overall</td>
<td>0.835</td>
<td>0.843</td>
<td>0.848</td>
<td>0.853</td>
<td>0.826</td>
<td>0.833</td>
<td>0.824</td>
<td>0.855</td>
<td>0.838</td>
</tr>
</tbody>
</table>

The results in Table 4 show that, similarly to metafrontier results, when firms are assessed with regard to own region frontier, the most CSR efficient region is Latin America and Caribbean (average indicator of 0.881), and the least efficient are Asian-Pacific firms (average indicator of 0.817). Therefore, given the regional frontier, firms in Latin America and Caribbean could increase its CSR outputs by approximately 12%, while in Asia-Pacific the potential to increase CSR is of approximately 18%. That implies that Latin America and Caribbean firms obtain the best results in the management of CSR initiatives, while Asian-Pacific firms have the highest levels of inefficiency related with shortcomings in managerial practices pertaining to CSR.
The results for regional gap in Table 5 show that the average gap between the levels of CSR efficiency on the frontier of each region and the metafrontier is the widest in the case of Asian Pacific firms, followed by the United States and Canada, Latin American and Caribbean and finally European firms. These results reveal that the average regional gap is the highest for Europe of 0.990 and the lowest in Asia Pacific of 0.869. In other words, the highest values of program efficiency in CSR are achieved by European firms, and the lowest by Asian Pacific firms. This means that the maximum CSR output that could be produced in Europe is about 99% of the maximum CSR output that could be obtained using metatechnology. The maximum CSR output using Asian-Pacific technology is only approximately 87% of this output that could be achieved by metafrontier. Worth noting is regional country gap for Europe in 2017 equal to 1. This indicates that for this year firms in this country obtained CSR output that placed them on the point of tangency between their region-specific frontier and the metafrontier.

To be able to assess the differences in CSR indicators between regions with statistical precision, we run Simar and Zelenyuk (2006) test\(^6\). The results of this test are presented in Table 6. The results of the tests show that there are significant differences in inefficiencies between regions, for metafrontier CSR indicators, region-specific CSR indicators and regional gap.

<table>
<thead>
<tr>
<th>Region</th>
<th>Metafrontier Statistic</th>
<th>Region-specific Statistic</th>
<th>Regional gap Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe – United States and Canada</td>
<td>8.479</td>
<td>2.22 x 10^-16</td>
<td>8.354</td>
<td>2.22 x 10^-16</td>
</tr>
<tr>
<td>Europe – Latin America and Caribbean</td>
<td>0.623</td>
<td>9.82 x 10^-2</td>
<td>4.562</td>
<td>4.0 x 10^-4</td>
</tr>
<tr>
<td>Europe – Asia-Pacific</td>
<td>88.436</td>
<td>2.22 x 10^-16</td>
<td>24.428</td>
<td>2.22 x 10^-16</td>
</tr>
<tr>
<td>United States and Canada – Latin America and Caribbean</td>
<td>3.202</td>
<td>9.40 x 10^-3</td>
<td>3.024</td>
<td>3.12 x 10^-2</td>
</tr>
<tr>
<td>United States and Canada – Asia-Pacific</td>
<td>99.231</td>
<td>2.22 x 10^-16</td>
<td>34.031</td>
<td>2.22 x 10^-16</td>
</tr>
<tr>
<td>Latin America and Caribbean – Asia-Pacific</td>
<td>59.987</td>
<td>2.22 x 10^-16</td>
<td>28.447</td>
<td>2.22 x 10^-16</td>
</tr>
</tbody>
</table>

5 Conclusions

The paper analyzed the performance of food and beverage manufacturing firms with regard to their CSR practices over the period 2011-2018. To differentiate from previous research, we analyzed CSR engagement of food and beverage firms worldwide, represented by the regions of Europe, the United States and Canada, Latin America and the Asia-Pacific. We applied the method of DEA-based composite indicator (BoD model) in order to aggregate three dimensions of CSR: environmental, social and governance. We further modified an original BoD model to account for non-convexities through FDH approach. The study found a considerable potential for improvement in CSR practices as revealed by the values of CSR composite indicators. The study also found differences in CSR efficiency between food and beverage firms in the regions considered.

Measuring CSR performance is necessary to guide sustainability improvements. Results of this study could be of interest to firms’ managers, CSR analysts and policy makers on how firms could improve CRS performance. Extensions of the results of this study could include the applications of ranking methods of firms’ CSR efficiency or extensions in the BoD model in order to account for the fact that CSR indicators are presented in the form of scores. Also, the development of composite indicators that would measure changes over time (productivity, efficiency and technology) within FDH approach is a potential line of future research. Alternative methods for the assessment of the differences in efficiencies between groups could be also considered in future studies.

\(^6\) In addition, we run Kneip et al. (2016) test of equivalence of mean efficiency across two groups. This test is not designed for indicators of technology gaps, so we run it only for region-specific and metafrontier indicators. The results of the test remained unchanged compared to Simar and Zelenyuk (2006) test, with an exception of region-specific indicator and differences between Europe and Asia. This difference in results between tests for this single case can be explained by the fact that Simar and Zelenyuk’s (2006) test compares the whole distributions of efficiencies between groups, while Kneip’s et al. (2016) test is designed to compare means of efficiencies, not the whole distributions.
Acknowledgements

Financial support for this article from the National Science Centre in Poland (decision number DEC-2016/23/B/H54/03398) is gratefully acknowledged. The calculations of adapted Li test were made at the Wroclaw Centre for Networking and Supercomputing (www.wcss.wroc.pl), grant no. 286.

References


