

CORPORATE SOCIAL RESPONSIBILITY OF DUTCH COMPANIES:  
BENCHMARKING, TRANSPARENCY AND ROBUSTNESS

BY

J.J. GRAAFLAND AND S.C.W. EIJJFINGER\*

*Summary*

This paper develops a method for benchmarking Corporate Social Responsibility (CSR) of Dutch companies. The benchmark includes economic, social and environmental aspects as well as national and international aspects of CSR. The overall benchmark is based on a weighted average of these aspects. The weights are based on the opinions of companies and NGO's. Using different methods of weighting, we find that the ranking for 58 large Dutch companies is very robust.

1 INTRODUCTION

Corporate Social Responsibility (CSR) has become an important theme. Due to an increasing interest of society in responsible behaviour of companies, many companies are nowadays concerned about values such as integrity and develop ethical codes to foster responsible behaviour of their employees. They feel that they must meet the triple P bottom line expressing the expectations of stakeholders with respect to the company's contribution to profit, planet and people in order to get a licence to operate. Firms who do not meet these expectations may see their reputation go down with a negative impact on market shares and profitability (McIntosh et al. (1998)). As people come to expect corporations to take a larger social role, companies will develop a social identity that is as important as brand identity. This changes the very nature of business and creates win-win situations.

This reputation mechanism will only be effective in stimulating companies to CSR if the social and environmental value creation of companies is transparent. If CSR is not transparent, companies may be able to build up a good CSR reputation by a successful communication strategy rather by really doing substantial effort to attain high ethical standards. On the other hand, companies that do pay high attention to CSR, may unjustly loose their good reputation as a result of one

\* Corresponding author: J.J. Graafland, Tilburg University, Room P2211, Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, the Netherlands; phone: +31 13 4662702, fax: +31 13 4662892, e-mail: J.J.Graafland@uvt.nl.

Respectively Department of Philosophy and Department of Economics, Tilburg University, The Netherlands. The authors thank N.C.G.M. Stoffele, A.M. Coldeweijer and H. Smid for their assistance in the research and two anonymous referees for their useful comments and remarks.

unlucky incident, if a more complete insight into its CSR performance is lacking. If companies are not transparent, customers, Non Governmental Organisations (NGO's) and the society at large cannot judge their contribution to CSR.

One way of improving the transparency of the CSR efforts of companies is benchmarking by independent institutes. Construction of an index that weights the contribution of companies into one number would clarify the position of individual companies and improve the comparability of their CSR efforts.<sup>1</sup> The publication of such an index can potentially enforce the reputation mechanism and provide a competitive advantage to companies that are indeed actively fostering social and ecological values. This would provide other companies a strong incentive to integrate CSR into the company's strategy (Graafland (2002a)).

This paper develops a method for benchmarking CSR. After a definition of corporate social responsibility, section 3 describes the methodology of our benchmark approach. In particular, we used a questionnaire that asks for about seventy concrete aspects of CSR. In addition, we asked for the values that companies and NGO's attach to these different aspects of CSR. The responses are used to construct weights for the calculation of the overall benchmark index. We also sent the questionnaire to a number of NGO's of which 6 responded. Section 4 describes the response to the questionnaire and the process of checking its reliability. Section 5 presents the outcomes for the overall benchmark index and investigates the robustness of this index using different weights. In particular, we compare four benchmarks based on respectively company based weights, sector based weights, total averaged based weights averaged over the four sectors and the weights reported by the NGO's and investigate whether these differ significantly. In this way, we also test the relevance of both the sector specific context of the company as well as the stakeholder specific context on the benchmark. Section 6 summarizes the main findings.

## 2 DEFINITION OF CSR

Any method of benchmarking corporate social responsibility should first consider the definition of CSR. There are several possibilities (Stoffele (2002)). We only mention two of them:

*Definition 1: A firm takes on a visible role in the society which goes beyond the core business and beyond what the law requires and which leads to added value for the company and the society. (source: Social Economic Council (2001))*

This is the definition that Mr. Ybema, Deputy-Minister for Economic Affairs, used when he asked the Social Economic Council (in Dutch 'Sociaal Economische

<sup>1</sup> An analogous approach in benchmarking is, for example, the development of an index for the transparency of central banks regarding monetary policymaking. See e.g. Eijffinger and Geraats (2002).

Raad' or briefly SER) in The Hague to advise about corporate social responsibility. Using this definition, CSR is restricted to a small set of activities, such as accountants cleaning up a children's farm or teaching at schools because of the shortage of teachers.

*Definition 2: Corporate social responsibility incorporates two elements.*

- 1 Sufficient focus by the enterprise on its contribution to public prosperity in the longer run*
- 2 the relationship with its stakeholders and society at large*

This is the definition used by the Social Economic Council (2001). The first element stresses that the enterprise can be viewed as a value creating entity. The long term value creation does not only relate to economic value, but concerns value creation in three dimensions which is called the Triple P bottom line:

- *Profit*: the economic dimension. This dimension refers to the creation of value through the production of goods and services and through the creation of employment and sources of income. The financial returns reflect the appreciation of consumers for the company's products and the efficiency with which factors of production are used. The long-term profit motive should provide the financial basis for the continuity of the company.
- *People*: the social dimension. This has a variety of aspects concerning the effects for human beings, inside and outside the organisation, such as good labour relations, safety etc.
- *Planet*: the ecological dimension. This dimension relates to the effects on the natural environment.

Another crucial aspect in the SER's definition is the relationship of the company with its stakeholders and the society at large. The enterprise is described as a form of cooperation of different stakeholders. The company should balance between partially conflicting interests. Good stakeholder relations also require that the firm is answering justified questions, is opening up its way of doing business and is willing to have a continuous dialogue with several interested parties. The SER distinguishes between primary and other stakeholders. Primary stakeholders are employees and shareholders, who have forms of structural consultation with the managers of the company. Other stakeholders include consumers, suppliers, competitors, the government and the society at large, all of whom have a certain interest in the companies' activities.

In our opinion, the definition of Deputy-Minister Ybema is too narrow, because CSR often also encompasses the core business of a company. If an oil company invests in alternative and less polluting petrol brand, then it contributes to the welfare of the society as a whole (Graafland (2001)). For this reason, we base our benchmark on the SER's definition of corporate social responsibility.

### 3 METHOD OF BENCHMARKING CSR

In this section, we describe several aspects of our method of benchmarking and compare them with other methods of benchmarking in the literature. First, we describe the selection of aspects included in the benchmark. Next, we discuss the method of data collection. Third, we consider the quantification and method of weighting in the construction of the overall index.

#### 3.1 *Selection of Aspects of CSR and Sectors*

Corporate social responsibility relates to a set of highly diverse aspects of the behaviour of companies. In their study of CSR in the Netherlands, Graafland et al. (2002) distinguish more than 60 concrete aspects. Each of these aspects can be further refined. For example, Kaptein (2001) gives 50 very concrete different examples of behaviour of employees that lack integrity.

As we strive to include both aspects of economic, social and ecological sustainability as well as national as international aspects of CSR, we had to make a selection in order to keep the benchmark relatively simple and transparent. Although there are many ways to increase the complexity of the benchmark, a highly comprehensive benchmark like proposed by Vlek et al. (2002) reduces the transparency and makes it more difficult to assign quantitative numbers to the scores of the companies. On the other hand, the benchmark should be systematic and complete. In order to keep a good balance between transparency and completeness, we distinguish between 68 concrete aspects of CSR, subdivided to 6 stakeholder groups (employees, suppliers, customers, the society at large, shareholders, competitors), and the organizational aspect of ethics.<sup>2</sup> Like Kleinfeld (2001), we included several instruments that facilitate responsible behavior of the company, such as the ISO standards, code of conduct, ethical committee and ethical training (Graafland et al. (2003b)).<sup>3</sup> Table 1 presents a summary overview of these aspects.

As can be seen from Table 1, the questionnaire comprises CSR aspects relating to economic values (in particular prevention of inside trading, profitability, measures to prevent collusion and bribery), ecological values (such as ecological effects of the production process and product of the company as well as of the supplier) and social values (such as the own labor conditions of the company, suppliers and the contribution to solving social problems). We are using both attention to performance indicators as well as to process indicators and instruments that measure procedures to foster ethical standards. This is in accordance with

2 This is more detailed than Graves et al. (2002), who base their benchmark on stockholder data (three-year average of total return to shareholders), pollution emissions, union relationships, employee benefits and philanthropic contributions. For a complete list of the 68 aspects, see Chapter 4 in Graafland et al. (2003a).

3 For more explanation of the instruments, see Graafland et al. (2003b).

TABLE 1 – ASPECTS OF CORPORATE SOCIAL RESPONSIBILITY PER STAKEHOLDER

<i>Employees</i>	Human rights / Dialogue with NGO's / Equal opportunity for women / Equal opportunity for minorities / training / safety and health / Participation / correct attitude / good fellowship among workers
<i>Suppliers</i>	Safety of product / environmental effect of product and production process / labour conditions of supplier / respect for supplier
<i>Customers</i>	Safety and quality of product / supply of sustainable alternative / respect for customer
<i>Society at large</i>	Environmental effects / active dialogue with environmental organisations / reintegration of disabled people / contribution to reduction of poverty in third world / contribution to local projects
<i>Shareholders</i>	Prevention of inside trade of stocks / profitability
<i>Competitors</i>	Respect for intellectual property of competitors / measures to prevent collusion / measures to prevent bribery
<i>Organization of ethics</i>	Various instruments, including code of conduct, ISO certification, external audits, social reports, social handbook, ethical committee and ethical training

principle P6.4 of the AA1000 standard (Jonker (2000)) that states that indicators should both reflect organisational processes as well as the results of these processes.

Of course, the number of concrete items could be increased further. For example, for future research it would be interesting to include the compliance to corporate governance measures proposed by the Code Tabaksblat. The concept of this code of the *Committee Corporate Governance* chaired by the former CEO of Unilever, Morris Tabaksblat, was presented in July 2003.<sup>4</sup> In February 2004 the Dutch government decided to adopt this new code for corporate governance for the Netherlands. The Code Tabaksblat consists of many proposals to improve the governance and management of Dutch firms, which are lagging behind Anglo-Saxon governance and management. The most important changes proposed by Tabaksblat et al. were regarding the control by the shareholders, the maximum compensation for discharging topmanagers to one-year annual salary (the so-called 'golden parachute'), and the maximum of five directorships per person (chairmanship counts double). Furthermore, Tabaksblat et al. proposed to appoint executive board members for a maximum term of four years with a possible re-appointment of another four years. The shareholders should approve also the remuneration in stocks and options *ex ante* and the reward by options should be

4 Since our research was carried out in the midst of 2002, well before the presentation of the concept Code Tabaksblat, we could not include these aspects of corporate governance in our research. We also lack information about the company's responses on the question how important these issues are. That means that we have no information about how strongly the issues covered by the Code Tabaksblat should be weighted in the overall benchmark.

depending on specific *ex ante* targets for the topmanagers. It is evident that the Committee Tabaksblat has clearly chosen for implementation of the Anglo-Saxon model of corporate governance for the Netherlands. In due course this model of corporate governance will affect many aspects of Corporate Social Responsibility (CSR) of firms in the Netherlands as well. However, it is too early to make an evaluation of the consequences of the Code Tabaksblat for the CSR of Dutch firms at this moment.

### 3.2 *Method of Data Sampling*

The CSR of companies can be investigated by different methods. Examples of methods designed to analyze the ethical standards of companies in great detail are case studies and participatory research like done by Graafland (2002b). The advantage of these methods is that much more insight is provided in the context, intention and problems of control of the specific company. On the other hand, close examination of a limited set of companies do not allow comparison between a larger number of companies nor general conclusions with respect to the quality of CSR.

Another way of sampling data is by examining what companies state about their policies in public documents. This method is applied by Krut and Munis (1998). They do not look at the activities the companies actually undertake. Implicitly they assume that transparency is sufficiently high to ensure that companies will not inflate the public material. At the end, when all the qualitative scores were assigned to the companies, each company was given an opportunity to review and comment on the qualitative scores. Although Krut and Munis do not explicitly mention it, this procedure demands conservative answering by the researchers. When a company remarks the score that it is given is too high, it will not have the incentive to downgrade this score. Reversed, when a score is too low, there will often be an incentive to correct it, because upgrading the score will be in the company's own interest.

Another disadvantage of this method is that public information on CSR efforts is relatively scarce. For that reason we use the questionnaire as an instrument to obtain information about the CSR efforts of companies. Of course, as the companies fill in questionnaires, there is a risk that the answers do not reflect the actual situation. This is even more likely in our research, because we communicated to the companies that the benchmark would be published. In order to prevent the bias in results, we have spent much energy in checking the results of the questionnaire using public information from newspapers, annual reports and other public sources (see below).

### 3.3 *Method of Weighting*

One of the methodological problems of benchmarking is the weighting of all the aspects into one overall index. As CSR is a highly complex phenomenon, a quantitative added value approach is not possible. The information required for performing cost benefit analysis on all the effects of the company is not available. Another fundamental problem is that it is very difficult to value ecological and social effects in monetary terms. For this reason, we use a pragmatic method by defining concrete CSR aspects that entrepreneurs find reasonable and relevant to what they care about. As Anderson (1993) argues, such a component-value strategy, in which the overall value is a weighted sum of its component values, can offer a successful solution to the multi-criterion evaluation problem if criteria are relevant for the practice to be evaluated. It represents evaluation as essentially a matter of calculation, with the aim of making the process precise and decisive. Similar component-value strategies are applied, for example, in decathlon scoring, in which times and distances in different events, such as the hundred-metre dash, the shot put and the long jump, are converted to a common point scale (Anderson (1993)).

In order to construct the relevant weights used for weighting the 68 aspects of CSR into one benchmark index, we investigate how much the respondents value these different aspects and based our weights in the total benchmark on their opinions. Basing the weights on the opinions of the companies is also consistent with principle P6.3 of the AA1000 standard, which states that the identification of indicators reflecting the company's performance must be based on the values of the company itself, the opinions of stakeholders and the society at large (Jonker (2000)). For this reason, we added a second question list to the questionnaire. Whereas in the first part the company has been asked questions both about the effort and procedures of the company to foster ethical standards as well as to the results of these efforts, the second part of the questionnaire is asking the company to give its opinion about the relevancy of the different aspects of CSR (they can chose a value of 0, ½ or 1.0 for each aspect). Also some Non Governmental Organisations (NGO's) have been sent this part, which enables us to know how these organisations value the various activities.<sup>5</sup> Although NGOs are only one type of stakeholder out of many others, using NGOs' responses provides a good test for the robustness of the benchmark, because NGOs have, on average, a more critical view on CSR than other stakeholders such as suppliers, customers, competitors, employees and the government. Hence, if we find that the benchmark is robust when the companies' weights are replaced by NGOs' weights, it is very likely that the benchmark will also be robust if weights preferred by other stake-

5 The following NGO's have sent in a completed questionnaire: Novib, Wereld Natuur Fonds, Milieudefensie, Hivos and one anonymous NGO.

holders would be used. The respondents are also asked to attach weights to the relative importance of different stakeholder groups.

Compared to other benchmark methods, this aspect of our methodology is very innovative. It enables us to take account of the context of the company. For example, whereas the construction and chemical companies attach a high priority to environmental aspects, these aspects are relatively unimportant for the retail sector and the financial and banking sector. Moreover, as we explicitly distinguish between six different stakeholder groups and ask the respondents about the relative importance of these stakeholder groups, we are also able to take account of different stakeholder perspectives of companies. In contrast, other researchers like Waddock and Graves (1997) and Graves et al. (2002) take an unweighted average of seven detailed parameters to yield a single score for each firm. The fact that the scale is unweighted means all stakeholders have equal status.

Another way of investigating the impact of the context on CSR is by focusing our study on four sectors: construction, retail, chemical companies and financial and banking services. The focus on four sectors allows comparing the results for different companies within one sector. Moreover, the selection facilitates detecting sector-specific characteristics.

### 3.4 *Quantification*

In order to quantify the scores, we distinguish three scores per aspect of CSR (valued respectively by 0,  $\frac{1}{2}$  and 1.0). This is similar to the method applied by Waddock and Graves (1997) and Graves et al. (2002). In their method, each company can have 'strengths' and 'concerns' for each category. These are rated on scales ranging from -2 (major concern) to 0 (neutral) to +2 (major strength). To arrive at a net score in a category, the concerns are subtracted from strengths. Thus a firm with a score of 'two' in employee strengths and 'one' in employee concerns would have a net score of 'one' in the employee category.

Other benchmark researchers such as Vlek et al. (2002) and Krut and Munis (1998) refrain from quantification of all CSR aspects. By using a qualitative measure, they avoid the problems involved with the quantitative measurement of CSR. However, this comes at the expense of deriving clear conclusions from their standard. Indeed, there is always a trade-off between providing sufficient depth and context for users on the one hand and providing easy comparable outcomes.

A very important aspect of the quantification is the choice of the cut-off values in the scores used for determining the mark per aspect of CSR. These cut-off values are based on the outcomes of Graafland et al. (2002). They investigate similar features of corporate social responsibility for 110 companies in Brabant and Zeeland. This study has focused on the construction sector, retail sector and financial and banking sector. Therefore, the outcomes of this research yield good indications about relevant cut-off values determining the valuations per concrete aspect of CSR. It should be noted, however, that for some questions the cut-off



values are a bit vague. For most (48 out of 68) questions the differences between the options is unambiguous causing no problems of interpretation by the respondents. In many questions, the options refer to two or three excluding answers (like: 'Does your company offer paid child care to employees?: yes / no'). In other cases, the options are stated in quantitative terms (such as: 'The budget for training of employees is: less than 2% of total wage costs; between 2% and 5% of total wage costs; more than 10% of total wage costs'). For 20 out of 68 questions, the difference between two (out of three) options is more open to alternative interpretation. This leaves room for manipulation. An example is the question: 'How much effort does your company make to prevent abuses on the working floor (think of teasing or sexual intimidation etc.): no; in some degree; much'. The description of the second and third option in this question leaves some room for interpretation. Overall, more than 80%  $((48+20/3)/68)$  of the questions is unambiguous. We find that acceptable. Since CSR comprises many aspects which are relatively difficult to quantify, it is hard to exclude all ambiguity in the options.

#### 4 RESPONSE AND TRANSPARENCY

##### 4.1 *Response to the Questionnaire*

The addresses of the companies were taken from publications of Price Waterhouse Coopers (2001, 2002a, 2002b and 2002c). In total we sent 378 questionnaires. The response rate varied from 20% in the construction and chemical sector to 14 % for the financial and banking sector and 9 % for the retail sector. On average, the response rate was 15.3 %. The sample of companies that sent in a complete questionnaire consists of 58 companies.

Very interestingly, the response rate was relatively high for the construction sector. This sector is currently plagued by a low ethical reputation because of collusion. The high response rate might indicate that the companies are aware that, in order to improve their reputation, attention to corporate social responsibility is necessary. Another surprise was that companies that are well known for their relatively pro-active attitude on CSR (and therefore would probably get a high score) did not respond to the questionnaire. In telephone calls one manager of the Koninklijke Shell Groep told us that his company does not cooperate with scientific research. This seems to contradict the open attitude that this company communicates in its publications on CSR. Another manager of Unilever N.V. declared that he did not have much confidence in benchmarking.

#### 4.2 Transparency

As noticed above, it is difficult to interpret the relatively low response to our questionnaire. Although one is inclined to believe that this is a sign of lack of openness of companies, there might also be other and more valid reasons why companies did not respond, such as a high work pressure and the large number of questionnaires that companies are asked to fill in.

The low response induced us to try to get more information by investigating public sources. For this purpose, we did a major effort. In particular, whereas one member of our staff was full time researching annual reports, newspaper articles and Internet sites to detect relevant information, 180 third years students of the International Business course also each checked one particular company from our sample. Notwithstanding this large effort, it appeared to be very hard to obtain information from public sources for all the CSR aspects distinguished in our questionnaire. Although some large companies are pro-actively fostering their transparency by social reporting, an integrated and uniform framework is, however, lacking. Large companies often acknowledge the importance of standardisation (which is, for example, supported by the the guidelines of Global Reporting Initiative (2002)), but in practice there are substantial differences in how companies apply these standards (Lamoen and Tulder (2001)). Table 2 presents an overview of the results of the average number of questions that could be checked by public documents for companies that send in the questionnaire:

TABLE 2 – AVERAGE NUMBER OF QUESTIONS FOR WHICH WE FOUND PUBLIC DOCUMENTS<sup>a</sup>

Construction	Retail	Chemical sector	Financial and banking	Total
3	6	5	8	5

<sup>a</sup> From a subset of 30 questions

As can be seen, the transparency of the companies is relatively low. This especially holds for companies in the construction sector. This implies that it is impossible to benchmark CSR efforts of companies on available public documents.

The information from public resources also allowed us to investigate the quality of the responses of the companies that have sent in a questionnaire. For this purpose, we have compared the public information with the responses from the questionnaire. Table 3 gives an overview of the number of answers diverging from the answers given by the companies as a share of the total number of answers that were found in the public sources.

TABLE 3 – SHARE OF ANSWERS FROM PUBLIC SOURCES THAT DIFFER FROM ANSWERS OF COMPANIES

Construction	Retail	Chemical	Financial and banking	total
22 %	29%	26%	22%	25%

The table shows a remarkable similarity between the different sectors. In all cases, the share of answers that differs from the answers given by the companies is within a range of 20% to 30 %.

Table 4 gives an impression of the type of differences that we found, each cell representing one example.

We found that in several cases the public documents give a more favourable picture than the information provided by the company. Since we can only compare the firms' responses with public information for 5-12% of the questions, it is difficult to draw a clear conclusion about the reliability of the information filled in by the companies. In order to retain the consistency in our data, we did not adjust the companies' responses using the information from public sources. Because of the low transparency, the average number of deviating answers per company is very small (namely 1.25 out of 68 answers). Correction on the basis of public information would therefore hardly affect the benchmark. Moreover, as the information that we sampled using public sources may also be subject to errors, we preferred to use the answers filled in by the companies themselves.

## 5 OUTCOMES OF THE BENCHMARK

In this section the outcomes of the benchmark procedure are described. There are four different benchmarks for all companies, which are based on different weights. First, we present the formulas used for calculating the benchmarks. Next, the outcomes of the benchmarks per sector are described. Third, we give an intersectoral view. Finally, we investigate the robustness of the different benchmarks graphically and by using statistical tests.

### 5.1 *The Benchmark Method: Four Methods of Weighting*

The overall benchmark (B) was constructed by using the following formula:

$$B = (\text{Sum}(w_i * b_i) + w_o b_o) / (\text{Sum}(w_i) + w_o)$$

$w_i$  denotes the weight per stakeholder and  $w_o$  is the weight for the use of instruments to organize ethics as filled in by the respondents.  $b_i$  is the benchmark per stakeholder and  $b_o$  the benchmark for the use of instruments.  $b_i$  is constructed by the following formula:

$$b_i = \text{Sum}(w_j * b_j) / \text{Sum}(w_j)$$

TABLE 4 – EXAMPLES OF DIFFERENCE BETWEEN ANSWERS OF COMPANIES AND PUBLIC DOCUMENTS

Construction	Retail	Chemical	Financial and banking
compliance ILO standards (freedom of union) – company: fully – public info: In most cases	deceptive promotion activities – company: some-times – public info: no	compliance ILO standards (working times) – company: in most cases – public info: fully	active dialogue with NGO's – company: incidentally – public info: on regular basis
environmental safety checked – company: Never – public info: On regular basis	environmental reporting of company? – company: Yes – public info: No	compliance to ILO standards (freedom of unions) – company: in most cases – public info: fully	percentage of women in higher management – company: No answer – public info: 0-10%
product-line as durable alternative for core products – company: No answer – public info: No	transportation plan – company: Slightly – public info: Yes	budget for courses for employees as % of total wage costs – company: more than 5% – public info: 0-2%	working conditions supplier checked – company: Sometimes – public info: On regular basis
handbook for employees? – company: Planned – public info: No	internal and written ethical code? – company: No – public info: Yes	environmental safety of supplier checked – company: Sometimes – public info: Never	deceptive promotion activities – company: No answer – public info: Sometimes
	public company code? – company: No – public info: Yes	working conditions supplier checked – company: Sometimes – public info: Never	finances for non-compliance with environmental norms – company: Never – public info: On regular basis
		social annual report – company: No – public info: Yes	independent audit of the company code? – company: No – public info: Yes

where  $w_j$  denotes the weight per aspect of CSR (ranging from 0 to 1) and  $b_j$  the value of the option per aspect filled in by the respondents. The benchmark for the use of instruments is calculated in a similar way.

As we asked the opinion of the respondents and NGO's about the weights to be used for the benchmark, we have four alternatives of weighting. First, we can

construct an overall benchmark index using for each company the weights reported by the *companies* themselves. The company then defines by itself the criteria that are relevant to judge its CSR performance. We will label the benchmarks with this weighting scheme as the *individual* benchmarks. The advantage of this method of weighting is that the benchmark will maximally reflect the context that the individual company perceives as most relevant. The disadvantage of this method is that the weight might be biased. In particular, companies might attach high values to CSR aspects for which they rank high.

A more objective method is the second approach in which we used averaged weights per *sector*. As companies operating in one particular sector face similar conditions, using sectoral weights will still take account of the context of the company, although not as specific as in the case of individual weights. We will name these benchmarks the *sectoral (based)* benchmarks. An advantage of this approach is that subjective judgments and the upward bias created by using individual weights will be filtered out.

A third method is to use average weights based on the *total* sample of companies. We will call these benchmarks the *total average* or *total based* benchmarks. Compared to the second approach, this method of weighting does not take account of context aspects related to typical sectoral characteristics. Still, it is interesting to test whether this general method of weighting produces completely different benchmark results.

In the last approach we will use the average weights of the *NGO's*. These benchmarks will be referred to as the *NGO (based)* benchmarks. This also provides insight into the robustness of the benchmark approach. In particular, it allows us to test how stakeholder related the benchmark is. Of course, we are very much interested to see whether and to which degree the individual, sectoral, average and NGO benchmarks will diverge from each other.

## 5.2 Sectoral Outcomes of the Benchmarks

Table 5 presents the four different benchmarks for the 10 best performing companies in the construction sector respectively of all companies (for the other sectors, see Graafland et al. (2003a)). For every different benchmark, the companies are presented in a descending order. Theoretically, the highest score of a company would be 1 and the lowest score 0. In practice, the range runs from 0.95 to 0.08 for all benchmarks and all sectors.<sup>6</sup>

Table 5 shows that the benchmarks for the construction sector are rather robust. If sectoral, total and NGO's based weights are used, Beheer- en Beleggingsmaatschappij Teerenstra B.V. achieves the highest mark. In case of individual

<sup>6</sup> It should be mentioned that the company with the lowest benchmark (Audax B.V.) only partly filled out the questionnaire. As blanc answers received a zero mark, this may explain the extremely low benchmark. The actual situation is probably better for this company.

TABLE 5 – BENCHMARK ORDERING FOR CONSTRUCTION SECTOR RESPECTIVELY ALL SECTORS<sup>a</sup>

	<i>Individual</i>	<i>Total</i>	<i>NGO</i>	<i>Sectoral</i>
	<i>Construction sector</i>			
1	Dura Vermeer Groep NV	0,85	Beheer- en Beleg- gingsmaatschappij Teerenstra B.V.	Beheer- en Beleg- gingsmaatschappij Teerenstra B.V.
2	Van Straten groep B.V.	0,80	Dura Vermeer Groep NV	Dura Vermeer Groep NV
3	Koninklijke Volker Wessels Stevin N.V.	0,78	Grontmij NV	Grontmij NV
4	Grontmij NV	0,78	Van Straten groep B.V.	Van Straten groep B.V.
5	Unidek Beheer B.V.	0,76	Koninklijke Volker Wessels Stevin N.V.	Koninklijke Volker Wessels Stevin N.V.
6	Simon Benus Beheer B.V.	0,75	Unidek Beheer B.V.	Unidek Beheer B.V.
7	Beheer- en Beleg- gingsmaatschappij Teerenstra B.V.	0,75	Verhoeven Bouw- groep B.V.	Verhoeven Bouw- groep B.V.
8	Verhoeven Bouw- groep B.V.	0,72	Hollandsche Beton Groep nv	Hollandsche Beton Groep nv
9	Hegeman Holding B.V.	0,69	Klaassen Beheer Maatschappij B.V.	Klaassen Beheer Maatschappij B.V.
10	Klaassen Beheer Maatschappij B.V.	0,67	Hegeman Holding B.V.	Van den Herik Beheer B.V.

TABLE 5 – (CONTINUED)

<i>All companies</i>		<i>Total</i>			<i>NGO</i>			<i>Sectoral</i>		
	<i>Individual</i>									
1	Cytec Industries B.V.	0,94	DSM N.V.	0,89	DSM N.V.	0,88	DSM N.V.	0,90	DSM N.V.	0,90
2	DSM N.V.	0,94	Cytec Industries B.V.	0,84	Cytec Industries B.V.	0,83	Cytec Industries B.V.	0,85	Cytec Industries B.V.	0,85
3	ING Bank	0,90	Beheer- en Belegingsmaatschappij Teerenstra B.V.	0,78	Beheer- en Belegingsmaatschappij Teerenstra B.V.	0,78	Beheer- en Belegingsmaatschappij Teerenstra B.V.	0,79	Beheer- en Belegingsmaatschappij Teerenstra B.V.	0,79
4	Dura Vermeer Groep NV	0,85	Fortis Bank	0,76	Fortis Bank	0,75	Fortis Bank Nederland (Holding) N.V.	0,76	Fortis Bank Nederland (Holding) N.V.	0,76
5	Fortis Bank	0,84	Dura Vermeer Groep NV	0,75	Dura Vermeer Groep NV	0,74	Dura Vermeer Groep NV	0,75	Dura Vermeer Groep NV	0,75
6	Beheer- en Belegingsmaatschappij Teerenstra B.V.	0,82	Stauffer	0,74	Stauffer	0,72	Stauffer Chemical B.V.	0,75	Stauffer Chemical B.V.	0,75
7	nv Bank Nederlandse Gemeenten	0,80	Grontmij	0,70	ING Bank	0,68	nv Bank Nederlandse Gemeenten	0,70	nv Bank Nederlandse Gemeenten	0,70
8	Van Straten Groep B.V.	0,80	ING Bank	0,69	Grontmij	0,68	ING Bank N.V.	0,69	ING Bank N.V.	0,69
9	Koninklijke Volker Wessels Stevin N.V.	0,78	Van Straten Groep B.V.	0,69	Van Straten Groep B.V.	0,67	Grontmij NV	0,69	Grontmij NV	0,69
10	Grontmij	0,78	nv Bank	0,69	Koninklijke Volker Wessels Stevin N.V.	0,67	Van Straten groep B.V.	0,68	Van Straten groep B.V.	0,68

<sup>a</sup> Ten best performing companies

weights Dura Vermeer has the highest benchmark, but this company takes the second place in the other benchmarks. Other companies that appear to belong to the top five are Grontmij NV, Van Straten groep B.V. and Koninklijke Volker Wessels Stevin N.V. For the other sectors, we found similar results.

On basis of these results, we conclude that the weights do not generally have a strong impact on the benchmarks. The ascending order is rather robust. In all sectors, the company receiving the highest mark is the same for weights based on sectoral averages, total averages or NGO's averages. Also the top five is very robust in each sector when different weights are used. This has very important implications:

- the sectoral specific context is not of decisive importance, since *sectoral* based weights hardly change the benchmark compared to total based weights
- also stakeholder specific opinions hardly matter, as the *NGO* based weights do not impact the benchmark outcome compared to total based weights

Only in the case of weights based on the *individual* answers of company, the company receiving the highest mark differs from the other methods. Since individual weights are relatively subjective and might be biased to the actual performance of companies, the weights based on sectoral averages may provide a more reliable guide since it still takes into account the sectoral context whilst being more inter-subjective in nature.

In order to test whether the differences in the overall benchmark are also robust when considering subcategories, Table 6 reports the outcomes for the 7 subcategories for the top-five and bottom-five companies. From this table we can see that in most cases the exact ranking per subcategory differs from the total ranking reported in the last column. For example, whereas DSM has the highest overall benchmark (as well as for the subcategories employees, customers, competitors and instruments), it is only ranked at the third place for supplier related items. Also for the lowest-five companies the ranking per subcategory differs from the overall ranking. However, when comparing the top-five with the bottom-five companies, we find that in almost all cases (except the three bold figures) each of the bottom-five companies receives (in most cases significant) lower benchmarks for each subcategory than each of the top-five companies.

### 5.3 An Intersectoral Comparison

The small difference between benchmarks based on sector weights and on total average weights (including all sectors) suggests that it is reasonable to compare the CSR performance of companies of different sectors. In order to compare the CSR performance of different sectors, Table 7 presents an overview of the average benchmark values per sector and per weight.

Table 7 yields some interesting conclusions. First, as one could expect, we find for all sectors that the average benchmark is lowest in the case of weights based on the responses of NGO's and highest in the case of weights based on the



TABLE 6 – SUBCATEGORIES FOR COMPANIES WITH HIGHEST AND LOWEST TOTAL BENCHMARK<sup>a</sup>

Top 5	Employees	Suppliers	Customers	Society	Shareholders	Competitors	Instruments	Total
DSM	0.81	0.94	1.00	0.85	0.75	1.00	0.94	0.90
Cytec	0.77	1.00	1.00	0.62	0.75	1.00	0.81	0.85
Beheer	0.69	0.81	1.00	0.95	1.00	1.00	0.30	0.79
Fortis	0.81	0.87	0.95	0.55	0.75	1.00	0.52	0.76
Dura	0.68	1.00	1.00	0.68	0.50	0.76	0.68	0.75
Bottom 5	Employees	Suppliers	Customers	Society	Shareholders	Competitors	Instruments	Total
Maison de Bonneterie	0.57	0.43	0.66	0.29	0.00	0.57	<b>0.41</b>	0.42
Carpet	0.36	0.30	0.68	0.31	0.50	0.48	0.23	0.41
Tokai	0.39	0.34	0.58	<b>0.59</b>	0.00	0.66	0.08	0.36
J&W Vastgoed	0.55	0.34	0.40	0.43	0.00	<b>0.82</b>	0.29	0.40
P&C	0.38	0.40	0.62	0.30	0.00	0.32	0.18	0.31

<sup>a</sup>Based on sectoral weights

TABLE 7 – AVERAGE BENCHMARKS PER SECTOR

Weights	Construction	Retail	Chemical	Financial and banking
Individual	0.651	0.464	0.703	0.696
Total average	0.596	0.440	0.675	0.595
NGO	0.582	0.430	0.662	0.579
Sectoral	0.595	0.445	0.685	0.603
Average number of employees of companies that responded	2273	555	11573	11817

individual responses of companies. This suggests that the discrepancy between what companies actually do and what they should do is larger according to the NGO's than to the companies. For example, whereas NGO's attach a high value to an active dialogue between companies and NGO's (with a weight of 0.9), companies value this aspect of CSR relatively low (with a weight of 0.3). As the actual performance of companies in this respect is relatively low (see Table 4.1 in Graafland et al. (2003a)), this lowers the average benchmark when using NGO's weights. On the other hand, individual companies tend to value different aspects of CSR more in line with what they actually do. This results in a relative high benchmark.

A second interesting finding is that in all cases the difference between the average benchmark based on total average weights, NGO average weights and sectoral average weights is rather small. This confirms the finding that, although NGO's are most critical, they nevertheless have quite similar opinions about what matters with respect to CSR.

Third, we find that, on average, the chemical sector appears to be most professional in CSR. This may partly be explained by the large size of chemical companies. As shown by Graafland et al. (2003b), large companies normally pay more attention to CSR, because they are more visible to the public and the media. This makes investments in responsible production and selling patterns relatively more important for large companies. This positive relationship between company size and CSR efforts may partly explain why the retail sector is paying the least attention to CSR. As shown in Table 7 the retail companies in our sample have by far the smallest size.

Fourth, an unexpected result is that the construction sector ranks for three out of four types of weighting as high as the financial and banking sector, which is normally known for its high attention to CSR. This is even more amazing if we take into account that the average size of the financial companies is much larger than the average size of the construction companies in our sample. This invokes the question whether the good reputation of the financial sector (compared to the bad reputation of the construction sector) is really based on good CSR perform-

ance or more due to a more effective communication strategy by financial and banking institutions. Likewise, this result suggests that the current bad reputation of the construction sector may be an overreaction to the recent news about the fraud in the construction sector and that people tend to forget that construction companies do pay much attention to other social and ecological issues. Another explanation might be that the upward bias might be somewhat higher for construction companies, because companies find it hard to admit major infringements or illegal actions in questionnaires.

It should be noted that the comparison of the exact outcomes of the benchmark may be distorted by biases resulting from the relative high non-response. In particular, we expect a certain upward bias, because companies with a low grade have probably been less prepared to cooperate with our research, as we had informed them that the results would be published. This bias does, however, not change our finding that the chemical and construction sector do on average rather well. For these sectors, the response rate is relatively high. If the response rate of the financial and retail companies had been as high as for the chemical and construction sector, the average benchmarks for the financial and retail sectors would probably have been even lower.

Table 8 presents an overview of the variation in benchmark results for the different methods of weighting. Again, using weights based on total averages, NGO averages or sectoral averages hardly affect the standard deviation. Only for individual weights the standard deviation is substantially higher.

TABLE 8 – STANDARD DEVIATION OF THE BENCHMARK

Individual	Total	NGO	Sectoral
0.1698	0.1348	0.1343	0.1346

#### 5.4 *Formal Tests of the Robustness of the Benchmark*

As already indicated, using different weights for the various CSR aspects does not have a major impact of the ordering of the CSR performance of different companies. To test this finding more systematically, this section provides both graphical and econometric evidence of the correlation between the benchmark results.

In Figures 1-3 the various benchmarks are plotted together. These figures show again that the differences between the sectoral benchmark, the NGO benchmark and the average benchmark are not large. In contrast, if the total average benchmark and individual benchmark are plotted together the points are less close to the reference line (45° line), which means that the differences between the two benchmarks are larger. The picture also makes clear that the difference between benchmark based on individual weights and benchmark based on total average

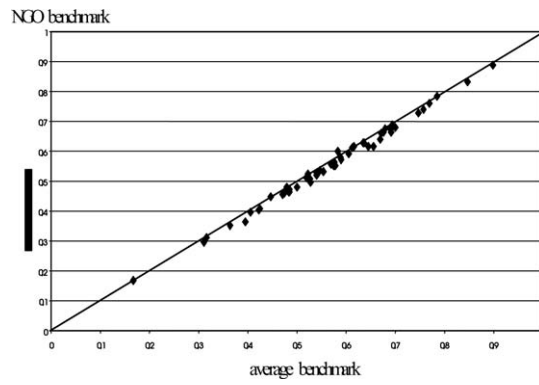


Figure 1 – NGO benchmark plotted with the total average benchmark

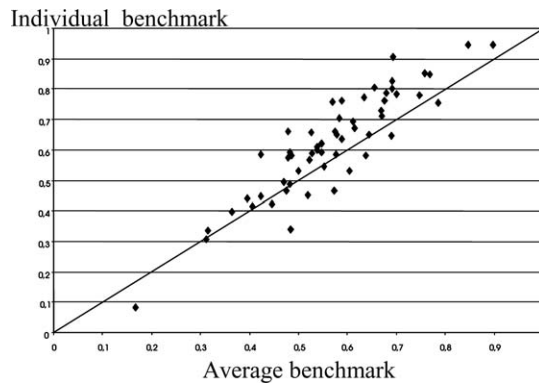


Figure 2 – Individual benchmark plotted with the total average benchmark

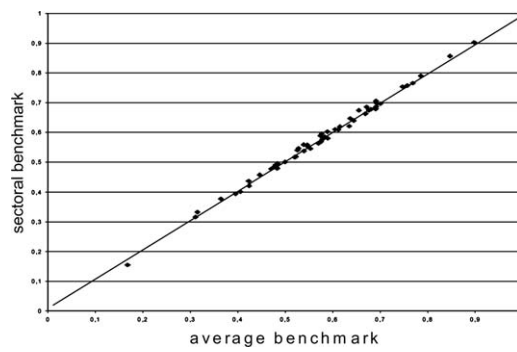


Figure 3 – Sectoral benchmark plotted with the total average benchmark

weight is more likely to be positive for companies with a high CSR performance than for companies with a low CSR performance. This might indicate that companies with a high CSR performance are more aware which aspects of CSR receive a high valuation in their company. As a result, they are more able to provide a valuation that is consistent with their performance.

Finally, the correspondence between the different benchmarks can also be tested by statistical correlations between the benchmarks. For this purpose, Table 9 presents Kendall's  $\tau$  and Spearman's  $\rho$  values. In both cases, the outcomes confirm that the total average benchmarks, the NGO benchmarks and the sectoral benchmarks are very strongly correlated and that the correlation between the individual benchmarks and the other benchmarks is smaller.

TABLE 9 – CORRELATION COEFFICIENTS<sup>a</sup>

<i>Kendall's</i> $\tau$	Individual	Total average	NGO	Sectoral
Individual	1.000	0.746*	0.746*	0.741*
Total average	0.746*	1.000	0.950*	0.951*
NGO	0.746*	0.950*	1.000	0.936*
Sectoral	0.741*	0.951*	0.936*	1.000
<i>Spearman's</i> $\rho$	Individual	Total average	NGO	Sectoral
Individual	1.000	0.897*	0.899*	0.896*
Total average	0.897*	1.000	0.995*	0.995*
NGO	0.899*	0.995*	1.000	0.991*
Sectoral	0.896*	0.995*	0.991*	1.000

<sup>a</sup> The asterisks (\*) in the tables mark the numbers that are significant at a 99% confidence interval.

## 6 SUMMARY OF THE MAIN FINDINGS

Through benchmarking CSR, companies are given a mark for their actions and achievements, which enables stakeholders to judge how responsible a specific company is. When a score is constructed, it is much easier for stakeholders to confront the company with its actions. This increases the transparency and accountability of companies.

Notwithstanding these and other advantages, benchmarking of CSR is also criticized because it tends to disregard the context of the company. In order to counter this problem as good as possible, we have developed a benchmark method with a focus on four sectors (construction sector, retail sector, chemical sector and financial and banking sector) based on 68 concrete aspects which are both related to the CSR efforts as well as to the results of these efforts. We have also asked for the company's opinion about the relevancy of these different aspects of CSR. Also some NGO's have been sent this part of the questionnaire in order to

know how these organisations value the various activities. Using this information, we have developed four alternative benchmarks using weights based respectively on opinions of individual companies, sectoral averages, total averaged weights and NGO opinions. Such a component-value strategy, in which the overall value is a weighted sum of its component values, can offer a successful solution to the multi-criterion evaluation problem if criteria are relevant for the actual situation to be evaluated. Basing the weights on the opinions of the companies and NGO's is also consistent with principle P6.3 of the AA1000 standard which states that the identification of indicators that reflect the company's performance must be based on the values of the company itself, the opinions of stakeholders and the society at large (Jonker (2000)).

In total we have sent 378 questionnaires to the largest Dutch companies in the construction, retail, chemical and financial and banking sector. The average response rate was 15 %. This rather low response induced us to try to get more information by investigating public sources. For this purpose, we did a major effort. We found out that it is very hard to obtain information from public sources for all the CSR aspects distinguished in our questionnaire. On average, we only found public information for 5 questions per company. This implies that it is impossible to benchmark CSR efforts of companies on available public information.

Comparison of the public information and the information from the questionnaires showed in 20-30 % different answers. Analysis of the type of differences showed that in some cases the exact meaning of the answer options is too crude. Another explanation is that persons who filled in were not fully informed about the actual situation of their company or that the public source does not provide reliable information about the actual situation of the company. As our investigation covers a limited period, it might well be that we missed some information related to issues in the past. A final explanation is that the company did not want to reveal sensitive information in the questionnaire, for example about fines for environmental infringements or deceptive promotion activities. Another interpretation is that the person who filled in the questionnaire interprets the cases different than public sources (like newspapers) do.

Analysis of the responses shows in some cases some clear sectoral patterns of CSR. For example, environmental aspects of suppliers are important for the construction and chemical sector, but do not receive a high priority in the retail and banking sector. This difference in effort is also reflected in the weights. Still, for many aspects we find a high similarity in weights between different sectors. For example, all aspects of employee relations are valued with an average weight of at least 0.5, except an active dialogue with NGO's. Of course, NGO's rate this aspect much higher, but for most other aspects the NGO's view does not differ very much from that of the companies.

Comparison of the four benchmarks shows that the various weights do not have a strong impact on the benchmarks. The ascending order is rather robust. In all sectors, the ordering of companies is very robust for weights based on the

sectoral averages, total averages or NGO's averages. This indicates that neither the sectoral context nor specific stakeholder perceptions have a major impact on the benchmark results. Only in the case of weights based on the individual answers of companies, the company receiving the highest mark changes. Since individual weights are relatively subjective and might be biased to the actual performance of companies, the weights based on sectoral averages may provide a more reliable guide because it still takes into account the sectoral context, whilst being more inter-subjective in nature.

Furthermore, we find that the average benchmark is lowest in the case of weights based on the responses of NGO's and highest in the case of weights based on the individual responses of companies. However, the difference between the average benchmark based on total average weights, NGO average weights and sectoral average weights is rather small. This is explained by the fact that, although NGO's are most critical, they nevertheless have quite similar opinions about what matters with respect to CSR.

Moreover, we find that, on average, the chemical sector appears to be most professional in CSR and the retail companies least professional, whereas the construction sector and financial and banking sector have an intermediate position. This suggests a positive relationship between company size and CSR efforts. It also indicates that reputation and actual CSR performance do not necessarily coincide.

Finally, it should be acknowledged that benchmarking will never provide a complete picture of the behaviour of the company. Benchmarking may even have a negative impact on corporate social responsibility if it leads companies to engage in more 'image management' or sets standards that narrow the scope of responsibility (i.e., what gets measured is what gets done). Therefore, it cannot be a substitute for other institutions that enforce companies to consider society's interests, and in particular the provisions laid down in legal requirements. Other sources of information, such as financial and social reporting, certifications, and in case of legal misdoings in-depth investigations remain necessary to make companies effectively accountable for their contribution to the society's interest. On the other hand, focusing only on specific infringements of the law may also produce an unbalanced view on the overall CSR performance of a company.

## REFERENCES

- Anderson, E. (1993), *Value in Ethics and Economics*, Cambridge/London, Harvard University Press
- Eijffinger, S.C.W. and P.M. Geraats (2002), How Transparent Are Central Banks?, CEPR Discussion Paper, No. 3188, Centre for Economic Policy Research, London.
- Global Reporting Initiative (2002), *Final Report of the Measurement Working Group*.
- Graafland, J.J. (2001), *Maatschappelijk ondernemen: Analyse, verantwoording en fundering*, Inaugural Lecture, Tilburg University.

- Graafland, J.J. (2002a), Modelling the Trade Off between Profits and Principles, *De Economist*, 150, pp. 129-154.
- Graafland, J.J. (2002b), Sourcing Ethics in the Textile Sector: The Case of C&A, *Business Ethics: A European Review*, 11(3), pp. 282-294.
- Graafland, J.J., B.W. van de Ven and N. Stoffele (2002), Wat betekent maatschappelijk ondernemen concreet?, *CMO Report*, Tilburg University, [www.uvt.nl/wijsbegeerte/cmo](http://www.uvt.nl/wijsbegeerte/cmo)
- Graafland, J.J., S.C.W. Eijffinger, N.C.G.M. Stoffele, H. Smid and A.M. Coldeweijer (2003a), Corporate Social Responsibility of Dutch Companies: Benchmarking and Transparency, CMO Report, Tilburg University, [www.uvt.nl/wijsbegeerte/cmo](http://www.uvt.nl/wijsbegeerte/cmo)
- Graafland, J.J., B.W. van de Ven and N. Stoffele (2003b), Strategies and Instruments for Organising CSR by Small and Large Business in the Netherlands, *Journal of Business Ethics*, 47, pp. 45-60.
- Graves, S.B., S.A. Waddock and J. Kelly (2002), 100 Best Corporate Citizens, *Business Ethics*, 11(2), pp. 8-13.
- Jonker, J. (ed.) (2000), *AA 1000 standaard: richtlijn voor maatschappelijk verantwoord ondernemen*, Van Gorcum, Assen.
- Kaptein, M. (2001), De integriteitsbarometer voor organisaties, *Bedrijfskunde*, 73(3), pp. 12-18.
- Kleinfeld, A. (2001), Benchmarking the Moral Decision-Making Strength of European Biotech Companies: A European Research Project, *Business Ethics: A European Review*, 10, pp. 122-139.
- Krut, R. and K. Munis (1998), Sustainable Industrial Development: Benchmarking Environmental Policies and Reports, *Greener Management International*, 21, pp. 88-98.
- Lamoen, C. van, and R. van Tulders (2001), Accepteren, aanmodderen of afwijzen, *Economisch Statistische Berichten*, pp. 932-934.
- McIntosh, M., D. Leipziger, K. Jones and G. Coleman (1998), Corporate Citizenship: Successful Strategies for Responsible Companies, *Financial Times Management*, London, pp. 61-68.
- PriceWaterhouseCoopers, Ondernemingsanalyses (2001), *Trends in de chemie*, Dordrecht, Elsevier bedrijfsinformatie B.V.
- PriceWaterhouseCoopers, Ondernemingsanalyses (2002a), *Trends in de bouw*, Dordrecht, Elsevier bedrijfsinformatie B.V.
- PriceWaterhouseCoopers, Ondernemingsanalyses (2002b), *Trends in de retail*, Dordrecht, Elsevier bedrijfsinformatie B.V.
- PriceWaterhouseCoopers, Ondernemingsanalyses (2002c), *Trends bij banken*, Dordrecht, Elsevier bedrijfsinformatie B.V.
- Social Economic Council (2001), *Corporate Social Responsibility: A Dutch Approach*, Assen, Van Gorcum.
- Stoffele, N.C.G.M., (2002), Corporate Social Responsibility in Noord-Brabant and Zeeland, Masters Thesis, Tilburg University.
- Vlek, C.A.J., E.M. Steg, D. Feenstra, W. Gerbens-Leenis, S. Lindenberg, H. Moll, A. Schoot Uiterkamp, F. Sijtsma en A. Van Witteloostuijn (2002), Een praktisch model voor duurzaam bedrijfspresteren, *Economisch Statistische Berichten*, 87(4369), pp. 524-527.
- Waddock, S.A. and S.B. Graves (1997), The Corporate Social Performance-Financial Performance Link, *Strategic Management Journal*, pp. 303-319.