R&D – the missing link between Corporate Social Performance and Financial Performance?

The relationship between corporate social performance (CSP) and financial performance (CFP) has been analyzed for decades. Despite these efforts, the results remain ambiguous. The omission of important variables in the econometrical estimation process is expected to be one reason for the mixed results. Accordingly, this study is focused on the role of R&D as pointed out by Siegel and McWilliams (2000). Therefore, a systematic literature and vote count review is conducted in order to evaluate the acceptance, significance and influence of R&D as a control variable for analyzing the relationship between CSP and CFP. The results of this review are contrary to the expectations of Siegel and McWilliams (2000). While a growing number of significant and positively directed R&D variables can be found in estimation models since their indication, the integration has rather led to an increase in findings with a positive CSP-CFP relation instead of findings with a neutral relation. These results have been used to derive implications for future research with regards to the materiality and operationalization of R&D.

1. Introduction

For decades, the relationship between corporate social performance (CSP)[1] and corporate financial performance (CFP) has been controversially discussed in academia. Different empirical models have been specified and tested. However, to date there are still questions whether this relationship is positive, negative or neutral (Friede et al. 2015). One explanation for the ambiguous results was suggested by Siegel and McWilliams (2000). They argued that the controversy is founded on a misspecification of the underlying research models. They hypothesize that significant variables are omitted from the analysis and, therefore, the estimations are biased and inconsistent. Predominately, their analysis is focused on the omission of corporate innovativeness measured by R&D. The analysis led to the conclusion that the impact of CSP on CFP is neutral when R&D is included in the model as R&D can be regarded as a kind of proxy for CSP. Accordingly, all previous estimation models that do not include a variable that accounts for R&D suffer from misspecification.

The research focus of this paper is directed towards the changes that have been induced by the article of Siegel and McWilliams (2000) and how the perception of R&D as a significant variable has developed. Consequentially, the following analysis aims to provide a systematic literature review on the evolution of estimation models since the indication of the misspecification problem. To guide this analysis the following research questions have been developed:

1. Did the empirical literature on the CSP-CFP relationship incorporate Siegel and McWilliams' suggestion to include R&D as a control variable?
2. Does the inclusion of R&D in fact yield neutral results regarding the CSP-CFP relation as predicted by Siegel and McWilliams?
3. Is the impact of R&D on the regression model statistically significant and does the sign of the coefficient point in a positive direction as predicted by Siegel and McWilliams?

To analyze these questions, the next chapter will provide an overview on important landmarks and developments in this debate. Afterwards, the research methodology and a detailed quantitative analysis of the reviewed literature shall be provided.

[1] Here, Corporate Social Performance will be used as an umbrella term that incorporates social and environmental aspects.

2. The business case for sustainability: empirical landmarks

Analyzing the link between CSP and CFP becomes increasingly important as a significant positive relation could justify the business case for sustainability (Schreck 2011). The business case for sustainability is one side of an overarching debate that is divided into two paradigms (Barnett and Salomon 2012). In this realm, CSP measures are valued as aspects that improve the corporate financial performance and overall "societal expectations" (Günther et al. 2012). Hence, a company should pursue sustainability measures to obtain competitive advantages (Busch and Hoffmann 2011; Orlitzky et al. 2003; López et al. 2007). On the other side, there is the paradigm that highlights the costs of sustainability measures. From this point of view, investing in CSP will not improve the financial performance, but rather lead to additional costs that reduce the financial performance (Friedman 1970). Accordingly, a company should only invest in CSP measures as far as it is obliged to, otherwise it would destroy value (McWilliams and Siegel 2000).

Based on these different views an extensive literature has evolved that led rather to a continuum between the two views than to a clear solution. Different quantitative and qualitative research methodologies have been employed. Basically, studies are focused on the validation of the basic causality between CSP and CFP (Endrikat et al. 2014; Günther et al. 2012). Therefore, empirical analyses remain at the core of the debate, but to date fail to present a clear evidence for one perspective (Barnett and Salomon 2012; Surroca and Tribo 2005). These results are categorized in the three categories: positive, negative and neutral relationship (Aras et al. 2010; Günther et al. 2012). Increasingly, research is focused on methodological problems of the empirical analyses to explain the differences in the results (King and Lenox 2001; Elsayed and Paton 2005).

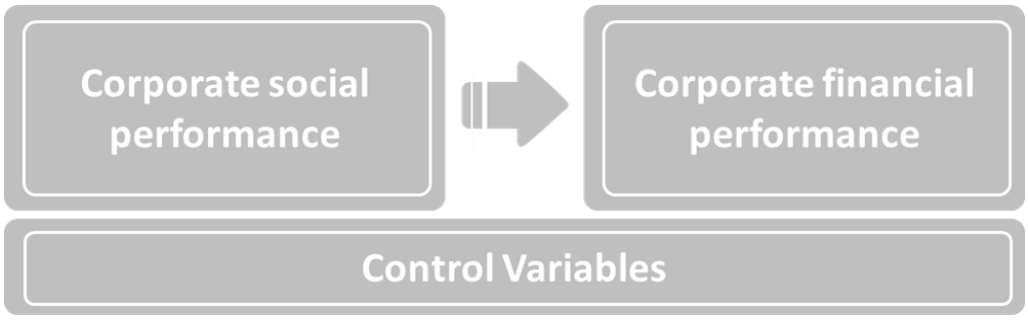

Illustration 1: Basic Econometric Model, own illustration.

The basic econometric models regress the impact of CSP on CFP as indicated by Illustration 1. Both variables are considered to be rather broad meta-constructs (Friede et al. 2015; Margolis and Walsh 2003). Varies operationalizations have been applied to measure them. CSP variables are generally categorized into individual and aggregated CSP measures (Tranfield et al. 2003; Makni et al. 2009). In that regard, social as well as environmental dimensions or indices combined out of both are used to describe CSP (López et al. 2007; Orlitzky et al. 2003; Surroca and Tribo 2005; Alvarez 2012). Again, also these constructs are rather broad and there is no standardized way available that defines clear indicators for social and environmental aspects. Thus, many different proxy variables like reputation indices, philanthropy measures or scores provided by sustainability analysts have been used by scholars (Fauzi 2009). One can see that CSP is a multidimensional construct with no clear underlying measurement method.

A growing focus can be seen with regards to the individual CSP measures in order to assess the specific aspects separately. These analyses are important in order to understand differences between social or environmental CSP measures and their impact on CFP (Mahoney and Roberts 2007; Schreck 2011). However, this multidimensionality increases the complexity of the debate and hampers the comparability as different studies use different constructs to measure CSP.

The measurement of CFP seems to be more straightforward as the financial performance of corporations has been assessed for decades. However, a similar problem has developed. There is no clear agreement on the right measure for CFP. Again, there are different approaches to grasp the financial performance. Mainly, the approaches are divided into accounting and market-based measures (Aras et al. 2010; Marti et al. 2015). Typical accounting based measures are return on assets and return on equity (Günther et al. 2012). The drawback of this approach is its retrospective nature and inconsistency as it is based on accounting principles (Aras et al. 2010). Therefore, measurement via market-based variables has become important in literature. Tobin's Q, Market Value Add or development of stock prices are commonly used indicators for market-based approaches (Margolis et al. 2009). These constructs are rather forward-looking while integrating the shareholder expectations into the debate (Aras et al. 2010).

Additionally, control variables are used to improve the model fit (Garcia-Castro et al. 2010; Andersen and Dejoy 2011). Traditionally, the specification is founded on a close connection to findings from financial econometrics. Based on that, initial models have employed variables to control for risk, size and industry (Ullmann 1985). This set up has often been replicated (Alvarez 2012; Lopatta and Kaspereit 2011; Margolis et al. 2009). However, over time more studies challenged the fitness of these control variables as model misspecifications could be the reason for the heterogeneous results (Surroca and Tribo 2005). Accordingly, more variables have been introduced and tested. Especially, intangible firm resources like R&D for innovativeness or human capital for organizational resources have been increasingly analyzed (McWilliams and Siegel 2000; Surroca and Tribo 2005). Hence, the status quo for the model specification has changed over the years.

Previous reviews have summarized the relationship between CSP and CFP while analyzing the role of control variables (Friede et al. 2015; Margolis and Walsh 2003). Van Beurden (2008) highlighted the relevance of these "moderating variables" and their impact on the significance of the CSP-CFP relation. However, these reviews encounter the impact of the R&D variable from a rather broad point of view based on a limited number of studies. The following analysis aims to improve the transparency specifically with regards to R&D; how they are used and to what results their inclusion leads.

3. Methodology

Based on the research questions outlined before, a systematic literature review has been conducted with specific focus on quantitative findings in the literature in order to analyze the relationship between R&D, CSP and CFP.

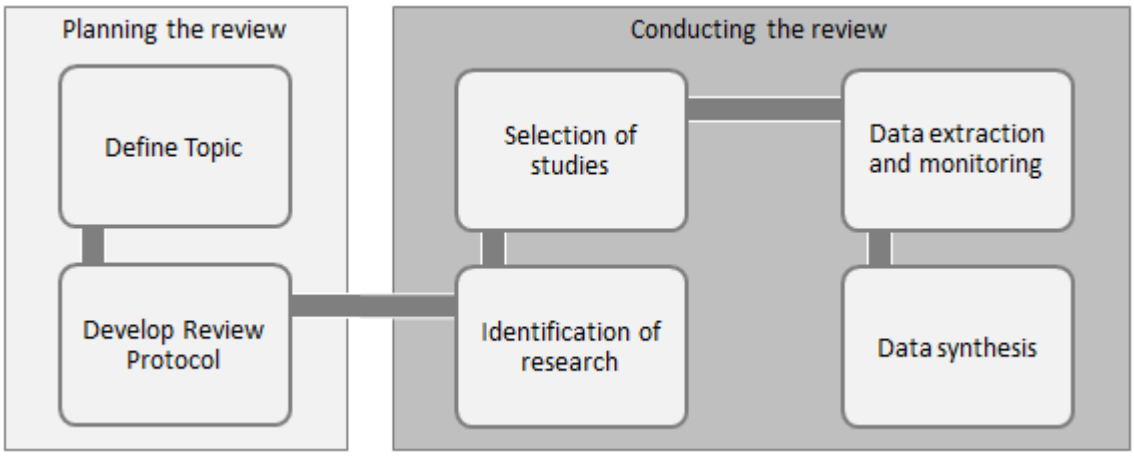

Illustration 2: Review Process, own illustration, based on Tranfield (2003).

Tranfield (2003) provided a practical process for systematic literature reviews. At the beginning, the research question and the focus of this article have been developed as already outlined before. Based on this framework a research protocol has been created that defines the search strategy, initial screening, advanced inclusion and exclusion criteria as well as a draft for the database model.

The search strategy can be divided into two parts: computerized search and manual search. For the computerized search, predefined keywords combined with Boolean algebra have been used to find material articles in a suitable search engine. ABI/INFORM by ProQuest has been used as the main search engine as it is the most complete database available and covers nearly all relevant business and management journals (McWilliams and Siegel 2000). Through predefined keywords, ABI/INFORM provides a detailed overview about relevant articles. The following keywords and syntax has been used to identify material publications:

> *(("Financial performance" AND ("Research and Development" OR R&D) AND (regression OR quantitative) AND ("Corporate Social Responsibility" OR CSR) OR ("Corporate Social Performance" OR CSP) OR (Environmental Social Governance OR ESG)) AND peer(yes)) AND la.exact("ENG")*

For the manual search, reference lists from relevant primary studies have been searched as well as relevant journals for further articles. Additionally, direct correspondence with other researchers have led to an enhance database. General parameters have been predefined for both search strategies in order to secure the quality of the reviewed articles. Therefore, all articles needed to be in English, published in peer-reviewed journals and be published after 2000 coherent with the time Siegel and McWilliams published their article.

Further, advanced evaluation criteria have been predefined to manage the complexity and extent of articles. First of all, screening criteria have been developed as a filter to optimize the number of primary results. In that regard, titles and abstracts have been reviewed whether they are located within the field of CSP/CFP relationship analysis and employ quantitative statistical approaches. This step aims to improve the quality of the included studies regarding the underlying research focus. Afterwards, a set of exclusion criteria has been created for a more profound analysis of the remaining articles. Accordingly, studies have been excluded that synthesize results (either in meta-analysis or literature reviews), do not use regression models (e.g. excluding factor analyses, structured equation models), estimate solely curvilinear models and do not follow the causality that CSP has an impact on CFP.

The literature review has been conducted based on the outlined review protocol. The initial keyword based search request at ABI/INFORM has led to 1,860 results. Additionally, 70 articles have been identified through manual search. In the first part of the selection process, these articles have been screened regarding their relevance based on title and abstract. Accordingly, unsuited publications have been excluded and the number of results decreased to 272 articles. Out of these articles, 97 publications have been identified as the final database after excluding further studies based on their econometrical fitness for this analysis as outlined above. The whole process has been conducted by two researchers in order to account for complementary validation (Tranfield 2003).

The conduction phase is complemented by the data extraction and synthesis. Parameters have been defined in order to systematically extract information on the model design and regression results from the selected articles. Based on the multidimensional character of this debate, most of the studies use more than one regression estimation model to test different specifications. However, not all models are suitable to be integrated into this review as the main focus is directed towards the influence of R&D. Accordingly; a decision tree has been developed in order to structure the selection of relevant models to identify representative models for this analysis. First, all models with multiple CFP measures have been selected. Afterwards, models with aggregated CSP variables have been preferred. If no

aggregated CSP measures have been available, models based on individual measures were taken into account. Models without any CSP measure have been excluded. In the next stage, the use of R&D has been evaluated. All models that included further a measure for R&D were included in the analysis and for models that have not a variable for R&D, the model with the highest R^2 has been selected. Based on this selection process a total of 197 models have been derived from the articles.

The extracted models have been analyzed and synthesized via general descriptive methods as well as by the vote-counting method. The vote counting technique has been often used by scholars to synthesize the findings in this debate (Lin et al. 2009). Hereby, the results of the individual models are counted and structured into three categories. These categories describe the relationship between CSP and CFP as positive, negative or neutral (Elsayed and Paton 2005).

This basic method is much disputed as it does not "correct for sampling and measurement errors" (Friede et al. 2015). Accordingly, the results provide weaker statistical evidence and lead to slightly biased results as more detailed meta-analysis (Wang et al. 2015). However, Elsayed and Paton (2005) as well as Friede et al. (2015) have not found any significant differences between the results of the vote counting compared to meta-analyses. Based on the aim of this review, the vote-counting might be very beneficial as a first step in order to discover new potential for further research. Therefore, it should be rather considered as a well substantiated starting point for an in-depth discussion.

4. Results

Based on the outlined review process 97 publications from 2000 to 2016 have been analyzed. Despite some exceptions, the number of publications has increased recognizable until reaching its peak in 2011 with 12 publications as highlighted in Figure 2. Since that the amount of publications has averaged out on a rather lower level, but gaining more prominence in 2016. This highlights the further importance of this debate in the academic discourse.

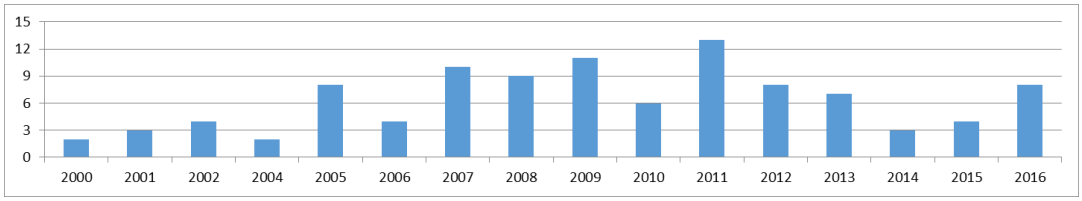

Figure 1: Number of publications per year from 2000-2016 (n = 97).

Out of these publications, 197 models have been extracted and reviewed. On average, around 60% of these models do not include R&D as a control variable while 40% consider R&D in the specification.

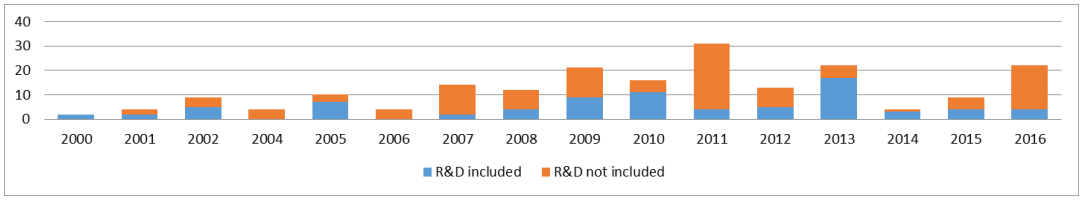

Figure 2: Integration of R&D in research models per year from 2000-2016 (n = 197).

The distribution of models that include R&D and models that exclude R&D does not appear to follow a certain trend as seen in Figure 3. From 2000 till 2005 on average more models included R&D as a control variable. Hence, the integration of R&D has been quite prominent shortly after Siegel and

McWilliams (2000) published their results. However, since 2006 on average more models have not integrated R&D variables. Especially in 2011 and 2016, almost all models do not consider R&D. In sum, one can see that the use of R&D has increased over the years indicating that R&D has become a considerable part of academic discourse. However, there is still a noticeable amount of studies that do not account for R&D. One possible explanation could be related to the data transparency. R&D information is not available for many companies. This often leads to a significant decrease in sample size which shall be avoided.

4.1 Impact of R&D on the Relationship between CSP and CFP

An assessment of the relationship between CSP and CFP as proposed by the second research question is founded on the general understanding that the estimation results of the individual models have been divided into three categories:

1. *Models that find a positive relationship between CSP and CFP*
2. *Models that find a neutral relationship between CSP and CFP*
3. *Models that find a negative relationship between CSP and CFP.*

The evaluation is based on the significance and direction of the CSP variable. A non-significant variable displays a neutral relationship and a significant positive or negative CSP variable vice versa the other categories. Additionally, another categorization has been developed to highlight the differences between the full sample and the sub-samples, where either R&D is included or excluded in the basic model specification.

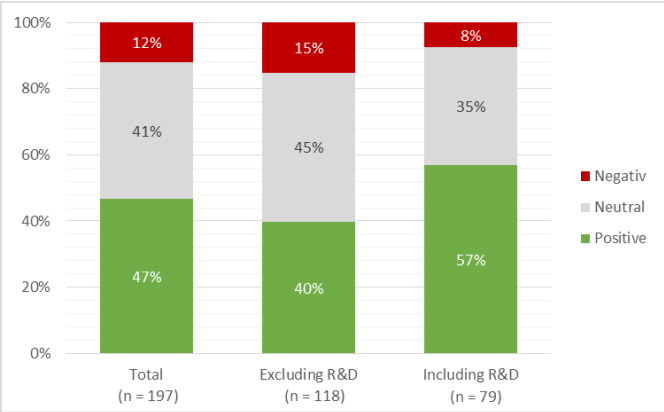

Figure 3: CSP and CFP relation in dependence of R&D.

When looking at the full sample, the relationship between CSP and CFP is consistent with other findings (Friede et al. 2015). 47% of the models indicate a positive relation between CSP and CFP. Further 41% account for a neutral relationship. The remaining 12% estimation models find a negative relationship between corporate social and financial performance. Hence, in general the association between CSP and CFP does not seem to destroy financial value.

Figure 4 highlights the changes in the distribution when R&D is included or excluded. For models where R&D is excluded from the set of control variables, only 40% of the statistical estimations find a positive relationship. Hence, omitting R&D leads to a decrease in positive findings. On the other side, there are slightly more findings that indicate neutral (45%) as well as negative (15%) relationships.

For models where R&D is included, the development is contrary. The integration of R&D leads to an increase in models that find a positive relationship (57%). Accordingly, the amount of neutral (35%) and negative (8%) findings decreases recognizably. These results are different to the prediction by Siegel and McWilliams. According to their proposition, the relationship between CSP and CFP should be neutral when R&D is included in the model. However, the results of this review suggest a different development as the inclusion of R&D leads to a decrease with regards to the neutral relationship. Hence, the correlation between R&D and CSP could be depended upon additional factors.

4.2 Significance and Direction of R&D

Based on the impact on the relationship between CSP and CFP, the models that include R&D (n = 79) have been analyzed more detailed in order to evaluate its robustness as indicated by the third research question. Therefore, in a first step, the significance of the R&D variable within these models has been analyzed. If a variable is considered as being significant, it is estimated to be different from zero and to add explanatory value to the model. Approximately, two third of the R&D variables are significant variables, while one third are non-significant. Hence, the inclusion of R&D seems to be useful based on the overall findings. Second, the directions of the significant R&D variables have been analyzed. Only significant variables have been reviewed as the non-significant variables have no explanatory impact on the model and can rather be considered as neutral. Intuitively, a positive sign for the coefficient of the R&D variable would be reasonable. Consequentially, an increase in R&D expenditures would be expected to lead to an increase in competitive advantage and, therefore, to an increase in financial performance (McWilliams and Siegel 2000). Hence, negative results would be rather counter-intuitive. Here, more than 70% of the significant R&D variables are positive. Thus, the findings are consistent with the theoretical expectations. Consequentially, the results for the R&D variable itself with regards to the significance and direction can be considered as being in accordance with the proposition of Siegel and McWilliams. Nevertheless, the partially different results should be mind for further analyses.

In addition, the significance of the R&D variables has been reviewed with regards to its impact on the relationship between CSP and CFP as outlined in 4.1. When considering only models with significant R&D variables (n = 53), almost two third of the models indicate a positive relationship between CSP and CFP. On the other hand, for models where R&D is non-significant (n = 24), about two thirds of the models are neutral. Hence, the results support the previous findings. Including substantiated R&D variables lead to positive relationships and only non-significant variables would improve the assumption of a neutral relationship.

4.3 Further Impacts on Model Specification

Further assessments have been made with regards to the basic model specification. Therefore, the operationalization of CSP and CFP has been evaluated and changes have been analyzed concerning the inclusion and exclusion of R&D. This analysis shall evaluate potential differences that could come from social or environmental CSP operationalizations as well as short or long-term finance measures.

The operationalization of CSP has been divided into three basic categories: environmental focused CSP; social focused CSP and an index based on social and environmental performance. For the full sample, the distribution between the three variables is quite even. 37% of the models use an index of social and environmental factors to create the CSP variable, where 34% use social and 30% environmental focused CSP measures. These numbers change recognizably, when R&D is excluded. The use of social focused CSP measures decreases by 10%. Hence, when R&D is not integrated in the model, rather environmental measures (36%) or indices of both (41%) are used. Vice versa the results change, when R&D is included in the model set-up. For these models mainly social focused CSP variables are used (48%) and the share of environmental (22%) and indices (30%) variables decreases noticeably. Hence, one can see clear differences in the specification of CSP with regards to the use of R&D.

In the same way, the operationalization of CFP has been analyzed. Here, CFP has been divided into accounting-based and market-based measures. Over all models, 54% of the CFP variables are accounting-based measures and 43% are market-based measures, while 3% of the models do not specifically state the CFP variable. Thus, the split between the two operationalizations is quite even. Complementary, the changes have been analyzed when R&D is included or excluded. However, the inclusion or exclusion does not lead to a change in the distribution among the different

operationalization. Consequentially, there seems to be no relation between the choice of CFP operationalization and R&D.

4.4 Impact of R&D on Research Setting

Furthermore, general research aspects have been analyzed in the review process in order to gain a deeper understanding about the connections between R&D and the research setting. For this reason, the region, CSP database, industry coverage and time horizon of each study has been evaluated.

For the regional coverage of a study, five geographical areas have been selected to understand the origin of the data used for the estimations. For the full sample, most of the studies are based on companies from the United States (49%). For the other regions (World, Europe, Asia, Rest of the World), there is no further trend foreseeable. All of these regions have a similar share slightly above 10%. The most remarkable change that comes from the inclusion or exclusion of R&D can be seen with regards to the United States. When R&D is excluded, the portion of U.S. based analyses decreases to 37%. When R&D is included this figure increases to 69%. This leads to the conclusion that almost all studies that account for R&D are based on the development of U.S. companies.

Additionally, the usage of specific CSP databases has been incorporated in the analysis. Basically, two findings shall be highlighted here. The single most used database is the KLD database (32%) provided by MSCI. Another point to consider is the increasing portion of KLD when the research model includes R&D. Here, 54% of the models use KLD as their database for measuring CSP. Besides, the field of database sources is very heterogeneous. Many different and often country specific databases have been used, which aggravates the comparability between the results.

For the industry coverage and the time horizon a clear trend can be extracted independent from the in- or exclusion of R&D. Almost all studies use datasets that cover multiple industries and more than two thirds of the models use longitudinal data.

5. Options for Methodological Enhancements

Based on the outlined results, new options for methodological enhancements have become apparent. The relationship between CSP and CFP is very complex and the model specification is one of the most important aspects when it comes to analyzing their relation. Many issues are still disputed, but the core of this literature review shall be directed to the role of R&D. The previous results have shown that R&D is an important variable and not all connections are evident or have been predicted correctly. Accordingly, further research is needed to gain more understanding on correct specification and its impact on the estimation model. In this section five complementary possible future research directions shall be outlined:

(1) Meta-analysis to calculate effect sizes and advance the results of this review
(2) Re-test Siegel and McWilliams findings with the same study design and updated longitudinal data
(3) Re-evaluate operationalization of R&D ratio
(4) Evaluate the impact of the CSP operationalization
(5) Create an enhanced model to control for more relevant variables with longitudinal data

Most recognizably, this literature review points out the distress between the predictions by Siegel and McWilliams (2000) and the published empirical results. The heterogeneity in the relationship between CSP and CFP has not eased through the inclusion of R&D. Instead of producing more neutral results, the inclusion of R&D rather leads to more positive results. Hence, a quite contrary view as postponed by Siegel and McWilliams. As this review is based on rather generic evaluation techniques, these findings could be further substantiated through a meta-analysis that calculates the effect sizes to

provide a more significant statistical analysis with regards to the impact of R&D on the relationship between CSP and CFP. Endrikat et al. (2014) followed a similar approach in order to assess the moderating effects of the control variables. However, according to their set-up R&D was rather one of many addressed issues and the changes that result from R&D have not been investigated in detail. Further the difference between models that include and exclude R&D has not been regarded as being significant although it almost reached the 5% acceptance interval (p-value = 6.5%). Nevertheless, there is no further evidence how the inclusion or exclusion of R&D impacts the estimation model or how it affects other parts of the research setting. Hence, such a meta-analysis should not only focus on the generic relationship between CSP and CFP, but rather on the role of a single specific control variable to account for all changes that come from the inclusion or exclusion of this variable.

In order to validate the results by Siegel and McWilliams, it would be an option to reproduce their initial study to validate the results. Therefore, the database should be recreated to allow for an identical model specification and regression analysis. The dataset of Siegel and McWilliams consists of average values for the applied variables from 1991-1996. These average values could be reproduced in a first step to carry out the analysis. In addition, another database could be created where average values are calculated based on a different time horizon. Thus, a time frame from 1991-2016 or smaller more recent sub samples of 5 years each could be possible to validate the findings of the initial sample. Furthermore, instead of calculating average values a panel design could be used to control for unobserved heterogeneity. Re-evaluating the validity of previous research studies is gaining prominence as results are often taken for granted. However, all results underlie certain contexts that could change over time or that are affected by unobserved methodological bias (Rost and Ehrmann 2015; Krämer 2011).

Based on the latest findings, the heterogeneity of results has not flattened. With regards to R&D this issues could also be related to the specific operationalization of R&D that has been proposed by Siegel and McWilliams. There, the R&D expenditures are divided by the sales of the company. This proposition almost has been unchallenged since then. Many scholars use the same variable set-up within their estimation models (Barnett and Salomon 2012; Garcia-Castro et al. 2010). Accordingly, the level of fit for this operationalization has been addressed very seldom within this debate. Surroca et al. (2010) and Andersen & Dejoy (2011) use two different approaches in order to include R&D. Surroca et al. (2010) divide R&D expenditures by the number of employees. Andersen & Dejoy (2011) do not use a ratio, but the total expenditures. However, the underlying reason is not further investigated, although a misspecification of the variable could also significantly influence the estimation results. This could also be a potential reason for the non-significance or negative direction of the R&D variables as indicated in some models and could be the starting point for further evaluations.

Additionally, this literature review highlights how the operationalization of the CSP variable differs between models where R&D is included and excluded. More social-based CSP measures are used for models where R&D is included. Presumably, this fact is related to the availability of data sources, as most of these studies are based on the data provided by KLD which offers rather social-based CSP information. However, the inclusion and exclusion of R&D should be further analyzed for different operationalizations of CSP in order to understand how the design of the variable impacts the relationship between CSP and CFP and how this relates to the significance of R&D. Accordingly, aggregate CSP measures should be used for initial analysis. Siegel and McWilliams constructed a dummy variable for CSP by evaluating whether a company is included in the Domini 400 Social Index or not, while other models use evaluations by third party rating agencies. Tracing back these individual approaches might be helpful in order to understand how the results are impacted by the design of CSP.

If all these further specification issues have been tested, the way is paved to carry out an additional robust panel estimation that accounts for the new aligned R&D variable as well as for the remaining

control variables that are considered to be significant for estimating the corporate financial performance. In the previous estimations, the models are considered to be more closely in alignment with the model specification Siegel and McWilliams, while since then further analyses have pointed out the relevance of other variables (e.g. multiple CFP variables). Accordingly, this basic model could be enhanced by a model that is based on a longitudinal time-series and accounts for all lessons learned from model specification issues.

6. Limitations

The limitations are mainly coming from two directions. First, the process of the systematic literature review follows certain assumptions and steps that could lead to biases. Second, the vote counting technique has some statistical weaknesses that shall not be neglected.

Limitations of Systematic Literature Reviews

Basically, like every review, this analysis underlies specific assumptions and characteristics. First, due to set of exclusion criteria, some studies are withheld from this review (e.g. articles published before 2000). Accordingly, some aspects from previous datasets are underestimated although some parts of this debate are deeply connected to the findings of these previous studies.

With regards to the selection process, standardized dual review strategies have been developed in order to identify all relevant publications. However, the screening is still based on manual decision making and there are potentially more articles that haven't been identified, as not all articles are available within the ABI/inform. In addition, a similar problem arises regarding the selection of estimation models. Simplifying a comprehensive set of models is difficult as all analyses underlie a different model specification and research design. Hence, due to the focus on the R^2 and the accordance to the basic research model outlined by Siegel and McWilliams, additional models could have been neglected.

Limitations of Vote Counting

Vote Counting is a very pragmatic approach in order to gain a systematic overview in a consistent and fast way. Therefore, this approach has its usefulness. However, the first problem comes with the selection of relevant models for the analysis. The model extraction has followed a standardized procedure, but still could underlie some bias caused by subjectivity (Lin et al. 2009). In addition, the synthesis of the results is not based on any further statistical aggregation methods. Breaking down complex statistical multivariate analyses into three distinct categories is a very generic approach. Here, a lot of information is neglected and purely reduced to a simple categorization. From statistical considerations, this information is necessary in order to account for the specificities of the individual estimations. Hence, differences in sample sizes or effect sizes are not reviewed, which could lead to biases in the synthesizing process (Lin et al. 2009; Elsayed and Paton 2005). Accordingly, scholars advocate to (at least additionally) conduct more sophisticated meta-analysis. One of the most prominent studies in that regard has been published by Saeidi et al. (2015). Hence, for a deeper analysis of the impact of R&D on the relationship between CSP and CFP the calculation of the effect sizes while controlling for additional factors like the sample size could be very beneficial in order to draw further substantiated results.

7. Conclusion

Decades of research have substantiated that the relationship between CSP and CFP is indeed complex. Over the years different additional aspects have been introduced into this debate in order to improve the transparency on this relationship. Guided by three complementary research questions, this systematic literature review is evaluating the role of R&D in order to review the propositions made by Siegel and McWilliams (2000) and identify potential connection points for further research.

Did the empirical literature on the CSP-CFP relationship incorporate Siegel and McWilliams' suggestion to include R&D as a control variable?

The results indicate mixed results with regards to the model specification. Directly after the indication of the relevance of R&D by Siegel and McWilliams, most of the publication included R&D in their research models. Over all years, around two third of the studies include R&D. Hence, it can be viewed as an important moderating variable. However, there is still a noticeable amount of models that do not consider R&D in their research design. Additionally, it seems as if recently less models account for R&D. Consequentially, one could argue that die model specification did change, but the change process is ongoing as the influence of R&D is still difficult to grasp.

Does the inclusion of R&D in fact yield neutral results regarding the CSP-CFP relation as predicted by Siegel and McWilliams?

It has been hypothesized that the inclusion of R&D would lead to a neutral relationship between CSP and CFP as found by Siegel and McWilliams. However, this analysis has shown that the integration of R&D rather leads to an increase in positive relationships and a noticeable decrease in neutral results (i.e. contrary to the assumptions). These findings become even stronger for models where R&D is one of the main significant variables. A small increase in the neutral relationship could only be obtained from models where R&D is excluded from the set of control variables. Consequentially, R&D leads to different results with regards to this debate, but in favor of a positive CSP-CFP relationship.

Is the impact of R&D on the regression model statistically significant and does the sign of the coefficient point in a positive direction as predicted by Siegel and McWilliams?

The important role of R&D is founded on the assumption that R&D has a significant positive impact on the corporate financial performance. The results of this review highlight that most of the R&D variables are significant as well as positive. Hence, in accordance with the general assumptions, R&D can be considered as a valuable moderating factor to add explanatory power to the estimation model. However, about one third of the models show ambiguous results. Here, more evidence is needed regarding the operationalization of R&D. Some scholars use R&D intensities to account for size effects, but use different size deflators (e.g. sales, employees), while other models include pure R&D costs.

Based on these findings, five additional options for methodological enhancements have been identified in order to improve the transparency with regards to the role of R&D. First, an enhanced meta-analysis might be suitable in order to foster the robustness of the results of this review. Second, the original analysis by Siegel and McWilliams could be replicated with updated longitudinal data to verify the results. Third and fourth, methodological aspects regarding the operationalization of the R&D and CSP constructs might be re-evaluated in order to understand the impact of the research design. And fifth, the findings of all these analyses could be re-integrated into an enhanced model that controls for all relevant moderating variables to re-evaluate the relationship between CSP and CFP in a robust model.

In sum, this review highlights the role of R&D in the CSP-CFP debate. The underlying analysis indicates deviant results regarding the impact of R&D in contrast to the propositions by Siegel and McWilliams (2000). There seem to be more influencing factors in these analyses. Accordingly, further research is needed in order to design a model that accounts for the quite evident need of a R&D variable, but carefully acknowledges potential problems in the operationalization to avoid additional noise and bias in the estimation process.

Bibliography

Alvarez, Isabel Gallego (2012): Impact of CO 2 Emission Variation on Firm Performance. In: *Bus. Strat. Env.* 21 (7), pp. 435–454. DOI: 10.1002/bse.1729.

Andersen, Margaret L.; Dejoy, John (2011): Corporate social and financial performance. The role of size, industry, risk, R&D and advertising expenses as control variables. In: *Business and society review* 116 (2), pp. 237–256.

Aras, Güler; Aybars, Aslı; Kutlu, Ozlem (2010): Managing corporate performance. In: *Int J Productivity & Perf Mgmt* 59 (3), pp. 229–254. DOI: 10.1108/17410401011023573.

Barnett, Michael L.; Salomon, Robert M. (2012): Does it pay to be really good? Addressing the shape of the relationship between social and financial performance. In: *Strat. Mgmt. J.* 33 (11), pp. 1304–1320. DOI: 10.1002/smj.1980.

Busch, Timo; Hoffmann, Volker H. (2011): How Hot Is Your Bottom Line? Linking Carbon and Financial Performance. In: *Business & Society* 50 (2), pp. 233–265. DOI: 10.1177/0007650311398780.

Elsayed, Khaled; Paton, David (2005): The impact of environmental performance on firm performance. Static and dynamic panel data evidence. In: *Structural Change and Economic Dynamics* 16 (3), pp. 395–412. DOI: 10.1016/j.strueco.2004.04.004.

Endrikat, Jan; Guenther, Edeltraud; Hoppe, Holger (2014): Making sense of conflicting empirical findings. A meta-analytic review of the relationship between corporate environmental and financial performance. In: *European Management Journal* 32 (5), pp. 735–751. DOI: 10.1016/j.emj.2013.12.004.

Fauzi, Hasan (2009): Corporate social and financial performance: Empirical evidence from American companies. In: *Globsyn Management Journal, Forthcoming*.

Friede, Gunnar; Busch, Timo; Bassen, Alexander (2015): ESG and financial performance. Aggregated evidence from more than 2000 empirical studies. In: *Journal of Sustainable Finance & Investment* 5 (4), pp. 210–233. DOI: 10.1080/20430795.2015.1118917.

Friedman, Milton (1970): The social responsibility of business is to increase its profits. In: *The New York Times Magazine*.

Garcia-Castro, Roberto; Ariño, Miguel A.; Canela, Miguel A. (2010): Does Social Performance Really Lead to Financial Performance? Accounting for Endogeneity. In: *J Bus Ethics* 92 (1), pp. 107–126. DOI: 10.1007/s10551-009-0143-8.

Günther, Edeltraud; Hoppe, Holger; Endrikat, Jan (2012): Corporate financial performance and corporate environmental performance: A perfect match?

King, Andrew A.; Lenox, Michael J. (2001): Does it really pay to be green? An empirical study of firm environmental and financial performance: An empirical study of firm environmental and financial performance. In: *Journal of Industrial Ecology* 5 (1), pp. 105–116.

Krämer, Walter (2011): The Cult of Statistical Significance – What Economists Should and Should Not Do to Make their Data Talk. In: *Schmollers Jahrbuch* 131 (3), pp. 455–468. DOI: 10.3790/schm.131.3.455.

Lin, Chin-Huang; Yang, Ho-Li; Liou, Dian-Yan (2009): The impact of corporate social responsibility on financial performance. Evidence from business in Taiwan. In: *Technology in Society* 31 (1), pp. 56–63. DOI: 10.1016/j.techsoc.2008.10.004.

Lopatta, Kerstin; Kaspereit, Thomas (2011): The World Capital Markets' Perception of Sustainability and the Impact of the Financial Crisis.

López, M. Victoria; Garcia, Arminda; Rodriguez, Lazaro (2007): Sustainable Development and Corporate Performance. A Study Based on the Dow Jones Sustainability Index. In: *J Bus Ethics* 75 (3), pp. 285–300. DOI: 10.1007/s10551-006-9253-8.

Mahoney, Lois; Roberts, Robin W. (2007): Corporate social performance, financial performance and institutional ownership in Canadian firms. In: *Accounting Forum* 31 (3), pp. 233–253. DOI: 10.1016/j.accfor.2007.05.001.

Makni, Rim; Francoeur, Claude; Bellavance, Francois (2009): Causality Between Corporate Social Performance and Financial Performance: Evidence from Canadian Firms. In: *Journal of business Ethics* 89 (3), pp. 409–422. DOI: 10.1007/s10551-008-0007-7.

Margolis, Joshua D.; Elfenbein, Hillary Anger; Walsh, James P. (2009): Does it pay to be good… and does it matter? A meta-analysis of the relationship between corporate social and financial performance.

Margolis, Joshua D.; Walsh, James P. (2003): Misery loves companies: Rethinking social initiatives by business. In: *Administrative science quarterly* 48 (2), pp. 268–305.

Marti, Carmen Pilar; Rovira Val, M. Rosa; Drescher, Lisa G. J. (2015): Are Firms that Contribute to Sustainable Development Better Financially? In: *Corporate Social Responsibility & Environmental Management* 22 (5), pp. 305–319. DOI: 10.1002/csr.1347.

McWilliams, Abagail; Siegel, Donald (2000): Corporate social responsibility and financial performance: correlation or misspecification? In: *Strategic Management Journal*, pp. 603–609.

Nelling, Edward; Webb, Elizabeth (2009): Corporate social responsibility and financial performance: the "virtuous circle" revisited. In: *Review of Quantitative Finance and Accounting* 32 (2), pp. 197–209.

Orlitzky, Marc; Schmidt, Frank L.; Rynes, Sara L. (2003): Corporate social and financial performance: A meta-analysis. In: *Organization studies* 24 (3), pp. 403–441.

Rost, Katja; Ehrmann, Thomas (2015): Reporting Biases in Empirical Management Research The Example of Win-Win Corporate Social Responsibility. In: *Business & Society*, S. 0007650315572858.

Saeidi, Sayedeh Parastoo; Sofian, Saudah; Saeidi, Parvaneh; Saeidi, Sayyedeh Parisa; Saaeidi, Seyyed Alireza (2015): How does corporate social responsibility contribute to firm financial performance? The mediating role of competitive advantage, reputation, and customer satisfaction. In: *Journal of Business Research* 68 (2), pp. 341–350. DOI: 10.1016/j.jbusres.2014.06.024.

Schreck, Philipp (2011): Reviewing the Business Case for Corporate Social Responsibility: New Evidence and Analysis. In: *Journal of business Ethics* 103 (2), pp. 167–188. DOI: 10.1007/s10551-011-0867-0.

Surroca, Jordi; Tribo, Josep A. (2005): The Corporate social and financial performance relationship: What's the ultimate determinant? In: *Academy of Management Annual Meeting Proceedings*, pp. E1-E6-E6. DOI: 10.5465/AMBPP.2005.18783433.

Surroca, Jordi; Tribó, Josep A.; Waddock, Sandra (2010): Corporate responsibility and financial performance. The role of intangible resources. In: *Strat. Mgmt. J.* 31 (5), pp. 463–490. DOI: 10.1002/smj.820.

Tranfield, David; Denyer, David; Smart, Palminder (2003): Towards a methodology for developing evidence-informed management knowledge by means of systematic review. In: *British journal of management* 14 (3), pp. 207–222.

Ullmann, Arieh A. (1985): Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance of US firms. In: *Academy of management review* 10 (3), pp. 540–557.

Wang, David Han-Min; Chen, Pei-Hua; Yu, Tiffany Hui-Kuang; Hsiao, Chih-Yi (2015): The effects of corporate social responsibility on brand equity and firm performance. In: *Journal of Business Research* 68 (11), pp. 2232–2236. DOI: 10.1016/j.jbusres.2015.06.003.