Impact of Peer Effect on ESG and Pillar scores of firms

Nicholas Nord Drange & Biplab Nath

Supervisor: Darya Yuferova
Assistant Professor, Department of Finance

PREFACE

The purpose of this thesis is to investigate the impact of peer effects on a firm's ESG performance. In order to quantify ESG performance we consider the enhanced ESG scores published by Thomson Reuters (now Refinitiv).

The preliminary data of ESG and pillar scores was downloaded from Datastream. The financial data was supplemented from Compustat. The data was cleaned and sorted using R programming and Excel. The regressions and related empirical testing were carried out on STATA.

The thesis was an excellent opportunity to utilize the skills that were imparted during the Masters degree. It has assisted us in understanding the rigours involved in creating datasets and empirical analysis thereof. We have also gained a deeper understanding of specific topics in corporate governance dealing with sustainability, CSR and ESG.

We would like to express our sincere gratitude to our supervisor, Assistant Professor Darya Yuferova, who has been a constant source of inspiration and guidance. During our initial months we were struggling to create the final panel data. But she expressed her confidence in us and inspired us to complete the task that we had set out for ourselves.

We would also like to express our thanks to Fredrik Kavli, researcher at NHH Bibliotek, who relentlessly pursued Thomson Reuters on our behalf so that we could gain access to the ESG database.

We hope that the thesis will be useful for all relevant stakeholders in this area of study and especially for those who passionately view ESG behaviour as a fundamental duty at the individual, organisational and national level.

Happy reading!

Bergen 2019

---------- ----------

(Nicholas Nord Drange) (Biplab Nath)

Impact of Peer Effect on ESG and Pillar scores of firms

Nicholas Nord Drange & Biplab Nath

Bergen 01 June 2019

ABSTRACT

This paper contributes to existing literature by demonstrating that the peer effect of ESG performance as well as pillar dimension-performance positively affect the ESG behaviour of firms. We show a positive and statistically significant relationship between the ESG performance of a firm with the ESG scores of its peers. We detect that the relationship is robust after controlling for fixed effects in our regression models. Similarly, we find a positive and statistically significant relationship in the E, S and G dimensions as well. Furthermore, we conducted two supplemental tests. In the first test, we find that out of three pillars scores, the Social and Governance scores have the largest and significant peer effect on a firm's ESG score. In the second test, we show that if a firm has an ESG score which is lower than the average scores of the peer group, then it experiences a higher peer effect than it would have if its score were above the peer average.

Table of Contents

Peer Effect on ESG and Pillar scores of firms

1 Introduction

Contemporary research has demonstrated that firms can derive competitive advantage from ESG[1] or from such socially responsible behaviour (Porter and Kramer, 2006; Du, Bhattacharya and Sen, 2011). This implies that, competition may have a bearing on a firm's ESG behaviour. Hence, a relevant question is whether the ESG performance of a firm is influenced by the ESG performance of its competitors, or, in other words, is there a peer effect of ESG behaviour? And furthermore, is there a peer effect of the ESG dimension[2] within the corresponding ESG dimension of a firm? This is the fundamental question that drives us to conduct this research.

This study investigates if a firm's ESG score is influenced by the ESG scores of its peers. It further examines if the three components[3] of the peers' ESG performance, influence a firm's respective ESG dimension.

We employ yearly ESG scores published by Thomson Reuters Financial & Risk Business unit (hereafter referred by its new name: Refinitiv)[4] as a metrics for evaluating a firm's ESG performance. A higher overall[5] ESG score reflects better ESG performance for that particular year and likewise with respect to E, S and G dimension[6] scores (hereafter referred as pillar scores). In order to measure peer effects, we employ the simple average[7] of ESG scores of a firm's peer group as a proxy for peer effect in this study.

The central finding in our analysis is that the average of the previous year's ESG scores of the peers of a firm has a statistically significant and positive effect on the ESG score of the firm in the present year. Likewise, we found positive and significant peer effects when we decompose ESG into its pillar dimension scores. Our findings are consistent with our hypotheses and also with existing literature. Cao et al (2018) and Liu & Wu (2016) conduct

[1] ESG is the acronym for Environmental, Social and (Corporate) Governance
[2] E, S and G dimensions refer to Environmental, Social and Governance pillars respectively.
[3] The three components namely: Environment, Social or Governance aspects (hereafter referred as ESG dimensions)
[4] Thomson Reuters Financial & Risk Business announces new company name: Refinitiv https://www.thomsonreuters.com/en/press-releases/2018/july/thomson-reuters-financial-and-risk-business-announces-new-company-name-refinitiv.html
[5] The term ESG score will always imply overall ESG score
[6] E, S and G dimensions refer to Environmental, Social and Governance pillars respectively.
[7] Hereafter, the term 'average' will always imply simple average unless otherwise specified as weighted average.

studies on the peer effect of CSR. They demonstrate that the socially responsible behaviour of competitors positively affects the CSR behaviour of a firm.

Our investigation employs control variables using financial data primarily collected from Datastream (Thomson Reuters) and Compustat. The main dataset was constructed using the Refinitiv ESG database for the period from 2009 to 2017. We employed the Fama-French 49 Industry classification scheme in order to create peer groups for each firm. Our main sample consists of ESG and financial data of 3450 public firms across all 49 industry groups.

For robustness checks, first, we repeat the tests using weighted average scores of the peers instead of simple averages. Second, we test the models employing another peer grouping method, namely, the Hoberg-Phillips text-based industry classification. Thus, we create a separate dataset. Third, we conduct tests by adding a second lag of the independent variable to the original model. The results of the robustness tests are discussed in Section 7.

Thereafter, having tested our hypotheses and conducted robustness checks, we performed some additional tests as a supplemental analysis. Subsequently, testing the original models that we had specified, and observing the results to be in conformity with our hypotheses, further questions were generated. First, we conducted another test to discern which of the three pillar scores of the peers drives the ESG score of a firm. Our findings show that the peer effect of the Social pillar is highest, followed closely by Governance pillar score.

Second, in the supplemental analysis, we estimate the peer effect when a firm has an ESG score above the peer group average. The results demonstrate that the peer effect experienced by a firm is smaller if a firm's ESG score is above the average of the peer-group. Ipso facto, if a firm's ESG score is less than the peer average then it experiences a higher peer effect to improve its own score. However, we do not find a similar relationship in the pillar scores.

The principal contribution of this study is that our estimated models clearly establish a peer effect in not only overall ESG performance, but also with respect to performance in ESG dimensions. Previous literature on ESG peer effects has mostly focussed on the financial and competitive positioning aspects (Leary and Roberts, 2014; Graves and Waddock, 1997; Ferrero-Ferrero et al, 2016).

As far as we know, there has been no prior research on our specific topic of peer effects of ESG pillar dimensions. However, we notice that three research papers bear some degree of resemblance to our study.

Ferrero-Ferrero et al (2016) decompose ESG in terms of pillar scores. But, the purpose of their research is fundamentally different as they examine ESG pillar scores with respect to firm performance; whereas our study focuses on peer effect of ESG pillar scores on ESG performance. Cao et al (2018) and Liu & Wu (2016) examine the peer effects of CSR behaviour, whereas our study deals with ESG performance with particular focus on the three pillar dimensions. Cao et al (2018) and Liu &Wu (2016) employ CSR ratings from the KLD[8] database without further examination in terms of pillar scores since the CSR ratings do not have further clarity in decomposition similar to the ESG dimensions.

The narrative flow of the paper is as follows. In Section 2 we conduct a review of extant literature. We formulate our research questions and hypotheses thereof in Sections 3. Thereafter, we present the four models corresponding to our four hypotheses and methodology in Section 4. This is followed with a description of the data collection and dataset construction in Section 5. Then we furnish the summary statistics and provide a detailed discussion of the results in Section 6. Thereafter, in Section 7, we test the robustness of our models and also include supplemental tests. And finally, we conclude with a discussion of the implications and limitations of our study in Section 8.

2.0 Literature Review

Of late, keeping in sync with the growing interest, substantial academic research has been conducted on ESG, CSR, sustainability and other such socially responsible behaviour. Much of the research on socially responsible behaviour has primarily focussed on financial benefits and firm performance aspects. It appears that not much research has been conducted on the peer effect of socially responsible behaviour and how it influences firms to conform or react in a certain manner. More specifically, we observe that there is no research available regarding the peer effect of ESG scores on a firm's ESG performance. Having identified this area as less researched, we investigate further in order to find a specific research gap.

In this section we will scrutinize extant literature which relate to the salient elements of our study. Accordingly, we have divided this section into three sub-sections and a summary. First, we will discuss the relationship of socially responsible behaviour and firm performance. Second, we will examine existing literature pertaining to peer effects. Third, will be a comparative analysis between ESG and CSR ratings.

[8] KLD (Kinder, Lydenberg, Domini). KLD Research & Analytics, Inc database now acquired by MSCI. Inc .

2.1 SRB and Firm performance

The increase of public awareness towards environmental, social and corporate governance issues in the past few decades has compelled firms to step forward and disclose their performance on various ESG parameters**.** The ESG ratings provided by various agencies[9] are sought by investors, institutions, NGOs, public and business managers for making investment decisions. Stakeholders across the globe have been explicit in their demands for better performance in terms of socially responsible behaviour.

The term 'socially responsible behaviour' implies all such activities that firms are involved with, which aim at enhancement of lives of their stakeholders. Such socially responsible activities refer to a wide spectrum of engagements such as welfare, social and economic improvement, environmental protection, customer concern, ethical business practices, employee welfare, direct philanthropy, internal audit etc. Therefore, it is assumed hereafter that, the term 'socially responsible behaviour' (referred as SRB) will subsume every such activity/ engagement associated with ESG, CSR, SRI, MRI, Sustainability, Ethical investing, behaviour[10] etc which is deemed as socially responsible.

Graves and Waddock (1990 & 1994) demonstrate through empirical analysis that there exists a positive correlation between CSR behaviour and firm performance and thus the shift from the traditional 'myopic institution theory'[11] to the institutional control of enterprises. This reinforced the stakeholder theory as it underscores the importance of institutional stakeholders (Freeman, 2004). Ferrell, Liang and Renneboog, (2016) refute the agency approach view of CSR by demonstrating that firms with higher CSR ratings have fewer agency problems. Fatemi et al (2015) provide evidence to support the claim that CSR activities can create value for a firm.

Friede, Busch and Bassen (2015) conduct an extensive analysis of over 2000 empirical studies which point towards a positive correlation between ESG and financial performance of firms. In this regard Graves and Waddock (1997) report of empirical evidence to support a positive correlation between CSR and financial performance. Furthermore, Gillan, Hartzell, Koch and

[9] Major ESG/ CSR rating agencies are Sustainalytics, MSCI (Morgan Stanley Capital International), Thomson Reuters (Refinitiv), FTSE Russell, Vigeo-EIRIS, RobecoSAM, ISS-oekom (Escrig-Olmedo et al, 2019)

[10] SRI (Socially Responsible Investment), MRI (Morally Responsible Investment), CSR (Corporate Social Responsibility), ERI (Ethically Responsible Investment)

[11] This refers to the hypothesis advanced by Drucker (1986); Graves (1988); Loescher (1984) and others, that managers make decisions for the short term and believed that if firms invested in socially responsible activities, they would be penalized by the investors.

Starks (2010) also find evidence that operating performance, efficiency, and firm value increase with stronger ESG performance.

Ferrero-Ferrero et al (2016) conduct extensive empirical analysis to demonstrate that firms investing in ESG activities not only improve their financial performance, but also can employ it as a competitive advantage. Furthermore, Lins, Servaes and Tamayo (2017) provide empirical evidence that during the 2008-09 financial crisis, high-CSR firms witnessed more growth and profitability than low-CSR firms.

2.2 Peer Effects

Human, as well as organisational interactions are consistently characterized by peer effects or social norms in various forms. Brown, Ivković, Smith and Weisbenner, (2008) provide evidence that the financial decision by an individual, of dealing in stocks, is influenced by that of the community, or in other words, peer influence. Zhang et al (2018) document how the peer effects of various aspects in our daily lives, such as household, neighbourhood and workplace influence our financial decision. Manski (1993) mentions that peer influences are associated with endogenous effects and provides various terms such as "bandwagon" , "contagion", "herd behaviour" while referring to the reflection problem. Although he refers to the issue of peer effects at an individual level, however, researchers extend the applicability when examining peer effects of firms. Leary& Roberts (2014) apply Manski's model and demonstrate that peer effects do influence a firm's capital structure and financial decisions. Hornuf &Schwienbacher (2018) provide evidence that investors make decisions based on similar decisions by other investors. Adhikari &Agrawal (2018) establish that a firm's dividend and share repurchase policy is influenced by the payout policies of its industry peers. Jenter &Kanaan (2015) demonstrate that CEOs are fired by the board of directors based on the performance of other CEOs in the industry even when the below par performance of the CEO is due to industry specific factors beyond the CEO's control.

Hence, we observe that social norms or peer effects, are all pervasive even at firm and organisational level. This led us to question whether the ESG performance of a firm is influenced by the ESG behaviour of its competitors; to what extent can peer effects explain the ESG performance of a firm?

2.3 ESG and CSR

The difference between CSR and ESG is much debated as there exists a degree of subjectivity in their context, and therefore in their definitions. Bowen (1953) is one of the first to propose

a definition for the 'social responsibilities of businessmen'. Carroll (1999) conducts a comprehensive review of literature tracing the genesis and proliferation of various definitions of CSR. He concludes that CSR has progressed into alternative themes such as corporate social performance (CSP), corporate social responsiveness, business ethics and such terms which have become commonplace. Van Marrewijk (2003) states that it is pointless to search for an all-inclusive definition for the terms CSR and Corporate Sustainability. He opines that, owing to the complexities and wide spectrum of organisations and their respective goals, it is prudent that definitions should be dynamic, and accordingly, modified for specific circumstances and requirements. Van Marrewijk (2003) also mentions that in the past, Corporate Sustainability and CSR were treated as different terms with the former universally accepted as primarily concerned with environmental issues while CSR referred to social dimensions. Subsequently, the two terms appear to have converged. This illustrates the subjectivity of the term CSR (Krüger, 2015); it has different interpretations for different stakeholders. CSR has also been understood to imply social and environmental dimensions while ESG has an additional dimension of corporate governance (Del Bosco & Misani, 2016). On the other hand, Stellner et al, (2015) state that, while there is no universally accepted definition of CSR, they prefer to regard CSR as encompassing the three dimensions of environmental, social and corporate governance factors.

We also observe a lack of convergence in terms of the accuracy and reconciliation for CSR and ESG scores awarded by various rating agencies. Dorfleitne, Halbritter, and Nguyen, (2015) conduct an empirical analysis of over 8500 global companies with ESG ratings from three prevalent ESG metrics providers[12]. They found that there is no convergence in the measurements of ESG scores of KLD, Bloomberg and ASSET4. Chatterji, Levine and Toffel, (2009) present evidence that KLD CSR scores on various parameters do not accurately reflect the true level of CSR compliance. They also raise concerns regarding the accuracy of KLD scores with respect to environmental parameters

Thus, it can be appreciated that, amongst the plethora of terms which encompass socially responsible behaviour (SRB), there is an absence of a universally accepted definition which can subsume all related themes of SRB. Moreover, since there is no single universally accepted definition for ESG and CSR, the difference between the two terms is rather blurred. It is further observed that ESG has comparatively more clarity than the term CSR, due to its specific connotation and well-defined dimensions. However, for the purpose of this paper we

[12] The three ESG ratings providers: ASSET4, Bloomberg and KLD

assume that the terms are synonymous. We feel that the main thought processes and intentions of mapping firms and assigning them numeric scores, such as CSR ratings or ESG scores, are the same.

2.4 Summary of Literature Review

The review of literature has revealed some interesting findings relevant to our study. Firstly, that ESG is indicated to be positively linked with firm performance. Secondly, the literature suggests that ESG is being adopted by firms in order to derive competitive advantage. This implies that if ESG is an attribute for competitiveness, then there should be a significant peer effect, which should influence a firm's own ESG endeavours. Thirdly, we discern a research gap with regards to the peer effect of ESG pillar dimensions. With these deductions in mind, we proceed to develop our hypotheses and investigate this research gap.

3.0 Hypotheses Development

Based on prior research mentioned, we expect a positive relationship between a firm's ESG performances with the past ESG scores of its peer group. In other words, the ESG scores of a firm should be positively correlated with the previous year's ESG scores of its peers, and, similar rationale applies with regards to pillar scores. We formulate our research questions as under:

3.1 Research Questions

(a) Does the average ESG score of a peer group in the preceding time-period influence an individual firm within the peer group to improve its ESG score in the present time period?

(b) Does the average pillar score of a peer group in the preceding time-period influence an individual firm within the peer group to improve its pillar score in the present time period?

3.2 Articulation of Hypotheses

The hypotheses are based on the findings from the review of literature and intuitive reasoning. If the peers of an individual firm have improved their average[13] ESG score, then the individual firm may face social backlash and pressure from various stakeholders. This will compel the firm to improve its ESG performance. In other words, the average ESG scores (of the preceding time period)[14] of the peer group will influence the firm to improve its own ESG performance in the present time period. Accordingly we propose our first hypothesis:

Hypothesis1: The average ESG score of a peer group significantly affects an individual firm within the peer group to improve its respective ESG score.

Similarly, we hypothesize that a firm's individual pillar score today will be significantly affected by the average pillar score of its peer group in the preceding time period. Accordingly, we propose the following hypotheses concerning pillar scores:

Hypothesis2: The average Environment pillar score of a peer group compels an individual firm within the peer group to improve its respective Environment pillar score.

Hypothesis3: The average Social pillar score of a peer group compels an individual firm within the peer group to improve its respective Social pillar score.

Hypothesis4: The average Environment pillar score of a peer group compels an individual firm within the peer group to improve its respective Governance pillar score.

4.0 Methodology

We examine the ESG behaviour of firms within their peer groups. A fixed effects model has been employed to control for time invariant effects in peer groups in the panel data. This is necessitated as the model compares firms across various industry segments and there could be significant fixed effects over time due to a number of reasons such as Industry- specific regulations, firm-specific ethos, culture etc. We have incorporated one year lag of the average ESG and pillar scores in the models in order to control for accumulated time effects if any. The first lag of the ESG/pillar scores is of primary interest in our study. In order to test Hypothesis1 we construct the following model (Model 1):

[13] Hereafter, the term 'average' implies both arithmetic as well as weighted average, unless specified otherwise.
[14] ESG rating agencies usually publish the scores on an annual basis. The ESG scores for a particular year of review are published usually in the first half of the subsequent year.

$$ESG\_score_{i,t} = \beta_0 + \beta_1(ESG\_peer)_{i,t-1} + \beta_2\, FirmSize_{i,t} + \beta_3 Debt\ Ratio_{i,t} + \beta_4 ROA_{i.t} + \beta_5(Price/Book)_{i,t} + \alpha_i + \varepsilon_{i,t} \quad \text{……..(Model 1)}$$

The dependent variable $ESG\_score_{i,t}$ is the overall ESG score of a $firm_i$ at year 't' . $(ESG\_peer)_{i,t-1}$ is the average[15] of the overall ESG scores of the peers of $firm_i$ at year 't-1'. $FirmSize_{i,t}$ is equal to the natural logarithmic value of total assets of a $firm_i$ at year 't' while $Debt\ Ratio_{i,t}$ is the ratio of total debt to total assets of a $firm_i$. $ROA_{i.t}$ is the Return of Assets, that is, the ratio of net income to total assets of a $firm_i$. $(Price/Book)_{i,t}$ is the ratio of the share price to the book value per share of a $firm_i$. Lastly, α_i is the unobserved firm specific time invariant fixed effect and $\varepsilon_{i,t}$ is the error term. We control for firm fixed effects in the models. Since, the independent variable is grouped according to industry peer groups using the Fama-French classification; the industry fixed effects will also be controlled for by controlling for firm fixed effects. We have considered employing time dummies in our models to control for time fixed effects. The discussion regarding the reason why we do not include time fixed effects is in Section 7.

For the estimation of the models for individual pillar score we have likewise considered the lag1 of the average pillar score of the peer group for every $firm_i$ in the panel. Accordingly, we will estimate the following three models, namely, Model 2, 3 and 4 in order to test Hypothesis2, 3 and 4 respectively[16]:

$$E_{score\,i,t} = \beta_0 + \beta_1\left(E_{peer}\right)_{i,t-1} + \beta_2\, FirmSize_{i,t} + \beta_3 Debt\ Ratio_{i,t} + \beta_4 ROA_{i.t} + \beta_5(Price/Book)_{i,t} + \alpha_i + \varepsilon_{i,t} \quad \text{……..(Model 2)}$$

$$S_{score\,i,t} = \beta_0 + \beta_1\left(S_{peer}\right)_{i,t-1} + \beta_2\, FirmSize_{i,t} + \beta_3 Debt\ Ratio_{i,t} + \beta_4 ROA_{i.t} + \beta_5(Price/Book)_{i,t} + \alpha_i + \varepsilon_{i,t} \quad \text{……..(Model 3)}$$

$$G_{score\,i,t} = \beta_0 + \beta_1(G_{peer})_{i,t-1} + \beta_2\, FirmSize_{i,t} + \beta_3 Debt\ Ratio_{i,t} + \beta_4 ROA_{i.t} + \beta_5(Price/Book)_{i,t} + \alpha_i + \varepsilon_{i,t} \quad \text{……..(Model 4)}$$

[15] For any $firm_i$ the weighted average ESG scores for the peer group will be calculated using the market cap of all firms (excluding$firm_i$) in the peer group. Likewise, the simple average is the sum of pillar scores of all firms (excluding $firm_i$) divided by the number of peers.

[16] Model 2 , 3 and 4 correspond to Hypotheses 2 , 3 and 4 respectively.

where, $E_{score_{i,t}}$, $S_{score_{i,t}}$, $G_{score_{i,t}}$, $(E\_peer)_{i,t-1}$, $(S\_peer)_{i,t-1}$, $(G\_peer)_{i,t-1}$ are the Environmental, Social and Governance scores and average[17] of lagged peer group pillar scores accordingly. The detailed description of variables is in Appendix A.

Consistent with standard econometric practice we use relevant firm fundamental variables in our models (Cheng, 2008; Miller & Triana, 2009). For firm size, we employ the logarithmic value of the total assets of the firm in order to control for the fluctuations of firm size in our panel. The debt ratio serves as a proxy for leverage while the ROA is a proxy for growth. It can be appreciated that our model (Model 1) essentially splits the ESG score of a firm into two main components; that is, a peer-induced component $\beta_1(ESG\_peer)_{i,t-1}$ and the remaining part which is the idiosyncratic component that is not influenced by the peer effects.

$(ESG\_peer)_{i,t-1}$ is the year one lag (lag1) of the independent variable. We employ a lag instead of a contemporaneous independent variable in our models. The reason is that a firm is more likely to benchmark itself against the previous year's ESG ratings of its peers and accordingly formulate its own ESG strategy. Moreover, ESG scores are published by rating agencies on an annual basis. Thus, the score published for a firm in the present year is actually the assessment of ESG performance of the previous year. Therefore, the previous year's ESG score of the peers is more likely to influence a firm's ESG performance than contemporaneous scores. Moreover, if we employ a contemporaneous variable, then there could be issues of reverse causality. We will not be able to distinguish if the ESG scores of the peers is causing a change in the ESG score of the firm or if it is vice versa. Therefore, we employ the lagged ESG and lagged pillar scores in our models, in order to reduce the likelihood of reverse causality. In order to address heteroscedasticity, we cluster robust standard errors according to firms, to account for the heteroscedasticity in the observations across firms in the panel.

4.1 Peer ESG performance scores

The variable $(ESG\_peer)_i$ will serve as a proxy for peer effects of ESG and for robustness check we will estimate all models using both simple average as well as weighted average using market capitalization (see Section 7.0). The simple average is the sum of ESG scores of all firms (excluding $firm_i$) divided by the number of peers in the group excluding $firm_i$.

[17] For any $firm_i$ the weighted average pillar scores for the peer group will be calculated using the market cap of all firms (excluding $firm_i$) in the peer group. Likewise, the simple average is the sum of pillar scores of all firms (excluding $firm_i$) divided by the number of peers.

We calculated the simple average of the ESG scores of the peer group (excluding $firm_j$) in order to arrive at the ESG performance of peer group for firm$_i$. $firm_i$.

$$\left(ESG_{peer}\right)_i \;=\; \frac{1}{n}\sum_{j=1}^{n} ESG_j$$

where,

$n =$ Total number of firms in the peer group excluding $firm_i$
$ESG_j =$ ESG score of $firm_j$
N = Total number of firms in the peer group
thus, $\mathrm{N} = n + 1$

$(\mathrm{ESG\_peer})_i$ denotes the simple average ESG score of the peers of $firm_i$ and j merely denotes the identification number of firms in the peer group excluding $firm_i$. Thus, we assume, for every$firm_i$, there are *n* number of peers excluding itself. Each peer $firm_j$ will have an ESG score denoted by ESG_j.

Similar process was adhered to for the pillar variables $(E\_peer)_i$, $(S\_peer)_i$, and $(G\_peer)_i$ which are the pillar scores of the peer group for every $firm_i$. We calculate the simple average pillar scores of the peer group (excluding $firm_j$) in order to arrive at each pillar score performance of peer group for $firm_i$.

4.2 Robustness

In the main panel we employ the Fama French 49 Industry classification scheme using the 4-digit SIC codes in order to construct peer groups. Hereafter, we also refer to our main panel as the 'Fama-French panel'. The tests, results and associated discussion of the main panel are in Section 6 and forms part of the main analysis. Furthermore, we apply robustness checks with weighted average peer scores instead of the simple averages on the Fama-French panel. In addition, we test the main panel by introducing lag2 of the independent variable in the models. After testing the Fama-French panel, we embark on constructing fresh peer groups using the Hoberg-Phillips textual-analysis based peer grouping method, and accordingly another panel dataset thereof. Thereafter, we test our models on this second panel (hereafter referred as the 'Hoberg-Phillips panel') as a robustness check. The testing, results and discussion for the Hoberg-Phillips panel are covered in the Robustness Section (Section 7).

5.0 Data: Collection and Construction

5.1 Dataset: Fama-French method

The preliminary step was to obtain ESG scores and pillar scores from Refinitiv in Datastream. The Refinitiv ESG data consists of ten category scores and an overall ESG score. These scores are awarded to a firm on a scale of 0 to 100. Refinitiv shared the updated methodology[18] of calculating the three pillar scores from the ten category scores. The Refinitiv ESG database identifies firms on the basis of ISIN[19]. This enabled us to identify them on Compustat and accordingly obtain their 4-digit SIC codes. These codes were used to segregate firms into industry peer groups using the Fama French 49 Industry Classification scheme. The final dataset is an unbalanced panel of 3,450 firms covering a period from 2009 to 2017 corresponding to 24,742 firm- year observations. Hereafter, we refer to this dataset as the main panel. The detailed method of construction of the dataset is in Appendix B.

5.2 Sample Selection

The Refinitiv ESG database cumulatively covers 6553 firms over a period from 2002 to 2018. However, we chose the period 2009 to 2017 for three reasons. Firstly, during the initial period of 2002 to 2008, the Refinitiv database had very few firms[20]. Furthermore, it was observed that Refinitiv had stopped rating a substantial number of them during this period. Secondly, from a macroeconomic perspective, in order to avoid the shock of the global financial crisis of 2007-08, the period was discarded. Thirdly, since ESG rating was a relatively new concept, it is assessed that the rating agencies were novices with an upward learning curve during the initial phase. We assume that the rating agencies have improved their metrics by removing biases, and, that the ESG scores now are a better reflection of ESG performance than in the initial period of 2002-08.

To avoid bias, we selected firms in the Refinitiv ESG universe irrespective of firm size or any financial/ performance metrics. The sole criterion of selection was that a firm should have at least 3 years of ESG scores since our estimated model considers two yearly lags. With this criterion, we retained 4858 firms. Thereafter we removed firms for which financial data or SIC codes or peer firms were not available. This eliminated 1254 firms. We further dropped 59 firms which had SIC code 9997 & 9998 because these did not correspond to the Fama-

[18] The detailed methodology for calculation of pillar scores from the ten category scores is explained online at: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf

[19] ISIN (International Securities Identification Number) is a unique 12 digit alphanumeric code allotted to internationally traded securities.

[20] Refinitiv database contains ESG data for 697 firms in year 2002 and 708 in 2003

French 49 Industry groups since they are listed as conglomerates. In all we were left with 3450 firms corresponding to 24742 firm-year observations.

5.3 Advantages of the Refinitiv database

Our study is perhaps one of the first academic endeavours to employ the enhanced ESG scores from Refinitiv as these scores have been recently released in 2018. The Refinitiv database has certain advantages over the ASSET4 legacy data as well as SRB data from other agencies.

The Thomson Reuters ASSET4 ESG legacy data has been found to have certain major drawbacks; and therefore necessitated upgrading by Refinitiv. Some of the key enhancements of the Refinitiv ESG data over the earlier ASSET4 legacy data are: (1) integration of controversial issues across all categories[21] and discounting overall ESG scores accordingly (2) the scores are adjusted for size and impact of each category unlike the ASSET4 data which was equal- weighted.

Furthermore, the KLD database employs a set of binary indicator variables, which are either positive (Strengths), or negative (Concerns) and are awarded scores of 0 or 1 accordingly. Whereas the Refinitiv ESG and pillar scores utilize a wide metrics scale from 0 to 100, and thus more accurately capture the subtle differences on each parameter. According to Galbreath (2013), the KLD ESG scores lack robust assessment with particular reference to corporate governance dimension[22].

6.0 Results and Analyses

The objective of this chapter is to analyse the sample data and thereafter discuss the findings. Initially, descriptive statistics of the dataset will be presented. The purpose of this section is to increase the understanding for the main analysis which will be presented in subsequent sub-sections. The analysis forms the basis of answering the research questions.

[21] The Refinitiv ESG data consists of ten category scores and an overall ESG score.

[22] Which, according to Galbreath (2013) , fail to capture certain aspects of board structure, committee independence, accountability, reporting and disclosure, and shareholder rights. Krüger (2015) has also expressed concern over the accuracy of the KLD corporate governance metrics.

6.1 Descriptive statistics

In this section, we scrutinize the characteristics of the Fama-French dataset. First, an overview of the ESG and pillar scores will be presented. This will be followed by a display of statistics to show the dispersion of ESG and pillar scores through the period of review.

6.1.1 Summary tables: Fama-French dataset

The dataset is an unbalanced panel consisting of 3450 unique GVKEYS (firms). Illustration 1 displays the year wise distribution of firms in the panel. In other words, this also reflects the number of firms which reported ESG scores in the Refinitiv universe and satisfied our criterion of having at least 3 year ESG data.

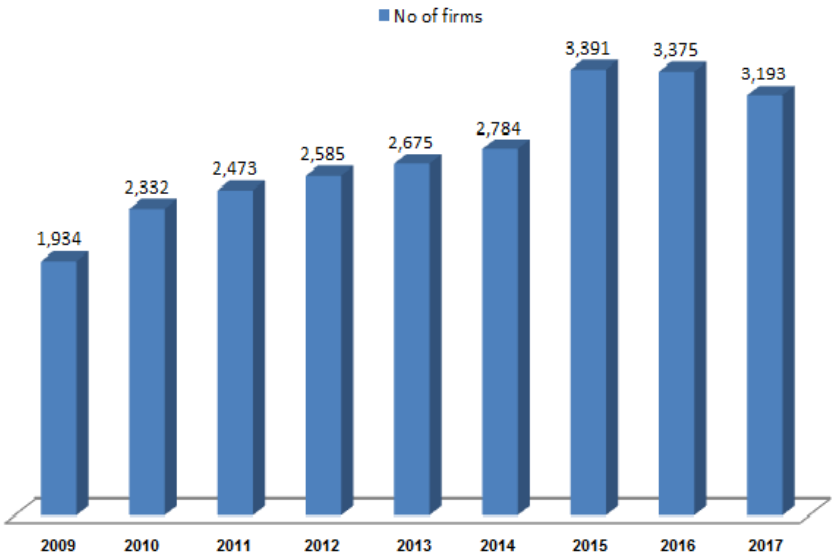


Illustration1: The bar chart depicts the year wise distribution of number of firms in the unbalanced panel comprising of a total of 3450 unique firms.

Table 1 reports the summary statistics of the control variables. We observe that the range and distribution of the firm size (total assets) is wide. On deduplication of the Industry group column we find that all 49 Industry groups are present in the dataset. These characteristics suggest that the sample is a fair representation of the market.

Table 1. Summary Statistics: Control variables

	N	mean	min	max	sd	p25	p50	p75
TOTALASSETS	24742	1420bn	0.002bn	437000bn	11900bn	3.072bn	13.3bn	113bn
MARKETCAP	24742	499bn	0.05bn	22400bn	2260bn	2.783bn	10.4bn	81.5bn
Debt Ratio	24742	0.244	0.000	1.022	0.174	0.104	0.234	0.357
P/B Ratio	24742	17.990	0.000	154916.800	1343.776	1.160	1.890	3.250
ROA	24742	0.046	-10.027	88.223	0.584	0.010	0.040	0.078

This table reports the descriptive statistics of the firm specific control variables. The sample consists of 3450 firms over the period 2009-2017 resulting in 24,742 firm-year observations. TOTALASSETS and MARKETCAP have been obtained from Datastream (Thomson Reuters) and are in billion USD ($). MARKETCAP has been winsorized at the 1st and 99th %percentiles. P/B Ratio and Debt Ratio have been obtained primarily from Datastream; missing values were made up from Compustat. For calculating ROA, the Net Incomes have been obtained from Compustat. Variable definitions are provided in Appendix A.

Table 2 depicts the summary statistics of ESG and pillar scores of firms and peers. We observe that the dispersion of ESG and pillar scores are high, with Environment pillar scores having the highest range and standard deviation.

Table 2. Summary Statistics: ESG scores

	N	mean	min	max	sd	p25	p50	p75
ESG	24742	51.83903	7.30925	97.48635	17.88437	37.81165	52.1239	65.6623
Environment	24742	52.22281	2.858824	99.24118	23.37948	31.77382	52.62559	71.57059
Social	24742	52.13981	2.501972	99.09	22.00879	35.19282	52.56275	69.48901
Governance	24742	51.06111	1.02	99.35541	21.03702	34.40984	51.3868	67.93459
simple_ESG	24742	51.83892	36.86041	76.02465	2.883999	50.37921	51.64314	53.33314
simple_Environment	24742	52.22263	29.53147	78.12353	3.692641	49.90434	51.92446	54.04411
simple_Social	24742	52.13959	33.68682	89.70676	3.624628	50.26026	51.89513	53.97398
simple_Governance	24742	51.06123	22.48074	79.62485	3.078708	49.375	51.1216	52.58696

This table reports the descriptive statistics of the independent variables, ie, ESG and pillar scores of firms and peers. Variable definitions are provided in Appendix A.

We examine the ESG and pillar score standard deviation within each Industry group (see Appendix C). We observe that the standard deviation within industry groups ranges from 7.84 (Guns) to 22.05 (Soda) and that the rest of the industry groups have an ESG score standard deviation between 13.05 and 19.93. The between and within standard deviations for the overall ESG score and the pillar scores of firms is in Appendix I.

Table 3 depicts the development of the average ESG and pillar scores of firms in the sample. We observe that both the average overall ESG score as well as the pillar scores for firms does not change much from year to year during the period between 2009 to 2015. From 2015 onwards the average ESG score as well as the average pillar scores develops with at a larger growth rate than the previous years. The reason for the sudden change in growth rate could be

due to the large increase in firms in the sample in 2015. This may have increased the competition in the peer-group.

Table 3. Development of average ESG and pillar scores

		2009	2010	2011	2012	2013	2014	2015	2016	2017
ESG	Average score	51,26	51,02	51,02	50,96	51,01	51,25	50,98	53,41	54,58
	Average growth		-0,005	0,000	-0,001	0,001	0,005	-0,005	0,048	0,022
Environment	Average score	51,79	51,48	51,44	51,50	51,26	51,45	51,00	54,12	55,21
	Average growth		-0,006	-0,001	0,001	-0,005	0,004	-0,009	0,061	0,020
Social	Average score	51,69	51,31	51,21	51,09	51,23	51,55	51,37	53,58	55,17
	Average growth		-0,007	-0,002	-0,002	0,003	0,006	-0,004	0,043	0,030
Governance	Average score	50,18	50,18	50,34	50,50	50,46	50,69	50,53	52,43	53,20
	Average growth		0,000	0,003	0,003	-0,001	0,005	-0,003	0,038	0,015

Illustration 2 shows how the average ESG and pillar scores in the sample develop from 2009 to 2017.

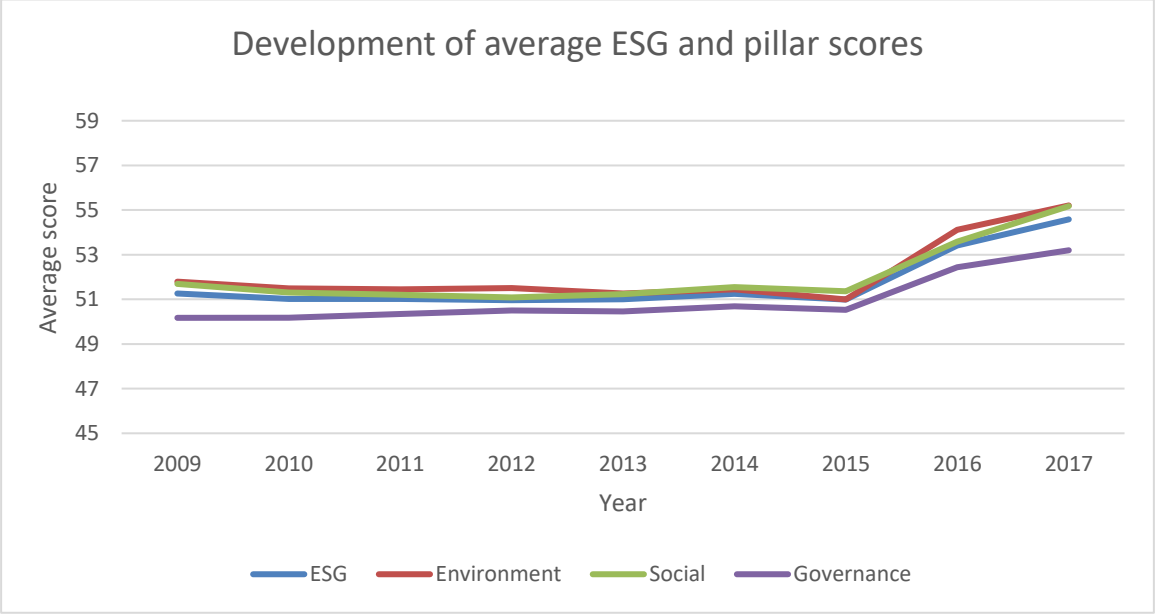


Illustration2

In Table 4, we compare average ESG score for three, five and nine years respectively. Accordingly, the corresponding average growth rate of the respective year, as well as the average units change, have been calculated.

Table 4. Development of average ESG and pillar scores

	Period	Average score	Average growth rate	Average units change per year
	Last 3 years	52,99	0,021	1,136
	Last 5 years	52,25	0,014	0,731
ESG	Last 9 years	51,72	0,008	0,415
	Last 3 years	53,44	0,024	1,292
	Last 5 years	52,61	0,014	0,752
Environment	Last 9 years	52,14	0,008	0,430
	Last 3 years	53,37	0,023	1,233
	Last 5 years	52,58	0,016	0,824
Social	Last 9 years	52,02	0,008	0,433
	Last 3 years	52,05	0,016	0,853
	Last 5 years	51,46	0,011	0,544
Governance	Last 9 years	50,94	0,007	0,377

We further investigate the distribution of firm size, by way of a scatter plot of total assets as well as market capitalization. We observe the dispersion of outlier data points, which correspond to very large firms (see scatter plot in Appendix D)

It is observed that about 90% of the data points lie below the 1 trillion dollar mark in terms of total assets. The top 10%of the data points lie between 1 trillion and 437 trillion dollars. We examine the data in terms of market capitalization (winsorized). It is observed that a considerable number of outliers are above the $10 billion dollar mark.

6.2 Main Analysis

In Table 5, Hypothesis 1 is tested, where we investigate if there is a peer group effect related to the overall ESG score. In Table 6 Hypotheses 2, 3 and 4 are tested, wherein we investigate further if there are peer group effects for the ESG pillar dimensions and which of them significantly affect a firm's pillar score.

6.2.1 Hypothesis 1: Peer-effect on overall ESG score

Table 5 reports the regression results for estimation of Model 1 regarding peer-effect experienced by a firm on its ESG score as a result of an increase in the average ESG score of its peer-group.

Table 5: Regression Model 1

VARIABLES	(1) ESG
ESGpeerSiL1	0.506***
	(0.0467)
Debtratio	0.756
	(1.317)
P/B ratio	2.21e-05***
	(2.30e-06)
ROA	-2.568***
	(0.857)
TA	5.953***
	(0.350)
Constant	-74.07***
	(6.043)
Observations	21,291
Number of firms	3,450
R-squared	0.077
Firm FE	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The results suggests a positive effect on firms current ESG scores as a result of an increase in the previous year's average ESG score of its peer group. We observe this relationship to be significant at the 1% level, indicating that the current ESG score of a firm increases by 0.506 when the previous year's average ESG score of its peer-group increases by 1 unit. This is in conformity with Hypothesis 1, suggesting that firms care about how their peers' ESG score develops on a year to year basis and due to peer-effect tries to improve its own. The explanatory power indicates that the model explains 7.7% of the total variation in observations.

In Section 6.1 we observe from the descriptive statistics that the average overall ESG score does not change much from year to year, suggesting that the average growth rate for all nine years is equal to 0.8%. Based on this, we find that the expected average unit change is suggested to be 0.414 instead of 1 unit, which implies the peer-effect estimated in the model to equal 0.209. However, when looking at the development over the last three years, we observe that the growth rate increases considerably. Assuming that the recent year's growth rate projects the future development in ESG score more accurately than the earlier years, it is reasonable to expect an average unit change close to 1 for the next years to be likely. Using the average of the last five years as the benchmark, we find that the expected average unit

change per year would equal 0.731 unit. Consequently, the realistic peer-effect would equal 0.369 unit.

6.2.2 Sub-conclusion: Hypothesis 1

The results from this model is in accordance with Hypothesis 1, suggesting that firms are sensitive to their peers' ESG scores in the previous year and this effect is significant at the 1% level.

6.2.3 Hypotheses 2, 3 and 4: Peer-effects of the pillar scores

In this section we estimate Models 2, 3 and 4 and present the regression results. This will be followed by a discussion and sub-conclusion.

Table 6 compares the peer-effect experienced by a firm with respect to its pillar scores.

Table 6: Regression Model 2, 3 & 4

VARIABLES	(1) Environment	(2) Social	(3) Governance
EnvironmentpeerSiL1	0.369***		
	(0.0461)		
SocialpeerSiL1		0.449***	
		(0.0541)	
GovernancepeerSiL1			0.150***
			(0.0570)
Debtratio	-0.247	-2.169	4.984**
	(1.776)	(1.700)	(2.013)
P/B ratio	-6.22e-05***	2.70e-05***	0.000122***
	(3.21e-06)	(2.34e-06)	(7.30e-06)
ROA	-3.868***	-3.126***	-0.659
	(0.998)	(1.169)	(0.683)
TA	7.979***	7.197***	2.721***
	(0.479)	(0.476)	(0.485)
Constant	-100.5***	-91.22***	-3.204
	(8.119)	(8.147)	(8.434)
Observations	21,291	21,291	21,291
R-squared	0.071	0.059	0.008
Number of firms	3,450	3,450	3,450
Firm FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In column (1) we find that the current Environment score of a firm increases by 0.369 when last year's average environment score of the peer group increases by 1 unit. This finding indicates that there is a positive relationship between the current eEvironment score of a firm and the score of its peers one year ago. The estimated change of 0.369 in the current Environment score is significant at the 1% level and the explanatory power indicates that the model explains 7.1% of the total variation in the observations.

Further on, in column (2) we find that the current Social score of a firm increases by 0.449 when last year's average Social score of the peer group increases by 1 unit. This finding indicates that there is a positive relationship between the current Social score of a firm and the score of its peers one year ago. The estimated change of 0.449 in the current Social score is significant at the 1% level and the explanatory power indicates that the model explains 5.9% of the total variation in the observations.

Likewise, we find that the results in column (3) also are in line with the relationship found for the Environment score as well as the Social score. The results suggests that the current Governance score of a firm increases by 0.150 when last year's average Governance score of the peer group increases by 1 unit. This finding indicates that there is a positive relationship between the current Governance score of a firm and the score of its peers one year ago. The estimated change of 0.150 in the current Governance score is also significant at the 1% level. The explanatory power indicates that the model explains 0.8% of the total variation in the observations.

From Section 6.1.1 it is observed that the historical development of the average pillar scores are approximately the same as suggested for the average overall ESG score, the growth rate being the largest for the Environment and Social score. Assuming that the average score development for the last five years is a good prediction for the score development over the years to come, it is indicated that a 0.751 average unit change is reasonable for the Environment score, 0.823 for the Social score and 0.543 for the Governance score. Hence, using these numbers as basis for the estimated effects in the three models, it is suggested that an estimated effect equal to 0.277 is likely for the Environment score, 0.369 for the Social score and 0.814 for the Governance score.

6.2.5 Sub-conclusion: hypothesis 2, 3 and 4

The results are in line with Hypotheses 2, 3 and 4 suggesting a positive effect on a firm's current Environment score, Social score and Governance score because of an increase in the previous year's respective pillar scores of its peer group. We observe this relationship to be significant, indicating that firms care about how its peers pillar scores develops from year to year and due to peer-effect tries to improve its own. We find that the Social score represents the strongest peer-effect of the three pillar scores, slightly stronger than the Environment score. The Governance pillar score is indicated to experience a significantly lower peer-effect than the two other pillar scores.

7.0 Robustness

In this section we present various robustness tests to validate our models. First, we repeat the regressions of the models as performed in the main analysis, but instead of using simple averages of ESG scores of the peer group, we employ weighted-averages. Second, we run the original models but we introduce lag2 along with lag1. Likewise the same procedure is applied for the pillar score Models 2, 3 and 4. Next, we generate time dummies and estimate the model using time fixed effects. Afterward, we conduct the regressions as performed in the main analysis, but on another sample using the Hoberg-Phillips peer-grouping method instead of the Fama-French method. Lastly, we run two tests as supplemental analysis.

7.1 Simple average and weighted average

In this sub-section we examine whether there is a difference from the results found in the main analysis using simple average compared to using weighted average. Firstly, the methodology will be explained, thereafter we compare the results obtained for the overall ESG score as well as the pillar scores.

7.1.1 Weighted average method

As per the model of Leary and Robert (2014) we define the explanatory variable $(ESG\_peer)_i$ as the weighted average of ESG scores of the peers using market capitalization. Therefore, we will calculate the weighted average of the ESG scores of the peer group (excluding $firm_j$) in order to arrive at the ESG performance of peer group for $firm_i$. Similar process was adhered to for the variables $(E\_peer)_i$, $(S\_peer)_i$, and $(G\_peer)_i$ which are the

pillar scores of the peer group for every $firm_i$ we calculate the weighted average pillar scores of the peer group excluding the $firm_i$.

$$\left(ESG_{peer}\right)_i = \frac{1}{w}\sum_{j=1}^{n} w_j ESG_j$$

where,

$$w = \sum_{j=1}^{n} w_j$$

where $(\text{ESG\_peer})_i$ denotes the weighted average ESG score of the peers of $firm_i$ and j merely denotes the identification number of firms in the peer group excluding $firm_i$. Thus, we assume, for every $firm_i$, there are *n* number of peers excluding itself. Each peer $firm_j$ will have a market cap of w_j and an ESG score denoted by ESG_j. Thus, *w* is the sum of individual market cap of each peer $firm_j$.

7.1.2 Hypothesis 1: Peer effect on overall ESG score

In Table 7 we compare the results from using simple average and weighted average for the overall ESG score. Column (1) corresponds to the simple average method whilst column (2) corresponds to the weighted average method.

Table 7: Regression Model 1: Simple vs Weighted

VARIABLES	(1) ESG	(2) ESG
ESGpeerSiL1	0.506***	
	(0.0467)	
ESGpeerWtL1		0.0694***
		(0.0145)
Debtratio	0.756	0.506
	(1.317)	(1.335)
P/B ratio	2.21e-05***	2.58e-05***
	(2.30e-06)	(2.16e-06)
ROA	-2.568***	-2.682***
	(0.857)	(0.906)
TA	5.953***	6.272***
	(0.350)	(0.356)
Constant	-74.07***	-57.41***
	(6.043)	(5.943)
Observations	21,291	21,291
R-squared	0.077	0.067
Number of firms	3,450	3,450
Firm FE	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The results from column (2) are in conformity with Hypothesis 1 suggesting a positive effect on a firm's current ESG scores as a result of an increase in the previous year's ESG score of its peer group. We observe this relationship to be significant at the 1% level, indicating that firms care about how their peers' ESG score develops from year to year and due to peer-effect tries to improve its own scores. Comparing the results between the model using a simple average and the one using weighted average, we find that the relationship is directionally the same, but that the effect is suggested to be approximately 0.20 stronger in the model using a weighted average. The explanatory power indicates to be almost equal, where the model using simple average is 1% stronger than the one using weighted average.

7.1.3 Sub-conclusion: Hypothesis 1

From the comparison above, we find that the model used in the main analysis is robust to the model using weighted average. They are both indicating the directionally same relationship, but the weighted average is estimated to have a stronger effect. We will now do the same comparison for the pillar models related to hypothesis 2, 3 and 4.

7.1.4 Hypothesis 2, 3 & 4: Peer effects of the pillar scores

Table 8 compares the regressions of the two methods, i.e., simple average and weighted average for the individual pillar scores. In column (1) and (2) the table displays the results of the Environment pillar using simple and weighted averages respectively. Likewise, column (3) and (4) show the results for the Social pillar and columns (5) and (6) the results for the Governance pillar.

Table 8: Regression : Model 2, 3 & 4 : Simple vs Weighted

VARIABLES	(1-Si) Environment	(2-Wt) Environment	(3-Si) Social	(4-Wt) Social	(5-Si) Governance	(6-Wt) Governance
EnvironmentpeerSiL1	0.369***					
	(0.0461)					
EnvironmentpeerWtL1		0.125***				
		(0.0160)				
SocialpeerSiL1			0.449***			
			(0.0541)			
SocialpeerWtL1				0.0597***		
				(0.0162)		
GovernancepeerSiL1					0.150***	
					(0.0570)	
GovernancepeerWtL1						-0.0279*
						(0.0159)
Debtratio	-0.247	-0.493	-2.169	-2.365	4.984**	4.900**
	(1.776)	(1.773)	(1.700)	(1.713)	(2.013)	(2.007)
P/B ratio	-6.22e-05***	-6.44e-05***	2.70e-05***	2.73e-05***	0.000122***	0.000129***
	(3.21e-06)	(3.30e-06)	(2.34e-06)	(2.46e-06)	(7.30e-06)	(7.09e-06)
ROA	-3.868***	-3.786***	-3.126***	-3.248***	-0.659	-0.708
	(0.998)	(1.023)	(1.169)	(1.238)	(0.683)	(0.683)
TA	7.979***	8.087***	7.197***	7.409***	2.721***	2.854***
	(0.479)	(0.479)	(0.476)	(0.483)	(0.485)	(0.484)
Constant	-100.5***	-90.64***	-91.22***	-75.07***	-3.204	3.727
	(8.119)	(8.016)	(8.147)	(8.005)	(8.434)	(8.120)
Observations	21,291	21,291	21,291	21,291	21,291	21,291
R-squared	0.071	0.070	0.059	0.053	0.008	0.007
Number of firms	3,450	3,450	3,450	3,450	3,450	3,450
Firm FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From the results, we observe that both the Environment score as well as the Social score are suggesting the same relationship as for the model using simple average. The estimated effect is significant at the 1% level for both scores. However, we see that the estimated effect is larger for simple average compared to weighted average for the Environment score, but the opposite being the case for the social score where the weighted average model suggests the strongest effect. For the Environment score we observe that the estimated effect is approximately three times stronger for the simple average model. The difference in the estimated effect between the two models observed for the Social score is nonetheless smaller and equal to a difference of approximately 0.10. When it comes to the Governance score, the

results from the weighted average suggests a reverse effect and the opposite relationship compared to the model using simple average. This is not in line with hypothesis 4 and questions the robustness of the results obtained when using the simple average model. At the same time the effect estimated for the weighted average model, equal to -0.0279, is only significant at the 10 % level.

7.1.4 Sub-conclusion Hypothesis 2, 3 & 4

The results for the Environment and the Social score when using the weighted average are directionally consistent with the relationship found when using simple average. However, the estimated effect for the Environment scores is larger for the simple average, whilst for the social score the weighted average has the largest effect for the pillar. The results for these two pillars are in line with the models using simple average and hypothesis 2 and 3, we find that the results for the Governance score using weighted average are suggesting a reverse effect. Based on this finding we should question the robustness of the results from our initial model using simple average, although the estimated effect for the Governance pillar is only significant at the 10% level. The implications of the difference between the models will be addressed in section 7.1.4.

7.1.5 Discussion and Conclusion: Simple average and weighted average

The results of the models employing weighted average are not consistent with the results using simple average. This raises an element of doubt regarding the robustness of the results using simple averages in the models.

The reason for the difference in peer-effect between the two methods could be that firms, above the third quartile in terms of market capitalization, dominate the results in the weighted model and consequently distort the actual peer-effect experienced by the firms. In Section 6.1.1, we mentioned presence of large size outliers in terms of market capitalization in the sample. It appears that these very large firms may be causing the skewness in weighted averages. By conducting regressions using the simple average, we overcome this problem by estimating the peer-effect on firms regardless of their market capitalization. On the other hand it can be argued that the weighted average models are more accurate to use, as they display the peer-effect experienced by the biggest players in the market who represent the major share of the market. This implies that these large firms have more influence by their size alone. However, the scope of our research, necessitates that we adopt such a model which can

provide equal representation for all firms irrespective of size. This approach is intuitively reasonable as we are primarily focussed on the study of ESG behaviour per se and not firm performance whereas market cap is a direct function of firm performance. Therefore, we conclude that simple average method is more aligned with the focus of our research. Based on this, we argue that the simple average model is the preferred method giving more accurate estimates. Hence, further analysis will focus on models using simple average.

7.2 Inclusion of lag2

In this section we include lag2 into the original model. We run the regressions using the Fama-French sample. The intuition behind including lag2 is that the ESG scores of the peers two years ago potentially could influence a firm's current ESG score. The rationale is to check that our model is correctly specified and rule out omitted variable bias. All four hypotheses are tested by including lag2 in the models. We will only discuss the results in this section however the tables will be reported in Appendix E

7.2.1 Results

In Table 9 the results from the modified model including lag1 and lag2 is displayed in column (1). Column (2) displays the results from the original model

Table 9: Regressions: Model 1: with and without lag2

VARIABLES	(1) ESG	(2) ESG
ESGpeerSiL1	0.670***	0.506***
	(0.0533)	(0.0467)
ESGpeerSiL2	-0.185***	
	(0.0488)	
Debtratio	1.947	0.756
	(1.448)	(1.317)
P/B ratio	4.69e-07	2.21e-05***
	(2.79e-06)	(2.30e-06)
ROA	-1.294**	-2.568***
	(0.655)	(0.857)
TA	5.533***	5.953***
	(0.385)	(0.350)
Constant	-65.91***	-74.07***
	(7.196)	(6.043)
Observations	17,841	21,291
R-squared	0.075	0.077
Number of firms	3,444	3,450
Firm FE	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We observe that in all four regressions, lag2 has negative coefficients while lag1 is consistent in sign and significance as seen in the main models so far. The reverse effects of lag2 are significant at the 1% level. This is not in conformity with hypothesis 1, 2, 3 and 4 or economic intuition. Assuming that a firm experiences a peer-effect from its peers' average ESG score one year ago, but at the same time the average ESG score of its peers two years ago compels it to decrease its own ESG score, is counter-intuitive and does not make business sense. Hence, it is necessary to investigate the reverse effect for lag2.

7.2.2 Reverse effect lag2

There can be several reasons why lag2 of the average ESG and pillar scores of peers suggests a reverse effect. One of the reasons could be that there is a high correlation between lag1 and lag2 potentially leading to multicollinearity. In such a case, the precision of the estimate coefficients reduces and weakens the statistical power of the regression model. To test this, we examine if there is correlation between lag1 and lag2.

7.2.2.1 Multicollinearity: lag1 and lag2

The correlation matrix below shows the correlation between the variables for lag1 and lag2 of the average ESG score of the peer-group. We observe high correlation between lag1 and lag2 at 0.827. This finding suggests that there might be a issue of multicollinearity in the model.

	ESGpeerSiL1	ESGpeerSiL2
ESGpeerSiL1	1	
ESGpeerSiL2	0.827	1

Investigating further for the lag1 and lag2 of the average peer pillar scores (see Appendix F for correlation matrix), the results demonstrates that lag1 and lag2 are highly correlated for all three pillars. The correlation coefficient ranges from a minimum of 0.76 represented by the Governance pillar and a maximum of 0.86 represented by the Social pillar. This finding implies that we should consider excluding lag1 or lag2 from the model in order to obtain accurate estimates.

We further investigate lag1 and lag2 for multicollinearity by regressing lag1 on lag2. If we detect the estimated effect of lag2 to be positive and significant when holding lag1 as the

dependent variable, it is suggested that lag2 predict lag1 and that multicollinearity can be present between the two variables. In Table 10 the results from the regression is displayed in column (1).

Table 10: Regression: lag1 on lag2

	(1)
VARIABLES	ESGpeerSiL1
ESGpeerSiL2	0.303***
	(0.0179)
Debtratio	-0.393
	(0.260)
P/B ratio	7.54e-06***
	(6.27e-07)
ROA	-0.285*
	(0.152)
TA	0.777***
	(0.0709)
Constant	22.89***
	(1.495)
Observations	17,841
R-squared	0.084
Number of firms	3,444
Firm FE	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From the results we observe that in column (1) lag2 is positive and significant at the 1 % level. This finding indicates that there are some interplay between lag1 and lag2. Consequently, the suspicion towards multicollinearity strengthens and one of the lags should be excluded from the model used to test our hypotheses. In order to decide which of the lags to be included and accepted as the correct model, we run two regressions; ESG on lag1 and ESG on lag2. In Table 11 reports the results from regressing ESG on lag1 is displayed in column (1) and in column (2) the results from regressing ESG on lag2.

Table 11: Regressions: Model 1 : with lag1 : with lag2

VARIABLES	(1) ESG	(2) ESG
ESGpeerSiL1	0.506***	
	(0.0467)	
ESGpeerSiL2		0.0182
		(0.0517)
Debtratio	0.756	1.684
	(1.317)	(1.478)
P/B ratio	2.21e-05***	5.52e-06**
	(2.30e-06)	(2.64e-06)
ROA	-2.568***	-1.486**
	(0.857)	(0.703)
TA	5.953***	6.054***
	(0.350)	(0.397)
Constant	-74.07***	-50.57***
	(6.043)	(7.143)
Observations	21,291	17,841
R-squared	0.077	0.055
Number of firms	3,450	3,444
Firm FE	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From column (1) it is observed that lag1 is positive and significant at the 1 % level. In column (2) lag 2 is also suggested to have a positive effect, but is insignificant. Based on this finding it is reasonable to keep lag1 as the main model and discard lag2. From a business perspective it is also more reasonable to include lag1 if choosing between the two, as only including lag2 implies that firms care more about the ESG and pillar scores of their peers two years ago than the most recent scores available for one year ago. Hence, including lag1 and discarding lag2 is reasonable.[23]

Furthermore, just to check whether the total estimated effect is significantly different between our original model including only lag1 and the modified model including both lags, we calculate the difference in the total estimated effect by running our models including lag1 and

[23] Preferably we would like to run a VIF-test to check for multicollinearity, but VIF-test cannot be conducted on panel data (xtreg). By adding dummies for GVKEY, OLS should give identical estimates to xtreg and allow us to run a VIF-test. Due to limited matsize (upper limit is 11000) in Stata this operation was not possible as we would get more variables than the upper limit.

lag2. Table 12 compares the regression results for the overall ESG score from the modified model in column (1) with the original model in column (2).

Table 12: Regressions: Model 1: with and without lag2

VARIABLES	(1) ESG	(2) ESG
ESGpeerSiL1	0.670***	0.506***
	(0.0533)	(0.0467)
ESGpeerSiL2	-0.185***	
	(0.0488)	
Debtratio	1.947	0.756
	(1.448)	(1.317)
P/B ratio	4.69e-07	2.21e-05***
	(2.79e-06)	(2.30e-06)
ROA	-1.294**	-2.568***
	(0.655)	(0.857)
TA	5.533***	5.953***
	(0.385)	(0.350)
Constant	-65.91***	-74.07***
	(7.196)	(6.043)
Observations	17,841	21,291
R-squared	0.075	0.077
Number of firms	3,444	3,450
Firm FE	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Comparing the total estimated effect for the two models, we find that the difference only equals 0,021. This difference is found by adding the estimated effect of lag1 and lag2 in column (1) and subtracting the estimated effect for lag one in column (2). The same relationship is also true for the individual pillar scores. Finding the difference in the total estimated effect to be small, we are more confident in discarding lag2.

7.2.3 Sub-conclusion: Inclusion of lag2

From the results including lag2 in the original model for all hypotheses, we found lag2 to be significant and suggesting a negative effect. Investigating the reverse effect of lag2, we detected high correlation between lag1 and lag2 in the modified model. This finding indicates the presence of multicillinearity. Furthermore, we observe a positive and significant relationship when regressing lag1 on lag2 strengthening the suspicion of multicollinearity. Based on this we found it appropriate to include only one of the lags and tested which of the

lags to include, by regressing ESG on lag1 and ESG on lag2. The results suggests only lag1 to be significant and in combination with economic intuition, we decided to keep our original model as correct and discard the modified model with both lags. Lastly, we compare the difference in the estimated total effect between the original and modified model; finding that the difference is just 0,021. We observe similar small differences in magnitude in the pillar models as well.

7.3 Time fixed effects

Table 13 reports the regression results when we include yearly time dummies in our model to control for time fixed effects.

Table 13: Regression: Model 1: with time dummies

VARIABLES	(1) ESG
ESGpeerSiL1	-0.0488
	(0.0537)
2011.YEAR	0.143
	(0.171)
2012.YEAR	0.352*
	(0.194)
2013.YEAR	0.626***
	(0.205)
2014.YEAR	1.264***
	(0.230)
2015.YEAR	3.828***
	(0.246)
2016.YEAR	5.751***
	(0.265)
2017.YEAR	7.002***
	(0.306)
Debtratio	-0.0457
	(1.175)
P/B ratio	4.36e-05***
	(3.23e-06)
ROA	-0.619
	(0.716)
TA	2.760***
	(0.315)
Constant	5.894
	(5.938)
Observations	21,291
Number of firms	3,450
R-squared	0.187
Firm FE	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We observe that by including time fixed effects, the lag1 independent variable is insignificant, while most of the time dummies are significant at the 1% level.

However, our contention is that we have an unbalanced panel where the number of firms is more in the latter years as compared to the former years. Therefore, if we include time dummies, the model may incorrectly capture differences in yearly cross section size and any underlying time trend. Hence, we discard time dummies in our model.

7.4 Hoberg- Phillips

In this section, we estimate Models 1, 2, 3 and 4 using the Hoberg-Phillips panel. First we examine the characteristics of the Hoberg-Phillips dataset. Thereafter, we furnish an overview of relevant statistics and then show the dispersion of ESG and pillar scores through the period of review. This will be followed by a discussion of the results. The same methodology, as explained in Section 4, will also be applied for the Hoberg-Phillips panel regressions.

7.4.1 Dataset: Hoberg – Phillips method

The firms were identified on Compustat based on GVKEYS[24] as the Hoberg-Phillips peer groupings are based on GVKEYS. The detailed description of the textual analysis based industry classification is in Appendix G. The peer groupings were incorporated while computing the simple average of the ESG scores of peers. The dataset is an unbalanced panel of 1299 firms over a period from 2009 to 2017 consisting of 7960 firm-year observations. The detailed method for sample selection and construction of the dataset is in Appendix G.

7.4.2.Summary statistics: Hoberg Phillips dataset

The dataset is an unbalanced panel consisting of 1299 unique GVKEYS (or firms) corresponding to 7960 firm-year observations. This panel is constructed from the Refinitiv ESG database of 6553 firms and is a sub-set of the Fama French panel dataset. Illustration 3 displays the year wise distribution of firms in the panel.

[24] The GVKEY or Global Company Key is a unique 6 digit identification number allotted by Compustat to a firm.

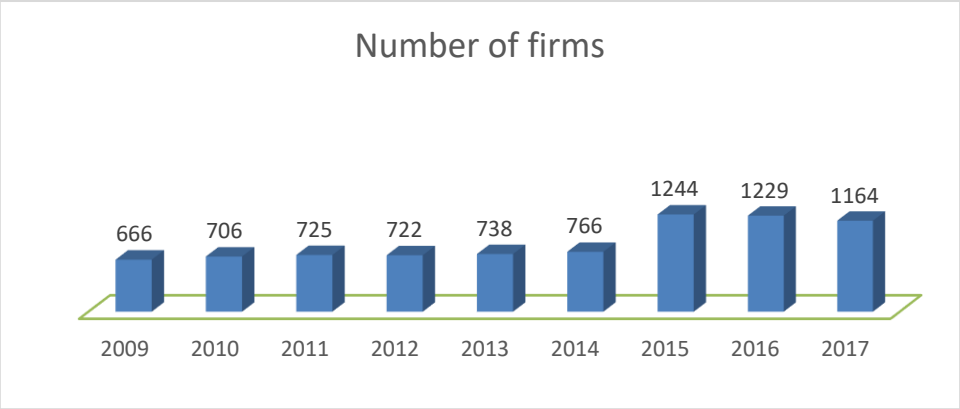


Illustration3: The bar chart depicts the year wise distribution of number of firms in the Hoberg- Phillips unbalanced panel comprising of a total of 1299 unique firms.

Table 14 reports the summary statistics of the firm-specific control variables.

Table 14. Summary Statistics

	N	mean	min	max	sd	p25	p50	p75
TOTALASSETS	7960	31.1bn	0.046bn	2570bn	133bn	2.657bn	6.48bn	17.8bn
MARKETCAP	7960	17bn	0.0031bn	868bn	42.4bn	2.681bn	5.401bn	13.7bn
Debt ratio	7960	.2650095	0	1.022116	.1866501	.1135877	.2526824	.3867722
P/B ratio	7960	4.563599	0	1410.85	23.06597	1.45	2.33	3.87
ROA	7960	.0429477	-1.759041	.6812982	.1026354	.0115643	.0409658	.0808686

This table reports the descriptive statistics of the firm specific control variables. The sample consists of 1299 firms over the period 2009-2017 resulting in 7960 firm-year observations. TOTALASSETS and MARKETCAP have been obtained from Datastream (Thomson Reuters) and are in billion USD ($). MARKETCAP has been winsorized at the 1st and 99th %iles. P/B Ratio and Debt Ratio have been obtained primarily from Datastream; missing values were made up from Compustat. For calculating ROA, the Net Incomes have been obtained from Compustat. Variable definitions are provided in Appendix A.

Table 15 depicts the summary statistics of ESG and pillar scores of firms and peers.

Table 15. Summary Statistics

	N	mean	min	max	sd	p25	p50	p75
ESG	7960	50.17334	8.0165	97.48635	16.84964	36.78243	48.4339	63.02027
Environment	7960	47.63149	2.875294	98.82765	23.10279	27.53662	43.98912	66.2475
Social	7960	51.47535	4.746761	98.57056	19.46895	36.21585	49.5157	65.83085
Governance	7960	51.49143	3.431475	98.34656	20.75076	35.64025	52.06361	67.96484
simple_ESG	7960	50.94946	19.97885	91.25285	5.462775	47.7035	50.52766	53.46345
simple_Environment	7960	48.51435	6.642059	95.38471	7.531576	43.98542	48.23718	52.14
simple_Social	7960	52.3441	20.70056	97.45986	6.434964	48.33607	51.52234	55.42675
simple_Governance	7960	52.04073	10.18197	93.05672	6.729022	48.44219	51.92586	55.68192

This table reports the descriptive statistics of the independent variables, ie, ESG and pillar scores of firms and peers. Variable definitions are provided in Appendix A.

We further investigate the distribution of firm size, by way of a scatter plot of total assets as well as Market Capitalization (scatterplot in Appendix H). We observe the dispersion of outlier data points which correspond to very large firms. About 90% of the data points lie below the 500 billion dollar mark in terms of total assets. We examine the data in terms of market capitalization (winsorized). It is observed that a considerable number of outliers are above the $100 million dollar mark.

Next step would be to examine the development of ESG and pillar scores over time. Illustration 4 shows how the average ESG and pillar scores in the sample develop from 2009 to 2017. We observe a different pattern as compared to the Fama-French panel. We see a sudden dip in the mean scores in 2014-15,

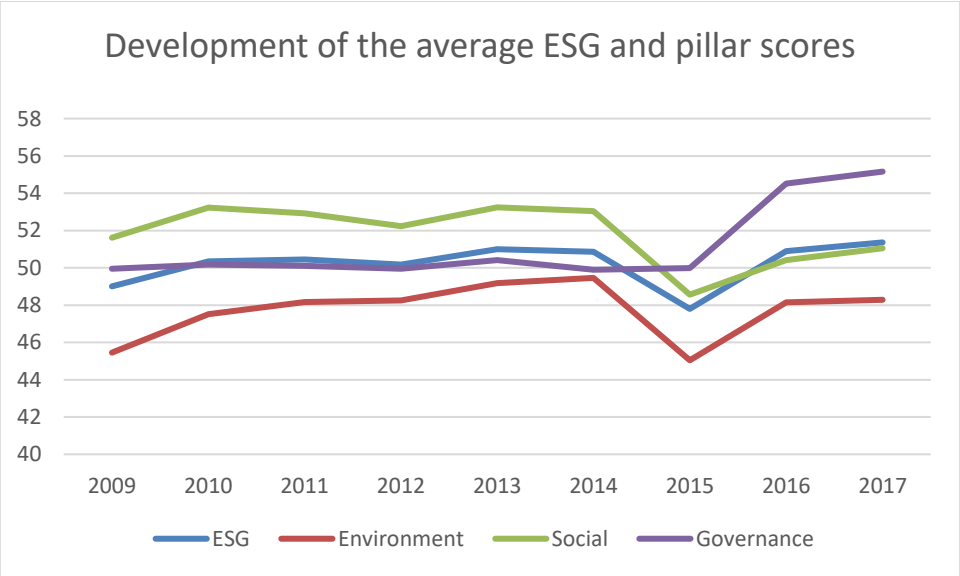


Illustration 4

7.4.3 Hypothesis 1

In Table 16 we compare the results for the overall ESG score between the Fama-French panel and the Hoberg-Phillips panel. Column (1) corresponds to the Fama-French panel whilst column (2) corresponds to the Hoberg-Phillips panel.

Table 16: Regressions: Model 1: FF vs HP

VARIABLES	(1) ESG (FF)	(2) ESG (HP)
ESGpeerSiL1	0.506***	-0.0867***
	(0.0467)	(0.0308)
Debtratio	0.756	8.468***
	(1.317)	(2.311)
P/B ratio	2.21e-05***	-0.0120*
	(2.30e-06)	(0.00704)
ROA	-2.568***	-0.696
	(0.857)	(1.591)
TA	5.953***	5.912***
	(0.350)	(0.661)
Constant	-74.07***	-40.29***
	(6.043)	(10.40)
Observations	21,291	6,661
R-squared	0.077	0.078
Number of firms	3,450	1,282
Firm FE	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From the table we see that the results in column (2) are not in line with hypothesis 1, suggesting a negative relationship between the current ESG score of a firm and an increase in the lag1. When using Hoberg-Phillips the effect on a firm's current ESG score as a result of an increase in the lag1 ESG score of its peer-group, indicates a negative effect equal to 0.0867. The estimated change is significant at the 1 % level and the explanatory power of the model suggested to be 7.8%.

7.4.4 Sub-conclusion Hypothesis 1

The results of the Hoberg Phillips panel are not consistent with the results found when using the Fama- French panel. Based on this finding we should question the robustness of the results from our initial model using the Fama-French panel. This will be addressed in section 7.4.7

7.4.5 Hypothesis 2, 3 & 4

In Table 17 we compare the results for the overall individual pillar scores between the Fama-French panel and the Hoberg-Phillips panel. In column (1) and (2) the table displays the results of the Environment pillar for the Fama-French panel and Hoberg-Phillips respectively.

Likewise, column (3) and (4) show the results for the Social pillar and columns (5) and (6) the results for the Governance pillar.

Table 17: Regressions: Model 2, 3 & 4: FF vs HP

VARIABLES	(FF-1) Environment	(HP-2) Environment	(FF-3) Social	(HP-4) Social	(FF-5) Governance	(HP-6) Governance
EnvironmentpeerSiL1	0.369***	-0.0534**				
	(0.0461)	(0.0268)				
SocialpeerSiL1			0.449***	-0.116***		
			(0.0541)	(0.0289)		
GovernancepeerSiL1					0.150***	0.145***
					(0.0570)	(0.0392)
Debtratio	-0.247	9.764***	-2.169	6.257**	4.984**	8.779**
	(1.776)	(2.956)	(1.700)	(2.723)	(2.013)	(3.664)
P/B ratio	-6.22e-05***	-0.0171**	2.70e-05***	-0.00634	0.000122***	-0.0165
	(3.21e-06)	(0.00702)	(2.34e-06)	(0.00554)	(7.30e-06)	(0.0136)
ROA	-3.868***	0.417	-3.126***	2.699	-0.659	-6.541**
	(0.998)	(2.422)	(1.169)	(2.176)	(0.683)	(2.635)
TA	7.979***	8.016***	7.197***	6.369***	2.721***	2.791***
	(0.479)	(0.923)	(0.476)	(0.749)	(0.485)	(1.000)
Constant	-100.5***	-78.10***	-91.22***	-44.16***	-3.204	-1.354
	(8.119)	(14.36)	(8.147)	(11.83)	(8.434)	(15.61)
Observations	21,291	6,661	21,291	6,661	21,291	6,661
R-squared	0.071	0.076	0.059	0.064	0.008	0.016
Number of firms	3,450	1,282	3,450	1,282	3,450	1,282
Firm FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From the results, we observe that both the Social score as well as the Governance score are suggesting the same relationship as for the model using the Fama-French panel. The estimated effect is significant at the 1% level for both scores using the Hoberg-Phillips panel. We observe from column (3) and (4) that the estimated effect is suggested to be almost four times larger for the Fama-French compared to Hoberg-Phillips for the social score. The Governance score though is estimated to be represented by almost the same effect for both panels. However, when it comes to the Environment score, the results from the Hoberg-Phillips panel suggests a reverse effect and the opposite relationship compared to the model using Fama-French. This is not in line with Hypothesis 2 and questions the robustness of the results obtained when using the Hoberg-Phillips panel. The effect estimated for the Environment score using the Hoberg-Phillips panel, equal to -0.0534, is significant at the 5 % level.

7.4.6 Sub-conclusion Hypothesis 2, 3 & 4

The results for the Social score and the Governance score when using the Hoberg-phillips panel are directionally consistent with the relationship found when using the Fama-French panel. However, the estimated effect for the Social score is much larger for the Fama-French panel. The Governance score is suggesting an almost similar effect for both panels. Despite these two pillars being in line with the Fama-French panel and hypothesis 3 and 4, we find that the results for the environment score using the Hoberg-Phillips panel are suggesting a reverse effect. Based on this finding we should question the robustness of the results from our initial model using the Fama-French panel. This will be addressed in section 7.4.7.

7.4.7 Discussion and conclusion: Hoberg-Phillips panel

The results of Hoberg-Phillips are not consistent with the results using Fama- French. This raises an element of doubt regarding the robustness of the Fama French panel models. Illustration 5 shows a graphical comparison of the development of the average growth rates of ESG and pillar scores for both panels. We observe that the Hoberge-Phillips panel witnesses more pronounced fluctuations. On closer inspection we notice that the Governance score has lesser fluctuation than the other two pillars and the ESG score. This could explain the negative coefficients of the ESG score and the Social and Environment pillars in the previous regressions.

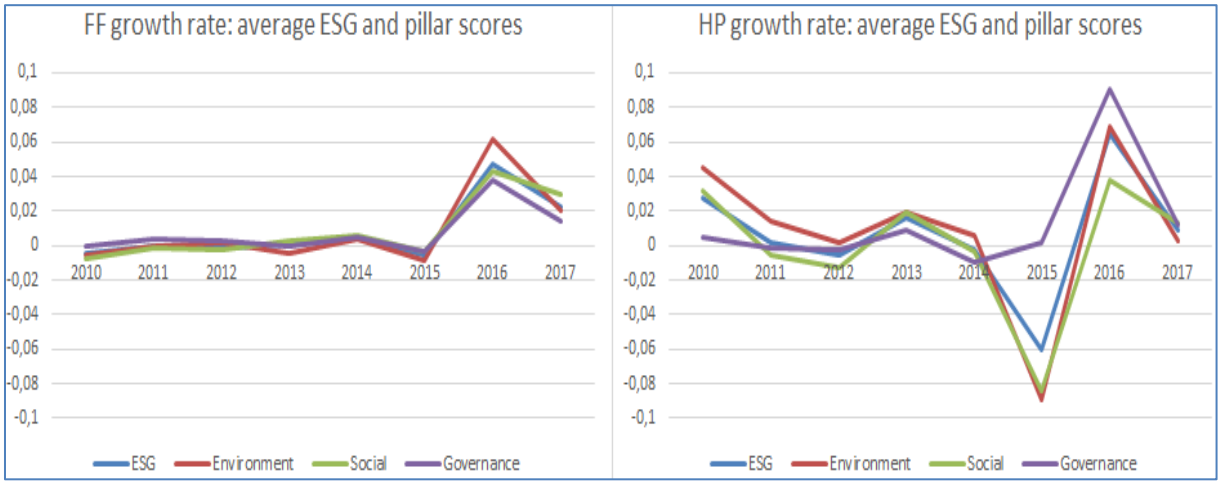


Illustration 5 : FF=Fama-French; HP= Hoberg-Phillips

It is important to note that the sample used in the Hoberg-Phillips model has reduced the Fama French sample to less than one third[25] in sample size, which can render the sample to be

[25] The Fama-French panel has 24742 observations while the Hoberg- Phillips panel has 7960 observations

biased. On closer inspection of the descriptive statistics of both datasets (Table 18), we observe that the means of **TOTALASSETS** and **MARKETCAP** of the Hoberg-Phillips sample have reduced considerably as compared to the Fama-French panel. This implies that the Hoberg-Phillips sample has missed out large-sized firms. We further investigate by examining the quartiles. We observe that the values of the third quartile in the Hoberg-Phillips panel have changed considerably. The Hoberg-Phillips sample appears to be disproportionately biased towards smaller sized firms in terms of total assets and market capitalization.

Table 18: Comparison of samples: FF vs HP

Panel	Variable	N	mean	min	max	sd	p25	p50	p75
FF	TOTALASSETS	24742	1420bn	0.002bn	437000bn	11900bn	3.072bn	13.3bn	113bn
	MARKETCAP	24742	499bn	0.05bn	22400bn	2260bn	2.783bn	10.4bn	81.5bn
HP	TOTALASSETS	7960	31.1bn	0.046bn	2570bn	133bn	2.657bn	6.48bn	17.8bn
	MARKETCAP	7960	17bn	0.0031bn	868bn	42.4bn	2.681bn	5.401bn	13.7bn

FF=Fama-French; HP= Hoberg-Phillips

As mentioned before, the Hoberg-Phillips industry classification scheme is based on textual analysis of 10K filings. The scheme relies on product similarity and is considered unique in capturing the dynamic nature of the competitive landscape. However, it is pertinent to note that while the Fama-French classification scheme has been used by academia as well as the industry for last two decades, the Hoberg-Phillips scheme is relatively new as it was introduced in 2016.

Despite this, Grullon, Larkin and Michaely, (2018) question the assumption of fixed industry classifications implying that the industry structure is static. They describe that firms often introduce new products or improve existing products. Firms at times discontinue certain product lines or follow paths of differentiation / diversification. This implies that firms are not static entities in the marketspace but are constantly venturing into or out of various industries. They state that this is a major limitation of fixed industry classification, which thus fail to capture the dynamic nature of the competitive landscape. Grullon et al (2018) further mention that the Hoberg-Phillips text based industry classification has certain advantages over the fixed industry classifications such as the SIC11, NAICS12 etc. They state that the text based analysis is more realistic since it measures the time variant degree of competition based on product similarity, and thus, offers a larger peer base.

However, Fang et al (2013) draw attention to major limitations of the Hoberg-Phillips industry classification. They state that the peer grouping method employs key words such as nouns and proper nouns from the 10K filings. They argue that the use of unique words instead of topic features in text analysis introduce the element of high dimensionality which leads to inaccurate peer groupings. They further argue that the Hoberg - Phillips method considers business activities and neglects the scale of business, which they point out as a major concern. Moreover, one must note that the Hoberg Phillips classification considers the similarity of product lines as the sole criterion for peer grouping, whereas in the Fama-French fixed classification, the similarity of production processes is the central consideration. It could be that production processes in certain industries may be more closely associated with ESG concerns than the product itself, as compared to other industries. This could be a reason for the inability of the Hoberg-Phillip classification scheme to correctly capture the peer groups in our sample.

As we see there are arguments in favor of both methods and based on the information at hand it is difficult to tell which of them gives the most accurate estimates. However, in this case as the sample has become biased towards smaller firms, we discard the Hoberg-Phillips panel.

7.5 Supplemental analyses

After having articulated a set of hypotheses and estimated the models accordingly we observe that the results prove our hypotheses. We are now interested to further investigate relevant questions and thus refer this section as 'supplemental analysis'. This analysis is of an exploratory and inquisitive character where we want to see, based on the results from the main analysis, if there are differences in the peer-effect within our sample.

First, we test if there is a driving pillar for the overall ESG score. Second, we test whether there is a difference in the peer-effect experienced by firms which have an ESG score above the average of its peer-group as compared to firms which have an ESG score below the average of its peers. Third, we conduct the same tests on the pillar scores.

7.5.1 Driving pillar on ESG overall score

In Table 19 we investigate whether there is a driving pillar for the overall ESG score. In order to test this we have conducted a regression where we run the ESG score of a firm against lag1

average pillar score of the peer-group for each individual pillar, controlling for the same set of financial variables as in the main analysis.

Table 19: Regression: ESG on E, S & G

VARIABLES	(1) ESG
EnvironmentpeerSiL1	0.0927**
	(0.0462)
SocialpeerSiL1	0.222***
	(0.0534)
GovernancepeerSiL1	0.217***
	(0.0383)
Debtratio	0.746
	(1.316)
P/B ratio	1.97e-05***
	(2.65e-06)
ROA	-2.539***
	(0.853)
TA	5.920***
	(0.350)
Constant	-74.76***
	(6.038)
Observations	21,291
Number of firms	3,450
R-squared	0.078
Firm FE	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The result demonstrates that the peer-effects for all three pillars are significant. The Environment pillar represents the lowest peer-effect on the overall ESG score of a firm at the 5% significance level whereas the other two pillars are significant at the 1% level. Furthermore, these two pillars, that is, Social and Governance are considerably higher in magnitude whereas the estimated effect for the Environment pillar is equal to 0.0927 when the average Environment score of its peer-group increases by 1 unit. The regression reveals that a firm experiences the largest peer-effect if its peers increase their Social score, closely followed by the Governance score. The effects being equal to 0.222 for the Social score and 0.217 for the Governance score at the 1 % significance level.

7.5.2 Sub-conclusion: Driving pillar on ESG overall score

The analysis suggests that the Social and Governance pillars have more than double the effect on the overall ESG score of a firm as compared to the Environment score. In other words, there is no clear individual driving pillar for the overall ESG score of a firm, but rather two. The reason for why the Social and Governance pillars are driving the overall ESG score of a firm compared to the Environmental pillar is an interesting topic. The economic intuition could be that the management of a public company would first like to invest in Social and Governance-related issues due to demand from stakeholders. It could be that shareholders are more concerned that a firm is better governed with, say, more independent directors in order to reduce agency problems. Management structure, employee relations and compensation have a direct bearing on shareholders' interest of maximizing profits. On the Social dimension it can be argued that shareholders care more for consumer-protection and allied litigation or diversity of the employees as these directly affect the firm, than environment issues. There could be a number of reasons to explain why the peer effect of Social and Governance pillar scores is higher than Environment. However, we leave this topic open for future research as it is beyond the scope of this paper.

7.5.3 ESG score below vs. above peer-group average

In Table 20 we investigate whether the peer-effect experienced by a firm is lesser when firms have an ESG score above its peers' average compared to firms which have ESG scores below its peers' average score. In order to test if the relationship holds, we generate a dummy 'above' (above = 1 when ESG> ESGpeerSiL1) and accordingly create an interaction term 'abovexESGpeerSiL1', which is the product of the dummy and the independent variable ESGpeerSiL1. In column (1) we test this relationship.

Table 20: Regression: overall ESG: using dummy and interaction tern

VARIABLES	(1) ESG
ESGpeerSiL1	0.665***
	(0.0540)
above	20.86***
	(3.199)
abovexESGpeerSiL1	-0.128**
	(0.0618)
Debtratio	0.327
	(0.932)
P/B ratio	3.18e-05***
	(2.11e-06)
ROA	-1.613***
	(0.470)
TA	3.865***
	(0.249)
Constant	-54.49***
	(4.771)
Observations	21,291
Number of firms	3,450
R-squared	0.391
Firm FE	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The results in Table 20 suggests that the peer-effect experienced by firms with an ESG score above the average of its peers are smaller than for firms which have an ESG score below its peers. From the interaction term 'abovexESGpeerSiL1' it is indicated that the estimated effect is 0.128 lower for firms which have an ESG score above the average of its peers compared to firms which are in the opposite situation. The estimated effect is significant at the 5 % level and the explanatory power equal to 39.1%.

7.5.4 Sub-conclusion: ESG score below vs. above peer-group average

The results suggest that the peer-effect experienced by firms with an ESG score above the average of its peers are smaller than for firms which have an ESG score above its peers.

7.5.5 Pillar scores below vs. above peer-group average

Table 21 reports the test for the same relationship as in section 7.3.4, but for the individual pillar scores of a firm. We investigate whether the peer-effect experienced by a firm with

regard to their individual pillar scores is stronger when the respective individual pillar score of a firm is below its peers' average compared to firms which have individual pillar scores above its peers' average score. In order to test if the relationship holds, we generate a dummy 'above' (above = 1 when Environment> EnvironmentpeerSiL1) and accordingly an interaction term 'aboveENVxEnvironmentpeerSiL1' which is the product of the dummy and the independent variable EnvironmentpeerSiL1. The same process is conducted for the two other pillars, Social and Governance. In column (1) the results for the Environment score is displayed, column (2) the results for Social score and in column (3) the results for the Governance score.

Table 20: Regression: pillars: using dummy and interaction term

VARIABLES	(1) Environment	(2) Social	(3) Governance
EnvironmentpeerSiL1	0.507***		
	(0.0557)		
aboveENV	24.17***		
	(3.453)		
aboveENVxEnvironmentpeerSiL1	-0.0744		
	(0.0658)		
SocialpeerSiL1		0.652***	
		(0.0600)	
aboveSOC		24.97***	
		(3.699)	
aboveSOCxSocialpeerSiL1		-0.118*	
		(0.0712)	
GovernancepeerSiL1			0.479***
			(0.0561)
aboveGOV			24.67***
			(3.386)
aboveGOVxGovernancepeerSiL1			-0.0360
			(0.0668)
Debtratio	-1.378	-1.712	2.066
	(1.252)	(1.201)	(1.303)
P/B ratio	1.54e-05***	-1.24e-06	5.67e-05***
	(2.87e-06)	(1.96e-06)	(5.29e-06)
ROA	-2.188**	-2.425***	-0.280
	(0.850)	(0.751)	(0.566)
TA	4.680***	4.614***	1.519***
	(0.343)	(0.331)	(0.316)
Constant	-62.38***	-68.16***	-10.79*
	(6.192)	(6.216)	(5.837)
Observations	21,291	21,291	21,291
R-squared	0.418	0.401	0.483
Number of Firms	3,450	3,450	3,450
Firm FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From Table 20 we find that for all the three individual pillar scores the estimated effect is directionally the same as found for the overall ESG score in section above. Despite this, the results indicate that only the Social score is significant, but only at the 10 % level. The estimated effect for the Environment score and the Governance score are both suggested to be insignificant. Thus, we conclude that the results for the pillar scores in this analysis cannot be interpreted in a meaningful way.

7.5.6 Sub-conclusion: Pillar scores below vs. above peer average

Based on the fact that neither of the estimated effects of the three pillar scores are suggested to be signifcant within the 5 % level, we cannot draw a conclusion.

8.0 Implications and discussion

The results from the main analysis suggest that there are significant peer-effects related to the overall ESG scores as well as the individual pillar scores of firms. This implies that firms care about the ESG performance of its peers and when the firms within the peer-group increases its score, each individual firm also feels compelled to improve its own score. Our findings can be read in conjunction with existing literature which shows that higher ESG scores correlate with better financial performance as well as represents a competitive advantage. These aspects can be considered as incentives for a firm to improve its ESG score.

We also find that depending on whether a firm has an ESG or individual pillar score below or above the average score of its peer-group, the peer-effect differs. The results suggests that firms which have a score below the average score of its peer-group, experiences a stronger peer-effect as compared to firms which have a score above the average of its peer-groups.
It is assessed that these findings may have an implicit connotation for relevant stakeholders. For instance, governments and regulatory bodies may recognize the importance of ESG rating agencies and their implicit service. If more firms can be brought under the coverage of rating agencies, then it will increase the peer pressure on firms and increase the competition in the pool. The peer effect will compel each firm to improve its ESG performance. This will also exert pressure on firms which are hesitant to disclose ESG information and be evaluated by rating agencies. Furthermore, the findings may be of interest for investors as well. Amel-Zadeh and Serafeim (2018) demonstrate, through a global survey, that investors primarily seek ESG information for investment decisions in order to discern the risk and not the

competitive positioning of a firm. One argument could be that, if ESG performance of a firm is a function of the ESG performance of its peers, then investors might wrongly attribute a firm's ESG performance as genuine and thus less risky. Although, it is pertinent to point out the counter-argument, that a firm's mere intent to improve its ESG performance due to peer pressure, is not mala fide per se. However, it is assessed that reading Amel-Zadeh and Serafeim (2018) in conjunction with our findings, implies that investors may need to be attentive to distinguish between a genuine intent of a firm to improve its ESG performance with improvement due to peer pressure. In this connection it is important to also note that firms tend to adopt ESG measures more if they operate in an environment where the regulatory and legal framework demands ESG compliance (Liang and Renneboog, 2017).

Our findings contribute to existing literature by showing that firms experience a significant peer-effect related to, not only their ESG score, but also for the individual pillar scores as well. Furthermore, the results from the supplemental analysis open avenues for future research.

8.1 Limitations

We acknowledge that our study has certain limitations. Firstly, we have confined our study to the Refinitiv ESG database and the firms thereof. It can be argued that this leads to a bias in our sample as the study does not cover ESG data from other rating agencies. We accept this line of argument. Ideally, if we could incorporate ESG data from databases of other ESG rating agencies, it would make our study more robust. Secondly, the firms in the Refinitiv database are restricted to public companies as Refinitiv predominantly relies upon data disclosed by the firms in their annual reports, websites and stock exchange filings. We accept the fact that our data is restricted to public companies, however we wish to point out that private firms have lesser degree of obligation to disclose such ESG information. Therefore, we restricted our study to public firms. Lastly, we acknowledge the limitations within the ESG data itself. Although, the Refinitiv database covers more than 6500 firms, but this is a cumulative figure from 2002 to 2017. The drawback being that a large number of these had to be dropped from the sample as they have less than 3 year ESG scores.

REFERENCES

Adhikari, B.K. and Agrawal, A., 2018. Peer influence on payout policies. *Journal of Corporate Finance*, *48*, pp.615-637.

Bowen, H.R. and Johnson, F.E., 1953. *Social responsibility of the businessman*. Harper.

Brown, J.R., Ivković, Z., Smith, P.A. and Weisbenner, S., 2008. Neighbors matter: Causal community effects and stock market participation. *The Journal of Finance*, *63*(3), pp.1509-1531.

Cao, J., Liang, H. and Zhan, X., 2018. Peer effects of corporate social responsibility. *Available at SSRN 2831834*.

Carroll, A.B., 1999. Corporate social responsibility: Evolution of a definitional construct. *Business & society*, *38*(3), pp.268-295.

Chatterji, A.K., Levine, D.I. and Toffel, M.W., 2009. How well do social ratings actually measure corporate social responsibility?. *Journal of Economics & Management Strategy*, *18*(1), pp.125-169.

Cheng, S., 2008. Board size and the variability of corporate performance. *Journal of financial economics*, *87*(1), pp.157-176.

Del Bosco, B. and Misani, N., 2016. The effect of cross-listing on the environmental, social, and governance performance of firms. *Journal of World Business*, *51*(6), pp.977-990.

Dorfleitner, G., Halbritter, G. and Nguyen, M., 2015. Measuring the level and risk of corporate responsibility–An empirical comparison of different ESG rating approaches. *Journal of Asset Management*, *16*(7), pp.450-466.

Du, S., Bhattacharya, C.B. and Sen, S., 2011. Corporate Social Responsibility and Competitive Advantage: Overcoming the Trust Barrier. *Management Science*, *57*(9).

Escrig-Olmedo, E., Fernández-Izquierdo, M. Á., Ferrero-Ferrero, I., Rivera-Lirio, JM, & Muñoz-Torres, MJ (2019). Rating the Raters: Evaluating how ESG Rating Agencies integrate sustainability principles. *Sustainability* , *11* (3), 915.

Fang, F., Dutta, K. and Datta, A., 2013. LDA-based industry classification. In *International Conference on Information Systems (ICIS 2013): Reshaping Society Through Information Systems Design* (Vol. 3, pp. 2500-2509)

Fatemi, A., Fooladi, I. and Tehranian, H., 2015. Valuation effects of corporate social responsibility. *Journal of Banking & Finance*, *59*, pp.182-192.

Ferrell, A., Liang, H. and Renneboog, L., 2016. Socially responsible firms. *Journal of Financial Economics*, *122*(3), pp.585-606.

Ferrero-Ferrero, I., Fernández-Izquierdo, M. and Muñoz-Torres, M., 2016. The effect of environmental, social and governance consistency on economic results. *Sustainability*, *8*(10), p.1005.

Freeman, R.E., 2004. The stakeholder approach revisited. *Zeitschrift für Wirtschafts-und Unternehmensethik*, *5*(3), pp.228-254.

Friede, G., Busch, T. and Bassen, A., 2015. ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, *5*(4), pp.210-233.

Galbreath, J., 2013. ESG in focus: The Australian evidence. *Journal of business ethics*, *118*(3), pp.529-541.

Gillan, S., Hartzell, J.C., Koch, A. and Starks, L., 2010. Firms' environmental, social and governance (ESG) choices, performance and managerial motivation. *Unpublished working paper*.

Graves, S.B. and Waddock, S.A., 1990. Institutional ownership and control: Implications for long-term corporate strategy. *Academy of Management Perspectives*, *4*(1), pp.75-83.

Grullon, G., Larkin, Y. and Michaely, R., 2018. Are US industries becoming more concentrated?. *Forthcoming, Review of Finance*.

Halbritter, G. and Dorfleitner, G., 2015. The wages of social responsibility—where are they? A critical review of ESG investing. *Review of Financial Economics*, *26*, pp.25-35.

Hoberg, G. and Phillips, G., 2010. Product market synergies and competition in mergers and acquisitions: A text-based analysis. *The Review of Financial Studies*, *23*(10), pp.3773-3811.

Hoberg, G. and Phillips, G., 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, *124*(5), pp.1423-1465.

Hornuf, L. and Schwienbacher, A., 2018. Market mechanisms and funding dynamics in equity crowdfunding. *Journal of Corporate Finance*, *50*, pp.556-574.

Jenter, D. and Kanaan, F., 2015. CEO turnover and relative performance evaluation. *the Journal of Finance*, *70*(5), pp.2155-2184.

Keil, J., 2017. The trouble with approximating industry concentration from Compustat. *Journal of Corporate Finance*, *45*, pp.467-479.

Krüger, P., 2015. Corporate goodness and shareholder wealth. *Journal of financial economics*, *115*(2), pp.304-329

Leary, M.T. and Roberts, M.R., 2014. Do peer firms affect corporate financial policy?. *The Journal of Finance*, *69*(1), pp.139-178.

Liang, H. and Renneboog, L., 2017. On the foundations of corporate social responsibility. *The Journal of Finance*, *72*(2), pp.853-910.

Lins, K.V., Servaes, H. and Tamayo, A., 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance*, *72*(4), pp.1785-1824.

Liu, S. and Wu, D., 2016. Competing by conducting good deeds: The peer effect of corporate social responsibility. *Finance Research Letters*, *16*(C), pp.47-54.

Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. *The review of economic studies*, *60*(3), pp.531-542.

Miller, T. and del Carmen Triana, M., 2009. Demographic diversity in the boardroom: Mediators of the board diversity–firm performance relationship. *Journal of Management studies*, *46*(5), pp.755-786.

Parmar, B.L., Freeman, R.E., Harrison, J.S., Wicks, A.C., Purnell, L. and De Colle, S., 2010. Stakeholder theory: The state of the art. *The academy of management annals*, *4*(1), pp.403-445.

Porter, M.E. and Kramer, M.R., 2006. The link between competitive advantage and corporate social responsibility. *Harvard business review*, *84*(12), pp.78-92.

Stellner, C., Klein, C. and Zwergel, B., 2015. Corporate social responsibility and Eurozone corporate bonds: The moderating role of country sustainability. *Journal of Banking & Finance*, *59*, pp.538-549.

Van Marrewijk, M., 2003. Concepts and definitions of CSR and corporate sustainability: Between agency and communion. *Journal of business ethics*, *44*(2-3), pp.95-105.

Waddock, S.A. and Graves, S.B., 1997. The corporate social performance–financial performance link. *Strategic management journal*, *18*(4), pp.303-319.

Zhang, A.C., Fang, J., Jacobsen, B. and Marshall, B.R., 2018. Peer effects, personal characteristics and asset allocation. *Journal of Banking & Finance*, *90*(C), pp.76-95.

APPENDIX A

VARIABLES

The list and description of variables used in the four model equations are tabulated below:

Serial No	Variable	Description
1	$ESG\_score_{i,t}$	overall ESG score of a $firm_i$ at year *'t'*
2	$(ESG\_peer)_{i,t-1}$	the average of the overall ESG scores of the peers of $firm_i$ at year *'t-1'*
3	$(ESG\_peer)_{i,t-2}$	the average of the overall ESG scores of the peers of $firm_i$ at year *'t-2'*
4	$FirmSize_{i,t}$	Equal to the natural logarithmic value of total assets of a $firm_i$ at year *'t'*
5	$Debt\ Ratio_{i,t}$	Ratio of total debt to total assets of a $firm_i$.
6	$(Price/Book)_{i,t}$	Ratio of the share price to the book value per share of a $firm_i$.
7	$ROA_{i,t}$	Return of Assets, that is, the ratio of net income to total assets of a $firm_i$
8	α_i	the unobserved firm specific time invariant fixed effect
9	$\varepsilon_{i,t}$	error term
10	$E\_score_{i,t}$	Environmental pillar score of a $firm_i$ at year *'t'*
11	$S\_score_{i,t}$	Social pillar score of a $firm_i$ at year *'t'*
12	$G\_score_{i,t}$	Governance pillar score of a $firm_i$ at year *'t'*
13	$(E\_peer)_{i,t-1}$	The average Environmental pillar score of peers of a $firm_i$ at year *'t-1'*
14	$(E\_peer)_{i,t-2}$	The average Environmental pillar score of peers of a $firm_i$ at year *'t-2'*
15	$(S\_peer)_{i,t-1}$	The average Social pillar score of peers of a $firm_i$ at year *'t-1'*

The list and description of variables generated for Stata regression outputs are tabulated below:

S No	Variables	Description
1	ESG	ESG score of a firm
2	Environment	Environment pillar score of a firm
3	Social	Social pillar score of a firm
4	Governmental	Governance pillar score of a firm
5	Debtratio	Ratio of total debt to total assets of a $firm_i$.
6	P/B	Ratio of the share price to the book value per share of a $firm_i$.
7	ROA	Return of Assets, that is, the ratio of net income to total assets of a $firm_i$
8	TA	Equal to the natural logarithmic value of total assets of a $firm_i$ at year *'t'*
17	ESGpeerWtL1	Weighted average ESG score of peers Lag1
18	ESGpeerWtL2	Weighted average ESG score of peers Lag2
19	EnvironmentpeerWtL1	Weighted average Environment score of peers Lag1
20	EnvironmentpeerWtL2	Weighted average Environment score of peers Lag2
21	SocialpeerWtL1	Weighted average of Social score of peers Lag1
22	SocialpeerWtL2	Weighted average of Social score of peers Lag2
25	GovernancepeerWtL1	Weighted average Governance score of peers Lag1
26	GovernancepeerWtL2	Weighted average Governance score of peers Lag1
27	ESGpeerSiL1	Simple average ESG score of peers Lag1
28	ESGpeerSiL2	Simple average ESG score of peers Lage2
29	EnvironmentpeerSiL1	Simple average Environment score of peers Lag1
30	EnvironmentpeerSiL1	Simple average Environment score of peers Lag2
31	SocialtpeerSiL1	Simple average Social score of peers Lag1
32	SocialpeerSiL2	Simple average Social score of peers Lag2
33	GovernancepeerSiL1	Simple average Governance score of peers Lag1
34	GovernancepeerSiL1	Simple average Governance score of peers Lag2
35	above	Dummy = 1 if ESG>ESGpeerSiL1
36	abovexESGpeerSiL1	'above' * 'ESGpeerSiL1'
37	aboveENV	Dummy = 1 if Environment>EnvironmentpeerSiL1
38	aboveENVxEnvironmentpeerSiL1	'aboveENV'*'EnvironmentpeerSiL1'

39	aboveSOC	Dummy = 1 if Social>SocialpeerSiL1
40	aboveSOCxSocialpeerSiL1	'aboveSOC'*'SocialpeerSiL1
41	aboveGOV	Dummy = 1 if Governance>GovernancepeerSiL1
42	aboveGOVxGovernancepeerSiL1	'aboveGOV'*'GovernancepeerSiL1'

APPENDIX B

Construction of dataset: Fama Frenchpeer group method

The Fama French 49 Industry Classification scheme was downloaded from Kenneth French Data Library website[26]. As stated earlier, the initial panel had 4858 firms which had a minimum of three years ESG data which we obtained from the Refinitiv database. We deleted firms which did not have SIC codes or if the SIC codes did not match with the Fama French Industry classification. We were left with 4401 firms. Thereafter we arranged the 4401 firms on the basis of the four-digit SIC codes and sorted them into the 49 Industry groups. Thereafter, we created all possible combinations of pair- wise peer groupings within the 49 Industry groups using R programming. Subsequently, we merged them on R Studio with actual years of ESG data in the initial unbalanced panel, keeping the integrity of the 49 Industry groups intact. Thereafter we removed the firms for which there were no peers available or no financial data available. In the final dataset, we were left with an unbalanced panel of 3,545 firms covering a period from 2009 to 2017 corresponding to 25,747 firm- year observations.

The preliminary step was to obtain ESG scores and pillar scores from Refinitiv in Datastream. The Refinitiv ESG data consists of ten category scores and an overall ESG score. These scores are awarded to a firm on a scale of 0 to 100 on an annual basis. Refinitiv shared the updated methodology[27] of calculating the three pillar scores from the ten category scores. The Refinitiv ESG database identifies firms on the basis of ISIN[28]. This enabled us to identify them on Compustat and accordingly obtain their 4-digit SIC codes. These codes were used to segregate each firm into its industry peer group using the Fama French49 Industry Classification scheme. The final dataset is an unbalanced panel of 3,545 firms covering a period from 2009 to 2017 corresponding to 25,747 firm- year observations.

[26] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
[27] The detailed methodology for calculation of pillar scores from the ten category scores is explained on Refinitiv website: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf
[28] ISIN (International Securities Identification Number) is a unique 12 digit alphanumeric code allotted to internationally traded securities.

APPENDIX C

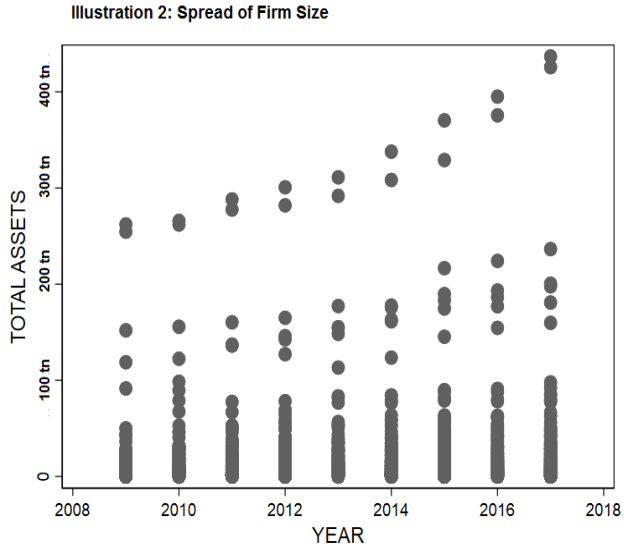
Mean & Standard Deviation: in Industry groups

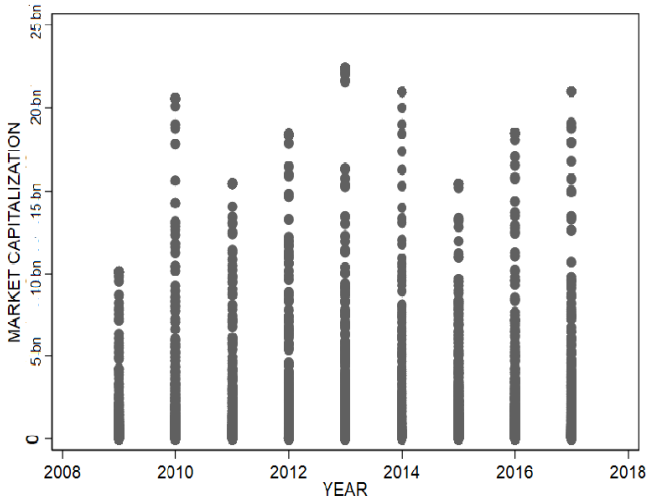
Industry Group	Avg ESG	StdDev ESG	Avg E	StdDev E	Avg S	StdDev S	Avg G	StdDev G
1 Agric	50,85	16,01	47,96	24,57	54,37	20,15	49,97	20,99
10 Clths	57,12	17,24	57,02	21,70	59,28	22,40	54,73	18,64
11 Hlth	57,25	16,05	53,40	19,29	59,10	18,67	59,40	21,62
12 MedEq	49,13	18,82	48,82	24,01	50,20	22,86	48,23	20,80
13 Drugs	52,17	17,94	53,22	23,78	53,11	21,93	49,91	21,30
14 Chems	51,37	18,10	52,75	23,63	50,35	22,80	51,04	22,01
15 Rubbr	56,76	14,61	62,99	22,74	54,70	18,17	52,21	19,94
16 Txtls	53,84	19,70	57,33	23,87	52,67	21,51	51,30	25,06
17 BldMt	54,66	16,55	56,94	22,35	54,58	21,41	52,21	20,39
18 Cnstr	52,20	18,42	54,05	23,71	52,28	22,18	50,05	21,08
19 Steel	52,45	19,79	54,14	23,75	51,71	23,98	51,42	22,87
2 Food	47,85	17,16	48,56	22,74	46,48	22,28	48,65	20,06
20 FabPr	67,00	16,66	63,49	24,54	74,21	17,86	62,51	20,84
21 Mach	50,19	17,56	51,79	23,49	49,56	22,69	49,15	20,08
22 ElcEq	51,07	17,21	52,20	21,89	50,22	22,48	50,80	22,54
23 Autos	54,94	19,67	57,71	24,08	57,03	24,51	49,42	21,17
24 Aero	51,46	18,48	50,90	24,31	48,91	21,46	55,05	20,34
25 Ships	53,28	18,43	48,25	22,50	54,88	19,66	57,02	18,78
26 Guns	50,21	7,84	46,40	13,09	59,75	16,73	43,37	18,71
27 Gold	49,90	14,85	46,85	21,43	52,21	19,92	50,60	18,15
28 Mines	50,16	16,35	48,68	19,97	51,64	20,81	50,09	19,64
29 Coal	44,02	15,97	42,61	19,49	45,64	23,55	43,70	17,86
3 Soda	54,10	22,05	57,89	26,40	48,74	23,75	56,12	26,70
30 Oil	53,41	17,31	53,42	22,84	53,22	21,38	53,63	19,57
31 Util	51,68	18,52	52,02	24,53	52,65	22,27	50,19	20,54
32 Telcm	54,24	17,82	55,05	23,82	56,90	22,30	50,23	20,64
33 PerSv	54,85	20,10	52,18	27,18	56,40	21,17	56,03	21,99
34 BusSv	50,89	17,35	49,99	23,44	51,15	20,22	51,58	21,26
35 Comps	49,83	20,70	53,37	24,91	47,67	22,97	48,40	23,81
35a Cmptr SW	55,24	15,78	57,62	21,66	54,89	17,96	52,98	21,33
36 Chips	47,30	19,58	48,27	24,54	46,51	25,17	47,14	21,40
37 LabEq	51,36	17,19	52,60	23,06	49,06	19,26	52,67	21,36
38 Paper	50,90	19,93	48,13	26,93	52,76	22,33	51,83	20,12
39 Boxes	59,16	17,94	60,66	24,10	60,10	22,31	56,41	20,46
4 Beer	56,32	13,95	65,73	20,85	58,20	16,82	43,66	14,67
40 Trans	51,47	17,93	52,25	22,67	50,42	21,57	51,80	21,68
41 Whlsl	50,42	16,16	50,27	22,64	51,43	20,05	49,41	20,37
42 Rtail	52,40	17,63	51,94	22,01	52,58	21,65	52,68	21,49
43 Meals	54,05	16,72	54,02	21,99	54,38	19,86	53,70	18,62
44 Banks	51,28	17,89	50,57	23,26	51,40	21,82	51,93	21,70
45 Insur	52,86	17,65	52,06	23,76	52,92	22,52	53,66	18,43
46 RlEst	52,61	17,66	52,74	23,51	52,34	23,08	52,79	20,63
47 Fin	51,55	17,93	52,38	23,57	52,46	21,35	49,58	21,24
48 Other	49,37	16,22	49,59	21,94	48,49	22,63	50,15	18,78
5 Smoke	56,62	11,44	49,08	16,01	62,35	14,81	58,36	19,94
6 Toys	54,62	18,10	58,81	25,36	56,64	22,40	47,58	21,33
7 Fun	50,44	16,05	52,26	21,80	53,06	21,63	45,35	21,27
8 Books	59,15	14,55	60,64	18,41	60,39	18,29	56,03	17,79
9 Hshld	52,53	18,68	53,62	23,37	51,55	23,59	52,45	22,60

Abbreviations
Avg: average
StdDev: standar deviation
E: Environment
S: Social
G: Governance

APPENDIX D

Scatter plot: firm size Fama-French panel





APPENDIX E

VARIABLES	(1) Environment	(2) Environment	(3) Social	(4) Social	(5) Governance	(6) Governance
EnvironmentpeerSiL1	0.554***	0.369***				
	(0.0560)	(0.0461)				
EnvironmentpeerSiL2	-0.241***					
	(0.0445)					
SocialpeerSiL1			0.636***	0.449***		
			(0.0628)	(0.0541)		
SocialpeerSiL2			-0.195***			
			(0.0519)			
GovernancepeerSiL1					0.188***	0.150***
					(0.0614)	(0.0570)
GovernancepeerSiL2					-0.0435	
					(0.0615)	
Debtratio	1.007	-0.247	-2.048	-2.169	7.421***	4.984**
	(1.966)	(1.776)	(1.859)	(1.700)	(2.227)	(2.013)
P/B ratio	-5.60e-05***	-6.22e-05***	-1.43e-05***	2.70e-05***	9.34e-05***	0.000122***
	(3.64e-06)	(3.21e-06)	(3.04e-06)	(2.34e-06)	(7.72e-06)	(7.30e-06)
ROA	-2.560***	-3.868***	-1.522	-3.126***	0.248	-0.659
	(0.743)	(0.998)	(1.088)	(1.169)	(0.666)	(0.683)
TA	7.660***	7.979***	6.929***	7.197***	2.036***	2.721***
	(0.538)	(0.479)	(0.529)	(0.476)	(0.569)	(0.485)
Constant	-92.11***	-100.5***	-86.08***	-91.22***	8.332	-3.204
	(9.511)	(8.119)	(9.484)	(8.147)	(10.40)	(8.434)
Observations	17,841	21,291	17,841	21,291	17,841	21,291
R-squared	0.070	0.071	0.058	0.059	0.007	0.008
Number of firms	3,444	3,450	3,444	3,450	3,444	3,450
Firm FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

APPENDIX F

CORRELATION MATRIX

	ESGpeerSiL1	ESGpeerSiL2	EnvironmentpeerSiL1	EnvironmentpeerSiL2	SocialpeerSiL1	SocialpeerSiL2	GovernancepeerSiL1	GovernancepeerSiL2
ESGpeerSiL1	1							
ESGpeerSiL2	0.827	1						
EnvironmentpeerSiL1	0.8546	0.7169	1					
EnvironmentpeerSiL2	0.6861	0.8501	0.8221	1				
SocialpeerSiL1	0.8956	0.7753	0.6901	0.5744	1			
SocialpeerSiL2	0.7672	0.896	0.6072	0.6853	0.8698	1		
GovernancepeerSiL1	0.6499	0.4842	0.2961	0.1901	0.4137	0.3223	1	
GovernancepeerSiL2	0.4956	0.6158	0.1982	0.2493	0.347	0.379	0.7604	1

APPENDIX G

Construction of dataset: Hoberg Phillips peer group

The Hoberg - Phillips peer groupings were obtained from the open source website called Hoberg Phillips Data Library. The downloaded text file (hereafter referred as TNIC2)[29] pertains to the mapping of firms which execute 10K filings[30] in USA. The data contained over 32.2 million rows as each firm in the Hobert- Phillips database is mapped against the remaining firms on the basis of product similarity. On deduplication we observed that the file contained 13808 unique GVKEYS[31]. We downloaded the financial data of the firms from Refinitiv on Datastream. Additional financial parameters were obtained from Compustat.

In order to create peer-groupings we have employed a text based industry classification developed by Hoberg and Phillips (2010)[32]. The Text-based Network Industry Classification (TNIC) defines the peer group of any given firm in an industry.

Our initial sample consists of 6554 firms from the Refinitiv ESG database covering a time period of 2009 to 2017. The relevant financial data was downloaded from Datastream and supplemented with financial data from Compustat. We started the exercise with this Refinitiv ESG universe of 6554 firms and taking the intersection with over 75000 firms[33] from the Compustat data and the 13808 firms of the Hoberg- Phillips universe. Thereafter, we retained the firms which had at least 3 years continuous ESG scores from 2009 to 2017 in order to construct a final balanced panel of 1299 firms. Afterwards we merged the GVKEYS of these 1299 firms in R with the TNIC2 data. After selecting the rows pertaining to 2009 -17 we were left with over 600,000 peer -grouping observations from the original 32.2 million observations in the Hoberg-Phillips universe. We incorporated these firm-year-peer grouping observations in order to calculate the simple and weighted average ESG scores of the peer group as well as for pillar scores of the peer groups for each firm. The same procedure was

[29] TNIC2 is a .txt file consisting over 32 million rows showing peer groupings of 13808 firms. It maps each of the 13808 firms with the others based on product- similarity textual analysis. Only if the similarity is above a certain baseline value then the two firms are deemed as peers. These peer groupings are then graded on a scale of 0 to 10 till four places after the decimal.

[30] 10K is an annual report that is mandatory for a firm to file with the US Securities and Exchange Commission (SEC). It contains information of the firm including its equities, executive compensation, financial statements, products offered etc.

[31] The GVKEY or Global Company Key is a unique 6 digit identification number allotted by Compustat to a firm.

[32] The Hoberg-Phillips TNIC uses textual analysis of time- variant product descriptions in 10K filings to create peers. The rationale is that a firm A is a competitor of firm B if both have similarity of product lines. The strength of the peer relationship is proportional to the strong tendency of product vocabulary in the business description section of the 10K report. Thus the TNIC peer groupings are dynamic unlike fixed industry groupings such as NAICS, SIC etc which focus on production processes rather than product lines.

[33] We created a masterbank of 75000+ firms which had a GVKEY allotted by Compustat as a unique identifier.

adopted for obtaining the simple average scores as well. These scores were integrated into the final panel.

Hoberg-Phillips: Sample Selection

We used the same selection method as was used for the Fama French panel mentioned above. In order to avoid bias we retained firms which had ESG scores and financial data irrespective of firm size/ performance parameters. Since the selection criteria of minimum 3 year data is consistent we commenced with the initial panel of 4858 firms. Thereafter we excluded firms for which financial data was not available. Then we proceeded to identify the remaining firms on the basis of GVKEY which is mandatory in order to generate Hoberg-Phillips peer groupings. This is important to note as a limiting factor while applying the Hoberg-Phillips methodology as it reduced the panel to less than one-third in size. In the process our final dataset was left with 7960 firm-year observations whereas, the main FF panel has 24742 observations.

Text based industry classification

The Hoberg-Phillips text based industry classification (TNIC)[34] data can be obtained from the open source website Hoberg-Phillips Data Library[35].The Hoberg-Phillips TNIC uses textual analysis of time- variant product descriptions in 10K[36] filings to create peers. The rationale is that a firm A is a competitor of firm B if both have similarity of product lines. The strength of the peer relationship is proportional to the strong tendency of product vocabulary in the business description section of the 10K report. Thus, the TNIC peer groupings are dynamic unlike fixed industry groupings such as NAICS, SIC etc which focus on production processes rather than product lines.

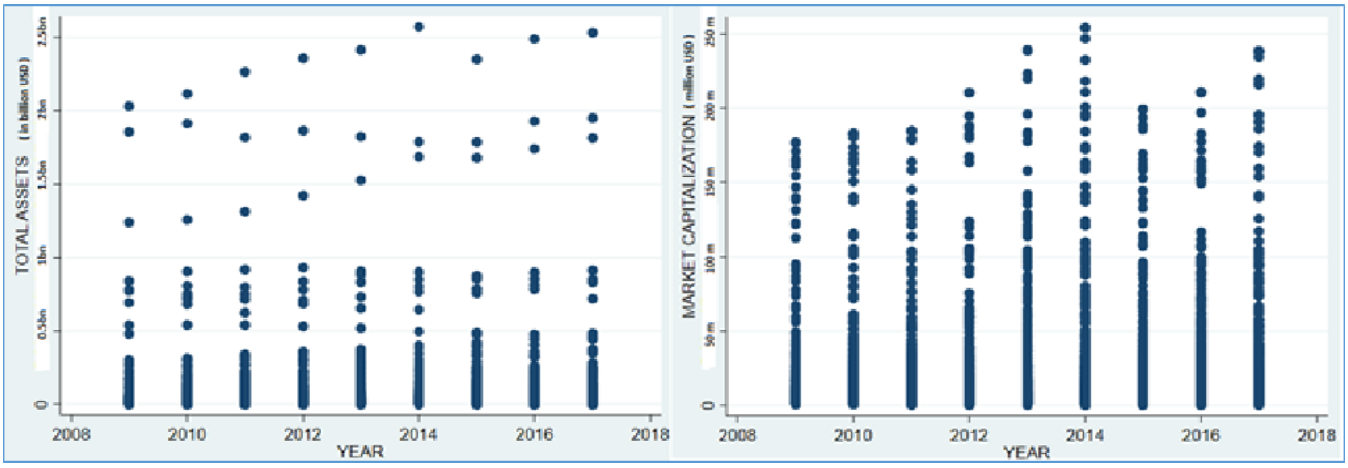
Grullon, Larkin and Michaely, (2018) describe that firms often introduce new products or improve existing products. Firms at times discontinue certain product lines or follow paths of differentiation / diversification. This implies that firms are not static entities in the marketspace but are constantly venturing into or out of various industries. They question the

[34] Text-based Network Industry Classification
[35] http://hobergphillips.tuck.dartmouth.edu/
[36] 10K is a mandatory comprehensive annual report of financial performance that a public company operating in USA needs to file with the Securities and Exchange Commission (SEC). It contains a section (Part I, Item 1) dedicated towards business description. These disclosures are utilized in the TNIC to arrive at degree of product similarity, which is the basis for creating peer groupings in the Hoberg Phillips TNIC.

assumption of fixed industry classifications that the industry structure is static. They state that it is a major limitation of fixed industry classification, and thus, fail to capture the dynamic nature of the competitive landscape. Grullon et al (2018) further mention that the Hoberg-Phillips text based industry classification has certain advantages over the fixed industry classifications such as the SIC[37], NAICS[38] etc. They state that the text-based analysis is more realistic since it measures the time variant degree of competition based on product similarity, and thus, offers a larger peer base.

APPENDIX H

Scatter plot: firm size Hoberg-Phillips panel



[37] Standard Industrial Classification
[38] North American Industry Classification System

APPENDIX I

Between and within Standard Deviation

Variable		Mean	Std. Dev.	Min	Max	Observations
ESG	overall	51.83903	17.88437	7.30925	97.48635	N = 24742
	between		16.34465	11.0889	91.38843	n = 3450
	within		6.880531	-9.733763	95.7259	T bar = 7.17159
Environment	overall	52.22281	23.37948	2.858824	99.24118	N = 24742
	between		21.5473	4.480588	97.67928	n = 3450
	within		8.9723	-8.632977	119.3815	T bar = 7.17159
Social	overall	52.13981	22.00879	2.501972	99.09	N = 24742
	between		19.91305	5.096291	97.51531	n = 3450
	within		8.870677	-15.20502	103.0579	T bar = 7.17159
Governance	overall	51.06111	21.03702	1.02	99.35541	N = 24742
	between		18.14991	5.712664	92.83627	n = 3450
	within		10.99212	-4.592683	102.1729	T bar = 7.17159