Master's Thesis - master
Innovation Sciences

ESG: Green Intentions or Green Illusions

Influence of ESG Ratings on Firms' Innovative Performance in the European Union

Jolan Keijzer

Utrecht University | Innovation Sciences

Utrecht University

ESG: Green Intentions or Green Illusions

Influence of ESG Ratings on Firms' Innovative Performance in the European Union

by

Jolan Keijzer

Student Name	Student Number
Jolan Keijzer	6920012

UU Supervisor:	Iris Wanzenböck
Date:	24-05-2023
Programme:	Innovation Sciences
University:	Utrecht University

Utrecht University

Abstract

Environmental Social Governance (ESG) has become increasingly important to incorporate non-financial information in investment decisions. ESG disclosure requirements are gaining momentum in Europe, reinforced by the Corporate Sustainability Reporting Directive (CSRD). These policies are introduced to promote sustainable development, while simultaneously creating shareholder value. However, most European ESG research solely focuses on the financial performance of firms. This finance literature shows that ESG leads to increased access to capital at lower cost, which results in increased R&D spending. However, the literature fails to examine potential non-financial benefits of ESG. This leads to a lack of understanding of why firms engage in ESG and what benefits this results in. Legitimacy and signalling theory provide opposing explanations for why firms engage in ESG. Legitimacy theory posits that poorly-performing firms with regards to ESG engage with ESG in order to legitimise other practices that misalign with societal values. Whereas, signalling theory posits that well-performing firms use disclosure to highlight their superior ESG performance. Neither of these theories provide conclusive evidence on non-financial benefits of ESG. This paper zooms in on non-financial benefits of ESG through green innovative performance, a component of sustainable performance, and distinguishes between green and non-green innovation in order to examine the incongruence of legitimacy and signalling theory. Concretely, this thesis asks: *"How do changes in firms' ESG rating affect the green and non-green innovative performance of firms in the European Union and how do these effects relate to each-other"?*

To answer this, negative binomial regression analysis and the Seemingly Enrelated Estimation method are applied to patent and ESG rating data, which measure firms' innovative and ESG performance. The sample consists of large European firms in the period from 2014-2017 and a distinction is made in green and non-green innovation through CPC class Y02. As expected, ESG positively influences both green and non-green innovation, which highlights the presence of the effects of both legitimacy and signalling theory. For green innovation this effect is especially evident for smaller firms. Surprisingly, it is shown that the effect of ESG is stronger on non-green innovation than on green innovation. These findings provide implications for the CSRD. First, CSRD extends the reporting requirement to smaller firms, which show a larger increase in green innovation. Second, larger firms seem to increase their ESG without improving their underlying performance. The audit requirement of the CSRD might reduce this discrepancy between the ESG rating and underlying performance.

Contents

1 Introduction

From 2023 onward, the Corporate Sustainability Reporting Directive (CSRD) issued by the European Commission (EC) comes into effect (EC, 2021a). This directive extends existing Environmental Social Governance (ESG) reporting requirements set by the Non-Financial Reporting Directive (NFRD). Concretely, CSRD extends the scope of companies that are required to report, requires audits of reported information, and introduces more detailed requirements and standards (EC, 2021b). According to the EC (2021b) the goal of CSRD is to *"enable investors to re-orient investments towards more sustainable technologies and businesses"*. Most European ESG-related literature is of a financial nature. As a result, European literature primarily focuses on the effect that changes in ESG ratings have on financial (De Lucia et al., 2020; Landi & Sciarelli, 2018; Taliento et al., 2019) and stock performance (Giese et al., 2019). Hence, existing research fails to examine potential non-financial benefits. With the context of the upcoming CSRD, it is important to consider whether these financial benefits merely result in increased shareholder value or manage to result in the development of sustainable technologies. This is underlined by the EC's call for research into the influence of ESG on innovation (EC, 2011).

In contrast to Europe, literature exists on the effects of changes of ESG ratings on firms' (green) innovative performance in Japan (Broadstock et al., 2020) and China (Liu & Lyu, 2022; Tang, 2022; Xu et al., 2020; Yin & Wang, 2018; Zhang et al., 2020; Zheng et al., 2022). These papers show a positive relationship between firms' ESG rating and their (green) innovative performance. There exists some literature which examines the effect of mandatory ESG disclosure on innovation that includes European firms in their sample, though not as a primary focus. For example, Ioannou and Serafeim (2017) find that firms with improved ESG ratings enjoy easier access to funding. Similarly, Gibbons (2020) indicates a positive effect between mandatory ESG disclosure policies and research & development (R&D) spending due to relaxation of capital rationing by investors. However, these papers examine the effect of the binary variable on whether policy is present on country-level (Gibbons, 2020; Ioannou & Serafeim, 2017). In other words, these papers only examine whether the introduction of a mandatory ESG disclosure policy in a country influences innovative performance. Consequently, the effects of changes in ESG ratings on green innovation at firm-level in Europe is not addressed. It is interesting to bridge the differences between financial literature and innovation literature by zooming in on effects of changes in ESG ratings on green innovation.

Noting the lack of focus on potential non-financial effects of ESG disclosure in finance and European innovation literature, other frameworks are consulted. Two heavily related theories to ESG disclosure are legitimacy and signalling theory (An et al., 2011; Cho & Patten, 2007; Deegan, 2002; Hasseldine et al., 2005; Hummel & Schlick, 2016; O'donovan, 2002). These theories postulate opposing reasons for why firms may, or may not engage in ESG disclosure. First, legitimacy theory posits that firms with activities conflicting with social values are likely to use ESG disclosure for legitimisation (O'donovan, 2002). Whereas, according to signalling theory, high-performing firms with regards to ESG are likely to use ESG disclosure to signal this otherwise inobservable quality (Connelly et al., 2011). Therefore, two counteracting effects occur where both good and bad performing firms engage in ESG disclosure.

This questions whether easier access to funding of firms with high ESG ratings results in the further development of sustainable technologies.

Even though Asian research suggests that the CSRD would have a positive effect on (green) innovative performance, this relation has not yet been assessed in Europe. For this reason, it is beneficial to determine whether these effects are also prevalent in the European institutional context. Moreover, the juxtaposition of legitimacy and signalling theory makes it interesting to determine the (relative) effect on both green and non-green innovative performance. The following research question culminates from this:

"How do changes in firms' ESG rating affect the green and non-green innovative performance of firms in the European Union and how do these effects relate to each-other?"

To answer this question, secondary ESG rating and patent data is collected, wherein ESG ratings represent the ESG performance, and patent counts represent innovative performance. This data is analyzed with negative binomial regression models and the Seemingly Unrelated Estimation (SUEST) method. First, the relationship between firms' ESG rating and the two dependent variables green and non-green innovative performance are modeled separately. A distinction between green and non-green innovation is made upon Y02 patent cooperative patent classification (CPC) classes, which represent *'technologies for adaptation to climate change'* (EPO, 2022a). Subsequently, the strengths of these effects are compared using the SUEST method to determine which effect is more prevalent. This analysis will help determine the effect that increased focus on ESG has on the sustainable development of technologies in the EU. This helps to discover the usefulness of the upcoming CSRD policy for sustainable development and helps bridge the gap between financial and innovation literature.

With this approach, this paper can make several empirical, methodological, theoretical and practical contributions. First, this paper extends the collective understanding of how ESG ratings influence green and non-green innovative performance in the European institutional context. Second, a methodological contribution is made by employing the SUEST method in innovation sciences. The SUEST method allows for meaningful comparison of coefficients over models with varying dependent variables. The adoption of this method in innovation sciences could enhance the methodological robustness of the field, because this method is rarely used in this field. Third, the comparison of the effect of ESG on green and non-green innovation with the SUEST method provides an empirical insight into the relative strength of these effects. An understanding of the relative strength of these effects helps shed a light on the presence of the legitimacy and signalling effect in this relationship. Fourth and final, this research provides policymakers with a more thorough understanding of how the CSRD and potential future policies can promote or inhibit innovation and sustainable development.

2
Conceptual and theoretical background

2.1. ESG

The Environmental Social Governance (ESG) framework dictates around which topics firms report on their non-financial performance (Deloitte, 2022). This non-financial performance of firms is centered around how firms incorporate the three pillars of environmental, social and governance into their business (Gillan et al., 2021). This allows investors to consider this performance in investment decisions and provides society with information on firms' sustainability performance (Matos, 2020). In practice, firms communicate about their ESG performance through ESG/Corporate Social Responsibility (CSR) reports or a chapter in their annual reports (Gillan et al., 2021). In ESG-related literature, reporting on ESG performance is also referred to as ESG disclosure (Fatemi et al., 2018), which can take the form of voluntary and mandatory disclosure due to policies that mandate ESG reporting (Ioannou & Serafeim, 2017). Typically, specialised ESG rating agencies assign ESG ratings to firms to indicate their ESG performance, which in turn are used by investors to assess ESG performance (Dorfleitner et al., 2015). However, different ESG rating agencies can assign different ESG ratings and there is to date no standardized method to calculate these ratings (Gibson Brandon et al., 2021).

The concept ESG was first introduced in public discourse in 2004 by the 'United Nation's Principles for Responsible Investment' (UNPRI) report (Gillan et al., 2021). This report gathered recommendations from the financial industry on incorporating ESG information in investing. This document provides a non-exhaustive outline of what can be considered ESG information under the three pillars (UNGC, 2004, p. 6), as shown in table 2.1. The financial institutions conclude that increased focus on ESG is key because firms that perform better on ESG are able to increase shareholder value and simultaneously contribute to sustainable societies (UNGC, 2004). In light of the reasoning behind the new European CSRD policy, it is important to consider what the effects of increased ESG activity are.

Notably, in existing literature the terms ESG and CSR disclosure/reporting are often used interchangeably with minor differences in their interpretation (Gillan et al., 2021; Tang, 2022). ESG is posed as an extension of CSR by including governance more explicitly, whereas, CSR includes governance indirectly through its relation to environmental and social topics at hand (Gillan et al., 2021). For this reason, this research regards the effects that are found in CSR literature to also extend to ESG disclosure. Wherein, ESG disclosure shows the same effects on innovative and sustainability performance as CSR disclosure. In this paper, the term ESG is used in these situations to prevent ambiguity. Moreover, in both the literature on ESG disclosure and CSR disclosure, the research gap into the distinction between the effect of ESG on green and non-green innovative performance in Europe exists.

Table 2.1: Description of ESG information (UNGC, 2004)

Pillar	Environmental	Social	Governance
	Climate change and related risks	Workplace health and safety	Board structure and accountability
	Reduction of toxic releases and waste	Community relations	Accounting and disclosure practices
	New regulation expanding boundaries of environmental liability	Human rights at company and partners	Audit committee structure and independence of auditors
	Pressure by civil society to improve performance, transparency and accountability	Pressure by civil society to improve performance, transparency and accountability	Executive compensation
	Emerging markets for environmental products and services	Government and community relations while operating in developing countries	Management of corruption and bribery issues

In finance literature, ESG has been increasingly prevalent (Gillan et al., 2021). This body of research has primarily focused on incorporating ESG in assessing corporate financial performance (Friede et al., 2015). Friede et al. (2015) found that only 10% of finance literature focused on the relationship between ESG and financial performance shows a negative relationship. Moreover, increased ESG disclosure has been shown to result in reduced capital rationing by investors (Gibbons, 2020), decreased cost of capital (El Ghoul et al., 2011), increased access to capital (Cheng et al., 2014; García-Sánchez et al., 2019), improved analysts' earnings forecasts (Krueger et al., 2021), and reduced investor risk (Lopez-de-Silanes et al., 2020).

Clearly, this increased focus on ESG in financial markets does not provide insight into the potential non-financial benefits of increased ESG activity. This observation touches upon the main voiced critique on ESG disclosure. Namely, that investors merely use ESG disclosed information to avoid their exposure to risk and maximize profits (Matos, 2020). This is further underlined by Amel-Zadeh and Serafeim (2018) who conclude that investment professionals use ESG information primarily for financial over ethical motives. Moreover, critics voice that many of the best profit-driven social impact firms do not achieve good ESG ratings despite their positive non-financial impacts (Porter et al., 2019). This might be due to ESG green washing, wherein firms disclose large amounts of ESG data even though they perform poorly in ESG aspects (Yu et al., 2020). This is possible because ESG reports are not audited and therefore this information is not always reliable.

These voiced critiques point towards the question of what the added value of ESG is for the sustainable development of society. In this regard, it should be challenged whether these firms engage with ESG out of their social conscience or whether they merely do so to increase profits. Nonetheless, this does not have to exclude societal benefits from being a result of engaging in ESG.

This dive into financial literature shows a discordance between topics in financial literature and the explicit non-financial goals stated in the UNPRI report and CSRD policy. Therefore, it is important to consult other literature strands to discover their conclusions on the non-financial effects of ESG disclosure.

2.2. Legitimacy and signalling theory

Two theories that are heavily related to ESG disclosure are legitimacy theory and signalling theory (An et al., 2011; Cho & Patten, 2007; Deegan, 2002; Hasseldine et al., 2005; Hummel & Schlick, 2016; O'donovan, 2002). These theories postulate opposing reasoning for why firms would or would not engage in ESG disclosure. An understanding of the motivations for firms to (not) engage in ESG disclosure might help elucidate whether ESG can stimulate sustainable development.

On the one hand, legitimacy theory builds upon the concept of organisational legitimacy, which pertains to an organisation's alignment with the value system of society (Dowling & Pfeffer, 1975). In the case of disparity between the value system of the organisation and the social system, a threat to organisational legitimacy might occur (Dowling & Pfeffer, 1975). As a response to these threats to organisational legitimacy, legitimisation strategies are employed by firms (Suchman, 1995). Based on this line of reasoning, legitimacy theory posits that the greater the likelihood is that a firms' activities are in conflict with social values, the more likely they are to engage in legitimisation strategies (O'donovan, 2002). Thus, from this point of view, firms that perform poorly on ESG would be more likely to increase their ESG disclosure as a form of legitimisation (O'donovan, 2002).

On the other hand, signalling theory builds upon the concept of information asymmetries that occur between holders and non-holders of information because of differences in public and private knowledge (Stiglitz, 2002). In the case of information asymmetries about latent or opaque qualities, it can be especially difficult for high-performing firms to receive payoff for their high quality (Connelly et al., 2011). In these situations, it is beneficial for high-performing firms to signal their quality. This is often realised by utilising reporting to signal their otherwise in-observable qualities (Toms, 2002). Consequently, according to signalling theory, a firm that performs well on ESG would be more likely to increase their ESG disclosure to distinguish itself (Cho & Patten, 2007; Yin & Wang, 2018).

The juxtaposition of these theories does not result in unanimous *a priori* assumptions about the contribution of ESG disclosure to non-financial goals. According to signalling theory, firms with good ESG performance use ESG disclosure to distinguish themselves. Whereas, legitimacy theory presumes firms with bad ESG performance use ESG disclosure as a legitimisation strategy. This points to two counteracting effects that occur wherein both good and bad performing firms increase their ESG disclosure to either promote or mask their performance on ESG aspects.

In similar fashion, an empirical inquiry by Hummel and Schlick (2016) contrasts legitimacy theory and voluntary disclosure theory to analyze the relation between disclosure and sustainability performance. This paper concludes that both effects are present and that poorly performing firms disclose low-quality sustainability information as a legitimisation strategy and well performing firms disclose high-quality sustainability information to signal their market value (Hummel & Schlick, 2016). This shows that both well and poor performing firms engage in disclosure strategies. This is problematic because ESG ratings are calculated based both on the amount and quality of disclosure (Dorfleitner et al., 2015; Yu et al., 2020). Due to this calculation method and the lack of audit of the disclosed information, both types of disclosure can result in improved ESG ratings.

Thus, even though increased ESG disclosure leads to increased access to capital (Cheng et al., 2014; García-Sánchez et al., 2019) at lower cost (El Ghoul et al., 2011), it is not clear whether these increased funds end up at well-performing firms. This again underlines the question whether increased funds to firms with high ESG ratings would result in more green innovation.

2.3. ESG and Innovation

Due to the inconclusiveness of scientific discourse on non-financial effects of ESG, it is important to consider how to determine these non-financial effects. The non-financial nature of these effects render it difficult to adequately measure them (Aupperle et al., 1985). To tackle this issue, topics of sustainable development and sustainable performance are often measured through proxies representing underlying mechanisms (Parris & Kates, 2003). Even though, ESG ratings are regarded as a proxy for the sustainable performance of a firm (Herbohn et al., 2014), existing literature does not come to conclusive answers on the reasoning for increased disclosure. In the technology and innovation management literature, a common proxy for sustainability performance is the green innovative performance of a firm (Schiederig et al., 2012). It has been repeatedly researched how (green) innovation has positive effects on sustainable performance (Adams et al., 2016; Sezen & Cankaya, 2013) on both environmental, economic and social dimensions (Boons & Lüdeke-Freund, 2013; Kuzma et al., 2020). Boons et al. (2013) even argue that innovation is a prerequisite to reach sustainability goals. Therefore, assessing firms' green innovative performance helps to determine non-financial effects of ESG. The following paragraphs provide theoretical grounding for how ESG is expected to influence (green) innovation, which culminates in a set of hypotheses about this relation.

As mentioned before, increased ESG disclosure has been shown to result in easier access to capital (Cheng et al., 2014; García-Sánchez et al., 2019) at lower costs (El Ghoul et al., 2011). According to both Cheng et al. (2014) and García-Sánchez et al. (2019) this effect occurs because firms with better ESG performance show increased stakeholder engagement, which decreases short-term opportunistic behavior and loosens financial constraints. Whereas, the lower cost originates from a decrease in information asymmetry that leads to a decrease in the firm's perceived risk by investors (El Ghoul et al., 2011). Furthermore, Gibbons (2020) shows that due to a reduction in capital rationing by investors, firms divert more resources to R&D spending. Unfortunately, this research examines the effects of the binary variable of whether a policy is present in a country and in doing so fails to examine the effects of changes in ESG rating on innovative performance (Gibbons, 2020). Moreover, it is expected that improved ESG ratings would attract stakeholders that are especially interested in ESG as driver of business value (Lyon & Maxwell, 2011). Consequently, the attraction of these stakeholder would result in firms being more inclined to make sustainable investments (Herremans et al., 1993). Due to the stakeholder pressures towards sustainable investments regardless of firms' current actual ESG performance, it can be argued that improved ESG ratings will have a positive effect on R&D spending for both good and bad performing firms. Therefore, it is postulated that the ESG rating of a firm has a positive relation to R&D activity and in turn this might lead to increased innovative performance.

This theoretical deduction is further underlined by existing empirical research in Japan (Broadstock et al., 2020) and China (Liu & Lyu, 2022; Tang, 2022; Xu et al., 2020; Yin & Wang, 2018; Zhang et al., 2020; Zheng et al., 2022) that shows a positive relationship between ESG disclosure and green innovative performance. For example, Broadstock et al. (2020) states that increased ESG ratings result in an extension of a firms' technological production frontier, which leads to changes in innovation processes and increases R&D spending. Moreover, Liu and Lyu (2022) and Yin and Wang (2018) show that improved ESG ratings result in decreased capital constraints and enhanced legitimacy of green innovation investments amongst stakeholders. These developments result in increased green patenting. In similar fashion, Tang (2022) argue that improved ESG ratings result in decreased financing constraints and make managers focus more on long-term firm development. This effect has shown an increase in innovation output. Finally, Xu et al. (2020) indicate that improved ESG ratings result in increased investor attractiveness and employee satisfaction that resulted in increased green innovation. Moreover, this paper shows that ESG performance has a moderating effect on the influence of R&D on green innovation because it relieves stakeholder concerns on green innovation investments.

Thus, improved ESG ratings can not only lead to more innovative inputs in the form of R&D spending, but also result in increased innovative outputs. On the other hand, the firms in these samples operate in a distinctive institutional environment from the European institutional context under which the CSRD policy operates. This renders it important to determine whether this effect can also be observed in Europe. Signalling theory posits that firms with high ESG performance are more likely to disclose ESG information to reap benefits of these latent and in-observable qualities. Therefore, if the financial benefits of an improved ESG rating indeed lead to increased R&D spending as proposed by Broadstock et al. (2020) and Gibbons (2020). Improved ESG ratings would indeed lead to improved green innovative performance, because the well performing firms are expected to continue to invest this money in green R&D projects.

It follows that the *a priori* expectation of this research is that changes in ESG rating have a positive relationship with firms' green innovative performance. The following hypothesis is formulated based upon these expectations.

Hypothesis (H1): *For firms operating under EU legislation, a firms' ESG rating is positively related to a firms' green innovative performance.*

The conundrum that originated from the juxtaposition of legitimacy and signalling theory makes it interesting to also look into 'non-green' innovation. This could help provide understanding on whether firms use ESG as a signalling or legitimisation strategy. Legitimacy theory states that firms with low ESG performance are more likely to disclose ESG information to legitimize their non-green activities. Therefore, if the financial benefits of an improved ESG rating indeed lead to increased R&D spending in poor performing firms, improved ESG ratings would not lead to improved green innovative performance, but instead lead to more non-green innovative performance.

Therefore, a positive relation between a firms' ESG rating and their non-green innovative performance is expected. This is further underlined as improved ESG ratings increase the focus on long-term behavior more generally (Cheng et al., 2014; García-Sánchez et al., 2019). The second hypothesis is as follows:

Hypothesis (H2): *For firms operating under EU legislation, a firms' ESG rating is positively related to a firms' non-green innovative performance.*

If both hypotheses are not rejected, it suggests that both expected effects from signalling and legitimacy theory occur at the same time. Still, this research expects a stronger relation to be present between a firms' ESG rating and their respective green innovation performance as opposed to non-green innovation performance. It is expected that increased ESG performance not only results in increased stakeholder engagement and more long-term orientation (Cheng et al., 2014; García-Sánchez et al., 2019), but this engagement to be explicitly aimed at green innovation (Liu & Lyu, 2022; Yin & Wang, 2018). As a result, the relation to green innovation is expected to be stronger. If this is proven, it shows that although ESG is used as both legitimisation and signalling strategy, ESG results in relatively more green innovation than non-green innovation. The following hypothesis originates from this:

Hypothesis (H3): *For firms operating under EU legislation, a firms' ESG rating is more positively related to a firms' green innovative performance than to a firms' non-green innovative performance.*

3
Method

3.1. Research design

This research attempts to determine the relationship between firms' ESG ratings and their innovative performance. Consequently, a firm-level analysis is warranted. This research is confirmatory, since it attempts to test *a priori* formulated hypotheses to determine the relations between the (in)dependent variables (Jaeger & Halliday, 1998). Considering this confirmatory nature and interest in changes over time, a quantitative longitudinal research design is employed. Secondary data sources are used to source the relevant data. This research design makes it difficult to make causal claims about how X causes Y because *'correlation does not imply causation'* (Kornbrot, 2005, p. 1). However, methods are available that can help to achieve causal inference from observational data without intervening (Rohrer, 2018). This research employs a directed acyclic graph (DAG) approach to causal inference. The DAG approach attempts to create a systematic representation of the effects in the structural causal model (SCM) (Textor et al., 2016). Using a set of rules, also known as d-separation rules, the SCM informs decisions about what variables need to be controlled for to estimate causal effects (Textor et al., 2016). A more detailed explanation of this approach is found in appendix A.

3.2. Sample

Given the interest in CSRD policy in Europe the sample is a subset of European firms. Firms listed on primary market indices are selected as these are required to report on ESG information, which reduces selection bias (Di Simone et al., 2022). The NFRD policy has resulted in a requirement for large firms that are listed on European regulated markets to disclose ESG information from 2014 onward (EC, 2014). For this reason, the sample is demarcated to firms that are part of the STOXX600 index, representing the 600 largest European companies. For several reasons[1] the **final sample is reduced to 364 companies** (see footnote and 3.3.4).

The selected time-period of analysis is from 2014 till 2017. The reason for this is that from 2014

[1]Under NFRD a firm is considered large if it has more than 500 employees (EC, 2014) and thus some of the STOXX600 firms are not required to disclose ESG information. Furthermore, counterintuitively, the STOXX600 consists of 570 unique companies due to several companies being listed on multiple exchanges. In addition, this index spreads across European countries, whereas this research is interested in companies that operate under EU legislation. Thus, for some edge cases it is considered whether they are included in the sample. First of all, some firms in the STOXX600 are listed in Norway, which is not part of the European Union. However, the Norwegian stock exchange is considered an EU regulated market (ESMA, 2022). Due to this, Norwegian listed firms fall under the NFRD policy (EC, 2014) and are thus included in the sample. In contrast, Swiss exchanges are are not considered EU regulated markets (ESMA, 2022) and thus Swiss firms that are only listed on Swiss exchanges are not considered in the sample. Moreover, a subset of the firms that are part of the STOXX600 index are firms that are based in the United-Kingdom (UK). During 2020, the departure of the the UK from the European Union was finalised and for this reason from 2021 onward the UK does fall under the EU institutional framework (Rijksoverheid, 2022). However, noting the selected time-period of analysis, these UK listed firms are included in the sample. The reason for this is that these firms operated under the EU institutional framework during the selected time-period. Regrettably, not all companies have been part of this index for the entire sample period, which further reduces the available sample size. Furthermore, some company names returned so many OECD HAN application names that is was infeasible to check which HAN names represented the actual company. Finally, some companies had too many missing ESG ratings in the selected time-period

onward the NFRD policy came into effect (EC, 2014), which greatly increases data availability. Even though, the legal delay for publishing is 18 months after application in most patent offices (OECD, 2004), figure 3.1 shows a steep decline in average number of green and non-green patents after 2018. This discrepancy seems to be explained by the 36 month publishing target of the EPO (EPO, 2022b). Moreover, this target is only achieved on 80% of patents, which further increases the patent backlog (EPO, 2022b). Additionally, as described in chapter 3.5, a one year time lag is introduced to the effect of ESG on innovative performance, which reduces the end of the sample period to 2017.

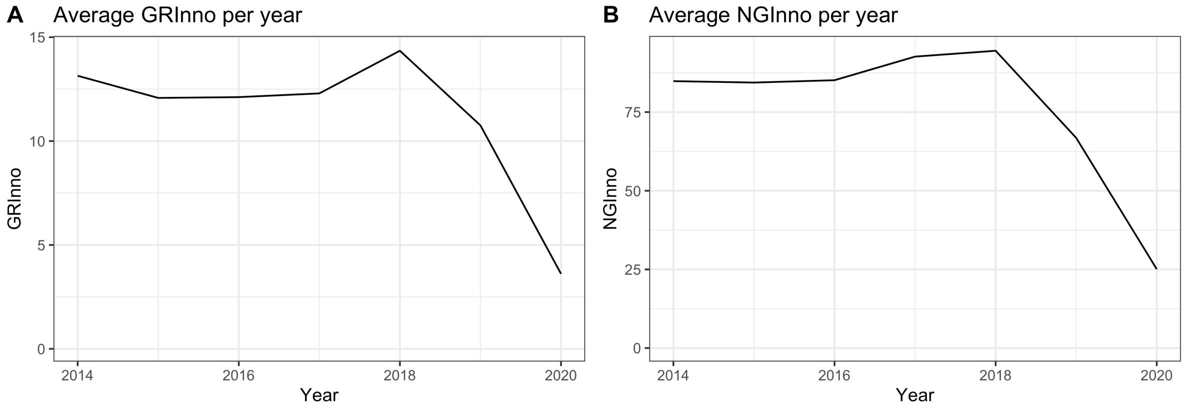

Figure 3.1: Average patents per year for green patents (A) and non-green patents (B)

3.3. Operationalization of variables and data collection

3.3.1. ESG

This research aims to determine the relationship between ESG ratings and innovative performance. These ESG ratings are calculated by ESG rating agencies that focus on translating ESG-related information that firms disclose into ratings (Berg et al., 2022). ESG rating agencies use different rating methodologies, which results in differing ESG ratings for the same firm (Li & Polychronopoulos, 2020). This is problematic as it hinders the generalizability of research performed on different datasets. Commonly, the ESG ratings are calculated based upon various underlying indicators that together determine the resulting ESG rating (Berg et al., 2022).

The only available license [2] at Utrecht University that provides access to an ESG dataset is Factset. Factset provides access to overall ratings of the FTSE Russel ESG rating dataset and a large set of financial indicators. The latter is potentially useful to gather data on control variables. The ESG dataset contains ratings on 7200 publicly listed firms and is calculated based upon a set of over 300 indicators (FTSE-Russell, 2022a). These 300 indicators are grouped into 14 themes that each fall under one of the three ESG pillars and are shown in figure 3.2. To calculate ratings, FTSE Russell solely uses publicly disclosed information (FTSE-Russell, 2022a) aligning well with this research.

[2]The market for ESG rating data is highly privatized and monetized by the ESG rating agencies. As a result, most of the ESG rating-related data is not publicly available and is expensive to gain access to. For this reason, researchers are reliant on the licenses that are available to them to collect ESG rating data.

Figure 3.2: FTSE Russel ESG rating themes (FTSE-Russell, 2022a)

In line with the sample time period, the ESG dataset provides historical ESG ratings from 2014 onwards. To access this data, a Factset Excel plug-in is used to browse the dataset based upon company names.

Concretely, the historical ESG ratings per year of a firm are used as an indicator of the ESG performance of a firm. The resulting variable name is ESG.

3.3.2. Innovative performance

In innovation literature, several metrics of innovative performance exist. The most commonly used measurements are *"R&D inputs, patent counts, patent citation, or counts of new product announcements"* or a composite of multiple indicators (Hagedoorn & Cloodt, 2003, p. 1). These measurements have their respective (dis)advantages (Archibugi, 1992). Hagedoorn and Cloodt (2003) argue that due to the statistical overlap between these indicators, any of these indicators can be used to measure innovative performance. Additionally, the ideal indicator is contingent on the research context (Hagedoorn & Cloodt, 2003).

This research investigates the relationship between ESG performance and both green and non-green innovation. Therefore, the indicator should allow for distinction between green and non-green innovation. This distinction is impossible for R&D spending due to the lack of availability of (granular) R&D data[3]. Opposingly, for patents a distinction can be made between green and non-green patents. This distinction is made in the Cooperative Patent Classification (CPC) with category Y02 (Veefkind

[3]To check this assumption the annual reports of 2021 of the constituents of the Dutch large-cap index (AEX25) are analyzed by the researcher. The occurrences and context of a set of terms are analyzed to determine the availability of information. These terms are as follows: R&D, research, development, investment, expenditure, spending and innovation. This analysis shows that only two firms (Philips and Signify) of the 25 companies that are part of this index communicate their R&D at the required granularity. Whereas, eight firms merely communicate overall R&D spending (ArcelorMittal, ASM International, ASML, BE Semiconductor, DSM, Shell, Unilever and Wolters Kluwers) and the remaining 15 firms did not report any information on R&D spending (Adyen, Aegon, Ahold Delhaize, Akzo Nobel, Heineken, IMCD, ING, Just Eat Takeaway, KPN, NN group, Prosus, Randstad, RELX, Unibail-Rodamco-Westfield and UMG). This shows that R&D spending information is not widely available at the green and non-green granularity and accentuates the general difficulty of incorporating R&D spending information in innovation research at firm level.

et al., 2012) containing *'technologies for adaptation to climate change'* (EPO, 2022a). Patents provide firms with exclusive rights for the commercial exploitation of an invention for a set time and thus grants a legal monopoly to a firm (Mansfield, 1986). Using patents as indicator of innovative performance has several (dis)advantages. Patents often represent outputs of the inventive process that are likely to result in applications due to the high cost of acquiring patents (Archibugi, 1992). However, not all inventions are (eligible to be) patented and differences exist in patenting propensity between industries and over time (Archibugi, 1992; Mansfield, 1986). Hence, even though patents are able to capture a lot of information on innovative performance, the effectiveness of this indicator is not guaranteed. Nevertheless, patents are useful for this research and annual patent count on firm level is used as an indicator of a firms' innovative performance. Patents classified as Y02 are regarded as green innovation, whereas a firms' remaining patents are regarded as non-green innovation.

To gather patent data, secondary datasets of OECD are used. Specifically, OECD-Regpat is used as overall patent dataset and OECD-HAN is used to link patents to firms (Maraut et al., 2008). The OECD-HAN data is useful for firm-level analysis since patents are pre-matched to company names (OECD, 2022). Unfortunately, the quality of this pre-matching is not high and the researcher is required to check long lists of potential company names per company. This is mainly due to the high number of misspellings still present in the data and the difficulty of multiple companies operating under the same acronyms or similar names. Next to the company linkages found in OECD-HAN, OECD Regpat contains other useful information. First, OECD-Regpat includes the priority date, which represents the filing date of the patent and is commonly used in innovation research (Dernis & Khan, 2004). Second, OECD-Regpat provides CPC information to distinguish between green and non-green innovation (OECD, 2022).

The patents of a firm are counted per year and grouped as either green or non-green. The resulting operationalizations are GRInno and NGInno, where GRInno represents green innovation performance and NGInno represents non-green innovation performance.

3.3.3. Control variables

This section describes the relevant variables and the assumptions about their effects on other variables in the structural causal model (SCM). These assumptions are combined in a DAG to determine the valid adjustment set (set of variables that need to be controlled for (Pearl, 1988)). In describing the SCM, no distinction is made between green and non-green innovative performance since it is assumed that both effects originate from the same SCM.

R&D spending

The ESG rating of a firm has a positive effect on R&D spending (Broadstock et al., 2020; Xu et al., 2020). In turn, increased R&D spending positively influences innovative performance (Cohen & Levinthal, 1989). Therefore, some of the effect that ESG has on innovative performance is an indirect effect through the effect of ESG on R&D spending. Hence, controlling for R&D spending would result in the blocking of the transmission of this causal information and is undesirable. Moreover, for the majority of companies no R&D data is publicly available. Consequently, the usage of R&D as a control variable is unwarranted and inhibited.

Revenue

ESG disclosure is a costly practice and firms with high revenues are able to more easily engage in ESG disclosure (García-Sánchez et al., 2019). Larger firms are shown to have higher ESG ratings (Drempetic et al., 2020) and larger firms with more revenue have a higher propensity to patent their innovations (Arundel & Kabla, 1998; Holgersson, 2013). Subsequently, revenue has a confounding effect on ESG and innovative performance and has to be included as a control variable.

Additionally, it has been shown that increased revenues do not result in an increase in R&D spending (Morbey, 1988). For this reason, no statistical path exists between revenue and R&D spending.

Year

Patenting propensities of firms change over time with varying trends in different industries (Bessen & Hunt, 2007; Hall & Ziedonis, 2001). It is expected that as time progresses firms improve their ESG performance as this is incentivized by financial markets (Ioannou & Serafeim, 2017). In this regard, time has a confounding effect on this relationship and is required as a control variable.

However, year also has an effect on the revenue of a firm as this revenue varies over time. It is important to consider whether under these circumstances year and revenue should still be included as a control variables. This is determined upon the completion of the final DAG.

Industry

A firms' industry influences its patenting propensity (Mansfield, 1986). Simultaneously, a firms' industry has a large effect on the ESG rating of a firm, due to the nature of their activities. Therefore, industry has a confounding effect on ESG and innovative performance and is required as a control variable.

Directed Acyclic Graph (DAG)

The resulting DAG is shown in figure 3.3. As becomes apparent, the valid adjustment set consists of Industry, Revenue and Year. These variables are used as the control variables. Opposingly, R&D is not used as a control variable because R&D can be considered as a mediator between ESG and Inno and controlling would block the transmission of causal information of the effect of ESG on Inno.

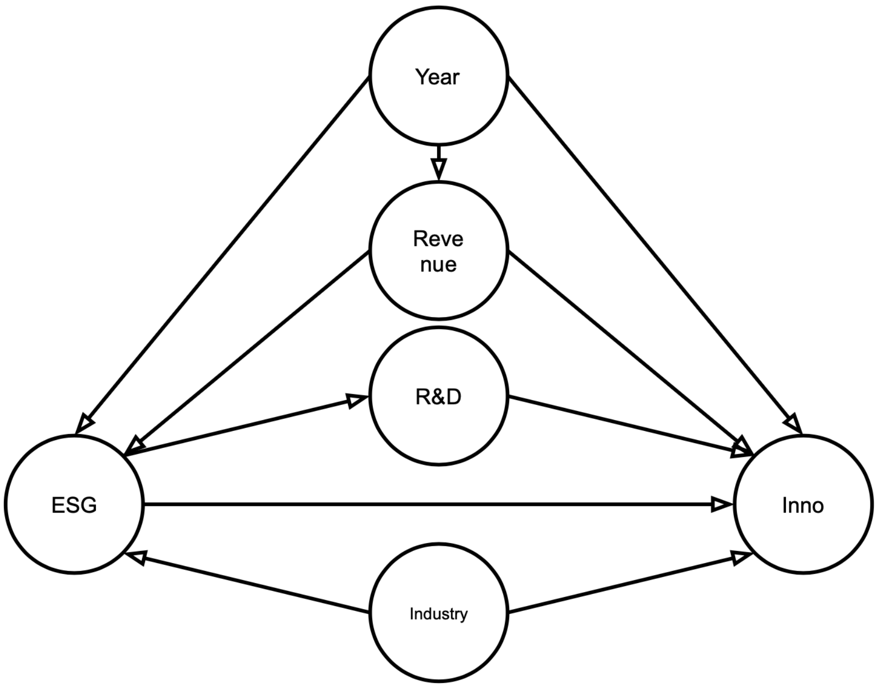

Figure 3.3: Directed Acyclic Graph that represents the Structural Causal Model of this research

The resulting operationalization and data sources are shown in table 3.1.

Table 3.1: Operationalization of variables

Variable	Type	Operationalization	Category	Measure	Source
GRInno	Dependent	Number of green patents (Y02) of firm i at year t	Discrete	Count	OECD Regpat and OECD HAN
NGInno	Dependent	Number of non-green patents (non-Y02) of firm i at year t	Discrete	Count	OECD Regpat and OECD HAN
ESG	Independent	Overall ESG rating of firm i at year t	Continuous	Float	FTSE Russel ESG Ratings (Factset)
Revenue	Control	Sales of firm i at year t	Continuous	Millions of Euros	Factset
Year	Control	Year of measurement (for patents: Priority year)	Categorical	Category	Factset
Industry	Control	Main industry category of firm i	Categorical	Category	Factset

3.3.4. Missingness and data imputation

After data collection for ESG and Revenue data, a large part of the company sample shows missing data in some years for ESG and Revenue data. This is especially evident for ESG data with a maximum of 219 missing values (out of 547 companies) in 2014. Due to the relatively small number of companies in the sample and the short time period of analysis it is paramount to maintain as much data as possible. Therefore, it is analysed whether any discernable patterns in the missingsness are observed and afterwards data imputation methods are considered.

In the sample, 66 companies have no available ESG data. Consequently data imputation is not useful for these companies and these rows are dropped. Nevertheless, the missingness of this data is analysed and no discernable pattern is found in terms of country of origin. Moreover, both Revenue and employee counts are determinants of whether a company falls under the NFRD. No pattern is found because these companies have revenues and employee counts (not part of dataset, but manually retrieved from Google) above the NFRD threshold.

In the remaining 481 companies, large differences exist in number of years of missing data. To prevent unreasonable amounts of data imputation companies with more than 2 missing values for ESG in the selected time period are dropped. This reduces the sample to 370 companies and as a result there are no more missing values for Revenue. Further analysis of the ESG missingness shows no clear patterns in terms of revenue, employee count or country of origin. Due to the lack of a discernable pattern, it is not appropriate to assign the missingness problem as either Missing At Random (MAR) or Missing Completely At Random (MCAR) (Rubin, 1976). This is because the data is missing due to reasons that are unknown to the researcher and therefore the missingness problem is Missing Not At Random (MNAR) (Van Buuren, 2018). Unfortunately, this means that any form of data imputation will introduce some form of bias (Van Buuren, 2018).

Be that as it may, to prevent further reduction of the sample size, data imputation is employed. Specifically, back fill data imputation is applied for the following two reasons. First of all, due to the MNAR nature of the missingness problem, any method will introduce bias (Van Buuren, 2018) and therefore a simple method is desirable. Second, upon observation of the ESG data, companies often have the same ESG rating in consecutive years, which makes it more reasonable to apply backfill. The final sample consists of 364 companies.

3.4. Reclassification of industries

To reflect on whether the selected sample is representative of the STOXX600, table 3.2 and 3.3 show descriptive statistics about *Country* and *Industry*. In these tables, the *_raw* suffix represents the unfiltered dataset. Table 3.2 shows that for most countries no large differences exist in the percentage of companies in the sample and the original dataset. However, Great Britain already represented 27% of companies, which increased to 33% and the percentage of Swedish companies is reduced from 12% to 6%. The reduced sample of Swedish companies is mainly due to the high number of Swedish companies without any *ESG* data (22). This lack of ESG data is attributed to the relatively low *Revenue* of these companies, which indicates that these firms are smaller and do not fall under the NFRD.

In similar fashion, the *Industry* descriptive statistics shown in table 3.3 highlight that representativeness of the sample. The Technology sector is underrepresented with a 50% reduction from 6% to 3%. On the other hand, the *Industry* variable helps to reduce the impact of this imbalance for the other industries.

Table 3.2: Count, percentage of total and percentage difference for *Country* in the filtered and raw dataset

Country	count	percent	count_raw	percent_raw	percent_diff
Austria	5	0.01	8	0.01	0
Belgium	10	0.03	17	0.03	0
Denmark	14	0.04	23	0.04	0
Finland	13	0.04	18	0.03	0.01
France	59	0.16	75	0.14	0.02
Germany	46	0.13	69	0.13	0
Great Britain	119	0.33	146	0.27	0.06
Ireland	6	0.02	8	0.01	0.01
Italy	19	0.05	33	0.06	−0.01
Luxembourg	2	0.01	2	0	0.01
Netherlands	13	0.04	27	0.05	−0.01
Norway	7	0.02	17	0.03	−0.01
Poland	6	0.02	9	0.02	0
Portugal	3	0.01	4	0.01	0
Spain	21	0.06	24	0.04	0.02
Sweden	21	0.06	67	0.12	−0.06

Table 3.3: Count, percentage of total and percentage difference for *Industry* in the filtered and raw dataset

Industry	count	percent	count_raw	percent_raw	percent_diff
Automobiles and Parts	9	0.02	14	0.03	−0.01
Banks	34	0.09	39	0.07	0.02
Basic Resources	12	0.03	18	0.03	0
Chemicals	17	0.05	19	0.03	0.02
Construction and Materials	11	0.03	20	0.04	−0.01
Consumer Products and Services	19	0.05	34	0.06	−0.01
Energy	14	0.04	18	0.03	0.01
Financial Services	17	0.05	32	0.06	−0.01
Food, Beverage and Tobacco	18	0.05	26	0.05	0
Health Care	26	0.07	47	0.09	−0.02
Industrial Goods and Services	61	0.17	90	0.16	0.01
Insurance	22	0.06	25	0.05	0.01
Media	9	0.02	14	0.03	−0.01
Personal Care, Drug and Grocery Stores	11	0.03	15	0.03	0
Real Estate	18	0.05	32	0.06	−0.01
Retail	8	0.02	10	0.02	0
Technology	12	0.03	31	0.06	−0.03
Telecommunications	15	0.04	20	0.04	0
Travel and Leisure	10	0.03	15	0.03	0
Utilities	21	0.06	28	0.05	0.01

In comparing industries, large differences exist in patenting propensities (Mansfield, 1986). This difference is especially evident between manufacturing and service industries (Morikawa, 2019). According to López and Roberts (2002), this is due to the weak appropriability regimes present in service industries. As shown in table 3.3, a large percentage of the sample (45%) operates in a service based industry. Therefore, the dataset appears skewed towards low patenting propensity industries and additional scrutiny is warranted. To deal with this, the industries are reclassified into low, medium and high propensity mainly according to work by Arundel and Kabla (1998) and other prior research (Boscaljon et al., 2006; Hunt, 2010; Kortum & Putnam, 1997; La Belle & Schooner, 2013; López & Roberts, 2002). The counts per propensity category are shown in table 3.4 and below the new classification of industries are found in table 3.5. A more detailed explanation of the reclassification and the methodological limitations of this reclassification is found in appendix B.

As becomes apparent from table 3.4, the distribution between propensity categories is not highly skewed with similar counts in the low and high category. However, the medium category is slightly underrepresented. This is in contrast with the expectation that the distribution would be skewed towards low propensity industries. Nonetheless, this new categorization provides a meaningful improvement in terms of interpretability and is used in the regressions to denote the patenting propensities of industries.

Table 3.4: Propensity category counts and percentages

Propensity category	count	percent
Low	137	0.38
Medium	88	0.24
High	139	0.38

Table 3.5: Industry reclassification to propensity categories

Industry	Propensity category
Automobiles and Parts	Medium
Banks	Low
Basic Resources	High
Chemicals	High
Construction and Materials	High
Consumer Products and Services	Low
Energy	Medium
Financial Services	Low
Food, Beverage and Tobacco	Medium
Health Care	High
Industrial Goods and Services	High
Insurance	Low
Media	Low
Personal Care, Drug and Grocery Stores	Medium
Real Estate	Low
Retail	Low
Technology	High
Telecommunications	Medium
Travel and Leisure	Low
Utilities	Medium

3.5. Expected time lag effect of ESG on innovative performance

To estimate a cause and effect relationship, a treatment should predate the effect that is estimated (i.e. taking an aspirin can not negate a headache in retrospect). This basic assumption is magnified in research with patent counts as the dependent variable, because a patent is an output of the innovative process which takes time. It is key to consider the temporal nature of how a treatment affects patenting. Existing research indicates that a time-lag of up to 3 years for the relationship between a treatment and patenting is reasonable (Hall et al., 1984). Therefore, it is expected that the optimal time lag for the effect of ESG on innovative performance is between 0 and 3 years. On the other hand, Hall et al. (1984) also argues that most of the effect occurs in the first year divided over a lag of *t+0* and *t+1*.

To determine the optimal time lag, multiple negative binomial models are fitted with time lagged patents varying from 0 to 3 years. These models are evaluated according to their variable significance and AIC. Due to the difficulty of estimating relative overall significance of categorical variables, categorical variables are excluded. Moreover, the intercept and Year variable are excluded from the plots as these are not significant, which hinders the interpretability.

As shown in figure 3.4, the variable significances for both the GRInno and NGInno model worsen as the time lag is increased. This effect becomes especially evident at a time lag of 2 years for GRInno and 3 years for NGInno. This shows that the best fit is achieved without introducing a time lag to either GRInno and NGInno. Opposingly, the observed AIC described in the left hand side table of 3.6, which decreases upon increasing the time lag, indicates that more of the variation is explained by a higher time lag. However, as the time lag increases less data remains and this reduces the amount of variance. When controlling for this by filtering the data to the minimum amount for all fitted models this effect is greatly reduced as shown on the right hand side of table 3.6. Taking both the variable significance and AIC values into account, it becomes evident that the optimal lag for this research is either 0 or 1. This is in line with the argument of Hall et al. (1984), who states that most of the effect occurs in the first year at *t+0* and *t+1*.

Apart from what is observed in the data, it is reasonable to assume that there is at least some lag between efforts towards innovation and innovative output. Moreover, adding a lag makes theoretical sense to prevent unwarranted reverse causation effects wherein *ESG* might influence *GRInno* and *NGInno* (Zapf et al., 1996). This is dealt with since the time lag guarantees that the cause predates the effect. For these reasons, this research employs a time lag for GRInno and NGInno of 1 year.

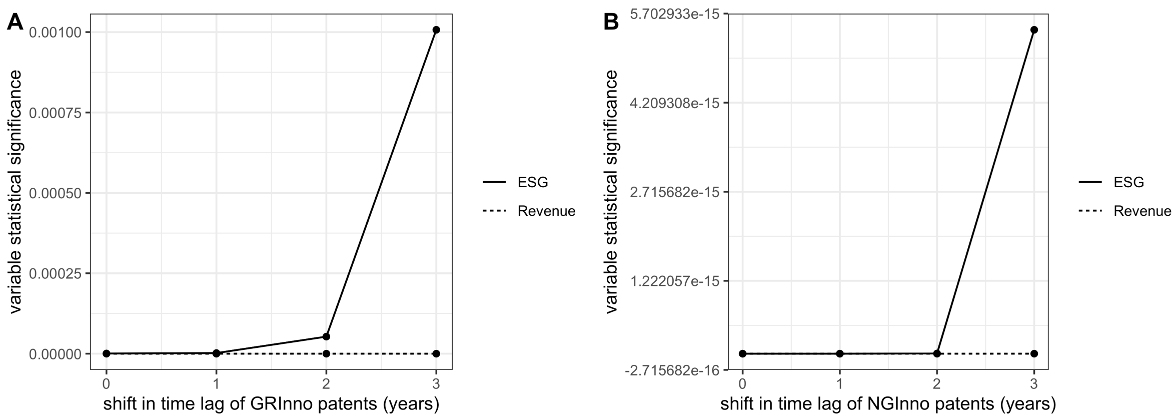

Figure 3.4: Time lag plots for green patents (A) and non-green patents (B)

Table 3.6: AIC of time lagged GRInno and NGInno models for varying data (left) and constant data (right)

lag	AIC GRInno	AIC NGInno	AIC GRInno constant data	AIC NGInno constant data
0	6, 715.345	11, 942.210	2, 717.214	4, 714.358
1	5, 361.882	9, 570.322	2, 682.566	4, 712.918
2	4, 004.668	7, 226.532	2, 623.034	4, 772.609
3	2, 685.350	4, 853.802	2, 685.350	4, 853.802

3.6. Analysis

This research is interested in the effects of the independent variables on two separate dependent variables (H1 and H2) and how these effects relate (H3). Therefore, two separate regression models are created and afterwards the SUEST method is applied to compare effect strengths on multiple dependent variables (Weesie et al., 2000).

3.6.1. Regression model selection

To apply SUEST both separate regression models need to be of the same type (Mize et al., 2019). Both dependent variables are count data and can only take non-negative discrete values (i.e. a firm can not have negative patents). With count data, the data follows either a (quasi-)Poisson or negative binomial distribution (Hausman et al., 1984). Determining the exact type of distribution *ex ante* is not possible because this is dependent on the actual distribution of the data (Gardner et al., 1995).

First of all, a Poisson model assumes equidispersion as it assumes that the mean and variance are equal (Gardner et al., 1995). In the case of GRInno and NGInno, both variables show strong overdispersion as is confirmed by two dispersion tests which are highly significant with a dispersion value of 46.55 and 235.22 respectively. Second, for overdispersed data, both a quasi-Poisson and negative binomial model are applicable. The choice between these models is determined *ex post* based upon whether the theta parameter that is calculated upon fitting a negative binomial model is significant (Verhoef & Boveng, 2007). Both GRInno and NGInno show a significant theta parameter, which shows that a negative binomial model is preferred. Third, a large part of the Company Year observations have a zero value for GRInno (71%) and NGInno (52%). It is important to control for zero-inflation, which occurs when there are too many zeros in the data. It is tested whether a fitted negative binomial model is able to correctly model the zero values in the data. To do this, the check zeroinflation test is applied, which shows that the zeros for GRInno and NGInno are modelled correctly by 98 and 94 percent respectively. The 94 percent value for NGInno is slightly below the default limit of 95 percent. However, noting that the SUEST method requires identical models, this value is deemed acceptable. The final model that is used is a standard negative binomial model.

3.6.2. Seperate regressions for H1 and H2

First, regression Model 1 (Equation 3.1) is created to test hypothesis 1 to determine the effect that ESG performance has on green innovative performance. The regression model equation is:

$$GRInno_{i,t+1} = \beta_0 + \beta_1 ESG_{i,t} + \beta_2 Revenue_{i,t} + \beta_3 Year_{i,t} + \beta_4 Industry_{i,t} + \epsilon_{i,t} \tag{3.1}$$

where $GRInno_{i,t+1}$ represents the dependent variable count of green patents of firm i at time $t+1$ and $ESG_{i,t}$ represents the independent variable ESG rating. This difference in t and $t+1$ originates from the time lag in the relationship between ESG and innovation. In this equation β represents the coefficient values, $Revenue_{i,t}$, $Year_{i,t}$ and $Industry_{i,t}$ represent the control variables, and ϵ captures unexplained residuals.

Second, regression Model 2 (Equation 3.2) is created to test hypothesis 2 to determine the effect that ESG performance has on non-green innovative performance. This model only varies in terms of dependent variable $NGInno_{i,t+1}$, which represents the count of non-green patents of a firm i at time $t+1$. The regression model equation is:

$$NGInno_{i,t+1} = \beta_0 + \beta_1 ESG_{i,t} + \beta_2 Revenue_{i,t} + \beta_3 Year_{i,t} + \beta_4 Industry i, t + \epsilon_{i,t} \tag{3.2}$$

3.6.3. Comparing effects multiple dependent variables with SUEST for H3

In comparing effects on multiple dependent variables it is incorrect to directly compare coefficients (Mize et al., 2019). To deal with this, the SUEST method developed by (Weesie et al., 2000) and extended by (Mize et al., 2019) is applied. In SUEST, effects on dependent variables are first modeled separately and afterwards modeled jointly (Mize et al., 2019). This method is currently implemented in Stata (Weesie et al., 2000) and this research implements SUEST in R. A conceptual explanation of SUEST is provided below (for more information see (Mize et al., 2019)). SUEST is applied by 'stacking' data such that the two dependent variables become one column as shown in the matrices below. In these matrices ESGModel1 and ESGModel 2 provide the ESG rating of that observation separately for model 1 and 2. Moreover, ESGModel 1 and ESGModel2 are set to 0 if the data is sourced from the NGInno or GRInno respectively. Finally, isModel1 is a dummy variable that indicates whether the dependent variable is sourced from GRInno or NGInno.

$$\begin{bmatrix} GRInno & NGInno & ESG & ID \\ 3 & 6 & 2.4 & 1 \\ 5 & 9 & 3.2 & 2 \\ \vdots & \vdots & \vdots & \vdots \\ 7 & 13 & 3.8 & N \end{bmatrix}$$

$$\begin{bmatrix} Y & ESGModel1 & ESGModel2 & isModel1 & ID \\ GRInno = 3 & 2.4 & 0 & 1 & 1 \\ GRInno = 5 & 3.2 & 0 & 1 & 2 \\ \vdots & \vdots & \vdots & \vdots & \\ GRInno = 7 & 3.8 & 0 & 1 & N \\ NGInno = 6 & 0 & 2.4 & 0 & 1 \\ NGInno = 9 & 0 & 3.2 & 0 & 2 \\ \vdots & \vdots & \vdots & \vdots & \\ NGInno = 13 & 0 & 3.8 & 0 & N \end{bmatrix}$$

Upon the estimation of regression model 3.3, the coefficients of $ESGModel1$ and $ESGModel2$ are the same as the separate models. Whereas, the coefficient of dummy variable $isModel1$ represents a multiplication factor for coefficient $ESGModel1$. Wherein, $ESGModel1$ is multiplied by $1 + isModel1$. Thus a positive $isModel1$ strengthens the effect of $ESGModel1$ and vice-versa. After this multiplication the coefficients are compared.

$$Y_{i,t+1} = \beta_0 + \beta_1 ESGModel1_{i,t} + \beta_2 ESGModel2_{i,t} + \beta_3 isModel1_{i,t} + \epsilon_{i,t} \tag{3.3}$$

3.6.4. Additional regression analyses

To better understand the relationships within the SCM, several additional regression analyses are performed. This study examines whether an interaction effect exists between the relationship of ESG and Revenue on innovative performance. This helps determine whether the effect of ESG on innovative performance is dependent on the size of a firm. Additionally, this research introduces non-linear effects (i.e. quadratic terms) to the negative binomial model. This analysis aims to determine whether the relationship between ESG and innovative performance is linear or curvilinear. More specifically, this examines whether an optimal level of ESG rating and Revenue exists. By applying these additional analyses, this study provides a more comprehensive understanding of the underlying mechanisms in the relationship between ESG and innovation.

3.7. Quality criteria of methods

This research aims to maximize validity and reliability, but also explicates the existing challenges. First, internal validity is improved through usage of the DAG approach to isolate the estimation of causal effects (Textor et al., 2016). Moreover, employing the SUEST method instead of comparing coefficients of different models directly improves internal validity (Mize et al., 2019). On the other hand, ESG ratings and patents have weaknesses in how they measure their underlying concepts (Archibugi, 1992; Gibson Brandon et al., 2021), which decreases concept validity. Second, external validity is improved as this research focuses on all large European firms, which increases generalizability and reduces selection bias (Di Simone et al., 2022). Third, internal reliability is improved by analyzing the effects over several time periods and controlling for the temporal component (Bryman, 2016). Fourth, external reliability is improved by using secondary data sources, which strengthens reproducibility. However, even though the patent data and control variables are publicly available, the ESG is not, which hinders the reproducibility.

4
Results

4.1. Descriptive statistics

In this section, the descriptive statistics of the data are described. First, the numerical variables are discussed and afterwards the distribution of patent counts are analysed per patenting propensity category.

In terms of *ESG*, table 4.1 shows that nearly the entire range of possible ESG ratings is present in the the data. The quantiles in the data are evenly distributed, which confirms the even distribution of the data. Surprisingly, there are *Company* and *Year* combinations with a *Revenue* of 0. This is surprising because the STOXX600 consists of the largest European companies. Upon further inspection, it becomes evident that this solely pertains to Porsche in 2014-2016 and is due organisational re-structuring. Moreover, *Revenue* has a mean that is much higher than the median value. This indicates that the *Revenue* data is highly skewed. The median value is still high at 7 billion annual *Revenue*, which is expected due to the nature of the STOXX600. Similar to *Revenue*, *GRInno* and *NGInno* show a highly skewed distribution. It becomes apparent that more than half of the *Company* and *Year* combinations have 0 green or non-green patents. The maximum values are 698 for *GRInno* and 3213 for *NGInno*. Both these maximum values are held by Siemens.

Table 4.1: Descriptive statistics

Statistic	N	Mean	Min	Pctl(25)	Median	Pctl(75)	Max
ESG	1,456	3.250	0.500	2.700	3.300	3.900	5.000
Revenue	1,456	17,154.030	0.000	2,268.825	7,053.500	18,548.040	266,553.100
GRInno	1,456	12.708	0	0	0	2	698
NGInno	1,456	89.163	0	0	0	27	3,213

The strongest correlation described in table 4.2 is between *GRInno* and *NGInno*. This shows that innovative firms tend to be innovative both in terms of green and non-green innovations. This implies an increased difficulty of dissecting these separate effects. Omitting the correlation between *GRInno* and *NGInno*, *Revenue* shows the strongest correlation with all other variables. This highlights the importance of *Revenue* in this SCM and confirms the *a priori* expectation that *Revenue* is related to most variables in the SCM. Notably, *Year* is only correlated to *ESG*, which indicates the lack of over-time patterns in the other variables. This is surprising because *Revenue*, *GRInno* and *NGInno* all show an increasing trend over time as shown in figure 4.1. These positive trends might be explained by other factors (i.e. *ESG* and/or *Revenue*). Finally, *ESG* shows some correlation with all other variables. The weakest correlations are with *GRInno* and *NGInno*, which hints at a weak influence of *ESG* on innovative performance. However, as mentioned before, *'correlation does not imply causation'* (Kornbrot, 2005, p. 1). Furthermore, this is in line with what is theorised, wherein confounding effects occur in the SCM.

Table 4.2: Correlation matrix

	ESG	Revenue	Year	GRInno	NGInno
ESG	1				
Revenue	0.23	1			
Year	0.17	0.01	1		
GRInno	0.07	0.27	0.01	1	
NGInno	0.14	0.19	0.01	0.75	1

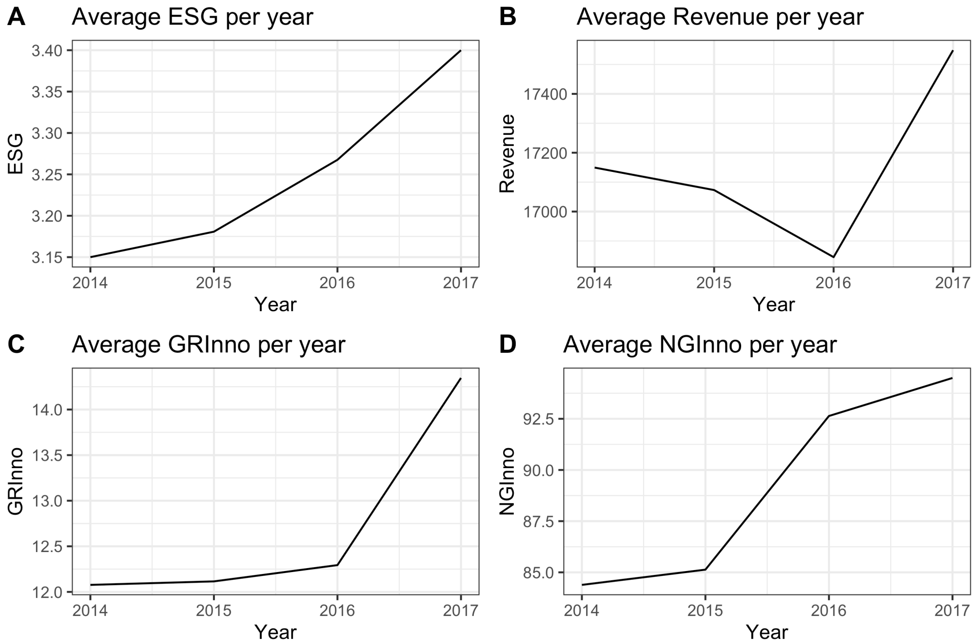

Figure 4.1: Development over time of the average *ESG*, *Revenue*, *GRInno* and *NGInno*

As highlighted in figure 4.2, a clear distinction exists between the number of green and non-green patents of the different patent propensity categories. Especially for the low propensity category a very small number of company-year combinations show non-zero counts. To clarify, a value of 1 on the x-axis represents a zero-value, because these axes denote GRINnno + 1 and NGInno + 1. This is necessary due to the log 10 transformation of the axes, because zero values become non-finite in these transformations. Moreover, in terms of non-green innovation (*NGInno*), the differences between the medium and high category are smaller. Allthough, the counts for the high category are slightly higher. As a result, it is expected that no large difference exists in how the medium and high propensity category influence green or non-green innovation.

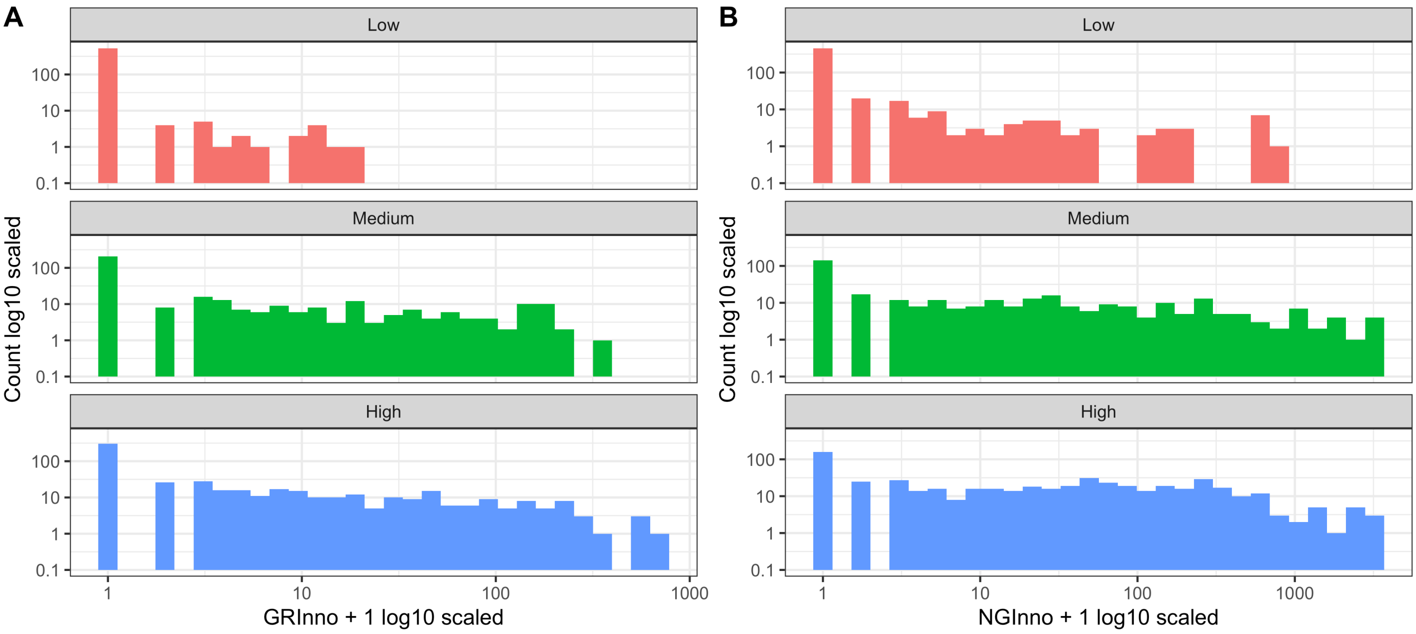

Figure 4.2: Counts of green patents (A) and non-green patents (B) for company-year combinations per propensity category with log10 scaled x and y axes

4.2. Regression results

In this section, the regression results are described. First, hypothesis 1 and 2 are discussed according to table 4.3. Subsequently, hypothesis 3 is reflected upon according to table 4.4. All models shown in table 4.3 and 4.4 are negative binomial models fitted on 364 companies that are part of the STOXX600 index. However, in table 4.4 some of the negative binomial models are modelled with the SUEST method. The time period of analysis is the years 2014 to 2017, which results in 1456 company-year combinations. In addition, a time lag of 1 year is applied to the effect of ESG on innovative performance.

4.2.1. Hypothesis 1 & 2: The effect of ESG on green and non-green innovation

The theory section postulates that a firms' ESG rating is positively related to both a firms' green (hypothesis 1) and non-green (hypothesis 2) innovative performance. This subsection investigates these relationships. Table 4.3 shows a control model and final model for the effects on both green innovation ($GRInno$) and non-green innovation ($NGInno$). In the control models (1 and 3), the effects of the control variables are modeled separately and in the final models (2 and 4) the effect of the dependent variable ESG is added.

In both control models (1 and 3) in table 4.3, it becomes apparent that the structure of the SCM is in line with the underlying SCM, which enhances the validity and reliability of the results. Surprisingly, the $Year$ dummy variable has no significant effect. Even though, line plot C and D in figure 4.1 show a clear positive trend over time. This indicates that this trend over time is explained by the other variables. In comparison to the final model, it becomes clear that the same effects hold. This emphasizes the correct formulation of the SCM and shows that the effects hold in both models.

The addition of ESG as a predictor in model 2 and 4 of table 4.3 show that ESG has a significant positive effect on both green innovation ($GRInno$) and non-green innovation ($NGInno$). This elucidates that an improved ESG rating has a positive effect on both green and non-green innovation of European firms. Consequently, for both hypothesis 1 and 2, the null hypothesis that states that no effect exists between ESG and green (H1) and non-green (H2) is rejected.

Moreover, the strong similarity between the regression outputs of model 2 and 4 verify the *a priori* expectation that green and non-green innovation follow the same SCM. The control variables in the SCM are discussed in further detail below. First of all, *Revenue* shows a significant positive effect in all models in table 4.3. As a result, it is confirmed that increased *Revenue* results in an improvement in both green and non-green innovative performance. The effect of *Revenue* is larger in the control models than in the final models. This indicates that *Revenue* is indeed a confounder for the effect of *ESG* on both green and non-green innovation. In other words, the analysis demonstrates that part of the effect that *Revenue* has on green and non-green innovation is achieved through improved *ESG* ratings. Secondly, as expected, the patenting *propensity category* of a firm has a large influence on the number of both green and non-green patents of a firm. The number of patents tend to increase from the low to medium and medium to high category. Opposingly, control model 3 shows that without the inclusion of *ESG*, the medium propensity category is more likely to patent non-green innovation than the high category. This is in line with expectations, because in figure 4.2 a large overlap exists between the green innovative patenting behavior of the medium and high propensity category.

In terms of model fit, models with the same dependent variable are compared according to the Akaike Information Criterion (AIC) and Log Likelihood. In both these measures model complexity is taken into account and a value closer to zero signals a superior model (Akaike, 1974). Upon comparison of the AIC for the control and final models, it becomes apparent that the addition of *ESG* as a predictor improves the model fit. Moreover, all models have a significant *theta* parameter, which justifies the usage of negative binomial models.

Table 4.3: Regression results for GRInno and NGInno

	Dependent variable:			
	GRInno		NGInno	
	(1)	(2)	(3)	(4)
ESG		0.523***		1.040***
		(0.115)		(0.090)
Revenue	0.00002***	0.00001***	0.00003***	0.00001***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Propensity category High	5.056***	5.175***	2.799***	3.166***
	(0.228)	(0.233)	(0.165)	(0.161)
Propensity category Medium	4.371***	4.323***	2.907***	2.478***
	(0.253)	(0.256)	(0.195)	(0.190)
Year	Included	Included	Included	Included
Country	Included	Included	Included	Included
Constant	-0.781	−2.625***	0.289	−2.251***
	(0.673)	(0.747)	(0.615)	(0.637)
Observations	1,456	1,456	1,456	1,456
Log Likelihood	-2,472.558	-2,465.041	-4,617.070	-4,572.233
θ	0.132*** (0.008)	0.135*** (0.008)	0.146*** (0.006)	0.159*** (0.007)
Akaike Inf. Crit.	4,989.116	4,976.082	9,278.141	9,190.465

Note: *p<0.1; **p<0.05; ***p<0.01

4.2.2. Hypothesis 3: Relative effects on green and non-green innovation

The results of hypothesis 1 and 2 indicate that the *a priori* assumption that ESG has a positive effect on both green and non-green innovation is confirmed. This does not shed light on whether ESG is mainly used as a signalling or legitimisation strategy. For this reason, it is important to determine the relative strengths of these effects. Solely interpreting the model coefficients indicates that the effect of ESG is stronger on non-green ($NGInno$) than on green innovation ($GRInno$), which would show that legitimisation is more prevalent. However, directly comparing coefficient of models with varying dependent variables can be problematic (Mize et al., 2019). To deal with this, the SUEST method is implemented to compare the effects of variables on varying dependent variables.

In table 4.4 the regression results for the SUEST method are described. In this overview, in model 1 $GRInno$ is regressed on solely ESG, in model 2 $NGInno$ is regressed solely on ESG and in model 3 both $GRInno$ and $NGInno$ are regressed on ESG in the stacked data set. In this latter model, the coefficient ESG_m1 represents the coefficient of ESG for $GRInno$, whereas coefficient ESG_m2 is the coefficient of ESG for $NGInno$. Finally, $ismodel1$ represents whether ESG_m1 is statistically different from the separate model. This first set of models is primarily included to determine whether the method works as planned because the example found in Mize et al. (2019) is also univariate. To elaborate, for models that only contain the main predictor of interest, the coefficients ESG_m1 and ESG_m2 should be approximately the same as the ESG coefficients of the separate models. Subsequently, model 4, 5 and 6 represent the same models with the inclusion of the control variables.

As shown by model 1-3, the R implementation of SUEST achieves the same results as the implementation of Mize et al. (2019). Noting that the coefficients are the same in the separate models and the combined model, it is concluded that SUEST is implemented correctly. Interestingly, $ismodel1$ is not statistically significant, which demonstrates that the coefficients ESG_m1 and ESG_m2 can be compared directly. Therefore, model 3 shows that the effect of ESG is stronger on non-green innovation ($NGInno$) than on green innovation ($GRInno$). This result is confirmed by model 6, which includes the control variables. In this final model, the difference in strength between the effects of ESG on green innovation ($GRInno$) and non-green innovation ($NGInno$) is even larger. The coefficient for $NGInno$ is even larger than in model 3, which again confirms the evident confounding role of *Revenue*, patenting *propensity category* and *country*. This shows that ESG has a stronger effect on non-green innovation than on green innovation. Consequently, for hypothesis 3, the null hypothesis that states that the effect of ESG is not stronger on green innovation than on non-green innovation cannot be rejected.

It is argued that the underlying mechanisms of how ESG influences green and non-green innovation stems from signalling and legitimacy theory. This juxtaposition of theories assumes two simultaneous counteracting effects. Signalling theory poses that well performing firms use ESG disclosure as a means of signalling their superior quality (Cho & Patten, 2007). Legitimacy theory posits that poor performing firms use ESG as a legitimisation strategy to legitimize a firm's activities that conflict with the values of society (O'donovan, 2002). The results from hypothesis 1 and 2 appear to confirm that indeed both these effects are occurring simultaneously. However, surprisingly, this research suggests that the legitimisation effect is more powerful. The reason for this is that previous research has shown that improved ESG ratings lead to increased R&D spending. Whereas, this research suggests that improved ESG ratings result in more non-green innovation than green innovation. If this is indeed due to increased R&D spending, this shows that many of the accrued benefits from improved ESG ratings are spent on non-green innovative endeavours. This bears consequences for how the societal contribution of increased focus on ESG disclosure should be assessed. The consequences of these results are discussed in further detail in the discussion section.

Table 4.4: Regression results for combined models with SUEST method based on Weesie et al. (2000)

	Dependent variable:					
	GRInno	NGInno	Combined	GRInno	NGInno	Combined
	(1)	(2)	(3)	(4)	(5)	(6)
ESG	0.604***	0.946***		0.523***	1.040***	
	(0.116)	(0.094)		(0.115)	(0.090)	
ESG_m1			0.603***			0.525***
			(0.102)			(0.099)
ESG_m2			0.946***			1.046***
			(0.102)			(0.091)
ismodel1			-0.685			-0.548
			(0.486)			(0.442)
Revenue				0.00001***	0.00001***	0.00001***
				(0.00000)	(0.00000)	(0.00000)
Propensity category High				5.175***	3.166***	3.758***
				(0.233)	(0.161)	(0.128)
Propensity category Medium				4.323***	2.478***	3.041***
				(0.256)	(0.190)	(0.149)
Year	Not included	Not included	Not included	Included	Included	Included
Country	Not included	Not included	Not included	Included	Included	Included
Constant	0.503	1.191***	1.191***	−2.625***	−2.251***	−2.187***
	(0.390)	(0.315)	(0.343)	(0.747)	(0.637)	(0.527)
Observations	1,456	1,456	2,912	1,456	1,456	2,912
Log Likelihood	-2,713.367	-4,823.770	-7,554.568	-2,465.041	-4,572.233	-7,085.900
θ	0.067*** (0.004)	0.101*** (0.004)	0.085*** (0.003)	0.135*** (0.008)	0.159*** (0.007)	0.143*** (0.005)
Akaike Inf. Crit.	5,430.733	9,651.541	15,117.140	4,976.082	9,190.465	14,221.800

Note: *p<0.1; **p<0.05; ***p<0.01

4.3. Additional regression analyses

To better understand how the dependent variables relate to the innovative performance of firms, some additional analyses are performed. First, a potential interaction effect between ESG and revenue is analysed. Second, it is analysed whether quadratic terms are present to determine whether the effect of ESG and revenue on innovation are constant or show diminishing or increasing returns.

4.3.1. Interaction effects

To determine whether the effects of *ESG* and *Revenue* on (non-)green innovation depend on each other, table 4.5 shows two regression models with added interaction effects between *ESG* and *Revenue*.

Notably, model 1 that estimates effects on green innovation (*GRInno*) indicates a negative significant interaction effect between *ESG* and *Revenue*. This signals that for firms with higher levels of *Revenue*, an increased *ESG* rating has a less prevalent positive effect on green innovation (*GRInno*). In other words, for larger firms an improved *ESG* rating tends to increase green innovation (*GRInno*) less than for smaller firms. It can be argued that for large firms the *ESG* rating is a relatively less strong predictor for green innovative performance (*GRInno*) than for smaller firms. Perhaps, this result points at the ability of larger firms to dedicate more resources to improving their *ESG* disclosure, which ultimately leads to them accomplishing higher *ESG* ratings. If this is the case, this would indicate that increased *Revenue* allows for improving *ESG* ratings without consequently improving green innovative performance (*GRInno*) as much as smaller firms.

The opposite interpretation of the interaction effect also exists. Wherein, at higher levels of *ESG*, the effect of *Revenue* on green innovation (*GRInno*) becomes less strong. However, the coefficient of *ESG* is higher in the model that includes the interaction effect as opposed to model 1 in table 4.10, which signals that at lower levels of *Revenue*, the importance of *ESG* ratings for green innovation performance is larger. This result would indicate that for firms with lower *Revenue* (i.e. smaller firms) a high *ESG* rating is an extra strong predictor of green innovative performance. Similar to the previous point, it may be that this points at the ability of larger firms to dedicate more resources to *ESG* disclosure, which allows them to accomplish higher *ESG* ratings, without increasing their underlying green innovative performance as much as smaller firms. This would help them accrue legitimisation effects of an increased *ESG* rating for their non-green activities. In similar fashion, it could be possible that for smaller firms *ESG* ratings more accurately signal a firms' actual green innovative performance.

In contrast to model 1, model 2 shows no significant interaction effect between *ESG* and *Revenue*. This shows that for non-green innovation (*NGInno*) no interaction effect occurs between *ESG* and *Revenue*. Therefore, the effects of neither *ESG* nor *Revenue* on non-green innovation are dependent on one another.

Table 4.5: Regression results additional analysis adding interaction effects between *ESG* and *Revenue*

	Dependent variable:	
	GRInno	NGInno
	(1)	(2)
ESG	0.935***	1.076***
	(0.133)	(0.101)
Revenue	0.0001***	0.00002
	(0.00002)	(0.00001)
Propensity category High	5.048***	3.154***
	(0.236)	(0.162)
Propensity category Medium	4.589***	2.501***
	(0.258)	(0.190)
Year	Included	Included
Country	Included	Included
ESG:Revenue	−0.00003***	-0.00000
	(0.00000)	(0.00000)
Constant	−4.129***	−2.345***
	(0.775)	(0.653)
Observations	1,456	1,456
Log Likelihood	-2,456.790	-4,572.123
θ	0.137*** (0.008)	0.159*** (0.007)
Akaike Inf. Crit.	4,961.580	9,192.246
Note:		*p<0.1; **p<0.05; ***p<0.01

4.3.2. Quadratic terms

Even though table 4.5 seems to indicate that an interaction effect exists between *ESG* and *Revenue*, it is important to consider other explanations. A negative significant interaction effect could also be explained by the existence of a negative quadratic term for either *ESG* or *Revenue*. To test this, table 4.6 show regression results for several models that include quadratic terms for *ESG* and *Revenue* and an interaction effect between *ESG* and *Revenue*

First of all, model 1 and 2 include quadratic *ESG* terms for the effect on green innovation (*GRInno*) and non-green innovation (*NGInno*) respectively. It becomes apparent that model 1 shows a significant positive *ESG* effect and a significant negative quadratic *ESG* term. This combination shows that the effect of *ESG* on green innovation (*GRInno*) follows an inverted u-shape trend. Wherein, at lower levels of *ESG* the effect of increased *ESG* ratings has a more positive effect on green innovative performance than at higher levels of *ESG* and the effect eventually turns negative. It is important to consider whether an interaction effect is still prevalent upon taking this quadratic effect into account.

In contrast to model 1 (*GRInno*), model 2 (*NGInno*) shows no significant quadratic *ESG* term, which was not necessarily expected either since no significant interaction effect was shown either.

Opposingly, both model 3 and 4 show a significant positive coefficient for *Revenue* and a significant negative quadratic *Revenue* effect. Similar to *ESG* in model 1, this shows that the effect of *Revenue* on both green and non-green innovation follows an inverted u-shape trend. Wherein, at lower levels of *Revenue* an increase in *Revenue* has a more positive effect on (non-)green innovation than at higher levels of *Revenue* and the effect eventually turns negative.

Moreover, model 5 and 6 show that these same trends hold upon the inclusion of both a quadratic *ESG* term and a quadratic *Revenue* term. This has several implications for the shape of the trend that is being modeled and it is key to determine whether these effects still hold upon the inclusion of both quadratic terms and the interaction effect.

To check this, model 7 includes both the quadratic terms and the interaction effect. Upon inspection of the coefficients of this final model, it becomes evident that the quadratic *ESG* term is no longer significant. Therefore the relationship between *ESG* and green innovation (*GRInno*) does not seem to follow an inverted u-shape trend upon the consideration of the interaction term between *ESG* and *Revenue*. The coefficient for *Revenue* is much higher in this model. However, the quadratic *Revenue* term is also higher and therefore at very high levels of *Revenue* an increase in *Revenue* leads to a larger decrease in green innovation. Finally, this output shows that the interaction term between *ESG* and *Revenue* is still present.

Table 4.6: Regression results additional analysis with inclusion of quadratic terms for *ESG* and *Revenue*

	Dependent variable:						
	GRInno	NGInno	GRInno	NGInno	GRInno	NGInno	GRInno
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ESG	1.964***	0.732	0.421***	0.917***	2.057***	0.932**	1.187*
	(0.651)	(0.474)	(0.118)	(0.092)	(0.634)	(0.473)	(0.635)
Rev. norm.	0.716***	0.716***	2.208***	2.050***	2.166***	2.055***	9.118***
	(0.161)	(0.140)	(0.353)	(0.298)	(0.354)	(0.298)	(1.088)
ESG2	−0.239**	0.052			−0.270***	-0.003	-0.050
	(0.102)	(0.076)			(0.100)	(0.076)	(0.102)
Rev. norm.2			−0.468***	−0.443***	−0.473***	−0.445***	−0.610***
			(0.110)	(0.096)	(0.110)	(0.096)	(0.119)
Prop. cat. H.	5.132***	3.184***	5.136***	3.068***	5.128***	3.067***	5.093***
	(0.235)	(0.162)	(0.231)	(0.160)	(0.233)	(0.161)	(0.234)
Prop. cat. M.	4.477***	2.438***	4.256***	2.386***	4.371***	2.393***	4.457***
	(0.258)	(0.191)	(0.252)	(0.189)	(0.255)	(0.189)	(0.254)
Year	Included	Included	Included	Included	Included	Included	Included
Country	Included	Included	Included	Included	Included	Included	Included
ESG:Rev. norm.							−1.838***
							(0.262)
Constant	−4.631***	−1.858**	−2.506***	−2.102***	−4.794***	−2.116**	−4.619***
	(1.200)	(0.916)	(0.739)	(0.632)	(1.169)	(0.913)	(1.151)
Observations	1,456	1,456	1,456	1,456	1,456	1,456	1,456
Log Likelihood	-2,462.546	-4,572.077	-2,459.556	-4,564.729	-2,458.183	-4,564.753	-2,450.692
θ	0.135*** (0.008)	0.159*** (0.007)	0.137*** (0.008)	0.161*** (0.007)	0.138*** (0.008)	0.161*** (0.007)	0.141*** (0.008)
Akaike Inf. Crit.	4,973.091	9,192.155	4,967.113	9,177.458	4,966.366	9,179.505	4,953.385

Note: *p<0.1; **p<0.05; ***p<0.01

4.4. Robustness checks

In this section a set of robustness checks is performed to determine whether changes to the model affect the hypothesis results. The following robustness checks are performed: varying time lag effect between ESG and innovative performance, the usage of $Industry$ categories instead of $Propensity_category$, usage of Poisson models instead of negative binomial, and finally different types of variable scaling are applied. In short, the robustness checks show that the result of the hypotheses hold under the performed robustness checks, which strengthens the results found in this research.

4.4.1. Varying time lag effect of ESG on innovative performance

To determine whether the selected time lag between the effect of ESG on innovative performance influences the hypotheses, models with varying time lags are fitted. Table 4.7 contains the results of these fitted models. In this table, the 0 models represent the previously used time lag of 1. Whereas, the -1 and +1 models represent the models with a 0 and 2 year time lag respectively.

Upon inspection of these results, it becomes clear that the overall results are the same as in the main results section. Accordingly, it is concluded that this robustness check does not result in different views on the hypotheses.

With regards to model fit, it is exhibited that an increase in time lag improves the model fit for both the green innovation ($GRInno$) and non-green innovation ($NGInno$) model. This result is however biased, since for the models with increased time lag the number of observations is smaller. Due to this, these models are required to explain a smaller variance, which in turn results in the improved model fit.

4.4.2. Usage of industry instead of patenting propensity categories

In table 4.8 regression results are shown for models that include the raw $Industry$ variable instead of the reclassified *Propensity category*. In terms of hypothesis 1 and 2 it is confirmed that under this robustness check ESG still has a significant positive effect on both green innovation ($GRInno$) and non-green innovation ($NGInno$). Therefore, in this robustness check for hypothesis 1 and 2 the null hypotheses are still rejected. Nonetheless, in contrast to the previous models, the ESG coefficient for the green innovation ($GRInno$) model is now higher than the coefficient of the non-green innovation ($NGInno$) model. At face value, this indicates that the effect of ESG is stronger on green innovation than on non-green innovation. However, this example shows the benefit of the SUEST method, because this method shows that the effect of ESG on green innovation ($GRInno$) is statistically significantly lower than on non-green innovation ($NGInno$). As a consequence, the same result for hypothesis 3 is found in this robustness check as in the main results section.

For both separate models and the combined SUEST model, the models with $Industry$ instead of *Propensity category* show a lower AIC value. Thus, the replacement of *Propensity category* with $Industry$ leads to an improved model fit, which is unsurprising because the $Industry$ variable allows for modelling at increased granularity.

Table 4.7: Regression results robustness check varying timelag of effect of ESG on innovative performance

	Dependent variable:					
	GRInno			NGInno		
	(-1)	(0)	(+1)	(-1)	(0)	(+1)
ESG	0.579***	0.523***	0.493***	1.129***	1.040***	1.005***
	(0.105)	(0.115)	(0.133)	(0.082)	(0.090)	(0.102)
Revenue	0.00001***	0.00001***	0.00002***	0.00001***	0.00001***	0.00001***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Propensity category High	5.051***	5.175***	5.137***	3.264***	3.166***	3.100***
	(0.205)	(0.233)	(0.269)	(0.144)	(0.161)	(0.184)
Propensity category Medium	4.173***	4.323***	4.266***	2.519***	2.478***	2.460***
	(0.226)	(0.256)	(0.297)	(0.170)	(0.190)	(0.219)
Year	Included	Included	Included	Included	Included	Included
Country	Included	Included	Included	Included	Included	Included
Constant	−2.799***	−2.625***	−2.492***	−2.590***	−2.251***	−2.081***
	(0.679)	(0.747)	(0.864)	(0.574)	(0.637)	(0.722)
Observations	1,820	1,456	1,092	1,820	1,456	1,092
Log Likelihood	-3,091.803	-2,465.041	-1,843.695	-5,698.954	-4,572.233	-3,453.448
θ	0.134*** (0.007)	0.135*** (0.008)	0.130*** (0.009)	0.161*** (0.006)	0.159*** (0.007)	0.161*** (0.008)
Akaike Inf. Crit.	6,231.606	4,976.082	3,731.389	11,445.910	9,190.465	6,950.895

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.8: Regression results robustness check industry instead of propensity categories

	Dependent variable:		
	GRInno	NGInno	Combined
	(1)	(2)	(3)
ESG	0.918***	0.687***	
	(0.126)	(0.094)	
ESG_m1			0.665***
			(0.103)
ESG_m2			0.794***
			(0.093)
ismodel1			−1.760***
			(0.433)
Revenue	0.00002***	0.00002***	0.00002***
	(0.00000)	(0.00000)	(0.00000)
Year	Included	Included	Included
Industry	Included	Included	Included
Country	Included	Included	Included
Constant	4.458***	3.036***	4.169***
	(0.819)	(0.686)	(0.566)
Observations	1,456	1,456	2,912
Log Likelihood	-2,368.376	-4,420.518	-6,864.160
θ	0.176*** (0.011)	0.213*** (0.010)	0.186*** (0.007)
Akaike Inf. Crit.	4,816.752	8,921.037	13,812.320
Note:			*p<0.1; **p<0.05; ***p<0.01

4.4.3. Usage of Poisson instead of negative binomial

It is determined that a negative binomial model is the best fit for the data set of this research, due to the presence of overdispersion and a significant theta parameter. On the other hand, it is beneficial to determine whether the same effects hold with the usage of a Poisson model. The regression results of the Poisson models are shown in table 4.9.

In terms of the coefficients, the Poisson models are highly similar to the negative binomial model results shown in 4.4. The direction and statistical significance of the coefficient are the same between the Poisson and negative binomial models. In contrast, the strength of the coefficients is slightly lower and the relative difference in strength of the effect of ESG ratings on green and non-green patenting is larger. Be that as it may, it is important to consider that directly comparing coefficients between different model types is not desirable. Altogether, the Poisson model results do not shed doubt on the results found in the main results section for either hypothesis 1, 2 or 3.

Finally, the AIC values of the Poisson models indicate that the model fit of the Poisson models is much worse than the negative binomial models. Therefore, the utilization of negative binomial models in this research is reaffirmed.

Table 4.9: Regression results robustness check Poisson instead of negative binomial

	Dependent variable:		
	GRInno	NGInno	Combined
	(1)	(2)	(3)
ESG	0.332***	0.717***	
	(0.011)	(0.005)	
ESG_m1			0.293***
			(0.010)
ESG_m2			0.725***
			(0.005)
ismodel1			−0.394***
			(0.040)
Revenue	0.00001***	0.00001***	0.00001***
	(0.00000)	(0.00000)	(0.00000)
Propensity category High	4.507***	2.303***	2.433***
	(0.091)	(0.013)	(0.012)
Propensity category Medium	3.835***	2.001***	2.087***
	(0.091)	(0.013)	(0.013)
Year	Included	Included	Included
Country	Included	Included	Included
Constant	−3.318***	−0.764***	−0.890***
	(0.140)	(0.050)	(0.046)
Observations	1,456	1,456	2,912
Log Likelihood	-24,116.440	-156,290.900	-182,738.700
Akaike Inf. Crit.	48,278.890	312,627.700	365,527.400
Note:			*p<0.1; **p<0.05; ***p<0.01

4.4.4. Variable scaling

Due to the high number of zeros in patents in the data, log transformations are not applicable. This is solvable through the additional transformation of adding a small number to the patents. Unfortunately, in this use case, this is only solved when a full patent is added (i.e. $\log(GRInno/NGInno + 1)$) and this transforms the problem to be modelled significantly. As a consequence, no transformations are applied to the patent variables.

Opposingly, in table 4.10, regression results are shown wherein the *Revenue* variable is normalised to allow for easier interpretation of its relative strength in comparison to *ESG*. For this variable, a MinMax normalisation is applied to normalise the *Revenue* data to the same scale as the *ESG* ratings (i.e. 0-5). The normalisation does not have a large effect on the coefficients of the other variables and there is no reason to re-evaluate the hypotheses. Interestingly, in model 1 for green innovation (*GRInno*) the coefficient of *Revenue* is slightly larger than *ESG*. Whereas, in model 2 for non-green innovation (*NGInno*) the coefficient of *Revenue* is slightly lower than *ESG*. It might be the case that in comparison to *ESG* ratings, *Revenue* is more important in green patenting (*GRInno*) than in non-green patenting (*NGInno*). Conversely, model 1, 2 and 3 all show highly similar coefficients for *Revenue* and therefore this points to the fact that *Revenue* has a highly similar effect in both SCMs. This means that *Revenue* has a similar positive effect on both green and non-green patenting behavior, which is in line with *a priori* assumptions.

Table 4.10: Regression results robustness check variable scaling with MinMax scaled Revenue between 0 and 5

	Dependent variable:		
	GRInno	NGInno	Combined
	(1)	(2)	(3)
ESG	0.523***	1.040***	
	(0.115)	(0.090)	
ESG_m1			0.525***
			(0.099)
ESG_m2			1.046***
			(0.091)
ismodel1			-0.548
			(0.442)
Revenue normalised	0.785***	0.722***	0.710***
	(0.161)	(0.140)	(0.106)
Propensity category High	5.175***	3.166***	3.758***
	(0.233)	(0.161)	(0.128)
Propensity category Medium	4.323***	2.478***	3.041***
	(0.256)	(0.190)	(0.149)
Year	Included	Included	Included
Country	Included	Included	Included
Constant	−2.625***	−2.251***	−2.187***
	(0.747)	(0.637)	(0.527)
Observations	1,456	1,456	2,912
Log Likelihood	-2,465.041	-4,572.233	-7,085.900
θ	0.135*** (0.008)	0.159*** (0.007)	0.143*** (0.005)
Akaike Inf. Crit.	4,976.082	9,190.465	14,221.800
Note:			*p<0.1; **p<0.05; ***p<0.01

5
Conclusion

The objective of this study is to shed a light on how ESG rating development influences both green and non-green innovation. Furthermore, this research attempts to determine the relative strength of these effects. This helps unravel how firms use the benefits attributed to improved ESG ratings and how an increased focus on ESG can potentially help to encourage sustainable development. Existing published research of the influence of ESG on patenting behavior exclusively focuses on Asian markets. Moreover, European ESG literature is primarily of financial nature and focuses on the effects ESG has on firms' financial performance. This leads to a lack of understanding about the relationship between ESG and innovative performance in the European Union. This lack of understanding is surprising, noting the upcoming CSRD and the call by the European Commission for research into the influence of ESG on innovation (EC, 2011). This research attempts to fill that existing gap. To concretize, this study aims to answer: *"How do changes in firms' ESG rating affect the green and non-green innovative performance of firms in the European Union and how do these effects relate to each-other?"*. To answer this question, patent data of European firms that are part of the STOXX600 index, representing the 600 largest European companies, between 2014 and 2017 was analyzed. This analysis was conducted using negative binomial regression analysis. In addition, the relative strength of effects were analysed using the Seemingly Unrelated Estimation (SUEST) method put forward by Weesie et al. (2000).

The *a priori* expectations of this study were that a positive relationship exists between ESG rating development and both green and non-green innovative performance of a firm. This was expected due to the simultaneous effects of legitimacy and signalling theory, wherein both well and poor performing firms engage in ESG disclosure (Cho & Patten, 2007; O'donovan, 2002). The findings reveal that ESG ratings indeed positively influence the green and non-green innovative performance of European firms. This shows that the positive relationship between ESG and (green) innovation that is found in Asia (Broadstock et al., 2020; Liu & Lyu, 2022; Tang, 2022; Xu et al., 2020; Yin & Wang, 2018; Zhang et al., 2020; Zheng et al., 2022) also exists in Europe. This extends the collective understanding of the relationship between ESG and innovation in Europe. In the case of green innovation, the strength of this effect is especially evident for firms with lower revenues (i.e. smaller firms) and for larger firms this effect is less strong. Nonetheless, at larger firms an improved ESG rating still leads to an increase in green innovative performance. This reduction in the strength of the effect of ESG on innovative performance is not found for non-green innovation. For both green and non-green innovation, the size of a firm is shown to be related to innovative performance in an inverted u-shape pattern. This exemplifies that as revenue increases, a further increase in revenue has a less strong effect on innovative performance. If firms get large enough this trend even reverses and increased revenues negatively influence innovative performance, but this only occurs at extremely high revenue levels.

Moreover, it was expected that a firms' ESG rating would have a relatively stronger effect on green innovation than on non-green innovation. This originates from the expectation that firms that use ESG disclosure as a legitimisation tool also experience increased pressures to improve their underlying ESG performance (Liu & Lyu, 2022; Yin & Wang, 2018). This study suggests that the opposite is true.

Wherein, a firms' ESG rating is more strongly related to non-green innovation than green innovation.

In conclusion, the answer to the research question is as follows. Improvements of EGG ratings of firms in the European Union positively influence both green and non-green innovation. It is concluded that the legitimisation and signalling effect are occurring simultaneously. For green innovation, this effect is especially strong for smaller firms and less strong for larger firms. Whereas, for non-green innovation the strength of the effect of ESG is not dependent on firm size. This indicates that for smaller firms an ESG rating is a stronger predictor of underlying green innovative performance. In addition, this could signal the green-washing ability of larger firms, which helps them increase their ESG rating without improving their underlying performance. Counter to *a priori* expectations, improvements of ESG ratings generally have a stronger effect on non-green innovation than on green innovation. This unexpected result might be explained by ESG entailing more than the environmental pillar and the other pillars also being conducive for innovation. This alternative explanation is explored further in the discussion section.

This research shows that society should be critical about how ESG ratings contribute to sustainable development. It seems that the increased access to funding at lower cost for firms with higher ESG ratings may not solely contribute to sustainable endeavours. This is emphasized even more by the confirmation that ESG disclosure is also used as a tool to legitimise activities that are in in-congruence with societal values. These findings are especially interesting noting the lack of focus on non-financial effects of ESG disclosure in European finance literature. These theoretical and practical implications are elaborated on in further detail in the discussion chapter.

6
Discussion

In this chapter, the conclusions described in this study are interpreted and discussed in relation to existing literature. Furthermore, the implications of the findings and the contributions of this paper are described. Finally, a critical review of the limitations of the research design is outlined. In these three sub-chapters, some recommendations for further research are found.

6.1. Interpretation of conclusion and theoretical embedding

The findings of this research indicate that ESG ratings have a positive influence on green innovation. This is consistent with the theory section, which posits that an improved ESG rating leads to easier access to capital (Cheng et al., 2014; García-Sánchez et al., 2019) at lower cost (El Ghoul et al., 2011) and attracts stakeholders interested in sustainable investments, thereby increasing R&D spending (Broadstock et al., 2020; Herremans et al., 1993). In addition, the finding that ESG also positively influences non-green innovation provides evidence for the occurrence of both the signalling and legitimisation effects, aligning with expectations. Signalling theory posits that well-performing firms engage in ESG disclosure to signal their superior performance (Cho & Patten, 2007), while legitimacy theory suggests that poor-performing firms would utilize ESG as a legitimisation strategy (O'donovan, 2002). Consequently, it was anticipated that higher ESG ratings would result in increased R&D spending, regardless of their underlying rationale; however, not all of this spending would be allocated towards green innovation. In contrast to *a priori* expectations, this research found a larger effect of ESG on non-green innovation than on green innovation. The following paragraphs provide a more thorough interpretation of the findings of this paper.

It should be noted that signalling and legitimisation theory are not mutually exclusive and these theories can simultaneously impact the relationship between ESG and innovation. The extent to which firms use ESG either as a legitimisation or signalling tool are dependent on specific circumstances and this is not necessarily a zero-sum game. Besides, even though ESG was initially put forward to promote sustainable development (UNGC, 2004), non-green innovation can also be beneficial to society (Ahlstrom, 2010). Hence, even though innovation can be marked as non-green, this does not necessarily imply that the innovation is non-beneficial to society. The definition of non-green in this research is solely based upon whether a patent is part of a Y02 CPC patent class: *'technology for adaptation to climate change'* (EPO, 2022a). Thus, non-green patents should not all be considered as working counter-actively to green patents. It is shown that the effect on non-green innovation is larger. This shows that some of the benefits that are provided to firms that improve their ESG rating seem to be put to use for non-green innovation. Future research can help to more accurately determine how the benefits that firms accrue due to an improved ESG rating are utilized. This follow up research could best entail a systematic analysis of financial data, coupled with a qualitative inquiry into the factors that influence R&D investment decisions of firms that engage in ESG disclosure.

The finding that ESG has a stronger effect on non-green innovation than on green innovation is unexpected. It is important to consider theoretical explanations for this finding. An explanation for this might be that ESG also considers Social and Governance and this research does not make this distinction. As shown by earlier research, good social conditions for workers are conducive for increased innovation (Landry et al., 2002). In addition, previous research by (Belloc, 2012) indicates that proper governance structures can positively influence innovative performance. This is underlined by the percentage distribution of how the individual pillars contribute to the overall ESG rating. Especially noting that for the FTSE-Russell (2022b) ESG dataset this distribution is dependent on the industry. This can result in high percentages of the overall ESG rating being attributed to the social or governance rating. Hence, this might explain why the effect of ESG is stronger on non-green innovation than on green innovation. For future research, it would be beneficial to assess the effects of the pillar specific ratings on green and non-green innovation.

Another key finding is the interaction effect that is found between ESG and revenue on green innovation. The existing body of literature does not look into possible interaction effects between revenue/size and ESG on innovative performance (Broadstock et al., 2020; Liu & Lyu, 2022; Tang, 2022; Xu et al., 2020; Yin & Wang, 2018; Zhang et al., 2020; Zheng et al., 2022). For this reason, this provides novel insights into the relationship between ESG and innovative performance. This interaction effect shows that the positive effect of improved ESG ratings on innovation is strongest for smaller firms and weaker for larger firms. There could be multiple factors that contribute to this effect. First of all, this might signal that larger firms are more capable at increasing their ESG rating without increasing their underlying green innovative performance. This could mean that larger firms might be more proficient in their ability to use ESG as a legitimisation strategy. This suggestion is in line with existing research by Yu et al. (2020) that highlights how larger firms are more adept at green-washing their activities. Second, this could also be related to the inverted u-shaped relationship between revenue and innovative performance. However, this interaction effect is still prevalent upon taking this u-shaped pattern into consideration. Third, it could also be explained by large firms generally having a higher ESG rating (Drempetic et al., 2020), which results in the ESG rating being a less strong indicator of actual underlying ESG performance.

It is concluded that the effect of revenue on both green and non-green innovation follows an inverted u-shaped trend. This is in line with a common trend in innovation literature that often shows an inverted u-shape relationship between variables and innovative performance. For example, these inverted u-shape relationships are found between innovation and competition (Aghion et al., 2005), and innovation and organizational slack (Nohria & Gulati, 1996). This shows that similar to other effects, increases in revenue only improves innovative performance up to an optimal level of revenue. In these edge cases, ESG would become a more important predictor, but the interaction effect between ESG and revenue reduces this.

6.2. Implications and contributions

This research proposes that for smaller firms, ESG is a stronger predictor of green innovative performance than for larger firms. It is shown that as firms become larger an improved ESG rating has less strong positive effects on green innovative performance. This finding has several practical implications. First of all, the upcoming CSRD forces more smaller firms to disclose information about their ESG performance. This finding suggests that especially smaller firms increase their green innovation upon improved ESG ratings. Therefore, this increased focus on ESG by smaller firms will ultimately increase their green innovative performance, which will most likely have a positive impact on sustainable development within the European Union. Secondly, the observation that for larger firms an improved ESG ratings tends to have less strong positive effects on green innovative performance should signal to policy makers that the disclosed ESG information of large firms should be subject to additional scrutiny. In similar fashion to the previous point, the upcoming CSRD introduces an auditing requirement for ESG information reported by larger firms. This increased audit requirement seems to tackle some of the challenges about ESG reporting put forward in this research. Noting that the small firms in our sample are still larger than many others, it is paramount to investigate whether the observed effects also extend to truly small firms. Such insights would assist policymakers in determining whether the CSRD is also beneficial in these situations.

Furthermore, the application of the SUEST method developed by Weesie et al. (2000) and implementation of the method in R provides a significant methodological contribution in innovation sciences. This method allows for more accurate comparison of the strength of effects across multiple dependent variables. Even though a handful of innovation sciences papers use this method, it is not broadly used and other scholars use the Stata implementation. More broad usage of the SUEST method in innovation sciences could help enhance the methodological robustness of the field. In addition, previous research into the relationship between ESG and innovation has not addressed the comparison between green and non-green innovation. This renders the application of the SUEST method particularly useful in the advancement of the theoretical understanding of this relation.

The finding that the effect of ESG is stronger on non-green innovation than on green innovation represents an interesting and novel empirical contribution. This result challenges the *a priori* expectation that ESG more strongly affects green innovation. Theoretically, this suggests that ESG has broader effects on firms' overall innovative capacity and confirms that ESG positively influences both green and non-green innovation. Practically, this finding provides evidence that firms can leverage their ESG efforts to enhance their overall innovative capacity. This finding raises questions on the underlying mechanisms of how ESG influences innovation. Future research should attempt to gain a better understanding of these mechanisms. A clearer understanding of the underlying mechanisms of how ESG influences innovation could provide meaningful insights to policy makers on how to foster an innovative business environment. Nevertheless, this finding already provides policy makers with additional arguments in favor of the implementation of ESG policies.

Finally, the existing literature on ESG is primarily of a financial nature and related to the financial performance of firms. The focus on financial benefits has overshadowed the non-financial effects of ESG. This scope has started to broaden to encompass its impact on other firm metrics. This research bridges the gap in European finance and innovation literature by focusing on the non-financial effects of ESG. Noting the novel findings of this paper, it can be concluded that bridging this gap further can lead to interesting findings. Further incorporation of these literature strands could improve the collective understanding of innovation from a new viewpoint. To concretize, these studies could look into the relation between ESG and more specific innovation topics. For example, open innovation, innovative network formation, innovation culture or employee creativity.

6.3. Limitations

While efforts have been made to reduce methodological constraints, this research is not exempt from limitations. The limitations of this study follow the following four categories: constraints related to ESG data, measuring innovation and ESG in general, sample size, and data imputation. These four categories are discussed below.

For example, working with ESG data results in the following three limitations. First, ESG data is not publicly nor freely available, which reduces the comparability of this research. Secondly, ESG ratings vary per ESG rating agency. This makes results over studies that use ESG ratings from different ESG rating agencies less comparable. Furthermore, this points at an inherent weakness of calculated ESG ratings that originates from a lack of standardization of calculation methods of different rating agencies. Nevertheless, ESG ratings provide researchers with condensed measures for the underlying ESG performance. This prevents researchers from having to investigate countless ESG reports themselves. Third, the unaudited nature of disclosed ESG information bears opportunity for induced framing bias by firms that disclose information (Hazen, 2020; Yu et al., 2020). Consequently, firms are able to use ESG disclosure as a means to highlight good aspects and underemphasize negative aspects. This negative consequence is strengthened further by the inherent framing bias of less sophisticated investors who are more likely to act upon positively framed information (Tan et al., 2014). The upcoming CSRD will partly deal with this by increasing auditing requirements for ESG information disclosed by large firms. In the future, this increased audit obligation will most likely lead to a data quality improvement.

The demarcation of this research creates further limitations, particularly in relation to more holistic views of what ESG and innovation denote. Foremost, patenting behavior only captures a part of the innovative performance of a firm. In doing so, this research does not provide insights into details of the innovative process, but merely looks at innovative outputs. Coupled with that, patenting behavior primarily relates to the Environmental pillar of ESG through technological development and focuses less on the Social and Governance pillars. Many ESG rating agencies also provide ratings on a 'pillar-specific level' which represent firms' performance at the Environmental (E), Social (S) and Governance (G) level (Christensen et al., 2022). Preferably, this research would have used an ESG rating dataset that contains ratings at this finer granularity. This would allow for a more comprehensive analysis. Due to the non-public nature of ESG rating data, this data was not available to the researcher. Future research into the effect of pillar specific ratings on innovative performance would make a valuable addition to the existing body of literature. Finally, this research employs a rather linear perspective on the development of innovation because it solely looks at the patenting behavior of firms. For that reason, this research takes a simplified stance on the process of innovation, which might not fully capture the existent complexity and dynamics. A follow-up research into underlying mechanisms of the complex relationship between ESG and innovation is advised. For example, a case study or case comparison that analyses ESG documents in-depth paired with interviews could elaborate on the innovation dynamics at specific firms.

Moreover, the sample size and short time period of analysis can pose a limitation. Large firms in the European Union have only been subject to mandatory ESG disclosure since 2014. For more recent years, patent data is not yet fully published due to large backlogs at patenting offices. This combination results in a short time period of analysis. This selection method for the sample results in a reduced selection bias, due to the imposed reporting requirement. Thus, even though a larger sample would be useful, this larger sample would result in different biases in the data. Similarly, patent performance can be quite random year-on-year and it can be beneficial to look into multi-year windows to discover longer term trends. The short time period of available data renders it difficult to analyse the data at decreased granularity.

This research applies backfill data imputation to a missing not at random (MNAR) missingness problem and this results in induced bias in the dataset. Be that as it may, any data imputation applied to a MNAR missingness problem results in bias. The employment of data imputation decreases the reduction in sample size, which positively influences the internal validity of this research.

References

Adams, R., Jeanrenaud, S., Bessant, J., Denyer, D., & Overy, P. (2016). Sustainability-oriented innovation: A systematic review. *International Journal of Management Reviews*, *18*(2), 180–205.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and innovation: An inverted-u relationship. *The Quarterly Journal of Economics*, *120*(2), 701–728.

Ahlstrom, D. (2010). Innovation and growth: How business contributes to society. *Academy of Management Perspectives*, *24*(3), 11–24.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use esg information: Evidence from a global survey. *Financial Analysts Journal*, *74*(3), 87–103.

An, Y., Davey, H., & Eggleton, I. R. (2011). Towards a comprehensive theoretical framework for voluntary ic disclosure. *Journal of Intellectual Capital*, *12*(4), 571–585.

Archibugi, D. (1992). Patenting as an indicator of technological innovation: A review. *Science and Public Policy*, *19*(6), 357–368.

Arundel, A., & Kabla, I. (1998). What percentage of innovations are patented? empirical estimates for european firms. *Research Policy*, *27*(2), 127–141.

Aupperle, K. E., Carroll, A. B., & Hatfield, J. D. (1985). An empirical examination of the relationship between corporate social responsibility and profitability. *Academy of Management Journal*, *28*(2), 446–463.

Belloc, F. (2012). Corporate governance and innovation: A survey. *Journal of Economic Surveys*, *26*(5), 835–864.

Ben-Gal, I. (2008). Bayesian networks. *Encyclopedia of statistics in quality and reliability*, *1*.

Berg, F., Koelbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of esg ratings. *Review of Finance*, *26*(6), 1315–1344.

Bessen, J., & Hunt, R. M. (2007). An empirical look at software patents. *Journal of Economics & Management Strategy*, *16*(1), 157–189.

Boons, F., & Lüdeke-Freund, F. (2013). Business models for sustainable innovation: State-of-the-art and steps towards a research agenda. *Journal of Cleaner Production*, *45*, 9–19.

Boons, F., Montalvo, C., Quist, J., & Wagner, M. (2013). Sustainable innovation, business models and economic performance: An overview. *Journal of Cleaner Production*, *45*, 1–8.

Boscaljon, B., Filbeck, G., & Smaby, T. (2006). Information content of business methods patents. *Financial Review*, *41*(3), 387–404.

Broadstock, D. C., Matousek, R., Meyer, M., & Tzeremes, N. G. (2020). Does corporate social responsibility impact firms' innovation capacity? the indirect link between environmental & social governance implementation and innovation performance. *Journal of Business Research*, *119*, 99–110.

Bryman, A. (2016). *Social research methods*. Oxford University Press.

Cheng, B., Ioannou, I., & Serafeim, G. (2014). Corporate social responsibility and access to finance. *Strategic Management Journal*, *35*(1), 1–23.

Cho, C. H., & Patten, D. M. (2007). The role of environmental disclosures as tools of legitimacy: A research note. *Accounting, Organizations and Society*, *32*(7-8), 639–647.

Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is corporate virtue in the eye of the beholder? the case of esg ratings. *The Accounting Review*, *97*(1), 147–175.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: The two faces of r & d. *The Economic Journal*, *99*(397), 569–596.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39–67.

Dablander, F. (2020). An introduction to causal inference.

De Lucia, C., Pazienza, P., & Bartlett, M. (2020). Does good esg lead to better financial performances by firms? machine learning and logistic regression models of public enterprises in europe. *Sustainability*, *12*(13), 5317.

Deegan, C. (2002). Introduction: The legitimising effect of social and environmental disclosures–a theoretical foundation. *Accounting, auditing & accountability journal*, *15*(3), 282–311.

Deloitte. (2022). *Esg explained: What is esg?* Retrieved November 4, 2022, from https://www2.deloitte.com/ce/en/pages/global-business-services/articles/esg-explained-1-what-is-esg.html

Dernis, H., & Khan, M. (2004). Triadic patent families methodology. *2004*(2).

Di Simone, L., Petracci, B., & Piva, M. (2022). Economic sustainability, innovation, and the esg factors: An empirical investigation. *Sustainability*, *14*(4), 2270.

Dorfleitner, G., Halbritter, G., & 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), 450–466.

Dowling, J., & Pfeffer, J. (1975). Organizational legitimacy: Social values and organizational behavior. *Pacific Sociological Review*, *18*(1), 122–136.

Drempetic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the esg score: Corporate sustainability ratings under review. *Journal of Business Ethics*, *167*(2), 333–360.

EC. (2011). *Towards a greater understanding of the changing role of business in society.* Retrieved November 9, 2022, from https://op.europa.eu/en/publication-detail/-/publication/a0893c6d-8526-4a9b-87ff-7392cdea2667

EC. (2014). *Directive 2014/95/eu of the european parliament and of the council.* Retrieved November 16, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%5C%3A32014L0095

EC. (2021a). Questions and answers: Corporate sustainability reporting directive proposal. Retrieved October 19, 2022, from https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1806

EC. (2021b). Sustainable finance package. Retrieved October 19, 2022, from https://finance.ec.europa.eu/publications/sustainable-finance-package_en

El Ghoul, S., Guedhami, O., Kwok, C. C., & Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance*, *35*(9), 2388–2406.

Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology*, *40*, 31.

EPO. (2022a). Classification search. Retrieved November 25, 2022, from https://worldwide.espacenet.com/patent/cpc-browser#!/CPC=Y02A

EPO. (2022b). *Quality report 2021*. Retrieved March 1, 2023, from https://documents.epo.org/projects/babylon/eponet.nsf/0/BBA26B12012E8BB3C12588550049A1BD/$FILE/quality_report_2021_en.pdf

ESMA. (2022). *Regulated market registry*. Retrieved November 15, 2022, from https://registers.esma.europa.eu/publication/searchRegistercore=esma_registers_upreg#

Fatemi, A., Glaum, M., & Kaiser, S. (2018). Esg performance and firm value: The moderating role of disclosure. *Global Finance Journal*, *38*, 45–64.

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

FTSE-Russell. (2022a). Esg ratings. Retrieved January 20, 2023, from https://www.ftserussell.com/data/sustainability-and-esg-data/esg-ratings

FTSE-Russell. (2022b). Guide to ftse and third party sustainable investment data used in ftse indices. Retrieved January 20, 2023, from https://research.ftserussell.com/products/downloads/Guide_to_FTSE_and_Third_Party_Sustainable_Investment_Data_used_in_FTSE_Russell_Indices.pdf

García-Sánchez, I.-M., Hussain, N., Martínez-Ferrero, J., & Ruiz-Barbadillo, E. (2019). Impact of disclosure and assurance quality of corporate sustainability reports on access to finance. *Corporate Social Responsibility and Environmental Management*, *26*(4), 832–848.

Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed poisson, and negative binomial models. *Psychological Bulletin*, *118*(3), 392.

Gibbons, B. (2020). Environmental and social disclosure and firm-level innovation. *Available at SSRN 3658415*.

Gibson Brandon, R., Krueger, P., & Schmidt, P. S. (2021). Esg rating disagreement and stock returns. *Financial Analysts Journal*, *77*(4), 104–127.

Giese, G., Lee, L.-E., Melas, D., Nagy, Z., & Nishikawa, L. (2019). Foundations of esg investing: How esg affects equity valuation, risk, and performance. *The Journal of Portfolio Management*, *45*(5), 69–83.

Gillan, S. L., Koch, A., & Starks, L. T. (2021). Firms and social responsibility: A review of esg and csr research in corporate finance. *Journal of Corporate Finance*, *66*, 101889.

Hagedoorn, J., & Cloodt, M. (2003). Measuring innovative performance: Is there an advantage in using multiple indicators? *Research Policy*, *32*(8), 1365–1379.

Hall, B. H., Griliches, Z., & Hausman, J. A. (1984). *Patents and r&d: Is there a lag?* (Tech. rep.). National Bureau of Economic Research.

Hall, B. H., & Ziedonis, R. H. (2001). The patent paradox revisited: An empirical study of patenting in the us semiconductor industry, 1979-1995. *rand Journal of Economics*, 101–128.

Hasseldine, J., Salama, A. I., & Toms, J. S. (2005). Quantity versus quality: The impact of environmental disclosures on the reputations of uk plcs. *The British Accounting Review*, *37*(2), 231–248.

Hausman, J. A., Hall, B. H., & Griliches, Z. (1984). Econometric models for count data with an application to the patents-r&d relationship.

Hazen, T. L. (2020). Social issues in the spotlight: The increasing need to improve publicly-held companies' csr and esg disclosures. *U. Pa. J. Bus. L.*, *23*, 740.

Herbohn, K., Walker, J., & Loo, H. Y. M. (2014). Corporate social responsibility: The link between sustainability disclosure and sustainability performance. *Abacus*, *50*(4), 422–459.

Herremans, I. M., Akathaporn, P., & McInnes, M. (1993). An investigation of corporate social responsibility reputation and economic performance. *Accounting, Organizations and Society*, *18*(7-8), 587–604.

Holgersson, M. (2013). Patent management in entrepreneurial smes: A literature review and an empirical study of innovation appropriation, patent propensity, and motives. *R&D Management*, *43*(1), 21–36.

Hummel, K., & Schlick, C. (2016). The relationship between sustainability performance and sustainability disclosure–reconciling voluntary disclosure theory and legitimacy theory. *Journal of Accounting and Public Policy*, *35*(5), 455–476.

Hunt, R. M. (2010). Business method patents and us financial services. *Contemporary Economic Policy*, *28*(3), 322–352.

Ioannou, I., & Serafeim, G. (2017). The consequences of mandatory corporate sustainability reporting. *Harvard Business School research working paper*, (11-100).

Jaeger, R. G., & Halliday, T. R. (1998). On confirmatory versus exploratory research. *Herpetologica*, S64–S66.

Jensen, F. V., et al. (1996). *An introduction to bayesian networks* (Vol. 210). UCL press London.

Kornbrot, D. (2005). Correlation. *Encyclopedia of statistics in behavioral science*, *1*, 398–400.

Kortum, S., & Putnam, J. (1997). Assigning patents to industries: Tests of the yale technology concordance. *Economic Systems Research*, *9*(2), 161–176.

Krueger, P., Sautner, Z., Tang, D. Y., & Zhong, R. (2021). The effects of mandatory esg disclosure around the world. *European Corporate Governance Institute–Finance Working Paper*, (754), 21–44.

Kuzma, E., Padilha, L. S., Sehnem, S., Julkovski, D. J., & Roman, D. J. (2020). The relationship between innovation and sustainability: A meta-analytic study. *Journal of Cleaner Production*, *259*, 120745.

La Belle, M. M., & Schooner, H. M. (2013). Big banks and business method patents. *U. Pa. J. Bus. L.*, *16*, 431.

Landi, G., & Sciarelli, M. (2018). Towards a more ethical market: The impact of esg rating on corporate financial performance. *Social Responsibility Journal*, *15*(1), 11–27.

Landry, R., Amara, N., & Lamari, M. (2002). Does social capital determine innovation? to what extent? *Technological Forecasting and Social Change*, *69*(7), 681–701.

Li, F., & Polychronopoulos, A. (2020). What a difference an esg ratings provider makes. *Research Affiliates*, *15*.

Liu, H., & Lyu, C. (2022). Can esg ratings stimulate corporate green innovation? evidence from china. *Sustainability*, *14*(19), 12516.

López, L. E., & Roberts, E. B. (2002). First-mover advantages in regimes of weak appropriability: The case of financial services innovations. *Journal of Business Research*, *55*(12), 997–1005.

Lopez-de-Silanes, F., McCahery, J. A., & Pudschedl, P. C. (2020). Esg performance and disclosure: A cross-country analysis. *Singapore Journal of Legal Studies*, (Mar 2020), 217–241.

Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics & Management Strategy*, *20*(1), 3–41.

Mansfield, E. (1986). Patents and innovation: An empirical study. *Management Science*, *32*(2), 173–181.

Maraut, S., Dernis, H., Webb, C., Spiezia, V., & Guellec, D. (2008). The oecd regpat database: A presentation.

Matos, P. (2020). Esg and responsible institutional investing around the world: A critical review.

Mize, T. D., Doan, L., & Long, J. S. (2019). A general framework for comparing predictions and marginal effects across models. *Sociological Methodology*, *49*(1), 152–189.

Morbey, G. K. (1988). R&d: Its relationship to company performance. *Journal of Product Innovation Management*, *5*(3), 191–200.

Morikawa, M. (2019). Innovation in the service sector and the role of patents and trade secrets: Evidence from japanese firms. *Journal of the Japanese and International Economies*, *51*, 43–51.

Nohria, N., & Gulati, R. (1996). Is slack good or bad for innovation? *Academy of Management Journal*, *39*(5), 1245–1264.

O'donovan, G. (2002). Environmental disclosures in the annual report: Extending the applicability and predictive power of legitimacy theory. *Accounting, Auditing & Accountability Journal*, *15*(3), 344–371.

OECD. (2004). *Oecd patent indicators*. https://www.oecd.org/science/inno/33776061.pdf

OECD. (2022). Oecd han database. https://www.wipo.int/edocs/mdocs/cws/en/cws_wk_ge_16/cws_wk_ge_16_eapo.pdf

Parris, T. M., & Kates, R. W. (2003). Characterizing and measuring sustainable development. *Annual Review of Environment and Resources*, *28*(1), 559–586.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan kaufmann.

Porter, M., Serafeim, G., & Kramer, M. (2019). Where esg fails. *Institutional Investor*, *16*(2).

Rijksoverheid. (2022). Brexit - eu-vk akkoord: Nieuwe regels. Retrieved November 16, 2022, from https://www.rijksoverheid.nl/onderwerpen/brexit/brexit-stand-van-zaken

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592.

Russel, F. (2023). Industry classification benchmark (equity) - ftse russell. https://research.ftserussell.com/products/downloads/ICB_Rules_new.pdf

Schiederig, T., Tietze, F., & Herstatt, C. (2012). Green innovation in technology and innovation management–an exploratory literature review. *R&D Management*, *42*(2), 180–192.

Sezen, B., & Cankaya, S. Y. (2013). Effects of green manufacturing and eco-innovation on sustainability performance. *Procedia-Social and Behavioral Sciences*, *99*, 154–163.

Stiglitz, J. E. (2002). Information and the change in the paradigm in economics. *American Economic Review*, *92*(3), 460–501.

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, *20*(3), 571–610.

Taliento, M., Favino, C., & Netti, A. (2019). Impact of environmental, social, and governance information on economic performance: Evidence of a corporate 'sustainability advantage'from europe. *Sustainability*, *11*(6), 1738.

Tan, H.-T., Ying Wang, E., & Zhou, B. (2014). When the use of positive language backfires: The joint effect of tone, readability, and investor sophistication on earnings judgments. *Journal of Accounting Research*, *52*(1), 273–302.

Tang, H. (2022). The effect of esg performance on corporate innovation in china: The mediating role of financial constraints and agency cost. *Sustainability*, *14*(7), 3769.

Textor, J., Van der Zander, B., Gilthorpe, M. S., Liśkiewicz, M., & Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: The r package 'dagitty'. *International Journal of Epidemiology*, *45*(6), 1887–1894.

Toms, J. S. (2002). Firm resources, quality signals and the determinants of corporate environmental reputation: Some uk evidence. *The British Accounting Review*, *34*(3), 257–282.

UNGC. (2004). Who cares wins: Connecting financial markets to a changing world.

Van Buuren, S. (2018). *Flexible imputation of missing data*. CRC press.

Veefkind, V., Hurtado-Albir, J., Angelucci, S., Karachalios, K., & Thumm, N. (2012). A new epo classification scheme for climate change mitigation technologies. *World Patent Information*, *34*(2), 106–111.

Verhoef, J. M., & Boveng, P. L. (2007). Quasi-poisson vs. negative binomial regression: How should we model overdispersed count data? *Ecology*, *88*(11), 2766–2772.

Weesie, J., et al. (2000). Seemlingly unrelated estimation and the cluster-adjusted sandwich estimator. *Stata Technical Bulletin*, *9*(52).

Xu, J., Liu, F., & Shang, Y. (2020). R&d investment, esg performance and green innovation performance: Evidence from china. *Kybernetes*, *50*(3), 737–756.

Yin, J., & Wang, S. (2018). The effects of corporate environmental disclosure on environmental innovation from stakeholder perspectives. *Applied Economics*, *50*(8), 905–919.

Yu, E. P.-y., Van Luu, B., & Chen, C. H. (2020). Greenwashing in environmental, social and governance disclosures. *Research in International Business and Finance*, *52*, 101192.

Zapf, D., Dormann, C., & Frese, M. (1996). Longitudinal studies in organizational stress research: A review of the literature with reference to methodological issues. *Journal of Occupational Health Psychology*, *1*(2), 145.

Zhang, Q., Loh, L., & Wu, W. (2020). How do environmental, social and governance initiatives affect innovative performance for corporate sustainability? *Sustainability*, *12*(8), 3380.

Zheng, M., Feng, G.-F., Jiang, R.-A., & Chang, C.-P. (2022). Does environmental, social, and governance performance move together with corporate green innovation in china? *Business Strategy and the Environment*.

A
Appendix A: DAG approach explained

In the DAG approach, it is attempted to create a systematic representation of the underlying structural causal model (SCM) and the effects between the variables based upon expert knowledge (Textor et al., 2016). Thereafter, the DAG helps determine which variables need to be controlled for to allow for the estimation of causal effects based upon a set of rules (Textor et al., 2016). The DAG is a network wherein the nodes represent variables and the edges represent unidirectional causal relationships (Ben-Gal, 2008). In turn, the DAG highlights which statistical dependencies are to be expected, which is highly contingent of the assumption that the DAG is accurate (Jensen et al., 1996). In doing so, it provides insight into which variables to control for by looking at triplets of nodes and assessing their interrelationships (Elwert & Winship, 2014). Three kind of triplets provide reasoning for whether to control for the third variable (Z) in estimating the causal effect of ESG (X) on Innovation (Y) (Rohrer, 2018), which are shown in figure A.1. To simplify, only forks should be controlled for unless a researcher is solely interested in direct effects in which case a chain should also be controlled for. First, in the chain structure ESG influences Z and this in turn influences Innovation (known as mediation) (Rohrer, 2018). Here, conditioning on Z blocks the transmission of (some of) the causal information of how ESG influences Innovation (Pearl, 1988). This is detrimental to the correct estimation of the causal effect and should be avoided. Second, in the fork structure Z causes both ESG and Innovation (known as confounding) (Rohrer, 2018). This results in the transmission of non-causal information if it is not controlled for (Rohrer, 2018). Therefore, it is key to control for a confounder. Finally, in a collider structure ESG and Innovation are not correlated but both cause Z (Rohrer, 2018). Clearly, there is no causal effect between ESG and Innovation. However, if this is controlled for, spurious association between ESG and Innovation is introduced also known as collider bias, which should be prevented (Dablander, 2020). To summarize, controlling for mediators and colliders has undesirable effects. In the control variables section, the SCM is extended from ESG and Innovation to include additional variables and their expected relationships.

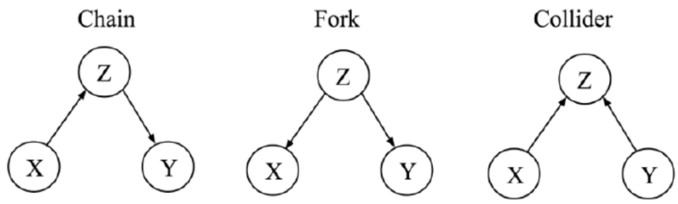

Figure A.1: Triplet types in DAGS (Dablander, 2020) with added type titles based on Rohrer (2018)

B

Appendix B: Reclassification of industries explained

This appendix describes how this paper reclassifies industries into low, medium and high propensity categories. This paper consults existing research on patenting propensities and primarily bases the reclassification on work by (Arundel & Kabla, 1998). Arundel and Kabla (1998) analysed several industries on their patenting propensity and puts forward a propensity score per industry. This propensity score represents the weighted percentage of service and product innovations that are patented within each industry. The industry categories between their work and this paper do not directly match. However, further investigation elucidates that the FTSE industry categories retrieved from Factset actually represent supersectors (Russel, 2023). The overarching industry classifications more accurately represent the categories found in Arundel and Kabla (1998). As an exception, the automobiles industry is still represented as the supersector, because this is directly represented in Arundel and Kabla (1998). Table B.1 shows how the overarching industries are linked to the industries in Arundel and Kabla (1998) and shows the average patenting propensity. Based on the patenting propensity the industries are reclassified into low (0-20), medium (20-30) and high (30+) patenting propensity categories. The financial and real estate industries are not represented in Arundel and Kabla (1998). However, several studies show that these industries have a low patenting propensity (Boscaljon et al., 2006; Hunt, 2010; Kortum & Putnam, 1997; La Belle & Schooner, 2013; López & Roberts, 2002).

Industry (supersector)	Overarching industry	Prop. Cat.	Propensity	Link industry Arundel and Kabla (1998)
Technology	Technology	High	39,45	Communication equipment, electrical equipment, office & computing equipment
Telecommunications	Telecommunications	Medium	27,1	Communication equipment and transport & telecom services
Health Care	Health Care	High	74	Pharmaceuticals
Banks	Financials	Low		No
Insurance	Financials	Low		No
Financial Services	Financials	Low		No
Real Estate	Real Estate	Low		No
Automobiles and Parts	Consumer Discretionary	Medium	25,2	Automobiles
Consumer Products and Services	Consumer Discretionary	Low	14,26	Transport and telecom, other transport equipment, textiles & clothing
Media	Consumer Discretionary	Low	14,26	Transport and telecom, other transport equipment, textiles & clothing
Retail	Consumer Discretionary	Low	14,26	Transport and telecom, other transport equipment, textiles & clothing
Travel and Leisure	Consumer Discretionary	Low	14,26	Transport and telecom, other transport equipment, textiles & clothing
Food, Beverage and Tobacco	Consumer Staples	Medium	25,3	Food beverages and tobacco
Personal Care, Drug and Grocery Stores	Consumer Staples	Medium	25,3	Food beverages and tobacco
Construction and Materials	Industrials	High	40,45	Electrical equipment, machinery, rubber and plastics, glass & clay & ceramics
Industrial Goods and Services	Industrials	High	40,45	Electrical equipment, machinery, rubber and plastics, glass & clay & ceramics
Basic Resources	Basic Materials	High	34,3	Mining, chemicals and basic metals
Chemicals	Basic Materials	High	34,3	Mining, chemicals and basic metals
Energy	Energy	Medium	25,1	Petroleum refining
Utilities	Utilities	Medium	26,7	Power utilities

Figure B.1: Reclassification scheme based upon (Arundel & Kabla, 1998)