YILAN CUI

Business School, The University of Queensland

St Lucia, Brisbane, Australia 4067 Email: y.cui2@uq.edu.au

PhD Candidate in Finance (Jul 2022 – June 2027 (expected))

Research interest: corporate governance, ESG, ESG rating, and DEI

Supervisors: Dr Nicolas Eugster and Dr Hoang Luong

Review Committee: Dr Dewan Rahman (chairman), A. Prof. Kelvin Tan (reader), and Dr Hasibul Chowdhury (reader)

Referees:

Dr Nicolas Eugster: Senior Lecturer in Finance at The University of Queensland

Email: n.eugster@business.uq.edu.au

Dr Hoang Luong: Lecturer in Finance at The University of Queensland

Email: h.luong@business.uq.edu.au

EMPLOYMENT

ACADEMIC TEACHING

Aug 2022 – Now The University of Queensland

Casual tutor: currently teaching Corporate Finance and Financial Management for Decision Makers; taught Intermediate Financial Accounting

Sep 2021 -- July 2023 Ulster University (Northern Ireland) & Dongbei University of Finance and Economics (China)

Lecturer: Taught accounting, finance, and introductory economics courses for an international program jointly run by the two universities (online delivery due to COVID)

Sep 2019 -- June 2021 Melbourne Polytechnic

Lecturer: Taught accounting and law courses (company law and business law/hospitality law)

Oct 2015 -- Sep 2019 International Finance College, Shanghai Lixin Accounting and Finance University

Lecturer: Taught accounting courses

Feb 2010 -- Nov 2013 The University of Queensland

Casual Tutor in Business School; PASS Leader in Economics School

INDUSTRY AND OTHER

Sep 2023 – Feb 2024 Business School, The University of Queensland

Member (volunteer) of the HDR Development Working Group

Oct 2002 – Jun 2006 Frasers Property (Shanghai, China)

Administrative/Marketing & Sales

Jul 2001 – Jul 2002 Shanghai Skynet Software Development Co., Ltd

Translator

EDUCATION

Jul 2022 – Now The University of Queensland

PhD Program in Finance

Jul 2020 – Jan 2022 The University of Queensland

Graduate Diploma in Research Methods

Jul 2008 – Jul 2010 The University of Queensland

Master of Commerce (Accounting and Finance)

Sep 1997 – Jul 2001 Shandong University

Bachelor of Arts

Can Board Gender Diversity Resolve ESG Rating Disagreement?

Yilan Cui, Nicolas Eugster, Hoang Luong *

Abstract

This paper studies ESG rating disagreement with a focus on its antecedents by investigating whether board gender diversity (BGD) can resolve ESG rating disagreement. Using the U.S. data, we find a negative association between BGD and ESG rating disagreement—that is, greater gender diversity on boards is linked to higher agreement between ESG ratings. Our findings are robust to a battery of sensitivity checks. To address endogeneity concerns, we use instrumental variable (IV) and quasi-natural experiment difference-in-differences (DiD) approaches, and the results consistently point to a causal effect of BGD on ESG rating disagreement. We also document a moderating role of BGD in the relationship between ESG rating disagreement and two firm-level characteristics: ESG controversies and conservatism. Furthermore, we find that BGD's effect on ESG rating disagreement operates through improvements in firms' ESG performance, managerial ability, and voluntary disclosure.

Keywords: CSR, ESG, ESG rating, ESG disagreement, ESG divergence, ESG dispersion, ESG discrepancy, DEI, diversity, board gender diversity, female director

1. Introduction

Recently, growing attention has been paid to the lack of consistency in environmental, social, and governance (ESG) ratings—often referred to as ESG rating disagreement, which captures the divergence in scores assigned by different ESG rating agencies.[2] Research shows that correlations between ESG ratings are generally low (Berg et al., 2022; Gibson Brandon et al., 2021). This lack of alignment has raised concerns among both policymakers and academic researchers, particularly as ESG ratings have become a widely used tool among investors. Institutional shareholders, in particular, rely on them for equity screening and evaluating the ESG performance of their portfolio firms. A CFA study (2020) of 2,800 investment

* Yilan Cui, PhD student at The University of Queensland, y.cui2@uq.edu.au; Nicolas Eugster, The University of Queensland, n.eugster@business.uq.edu.au; Hoang Luong, The University of Queensland, h.luong@business.uq.edu.au. The authors thank Kelvin Tan (reader of the PhD Review Committee) and Hasibul Chowdhury (reader of the PhD Review Committee) for helpful comments.

[2] We use ESG, CSR (corporate social responsibility), and sustainability interchangeably. Similarly, we use disagreement, divergence, dispersion, and discrepancy to describe differences across ESG ratings.

practitioners found that 85% of institutional investors have already incorporated ESG factors into their investment decisions, largely in response to client demand.

A substantial body of literature has explored the consequences of ESG rating disagreement, examining its effects on capital markets (Avramov et al., 2022; Ehling et al., 2023; Gibson Brandon et al., 2021; Serafeim & Yoon, 2023) and on firm behavior (Hu et al., 2023; Mao et al., 2024). However, relatively few studies have investigated the determinants of ESG rating disagreement. To the best of our knowledge, only two studies—Christensen et al. (2022) and Kimbrough et al. (2024)—have focused on this question, examining the role of firms' ESG disclosures and reporting mixed findings.

In this study, we investigate whether specific firm-level factors—in particular, board characteristics—can help explain and potentially reduce ESG rating disagreement. Our motivation builds on the taxonomy proposed by Berg et al. (2022), who identify three primary sources of ESG rating disagreement: scope (38%), measurement (56%), and weight (6%). According to their framework, *scope divergence* arises when ESG rating agencies select different sets of attributes to evaluate a firm's ESG performance; *measurement divergence* occurs when agencies use different indicators to assess the same attribute; and *weight divergence* reflects differences in how agencies prioritize and weight the importance of various attributes.

Among these, *scope* and *weight* are largely driven by the discretion of rating agencies, leaving firms with limited influence. In contrast, *measurement*—the main contributor to disagreement—can be indirectly shaped by firms through their ESG performance and disclosure practices. This suggests that firm- or board-level characteristics may play a significant role in explaining ESG rating divergence. Furthermore, Berg et al. (2022) highlight a rater effect, whereby the evaluation of a firm's ESG profile is influenced by the rater's overall perception of the firm, shaped in part by the firm's information environment. This further supports the idea that internal governance factors, especially those affecting disclosure quality, may play a role.

Among governance characteristics, we focus on board gender diversity (BGD) for two key reasons. First, there has been a growing push—through regulation, social movements (e.g., MeToo), and institutional investor campaigns (e.g., State Street's "Fearless Girl")—to increase female representation on boards. These pressures have prompted changes in board composition

in many jurisdictions, with recent research documenting the impact of both public awareness (Giannetti & Wang, 2023) and institutional investor pressure (Gormley et al., 2023).

Second, a well-established literature shows that gender-diverse boards positively influence firm outcomes, including ESG performance and disclosure quality. For example, Gul et al. (2011, 2013) show that gender diversity improves the firm's information environment and enhances analysts' earnings forecast accuracy. Kimbrough et al. (2024) find that voluntary ESG disclosure reduces ESG rating disagreement, particularly when reports are assured or follow Global Reporting Initiative (GRI) standards. Since voluntary disclosure reduces information asymmetry and disagreement among capital market participants, gender-diverse boards—through their improved monitoring and emphasis on transparency—may help reduce ESG rating disagreement as well.

This effect could be especially important when firms face ESG controversies, which increase ambiguity and subjectivity in interpretation. While controversies introduce complexity into disclosures and complicate assessment by rating agencies (Schiemann & Tietmeyer, 2022), boards with stronger monitoring capabilities—including those with greater gender diversity—are better positioned to manage such challenges (Kuzey et al., 2024; Khalid et al., 2024).

Given the increasing importance of ESG reporting and the demonstrated role of BGD in enhancing information environments, we ask whether board gender diversity can facilitate the transmission of information to the market—thereby reducing ESG information asymmetry and, in turn, ESG rating disagreement. This question remains largely unexplored, and this study aims to fill that gap.

To test our hypothesis, we construct a panel dataset of 6,836 firm-year observations from 2009 to 2019, covering 1,000 unique U.S. listed firms. We measure ESG rating disagreement as the absolute difference in ESG scores provided by two leading ESG rating agencies: Sustainalytics and Refinitiv. Our main explanatory variable is board gender diversity (BGD), defined as the proportion of female directors on the board. Using a pooled OLS regression model with year and either firm or industry fixed effects and controlling for a wide range of firm- and board-level characteristics, we find a negative association between BGD and ESG rating disagreement—that is, greater gender diversity on boards is linked to higher agreement between ESG ratings. Specifically, a one standard deviation increase in BGD corresponds to a 1.176-point reduction in ESG rating disagreement under the firm fixed effects specification,

which represents a 7.61% decrease relative to the sample mean and 11.41% of the sample standard deviation.

Our findings are robust to propensity score matching (PSM) analysis, several alternative model specifications, including Newey-West standard errors, generalized linear models (GLM), and weighted least squares (WLS). They also hold when using alternative measures for our dependent and independent variables, and when lagging the independent variables. To address endogeneity concerns, we use instrumental variable (IV) and quasi-natural experiment difference-in-differences (DiD) approaches. Across all these methods, our results consistently point to a causal effect of BGD on ESG rating disagreement.

In the final analyses, we examine the mechanisms behind this relationship. We document a moderating role of BGD in the relationship between ESG rating disagreement and two firm-level characteristics: ESG controversies and conservatism. Furthermore, we find that BGD's effect on ESG disagreement operates through improvements in firms' ESG performance, managerial ability, and voluntary ESG disclosure.

We make several important contributions to the literature. First, to the best of our knowledge, this is the first study examining corporate boards as a setting to explore firm-level determinants of ESG rating disagreement. While most existing research focuses on the consequences of ESG rating disagreement, our study contributes to the emerging literature by shifting the focus to its antecedents. Second, we extend the literature on board diversity by providing novel evidence on the benefits of greater gender diversity at the board level—specifically, its role in reducing ESG rating disagreement. Third, our findings offer practical implications for firms, particularly those with poor ESG performance and/or low managerial ability. By increasing board gender diversity, firms can reduce uncertainty related to ESG assessments, even in the absence of standardized ESG ratings or regulatory oversight. Finally, we contribute to policy discussions by highlighting the role of board gender diversity in improving the non-financial information environment. Our results suggest that mandatory board gender quotas may help mitigate information asymmetries between management and external stakeholders by promoting more consistent ESG disclosures.

The remaining parts are organized as follows. Section 2 reviews the literature and develops a hypothesis. Section 3 explains data and research methodology. Section 4 reports empirical results. Section 5 investigates the mechanism. Section 6 conducts robustness tests. Section 7 concludes.

2. Literature Review and Hypothesis Development

2.1 ESG rating Disagreement

One of the earliest studies that question the ESG rating convergence is by Chatterji et al. (2016), which analyses the social ratings across six raters: KLD, Asset4, Innovest, DJSI, FTSE4Good, and Calvert. An influential study, Berg et al. (2022), investigates potential drivers of ESG rating divergence using data from six key ESG rating agencies: Kinder, Lydenberg, and Domini (KLD), MSCI, Sustainalytics, Refinitiv (Asset4), Moody's ESG (Vigeo-Eiris), and S&P Global (RobecoSAM), and documents that correlations at the ESG level range between 0.38 and 0.71, a wide divergence. They find that ESG rating divergence is driven mainly by measurement (56%) and scope (38%), both of which far outweigh weight (6%). Among these drivers, measurement divergence is driven in part by a "rater effect", whereby the same rater tends to give a firm high scores in all other categories if it has already given it a high score in one category.

Most of the existing literature focuses on the impacts of ESG rating disagreement. One stream focuses on capital markets. For example, Gibson Brandon et al. (2021) document a positive relation between ESG rating disagreement and stock returns while Ehling et al. (2023) provide evidence that ESG rating disagreement yields significantly different optimal portfolios. The other stream focuses on firm behaviors. For example, Hu et al. (2023) find that the (E) disagreement makes Chinese listed firms more likely to greenwash through agency cost and information opacity. Mao et al. (2024) show that high (low) ESG rating disagreement leads to a high (low) degree of earnings management. Zhou et al. (2024) find that the positive effect of ESG rating disagreement is stronger in firms with greater media coverage and resource advantages of independent directors.

Despite such evidence, the literature investigating the determinants of ESG rating disagreement remains sparce. At present, one of the most influential papers that examines the determinants of ESG rating disagreement is by Christensen et al. (2022), who show that ESG disclosure (as measured by the Bloomberg ESG disclosure score) is positively associated with the degree of future ESG rating disagreement. Another study, by Kimbrough et al. (2024), uses textual analysis to measure ESG disclosure and finds that lengthy ESG reports can reduce ESG rating disagreement, but reports with positive tones or sticky words can increase it. Furthermore, the effect of ESG disclosure is stronger for firms whose ESG reports are certified by a third party or for firms that show greater compliance with GRI reporting standards.

2.2 Board Gender Diversity

PwC's Annual Corporate Directors Surveys over the past five years emphasize that board gender diversity plays a fundamental role—ranking higher than other diversity types such as age, race, and socioeconomic background—in fostering diversity of thoughts within the boardroom.

2.2.1 Traits of Women

The earliest literature on traits of women can be traced back to the *gender socialization theory* in a book by Nancy J. Chodorow published in 1978 and *social role theory* in a book by Alice H. Eagly published in 1987. We refer to these two theories in their latest editions as Chodorow (2023) and Eagly (2013), respectively. The *gender socialization theory* posits that gender identity is shaped in childhood and women are raised being more nurturing and sympathetic. Treated differently in their childhoods, women internalize gender identity through their parents' model roles and social interactions. The *social role theory* presents sex differences in the social context. This theory posits that women's responsibility of raising children and taking care of household work makes them more concerned for others. Studies also show that women rank altruism as more important than men (Dietz et al., 2002), are more risk-averse than men (Eckel & Grossman, 2008; Jianakoplos & Bernasek, 1998). Ho et al. (2015) find that female CEOs are more conservative.

2.2.2 Board Gender Diversity and Information Transparency/Asymmetry

A strand of the literature on information transparency has documented a positive impact of board gender diversity on firms' information environments (Upadhyay & Zeng, 2014). Studies in the finance literature have shown that firms with greater gender diversity are associated with higher earnings quality (Srinidhi et al., 2011) and a lower likelihood of financial restatement (Abbott et al., 2012), attributing these effects to female directors' better monitoring roles. Lai et al. (2023) find that firms' disclosure environment is improved through decreased earnings management following the appointment of female directors by institutional investors.

Gul et al. (2011) examine how board gender diversity improves firms' information environments through, for example, increased public voluntary disclosure in large firms (as proxied by the frequency of "other events" category disclosure in the 8-K filings) and accelerated collection of private information in small firms. In a subsequent study, Gul et al. (2013) focus on analyst forecast accuracy and find that board gender diversity is positively (negatively) related to analysts' earnings forecast accuracy (dispersion), arguing that board

diversity improves transparency and accuracy of financial reports, thus allowing earnings to be forecasted more accurately.

In the non-financial context, some studies focus on the relationship between board gender diversity and firms' ESG disclosure. Using the carbon emissions data from the Carbon Disclosure Project (CDP), Liao et al. (2015) find that increased board gender diversity leads to a greater voluntary disclosure of greenhouse gas (GHG) emissions. Katmon et al. (2019) document a positive effect of board gender diversity on the quality of ESG disclosure in emerging markets.

2.3 Hypothesis

We conjecture that the positive impact of board gender diversity might extend to other important aspects of firms' outcomes. Our argument rests on the premise that a more gender-diverse board is more likely to take actions that they believe can enhance firm value and reputation through improving the quality and timeliness of ESG disclosures, and/or mitigating firm risks, especially ESG-related litigation and reputation risks. Such improved ESG disclosure can reduce information asymmetry between firms and outside stakeholders, thereby mitigating ESG rating disagreement.

First, board gender diversity promotes a better monitoring role, which contributes to resolving ESG rating disagreement. Adams and Ferreira (2009) find that female directors exert greater monitoring effort with better attendance records and demand greater accountability from managers. Recent research (e.g., Kang et al., 2022) shows that board gender diversity enhances cognitive diversity, which improves board monitoring effectiveness and firm value. |The trickle-down effect of Gould et al. (2018) implies that managers monitored by a more effective board are more likely to act ethically, engage in prosocial behavior, and adopt stronger ESG-related practices (e.g., separate ESG committees for ethics or social responsibility (Isidro & Sobral, 2015)).

Firms with board gender diversity, and hence a better monitored board, are also more likely to have greater voluntary disclosure (Liao et al., 2015) and a higher-quality ESG disclosure (Katmon et al., 2019). A recent study by Cai et al. (2024) finds that the percentage of female directors on the board is positively related to the firm's diversity target disclosure.

Given the values that board gender diversity adds to a firm's ESG-related practices and quality of ESG disclosure, a more effective board can improve its information transparency, which in

turn facilitates ESG agencies' rating process and thus reduces ESG rating disagreement across different raters.

Second, board gender diversity brings more perspectives, discussions, and communications, all of which contribute to the resolution of ESG rating disagreement. According to the resource dependence theory, a firm should benefit from board diversity because a diversified board can introduce more insights, networks, and work and life experiences, thus making the firm more competitive (Barney, 1991; Pfeffer, 2019; Pfeffer & Salancik, 2015; Rixom et al., 2023). Rixom et al. (2023) argue that resource dependence theory is central to investigating the board efficacy as it focuses on firms' specific resources and the association with firm performance.

Research shows that diverse perspectives and opinions add value in a way that fosters a learning and inclusive environment, which enhances firms' problem-solving capabilities. Compared to men, women present different decision-making and leadership styles (Peterson & Philpot, 2007; Trinidad & Normore, 2005). Kang et al. (2022) document that having female directors or directors with diverse qualifications and skillsets on the board can improve cognitive diversity. Aggarwal et al. (2019) document that cognitive style diversity has an indirect, positive effect on team learning through collective intelligence. Duchek et al. (2020) develops a theoretical framework and proposes that diversity affects organizations' sensemaking and problem-solving capabilities.

On the other hand, board discussions add value in a way that facilitates informed decision-making. By virtue of their traits, women demand different information, which allows them to improve the depth and breadth of board discussions (Srinidhi et al., 2011). Such effective discussions help firms better understand ESG-related issues, thus making better and more informed decisions.

Furthermore, communications add value in a way that involves more collaboration and improves transparency. Hillman et al. (2007) argue that women have more potential to collaborate with different constituencies than men. They also posit that women provide important ties with various external organizations, which can reduce uncertainties and help obtain information and resources. More effective communications help firms disseminate information to various external parties, especially ESG rating agencies. Therefore, the firm's transparency is improved through ESG disclosure.

Prior studies show that voluntary disclosure reduces disagreement among capital market intermediaries. For instance, conference calls improve analyst forecast accuracy and reduce

forecast dispersion (Bowen et al., 2002), while disclosing accounting policies decreases uncertainty around earnings forecasts (Hope, 2003). Kimbrough et al. (2024) find that voluntary ESG disclosure lowers ESG rating disagreement, particularly when it is assured by a third-party or follows the GRI standards.

Taken together, with benefits from gender-diversified boards, firms are better equipped with enhanced problem-solving capabilities, informed decision making, and improved transparency, all of which enable firms to better tackle complicated ESG issues, make better ESG-related decisions, and adopt best ESG practices. Importantly, firms' better communication and improved information transparency facilitate the ESG rating process, thus leading to a reduction in ESG rating disagreement.

Third, board gender diversity mitigates firm risks, which contributes to the resolution of ESG rating disagreements. Since females are more risk-averse (Eckel & Grossman, 2008; Jianakoplos & Bernasek, 1998; Sugarman & Straus, 1988), they are more sensitive to uncertainties and negative effects such as ESG controversies, which give rise to divergent beliefs and expectations. Such divergence in beliefs and expectations can, in turn, introduce obfuscation and subjective judgments, which can lead to a significant leeway for interpretation, thus resulting in greater disagreement among ESG rating agencies.

Prior research documents that the level of ESG controversies in firms is associated with their information asymmetry (Cui et al., 2018; Schiemann & Tietmeyer, 2022). Schiemann and Tietmeyer (2022) argue that controversies and ESG disclosure are closely linked: under the legitimacy theory, firms facing serious controversies are more likely to increase disclosure in response to social or political pressure. However, this reactive and often complex disclosure can exacerbate information asymmetry, making it even harder for ESG agencies to assess and compare firms accurately because a firm's real ESG performance is entangled with potential greenwashing behaviors that some ESG rating agencies can detect, but others cannot.

Nonetheless, since firms with more risk-averse female directors are more likely to respond to ESG controversies, they tend to take proactive initiatives to mitigate corresponding risks. For example, women tend to reduce perceived environmental risks (Bord & O'Connor, 1997; Fukukawa et al., 2007) and firms with powerful female executives face fewer lawsuits (Adhikari et al., 2019).

In light of the analyses above, we argue that female directors' risk-aversion trickles down through oversight to managers, inducing the latter to adopt better risk management practices to

avoid ESG-related reputation, litigation, or physical environment risks, thus making more ESG-compliant decisions. As a result, these ESG-compliant decisions can lead to fewer ESG controversies and uncertainties, thus minimising leeway for interpretation by ESG raters and hence reducing subjective elements and ESG rating disagreement.

In the context of financial reporting, the monitoring function of gender-diverse boards has been well documented in the literature. For instance, board gender diversity facilitates the collection of private information in small firms (Gul et al., 2011) and enhances the accuracy of analysts' earnings forecasts (Gul et al., 2013). Cao et al. (2024) also find that executive female independent directors improve firms' financial reporting quality. Therefore, we expect, in the ESG reporting context, that board gender diversity plays a similar role in reducing ESG information asymmetry, and propose the following main hypothesis:

H1: Board gender diversity (BGD) is negatively associated with ESG rating disagreement, meaning that high BGD is linked to greater agreement in ESG ratings.

3. Data and Methodology

3.1 Data and Sample Selection

We obtain corporate board data from Refinitiv. We also source data on CSR voluntary reporting, CSR committees, and management scores to measure firm's managerial ability from this database. We use I/B/E/S and Thomson Reuters Institutional (13f) Holdings databases to measure analyst coverage and institutional ownership, respectively. We obtain firm fundamental data from Compustat North America.

ESG data are collected from two leading ESG rating agencies: Sustainalytics and Refinitiv. Although our sample is restricted by the intersection of both ESG data by both Sustainalytics and Refinitiv rating providers, our sample firms are still representative of U.S. listed firms since they represent most of the major US companies (i.e., those included in the S&P 500 (2009-2019), Russell 1000 (2011-2019), and Russell 2000 (2017-2019) indexes.

We start with the universe of Compustat North America firms. After merging these data with Sustainalytics ESG rating, I/B/E/S analyst following, Thomson Reuters Institutional Ownership, and Refinitiv ESG data via multiple linking tables, we end up with a sample of 7,522 firm-year observations from 2009 to 2019 for 1,207 unique firms. As is standard in the literature, we exclude financial firms (SIC 6000-6999), which typically are highly leveraged

and are subject to stricter regulations. Such a filter leaves us with a final sample of 1,000 unique firms and 6,836 firm-year observations. Table A2 in the Appendix shows the sample selection process.

3.2 Dependent Variable - ESG Rating Disagreement

Our dependent variable, ESG rating disagreement, is measured by the absolute difference in ESG scores provided by two leading ESG rating providers, Sustainalytics and Refinitiv, whose ESG scores are widely used by investors and academic researchers. A reason for choosing these raters is because we aim to minimize the impact of different methodologies used by different ESG rating providers. Sustainalytics and Refinitiv share more in their methodologies.

Berg et al. (2022) use Sustainalytics rating as a benchmark because it has the highest correlations with other five ratings. They document that the ESG score correlation between Sustainalytics and Refinitiv is 0.67. Further, ESG scores rated by Sustainalytics and Refinitiv are the closest in mean (Sustainalytics 56.4; Refinitiv 50.9) and median (Sustainalytics 55; Refinitiv 53.15). To facilitate a systematic comparison of the scope, they analyse each rater's methodologies and construct a taxonomy which shows that Refinitiv ranks first with the most individual indicators of 282, followed by Sustainalytics with 163 indicators. They also find that the divergence between Sustainalytics and Refinitiv is mainly driven by measurement (66%) and weight (22%), with scope contributing only 12%.

Sustainalytics rating started in 2009 with a coverage of various index constituents (e.g., S&P 500, Russell 1000, Russell 3000, MSCI World and MSCI ACWI). However, it ended its legacy data in 2019, with an enhanced replacement product called "ESG Risk Rating" (ESG RR) data deployed in December 2018. The ESG legacy data was retired as Sustainalytics was considering possible upcoming regulations on ESG rating industry. Due to differences in methodologies between Sustainalytics legacy ESG rating and ESG Risk Rating, we choose not to include their ESG Risk Rating data in our study to ensure consistency. The initial Sustainalytics ESG data from 2009 to 2019 consists of 45,094 firm-year observations of 7,924 unique firms.

Refinitiv rating started in 2002 with an initial focus on firms included in the S&P 500 and NASDAQ 100 and an increased coverage of firms from the Russell 1000 in 2011 and Russell 2000 in 2017. We select its ESG coverage of US listed firms from 2009 to 2019 and initially obtain 15,696 firm-year observations of 2,874 unique firms.

Following Kimbrough et al. (2024), we use the absolute difference[3] between the ESG rating scores by the two ESG raters to measure ESG rating disagreement. Specifically, *ESG_Divergence* is the absolute difference between Sustainalytics's overall ESG scores and Refinitiv's overall ESG Combined scores (ESG controversies incorporated)[4]. The overall ESG scores provided by Sustainalytics incorporate firms' controversies[5]. We use this measure (i.e., *ESG_Divergence*) as the main dependent variable because it better reflects the ESG rating divergence when both overall scores incorporate controversies (i.e., more similar rating methods), hence making the ratings more comparable. It also helps mitigate a concern about differences in methodologies adopted by ESG raters. In a robustness test, we also use a second measure of ESG rating divergence, called *ESG_Divergence2*, which is defined as the absolute difference between Sustainalytics' overall ESG scores (ESG controversies incorporated) and Refinitiv's overall ESG scores (ESG controversies *not* incorporated). For the sake of comparison, we also include and discuss *ESG_Divergence2* in our descriptive statistics.

3.3 Methodology

As in Christensen et al. (2022), we use the following ordinary least squares (OLS) regression model to estimate the effect of board gender diversity on ESG rating disagreement:

$$ESG_Divergence_{it} = \beta_0 + \beta_1 BoardGenderDiversity_{it} + \beta_k Controls + FE + \varepsilon_{it}, \quad (1)$$

where *ESG_Divergence* measures the ESG rating disagreement of firm *i* in year *t* across two leading ESG rating providers, Sustainalytics and Refinitiv. *BoardGenderDiversity* (BGD), the percentage of female board members, measures the board gender diversity of firm *i* in year *t*. The coefficient of interest is β_1, which measures the change in ESG rating disagreement in response to the change in board gender diversity. If it is negative and significant, then we find support for the main hypothesis that greater board gender diversity is associated with less ESG rating disagreement.

We control for a set of both firm and board characteristics that might affect ESG rating disagreement. For firm characteristics, we follow Christensen et al. (2022) to control for performance (*ROA*), capital structure (*LEV*), growth opportunities (*BTM*), firm size (*FirmSize*), analyst coverage (*AF*) and the institutional ownership (*IO*). We log-transform the variables *FirmSize*, *AF*, and *IO*. We expect to find a negative coefficient on *ROA*, *BTM*, and *FirmSize*

[3] We also use the standard deviation of ESG scores to measure ESG rating disagreement. It produces similar results for our baseline regression model (available under request).
[4] Both Sustainalytics and Refinitiv scores ranges from 0 to 100.
[5] Sustainalytics' ESG Rating Research Methodology 2017. www.sustainalytics.com

because financial performance, growth opportunities, and firm size generally improve a firm's capability to tackle ESG issues. The coefficient on *LEV*, however, can be either positive or negative since leverage not only can provide firms with more financial resources but can also pose financial risks to them. For *AF* and *IO*, we expect the coefficients to be negative as financial analysts and institutional investors tend to improve firms' information environments.

We also control for the percentage of female executive members (*ExecGenderDiversity*), which is expected to have a negative effect as it can improve management diversity. For board characteristics, we follow recent literature (Biswas et al., 2023; Giannetti & Wang, 2023) and include the percentage of independent board members (*BoardIndep*), CEO-Chairman duality (*CEO_Duality*), board tenure (*BoardTenure*) and board size (*BoardSize*). We use natural log for *BoardTenure* and *BoardSize*. The coefficients on *BoardIndep* and *BoardTenure* are expected to be negative as board independence and tenure can enhance board effectiveness, leading to less ESG rating disagreement. On the contrary, *CEO_Duality* is expected to be positive because a dual role of both CEO and board chairman is a sign of poor corporate governance, thus leading to a less effective board. The coefficient on *BoardSize*, however, can be either positive or negative as a smaller or oversized board might be less effective.

To control for time trends in ESG rating disagreement and fixed firm or industry[6] factors that may drive the ESG rating disagreement, we also include year and either firm or industry fixed effects, respectively. For comparison, we report regression results with both industry and firm fixed effects in most specifications. All continuous variables are winsorized at the 1st and 99th percentiles. In all regressions, we cluster standard errors at the firm level to allow arbitrary correlation of errors within firms. All variable definitions and data sources are presented in the Appendix.

3.4 Descriptive Statistics

Panel A of Table 1 reports summary statistics on ESG scores as rated by each agency. In our sample, Sustainalytics tends to rate higher for both overall ESG score and each underlying pillar (environmental, social and governance) than Refinitiv. The mean of *ESG_Sus* is 56.40, while the mean of *ESG_Refi* and *ESGCom_Refi* is 47.74 and 44.85, respectively. However, Sustainalytics tends to have smaller variance (standard deviation of 8.75) in overall score than Refinitiv (standard deviation of 20.16 for *ESG_Refi* and 18.53 for *ESGCom_Refi*).

[6] Fama French ffi17 is used as the sample size is not large.

[Insert Table 1 here]

Panel B of Table 1 provides descriptive statistics of regression variables. Although *ESG_Divergence* incorporates controversies by both raters (more similar rating methodologies), it has a slightly larger mean (15.45), i.e., larger disagreement, than the mean of *ESG_Divergence2* (14.64). Reasonably, as *ESG_Avg* incorporates controversies, it has a smaller mean (50.64) than *ESG_Avg2* (52.07) which does not incorporate controversies. Larger *ESG_Divergence* may suggest that ESG raters have larger disagreements when it comes to "controversies". Interestingly, a similar relation is found in the forecast accuracy context. Schiemann and Tietmeyer (2022) use RepRisk data to measure ESG controversy and find that firms with a higher ESG controversy exposure have higher analyst forecast errors.

The mean of *BoardGenderDiversity* suggests that, on average, 17.84% of directors on board are female. The median indicates that about half of the firms have less than 18.18% female board representation. And a majority (75%) of the firms have less than 25% female board representation. These statistics suggest that female director representation on the board is quite low.

The statistics of *ExecGenderDiversity* show that the percentage of female executive members is even lower with a mean of 14.10% and a median of 12.5%. It is worth noting that 25% of the firms have no female executive members. However, the level of board independence (*BoardIndep*) is quite high with an average percentage of independent board members of 80.78%. For *CEO_Duality*, only 25% of the firms separate the CEO from the role of chairman.

Panel C of Table 1 provides the mean difference for the regression variables by high/low ESG rating disagreement, based on the median of *ESG_Divergence*. As expected, the group of high ESG rating disagreement firms tend to have lower ESG performance (mean of 44.85 for *ESG_Avg* and 46.64 for *ESG_Avg2*) than the group of low ESG rating disagreement firms (mean of 56.42 for *ESG_Avg* and 57.50 for *ESG_Avg2*). For board gender diversity, similar with ESG performance, the group of high ESG rating disagreement firms tend to have lower board gender diversity (mean of 15.93 for *BoardGenderDiversity*) than the group of lower ESG rating disagreement firms (mean of 19.75 for *BoardGenderDiversity*). All other firm and board characteristics exhibit highly significant differences between both groups.

We also conduct frequency analyses by year and by industry (Fama-French 17 classification) for our main variables of interest, namely, *ESG_Divergence*, *ESG_Avg*, and *BoardGenderDiversity*. Similarly, the number of observations of these three variables stably

increased over time until reaching a peak in 2015 and stabilized at a lower level afterward. However, the trends of the mean are different. ESG rating disagreement has an overall decreasing trend, indicating that ESG rating dispersion across different raters is narrowing over time in our sample. On the contrary, ESG performance develops in the opposite way, with an increasing trend, i.e., improving over time. It is worth noting that a large reduction in ESG rating disagreement in 2016 coincided with a sudden jump in ESG performance. For board gender diversity, its mean exhibits a very similar (i.e., parallel) trend with that of ESG performance, including also a sudden jump in 2016 over time in our sample period. Details are reported in Table A3 in the Appendix.

Finally, Table 2 presents the correlation matrix for the regression variables. All variable pairs exhibit correlations below 0.7 and the low VIF values[7] further suggest that our baseline regression does not suffer from multicollinearity issue. The correlation between *ESG_Divergence* and *BoardGenderDiversity* is -0.283, indicating a negative association between our explanatory and dependent variables. The correlation results in Column 1 show consistency with our initial expectation.

[Insert Table 2 here]

4. Empirical Results

4.1 Baseline Regression

Table 3 reports the baseline regressions controlling for year and industry fixed effects in Column 1 and year and firm fixed effects in Column 2. We find in both models that the coefficient on board gender diversity is negative and statistically significant at the 1% level, suggesting that firms with greater board gender diversity are associated with lower ESG divergence (i.e., higher ESG rating agreement among the raters). The effect of board gender diversity is also economically nontrivial. For example, Column 1 shows that the coefficient on board gender diversity (*BoardGenderDiversity*) is 0.116, implying that a one standard deviation increase in board gender diversity is associated with a reduction of 1.218 in ESG rating disagreement, which translates into a 7.88% increase relative to the sample mean and 11.83% of the sample standard deviation.[8] The firm fixed effects model (Column 2) yields a

[7] The VIF values range from 1.06 (*CEO_Duality*) to 2.72 (*FirmSize*).
[8] Variation of *ESG_Divergence*=Coefficient*SD of *BoardGenderDiversity*= 0.116*10.5=1.218; Variation of *ESG_Divergence*/Mean of *ESG_Divergence*:1.218/15.45=7.88%; Variation of *ESG_Divergence*/SD of *ESG_Divergence*:1.218/10.30=11.83%.

similar result that a one standard deviation increase in board gender diversity corresponds to an ESG rating disagreement reduction of 1.176, equivalent to a 7.61% increase relative to the sample mean and 11.41% of the sample standard deviation. These findings confirm our hypothesis.

[Insert Table 3 here]

Comparing the explanatory power of the two models, it appears that the firm fixed effects model (adjusted R2 of 0.533) explains a much greater amount of the variation in ESG rating disagreement than the industry fixed effects model (adjusted R^2 of 0.172). This suggests that much of the ESG rating disagreement about a firm is very persistent over time (Christensen et al., 2022).

We also find that board independence, separation of CEO from chairman, and long board tenure are also negatively associated with ESG rating disagreement, especially when including firm fixed effects. However, executive gender diversity has no effect on ESG rating disagreement, and board size has a weak negative relation to ESG rating disagreement only with industry fixed effects.

4.2 Endogeneity Tests

While our results have so far provided empirical support for our main hypothesis that greater gender diversity is associated with less ESG rating disagreement, board gender diversity can still be endogenous. For example, the choice of board gender diversity might suffer from self-selection bias because firms with less ESG rating disagreement might choose to have greater board gender diversity in the first place, which implies that board gender diversity might not necessarily bear any relationship with ESG rating disagreement. Furthermore, even though our results are robust to the inclusion of firm fixed effects, any time-varying, unobservable variables omitted from the regression could bias our inference if they are correlated with board gender diversity. Another possibility is reserve causality in which firms with less ESG rating disagreement and stronger ESG performance could attract more female directors to the board. In this section, we address these concerns by employing two approaches, namely, instrumental variables (IV) regressions and a quasi-natural experiment based on exogenous shocks to board gender diversity.

4.2.1 Instrumental Variables

Since board gender diversity could be endogenous, in this section, we construct an instrument variable for it and use the two-stage least squares (2SLS) approach as an identification strategy. An ideal instrument should capture the variation in board gender diversity that is exogenous to ESG rating disagreement. Since the exclusion restriction of an instrument variable is inherently untestable, we need to conceptually motivate it.

Following Knyazeva et al. (2013), we use two instruments: the focal firm's headquarter state average board gender diversity (*StateBGD*) and SIC-industry average board gender diversity (*IndBGD*). *StateBGD* and *IndBGD* are measured as the average percent of female directors on the boards across the firms in the same state where firms' headquarters are located and in the same industry using the SIC classification, respectively.

As board structure can be affected by the local supply of directors and state-level political orientation (e.g., democratic party versus republican party) is related to firms in the state (Knyazeva et al., 2013; Giannetti & Wang, 2023), we expect that firms in a more gender equality-friendly state are associated with a more gender diversified board. However, state-average board gender diversity is unlikely to directly affect ESG rating disagreement.

Moreover, since firms in the same industry are more likely to adopt similar gender diversity strategies and practices due to industry peer pressure (Hillman et al., 2007; Knippen et al., 2019), industry gender diversity is likely to be correlated with the gender diversity of an individual firm in the same industry. However, there is no reason to believe that industry-average gender diversity directly affects a firm's ESG rating disagreement.

Table 4 presents the results of the 2SLS regression. To check the relevance of our instrument variables, we report the first-stage results in columns 1 and 3. As shown, the coefficients on both instruments are statistically significant, indicating that both instruments capture well the variation in board gender diversity. The Cragg-Donald Wald F-statistics are significantly large, thus rejecting the null hypothesis of weak instruments. Columns 2 and 4 report the second-stage results, where we use the fitted values from the first-stage regressions. We find that the coefficients on instrumented BGD are negative and significant, further reinforcing our baseline results that board gender diversity impacts ESG rating disagreement.

[Insert Table 4 here]

4.2.2 Quasi-Natural Experiment

In this section, we use a quasi-natural experiment in a difference-in-difference (DiD) framework to further address potential endogeneity associated with board gender diversity. Following related literature, we use the "Fearless Girl" campaign by the Big Three institutional shareholders launched in 2017 as an exogeneous shock to board gender diversity.

Over the last few years, firms have been under growing pressure from institutional shareholders. For example, State Street first initiated a campaign known as the "Fearless Girl" in March 2017 in which they pressured firms with no female directors to add at least one female director to their boards[9]. State Street threatened that firms with inadequately diverse boards would have an adverse vote on re-electing the nominating/governance committee chair (Gormley et al., 2023). In July 2017, BlackRock followed suit by making similar announcements, which were formally stated in its revised proxy voting guidelines in February 2018. They expect at least two female directors on every board and may vote against nominating/governance committee members if female directors are not adequate.[10] In August 2017, Vanguard's CEO wrote an open letter and announced that they also value meaningful progress of board diversity which is informative to engagement and voting.[11]

We use the "Big Three" campaigns in 2017 as an exogenous shock to board gender diversity to further tease out the impact of board gender diversity on ESG rating disagreement. We argue that since the explicit goal of these campaigns is to increase board gender diversity, it cannot affect ESG rating disagreement other than through board gender diversity. Prior research highlights the role of institutional pressure in shaping firms' strategic responses (Goodstein, 1994) and board gender diversity (Browder, 1995; Hillman et al., 2007). More recent evidence by Gormley et al., (2023) confirms the effectiveness of the Big Three campaigns in promoting board gender diversity. Therefore, if we observe a significant decrease in ESG rating disagreement following these campaigns, the exogenous nature of the shock strengthens the case for a causal relationship between board gender diversity and ESG rating disagreement by mitigating endogeneity concerns.

We identify the treatment group (*Big3_FemaleLow*) as firms subject to the influence of the Big Three campaign—specifically, firms with ownership by at least one of the Big Three

[9] State Street Global Advisors, 2017. SSGA's guidance on enhancing gender diversity on boards, March 7 guidance-on-enhancing-gender-diversity-on-boards.pdf (kpmg.com)
[10] BlackRock, 2018. Proxy voting guidelines for U.S. securities, February.
[11] An open letter to directors of public companies worldwide 2017VanguardOpenLettertoBoards.pdf (wlrk.com)

institutional investors and low female board representation (i.e., no more than one female director).[12] *Post2016* is a dummy variable equal to one for the years 2017, 2018, and 2019, capturing the period after the campaigns were initiated. Finally, our main variable of interest, *Big3_FemaleLow*Post2016*, is the interaction term between *Big3_FemaleLow* and *Post2016* and captures the effect of the Big Three campaigns on ESG divergence for firms with low female board representation in their portfolios.

Panel A of Table 5 reports the results. The coefficient on the interaction *Big3_FemaleLow*Post2016* is significantly negative with both industry (Column 1) and firm fixed (Column 2) effects, suggesting that firms with Big three institutional ownership and low female board representation experience a reduction in ESG rating disagreement following the "Fearless Girl" campaigns by the Big Three institutions in Year 2017. The coefficient on *Big3_FemaleLow* in Column 1 is significantly positive, confirming that firms with low female board representation and Big Three ownership had higher ESG rating disagreement before the campaigns.

[Insert Table 5 here]

Panel B shows a trend analysis on low board female representation with Big Three institutional ownership and ESG rating disagreement from 2014 to 2019. The coefficients on the interaction terms (*Big3_FemaleLow*Year)* prior to 2017 are mainly insignificant, while they are negative and significant from 2017 onwards, particularly in the firm fixed effect model. These findings support establishing a causal relation between board gender diversity and ESG rating disagreement and mitigate the endogeneity concern.

5. Mechanism

5.1 Moderating Role of Board Gender Diversity (BGD)

First, we look at the role of board gender diversity in the relation between firms' ESG controversies and their ESG rating disagreement. A firm's exposure to ESG controversies can make ESG ratings more difficult to determine, as ESG controversies involve complication and obfuscation in the disclosure. Schiemann and Tietmeyer (2022) find that analyst forecast errors increase with firms' exposure to ESG controversies, proxied by RepRisk score, suggesting that

[12] *Big3_FemaleLow* is constructed as the interaction of *Big3* and *FemaleLow*. *Big3* is a dummy variable equal to one if a firm has institutional ownership by State Street, Black Rock, or Vanguard, and zero otherwise. Data on institutional ownership are sourced from Thomson Reuters Institutional (13f). *FemaleLow* is a dummy variable equal to one if there is none or only one female director on the firm's board, and zero otherwise.

such controversies complicate analysts' forecasting tasks. Similarly, Cui et al. (2018), using KLD data, show that more CSR strengths and fewer CSR concerns reduce information asymmetry.

Another strand of literature has shown that more effective boards play a positive role in mitigating ESG controversies. For example, Kuzey et al. (2024) find that greater board monitoring intensity reduces ESG controversies in risk-prone firms, as measured by Refinitiv controversy scores, though it may not prevent such ESG controversies entirely. In the Chinese market, Khalid et al. (2024), using data from WIND, report that the presence of female directors helps mitigate ESG-related disputes. Therefore, we expect that firms with higher levels of board gender diversity tend to face lower levels of ESG controversies and may have different levels of ESG rating disagreement. To test this, we conduct additional tests by interacting BGD with *ESGControversies*.

The *ESGControversies* variable, obtained from Refinitiv, captures a firm's exposure to ESG-related controversies and adverse events reported in global media. The measure is continuous, ranging from 0 (high controversy) to 100 (no controversy), and is based on 23 distinct ESG controversy topics. Importantly, the construction of this score accounts for market capitalization bias, recognizing that firms with larger market capitalizations are more likely to attract media scrutiny and, consequently, are more frequently associated with reported controversies.

Panel A of Table 6 presents the moderating role of board gender diversity for the relation between ESG controversies and ESG rating disagreement. In both Columns 1 and 2, the coefficients on the interaction term *BoardGenderDiversity*ESGControversies* are significantly negative and consistent with the signs of the coefficients on *ESGControversies*, which are also significantly negative. This suggests that board gender diversity strengthens the association between ESG controversies and ESG rating disagreement. That is, the negative relation between ESG controversy score and ESG rating disagreement is more pronounced when a firm's board gender diversity level is high.

[Insert Table 6 here]

Then we look at the role of board gender diversity in the relation between firms' conservatism and their ESG rating disagreement. As women are more risk-averse (Eckel & Grossman, 2008; Jianakoplos & Bernasek, 1998) and more conservative (Ho et al., 2015), they are more sensitive to adverse events. Firms with these directors tend to have fewer uncertainties and risks

(Adhikari et al., 2019; Bord & O'Connor, 1997; Fukukawa et al., 2007), and hence less ESG rating disagreement through reduced interpretation by ESG rating agencies.

Accordingly, we expect that firms with higher levels of board gender diversity tend to exhibit greater conservatism and may experience different levels of ESG rating disagreement. Cash holdings are commonly used in the literature as a proxy for corporate conservatism, with prior studies documenting a positive relationship between the two (Lin et al., 2018; Louis et al., 2012; Manoel & Moraes, 2022). Therefore, we examine whether the effect of board gender diversity varies with firms' conservatism by interacting *BGD* with *CashHolding.*

Panel B of Table 6 presents the moderating role of board gender diversity for the relation between firm-level conservatism and ESG rating disagreement. In both Columns 1 and 2, the coefficients on the interaction term *BoardGenderDiversity*CashHolding* are significantly positive, while the coefficients on both *BoardGenderDiversity* and *CashHolding* are significantly negative. This shows that the firm-level conservatism weakens the board gender diversity effect on ESG rating disagreement. It might be interpreted in the other way that the board gender diversity effect on ESG disagreement is more pronounced when a firm is less conservative. This indicates a substitution effect between board gender diversity and firms' conservatism on reducing ESG disagreement.

5.2 Channel Analysis

We further explore potential channels which drive our findings on the effect of board gender diversity on ESG rating disagreement.

First, we examine ESG performance as a channel. Much of the existing evidence in the literature shows that corporate board diversity is positively associated with ESG performance (Beji et al., 2021; Hafsi & Turgut, 2013; Isidro & Sobral, 2015; Liu, 2018). A recent study by Mallidis et al. (2024) finds that board gender diversity improves firms' environmental and social performance but negatively affects ESG controversies performance. Using board gender quotas as natural experiments, other studies also document a positive relation between board gender diversity and ESG performance (Barroso et al., 2024; Dyck et al., 2023; Ginglinger & Raskopf, 2023).

This can be explained by even earlier research. Beutel and Marini (1995) find that females are more altruistic (i.e., to express concern and responsibility for others and/or social issues) and less likely to be materialistic. Some literature suggests that females are more concerned with environmental issues than males (Mainieri et al., 1997; Wehrmeyer & McNeil, 2000). These

traits, together with their better monitoring role (Adams & Ferreira, 2009), make board gender-diversified firms have better ESG performance.

According to signal theory, high ESG performers tend to issue voluntary CSR reporting and/or have CSR committees. Therefore, this improves firms' information transparency and helps reporting users collect information. A high-quality disclosure makes ESG raters, key CSR report users, exercise less subjective interpretation as they are confronted with less complexity and obfuscation. This, in turn, facilitates the ESG rating progress and eventually helps different raters narrow the dispersion and achieve more agreement. The empirical results of our sub-sample based on high/low ESG performance, as well as other studies (Christensen et al., 2022; Kimbrough et al., 2024), support our conjecture that ESG performance serves as a channel for the effect of board gender diversity on ESG rating disagreement.

We, in the first step, regress *ESG_Avg* on *BoardGenderDiversity* to estimate the relation between ESG performance on board gender diversity (Columns 1 and 2) and then add *ESG_Avg* to the baseline regression (Columns 3 and 4). Panel A of Table 7 presents the result of the test for the channel of ESG performance. As shown in Columns 1 and 2, ESG performance (*ESG_Avg*) is positively associated with board gender diversity. When adding *ESG_Avg* to the baseline regression, the coefficient on board gender diversity becomes insignificant with industry fixed effects (Column 3) and weaker, both in magnitude and significance, with firm fixed effects (Column 4), compared to the coefficients in the baseline regression. This indicates that ESG performance plays a mediating role in the effect of board gender diversity on ESG rating disagreement.

[Insert Table 7 here]

Second, we focus on managerial ability as a channel. One stream of the literature investigates the relationship between director diversity and manager diversity, showing that there is a "trickle-down effect" between board diversity and executive (top management) diversity, i.e., increasing diversity at the board level leads to higher diversity at the management level. For example, Gould et al. (2018) find a positive, but nonlinear, relationship between female board member diversity and female executive diversity, a trickle-down effect, which is consistent with earlier evidence of a linkage between female director diversity and female manager diversity (Matsa & Miller, 2011, 2013). Other recent studies, for example, Biswas et al. (2023); (Cai, Grennan, et al., 2024) also find similar evidence that female director diversity is positively associated with female senior manager diversity.

Another stream of the literature examines the relationship between board gender diversity and managerial ability. For example, Baghdadi et al. (2023) find that board gender diversity improves firms' managerial ability, especially in firms where female directors assume a monitoring role. Moreover, improved managerial ability benefits firms' ESG performance (Edmans et al., 2023; Welch & Yoon, 2023). Altunbas et al. (2022) find that the percentage of female managers is negatively associated with CO_2 emissions.

We argue that the spillover of board gender diversity to managerial diversity helps improve managerial ability. As a more diversified board is more likely to have or appoint diversified managers who tend to adopt diversity policies and/or pursue diversity initiatives, managers' ability, particularly to tackle ESG issues, is improved, issuing voluntary ESG disclosure (Liao et al., 2015) and having ESG committees (Isidro & Sobral, 2015). These enable these firms to tackle complicated ESG issues more effectively and lead to enhanced ESG performance.

Importantly, managers with higher ability are more sensitive to new issues (e.g., ESG rating disagreement) and quicker to respond to the consequences through proactive strategies and practices to minimize negative effects on the firm. They are more likely to adopt best governance practices and comply with ESG reporting frameworks (e.g., GRI, TCFD). These are critical for firms' actual ESG performance and their information transparency, which can facilitate ESG agencies' rating process and hence help reduce the ESG rating disagreement.

We measure managerial ability by *MgtScore*, which is collected from Refinitiv and referred to as a firm's commitment and effectiveness towards following best practice corporate governance principles. Panel B provides the result of the test for the channel of managerial ability. In Columns 1 and 2, board gender diversity is significantly positive to managerial ability (*MgtScore*). Also, with both industry (Column 3) and firm (Column 4) fixed effects, the coefficients on board gender diversity become weaker in terms of both magnitude and significance when *MgtScore* is added to the model compared to the baseline results. This indicates that managerial ability plays a mediating role in the effect of board gender diversity on ESG rating disagreement.

Finally, we look at CSR reporting as a channel. The literature provides evidence of a positive relationship between board gender diversity and firms' ESG disclosure (Katmon et al., 2019; Liao et al., 2015). Thus, the enhanced ESG disclosure leads to an improved information environment, which, in turn, reduces disagreement among ESG rating agencies. A recent study by Cai et al. (2024) investigates how board gender diversity enhances firm diversity

performance and finds that the percentage of female directors is positively related to firms' diversity target disclosure. That is, firms with more gender diverse boards are more likely to disclose their diversity targets. They also show that the percentage of female directors is positively related to a firm's diversity performance.

CSR_Reporting is a dummy variable that equals one if a firm issues a separate CSR report or a CSR section in its annual report.[13] Voluntary ESG reporting can improve firms' information transparency, help information collection by ESG agencies, and hence reduce ESG rating disagreement. Panel C shows the result of the test for the channel of CSR reporting. In Columns 1 and 2, board gender diversity is significantly positively related to CSR reporting. When *CSR_Reporting* is added to the baseline regression (Columns 3-4), the coefficients on board gender diversity become weaker compared to the baseline results. This suggests a channel role of CSR reporting.

We conduct additional Sobel-Goodman mediation tests to assess the role of the three channels. The total mediation effects are as follows: ESG performance (54.4%), managerial ability (51.5%), and CSR reporting (5.7%). These results provide consistent support for the presence of channel effects. Furthermore, sub-sample analyses based on high vs. low ESG performance, high vs. low managerial ability, and the presence of a stand-alone voluntary ESG report or a dedicated section in the annual report also yield consistent findings. As reported in Appendix Table A4, the initial effect of *BGD* on *ESG_Divergence* is primarily observed among firms with low ESG performance, low managerial ability, and those without formal ESG Reporting.

6. Robustness and Other Tests

We conduct a battery of robustness tests to further check whether our baseline results remain robust using propensity score matching (PSM), alternative models, alternative measures for dependent and independent variables, and applying lagging variables.

6.1 Propensity Score Matching (PSM)

First, we use a propensity score matching (PSM) approach to alleviate self-selection bias in the choice of board gender diversity, and the bias due to potential unobservable omitted variables affecting both dependent and independent variables. We estimate a logit model in which the

[13] We also examine the presence of a CSR committee as a potential channel, given that such committees are typically associated with stronger governance practices and more effective ESG issue management, which may help reduce ESG rating disagreement. The results are similar to those obtained for *CSR_Reporting* and are therefore not reported for brevity. Moreover, *CSR_Reporting* and *CSR_Committee* are also strongly correlated.

dependent variable *BGD_High* is a dummy variable equal to one if *BGD* is above the median, and zero otherwise.[14] The independent variables include all controls from the baseline model, along with year and industry fixed effects. Firms with high board gender diversity (*BGD_High* =1) are in the treatment group, while those with low board gender diversity (*BGD_High* =0) are in the control group. Based on the propensity scores estimated from the logit regression, we use the nearest-neighbour matching technique to pair each treatment firm with a control firm with the closest propensity scores (absolute difference smaller than 0.005). This procedure produces a PSM-matched sub-sample of 2,589 firm-year observations.

Panel A of Table 8 shows the mean differences in firm and board characteristics between treatment and control groups in the PSM-matched sample. As we can see, the PSM match has successfully removed all systematic differences between treated and control groups so that treated and control firms are similar in characteristics that potentially can affect ESG rating disagreement.

[Insert Table 8 here]

Panel B reports the average treatment effect (ATT), showing the differences between the treatment and control groups before and after the PSM approach. The average effect remains significantly different in the PSM-matched sub-sample, i.e., board gender diversity still has a strong effect on solving the ESG rating disagreement after matching and ruling out the self-selection bias.

Finally, Panel C presents the results from re-running the baseline regression on the PSM-matched sample. The coefficients on board gender diversity remain negative and strongly significant for both industry (Column 1) and firm (Column 2) fixed effect models, indicating a consistent effect of board gender diversity on ESG rating disagreement after alleviating self-selection bias concern.

6.2 Alternative Models

We use alternative models as the Ordinary Least Squares (OLS) model requires some assumptions such as linear distribution, constant variance (no heteroscedasticity), and uncorrelation with each other in error terms. However, the relation can be exponential or curvilinear, and there might be a serial correlation in the time series model. Also, an unequal

[14] In non-tabulated results, we also use *Female_Dummy*—a dummy variable equal to one if there is at least one female on the board and zero otherwise—as an alternative dependent variable in the logit model and find similar results, albeit with a smaller matched sub-sample.

variance may be present. These violations can make our estimations in the OLS model biased. So, we use first a generalized linear model (GLM) to mitigate the concern about potential non-linear relations, then a weighted least squares (WLS) regression to get a better fit for potential unequal variances, and finally, a Newey-West model to address heteroscedasticity and autocorrelation in the error terms.

Table 9 demonstrates the estimations from using these three alternative models. The coefficients on board gender diversity in all specifications are negative and significant, mitigating the concerns about estimation bias due to potential violation of assumptions in the OLS model.

[Insert Table 9 here]

6.3 Alternative Measures

We employ four alternative proxies to ensure the robustness of our measure for board gender diversity (BGD). First, we use *BGD_High*, a dummy variable equal to one if board gender diversity is high (i.e., above the median for each year). Second, we use *Female_Dummy*, a dummy variable if there is at least one female on the board. Third, we use *FemaleHigh,* a dummy variable if board female representation is high (i.e., at least two female directors on board). Finally, we use *Female_Num*, the log of number of females on board.

Panel A of Table 10 provides results of using these alternative measures for board gender diversity (BGD) with industry (Columns 1,3, 5, and 7) and firm (Columns 2, 4, 6, and 8) fixed effects. In all specifications, the coefficients on *BoardGenderDiversity* are consistently negative and significant, supporting our baseline results. This confirms that our findings are not sensitive to an alternative definition of BGD and that the effect of board gender diversity on ESG rating disagreement still holds.

[Insert Table 10 here]

We then replace *ESG_Divergence* with our second measure for ESG disagreement, *ESG_Divergence2*, and re-estimate the regression. This measure reflects a factor of different rating methods adopted by two ESG agencies, i.e., this divergence is the absolute difference of the overall ESG score by Sustainalytics (controversies incorporated) and Refinitiv (controversies not incorporated).

Panel B shows that, with both industry (Column 1) and firm (Column 2) fixed effects, the coefficients on board gender diversity are consistently negative with slightly lower magnitude.

Moreover, the explanatory power has improved in both industry and firm fixed effects models, with adjusted R^2 of 0.207 and 0.582, respectively.

As *ESG_Divergence2* may have some disagreement elements caused by different rating methodologies adopted by different raters, the consistently strong results provide evidence that board gender diversity can reduce the ESG rating disagreement, even though different rating methods exist across different raters.

6.4 Lag Independent Variables

One of the assumptions of OLS is that all independent variables are uncorrelated with the error term. This assumption can be violated when there are omitted variables, reverse causality, or measurement error, and consequently, the estimates are biased. Using lagged independent variables helps mitigate these concerns. We re-estimate our baseline model by applying lagged independent variables in the regression model. Results are presented in Table 11.

[Insert Table 11 here]

The significant coefficients with both industry (Column 1) and firm (Column 2) fixed effects, although with a slightly smaller magnitude, support a consistently strong negative relation between board gender diversity and ESG rating disagreement.

6.5 Other Test: E/S/G Pillar

We also examine the effect of board gender diversity on the rating disagreement in each pillar environment (E), social (S) and governance (G). Refinitiv does not provide the breakdowns for each E, S and G pillar for their overall ESG combined score, which is used to calculate our main dependent variable *ESG_Divergence*. As a result, in our baseline model, we are unable to run a baseline regression for each E, S, and G pillar. However, for the second measure of ESG rating disagreement (i.e., *ESG_Divergence2*), each E, S, and G pillar breakdowns are provided, so we can calculate the disagreement on each pillar and estimate the effect of board gender diversity on *E_Divergence*, *S_Divergence*, and G*_Divergence*.

Table 12 shows that all coefficients for board gender diversity are negative. However, only the environment and governance pillars are statistically significant, while the social pillar is insignificant.

[Insert Table 12 here]

7. Conclusion

With an increasing reliance on ESG scores by investors, it is important to discover factors that drive the rating disagreement across different ESG agencies. Nevertheless, the reasons behind the ESG rating disagreement have been under explored. This study aims to bridge this gap by investigating potential factors driving ESG rating disagreement from the perspective of the rated firms.

We examine how board characteristics, and more particularly board gender diversity, influence the ESG rating disagreement. We find that there is a significant negative association between board gender diversity and ESG rating disagreement. In other words, board gender diversity helps ESG raters achieve consensus. Our findings are robust to propensity score matching (PSM) analysis, alternative models, alternative measures for our dependent and independent variables, and the use of lag variables in our baseline regression. We also mitigate endogeneity issues by using instrumental variables and quasi-natural experiments in a difference-in-difference approach. These additional analyses support a causal effect of board gender diversity on ESG rating disagreement. Further, we document a moderating role of BGD in the relationship between ESG rating disagreement and two firm-level characteristics: ESG controversies and conservatism. Also, we find that the effect of BGD on ESG rating disagreement is through the channels of ESG performance, managerial ability, and CSR reporting.

Our study is important as it contributes to the scarce literature on the determinants of ESG rating disagreement, particularly from the angle of the rated firms. Moreover, the findings highlight the importance of board gender diversity and have practical implications on firms, especially for those with poor ESG performance and/or lower managerial ability. However, this study has a limitation of a small sample size, so caution might be exercised when generalizing to other settings.

Future research can extend to other firm or board characteristics that can explain ESG rating disagreement. For example, board independence also appears to have an effect. We encourage future research to explore in depth. Also, it is worthwhile to investigate the role of other types of diversity (e.g., diversity of directors' experience or expertise).

References:

Abbott, L. J., Parker, S., & Presley, T. J. (2012). Female board presence and the likelihood of financial restatement. *Accounting horizons*, *26*(4), 607-629.

Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, *94*(2), 291-309.

Adhikari, B. K., Agrawal, A., & Malm, J. (2019). Do women managers keep firms out of trouble? Evidence from corporate litigation and policies. *Journal of Accounting and Economics*, *67*(1), 202-225.

Aggarwal, I., Woolley, A. W., Chabris, C. F., & Malone, T. W. (2019). The impact of cognitive style diversity on implicit learning in teams. *Frontiers in psychology*, *10*, 112.

Altunbas, Y., Gambacorta, L., Reghezza, A., & Velliscig, G. (2022). Does gender diversity in the workplace mitigate climate change? *Journal of Corporate Finance*, *77*, 102303.

Avramov, D., Cheng, S., Lioui, A., & Tarelli, A. (2022). Sustainable investing with ESG rating uncertainty. *Journal of Financial Economics*, *145*(2), 642-664.

Baghdadi, G. A., Safiullah, M., & Heyden, M. L. (2023). Do gender diverse boards enhance managerial ability? *Journal of Corporate Finance*, *79*, 102364.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, *17*(1), 99-120.

Barroso, R., Duan, T., Guo, S. S., & Kowalewski, O. (2024). Board gender diversity reform and corporate carbon emissions. *Journal of Corporate Finance*, *87*, 102616.

Beji, R., Yousfi, O., Loukil, N., & Omri, A. (2021). Board diversity and corporate social responsibility: Empirical evidence from France. *Journal of Business Ethics*, *173*, 133-155.

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

Beutel, A. M., & Marini, M. M. (1995). Gender and values. *American sociological review*, 436-448.

Biswas, P. K., Chapple, L., Roberts, H., & Stainback, K. (2023). Board gender diversity and women in senior management. *Journal of Business Ethics*, *182*(1), 177-198.

Bord, R. J., & O'Connor, R. E. (1997). The gender gap in environmental attitudes: The case of perceived vulnerability to risk. *Social science quarterly*, 830-840.

Bowen, R. M., Davis, A. K., & Matsumoto, D. A. (2002). Do conference calls affect analysts' forecasts? *The Accounting Review*, *77*(2), 285-316.

Browder, D. (1995). Shareholders are valuing diversity. *DIRECTORS AND BOARDS-AMERICAN EDITION-*, *19*, 12-12.

Cai, W., Chen, Y., Rajgopal, S., & Azinovic-Yang, L. (2024). Diversity targets. *Review of Accounting Studies*, *29*(3), 2157-2208.

Cai, W., Grennan, J., & Qiu, L. (2024). Do Diverse Directors Influence DEI Outcomes? *Available at SSRN 4232434*.

Cao, Z., Upadhyay, A., & Zeng, H. (2024). Are all female directors equal? Incentives and effectiveness of female independent directors. *Journal of Banking & Finance*, *162*, 107110.

Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal*, *37*(8), 1597-1614.

Chodorow, N. J. (2023). *The reproduction of mothering: Psychoanalysis and the sociology of gender*. Univ of California Press.

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.
Cui, J., Jo, H., & Na, H. (2018). Does corporate social responsibility affect information asymmetry? *Journal of Business Ethics*, *148*, 549-572.
Dietz, T., Kalof, L., & Stern, P. C. (2002). Gender, values, and environmentalism. *Social science quarterly*, *83*(1), 353-364.
Duchek, S., Raetze, S., & Scheuch, I. (2020). The role of diversity in organizational resilience: a theoretical framework. *Business research*, *13*(2), 387-423.
Dyck, A., Lins, K. V., Roth, L., Towner, M., & Wagner, H. F. (2023). Renewable governance: Good for the environment? *Journal of Accounting Research*, *61*(1), 279-327.
Eagly, A. H. (2013). *Sex differences in social behavior: A social-role interpretation*. Psychology Press.
Eckel, C. C., & Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of experimental economics results*, *1*, 1061-1073.
Edmans, A., Flammer, C., & Glossner, S. (2023). *Diversity, equity, and inclusion*.
Ehling, P., Haukø Lundeby, S. R., & Qvigstad Sørensen, L. (2023). Portfolio Choice with ESG Disagreement: Customizing Sustainability through Direct Indexing. *Journal of Beta Investment Strategies*, *14*(3).
Fukukawa, K., Shafer, W. E., & Lee, G. M. (2007). Values and attitudes toward social and environmental accountability: A study of MBA students. *Journal of Business Ethics*, *71*, 381-394.
Giannetti, M., & Wang, T. Y. (2023). Public attention to gender equality and board gender diversity. *Journal of Financial and Quantitative Analysis*, *58*(2), 485-511.
Gibson Brandon, R., Krueger, P., & Schmidt, P. S. (2021). ESG rating disagreement and stock returns. *Financial analysts journal*, *77*(4), 104-127.
Ginglinger, E., & Raskopf, C. (2023). Women directors and E&S performance: Evidence from board gender quotas. *Journal of Corporate Finance*, *83*, 102496.
Goodstein, J. D. (1994). Institutional pressures and strategic responsiveness: Employer involvement in work-family issues. *Academy of management journal*, *37*(2), 350-382.
Gormley, T. A., Gupta, V. K., Matsa, D. A., Mortal, S. C., & Yang, L. (2023). The big three and board gender diversity: The effectiveness of shareholder voice. *Journal of Financial Economics*, *149*(2), 323-348.
Gould, J. A., Kulik, C. T., & Sardeshmukh, S. R. (2018). Trickle-down effect: The impact of female board members on executive gender diversity. *Human Resource Management*, *57*(4), 931-945.
Gul, F. A., Hutchinson, M., & Lai, K. M. (2013). Gender-diverse boards and properties of analyst earnings forecasts. *Accounting horizons*, *27*(3), 511-538.
Gul, F. A., Srinidhi, B., & Ng, A. C. (2011). Does board gender diversity improve the informativeness of stock prices? *Journal of Accounting and Economics*, *51*(3), 314-338.
Hafsi, T., & Turgut, G. (2013). Boardroom diversity and its effect on social performance: Conceptualization and empirical evidence. *Journal of Business Ethics*, *112*, 463-479.
Hillman, A. J., Shropshire, C., & Cannella Jr, A. A. (2007). Organizational predictors of women on corporate boards. *Academy of management journal*, *50*(4), 941-952.
Ho, S. S., Li, A. Y., Tam, K., & Zhang, F. (2015). CEO gender, ethical leadership, and accounting conservatism. *Journal of Business Ethics*, *127*, 351-370.
Hope, O. K. (2003). Accounting policy disclosures and analysts' forecasts. *Contemporary Accounting Research*, *20*(2), 295-321.

Hu, X., Hua, R., Liu, Q., & Wang, C. (2023). The green fog: Environmental rating disagreement and corporate greenwashing. *Pacific-Basin Finance Journal*, *78*, 101952.
Isidro, H., & Sobral, M. (2015). The effects of women on corporate boards on firm value, financial performance, and ethical and social compliance. *Journal of Business Ethics*, *132*, 1-19.
Jianakoplos, N. A., & Bernasek, A. (1998). Are women more risk averse? *Economic inquiry*, *36*(4), 620-630.
Kang, J.-K., Kim, S., & Oh, S. (2022). Does board demographic diversity enhance cognitive diversity and monitoring? *The Accounting Review*, *97*(6), 385-415.
Katmon, N., Mohamad, Z. Z., Norwani, N. M., & Farooque, O. A. (2019). Comprehensive board diversity and quality of corporate social responsibility disclosure: Evidence from an emerging market. *Journal of Business Ethics*, *157*, 447-481.
Khalid, F., Naveed, K., Sun, X., & Srivastava, M. (2024). Unveiling the influence of institutional quality on board gender diversity and corporate environmental, social, and governance disputes in China. *Business Ethics, the Environment & Responsibility*.
Kimbrough, M. D., Wang, X., Wei, S., & Zhang, J. (2024). Does voluntary ESG reporting resolve disagreement among ESG rating agencies? *European Accounting Review*, *33*(1), 15-47.
Knippen, J. M., Shen, W., & Zhu, Q. (2019). Limited progress? The effect of external pressure for board gender diversity on the increase of female directors. *Strategic Management Journal*, *40*(7), 1123-1150.
Knyazeva, A., Knyazeva, D., & Masulis, R. W. (2013). The supply of corporate directors and board independence. *The Review of Financial Studies*, *26*(6), 1561-1605.
Kuzey, C., Al-Shaer, H., Uyar, A., & Karaman, A. S. (2024). Do board monitoring and audit committee quality help risky firms reduce CSR controversies? *Review of Quantitative Finance and Accounting*, *63*(3), 1007-1045.
Lai, K. M., Khedmati, M., Gul, F. A., & Mount, M. P. (2023). Making honest men of them: Institutional investors, financial reporting, and the appointment of female directors to all-male boards. *Journal of Corporate Finance*, *78*, 102334.
Liao, L., Luo, L., & Tang, Q. (2015). Gender diversity, board independence, environmental committee and greenhouse gas disclosure. *The British accounting review*, *47*(4), 409-424.
Lin, C.-M., Chan, M.-L., Chien, I.-H., & Li, K.-H. (2018). The relationship between cash value and accounting conservatism: The role of controlling shareholders. *International Review of Economics & Finance*, *55*, 233-245.
Liu, C. (2018). Are women greener? Corporate gender diversity and environmental violations. *Journal of Corporate Finance*, *52*, 118-142.
Louis, H., Sun, A. X., & Urcan, O. (2012). Value of cash holdings and accounting conservatism. *Contemporary Accounting Research*, *29*(4), 1249-1271.
Mainieri, T., Barnett, E. G., Valdero, T. R., Unipan, J. B., & Oskamp, S. (1997). Green buying: The influence of environmental concern on consumer behavior. *The Journal of social psychology*, *137*(2), 189-204.
Manoel, A. A. S., & Moraes, M. B. d. C. (2022). Accounting conservatism and corporate cash levels: Empirical evidence from Latin America. *Corporate governance: An international review*, *30*(3), 335-353.
Mao, Z., Wang, S., & Lin, Y. E. (2024). ESG, ESG rating divergence and earnings management: Evidence from China. *Corporate Social Responsibility and Environmental Management*.

Matsa, D. A., & Miller, A. R. (2011). Chipping away at the glass ceiling: Gender spillovers in corporate leadership. *American Economic Review*, *101*(3), 635-639.
Matsa, D. A., & Miller, A. R. (2013). A female style in corporate leadership? Evidence from quotas. *American Economic Journal: Applied Economics*, *5*(3), 136-169.
Peterson, C. A., & Philpot, J. (2007). Women's roles on US Fortune 500 boards: Director expertise and committee memberships. *Journal of Business Ethics*, *72*, 177-196.
Pfeffer, J. (2019). Size and composition of corporate boards of directors: The organization and its environment. In *Corporate governance* (pp. 53-64). Gower.
Pfeffer, J., & Salancik, G. (2015). External control of organizations—Resource dependence perspective. In *Organizational behavior 2* (pp. 355-370). Routledge.
Rixom, J. M., Jackson, M., & Rixom, B. A. (2023). Mandating diversity on the board of directors: Do investors feel that gender quotas result in tokenism or added value for firms? *Journal of Business Ethics*, *182*(3), 679-697.
Schiemann, F., & Tietmeyer, R. (2022). ESG controversies, ESG disclosure and analyst forecast accuracy. *International Review of Financial Analysis*, *84*, 102373.
Serafeim, G., & Yoon, A. (2023). Stock price reactions to ESG news: The role of ESG ratings and disagreement. *Review of Accounting Studies*, *28*(3), 1500-1530.
Srinidhi, B., Gul, F. A., & Tsui, J. (2011). Female directors and earnings quality. *Contemporary Accounting Research*, *28*(5), 1610-1644.
Sugarman, D. B., & Straus, M. A. (1988). Indicators of gender equality for American states and regions. *Social Indicators Research*, *20*, 229-270.
Trinidad, C., & Normore, A. H. (2005). Leadership and gender: a dangerous liaison? *Leadership & Organization Development Journal*, *26*(7), 574-590.
Upadhyay, A., & Zeng, H. (2014). Gender and ethnic diversity on boards and corporate information environment. *Journal of Business Research*, *67*(11), 2456-2463.
Wehrmeyer, W., & McNeil, M. (2000). Activists, pragmatists, technophiles and tree-huggers? Gender differences in employees' environmental attitudes. *Journal of Business Ethics*, *28*, 211-222.
Welch, K., & Yoon, A. (2023). Do high-ability managers choose ESG projects that create shareholder value? Evidence from employee opinions. *Review of Accounting Studies*, *28*(4), 2448-2475.
Zhou, J., Lei, X., & Yu, J. (2024). ESG rating divergence and corporate green innovation. *Business Strategy and the Environment*, *33*(4), 2911-2930.

Table 1. Descriptive Statistics

In Panel A, *ESG_Sus* is the overall ESG score rated by Sustainalytics. This score incorporates controversies because Sustainalytics includes assessing firms' ESG performance for their level of involvement in major controversies or incidents. Refinitiv issues two types overall ESG score: overall ESG score (*ESG_Refi*) and overall combined ESG score (*ESGCom_Refi*). *ESGCom_Refi* incorporates controversies while *ESG_Refi* does not. In Panel A, *ESG_Divergence* is the absolute difference between Sustainalytics overall ESG scores and Refinitiv overall ESG Combined scores (ESG controversies incorporated). *ESG_Divergence2* is the absolute difference between Sustainalytics overall ESG scores and Refinitiv overall ESG scores. *ESG_Avg* is the average of overall ESG scores by Sustainalytics and overall ESG combined scores by Refinitiv. *ESG_Avg2* is the average of overall ESG scores by Sustainalytics and Refinitiv. ESG Combined scores do not have individual E/S/S pillar scores and therefore no individual average scores. *BoardGenderDiversity* is the percentage of female on the board.

Panel A: ESG Scores by Each Rating Agency

Variable	N	Mean	SD	p25	p50	p75
ESG_Sus	6836	56.40	8.750	49.37	55.00	62.62
E_Sus	6528	53.33	12.70	43.14	51.40	61.94
S_Sus	6528	54.88	10.74	46.75	53.67	62.00
G_Sus	6528	65.42	7.760	60.00	65.30	70.94
ESGCom_Refi	6836	44.85	18.53	30.37	43.92	58.87
ESG_Refi	6836	47.74	20.16	31.37	46.86	64.48
E_Refi	6836	38.47	28.92	9.360	38.32	63.21
S_Refi	6836	49.46	22.22	31.57	47.84	67.02
G_Refi	6836	53.17	22.38	35.99	54.63	71.49

Panel B: Descriptive Statistics of Regression Variables

Variable	N	Mean	SD	p25	p50	p75
ESG_Divergence	6836	15.45	10.30	6.770	14.16	22.82
ESG_Divergence2	6836	14.64	10.10	6.200	13.11	21.71
ESG_Avg	6836	50.64	12.55	40.47	49.79	59.98
ESG_Avg2	6836	52.07	13.46	40.80	51.04	63.18
BoardGenderDiversity	6824	17.84	10.50	11.11	18.18	25.00
ExecGenderDiversity	6819	14.10	12.16	0	12.50	22.22
BoardIndep	6803	80.78	11.87	75.00	83.33	90.00
CEO_Duality	6824	0.510	0.500	0	1	1
BoardTenure	6795	2.100	0.450	1.880	2.160	2.380
BoardSize	6824	2.310	0.220	2.200	2.300	2.480
ROA	6835	0.0500	0.110	0.0200	0.0600	0.100
LEV	6818	0.600	0.240	0.460	0.600	0.720
BTM	6823	0.390	0.570	0.180	0.330	0.540
FirmSize	6835	8.980	1.310	8.040	8.890	9.860
AF	6642	2.610	0.590	2.330	2.720	3.020
IO	5946	18.75	1.110	18.05	18.74	19.49

Panel C: Mean Difference of Regression Variables-High/Low ESG Rating Disagreement

Variable	N (Low)	Mean (Low)	N (High)	Mean (High)	Mean-diff (Low-High)
ESG_Avg	3421	56.4199	3415	44.8533	11.5666***
ESG_Avg2	3421	57.4988	3415	46.6385	10.8603***
BoardGenderDiversity	3414	19.7502	3410	15.9296	3.8206***
ExecGenderDiversity	3412	15.2157	3407	12.9919	2.2238***
BoardIndep	3401	83.0223	3402	78.5302	4.4921***
CEO_Duality	3414	0.5360	3410	0.4900	0.0460***
BoardTenure	3403	2.1176	3392	2.0800	0.0376***
BoardSize	3414	2.3399	3410	2.2730	0.0670***
ROA	3420	0.0631	3415	0.0444	0.0187***
LEV	3411	0.6172	3407	0.5857	0.0314***
BTM	3420	0.3721	3403	0.4102	-0.0382***
FirmSize	3420	9.2047	3415	8.7471	0.4576***
AF	3325	2.6949	3317	2.5244	0.1704***
IO	3079	18.9318	2867	18.5542	0.3776***

Table 2. Correlation Matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) ESG_Divergence	1.000													
(2) ESG_Avg	-0.582	1.000												
(3) BoardGenderDiversity	-0.283	0.382	1.000											
(4) ExecGenderDiversity	-0.138	0.215	0.345	1.000										
(5) BoardIndep	-0.272	0.347	0.250	0.125	1.000									
(6) CEO_Duality	-0.031	0.037	0.071	0.015	0.180	1.000								
(7) BoardTenure	-0.038	-0.022	-0.042	-0.096	0.000	0.088	1.000							
(8) BoardSize	-0.196	0.355	0.216	0.138	0.188	0.060	0.047	1.000						
(9) ROA	-0.084	0.102	0.063	0.059	0.020	0.022	0.154	0.036	1.000					
(10) LEV	-0.085	0.142	0.156	0.088	0.100	0.057	-0.109	0.186	-0.060	1.000				
(11) BTM	0.045	-0.038	-0.036	-0.044	-0.042	-0.043	-0.009	-0.003	-0.048	-0.275	1.000			
(12) FirmSize	-0.220	0.468	0.226	0.132	0.171	0.137	-0.051	0.470	0.032	0.195	0.103	1.000		
(13) AF	-0.177	0.292	0.111	0.074	0.128	0.089	0.034	0.161	0.109	-0.030	-0.029	0.424	1.000	
(14) IO	-0.212	0.412	0.168	0.132	0.200	0.065	-0.080	0.329	0.050	0.100	-0.020	0.686	0.534	1.000

Table 3. Baseline Regressions of ESG Rating Disagreement on Board Gender Diversity

This table presents the baseline regressions of ESG rating disagreement on board gender diversity and other firm and board characteristics. The regressions include year, industry, or firm fixed effects. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The *t*-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) ESG_Divergence	(2) ESG_Divergence
BoardGenderDiversity	-0.116***	-0.112***
	(-4.79)	(-4.39)
ExecGenderDiversity	-0.00759	-0.0319
	(-0.39)	(-1.58)
BoardIndep	-0.173***	-0.161***
	(-7.41)	(-6.63)
CEO_Duality	0.213	1.502***
	(0.49)	(3.38)
BoardTenure	-0.871	-2.006***
	(-1.52)	(-2.65)
BoardSize	-2.636*	0.259
	(-1.93)	(0.21)
ROA	-5.273***	-7.760***
	(-2.66)	(-3.75)
LEV	-0.908	2.812*
	(-0.94)	(1.78)
BTM	-0.489	-0.0489
	(-0.70)	(-0.09)
FirmSize	-0.130	-1.227**
	(-0.35)	(-2.09)
AF	-2.031***	-0.919
	(-3.93)	(-1.44)
IO	-0.475	-0.930*
	(-1.30)	(-1.92)
Intercept	55.32***	62.87***
	(9.08)	(6.62)
Year_FE	Yes	Yes
Industry_FE	Yes	
Firm_FE		Yes
N	5757	5556
Adj. R^2	0.172	0.533

Table 4. Industry BGD & State BGD as Instrumental Variables

This table presents the results using a two-stage least square (2SLS) approach. Two instrumental variables are used: *StateBGD* and *IndBGD*. The year, industry, and firm fixed effects are included. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Variables	BGD	ESG_Divergence	BGD	ESG_Divergence
Predicted BGD		-0.175***		-0.141***
		(-4.09)		(-3.12)
StateBGD	0.375***		0.435***	
	(6.12)		(7.56)	
IndBGD	0.850***		0.783***	
	(31.01)		(26.99)	
Controls	Yes	Yes	Yes	Yes
Cragg-Donald Wald F-statistics	1684.80		1292.56	
Observations	5,544	5,544	5,349	5,349
Fixed Year	YES	YES	YES	YES
Fixed Firm	NO	NO	YES	YES
Fixed Industry	YES	YES	NO	NO
Adj. R^2	0.569	0.127	0.819	0.047

Table 5. Shock of Big Three "Fearless Girl" Campaign

This table presents the results using the Big Three "Fearless Girl" as a shock. Panel A presents a DiD analysis with firms with low female board representation and with Big Three ownership as treated firms. Panel B presents a trend analysis of low female board representation and with Big Three ownership and ESG rating disagreement from 2014 to 2019. The year, industry, and firm fixed effects are included. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: DiD Analysis – Low Female Presentation with Big Three Ownership - Post 2016

	(1) ESG_ Divergence	(2) ESG_ Divergence
Big3_FemaleLow*Post2016	-2.164***	-1.578***
	(-3.04)	(-2.69)
Big3_FemaleLow	3.510***	
	(6.08)	
Controls & Intercept	Yes	Yes
Year_FE	Yes	Yes
Industry_FE	Yes	
Firm_FE		Yes
N	5757	5556
Adj. R^2	0.178	0.531

Panel B: Trend Analysis 2014-2019 Low Female Presentation with Big Three Ownership

	(1) ESG_Divergence	(2) ESG_Divergence
Year2014	0.844*	0.485
	(1.73)	(1.04)
Year2015	-1.323**	-2.015***
	(-2.52)	(-4.27)
Year2016	-2.494***	-2.724***
	(-4.46)	(-5.09)
Year2017	-2.884***	-2.770***
	(-5.20)	(-5.09)
Year2018	-2.398***	-2.336***
	(-4.40)	(-4.42)
Year2019	-2.705***	-2.745***
	(-5.00)	(-4.91)
Big3_FemaleLow	3.618***	3.186***
	(5.32)	(5.39)
Big3_FemaleLow *Year2014	0.135	0.816
	(0.16)	(1.15)
Big3_FemaleLow *Year2015	-0.176	0.185
	(-0.22)	(0.25)
Big3_FemaleLow *Year2016	-0.839	-1.470*
	(-0.88)	(-1.77)
Big3_FemaleLow *Year2017	-1.482	-2.712***
	(-1.57)	(-3.17)

Big3_FemaleLow *Year2018	-1.859*	-3.031***
	(-1.88)	(-3.38)
Big3_FemaleLow *Year2019	-4.212***	-5.141***
	(-3.76)	(-4.68)
Controls & Intercept	Yes	Yes
Industry_FE	Yes	
Firm_FE		Yes
N	5757	5556
Adj. R^2	0.178	0.539

Table 6. Moderating Role of Board Gender Diversity

This table presents a moderating effect of board gender diversity on the association of ESG controversy score and conservatism with ESG divergence by interacting board gender diversity with *ESGControveries* and *CashHolding* (proxy for conservatism), respectively. The year, industry, and firm fixed effects are included. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Moderating Role of BGD – ESG Controversies

	(1) ESG_Divergence	(2) ESG_Divergence
ESGControversies	-0.0448***	-0.0574***
	(-3.22)	(-4.38)
BoardGenderDiversity	0.101*	-0.00526
	(1.85)	(-0.10)
ESGControversies* BoardGenderDiversity	-0.00261***	-0.00137**
	(-4.30)	(-2.30)
Controls & Intercept	Yes	Yes
Year_FE	Yes	Yes
Industry_FE	Yes	
Firm_FE		Yes
N	5757	5556
Adj. R^2	0.232	0.564

Panel B: Moderating Role of BGD - Conservatism/Cash Holding

	(1) ESG_Divergence	(2) ESG_Divergence
CashHolding	-1.047***	-0.662**
	(-2.79)	(-2.12)
BoardGenderDiversity	-0.494***	-0.302***
	(-5.56)	(-3.33)
CashHolding* BoardGenderDiversity	0.0627***	0.0314**
	(4.45)	(2.24)
Controls & Intercept	Yes	Yes
Year_FE	Yes	Yes
Industry_FE	Yes	
Firm_FE		Yes
N	5626	5427
Adj. R^2	0.181	0.534

Table 7. Channel Analysis

This table presents the channel effects in the effect of board gender diversity on ESG rating disagreement with controls for firm and board characteristics. Panel A presents the channel effect of ESG performance. Panel B presents the channel effect of managerial ability. Panel C presents the channel effect of ESG reporting. Panel D presents the channel effect of ESG committees. The year, industry, and firm fixed effects are included. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Channel - ESG Performance

	(1) ESG_Avg	(2) ESG_Avg	(3) ESG_Divergence	(4) ESG_Divergence
BoardGenderDiversity	0.204***	0.106***	-0.0250	-0.0510**
	(6.57)	(5.34)	(-1.19)	(-2.20)
ESG_Avg			-0.445***	-0.572***
			(-18.58)	(-16.20)
Controls & Intercept	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes
Industry_FE	Yes		Yes	
Firm_FE		Yes		Yes
N	5757	5556	5757	5556
Adj. R^2	0.430	0.851	0.347	0.609

Panel B: Channel - Managerial Ability

	(1) MgtScore	(2) MgtScore	(3) ESG_Divergence	(4) ESG_Divergence
BoardGenderDiversity	0.438***	0.567***	-0.0671***	-0.0542**
	(11.58)	(14.05)	(-2.95)	(-2.32)
MgtScore			-0.120***	-0.111***
			(-13.23)	(-13.67)
Controls & Intercept	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes
Industry_FE	Yes		Yes	
Firm_FE		Yes		Yes
N	14460	14225	5757	5556
Adj. R^2	0.339	0.657	0.248	0.569

Table 7. Channel Analysis (continued)

Panel C: Channel - ESG Reporting

	(1) CSR_Reporting	(2) CSR_Reporting	(3) ESG_Divergence	(4) ESG_Divergence
BoardGenderDiversity	0.00428***	0.00300***	-0.0903***	-0.105***
	(7.28)	(4.64)	(-3.96)	(-4.20)
CSR_Reporting			-5.190***	-3.110***
			(-10.19)	(-6.88)
Controls & Intercept	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes
Industry_FE	Yes		Yes	
Firm_FE		Yes		Yes
N	14380	14145	5732	5535
Adj. R^2	0.395	0.626	0.214	0.539

Table 8. PSM Analysis

This table presents the PSM Analysis. Panel A presents the mean difference between treated and control groups in the PSM-matched sample. Panel B presents the average treatment effect. Panel C presents the regression results using the PSM-matched sample. The year, industry, and firm fixed effects are included. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Mean Difference Between Treated and Controls in PSM-matched Sample

Variable	N (Controls)	Mean (Controls)	N (Treated)	Mean (Treated)	Mean-diff (Matched - Treated)
ExecGenderDiversity	1298	14.0968	1291	13.8342	0.2626
BoardIndep	1298	81.4110	1291	81.5803	-0.1693
CEO_Duality	1298	0.5223	1291	0.4965	0.0258
BoardTenure	1298	2.1128	1291	2.1114	0.0014
BoardSize	1298	2.3132	1291	2.3128	0.0004
ROA	1298	0.0576	1291	0.0539	0.0036
LEV	1298	0.6068	1291	0.6028	0.0040
BTM	1298	0.3721	1291	0.3760	-0.0039
FirmSize	1298	8.9709	1291	8.9266	0.0443
AF	1298	2.6048	1291	2.5884	0.0163
IO	1298	18.7290	1291	18.7021	0.0270

Panel B: Average Treatment Effect (ATT)

Variable	Sample	Treated	Controls	Difference	T-stat
ESG_Divergence	Unmatched	13.0196470	16.5377506	-3.51810363	-13.62
	Matched	12.9930427	14.0748541	-1.08181143	-2.61

Panel C: Baseline Regression Analysis on PSM-matched Sample

	(1) ESG_Divergence	(2) ESG_Divergence
BoardGenderDiversity	-0.111***	-0.150***
	(-3.82)	(-4.16)
Controls & Intercept	Yes	Yes
Year_FE	Yes	Yes
Industry_FE	Yes	
Firm_FE		Yes
N	2589	2383
Adj. R^2	0.152	0.517

Table 9. Alternative Models - GLM, WLS, and Newey-West

This table presents the results using alternative models for the baseline regression of ESG rating disagreement on board gender diversity and controls for firm and board characteristics. The year, industry, and firm fixed effects are included. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) ESG_ Divergence GLM	(2) ESG_ Divergence GLM	(3) ESG_ Divergence WLS	(4) ESG_ Divergence WLS	(5) ESG_ Divergence Newey-West	(6) ESG_ Divergence Newey-West
BoardGender Diversity	-0.116***	-0.112***	-0.117***	-0.112***	-0.116***	-0.112***
	(-4.80)	(-4.39)	(-8.67)	(-6.19)	(-6.00)	(-5.04)
Controls & Intercept	Yes	Yes	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry_FE	Yes		Yes		Yes	
Firm_FE		Yes		Yes		Yes
N	5757	5757	5757	5757	5757	5757
Adj. R^2			0.176	0.591		

Table 10. Alternative Measures

Panel A presents the results using alternative measures for the independent variable (*BGD_High, Female_Dummy, FemaleHigh, and Female_Num*) in the baseline regression of ESG rating disagreement on board gender diversity and controls for firm and board characteristics. The year, industry, and firm fixed effects are included. Panel B presents the results using an alternative measure for the dependent variable (*ESG_Divergence2*) in the baseline regression of ESG rating disagreement on board gender diversity and controls for firm and board characteristics. The year, industry, and firm fixed effects are included. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Alternative Measures for Board Gender Diversity (BGD) - *BGD_High, Female_Dummy, FemaleHigh, and Female_Num*

	(1) ESG_ Divergence	(2) ESG_ Divergence	(3) ESG_ Divergence	(4) ESG_ Divergence	(5) ESG_ Divergence	(6) ESG_ Divergence	(7) ESG_ Divergence	(8) ESG_ Divergence
BGD_High	-1.744***	-1.122***						
	(-4.07)	(-3.16)						
Female_Dummy			-3.507***	-3.934***				
			(-4.01)	(-5.62)				
FemaleHigh					-2.902***	-2.426***		
					(-5.50)	(-5.58)		
Female_Num							-1.979***	-1.823***
							(-3.74)	(-3.50)
Controls & Intercept	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry_FE	Yes		Yes		Yes		Yes	
Firm_FE		Yes		Yes		Yes		Yes
N	5757	5556	5757	5556	5757	5556	5229	5066
Ajd. R2	0.169	0.531	0.171	0.536	0.176	0.535	0.137	0.509

Table 10. Alternative Measures (continued)

Panel B: Alternative Measure for ESG Rating Disagreement - *ESG_Divergence2*

	(1) ESG_Divergence2	(2) ESG_Divergence2
BoardGenderDiversity	-0.106***	-0.0863***
	(-4.19)	(-3.27)
Controls & Intercept	Yes	Yes
Year_FE	Yes	Yes
Industry_FE	Yes	
Firm_FE		Yes
N	5757	5556
Adj. R^2	0.207	0.582

Table 11. Lag Independent Variables on ESG Rating Disagreement

This table presents the results using one-year lag independent variables for the regression of ESG rating disagreement on board gender diversity and controls for firm and board characteristics. The year, industry, and firm fixed effects are included. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) ESG_Divergence	(2) ESG_Divergence
Lag.BoardGenderDiversity	-0.106***	-0.0820***
	(-4.18)	(-2.97)
Lag.ExecGenderDiversity	0.00394	-0.00201
	(0.19)	(-0.09)
Lag.BoardIndep	-0.144***	-0.0720***
	(-5.94)	(-2.90)
Lag.CEO_Duality	-0.299	0.567
	(-0.64)	(1.15)
Lag.BoardTenure	-0.288	-1.498*
	(-0.50)	(-1.94)
Lag.BoardSize	-2.622*	0.870
	(-1.82)	(0.67)
Lag.ROA	-5.396**	-6.617***
	(-2.40)	(-2.99)
Lag.LEV	-1.212	2.541
	(-1.17)	(1.38)
Lag.BTM	-0.710	0.438
	(-0.95)	(0.70)
Lag.FirmSize	0.162	-1.007
	(0.41)	(-1.50)
Lag.AF	-1.990***	-0.500
	(-3.49)	(-0.74)
Lag.IO	-0.632	-1.456**
	(-1.61)	(-2.32)
Intercept	51.56***	58.92***
	(7.73)	(4.91)
Year_FE	Yes	Yes
Firm_FE		Yes
Industry_FE	Yes	
N	5229	5184
Adj. R^2	0.145	0.513

Table 12. Baseline Results of Board Gender Diversity on E/S/G Disagreement

This table presents the results of the regression of disagreement of each pillar (E/S/G) on board gender diversity and controls for firm and board characteristics. The year, industry, and firm fixed effects are included. All variable definitions are provided in the Appendix. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) E_ Divergence	(2) E_ Divergence	(3) S_ Divergence	(4) S_ Divergence	(5) G_ Divergence	(6) G_ Divergence
BoardGender Diversity	-0.143***	-0.0639*	-0.0489	-0.0314	-0.119***	-0.0746*
	(-3.60)	(-1.88)	(-1.49)	(-1.00)	(-3.26)	(-1.80)
ExecGender Diversity	0.0221	0.0198	0.0142	-0.0209	-0.0826***	-0.102***
	(0.70)	(0.76)	(0.60)	(-0.76)	(-2.95)	(-2.96)
BoardIndep	-0.0754**	-0.0334	-0.0546**	-0.0566*	-0.283***	-0.401***
	(-2.10)	(-1.11)	(-2.10)	(-1.92)	(-8.88)	(-9.98)
CEO_Duality	0.243	1.017*	0.133	1.300**	1.711***	3.596***
	(0.36)	(1.86)	(0.23)	(2.18)	(2.65)	(4.63)
BoardTenure	0.143	0.330	0.446	-1.555	-3.410***	-3.273***
	(0.17)	(0.33)	(0.71)	(-1.56)	(-4.23)	(-2.65)
BoardSize	-5.991***	2.136	-0.750	0.122	2.532	2.740
	(-3.04)	(1.37)	(-0.49)	(0.08)	(1.40)	(1.34)
ROA	-8.928**	-2.741	0.122	-1.077	-10.94***	-1.241
	(-2.56)	(-0.99)	(0.04)	(-0.31)	(-3.53)	(-0.36)
LEV	-2.702*	3.362	-0.716	0.698	-1.461	-1.817
	(-1.66)	(1.63)	(-0.47)	(0.35)	(-1.09)	(-0.77)
BTM	-1.220	-0.239	-0.247	-0.423	-1.691*	-0.470
	(-1.14)	(-0.30)	(-0.30)	(-0.63)	(-1.75)	(-0.45)
FirmSize	-2.630***	-3.536***	0.326	0.908	-0.660	-1.067
	(-5.28)	(-4.04)	(0.82)	(0.96)	(-1.30)	(-1.12)
AF	-2.317***	-1.668**	-2.326***	-0.794	-1.338	-1.390
	(-2.96)	(-2.05)	(-3.50)	(-1.06)	(-1.63)	(-1.39)
IO	-0.196	-0.0719	-0.244	-1.994**	-1.142*	-1.632*
	(-0.39)	(-0.11)	(-0.70)	(-2.47)	(-1.91)	(-1.84)
Intercept	79.24***	54.46***	30.28***	55.10***	78.89***	98.46***
	(9.19)	(3.86)	(5.04)	(3.58)	(8.34)	(5.51)
Year_FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry_FE	Yes		Yes		Yes	
Firm_FE		Yes		Yes		Yes
N	5494	5473	5494	5473	5494	5473
Adj. R^2	0.221	0.659	0.0482	0.481	0.155	0.467

APPENDIX

Table A1. Variable Definitions and Sources

Variable Name	Definition, description	Source
Dependent Variables:		
ESG_Divergence	Absolute difference between Sustainalytics overall ESG scores and Refinitiv overall ESG Combined scores (both with ESG controversies incorporated)	Refinitiv, Sustainalytics
ESG_Divergence2	Absolute difference between Sustainalytics overall ESG scores (ESG controversies incorporated) and Refinitiv overall ESG scores (ESG controversies not incorporated).	Refinitiv, Sustainalytics
E_Divergence	Absolute difference between Sustainalytics environment (E) score and Refinitiv environment (E) score which is not a combined score with controversies incorporated	Refinitiv, Sustainalytics
S_Divergence	Absolute difference between Sustainalytics social (S) score and Refinitiv social (S) score which is not a combined score with controversies incorporated	Refinitiv, Sustainalytics
G_Divergence	Absolute difference between Sustainalytics governance (G) score and Refinitiv governance (G) score which is not a combined score with controversies incorporated	Refinitiv, Sustainalytics
Independent Variables:		
BoardGenderDiversity (BGD)	Percentage of female on the board	Refinitiv
BGD_High	Dummy variable which equals 1 if a firm's *BoardGenderDiversity* (BGD) is above the median, otherwise 0	Refinitiv
BGD_Low	Dummy variable which equals 1 if a firm's *BoardGenderDiversity* (BGD) is blow the median, otherwise 0	Refinitiv
FemaleHigh	Dummy variable which equals 1 if the board female number is equal to or larger than two, otherwise 0	Refinitiv
FemaleLow	Dummy variable which equals 1 if the board female number is equal to or smaller than one, otherwise 0	Refinitiv
Female_Dummy	Dummy variable which equals 1 if at least one female on a firm's board, otherwise 0	Refinitiv
Female_Number	Number of females on board	Refinitiv
Female_Num	Log of the number of females on board (*Female_Number)*	Refinitiv

Controls:		
ExecGenderDiversity	Percentage of female executive members	Refinitiv
CEO_Duality	Dummy variable which equals 1 if the CEO simultaneously chairs the board in the current period, otherwise 0	Refinitiv
BoardTenure	Log of the average number of years each board member has been on the board	Refinitiv
BoardSize	Log of the total number of board members at the end of the fiscal year	Refinitiv
BoardIndep	Percentage of independent board members as reported by the firm	Refinitiv
ROA	Net income divided by total assets	Compustat
LEV	Total liabilities divided by total assets	Compustat
BTM	Book value per share divided by the fiscal year-end closing price	Compustat
FirmSize	Natural log of total assets (in US$ millions)	Compustat
AF	Natural log annual average number of analysts following	IBES
IO	Natural log of total institutional ownership	Thomson Reuters Institutional (13f)
Other Variables:		
ESG_Avg	Average of overall ESG scores by Sustainalytics and overall ESG combined scores by Refinitiv (both incorporate controversies)	Refinitiv, Sustainalytics
ESG_Avg2	Average of overall ESG scores by Sustainalytics and Refinitiv (Sustainalytics incorporates controversies, but Refinitiv does not)	Refinitiv, Sustainalytics
High_ESG	Dummy variable which equals 1 if a firm's ESG_Avg is above the median, otherwise 0	Refinitiv, Sustainalytics
High_ESG2	Dummy variable which equals 1 if a firm's ESG_Avg2 is above the median, otherwise 0	Refinitiv, Sustainalytics
High_ESG_Dis	Dummy variable which equals 1 if a firm's *ESG_Divergence* is above the median, otherwise 0	Refinitiv, Sustainalytics
High_ESG_Dis2	Dummy variable which equals 1 if a firm's *ESG_Divergence2* is above the median, otherwise 0	Refinitiv, Sustainalytics
MgtScore	A firm's commitment and effectiveness towards following best practice corporate governance principles	Refinitiv

High_Mgt	Dummy variable which equals 1 if a firm's *MgtScore* is above the median, otherwise 0	Refinitiv
CSR_Reporting	Dummy variable which equals 1 if a firm issues a separate CSR report or a section in its annual report, otherwise 0	Refinitiv
ESGControversies	ESG Controversy Score, continuous from 0 to 100 (no controversy), measures a firm's exposure to ESG controversies and negative events reflected in global media.	Refinitiv
CashHolding	Log of cash and short-term investments that measures the firm-level conservatism.	Compustat
Big3	Dummy variable which equals to 1 if firms with Big 3 (Vanguard, State Street, and Black Rock) ownership, otherwise 0	Thomson Reuters Institutional (13f)
Big3_FemaleLow	Interaction of *Big3* and *FemaleLow*	Thomson Reuters Institutional (13f); Refinitiv
StateBGD	Average board gender diversity in each US state	Refinitiv
IndBGD	Average board gender diversity in each industry using SIC classification	Refinitiv
PostYear	Dummy variable which equals 1 if it is after (does not include) the Year, otherwise 0	

Table A2. Sample Construction

Description	No. of firm year obs.	No. of firms
Compustat North America Fundamentals 2009-2019	124,946	19,214
Merged with Sustainalytics ESG coverage firms 2009-2019	18,965	3,855
Merged with IBES	14,893	3,269
Merged with Thomson Reuters IO	12,505	2,873
Merged with Refinitiv US ESG coverage firms	7,522	1,207
Excluding financial firms (Fama French ffi17=16)	6,836	1,000

Table A3. Frequency Analysis

Panel A: Frequency Analysis by Year

	ESG Divergence		ESG Avg		Board Gender Diversity	
Year	N	Mean	N	Mean	N	Mean
2009	488	17.56	488	46.32	488	13.88
2010	511	17.06	511	47.44	509	13.98
2011	541	17.33	541	48.79	539	14.36
2012	550	17.21	550	50.17	549	15.20
2013	559	17.49	559	50.26	558	15.86
2014	568	17.33	568	51.00	567	17.10
2015	925	16.89	925	47.67	923	16.52
2016	669	13.94	669	52.03	667	19.06
2017	663	12.47	663	53.40	662	20.25
2018	702	12.77	702	53.37	702	22.15
2019	660	11.68	660	55.34	660	24.72
Total	6836	15.45	6836	50.64	6824	17.84

Panel B: Frequency Analysis by Industry

Industry (ffi17)	ESG Divergence		ESG Avg		Board Gender Diversity	
Year	N	Mean	N	Mean	N	Mean
Food	289	11.47	289	57.77	289	22.41
Mining and Minerals	164	16.05	164	51.97	164	14.03
Oil and Petroleum Products	457	19.34	457	47.56	453	13.20
Textiles, Apparel & Footware	140	14.44	140	52.08	140	24.05
Consumer Durables	127	15.56	127	51.02	127	20.24
Chemicals	216	13.53	216	52.81	208	16.75
Drugs, Soap, Prfums, Tobacco	286	14.01	286	52.27	286	22.37
Construction & Construction Materials	306	15.72	306	49.72	306	16.32
Steel Works Etc	80	15.81	80	44.89	80	16.86
Fabricated Products	50	15.45	50	49.20	50	13.54
Machinery & Business Equipment	1075	15.41	1075	53.74	1075	16.16
Automobiles	194	16.24	194	47.36	194	14.08
Transportation	415	16.00	415	49.30	415	15.73

Utilities	435	14.10	435	56.10	435	21.45
Retail	455	14.57	455	50.20	455	21.50
Other	2147	15.82	2147	48.06	2147	17.89
Total	6836	15.45	6836	50.64	6824	17.84

Table A4. Sub-Sample Analysis

This table presents the sub-sample analysis based on high/low ESG performance (above/below the median each year in our sample), high/low managerial ability (above/below the median each year in our sample), and whether a firm issues voluntary ESG reporting or a separate section in the annual report. The year, industry, and firm fixed effects are included. All variable definitions are provided in Table A1. Standard errors are clustered at the firm level. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

The results show that the significant relation is only present in the sub-samples of firms with low ESG performance, low managerial ability, and firms without voluntary ESG reporting. This may suggest that the overall effect is driven by the firms with low ESG performance, low managerial ability, and firms without voluntary ESG reporting, and that ESG raters tend to disagree more about the ESG performance of these firms.

This is consistent with the previous descriptive statistics that high ESG performers have lower ESG disagreement, the positive relation between *ESG_Avg* and *ESG_Divergence* in the baseline regression, and the channel analysis. This highlights the importance of board gender diversity, especially for those low ESG performers, low managerial ability, and firms without voluntary ESG reporting.

Panel A: Board Gender Diversity on ESG Rating Disagreement - High/Low ESG Performance

	(1) ESG_Divergence (High ESG)	(2) ESG_Divergence (High ESG)	(3) ESG_Divergence (Low ESG)	(4) ESG_Divergence (Low ESG)
BoardGender Diversity	-0.0123	-0.0424	-0.0940***	-0.113***
	(-0.51)	(-1.42)	(-3.12)	(-3.45)
Controls & Intercept	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes
Industry_FE	Yes		Yes	
Firm_FE		Yes		Yes
N	2956	2864	2801	2581
Adj. R^2	0.0290	0.308	0.260	0.668

Panel B: Board Gender Diversity on ESG Rating Disagreement - High/Low Managerial Ability

	(1) ESG_Divergence (High Managerial Ability)	(2) ESG_Divergence (High Managerial Ability)	(3) ESG_Divergence (Low Managerial Ability)	(4) ESG_Divergence (Low Managerial Ability)
BoardGender Diversity	-0.0349	-0.0405	-0.123***	-0.0976***
	(-1.33)	(-1.27)	(-4.00)	(-2.62)
Controls & Intercept	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes

Industry_FE	Yes		Yes	
Firm_FE		Yes		Yes
N	2964	2801	2793	2602
Adj. R^2	0.0525	0.418	0.250	0.642

Panel C: Board Gender Diversity on ESG Rating Disagreement - With/Without ESG Reporting

	(1) ESG_Divergence (With ESG Reporting)	(2) ESG_Divergence (With ESG Reporting)	(3) ESG_Divergence (Without ESG Reporting)	(4) ESG_Divergence (Without ESG Reporting)
BoardGender Diversity	-0.0221	-0.0219	-0.151***	-0.179***
	(-0.88)	(-0.71)	(-4.71)	(-5.21)
Controls & Intercept	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes
Industry_FE	Yes		Yes	
Firm_FE		Yes		Yes
N	3044	2926	2688	2473
Adj. R^2	0.0509	0.342	0.254	0.690