ecgi

ESG rating disagreement and stock returns

Finance Working Paper N° 651/2020
January 2020

Rajna Gibson
University of Geneva, GFRI, and ECGI

Philipp Krueger
University of Geneva, SFI and GFRI

Nadine Riand
University of Geneva and GFRI

Peter S. Schmidt
University of Geneva and GFRI

ECGI Working Paper Series in Finance

ESG rating disagreement and stock returns

Working Paper N° 651/2020

January 2020

Rajna Gibson
Philipp Krueger
Nadine Riand
Peter S. Schmidt

We thank Oghuzan Karakas for his valuable comments. The paper also benefited from comments received during the PRI conference in Paris in September 2019, the ECGI and University of Cambridge Conference "Engage or Divest? Strategies for responsible Investing" in Cambridge in October 2019. Supported by the Swiss National Science Foundation (SNSF) within the framework of the National Research Programme "Sustainable Economy: resource-friendly, future-oriented, innovative" (NRP 73). The views expressed in this paper are solely those of the authors. We also remain responsible for any errors and omissions.

Abstract

Using a sample of S&P 500 firms between 2013 and 2017, we study the impact of ESG rating disagreement on stock returns. Building on the heterogeneous beliefs literature, we conjecture that for disagreement about environmental ratings, a risk–based explanation induces a positive relationship between disagreement and stock returns. In contrast, we hypothesize that for disagreement about the social and the governance dimension, the impact on stock returns is driven by mispricing considerations and also depends on whether the disagreeing rating providers are located in civil or common law jurisdictions. The idea is that civil (common) law rating providers are more apt at identifying material social (governance) information, and that disagreement by such rating providers results in overvaluation and thus lower subsequent stock returns. Our empirical findings support these hypotheses.

Keywords: Disagreement, non-financial information, ESG ratings dispersion, heterogeneous beliefs, stock returns, legal origins, sustainable finance

JEL Classifications: G12, G24, Q01

Rajna Gibson
Professor of Finance
University of Geneva, The Geneva Finance Research Institute
40 Boulevard du Pont d'Arve
1211 Geneva 4, Switzerland
phone: +41 223 798 983
e-mail: rajna.gibson@unige.ch

Philipp Krueger
Associate Professor of Finance
University of Geneva, Geneva School of Economics and Management
40, Bd du Pont-d'Arve
1211 Geneva 4, Switzerland
phone: +41 223 798 569
e-mail: philipp.krueger@unige.ch

Nadine Riand
Researcher
University of Geneva, School of Economics and Management
40 Boulevard du Pont d'Arve
1211 Genève 4, Switzerland
e-mail: nadine.riand@unige.ch

Peter S. Schmidt*
Researcher
University of Geneva, School of Economics and Management
40 Boulevard du Pont d'Arve
1211 Genève 4, Switzerland
e-mail: peter.schmidt@unige.ch

*Corresponding Author

s:fi
RESEARCH

Swiss Finance Institute
Research Paper Series
N°19-67

ESG Rating Disagreement and Stock Returns

Rajna Gibson
University of Geneva

Philipp Krueger
University of Geneva and Swiss Finance Institute

Nadine Riand
University of Geneva

Peter Steffen Schmidt
University of Zurich and University of Geneva

ESG rating disagreement and stock returns

Rajna Gibson, Philipp Krueger, Nadine Riand, and Peter S. Schmidt *

First version: February 2019
This version: December 2019

Abstract: Using a sample of S&P 500 firms between 2013 and 2017, we study the impact of ESG rating disagreement on stock returns. Building on the heterogeneous beliefs literature, we conjecture that for disagreement about environmental ratings, a risk–based explanation induces a positive relationship between disagreement and stock returns. In contrast, we hypothesize that for disagreement about the social and the governance dimension, the impact on stock returns is driven by mispricing considerations and also depends on whether the disagreeing rating providers are located in civil or common law jurisdictions. The idea is that civil (common) law rating providers are more apt at identifying material social (governance) information, and that disagreement by such rating providers results in overvaluation and thus lower subsequent stock returns. Our empirical findings support these hypotheses.

Keywords: Disagreement, non-financial information, ESG ratings dispersion, heterogeneous beliefs, stock returns, legal origins, sustainable finance

1 Introduction

Quantitative assessments of a firm's environmental, social, and governance (ESG) policies play an increasingly important role, both in academia and investment practice. For example, ESG (or non-financial) ratings (scores) are now commonly used in finance research

* University of Geneva, Swiss Finance Institute, Geneva School of Economics and Management (GSEM), and Geneva Finance Research Institute (GFRI); Bd du Pont d'Arve 40, 1211 Geneva 4, Switzerland; Rajna Gibson Brandon, *rajna.gibson@unige.ch*; Philipp Krueger, *philipp.krueger@unige.ch*; Nadine Riand *nadine.riand@unige.ch*; Peter S. Schmidt *peter.schmidt@unige.ch*. We thank Oghuzan Karakas for his valuable comments. The paper also benefited from comments received during the PRI conference in Paris in September 2019, the ECGI and University of Cambridge Conference "Engage or Divest? Strategies for responsible Investing" in Cambridge in October 2019. Supported by the Swiss National Science Foundation (SNSF) within the framework of the National Research Programme "Sustainable Economy: resource-friendly, future-oriented, innovative" (NRP 73). The views expressed in this paper are solely those of the authors. We also remain responsible for any errors and omissions.

(see, for instance, Hong and Kostovetsky, 2012; Lins et al., 2017; Liang and Renneboog, 2017; Dyck et al., 2019). In a similar spirit, such ratings also increasingly shape investment decisions of institutional investors representing trillions in assets under management (see Gibson et al. (2019) or (GSIA, 2016; USSIF, 2018; PRI, 2018).

Recently, the financial press (Mackintosh, 2018; Wigglesworth, 2018), academics (Chatterji et al., 2016), as well as policy-oriented think tanks (Doyle, 2018) have stressed that considerable disagreement can exist when comparing a firm's ESG rating across different data providers. In this paper, we provide a first step towards better understanding the nature of disagreement about a firm's ESG ratings and–more importantly–we analyze the impact of such disagreement on stock returns. To do so, we collect ESG ratings from six prominent ESG ratings providers[1] for a sample of S&P 500 firms between 2013 and 2017.

We start our analysis by documenting some very basic empirical facts concerning ESG rating disagreement (or divergence, dispersion). We show, for example, that the average correlation between the overall ESG ratings of six rating providers is about 0.46. Surprisingly, the average correlation is lowest for the governance (0.19) and highest for the environmental dimension (0.43). We also provide evidence that disagreement tends to be higher for the largest firms in the S&P 500 and for firms that do not have a credit rating. In contrast, more profitable firms tend to have lower ESG rating disagreement. We further show that ratings disagreement is generally more pronounced for the telecommunications and the consumer durables industry with the exception of disagreement about governance, which is highest in the financial services industry.

While it is certainly useful knowing that ESG ratings can disagree and quantifying the extent of disagreement, it is more important to examine whether there are real consequences associated with such disagreement. In this paper we focus on examining whether

[1] We use data from Asset 4, Sustainalytics, Inrate, Bloomberg, MSCI KLD and MSCI IVA.

disagreement has implications for stock returns.

To study the implications of ESG rating disagreement for stock returns and structure our analysis, we build on the heterogeneous beliefs literature (Miller (1977)). Following Atmaz and Basak (2018), we conjecture that a risk-based explanation should be at play for the relation between stock returns and disagreement in the environmental rating. More specifically, we expect a positive relationship between disagreement in the environmental rating and subsequent stock returns. In contrast, we argue that for the social and the governance dimension the impact of disagreement on stock returns is driven by mispricing considerations and should also depend on the legal origin of the rating providers. More specifically, we expect that when ESG rating providers that are headquartered in countries with a civil law origin disagree about social information, such disagreement should have a more pronounced impact on stock returns. The premise is that social issues at the firm-level play a more important role in civil law countries and rating providers from such countries are better at identifying financially relevant social issues. In contrast, the shareholder centric view that dominates the theory and governance of the firm in common law countries implies that non-financial governance information produced by information intermediaries with a common law background is likely to be more financially relevant. As a result, disagreement about firms' governance by information intermediaries located in common law countries should therefore be more financially relevant. Hence, combining considerations related to mispricing and legal origins, we postulate that the relationship between stock returns and ESG rating disagreement should be negative for social (governance) ratings when rating agencies from civil (common) law jurisdictions disagree about the ratings.

In order to test our hypotheses, we measure disagreement using the standard deviation of ESG ratings for a given firm at a given point in time.[2] We calculate the disagreement

[2] We also use the range, that is the difference between the maximum and minimum ESG rating for a given firm at a given point in time, which leads to similiar conclusions.

proxy for the overall ESG rating, separately for the E, S, and G dimension but also stratify by the legal orgin of the data providers (civil vs. common). We then relate monthly stock returns to our ESG ratings disagreement proxies, controlling for a number of known characteristics that predict the cross-section of stock returns.

Consistent with existing models of heterogeneous beliefs in financial markets and the above arguments based on legal origin theory, we find that ESG rating disagreement negatively predicts stock returns whenever information intermediaries located in civil (common) law countries disagree about social (governance) related non-financial information. In line with a risk-based explanation, we observe that for the environmental rating disagreement, the relationship with stock returns is positive and independent of the legal origin of the data-providers. Finally, we examine whether the the stock market cycle has a bearing on the relationship between ratings disagreement and stock returns and find that in times when the S&P 500 returns are high, the risk-based explanation dominates and leads to a positive relationship between ratings disagreement and stock returns for all three ESG pillars.

An important contribution of our paper is to move beyond simply studying disagreement about non-financial information but rather understanding the economics behind it. We add to the debate by suggesting that it is highly relevant to consider "which" kinds of rating providers disagree and about "what" kind of non-financial information. In other words, an important aspect in the debate on ESG rating disagreement is about "who" (civil versus common law) disagrees about "what" (social, environmental or governance). A second contribution of our study is to show that the arguments and predictions laid forward by the existing heterogeneous beliefs asset pricing literature can be extended to disagreement on non-financial information.

Our empirical results should help academics, institutional investors, policy-makers, and ultimately firms themselves to better understand that beyond ESG ratings, the disper-

sion of these scores – stratified by the legal origins of the data providers for the S and the G scores – can have an economically meaningful impact on stock returns and thus on firms' cost of capital.

The rest of the paper is organized as follows. The next section discusses the link with the literature and fleshes out our main hypotheses. Section 3 presents the data. Section 4 provides the main analysis and discusses our results. Section 5 concludes the paper.

The rest of the paper is organized as follows. The next section discusses the link with the literature and fleshes out our main hypotheses. Section 3 presents the data. Section 4 provides the main analysis and discusses our results. Section 5 concludes the paper.

2 Literature review and testable hypotheses

2.1 Literature review

2.1.1 Heterogeneous beliefs in finance

In this paper, we examine the nature and implications of disagreement about firm-level ESG ratings. We anchor our analysis in the literature on heterogeneous beliefs in financial markets. Studying heterogeneous beliefs has a long tradition in finance, going back at least to Miller (1977). While the majority of studies focus on disagreement about financial information, our contribution is to study heterogeneous beliefs about *non-financial* information. As such we are contributing to the relatively scant research that has examined disagreement about other sources of risk: for example, Basak (2000), studies heterogeneous beliefs about non-fundamental processes that drive non-fundamental firm risk.

Miller (1977) argues that, whenever there is disagreement about the financials of a firm, prices will predominantly reflect optimistic investors because pessimistic investors are kept out of the market due to high short-sales constraints. Thus, higher disagreement about the financials of a stock will increase future losses due to excessive optimism and

therefore lower future returns will ensue. Based on Miller (1977), Chen et al. (2002) formulate a stock market model with differences of opinion and short-sales constraints. More recently, Atmaz and Basak (2018)[3] develop a dynamic general equilibrium model populated by a continuum of CRRA investors who differ in their beliefs and consume at a single date. In this setting, they underline the dual role of belief dispersion pointing out that with increasing belief dispersion, on the one hand, future returns may decrease (increase) when the view on a given stock is optimistic (pessimistic). On the other hand, higher belief dispersion reflects higher uncertainty and should thus lead risk-averse agents to require higher future returns. When the view on the stock is sufficiently optimistic, the former effect dominates resulting in a negative relation between the dispersion in beliefs about a stock and the stock's future mean return. To our knowledge, our paper provides the the first attempt to transpose and test the predictions of the heterogeneous beliefs theoretical models' literature in the context of non-financial information.

Empirical studies have tested the relation between dispersion in beliefs and stock returns in a variety of settings. These studies typically use the dispersion in analyst earnings forecasts as a proxy for the extent to which a stock is subject to heterogeneous beliefs. These studies generally document a significant relation between heterogeneous beliefs and stock returns (Diether et al., 2002; Anderson et al., 2005; Yu, 2011). However, while some studies find this relationship to be negative (see, in particular, Diether et al., 2002; Chen et al., 2002; Yu, 2011; Lakonishok et al., 1994; La Porta, 1996; Skinner and Sloan, 2002), others find it to be positive (Anderson et al., 2005; David, 2008; Banerjee and Kremer, 2010).

Other researchers have used alternative proxies for belief dispersion. For instance, Chen et al. (2002) test their model using a proxy of belief dispersion based on a measure

[3] There exist numerous other theoretical papers about heterogeneous beliefs. See for instance, Harrison and Kreps (1978); Harris and Raviv (1993); Detemple and Murthy (1994); Zapatero (1998); Berrada (2006); Jouini and Napp (2007); Dumas et al. (2009).

of stock-level differences in mutual funds holdings. Carlin et al. (2014) use disagreement among Wall Street mortgage dealers about prepayment speeds to empirically examine the effect of differences of opinion on asset prices.

2.1.2 ESG ratings

The academic use of ESG ratings has increased considerably over the last two decades and such measures are now commonly used in economics, finance, and management research (see, for instance, Hong and Kostovetsky, 2012; Krüger, 2015; Lins et al., 2017; Liang and Renneboog, 2017; Gibson Brandon and Krüger, 2018; Dyck et al., 2019). Given the complexity of measuring non-financial or ESG information, the validity of these ratings has been debated critically (Chatterji et al., 2009; Bouten et al., 2017; Delmas et al., 2013). Chatterji et al. (2016), for instance, study the convergence of social ratings produced by six well-established information intermediaries. They document a lack of agreement across raters that comes mainly from two sources: the lack of both a common theorization and commensurability.[4] These findings point out that research conclusions are potentially dependent on the choice of ratings providers, which should be taken into account when drawing conclusions based on existing empirical studies.

More recently, Eccles and Stroehle (2018) explore the importance of the social construction with respect to ESG ratings. They classify ESG rating agencies into value- vs. values-based organizations based on their social origins. Furthermore, they point out how social origins can influence the way in which sustainability is conceptualized, how financial materiality is defined, the way ESG is measured, and how ESG information is sold to investors. We add to this literature by focusing on another classification of ESG information that so far has not been examined in the literature. Namely, we focus on the

[4] A common theorization corresponds to an agreement across raters on a common definition of CSR. The term commensurability means that different raters would obtain the same result when measuring the same feature.

legal origins of the country in which the information intermediary producing the ESG information is headquartered.

Given the heightened interest ESG rating disagreement has generated both in practitioner circles and in the financial press, the topic has recently also spurred some academic interest. We are aware of two contemporaneous papers, which analyze the dispersion of ESG ratings. Christensen et al. (2019) analyze the determinants of ESG rating disagreement and find that more disclosure leads to higher disagreement. In addition they point out that the relationship between a firm's average ESG rating and ESG rating disagreement is non-linear. Importantly, they rely on ESG data from three providers, while our paper makes us of six different data sources. Berg et al. (2019) pursue a more granular approach and focus more on the question of why ratings disagree. By subdividing ratings into finer categories, they identify three sources of ESG rating divergence. First, they highlight that raters use different categories, which can lead to disagreement. They refer to this as *scope divergence*. Secondly, they point out that raters measure identical categories differently, which they refer to as *measurement divergence*. Finally they highlight *weight divergence*, which results from raters attaching different weights to the different categories when generating an aggregated ESG rating. They find that most of the differences can be traced to measurement (53 percent) and scope divergence (44 percent), while weight divergence seems to play a minor role. In addition they find a *Rater Effect*: ratings of one provider are positively correlated across different categories.

2.1.3 Law and finance

In this paper, we also build on the influential economics and finance research concerned with how legal origins shape economic outcomes and financial decisions (La Porta et al., 1998). In the context of ESG, Liang and Renneboog (2017) show that there is a strong correlation between a firm's ESG rating and the legal origin of the country in which the

firm is headquartered. Building on this work, we hypothesize that in our setting, the legal origin of the information provider's headquarter country should also have a bearing on the definition and understanding of non-financial information and also on how financially material such information is likely to be. More specifically, civil law countries are known to have a strong view regarding labor issues and social protection (Botero et al., 2004). On the other hand, common law countries are generally regarded to emphasize investor protection, stronger protection of shareholders rights, and a stronger view on other governance issues (See for example Doidge et al., 2007; La Porta et al., 1998).

2.2 Testable hypotheses

Our main hypotheses focus on the relation between dispersion in beliefs about non-financial information and future stock returns. We argue that this relation depends primarily on the the type of information about which there is disagreement (e.g., environmental, social or governance) and which rating providers disagree about it. More specifically, we hypothesize for social and governance information that it matters whether data providers are located in civil or common law countries.

Transposed to the context of ESG information, the arguments in Miller (1977) suggest that high dispersion implies a too optimistic initial view on a firm's non-financial performance and thus lower future stock returns. A similar prediction obtains in the framework of Atmaz and Basak (2018), if we combine their risk-based and their mispricing conjectures and assume that an optimistic pricing bias dominates. Note that the alternative hypothesis, that is a positive relation between ESG rating disagreement and future stock returns, would be consistent with the risk-based arguments spelled out in Atmaz and Basak (2018) in the absence of any mispricing (or if excessive pessimism dominates). We conjecture that the latter may apply to disagreement in the environmental rating since the environmental performance of a firm is easier to quantify and is based predominantly

on hard information (e.g., energy use, CO2 emissions, water use). We thus we state our first hypothesis with respect to the relation between disagreement in the environmental rating and stock returns:

Hypothesis 1: *For the environmental rating, the risk-based explanation by Atmaz and Basak (2018) should prevail and we should thus observe a positive relationship between disagreement in the environmental rating and future stock returns.*

We next conjecture that mispricing – induced by excessive optimism – is more likely to affect the social and the governance ratings since they are more subjective and generally based on soft information. In parallel, we rely on insights from the literature on law and finance to formulate our second hypothesis about the relation between stock returns and disagreement in terms of a firm's social and governance ratings. Indeed, for those two dimensions, we expect that the legal origin of the country in which the data providers are domiciled plays a key role. Social issues matter more in civil law countries given that firms from these countries generally adopt a more stakeholder oriented approach to corporate governance (e.g., Germany or Japan). As a result, data providers from civil law countries are better positioned to identify financially relevant social issues. Consequently, disagreement between data providers from civil law countries should thus be more relevant when such disagreement concerns social issues.[5] In contrast, common law countries (e.g., U.S. or U.K.) are generally subject to a more shareholder-centric approach to governance. Thus, disagreement in governance ratings should matter more when such disagreement is among rating providers from common law countries. Given these combined arguments from the heterogeneous beliefs and the law and finance literature, we formulate our second hypothesis on the relation between stock returns and disagreement about social and governance ratings:

[5] While the environmental score is potentially also more important in civil law countries, we conjecture that–as mentioned before– the environmental score is less prone to mis-measurement.

Hypothesis 2: *The larger the disagreement about the social (governance) rating issued by civil (common) law data providers, the higher the market price of the stock relative to the true value of the stock, and thus the lower its future returns.*

3 Data

3.1 Sample selection

To test our hypotheses, we construct a representative and homogeneous sample over the longest possible time period. We face the challenge that the availability of ESG data is restricted in both the cross-section and the time-series. In other words, ESG data is often only available for the largest firms and for more recent years. To use a sample as homogeneous as possible and to maximize the number of available ESG ratings per firm, we restrict ourselves to firms belonging to the S&P 500. Since one of the ESG data providers (namely Inrate) does not produce ratings going further back than 2013, we consider a sample period of five years going from 2013 to 2017.

3.2 Financial data

We use financial data from the Center for Research in Security Prices (CRSP) and accounting data from Standard & Poor's Compustat. For each stock, we calculate idiosyncratic volatility, total volatility, and the stock market beta at the end of each month using up to 250 daily observations (we require a minimum of 60 daily observations). We calculate market capitalization as (adjusted) total shares outstanding times stock price, both at the end of the month. The momentum signal at time t is calculated as the continuously compounded returns from month $t-2$ to month $t-12$. Book value of equity is the sum of shareholders' equity, deferred taxes and investment tax credit minus preferred

stock.[6] Only firms with a positive book value are selected into the sample. Following Novy-Marx (2013), gross profitability is calculated as total revenues (revt) minus cost of goods sold (cogs), divided by total assets (at). In addition, we also match the dispersion in EPS forecasts for one year ahead earnings from IBES (Diether et al., 2002).

3.3 ESG data

3.3.1 Overview of ESG data providers

We collect data from six ESG data providers: (1) Asset 4,[7] (2) Sustainalytics, (3) Inrate, (4) Bloomberg,[8] (5) MSCI KLD,[9] and (6) MSCI IVA.[10]

Table 1 displays important features of these six data providers. Column (1) shows the legal origin for each provider. We classify providers based on their country of legal origin as either *common* or *civil law*. Liang and Renneboog (2017) show that legal origins have an important impact on a firm's ESG performance. Similarly, we hypothesize that legal origins play an important role in terms of how a data provider conceptualizes and analyzes non-financial issues. We discuss this notion in detail in section 2.[11]

In Column (2), we show the country of origin of each provider. Three providers are

[6] If available, we use the redemption value as preferred stock. Otherwise, we use the liquidating value or, if the liquidation value is also not available, the carrying value.

[7] Asset 4 was acquired by Thomson Reuters in 2009, but the name was not changed until recently, since then the name of the dataset is *Thomson Reuters ESG Scores*. However, since the name Asset 4 is widely known we use the old name for simplicity. Note that as of 2018, the ESG ratings data is part of Refinitiv, which is a company co-owned by Thomson Reuters (45%) and Blackstone Group LP (55%)

[8] The full name is *Bloomberg ESG*, we denote this dataset simply as Bloomberg.

[9] The KLD dataset was initially created by Kinder, Lydenberg, and Domini Inc., which got acquired by Riskmetrics in 2009. In 2010, MSCI acquired Riskmetrics. For details on the history of KLD, see Eccles et al. (2019).

[10] The IVA dataset was initially created by Innovest, which was also acquired by Riskmetrics in 2009 before Riskmetrics got taken over by MSCI. See also Eccles et al. (2019) for details.

[11] The main argument is that civil law origin providers take a more stakeholder centric view towards non-financial information and attach a more important role to social issues. In contrast, common law providers take a more traditional shareholder-centric view attaching more importance to governance issues.

US-based (Bloomberg, MSCI KLD, and MSCI IVA) whereas two providers are Swiss-based (Asset 4[12] and Inrate) and one provider has its origins in the Netherlands (Sustainalytics).

In Column (3), we show the rating scales originally used by each provider. Three providers apply a scale from 0 to 100 for their assessments (Asset 4, Sustainalytics, and Bloomberg) while one provider uses a scale from 0 to 10 (MSCI IVA) and another provider a scale from 1 to 12 (Inrate)[13]. Originally, MSCI KLD does not provide a genuine scale itself. However, most academic studies sum up KLD's strengths and concerns separately and scale both by the total number of strengths and concerns available. This course of action results in a scale from -1 to +1 (See for example Lins et al., 2017).[14]

Because the different rating scales differ not only in terms of their statistical support, but also in terms of the distribution across the statistical support, a simple re-scaling would not suffice to make the different ratings comparable. Therefore, we do the following to achieve comparability across rating providers: At each point in time, we sort all stocks on the ratings of the respective providers. Then we calculate the percentile ranks and use these as adjusted scores.[15] When there are ties, we assign each company the average rank. We then normalize ranks between 0 and 1.

Column (4) shows the average number of sample stocks per year for which we observe an ESG rating from a given data provider. Sustainalytics, MSCI KLD, MSCI IVA and Bloomberg have on average the best coverage (about 460 stocks). Inrate and Asset 4 have the least number of stocks on average with 432 and 439, respectively. However, the average number of stocks for all providers is rather high with well above 400 and therefore

[12] Even though Asset 4 was taken over by Thomson Reuters, we classify it as being Swiss based. We do so given that Asset 4 was founded as a Swiss company, its conceptualization of non-financial information is more likely to have been shaped by civil law origins.

[13] Note that this scale is based on sustainability assessments from D- to A+.

[14] We include all available strength and concern items except those from the norms-oriented categories related to alcohol, military, firearms, gambling, nuclear, and tobacco.

[15] Using ranked measures is also more consistent with investment practice in which investors compare the value of a given signal relative to the value of the signal for other firms.

we consider the sample as being representative for S&P 500 companies.[16]

The fifth and last column reports the pillar scores supplied by the providers. All providers supply a total score, an environmental score, a social score, and a (corporate) governance score. In addition, Inrate also provides a labor score. Since the labor score captures a social topic, we use for Inrate the average of the original social and the labor score as the social score.

[Table 1 about here.]

3.3.2 Dataset matching

A big challenge for constructing a dataset from many sub-datasets is to properly match the different datasets. We match on three identifiers: (1) CUSIP, (2) ISIN, and (3) company name. The CUSIP code is available for all providers, except Inrate.[17] However, since the ISIN code is available for Inrate, we extract the CUSIP code from the ISIN code. Note also, that we only use the first six CUSIP characters for matching (known as the *issuer* identifier). The characters seven and eight identify the specific issue (for example *10* indicates common equity), and the ninth character is a check digit. The ISIN code is available for all providers except MSCI KLD. For the CRSP/Compustat data we retrieve the ISIN number from the CUSIP code and the current ISO country code of incorporation (fic).[18] To do the merge with the company names, we first convert the original

[16] Note that the number of observations for the providers is not the same across the different ESG dimensions. In particular Bloomberg has a substantially lower number of observations in the environmental dimension (on average 399).

[17] For the MSCI KLD dataset there seems to be some issues with the CUSIP code. The codes do not always have the same number of characters, and it seems that leading zeros are often truncated. Therefore, we re-fill leading zeros if the number of characters is less then eight. Then we add the self-computed check digit to the code if the eighth number is not the would-be check digit if there would be an additional leading zero (in that case we add a leading zero) or the last two characters consist of commonly used issue codes.

[18] For US stocks the ISIN number is composed of the country code (first two characters), the CUSIP code (characters three to eleven), and a check digit.

names of the providers, by using some commonly used abbreviations to avoid rather trivial missmatches. We use the unique union of all three matching procedures to compile our sample.

To construct the sample we also require that at least two rating observations for each *legal origin* be available for each company. This choice provides us with an internally consistent sample and, in addition, it is not overly restrictive.

In addition, we use a monthly frequency for our sample. Asset 4, Sustainalytics and MSCI IVA already provide data at a monthly frequency; Inrate provides ratings update on a semi-annual basis for the years 2015 and 2016; and Bloomberg and MSCI KLD provide data on a yearly frequency. To convert from a semi-annual or annual frequency, we simply use the respective annual or semi-annual value for the whole time period. Note that most ratings (also for the providers with a monthly frequency) change rather infrequently, with most ratings being constant for about one year, but also for longer periods.[19]

4 Results

4.1 Descriptive statistics and correlations

[Table 2 about here.]

Table 2 shows summary statistics and Pearson correlations between the ratings from the six different data providers. We display the results for the total rating and the three E, S, and G pillars in separate panels. The different rating providers are ordered by their legal origin, civil law followed by common law. The first four columns display descriptive statistics for ranked ESG scores from the different providers. The following columns display the cross-correlations. We also display the average correlation between providers in the last row of each panel.

[19] Since the providers change their ratings at different points in time, we argue that for our purposes it makes sense to use a monthly frequency.

Each provider has a rather constant number of observations across the different scores they are issuing, with the exception of Bloomberg which has substantially lower coverage for environmental ratings. Regarding the correlations, we first observe that the correlations for the overall ESG ratings is 0.46 on average, which is much lower than correlations prevailing among credit rating providers which average 0.99 (between Moody's and S&P ratings) according to Berg et al. (2019). Also, the overall correlation between providers is lower for the E, S, and G subcategories (Panels B–D) than for the total pillar, which is probably due to discrepancies in aggregation procedures across the three pillars. Surprisingly, the average correlation is lowest for the governance and highest for the environmental ratings. Some other interesting features emerge in Panel C (Social pillar), in which we observe a weaker relation between the civil law providers compared to the other panels (especially the relation between the social ratings provided by Inrate and the other two providers). Furthermore, in Panel D, we observe a similar decrease in correlations between the governance scores, but for the common law ratings providers. These results are descriptive and are intended to provide initial insights for the following analysis.

4.2 Determinants of ESG rating disagreement

In this section, we briefly discuss whether ESG rating disagreement correlates with standard firm-level variables. We explore the role of variables falling in one of the following five categories: (i) Balance sheet related, (ii) Industry related, (iii) Investor transparency, (iv) Valuation, and (v) Price.[20]

We use pooled panel regressions in which the disagreement measures serve as de-

[20] **(i) Balance sheet related:** Tangibility (*TAN*), current ratio (*CR*), leverage (*LEV*), gross profitability (*GP*) (Novy-Marx, 2013); **(ii) Industry related:** Industry concentration measured by the Herfindahl-Hirschman index (*HHI*) based on book equity, multi-segment (*MSEG*); **(iii) Investor transparency:** Missing credit rating (*NCR*), institutional ownership (*IO*), number of analysts (*NoA*), the dispersion of analyst forecasts of the firm's one year ahead earnings forecasts (*StdDev*) (Diether et al., 2002); **(iv) Valuation:** Book-to-market (*BM*)(Fama and French, 1995); **(v) Price:** market cap (*ME*) (Banz, 1981), momentum (*MOM*) (Jegadeesh and Titman, 1993), and total volatility (*TVOL*) (Ang et al., 2006).

pendent variables.[21] We also include industry and time fixed effects. Standard errors are double clustered at the firm and month level. In Table 3 we display the results for disagreement about the total rating (column 1) and separately for disagreement about the E, S, and G dimension of the rating (columns 2-4).

[Table 3 about here.]

Essentially, three variables play a significant role in explaining ratings disagreement. First, more profitable firms are subject to lower ESG rating disagreement (see columns 1 and 2). Secondly, firms without a credit rating *(NCR)* exhibit higher disagreement (see columns 1 and 3), as do larger firms (see columns 1, 3, and 4).[22] These results seem intuitive: Profitable firms may be viewed less critical by ESG analysts, whereas firms without a credit rating are subject to a less transparent information environment, making their assessment in terms of ESG more difficult. Finally, larger firms might be more diversified and complex and are further analyzed more thoroughly, explaining why they exhibit higher ratings disagreement.[23]

[Figure 1 about here.]

In Figure 1, we look at whether disagreement varies at the industry-level. We plot average disagreement in each of the twelve Fama and French industries. There seems to be some industry heterogeneity when it comes do ESG rating disagreement. Disagreement

[21] We use here the standard deviation of ratings to measure the ESG rating disagreement, but we observe similar results for the range dispersion measure.

[22] The reader might wonder why S&P 500 firms do not have a credit rating. In general firms without a credit rating do not seem to be exceptional. For example, in a sample of 12,312 firms, Avramov et al. (2009) report that 9,051 firms do not have a credit rating. In our sample 194 firms have no credit rating for at least one month, from out of 553 firms in total. Whence 22 of the firms with not credit rating have zero leverage for at least one month.

[23] Tangibility *(TAN)* plays a role in explaining environmental rating disagreement in that firms with more tangible assets have lower disagreement in the environmental rating. Again, this is highly intuitive given that firms with more tangible assets are also likely to have a more easily inferred environmental score.

in the environmental and social ratings are highest in the consumer durables and telecommunications sectors (see subfigures 1b and 1c). In contrast, ESG data providers seem to disagree the most when it comes to governance ratings of financial services companies (see subfigure 1d).

4.3 ESG rating disagreement and stock returns

We now test our two main hypotheses by examining the relationship between stock returns and ESG rating disagreement. As in the previous section, we use pooled panel regressions with standard errors double clustered in the firm and time dimension. We use monthly stock returns as the dependent variable in the regressions. Besides our main disagreement related explanatory variables, we include industry-month fixed effects and also control for standard characteristics that have been found to explain the cross section of stock returns.[24] Conceptually this is similar to Fama and MacBeth (1973) regressions with industry dummies. Given that our sample period is relatively short, it is important to control for return differences at the industry-level.

We use two measures of ESG rating disagreement, namely the standard deviation of ratings available for a given firm at a given point in time and the range between the highest and the lowest rating. We denote the standard deviation based measure as *Std Dev* and the range based measure as *Range*. We calculate these disagreement measures using all ratings (*All*) and also separately using only ratings from data providers with a civil (*Civil*) or common law (*Common*) background.

[Table 4 about here.]

The regression results are displayed in Table 4. In column (1) the main explanatory

[24] We control for market capitalization (Banz, 1981), book-to-market (Fama and French, 1995), gross profitability (Novy-Marx, 2013), momentum (Jegadeesh and Titman, 1993), the dispersion of analyst forecasts of the firm's one year ahead earnings forecasts (Diether et al., 2002), the firm's beta (Frazzini and Pedersen, 2014) and total volatility (Ang et al., 2006)

variable is the standard deviation of all ratings. In column (2), we use ratings disagreement stratified by civil and common law providers separately. In Columns (3) and (4) we repeat the same procedure for the range based disagreement measures. In Panel A, we report results for the overall ESG rating, and in panels B, C, and D separately for the E, S, and G pillars.

We first test Hypothesis 1 which implies a positive relationship between disagreement in the environmental rating and future stock returns. Looking at panel B of Table 4, there seems to be some evidence that disagreement in environmental ratings is significantly positively related with future stock returns, when the disagreement is measured by the standard deviation. This evidence is consistent with the risk-based explanation (with or without excessive pessimism) advocated by Atmaz and Basak (2018).

Next, we test Hypothesis 2 which focuses on the role of ratings disagreement for stock returns when these ratings concern social and governance issues while also considering the legal origin of the country in which the data providers are based. The premise is that rating agencies with a civil law origin are more skilled in identifying relevant social issues. In a similar vein, rating agencies with a common law origin are more skilled in determining relevant governance issues. When there is disagreement in social (governance) ratings among civil (common) law rating providers, this would lead to lower future returns in the sense of Miller (1977) and Atmaz and Basak (2018) in the presence of excessive optimism. The empirical results support Hypothesis 2. Columns (2) and (4) in Panel C show that stock returns for firms subject to more disagreement in social ratings issued by civil law data providers tend to have lower stock returns. In a similar spirit, we also find support for Hypothesis 2 when looking at disagreement by common law data providers about governance ratings. Firms with more disagreement in governance ratings exhibit lower future stock returns.

In summary, our results indicate that the risk–based explanation advocated in Hy-

pothesis 1 is weakly confirmed by positive relationship between disagreement in environmental ratings and subsequent stock returns. In contrast, the mispricing explanation coupled with the legal origin of the data providers as stated in hypothesis 2 is able to explain the negative relationship between the social (governance) rating disagreement and future stock returns when these ratings are established by data providers domiciled in civil (common) law countries.

4.3.1 Role of the stock market cycle

Recent literature suggests that sustainable practices may pay off in bad times (e.g. recessions). For example, Lins et al. (2017) show that firms with high social capital had four to seven percentage points higher returns as firms with low social capital (proxied by KLD data) during the 2008–2009 financial crisis. Hence, a related question is whether ESG rating dispersion also displays a distinct effect on stock returns across stock market cycles. We conjecture that at the top of the stock market, when the probability of a bear market is high, the risk–based explanation (perhaps coupled with excessive pessimism) advocated by Atmaz and Basak (2018) may dominate. Since there are no major crises during our sample period, we pursue a slightly different approach. More specifically, we define a *bull dummy*, which is set to one if the return of the S&P 500 index in a given month is above the 75th percentile of the overall S&P500 return distribution during our sample period. We then add an interaction term between the dummy and our measures of ESG rating dispersion to the return regressions outlined above. Since we want to explore the time-series behavior of the relationship between stock returns and ESG disagreement, we do not include time dummies in the specifications.

The results for the standard deviation dispersion measure are displayed in Table 5. For all four ESG pillars, we find a significant and negative coefficient for ESG rating dispersion in non-bull states. This is consistent with investors being too optimistic (Hy-

pothesis 1). In contrast, in bullish states of the market (that is when the bull dummy is equal to one), we find a significantly different coefficient as in non-bull states. For example, for the total pillar the coefficient on the disagreement measure in the non-bull state is -5.8, and the coefficient of the interaction with the bull-dummy is 23.1, which means that bull-state coefficient is 17.3 (since the coefficients of the E, S, and G pillars are very similar, we do not discuss them in detail here). Therefore, in bull-states, we find evidence consistent with a risk-based story (i.e. investors are compensated with higher returns for buying stocks with higher ESG rating disagreement) for all three rating pillars. We find similar results for the range dispersion measure.[25]

[Table 5 about here.]

4.3.2 Additional tests

For completeness, we also analyze the relation between future stock returns and the average total ESG rating (and its individual E,S, and G components). In Table 4, we look at the average across all providers (column 5) and stratified by the legal origin of the providers (column 6). We do not find strong evidence that future stock returns depend on the average ESG ratings, apart from the average environmental rating, that turns out to be marginally significant and positive. This latter observation is consistent with the view that good environmental policies, as measured by environmental ratings, affects future stock returns positively. (Gibson Brandon and Krüger, 2018).

Past research has also examined the relationship between stock return volatility and heterogeneous beliefs in financial markets. This literature generally finds that that stock volatility is monotonically increasing with belief dispersion (e.g. Ajinkya and Gift, 1985; Anderson et al., 2005; Banerjee and Kremer, 2010; Atmaz and Basak, 2018). We also test this relationship in our setting, but do not observe a significant relation between stock

[25] Results are available on request.

return volatility and disagreement with respect to non-financial information.[26] In addition, we also examine if downside risk plays a role (we use the lower partial moment and the value at risk, see for example Hoepner et al. (2019)), but do not find a significant relation between downside risk measures and ESG ratings disagreement either.[27]

5 Conclusion

In this paper, we examine the relation between stock returns and ESG rating disagreement. Recently, the issue of ESG rating disagreement has received considerable attention, for instance, from the financial press and practitioner circles. In addition, ESG rating disagreement has important implications for the generalization of academic research findings and is creating challenges for asset managers in their efforts of implementing ESG investment strategies. To date, there is relatively little quantitative research on ESG ratings disagreement and we provide a first step towards a better understanding of the impact of ESG rating disagreement on stock returns.[28]

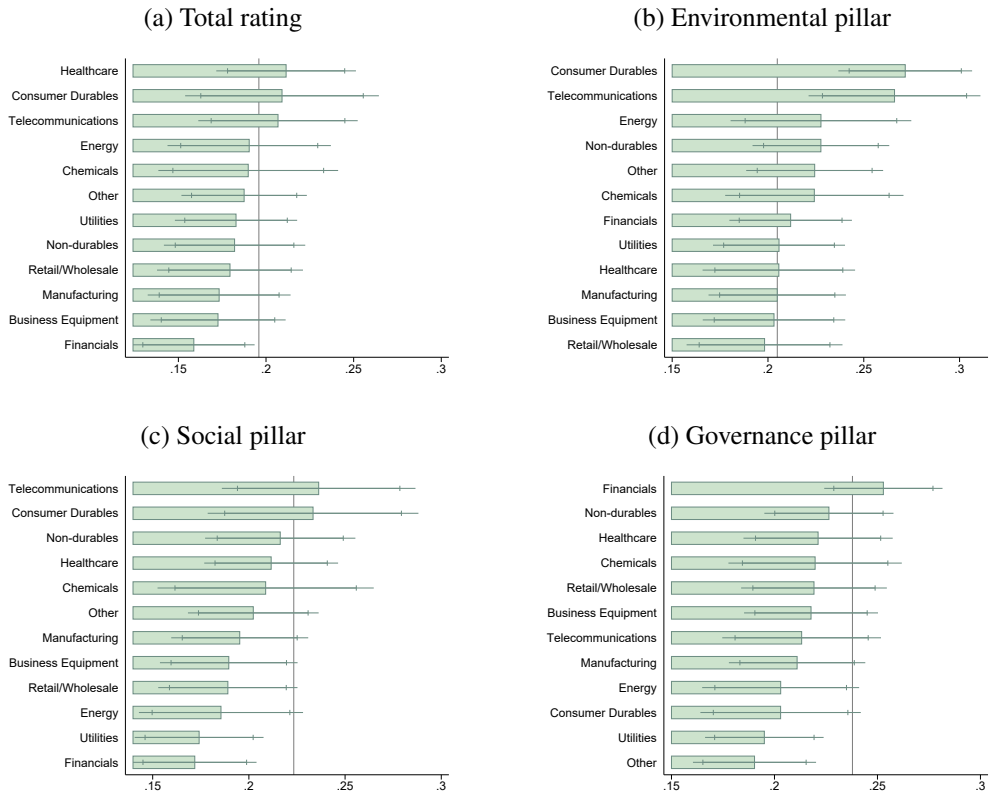
Using ESG ratings from six different information intermediaries, we document that the average correlation between the total ESG ratings from the six different providers is about 0.46. Surprisingly, the average correlation is lowest for governance and highest for environmental ratings. We also show that disagreement is higher for larger and less profitable firms as well as for firms that do not have a credit rating. We then examine the relation between stock returns and ESG rating disagreement. Motivated by theoretical arguments on the role of heterogeneous beliefs in financial markets combined with insights from the law and finance literature, we hypothesize that dispersion in ESG ratings should negatively predict stock returns whenever the ESG information is likely to be

[26] These results are not reported, but are available upon request.
[27] See Appendix A for detailed results.
[28] With the exception of Christensen et al. (2019), which focus on the determinants of ESG rating disagreement and Berg et al. (2019), which analyze ESG rating disagreement on a more granular level.

more financially relevant and there is excessive optimism. We argue that given the stakeholder centric view that shapes corporate governance in civil law countries and the more shareholder centric view typically found in common law countries, disagreement about social (governance) related non-financial information from civil (common) law information should lead to overvaluation of firms and thus low subsequent stock returns. For the environmental rating, we observe a positive relationship between rating disagreement and stock returns which suggests that a risk-based explanation advocated by Atmaz and Basak (2018) holds and may be even reinforced by excessive pessimism on this issue. We find evidence supporting these conjectures. Finally, we observe that in extreme bull market conditions, the risk–based explanation – perhaps coupled with excessive pessimism – prevails irrespective of the rating pillar considered and irrespective of the legal origin of the ESG ratings providers.

Our results are the first to document the subtle and multi-faceted implications of disagreement in ESG information on stock returns and thus on firms' cost of capital. They also have important consequences for responsible investors who rely on a one single data provider for ESG ratings used in their investment strategies but fail to account for ESG rating disagreement among data providers and the subtle and time-varying impact of disagreement on future stock returns.

Figure 1: ESG rating disagreement by Fama and French 12 industry



Note: This figure plots average disagreement for each of the twelve Fama and French industries. We report average industry-level disagreement for the total rating in subfigure 1a and the respective ESG components in subfigures 1b, 1c, and 1d. The vertical line represents the average disagreement across all industries. The horizontal line indicates the 90% confidence interval, while the whiskers indicate the 95% confidence interval.

Table 1: ESG data providers

Data provider	Legal origin	Origin	Rating scale	Number of stocks (sample)	Pillars
	(1)	(2)	(3)	(4)	(5)
Asset4	Civil law	CH	0 – 100	439	E, S, G, Total
Sustainalytics	Civil law	NL	0 – 100	460	E, S, G, Total
Inrate	Civil law	CH	1 – 12	432	E, L, S, G, Total
Bloomberg	Common law	US	0 – 100	456	E, S, G, Total
MSCI KLD	Common law	US	-1 – +1	457	E, S, G, Total
MSCI IVA	Common law	US	0 – 10	460	E, S, G, Total

Note: This table provides an overview of the ESG data providers which we use in this study. We list the name of the respective data provider (Data provider), the legal origin of each provider (Legal origin), the country in which the data provider has its origins (Origin), the rating scale used by the respective data provider (Rating scale), the average number of stocks per year in the sample for the total rating of each provider (Number of stocks (sample)), and the data dimensions (e.g., environmental, social, and governance) that are available from each provider. We refer to these data dimensions as Pillars.

Table 2: Descriptive statistics and correlations

	N	Mean	Median	StdDev	Pearson correlations				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Asset4	Sust.	Inrate	Bloom.	KLD
Panel A: Total rating									
Asset4	26,313	0.501	0.501	0.289					
Sustainalytics	27,592	0.498	0.499	0.289	0.768				
Inrate	25,945	0.501	0.534	0.284	0.233	0.303			
Bloomberg	27,349	0.501	0.501	0.289	0.747	0.719	0.122		
KLD	27,434	0.501	0.507	0.288	0.587	0.619	0.290	0.537	
MSCI IVA	27,587	0.501	0.502	0.289	0.428	0.469	0.319	0.316	0.469
Average correlation							**0.462**		
Panel B: Environmental Pillar									
Asset4	26,245	0.501	0.501	0.289					
Sustainalytics	27,516	0.501	0.501	0.289	0.714				
Inrate	25,880	0.501	0.518	0.286	0.305	0.488			
Bloomberg	23,941	0.501	0.501	0.289	0.650	0.569	0.206		
KLD	27,423	0.501	0.498	0.280	0.643	0.659	0.422	0.481	
MSCI IVA	27,522	0.501	0.502	0.289	0.154	0.313	0.403	0.120	0.276
Average correlation							**0.427**		
Panel C: Social Pillar									
Asset4	26,313	0.501	0.501	0.289					
Sustainalytics	27,592	0.501	0.504	0.289	0.618				
Inrate	25,945	0.501	0.522	0.288	0.133	0.143			
Bloomberg	27,261	0.501	0.508	0.288	0.679	0.541	0.061		
KLD	27,434	0.501	0.521	0.288	0.392	0.423	0.128	0.297	
MSCI IVA	27,587	0.501	0.500	0.289	0.299	0.330	0.236	0.208	0.390
Average correlation							**0.325**		
Panel D: Governance Pillar									
Asset4	26,313	0.501	0.501	0.289					
Sustainalytics	27,592	0.505	0.504	0.289	0.315				
Inrate	25,945	0.501	0.502	0.283	0.297	0.401			
Bloomberg	27,349	0.501	0.512	0.284	0.413	0.361	0.343		
KLD	27,434	0.501	0.489	0.230	-0.026	-0.040	0.083	0.009	
MSCI IVA	27,587	0.501	0.501	0.288	0.155	0.139	0.144	0.049	0.174
Average correlation							**0.188**		

Note: This table shows summary statistics and Pearson correlations between the ratings of the six different data providers. The results are displayed in separate panels for the Total rating and the E, S, and G components. The first four columns show the descriptive statistics of the different ESG providers' ranked scores (number of observations (*N*), mean (*Mean*, median (*Median*) and standard deviation (*StdDev*)). The following columns display the cross-correlations. We also display the average correlation between providers all providers (all), as well as for common law (common) and civil law (civil) providers in the last row of each panel.

Table 3: Determinants of ESG rating disagreement

Dependent variable:	ESG rating disagreement			
	(1)	(2)	(3)	(4)
Pillars:	Total	Environmental	Social	Governance
Balance Sheet related				
TAN	-0.007	**-0.051*****	-0.010	-0.018
	(0.018)	(0.018)	(0.018)	(0.018)
CR	0.017	-0.002	-0.001	0.011
	(0.013)	(0.012)	(0.013)	(0.011)
LEV	-0.015	-0.017	0.005	-0.006
	(0.012)	(0.012)	(0.012)	(0.010)
GP	**-0.038***	**-0.040****	-0.022	-0.015
	(0.020)	(0.018)	(0.018)	(0.017)
Industry				
HHI	0.024	0.017	0.002	-0.013
	(0.016)	(0.017)	(0.017)	(0.013)
MSEG	-0.004	-0.002	-0.001	-0.005
	(0.006)	(0.006)	(0.006)	(0.005)
Investor Transparency				
NCR	**0.027****	-0.015	**0.025****	-0.013
	(0.011)	(0.011)	(0.010)	(0.008)
IO	0.002	0.009	-0.012	**-0.021****
	(0.012)	(0.012)	(0.012)	(0.010)
NoA	-0.011	-0.005	0.002	0.005
	(0.011)	(0.012)	(0.012)	(0.010)
StdDev	0.006	0.002	**0.021****	-0.000
	(0.010)	(0.009)	(0.010)	(0.008)
Valuation				
BM	0.015	-0.006	**0.033****	0.014
	(0.016)	(0.015)	(0.015)	(0.013)
Price				
ME	**0.028***	0.011	**0.032****	**0.032****
	(0.015)	(0.014)	(0.015)	(0.012)
MOM	-0.008	-0.005	-0.003	-0.008
	(0.006)	(0.006)	(0.006)	(0.005)
TVOL	-0.009	-0.004	-0.006	-0.007
	(0.012)	(0.012)	(0.012)	(0.010)
Fixed Effects	Yes	Yes	Yes	Yes
N	21,199	21,160	21,199	21,199
Adjusted R^2	0.063	0.078	0.061	0.052

Note: This table display the results of pooled panel regressions separated into panels for T, E, S and G components. The dependent variables are our ESG ratings' dispersion measures respectively the standard deviation of all firm-level ratings available at a given point in time.The explanatory variables are the following: tangibility (*TAN*), current ratio (*CR*), leverage (*LEV*), gross profitability (*GP*) (Novy-Marx, 2013), Herfindahl-Hirschman index (*HHI*), multi-segment (*MSEG*), missing credit rating (*NCR*), institutional ownership (*IO*), number of analysts (*NoA*), the dispersion of analyst forecasts of the firm's one year ahead earnings forecasts (*StdDev*) (Diether et al., 2002), book-to-market (*BM*) (Fama and French, 1995), market cap (*ME*) (Banz, 1981), momentum (*MOM*) (Jegadeesh and Titman, 1993), and total volatility (*TVOL*) (Ang et al., 2006). We also include time-industry fixed effects. Clustered standard errors in parentheses.
Significance levels at 10% (*), 5% (**) and 1% (***).

Table 4: Stock returns and ESG rating disagreement

Dependent variable:	Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
Main explanatory variables:	Std Dev		Range		Average ESG	
Panel A: Total Pillar						
All	-0.291		-0.151		0.177	
	(0.519)		(0.192)		(0.316)	
Common		-0.372		-0.214		0.007
		(0.349)		(0.182)		(0.365)
Civil		0.214		0.116		0.168
		(0.369)		(0.190)		(0.389)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,913	25,913	25,913	25,913	25,913	25,913
Adjusted R^2	0.296	0.296	0.296	0.296	0.296	0.296
Panel B: Environmental Pillar						
All	**1.118***		0.341		**0.513***	
	(0.562)		(0.222)		(0.307)	
Common		0.475		0.224		**0.549***
		(0.361)		(0.194)		(0.288)
Civil		0.209		0.128		-0.008
		(0.430)		(0.225)		(0.270)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,850	25,850	25,850	25,850	25,850	25,850
Adjusted R^2	0.297	0.297	0.297	0.297	0.297	0.297
Panel C: Social Pillar						
All	-0.318		-0.148		-0.099	
	(0.553)		(0.208)		(0.318)	
Common		0.442		0.214		-0.249
		(0.384)		(0.202)		(0.283)
Civil		**-0.917****		**-0.475****		0.156
		(0.365)		(0.194)		(0.323)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,913	25,913	25,913	25,913	25,913	25,913
Adjusted R^2	0.296	0.297	0.296	0.297	0.296	0.296
Panel D: Governance Pillar						
All	-0.639		-0.282		0.393	
	(0.581)		(0.239)		(0.399)	
Common		**-0.819****		**-0.438****		0.076
		(0.364)		(0.188)		(0.287)
Civil		-0.089		-0.102		0.264
		(0.343)		(0.178)		(0.288)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,913	25,913	25,913	25,913	25,913	25,913
Adjusted R^2	0.296	0.297	0.296	0.297	0.296	0.296

Note: This table displays the results of pooled panel regressions of monthly stock returns on ESG rating disagreement (columns 1–4) as well as regressions of monthly returns on the average ESG rating (columns 5 and 6). The results are separated into four panels: Panel A reports the results for the total ESG rating and Panels B, C, and D report results for the E, S, and G component separately. The dependent variable *Returns* is the firm's monthly stock return. We use two ways of measuring ESG ratings disagreement. In columns 1 and 2, we use the standard deviation of ratings available for a given firm at a given point in time (*Std Dev*). In columns 3 and 4 we use the range between the highest and the lowest rating (*Range*). We calculate these disagreement measures using all ratings *(All)* and also separately using ratings issued only by data providers from civil law *(Civil)* or common-law *(Common)* countries. In columns 5 and 6, the main independent variables are the average ESG ratings. We also include time-industry fixed effects and control for standard characteristics that have been found to explain stock returns, namely market capitalization (Banz, 1981), book-to-market (Fama and French, 1995), gross profitability (Novy-Marx, 2013), momentum (Jegadeesh and Titman, 1993), the dispersion of analyst forecasts of the firm's one year ahead earnings forecasts (Diether et al., 2002), the firm's beta (Frazzini and Pedersen, 2014), and total volatility (Ang et al., 2006).

Double clustered standard errors in parentheses. Significance levels at 10% (*), 5% (**) and 1% (***).

Electronic copy available at: https://ssrn.com/abstract=3433728

Table 5: Role of the stock market cycle

Dependent variable:	Returns			
	(1)	(2)	(3)	(4)
Pillars:	Total	Environmental	Social	Governance
All	**-5.774*****	**-4.081*****	**-5.136*****	**-5.277*****
	(1.510)	(1.395)	(1.410)	(1.296)
All*Bull	**23.054*****	**22.663*****	**20.849*****	**19.940*****
	(2.301)	(2.264)	(2.053)	(1.945)
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Time FE	No	No	No	No
Observations	25,913	25,850	25,913	25,913
Adjusted R^2	0.085	0.088	0.088	0.090

Note: This table displays the results of pooled panel regressions of monthly stock returns on ESG rating disagreement (columns 1–4). We interact the ESG variable with a *bull market dummy*. This dummy is set to one if the return of the S&P 500 index is above its 75th percentile. This table reports results for each different pillar in columns 1–4. The dependent variable *Returns* is the firm's monthly stock return. We use here the standard deviation of ratings available for a given firm at a given point in time to measure ESG ratings disagreement. We calculate this disagreement measure using all ratings *(All)*. We also include industry fixed effects and control for standard characteristics that have been found to explain stock returns, namely market capitalization (Banz, 1981), book-to-market (Fama and French, 1995), gross profitability (Novy-Marx, 2013), momentum (Jegadeesh and Titman, 1993), the dispersion of analyst forecasts of the firm's one year ahead earnings forecasts (Diether et al., 2002), the firm's beta (Frazzini and Pedersen, 2014), and total volatility (Ang et al., 2006).
Double clustered standard errors in parentheses. Significance levels at 10% (*), 5% (**) and 1% (***).

References

Ajinkya, B. B. and Gift, M. J.: 1985, Dispersion of financial analysts' earnings forecasts and the (option model) implied standard deviations of stock returns, *Journal of Finance* **40**(5), 1353–1365.

Anderson, E. W., Ghysels, E. and Juergens, J. L.: 2005, Do heterogeneous beliefs matter for asset pricing?, *The Review of Financial Studies* **18**(3), 875–924.

Ang, A., Hodrick, R. J., Xing, Y. and Zhang, X.: 2006, The cross-section of volatility and expected returns, *The Journal of Finance* **61**(1), 259–299.

Atmaz, A. and Basak, S.: 2018, Belief dispersion in the stock market, *The Journal of Finance* **73**(3), 1225–1279.

Avramov, D., Chordia, T., Jostova, G. and Philipov, A.: 2009, Dispersion in analysts' earnings forecasts and credit rating, *Journal of Financial Economics* **91**(1), 83–101. **URL:** *http://www.sciencedirect.com/science/article/pii/S0304405X08001840*

Banerjee, S. and Kremer, I.: 2010, Disagreement and learning: Dynamic patterns of trade, *The Journal of Finance* **65**(4), 1269–1302.

Banz, R. W.: 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* **9**(1), 3–18.

Basak, S.: 2000, A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk, *Journal of Economic Dynamics and Control* **24**(1), 63–95.

Berg, F., Kölbel, J. F. and Rigobon, R.: 2019, Aggregate Confusion: The Divergence of ESG Ratings. Working Paper.

Berrada, T.: 2006, Incomplete information, heterogeneity, and asset pricing, *Journal of Financial Econometrics* **4**(1), 136–160.

Botero, J. C., Djankov, S., Porta, R. L., Lopez-de Silanes, F. and Shleifer, A.: 2004, The regulation of labor, *The Quarterly Journal of Economics* **119**(4), 1339–1382.

Bouten, L., Cho, C. H., Michelon, G. and Roberts, R. W.: 2017, CSR performance proxies in large-sample studies: 'Umbrella advocates', construct clarity and the 'validity police'. Working Paper.

Carlin, B. I., Longstaff, F. A. and Matoba, K.: 2014, Disagreement and asset prices, *Journal of Financial Economics* **114**(2), 226–238.

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

Chatterji, A. K., Levine, D. I. and Toffel, M. W.: 2009, How well do social ratings actually measure corporate social responsibility?, *Journal of Economics & Management Strategy* **18**(1), 125–169. 00262.

Chen, J., Hong, H. and Stein, J. C.: 2002, Breadth of ownership and stock returns, *Journal of Financial Economics* **66**(2), 171–205.

Christensen, D., Serafeim, G. and Sikochi, A.: 2019, Why is Corporate Virtue in the Eye of the Beholder? The Case of ESG Ratings. Working Paper.

David, A.: 2008, Heterogeneous beliefs, speculation, and the equity premium, *The Journal of Finance* **63**(1), 41–83.

Delmas, M. A., Etzion, D. and Nairn-Birch, N.: 2013, Triangulating environmental performance: What do corporate social responsibility ratings really capture?, *Academy of Management Perspectives* **27**(3), 255–267.

Detemple, J. and Murthy, S.: 1994, Intertemporal asset pricing with heterogeneous beliefs, *Journal of Economic Theory* **62**(2), 294–320.

Diether, K. B., Malloy, C. J. and Scherbina, A.: 2002, Differences of opinion and the cross section of stock returns, *The Journal of Finance* **57**(5), 2113–2141.

Doidge, C., Karolyi, A. G. and Stulz, R. M.: 2007, Why do countries matter so much for corporate governance?, *Journal of Financial Economics* **86**(1), 1–39.

Doyle, T. M.: 2018, Ratings that don't rate – the subjective view of ESG rating agencies, *Technical report*, ACCF American Council for Capital Formation.
URL: *https://bit.ly/2LBwvky*

Dumas, B., Kurshev, A. and Uppal, R.: 2009, Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility, *The Journal of Finance* **64**(2), 579–629.

Dyck, A., Lins, K. V., Roth, L. and Wagner, H. F.: 2019, Do institutional investors drive corporate social responsibility? International evidence, *Journal of Financial Economics* **131**(3), 693–714.

Eccles, R. G., Lee, L.-E. and Stroehle, J. C.: 2019, The social origins of ESG? An analysis of Innovest and KLD. Working Paper.

Eccles, R. G. and Stroehle, J. C.: 2018, Exploring social origins in the construction of ESG measures. Working Paper.

Fama, E. F. and French, K. R.: 1995, Size and book-to-market factors in earnings and returns, *The Journal of Finance* **50**(1), 131–155.

Fama, E. F. and MacBeth, J. D.: 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* **81**(3), 607–636.

Frazzini, A. and Pedersen, L. H.: 2014, Betting against beta, *Journal of Financial Economics* **111**(1), 1–25.

Gibson Brandon, R. and Krüger, P.: 2018, The sustainability footprint of institutional investors, *Swiss Finance Institute Research Paper* (17-05).

Gibson, R., Glossner, S., Krüger, P., Matos, P. and Steffen, T.: 2019, Responsible institutional investing around the world. University of Geneva and University of Virginia.

GSIA Global Sustainable Investment Alliance: 2016, Global sustainable investment review 2016, `https://bit.ly/2U1OuoF`.

Harris, M. and Raviv, A.: 1993, Differences of opinion make a horse race, *The Review of Financial Studies* **6**(3), 473–506.

Harrison, J. M. and Kreps, D. M.: 1978, Speculative investor behavior in a stock market with heterogeneous expectations, *The Quarterly Journal of Economics* **92**(2), 323–336.

Hoepner, A. G., Oikonomou, I., Sautner, Z., Starks, L. T. and Zhou, X.: 2019, ESG shareholder engagement and downside risk. Working Paper.

Hong, H. and Kostovetsky, L.: 2012, Red and blue investing: Values and finance, *Journal of Financial Economics* **103**(1), 1–19.

Jegadeesh, N. and Titman, S.: 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of Finance* **48**(1), 65–91.

Jouini, E. and Napp, C.: 2007, Consensus consumer and intertemporal asset pricing with heterogeneous beliefs, *The Review of Economic Studies* **74**(4), 1149–1174.

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

La Porta, R.: 1996, Expectations and the cross-section of stock returns, *Journal of Finance* **51**(5), 1715–1742.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. W.: 1998, Law and finance, *Journal of Political Economy* **106**(6), 1113–1155.

Lakonishok, J., Shleifer, A. and Vishny, R. W.: 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* **49**(5), 1541–1578.

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

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

Mackintosh, J.: 2018, Is Tesla or Exxon more sustainable? It depends whom you ask, *Wall Street Journal* .
URL: *https://on.wsj.com/2MQCC4m*

Miller, E. M.: 1977, Risk, uncertainty, and divergence of opinion, *The Journal of Finance* **32**(4), 1151–1168.

Novy-Marx, R.: 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* **108**(1), 1–28.

PRI Principles for Responsible Investment: 2018, Annual report 2018, https://bit.ly/2NsU8xv.

Skinner, D. J. and Sloan, R. G.: 2002, Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio, *Review of Accounting Studies* **7**(2), 289–312.

USSIF Forum for Sustainable and Responsible Investment: 2018, Report on US sustainable, responsible and impact investing trends 2018, https://bit.ly/2yZGHiG. Washington DC.

Wigglesworth, R.: 2018, Rating agencies using green criteria suffer from 'inherent biases', *Financial Times* .
URL: *https://on.ft.com/2H5E0kj*

Yu, J.: 2011, Disagreement and return predictability of stock portfolios, *Journal of Financial Economics* **99**(1), 162–183.

Zapatero, F.: 1998, Effects of financial innovations on market volatility when beliefs are heterogeneous, *Journal of Economic Dynamics and Control* **22**(4), 597–626.

A Appendix: Downside risk measures and ESG ratings disagreement

Table A.1: LPM and ESG rating disagreement

Dependent variable:	LPM					
	(1)	(2)	(3)	(4)	(5)	(6)
Main explanatory variables:	Std Dev		Range		Average ESG	
Panel A: Total Pillar						
All	0.085		0.017		**-0.090***	
	(0.118)		(0.045)		(0.051)	
Common		0.038		0.023		0.035
		(0.077)		(0.041)		(0.059)
Civil		0.019		-0.005		**-0.126****
		(0.084)		(0.046)		(0.063)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,924	25,924	25,924	25,924	25,924	25,924
Adjusted R^2	0.338	0.338	0.338	0.338	0.340	0.341
Panel B: Environmental Pillar						
All	-0.083		-0.038		-0.079	
	(0.123)		(0.046)		(0.054)	
Common		-0.067		-0.045		-0.073
		(0.076)		(0.041)		(0.063)
Civil		-0.031		-0.034		-0.010
		(0.085)		(0.045)		(0.062)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,860	25,860	25,860	25,860	25,860	25,860
Adjusted R^2	0.338	0.338	0.338	0.339	0.339	0.339
Panel C: Social Pillar						
All	0.076		0.030		-0.086	
	(0.114)		(0.044)		(0.055)	
Common		-0.006		-0.003		0.069
		(0.066)		(0.035)		(0.056)
Civil		0.071		0.026		**-0.153****
		(0.079)		(0.043)		(0.058)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,924	25,924	25,924	25,924	25,924	25,924
Adjusted R^2	0.338	0.338	0.338	0.338	0.340	0.342
Panel D: Governance Pillar						
All	0.043		0.016		**-0.170****	
	(0.125)		(0.046)		(0.068)	
Common		-0.046		-0.033		-0.004
		(0.073)		(0.038)		(0.052)
Civil		-0.075		-0.053		**-0.139*****
		(0.080)		(0.042)		(0.051)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,924	25,924	25,924	25,924	25,924	25,924
Adjusted R^2	0.338	0.338	0.338	0.339	0.343	0.344

Note: This table displays the results of pooled panel regressions of the monthly lower partial moment on ESG rating disagreement (columns 1–4) as well as regressions of monthly total volatility on the average ESG rating (columns 5 and 6). The results are separated into four panels: Panel A reports the results for the total ESG rating and Panels B, C, and D report results for the E, S, and G component separately. The dependent variable *LPM* is the monthly lower partial moment log transformed. We use two ways of measuring ESG ratings disagreement. In columns 1 and 2, we use the standard deviation of ratings available for a given firm at a given point in time (*Std Dev*). In columns 3 and 4 we use the range between the highest and the lowest rating (*Range*). We calculate these disagreement measures using all ratings *(All)* and also separately using ratings issued only by data providers from civil law *(Civil)* or common-law *(Common)* countries. In columns 5 and 6, the main independent variables are the average ESG ratings. We also include time-industry fixed effects and control for standard characteristics, namely market capitalization (Banz, 1981), book-to-market (Fama and French, 1995), gross profitability (Novy-Marx, 2013), momentum (Jegadeesh and Titman, 1993) and the dispersion of analyst forecasts of the firm's one year ahead earnings forecasts (Diether et al., 2002).
Double clustered standard errors in parentheses. Significance levels at 10% (*), 5% (**) and 1% (***).

Table A.2: VaR and ESG rating disagreement

Dependent variable:	VaR					
	(1)	(2)	(3)	(4)	(5)	(6)
Main explanatory variables:	Std Dev		Range		Average ESG	
Panel A: Total Pillar						
All	-0.005		-0.002		**0.002****	
	(0.003)		(0.001)		(0.001)	
Common		-0.000		-0.000		-0.001
		(0.002)		(0.001)		(0.001)
Civil		-0.003		-0.001		**0.004****
		(0.002)		(0.001)		(0.001)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,924	25,924	25,924	25,924	25,924	25,924
Adjusted R^2	0.371	0.371	0.371	0.371	0.373	0.375
Panel B: Environmental Pillar						
All	0.001		0.001		**0.003****	
	(0.003)		(0.001)		(0.001)	
Common		0.002		0.001		**0.003***
		(0.002)		(0.001)		(0.001)
Civil		-0.002		-0.001		0.000
		(0.002)		(0.001)		(0.001)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,860	25,860	25,860	25,860	25,860	25,860
Adjusted R^2	0.370	0.371	0.371	0.371	0.375	0.375
Panel C: Social Pillar						
All	-0.004		-0.002		0.001	
	(0.003)		(0.001)		(0.001)	
Common		-0.002		-0.001		-0.002
		(0.002)		(0.001)		(0.001)
Civil		-0.002		-0.001		**0.003****
		(0.002)		(0.001)		(0.001)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,924	25,924	25,924	25,924	25,924	25,924
Adjusted R^2	0.371	0.371	0.371	0.371	0.371	0.374
Panel D: Governance Pillar						
All	0.003		0.001		0.001	
	(0.003)		(0.001)		(0.002)	
Common		0.002		0.001		-0.001
		(0.002)		(0.001)		(0.001)
Civil		**0.003***		**0.002****		0.002
		(0.002)		(0.001)		(0.001)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	25,924	25,924	25,924	25,924	25,924	25,924
Adjusted R^2	0.371	0.372	0.371	0.374	0.370	0.372

Note: This table displays the results of pooled panel regressions of the monthly value at risk on ESG rating disagreement (columns 1–4) as well as regressions of monthly total volatility on the average ESG rating (columns 5 and 6). The results are separated into four panels: Panel A reports the results for the total ESG rating and Panels B, C, and D report results for the E, S, and G component separately. The dependent variable *VaR* is the monthly value at risk. We use two ways of measuring ESG ratings disagreement. In columns 1 and 2, we use the standard deviation of ratings available for a given firm at a given point in time (*Std Dev*). In columns 3 and 4 we use the range between the highest and the lowest rating (*Range*). We calculate these disagreement measures using all ratings *(All)* and also separately using ratings issued only by data providers from civil law *(Civil)* or common-law *(Common)* countries. In columns 5 and 6, the main independent variables are the average ESG ratings. We also include time-industry fixed effects and control for standard characteristics, namely market capitalization (Banz, 1981), book-to-market (Fama and French, 1995), gross profitability (Novy-Marx, 2013), momentum (Jegadeesh and Titman, 1993) and the dispersion of analyst forecasts of the firm's one year ahead earnings forecasts (Diether et al., 2002).
Double clustered standard errors in parentheses. Significance levels at 10% (*), 5% (**) and 1% (***).

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.org

ECGI Working Paper Series in Finance

www.ecgi.org\wp

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.org/wp) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html