

The Challenge of Disparities in ESG Ratings

Bruce I. Jacobs and Kenneth N. Levy

Bruce I. Jacobs

is the principal and co-chief investment officer of Jacobs Levy Equity Management in Florham Park, NJ.

bruce.jacobs@jlem.com

Kenneth N. Levy

is the principal and co-chief investment officer of Jacobs Levy Equity Management in Florham Park, NJ.

ken.levy@jlem.com

KEY FINDINGS

- Due to the lack of a common framework for ESG rating construction, ESG ratings on companies from different vendors show substantial disparities. These disparities have real consequences for asset owners, policymakers, academics, and asset managers.
- Investors need to be aware of and understand these disparities when choosing ESG rating data and implementing ESG integration.
- Despite noisy ESG ratings, disentangling may be suitable for assessing the efficacy of multidimensional ESG data in the presence of other characteristics impacting stock returns.

ABSTRACT

Environmental, social, and governance (ESG) considerations play an increasingly important role in investment decisions. Due to data vendors' lack of a common framework for creating ESG ratings, substantial disparities exist across vendors in their ESG ratings for the same company. ESG rating disparities make it difficult to assess whether ESG ratings are aligned with companies' ESG performance and how ESG investing affects investment performance. This presents challenges to asset owners, policymakers, academics, and asset managers. This article highlights the nature and sources of the ESG rating disparities and advises investors to understand these aspects of noisy ESG ratings and to exercise caution when implementing ESG integration.

With an estimated \$35 trillion of assets managed according to environmental, social, and governance (ESG) principles (Bloomberg Intelligence 2021), ESG has entered the mainstream of investing. Still, it bears underscoring that companies do not come with ESG labels investors can readily use to distinguish an ESG-caring firm from an ESG-neglecting firm. ESG characteristics are easy to describe in words (e.g., CFA Institute 2021), but much harder to measure compared to other firm characteristics, such as market capitalization, price-to-earnings ratio, or momentum. Yet many institutional investors, including pension funds, endowments, mutual funds, and ETFs, have integrated certain ESG principles into their investment decision making over the past decade, which has led to the rapid growth in ESG investing.¹

¹In 2021, the U.S. Department of Labor announced a proposed rulemaking that would afford ERISA plan fiduciaries more flexibility in considering ESG factors in investment decision-making (U.S. Department of Labor, October 14, 2021).

Asset managers implementing ESG integration first need to decide on a data vendor or vendors, acquire and process the data, and incorporate the processed data into the investment process.² Only when armed with reliable ESG data can asset managers ensure a rigorous implementation of an ESG mandate through a comprehensive research process.

Selecting an ESG rating vendor is challenging because there are numerous vendors and there is a great disparity in their ratings. The divergence in ESG ratings of different vendors has been documented by several authors (see Chatterji et al. 2016; Mackintosh 2018 and 2022; Wigglesworth 2018; Dimson, Marsh, and Staunton 2020; Matos 2020; Prall 2021; Berg, Koelbel, and Rigobon 2022; and Christensen, Serafeim, and Sikochi, 2022).

Mackintosh (2018) points out that, in 2018, Tesla was rated by MSCI at the top of the automotive industry on ESG issues, while FTSE ranked it at the bottom.³ A comprehensive study by Berg, Koelbel, and Rigobon (2022) using data of six major ESG rating vendors highlights the rating divergence from the different vendors.⁴ They analyze the rating divergence for the aggregate ESG rating and for each of the individual dimensions—environmental (E), social (S), and governance (G). They show that the correlations of the aggregate ESG ratings between different vendors are on average 0.54 and range from 0.38 to 0.71.⁵ For the individual E, S, and G dimensions, they find that the average correlations are 0.53, 0.42, and 0.30, respectively.

Among many studies on ESG rating divergence, Berg, Koelbel, and Rigobon (2022) provide insights into the underlying reasons of why ESG ratings disagree. They decompose rating divergence into three sources—divergences of scope, measurement, and weights. Scope divergence arises when ratings are based on different categories of attributes (examples of attributes are carbon emissions, labor practices, and business ethics), measurement divergence arises when raters measure the same category of attributes differently, and weights divergence arises when raters assess each category's relative importance differently in calculating the aggregate rating.

They find that the overall rating discrepancies are due primarily to divergences in scope and measurement, while weights divergence is much less important. Furthermore, they detect a rater effect, whereby a rater's assessment of one category for a firm is positively related to their assessment of other categories for the same firm. Thus, the assessment of ESG performance is influenced by a rater's general view of the firm, which suggests that averaging category ratings across vendors would be less effective against measurement divergence because the disagreements are not randomly distributed.

Brandon, Krueger, and Schmidt (2021) also analyze the magnitude of ESG rating disagreement and assess whether it correlates with a firm's financial and accounting characteristics. Their main focus is whether stock returns are related to ESG rating disagreement (the second moment of ESG ratings) and whether divergent ESG ratings have real consequences for firms and investors as a result.

Using ESG ratings for S&P 500 firms over the 2010–2017 period from seven major vendors, Brandon, Krueger, and Schmidt (2021) find that stock returns are

²The data vendor selection process involves evaluating the quality of data provided by each vendor, including coverage, accuracy, timeliness, and relevance (alpha opportunities), as well as the quality of each data vendor, including its willingness to be transparent about any potential biases.

³MSCI considered emissions from Tesla's cars, while FTSE considered emissions from Tesla's factories.

⁴The six ESG rating vendors examined in Berg, Koelbel, and Rigobon (2022) include KLD, Sustainability, Moody's ESG (Vigeo Eiris), S&P Global (RobecoSAM), Refinitiv (Asset4), and MSCI. See Eccles and Strohle (2018) for the history and establishment of the major ESG rating players.

⁵Compared to ESG ratings, debt credit ratings tend to have much stronger correlations among different vendors. For example, Prall (2021) reports that credit rating correlations range from 0.94 to 0.96 among the three main credit rating agencies (S&P, Moody's, and Fitch).

positively related to ESG rating disagreement and that the relationship is statistically significant.⁶ They provide two possible explanations for this finding. First, a disagreement may represent a risk for investors who demand a higher return to hold a stock when rating dispersion is high, in line with theories of heterogeneous beliefs in asset pricing. Second, disagreement about a firm's ESG rating may be a proxy for ESG uncertainty, and therefore captures a specific form of (Knightian) uncertainty. They also find that environmental rating disagreement is the sole driver of the return relationship, while disagreement about social and governance ratings is insignificantly related.

Both Berg, Koelbel, and Rigobon (2022) and Brandon, Krueger, and Schmidt (2021) provide substantive empirical evidence that ESG ratings from different vendors show significant disparities, and that they have nontrivial implications for investors. Portfolios based on different ESG ratings are likely to have different constituents, which could pose a problem for those who screen stocks based on ESG ratings.

Academic research has examined how adopting ESG principles can affect asset prices (see, e.g., Berg et al. 2021; Pastor, Stambaugh, and Taylor 2021a and 2021b; Pedersen, Fitzgibbons, and Pomorski 2021; and Avramov et al. 2022). Vendor rating divergence can affect research findings and in turn affect investment decisions.

Given the real-world consequences of vendors' rating divergence, there are growing calls for greater transparency. Berg, Koelbel, and Rigobon (2022) suggest that ESG raters should be more transparent in their definition of ESG in terms of scope, measurement, and weights. Comtois (2021) quotes CalPERS CEO Marcie Frost on the seriousness of the situation: "Sustainability is a cornerstone of the CalPERS investment program. And yet, we have found it challenging to effectively measure impact in our private equity portfolio because of the multitude of frameworks and definitions used by GPs and LPs."

Some ESG vendors are providing greater transparency into the design of their ESG scores by publishing in peer-reviewed journals. Lindeman (2022), which discusses ESG scores and sub-scores using Bloomberg's ESG data, is an example. Another promising initiative is the Aggregate Confusion Project at MIT,⁷ which seeks to address ESG ratings ambiguity by improving the quality of ESG measurement and decision making in the financial sector.

The Sustainability Accounting Standards Board (SASB) has identified industry-specific sustainability accounting standards, representing the subsets of environmental, social, and governance issues considered most relevant to financial performance in each industry (Sustainability Accounting Standards Board 2018). This could eventually provide a guide for ESG data standardization among the different vendors.

What can asset managers do right now about these consequential data disagreements? A typical asset manager may not subscribe to multiple vendors' data because the subscription and data management costs would be high and there would be no assurance that the disagreement effect will continue to generate excess returns.⁸

⁶The seven vendors include Refinitiv (Asset4), Sustainalytics, Inrate, Bloomberg, FTSE, KLD, and MSCI Intangible Value Assessment (IVA). Since rating scales are not uniform across these vendors, to make them comparable, Brandon, Krueger, and Schmidt (2021) convert the ratings of each vendor into ranked scores. Specifically, they sort all stocks according to the ratings of the respective providers, calculate the individual rating-specific percentile ranks, and normalize these ranks between 0 and 1. They measure rating disagreement—their main variable of interest—as the standard deviation of (normalized) ESG rating scores available for a given firm at a given point in time for the total rating and also separately for the environmental, social, and governance components.

⁷See <https://mitsloan.mit.edu/sustainability-initiative/aggregate-confusion-project>.

⁸Berg et al. (2021) develop a simple model that relates noisy ESG ratings to stock returns and show that the noisier the ESG rating, the lower its effect on stock returns. To tackle the downward bias caused by this noise problem, they propose using an instrumental variable approach for empirical analysis. Specifically, they instrument a rating of one vendor by the ratings of other vendors and apply two-stage least squares (2SLS) regressions. They document that the effect of ESG rating on stock returns is much stronger when the standard OLS is replaced by the 2SLS, suggesting that it is advantageous to rely on several complementary, although noisy, ratings.

For each ESG dataset considered, asset managers should evaluate its efficacy for their particular investment process. ESG ratings could be correlated with other firm characteristics. Quantitative managers can disentangle firms' ESG ratings or their individual ESG components from other firm characteristics included in their investment process.

Jacobs and Levy (1988, 2021) introduced a disentangling methodology that is implementable by systematic asset managers. We presented the idea of disentangling stock returns cross sectionally across numerous factors and identified the relationships between individual stock returns and firm characteristics. We also examined the benefits of using the resulting time series of returns to the disentangled factors for return forecasting. While noisy ESG ratings can attenuate regression coefficients, this methodology may nevertheless be suitable for assessing the efficacy of multidimensional ESG data in the presence of other characteristics impacting stock returns.

The disparity of ESG ratings across vendors makes it difficult to assess whether ESG ratings are aligned with companies' ESG performance and how ESG investing affects investment performance. This presents unique challenges to asset owners, policymakers, academics, and asset managers. Investors should understand the unique features of noisy ESG ratings data and exercise caution when implementing ESG integration.

ACKNOWLEDGMENTS

We appreciate the assistance of Arijit Banerjee and Sangwoo Lee, both from Jacobs Levy Equity Management.

REFERENCES

- Avramov, D., S. Cheng, A. Lioui, and A. Tarelli. 2022. "Sustainable Investing with ESG Rating Uncertainty." *Journal of Financial Economics* (forthcoming). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3711218.
- Berg, F., J. F. Koelbel, and R. Rigobon. 2022. "Aggregate Confusion: The Divergence of ESG Ratings." MIT Sloan School Working Paper, January 14, 2022. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3438533.
- Berg, F., J. F. Koelbel, A. Pavlova, and R. Rigobon. 2021. "ESG Confusion and Stock Returns: Tackling the Problem of Noise." MIT Sloan School Working Paper, November 19, 2021. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3941514.
- Bloomberg Intelligence. 2021. "ESG Assets Rising to \$50 Trillion Will Reshape \$140.5 Trillion of Global AUM by 2025, Finds Bloomberg Intelligence." <https://www.bloomberg.com/company/press/esg-assets-rising-to-50-trillion-will-reshape-140-5-trillion-of-global-aum-by-2025-finds-bloomberg-intelligence/>.
- Brandon, R. G., P. Krueger, and P. S. Schmidt. 2021. "ESG Rating Disagreement and Stock Returns." *Financial Analysts Journal* 77 (4): 104–127.
- CFA Institute. 2021. "What is ESG Investing?" <https://www.cfainstitute.org/research/esg-investing>.
- Chatterji, A. K., R. Durand, D. I. Levine, and S. Touboul. 2016. "Do Ratings of Firms Converge? Implications for Managers, Investors and Strategy Researchers." *Strategic Management Journal* 37 (8): 1597–1614.

Christensen, D. M., G. Serafeim, and A. (Siko) Sikochi. 2022. "Why Is Corporate Virtue in the Eye of the Beholder? The Case of ESG Ratings." *The Accounting Review* 97 (1): 147–175.

Comtois, J. 2021. "CalPERS, Carlyle Lead Initiative for Standardized ESG Reporting." *Pensions & Investments*, September 30, 2021. <https://www.pionline.com/private-equity/calpers-carlyle-lead-initiative-standardized-esg-reporting>.

Dimson, E., P. Marsh, and M. Staunton. 2020. "Divergent ESG Ratings." *The Journal of Portfolio Management* 47 (1): 75–87.

Eccles, R. G., and J. C. Strohle. 2018. "Exploring Social Origins in the Construction of ESG Measures." Working paper, August 1, 2018. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3212685.

Jacobs, B. I., and K. N. Levy. 1988. "Disentangling Equity Return Regularities: New Insights and Investment Opportunities." *Financial Analysts Journal* 44 (3): 18–43.

———. 2021. "Factor Modeling: The Benefits of Disentangling Cross-Sectionally for Explaining Stock Returns." *The Journal of Portfolio Management* 47 (6): 33–50.

Lindeman, A. J. 2022. "Esg Score Design and Portfolio Construction." *Journal of Impact and ESG Investing*, 2 (3). <https://jesg.pm-research.com/content/pmrjesg/early/2022/01/05/jesg.2022.1.037.full.pdf>.

Mackintosh, J. 2018. "Is Tesla or Exxon More Sustainable? It Depends Whom You Ask." *The Wall Street Journal*, September 17, 2018. <https://www.wsj.com/articles/is-tesla-or-exxon-more-sustainable-it-depends-whom-you-ask-1537199931>.

———. 2022. "Credit Suisse Highlights the Flaws of Trying to Quantify ESG Risks." *The Wall Street Journal*, January 18, 2022. <https://www.wsj.com/articles/credit-suisse-shows-flaws-of-trying-to-quantify-esg-risks-11642435816>.

Matos, P. 2020. *ESG and Responsible Institutional Investing around the World: A Critical Review*. CFA Institute Research Foundation. <https://www.cfainstitute.org/-/media/documents/book/rf-lit-review/2020/rflr-esg-and-responsible-institutional-investing.pdf>.

Pastor, L., R. F. Stambaugh, and L. A. Taylor. 2021a. "Sustainable Investing in Equilibrium." *Journal of Financial Economics* 142 (2): 550–571.

———. "Dissecting Green Returns." 2021b. Jacobs Levy Equity Management Center for Quantitative Financial Research Paper, September 7, 2021b. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3864502.

Pedersen, L. H., S. Fitzgibbons, and L. Pomorski. 2021. "Responsible Investing: The ESG-Efficient Frontier." *Journal of Financial Economics* 142 (2): 572–597.

Prall, K. 2021. "ESG Ratings: Navigating Through the Haze." CFA Institute. August 10, 2021. <https://blogs.cfainstitute.org/investor/2021/08/10/esg-ratings-navigating-through-the-haze/>.

Sustainability Accounting Standards Board. 2018. "SASB Codifies First-Ever Industry-Specific Sustainability Accounting Standards." Press release, November 7, 2018. <https://www.globenewswire.com/news-release/2018/11/07/1646736/0/en/SASB-Codifies-First-Ever-Industry-Specific-Sustainability-Accounting-Standards.html>.

U.S. Department of Labor, Prudence and Loyalty in Selecting Plan Investments and Exercising Shareholder Rights, 86 Fed. Reg. 57,272 (proposed October 14, 2021, to be codified at 29 C.F.R. Pt. 2550).

Wigglesworth, R. 2018. "Rating Agencies Using Green Criteria Suffer From 'Inherent Biases.'" *Financial Times*, July 20, 2018. <https://www.ft.com/content/a5e02050-8ac6-11e8-bf9e-8771d5404543>.