# How Company Size Bias in ESG Scores Impacts the Small Cap Investor

Osman T. Akgun, Thomas J. Mudge III, and Blaine Townsend

Bailard

# Bailard

ailard is an independent wealth and investment management firm in the San Francisco Bay Area. For individuals and institutions alike, Bailard proudly serves as a trusted partner focused on achieving long-term results aligned with client values and goals. The pursuit of excellence, in all its forms, drives everything we do. We believe in six core values—accountability, compassion, courage, excellence, fairness, and independence—and these values manifest in our strategies and client service. A tenured and independent firm founded in 1969, we offer wealth and institutional clients stability and candor. We stand committed to serving our clients with research-based insight, transparency, and compassion. *You are not a robot. Neither are we.*™ *

### Osman Akgun, PhD, CFA
Vice President, Domestic Equities | **Bailard, Inc.**

Osman serves as Vice President of Domestic Equities and as a portfolio manager for the Small Cap Value ESG Equity Strategy. His primary focus is quantitative research, model building, and portfolio optimization. Osman joined Bailard in 2012 after receiving his Ph.D. in Industrial Engineering and Operations Research at the University of California, Berkeley. Prior to his Ph.D. program, Osman received Masters' Degrees in Statistics as well as in Industrial Engineering and Operations Research, both from the University of California, Berkeley. During this time, Osman worked as a Research Assistant focusing on stochastic modeling and optimization, with applications to finance, service operations and computer communications. He was also a Teaching Assistant for Quantitative Methods for Business at the Presidio Graduate School. Prior to his graduate work, he received a B.S. in Industrial Engineering from Bogazici University, Turkey. Osman received his Chartered Financial Analyst® designation in 2017 and is a member of the CFA Institute and the CFA Society of San Francisco.

### Thomas J. Mudge III, CFA
Senior Vice President, Director, Equity Research | **Bailard, Inc.**

Tom heads Bailard's equity research and serves as a portfolio manager of the Small Cap Value ESG Equity Strategy. Tom earned his BA at Northern Michigan University in 1985 and the Chartered Financial Analyst® designation in 1994, and is a member of the CFA Institute and the CFA Society of San Francisco. He has also completed the certificate program in Investment Decisions and Behavioral Finance at Harvard University's John F. Kennedy School of Government.

### Blaine Townsend, CIMC®, CIMA®
Executive Vice President, Director, Sustainable, Responsible and Impact Investing | **Bailard, Inc.**

Blaine serves as an Executive Vice President and the Director of Bailard Wealth Management's Sustainable, Responsible and Impact Investing (SRII) group, as well as a portfolio manager of Bailard's Small Cap Value ESG Equity Strategy. Blaine is on both Bailard's fundamental and SRII investment committees, conducts social research, oversees corporate engagement efforts, and maintains client relationships.

Blaine began researching and writing about corporate social responsibility in the late 1980s. He started his career in Socially Responsible Investing in 1991 at the Muir Investment Trust, the nation's first environmentally screened bond fund. In 1996, he opened the California office for Trillium Asset Management, which he led for thirteen years. While at Trillium, Blaine managed socially responsible and sustainably focused portfolios, served on the firm's investment committee, and worked on corporate engagement efforts on a host of social and environmental issues from deforestation to media reform. Blaine also led the effort to create the "Joan Bavaria Awards for Building Sustainability in the Capital Markets", which are presented each year at the Ceres annual conference and was part of the working group that created OpenMic to address net neutrality and media reform. In 2009, he joined Nelson Capital Management, where he was a partner and a senior portfolio manager. He also served on the firm's leadership team and investment committee. Blaine chaired the company's corporate engagement committee and was on the Extraction-Free and Animal-Welfare model teams. Blaine joined Bailard in 2016.

Blaine currently serves as an Advisory Board member for *The Journal of Impact and ESG Investing*. He holds a BA from the University of California, Berkeley and CIMC® and CIMA® credentials. His writings on social investing have appeared in numerous publications including the *San Francisco Chronicle*, *Houston Chronicle*, *San Jose Mercury News,* and *London's Environmental Finance* magazine.

# How Company Size Bias in ESG Scores Impacts the Small Cap Investor

## Osman T. Akgun, Thomas J. Mudge III, and Blaine Townsend

**Osman T. Akgun**
is a vice president at Bailard, Inc., in Foster City, CA.
oakgun@bailard.com

**Thomas J. Mudge III**
is director of equity research and a portfolio manager at Bailard, Inc., in Foster City, CA.
tmudge@bailard.com

**Blaine Townsend**
is an executive vice president and director of the Sustainable, Responsible and Impact Investing (SRII) group at Bailard, Inc., in San Francisco, CA.
btownsend@bailard.com

### KEY FINDINGS

■ ESG size bias is not as significant in US small cap stocks as it is in US large cap stocks.

■ Stocks with high (low) ESG scores outperform (underperform) the market in the US small cap space.

■ It is possible for US small cap portfolio managers to use ESG as an alpha generating tool without taking size risk.

### ABSTRACT

The rising popularity of socially responsible investing (SRI) has increased interest in the relationship between traditional measures of corporate financial performance (CFP) and the emerging field of corporate social performance (CSP). SRI investors have tended to have a large capitalization (cap) stock focus that has served them well in the past, but that may be suboptimal in the future if we return to a period of small cap outperformance. Using environmental, social, and governance (ESG) scores as a proxy for CSP, our research supports past studies showing a large cap bias in measures of CSP. Narrowing our focus, we find that within the US small cap stock universe, the correlation between firm size and ESG scores is greatly reduced. We construct small cap ESG leaders and laggards portfolios and test their performance. We demonstrate that performance was not affected by neutralizing these portfolios with respect to firm size. Our results reinforce the idea that CSP proxies such as ESG scores have the potential to practically enhance portfolio performance in US small cap stocks.

### TOPICS

*ESG investing, analysis of individual factors/risk premia, portfolio construction, performance measurement**

When socially responsible investing (SRI) was introduced 50 years ago, it was greeted with skepticism from traditional Wall Street analysts and academics. The concept of limiting an investable universe based on values-based themes was at odds with traditional modern portfolio theory, and SRI did not yet have an established long-term track record.

Over time, as investor interest in SRI grew and its emphasis broadened from a largely ethical focus to include potential risk control and return benefits, SRI began to attract considerable attention from academics and practitioners. As a result, an

extensive body of literature has emerged studying the relationship between traditional measures of corporate financial performance (CFP) and the emerging field of corporate social performance (CSP).

There are a variety of ways to measure a firm's CFP. Accounting-based measures, such as Return on Assets or Net Income Growth, and market-value based measures, such as stock returns or book to price ratios, are some popular examples of CFP measures, but they all share one important feature: they can be measured objectively.

In contrast, CSP measures such as firm reputation or employee working conditions are not directly observable and require some subjective judgment in order to be measured.

This objective versus subjective measurement issue exists in many venues, including Olympic sports. The fastest runner, the longest jumper, or the team scoring the most goals can all be determined objectively—but determining the best gymnastics routine or figure skating performance requires the subjective opinions of a panel of judges.

One byproduct of a system of subjective measurement is the infrastructure required to implement it. To subjectively assess something in a responsible and consistent manner, you need detailed criteria for making decisions and judges with (preferably) extensive experience in making these types of judgments. Creating this infrastructure requires time and resources, and in the field of finance, may rely on criteria not easily disclosed by companies or not disclosed at all.

Another aspect of judging subjective measures is that bias can and often does creep in. Returning to the example of the Olympics, studies have shown that judges at the games show a perhaps unsurprising nationalistic bias in their scoring (Zitzewitz 2006; Emerson and Meredith 2011). The field of finance contains subjective biases as well, as a bad past experience with a company may result in a higher threshold for future investment consideration.

Initial interest in CSP came from socially responsible investors, often with differing priorities. As a result, early CSP measures tended to lack cohesion and consolidation. (See Orlitzky, Schmidt, and Rynes (2003) and Margolis, Elfenbein, and Walsh (2007) for the discussion of a broad list of these measures.)

Today, SRI investing has been largely eclipsed by environmental, social, and governance (ESG) investing, with an increased focus on corporate governance and an emerging taxonomy for CSP factors. The rapid growth trajectory of ESG investing since the early 2000s brought about demand for more uniform and coherent CSP constructs, and as a result, recent research has focused on using major ESG providers' data to study the CFP-CSP link. See Bouten et al. (2017) for a detailed list of recent research using different data providers.

While some progress has been made in standardizing measures of CSP, there remains substantial disagreement among the scores generated by ESG providers. Low ESG score correlation among different providers has been noted several times in the literature (Gibson, Krueger, and Schmidt 2020; Berg, Koelbel, and Rigobon 2020).

Some of these ESG score disagreements come from the criteria selected for measurement, some come from how the criteria are measured, and some come from the emphasis or weight placed upon each criterion selected.

Because of the variability of ESG scores among providers, the choice of which ESG provider to use becomes a crucial decision when attempting to measure CSP.

The relative weight or importance given to individual CSP measures by ESG score providers is of less concern than the CSP criteria selected and how they are measured, as long as a provider's formula is largely transparent. Just as a value stock manager may emphasize certain CFP performance measures such as earnings to price ratios more than a growth stock manager would, one would expect a CSP manager

with a focus on environmental issues to have a different ESG score emphasis than a manager with a religious values focus.

What is of concern are systematic biases in CSP data that are unrelated to investor priorities. There are three major tilts that can be observed in ESG providers' company scores (Deixonne 2019):

1. Industry Bias: Mature and heavily regulated industries (e.g., banks, wireless telecommunications) in general have more favorable ESG scores. There are also business activity related biases affecting ESG scores. Some industries, such as renewable energy, are less exposed to ESG risks, and hence score more favorably. Conversely, tobacco and gaming have greater ESG risk exposure.

2. Country Bias: Different regulations and restrictions around the world lead to significant discrepancies among ESG scores of firms from different regions. For instance, companies in Europe, which was one of the early adopters of ESG regulations and restrictions, have higher ESG scores on average than their counterparts in the United States.

3. Size Bias: Larger companies tend to have higher ESG scores. There are several possible explanations for this. Larger companies tend to receive greater scrutiny from investors (Burke et al. 1986), investment analysts, and the media, and as a result may feel pressured to provide greater ESG reporting and disclosures. Larger companies also may have the resources available to address ESG issues that smaller companies lack (Orlitzky 2001).

In this article, we focus on Size Bias.

Almost all major stock market indexes are capitalization weighted. The larger the market capitalization of a particular company, the larger its weight in the index. Due to the relatively massive size of the largest 50 stocks in the large cap universe, significant coverage of the total market capitalization of the entire index can be achieved by focusing upon them. In contrast, in the realm of small caps, the size difference between the largest and the smallest stock is much less. The largest 50 stocks make up 44.0% of the Russell 1000 Index, while the largest 50 stocks in the Russell 2000 Index account for only 10.6% of its total market value (see Exhibit 2 for detailed statistics).

Given these realities, early SRI investors creating a subjective measurement infrastructure logically focused on the largest, most widely held companies in order to get the greatest initial impact for their efforts. SRI investors needed to address widely held stocks that made up a substantial portion of most investors' portfolios, and the greater availability of CSP data from large companies who had the scale to disclose information with more regularity and depth made assessing them that much easier.

History bears this out.

In the early days of SRI (1970–1995), research was focused almost exclusively on large cap companies, and there was no SRI research coming from Wall Street. The early SRI research had more in common with the qualitative corporate social responsibility (CSR) research that emerged in the 1970s than it had to traditional "buy, sell, hold" research generated by Wall Street brokerage houses. It was not until SRI got a boost from the South African anti-Apartheid divestment movement in the late 1970s and 1980s that research in North America became more focused on the investable universe with tools geared specifically for asset managers.

The first broad-based SRI index attempting to track the S&P 500 was created in 1990, by the Boston-based research firm Kinder, Lydenberg, Domini & Co. (KLD). The KLD400 index first excluded 250 companies involved in industries associated with values-based SRI investing (tobacco, guns, alcohol, etc.), then added back 50 with

positive characteristics and then another 100 to redress sector industry exposures. The KLD400 was capitalization weighted, like the S&P 500, and was large cap dominated.

In 1991, KLD launched a research database called Socrates and established itself as an early provider of third-party research. SRI research at that time was labor intensive, analog and slow. Just as today, analysts depended heavily on corporate disclosure, but there was no standardized reporting on social or environmental performance. It was common to see reports on corporate engagement efforts of the established SRI firms (the subscribers) reflected in the company profiles. There was a very tight feedback loop between users and analysts. Updates were provided to subscribers monthly via floppy disc.

During the seismic shift from traditional SRI to the more European ESG framework, climate change risk increased the global demand for ESG research. Large firms like Dow Jones, Goldman Sachs, and MSCI, among others, tried to meet the demand for this research because large investors were now engaged. Other sophisticated and specialized ESG research operations like Sustainalytics and RobecoSam emerged. Divergent ESG research methodology began to develop. Data gathering began happening on a global, information-age basis by very large operations. However, one thing remained consistent: the bias toward research on larger companies and better ESG scores for larger companies.

This does not mean that smaller companies are necessarily worse when it comes to treating their employees well, or on other CSP issues. In fact, a recent study showed that large and small companies in the same industry tend to have similar ESG risk exposures, and that many of the ESG score differences between large and small companies are a matter of disclosure, with large companies being able and/or willing to disclose more CSP information (McCourt and Vandegrift 2020).

Beyond the understandable bias toward large cap companies exhibited by the pragmatic needs of early SRI investors, the average response to SRI inquiries and requests (disclosure) from large companies has been more complete than that of small companies, creating a feedback loop that further confirmed favoritism toward large cap stocks. As a result of these biases, investors with a primary CSP focus will find their portfolios naturally gravitating toward larger cap stocks due to the higher average and more robustly documented ESG scores they offer.

Today, relative to the large cap universe, ESG scoring in the small cap universe is still trying to catch up. The eVestment investment consultant database contains nearly 80 fully dedicated large cap ESG strategies and only 10 dedicated to ESG small cap. The depth of information on companies within the small cap universe still pales in comparison to large cap. As a result, the effects of ESG factors on the small and micro cap universes are just now being studied.

In recent years, coincident with a skyrocketing interest in SRI investing, the ESG score bias toward large cap has proven to be beneficial from a relative performance standpoint. The Russell 1000 Index (large cap) has outperformed the Russell 2000 Index (small cap) over the past 5- and 10-year periods, but over longer historical periods, this relationship is reversed, with small cap stocks coming out on top (see Exhibit 1).

Should small cap stocks return to an extended period of relative outperformance, SRI interest in small companies is likely to follow. If this becomes the case, ESG scoring biases and coverage gaps regarding small cap companies will command new interest.

**EXHIBIT 1**
**Historical Returns**

|  | 5 Years | 10 Years | 20 Years |
|---|---|---|---|
| Russell 1000 Index | 15.60% | 14.01% | 7.75% |
| Russell 2000 Index | 13.26% | 11.20% | 8.74% |
| Difference | 2.34% | 2.81% | –0.99% |

**NOTE:** This exhibit shows the annualized total returns for Russell 1000 Index and Russell 2000 Index and the difference between the two.

**SOURCE:** Morningstar Direct.

The remainder of the article is organized as follows. In the Data section we describe the datasets used for the analysis and present summary statistics. In the Results section we show our results in depth. We examine the relationship between size bias and ESG scores as well as the CFP-CSP link in US small cap stocks. Finally, in the Conclusion section we discuss the key takeaways of this study.

## DATA

One of the major challenges with ESG data from providers, as discussed in the previous section, is the rapidly evolving nature of their data. Because ESG investing is a relatively young and developing field, new measures of CSP, increased disclosure, and growing reporting standards result in the continual addition of new datapoints and frequent methodology changes by providers. It is quite common to see data irregularities such as big jumps in coverage or significant changes in the serial correlations of scores. As a result, it is not possible to do a robust historical back test using data from most of these providers. For our analysis, we chose MSCI ESG data as our CSP proxy, as it has reasonably consistent scores with adequate US small cap coverage from January 31, 2015, to October 30, 2020.[1]

We use weighted average ESG score,[2] as MSCI uses a proprietary weighted combination of individual Environmental, Social, and Governance scores to create a composite score for each company in their coverage universe. We use individual stock returns[3] to measure CFP and market capitalization to measure firm size. Our research examines all stocks in the Russell 3000 Index, which encompasses the largest 3,000 stocks trading on major US exchanges and is rebalanced periodically. This universe is further divided into large cap stocks as defined by membership in the Russell 1000 Index, and small cap stocks as defined by membership in the Russell 2000 Index. Factor return data for CAPM, Fama-French, and Carhart models come from Kenneth French's website (French 2021). For this analysis, we use the bottom two market capitalization quintiles sorted against the remainder of the Fama-French common factors to represent the small cap universe. Details can be found in the appendix section. The frequency for all the data used is monthly.

Summary statistics for the data can be seen in Exhibit 2. For each month, we calculate the number of stocks in each index; the percentage of stocks with a valid MSCI weighted average ESG Score; market capitalization; MSCI weighted average ESG score; and E, S, and G component scores averaged over all the stocks in each index. We then average these cross-sectional statistics over 70 months from January 31, 2015, to October 30, 2020. As can be seen in Exhibit 2, MSCI has reasonably good ESG score coverage in small cap stocks (ranging from 58.6% to 77.7%) over the period we examined. MSCI ESG scores were higher on average in the Russell 1000 Index than in the Russell 2000 Index, providing supporting evidence of the size-bias discussed in the previous section. While the raw average score difference between the two indexes may appear small, due to their narrow standard deviations, the difference is statistically quite significant.[4] While the size bias in ESG scoring between

---

[1] To measure consistency we use pair-wise month to month rank correlation of cross-sectional MSCI ESG scores. Average correlation for the period from January 31, 2015, to October 30, 2020, is 0.98 with a standard deviation of 0.02. However, correlation between December 2014 and January 2015 scores is 0.92. Similarly, the correlation between October 2020 and November 2020 scores is 0.83, indicating major methodology changes around these periods.

[2] We use February 2020 scores for March-July 2020 since we do not have access to weighted scores for that period.

[3] Return data comes from S&P Capital IQ.

[4] The value of the *t*-statistic for the paired sample test is 29.31.

**EXHIBIT 2**
Summary Statistics for January 2015–October 2020

| | Russell 1000 Index | | | | Russell 2000 Index | | | |
|---|---|---|---|---|---|---|---|---|
| | **Mean** | **Std. Dev.** | **Min** | **Max** | **Mean** | **Std. Dev.** | **Min** | **Max** |
| **Num. Stocks** | 996.9 | 20.6 | 968.0 | 1039.0 | 1980.7 | 22.0 | 1923.0 | 2033.0 |
| **Coverage %** | 89.6% | 0.7% | 87.8% | 91.8% | 68.9% | 3.8% | 58.6% | 77.7% |
| **Market Cap. (mil)** | 26275.2 | 3673.9 | 20213.3 | 34883.3 | 1119.3 | 99.3 | 851.9 | 1326.6 |
| **ESG Score** | 4.6 | 0.2 | 4.4 | 4.9 | 4.4 | 0.1 | 4.2 | 4.6 |
| **E Score** | 5.2 | 0.1 | 4.9 | 5.3 | 4.3 | 0.1 | 4.1 | 4.4 |
| **S Score** | 4.2 | 0.1 | 4.1 | 4.5 | 4.2 | 0.1 | 4.0 | 4.3 |
| **G Score** | 5.2 | 0.4 | 4.6 | 5.7 | 5.2 | 0.2 | 4.9 | 5.5 |
| **Top 50 Avg. Cap. (median multiple)** | 25.2 | 2.9 | 22.2 | 33.4 | 6.4 | 1.0 | 5.1 | 9.5 |
| **Top 50 Avg. Cap. (% of total)** | 44.0% | 2.3% | 40.6% | 50.6% | 10.6% | 0.9% | 9.0% | 13.4% |

NOTES: This exhibit shows the monthly average statistics of items in the first column for the Russell 1000 Index and Russell 2000 Indexes. Monthly Market Cap., ESG Score, E Score, S Score, and G Score show the cross-sectional average in the corresponding index. Coverage % shows the ratio of stocks with a valid MSCI score to the number of stocks in the corresponding index. The last two rows show the average market capitalization of the largest 50 stocks as a multiple of median market capitalization and as a percentage of total market capitalization in the corresponding index.

large cap and small cap stocks is of great interest, we also wanted to determine if this bias persists strictly within the small cap Russell 2000 Index, and if and how it affects the CFP-CSP relationship.

## RESULTS

We first study the relationship between MSCI ESG scores and market capitalization among both large cap stocks and small cap stocks. Exhibit 3 shows the pairwise cross-sectional correlations averaged over 70 months from January 31, 2015, to October 30, 2020. Market capitalization is positively correlated with the MSCI ESG score in both the Russell 1000 Index and the Russell 2000 Index. This result is in line with the ESG size-bias that has been observed many times in previous studies. However, this positive relationship seems to be driven mainly by the E component in large caps, whereas the G component has the largest correlation in small caps. Moreover, the size-bias is significantly less within the universe of small cap stocks.[5] Interestingly, we observed that the cross-sectional correlation in small cap stocks has increased recently, approaching the levels observed in large cap stocks. It is too early to tell whether this near-term change in trend is just statistical noise or if there is a fundamental shift corresponding with the growing popularity of ESG investment. It is possible that increasing pressure from ESG stakeholders has begun to alter the behavior of smaller companies.

### ESG and the Cross-Section of Stock Returns

This section focuses on the cross-sectional relation between CFP and CSP strictly within the universe of small cap stocks. First, we regress stock returns (CFP measure) onto MSCI ESG scores (CSP proxy) within the Russell 2000 Index, and test how firm size affects the regression results. Second, we attempt a certeris paribus analysis

---

[5] The value of the *t*-statistic for the paired sample test is 20.43.

## EXHIBIT 3

### Average Correlations

|  | Russell 1000 Index | | | | | Russell 2000 Index | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
|  | Market Cap | ESG Score | E Score | S Score | G Score | Market Cap | ESG Score | E Score | S Score | G Score |
| Market Cap | 1.00 |  |  |  |  | 1.00 |  |  |  |  |
| ESG Score | 0.15 | 1.00 |  |  |  | 0.08 | 1.00 |  |  |  |
| E Score | 0.26 | 0.42 | 1.00 |  |  | 0.01 | 0.38 | 1.00 |  |  |
| S Score | 0.04 | 0.63 | 0.03 | 1.00 |  | 0.02 | 0.59 | 0.05 | 1.00 |  |
| G Score | −0.05 | 0.27 | −0.16 | −0.06 | 1.00 | 0.08 | 0.24 | −0.22 | −0.21 | 1.00 |

NOTE: This exhibit shows the pairwise cross-sectional correlations averaged over the testing period for the Russell 1000 and the Russell 2000 Indexes.

## EXHIBIT 4

### Determinants of Future Return

|  | Dependent Variable: Next Month's Return | | | |
|---|---|---|---|---|
|  | Model 1 | Model 2 | Model 3 | Model 4 |
| ESG Score | 0.0026 | 0.0021 | 0.0023 | 0.022 |
|  | (3.01)*** | (2.69)*** | (3.35)*** | (1.52) |
| Size |  |  | −0.0006 | 0.0054 |
|  |  |  | (0.13) | (0.77) |
| ESG Score*Size |  |  |  | −0.0014 |
|  |  |  |  | (1.38) |
| Value |  | −0.0014 | −0.0006 | −0.0007 |
|  |  | (0.38) | (0.24) | (0.26) |
| Momentum |  | 0.0023 | 0.0033 | 0.0033 |
|  |  | (1.25) | (1.81)* | (1.8)* |
| Liquidity |  | −0.0001 | 0.0002 | 0.0003 |
|  |  | (0.06) | (0.07) | (0.08) |

NOTES: This exhibit shows the Fama-Macbeth regression results of four different models. The numbers shown in the first row for each variable are the average slope coefficients of monthly cross-sectional regressions within the Russell 2000 Index, and the numbers shown in parentheses for each variable are the t-statistics calculated by standard errors that are corrected using the Newey-West procedure with one lag. The dependent variable is the next month's return of the stock. ESG score is the MSCI weighted average ESG score. Size is the natural log of the market capitalization, in millions. ESG Score*Size is the interaction term between the two. Value is the natural log of the market value of equity over book value of equity. Momentum is the stock's return measured over the previous year excluding most recent month. Liquidity is the natural log of median daily trading dollar volume over previous six months in thousands. *, **, *** indicates statistical significance at the 10%, 5%, and 1% level, respectively.

by controlling for commonly accepted alpha generation factors including value, liquidity, and price momentum. Third, we introduce and control for firm size in addition to the other previously neutralized alpha generation factors. Finally, we examine the relationship between ESG scores and returns as they vary with firm size. Exhibit 4 shows the results for four different regression models. The first model captures the effect of the ESG score on future stock performance independent of firm size and other controlling factors. It appears that the ESG score is a significant linear predictor of future stock performance as a standalone factor. In the second model we add value, liquidity, and momentum, and the relationship between ESG score and future stock return stays significant at the 1% level. In the third model we add firm size, measured by market capitalization, and show that the significance of the ESG score as a return predictor is not affected by firm size. In the fourth model, we add the interaction term between firm size and ESG score to the mix and see that the significance of the slope coefficients changes considerably. This suggests that the slope of the relationship between the ESG score and future performance is dependent of firm size. This final result requires further analysis, which we address in the next section.

### Investable ESG Portfolios

Multiple linear regression is a commonly used statistical tool for analyzing causal relationships in finance. However, it might not be very practical from an investment professional's perspective. First of all, financial data tend to be very noisy, and it is unrealistic to expect the link between stock returns and ESG scores to be particularly strong given the influence that other factors like market capitalization have on the relationship. This is certainly the case with the relationship between the ESG score and firm size as shown in Exhibit 4. Secondly, there is often a disconnect between regression

**EXHIBIT 5**
Performance of ESG portfolios January 2015 to October 2020
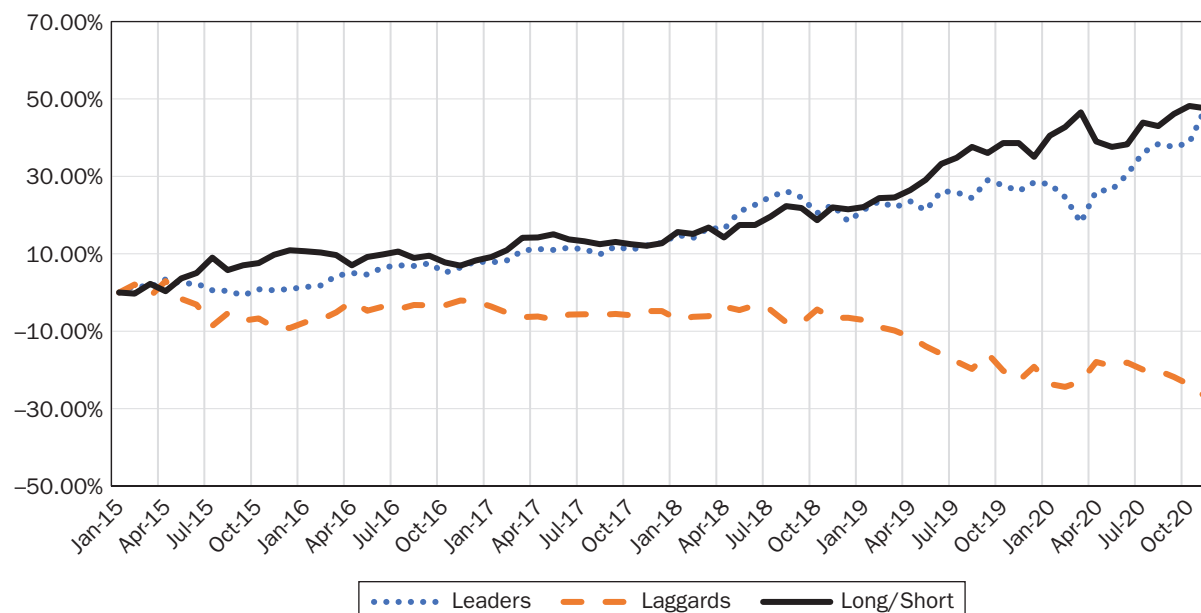


**NOTES:** This exhibit shows the cumulative performance of ESG portfolios. Leaders and laggards lines show the cumulative excess return over the Russell 2000 Index. Leaders (Laggards) portfolio is an equal weighted portfolio of stocks that rank in the top (bottom) 20% of according to MSCI ESG score. The Long/Short line shows the cumulative performance of a net neutral investment portfolio that buys the leaders and short sells the laggards portfolios. All portfolios were rebalanced every month.

results and implementable portfolios. Factor mimicking portfolios can be created using regression results (Fama 1976), but these portfolios require holding short as well as long stock positions, and necessitate investing in the entire universe, which in this case is approximately 2,000 small cap stocks.

In actual practice, portfolio managers tend to use ESG scores as a filtering tool, avoiding holding stocks with low scores in their portfolios. Therefore, we designed our tests from a practitioner's standpoint, looking at how a portfolio of stocks with low MSCI ESG scores (laggards) performs over time versus the rest of the market. We also tested the performance of a portfolio of high MSCI ESG scoring stocks (leaders) to see if the difference between leaders and laggards is in line with the strong linear relationship seen in Exhibit 4.

Exhibit 5 shows the cumulative excess performance of the bottom 20% (laggards) and the top 20% (leaders) of stocks selected according to their ranked MSCI ESG scores with monthly rebalancing. We also show the cumulative performance of a net neutral investment portfolio that buys high MSCI ESG scoring stocks and short sells low MSCI ESG scoring stocks. In our tests, the laggards portfolio underperformed the market by almost 30%, the leaders portfolio outperformed the market by almost 50%, and the net neutral long/short portfolio generated close to a 50% return (57 bps per month), suggesting a strong CFP-CSP link for these realistic portfolios. From this analysis, it appears that ESG scores may be used by portfolio managers as a risk mitigation tool as well as a potential source of alpha (outperformance).

We next check the robustness of the results presented in Exhibit 5 in order to determine if the relationship between ESG scores and stock returns that we observed can be influenced by biased exposures to other widely recognized fundamental factors used to explain cross-sectional asset returns. We regress the time-series returns of

## EXHIBIT 6
### ESG Long-Short Portfolio Returns

| | Dependent Variable: Long Short Return | | | |
|---|---|---|---|---|
| | Model 1: CAPM | Model 2: FF 3-Factor | Model 3: Carhart 4-Factor | Model 4: FF 5-Factor |
| Intercept | 0.0064 | 0.0044 | 0.0045 | 0.0046 |
| | (2.93)*** | (2.41)** | (2.77)*** | (2.47)** |
| Rm-Rf | –0.0593 | –0.039 | –0.0378 | –0.0415 |
| | (1.63) | (1.28) | (1.39) | (1.32) |
| SMB | | –0.4767 | –0.4174 | –0.4471 |
| | | (3.75)*** | (3.31)*** | (3.36)*** |
| HML | | –0.2701 | –0.1734 | –0.1729 |
| | | (5.35)*** | (3.76)*** | (2.22)** |
| UMD | | | 0.187 | |
| | | | (4.73)*** | |
| CMA | | | | –0.002 |
| | | | | –0.02 |
| RMW | | | | –0.071 |
| | | | | (0.98) |

NOTES: This exhibit examines the profitability of a trading strategy that buys the leaders portfolio and short sells the laggards portfolio. The leaders and laggards portfolios are equally weighted, and rebalanced every month. The resulting time-series returns on the long-short portfolio are regressed on widely recognized fundamental factors. Rm-Rf is the market return minus return on US Treasury bond. SMB is the return of a portfolio of small stocks minus the return of a portfolio of large stocks. HML is the return of a portfolio of stocks with high book-to-market ratio, minus the return of a portfolio of stocks with low book-to-market ratio. UMD is the return of a portfolio of stocks with high previous year return excluding the most recent month, minus the return of a portfolio of stocks with low previous year return excluding the most recent month. CMA is the return of a portfolio of stocks with low change in total assets in previous fiscal year, minus the return of a portfolio of stocks with high change in total assets in previous fiscal year. RMW is the return of a portfolio of stocks with high operating profitability minus the return of a portfolio of stocks with low operating profitability. *, **, *** indicates statistical significance at the 10%, 5%, and 1% level, respectively.

the net neutral long/short ESG score portfolio against the time-series returns of four different fundamental factor models: the CAPM model, the Fama French three-factor and five-factor models, and the Carhart four-factor model. If the profits of the ESG score long/short portfolio are explained by the commonly used factors of any of these models, then the estimated alpha (intercept term) of the respective regression model should be insignificant. Exhibit 6 shows the regression results for these models.

All four models indicate a positive ESG return premium even after controlling for market, size, book-to-market, momentum, profitability, and investment factors. The drop from the 57 bps per month baseline result of Exhibit 5 to the approximately 45 bps per month result in models 2–4 suggests that a portion of the observed ESG score return premium can be explained by some of these other fundamental factors. However, the residual ESG return premium is substantial, and remains significant at the 5% level for models 2 and 4 and at the 1% level for model 3.

In all the models, the coefficient on the size factor (SMB) is significantly negative. This aligns with the size bias we have discussed previously, that is, the net neutral long/short ESG score portfolio has a positive exposure to large stocks. Moreover, a significantly negative coefficient on the book-to-market factor (HML), and a significantly positive coefficient on the momentum factor (UMD), indicate that the long/short portfolio also has positive exposures to both growth and momentum stocks. In our analysis, neither profitability nor investment factors were significant in explaining ESG long/short returns.

Below, we further examine the firm size's effect on the CFP-CSP link in practice. Exhibit 7 shows the excess performance of size-neutral leaders/laggards portfolios over the Russell 2000 Index as well as the cumulative performance of a net neutral long/short ESG score portfolio that buys a size-neutral leaders portfolio and short sells a size-neutral laggards portfolio. Details of creating these portfolios can be found in the Appendix, but the processes can be thought of as representations of what portfolio managers do in practice. It's similar to tilting the weights of equally weighted top/bottom quintile portfolios by cutting down the position sizes of larger (smaller) market capitalization stocks and increasing the position sizes of smaller (larger) market capitalization stocks if the equally weighted portfolio is too large (small).

After size neutralization, the overall performance of the leaders portfolio did not substantially change, but the outperformance appears more volatile and less significant post 2018. On the other hand, size neutralization slightly heightened the underperformance of the laggards portfolio.

Finally, we perform another robustness check similar to the one described above. We neutralize the ESG score leaders and laggards portfolios to all the factors to which the long/short porfolio returns have significant exposure (namely size, value,

**EXHIBIT 7**

**Performance of Size Neutral ESG Portfolios January 2015 to October 2020**
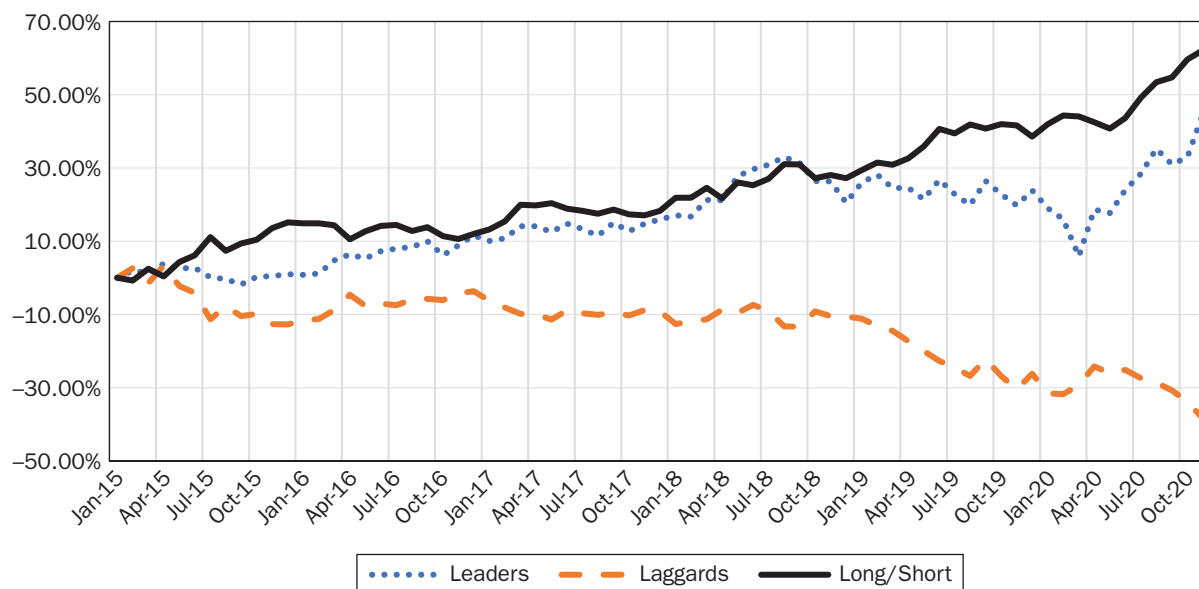


**NOTES:** This exhibit shows the cumulative performance of size neutral ESG portfolios. Leaders and laggards lines show the cumulative excess return over the Russell 2000 Index. Leaders (Laggards) portfolio is a portfolio of stocks that rank in the top (bottom) 20% according to the MSCI ESG score and weighted according to the optimization problem defined in the appendix. The Long/Short line shows the cumulative performance of a net neutral investment portfolio that buys the leaders and short sells the laggards portfolios. All portfolios were rebalanced every month.

and momentum) in Exhibit 6. Exhibit 8 shows the results. Similar to size neutralization, we tilt the weights of leaders and laggards portfolios via optimization to concurrently get zero exposures to market capitalization, price to book ratio, and previous year's return. Underperformance of the laggards portfolio is dampened in this case to closer to 20% over the test period. However the leaders portfolio still outperformed the index by over 50%, hence the net neutral long/short portfolio, even after controlling for common factor exposures, generated profits of 62 bps per month. This result is very similar to the baseline model's result in Exhibit 5 (57 bps per month). This analysis suggests that portfolio managers may be able to implement successful ESG score driven portfolio strategies across or within all portions of the market cap spectrum, including small cap, even when other common alpha factors are neutralized.

## CONCLUSION

For reasons of necessity and expedience detailed above, CSP investing pioneers and the ESG scoring methodology they helped to create have a large cap stock focus and bias that persists to this day. Coincidently, large cap stocks have outperformed small cap stocks in recent years, and SRI/ESG investors have benefited as a result.

Over a longer history, small cap stocks have tended to outperform large cap stocks, and it is reasonable to expect periods of small cap outperformance in the future.

When relative equity style performance rotates toward small cap stocks, SRI investors using measures of CSP will want unbiased ESG scores in order to fairly judge the attractiveness of these smaller companies. Unfortunately, unbiased scores

**EXHIBIT 8**

Performance of Factor Neutral ESG Portfolios January 2015 to October 2020
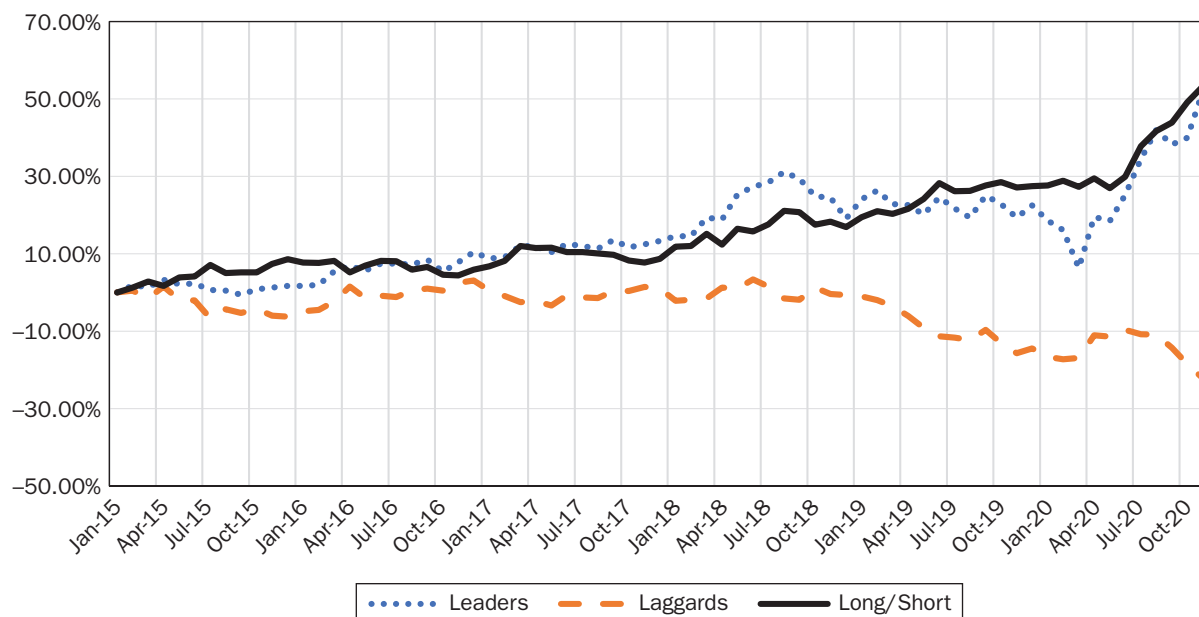


NOTES: This exhibit shows the cumulative performance of factor neutral ESG portfolios where size, momentum, and value are the factors used for neutralization. Leaders and laggards lines show the cumulative excess return over the Russell 2000 Index. The Leaders (Laggards) portfolio is a portfolio of stocks that rank in the top (bottom) 20% according to the MSCI ESG score and weighted according to the optimization problem defined in the appendix. Long/Short line shows the cumulative performance of a net neutral investment portfolio that buys the leaders and short sells the laggards portfolios. All portfolios were rebalanced every month.

do not yet exist, and due to ESG disclosure realities that favor larger companies, the gap may never be fully closed.

Fortunately, existing ESG measures when applied to small cap stocks appear able to successfully differentiate CSP risk and return characteristics between high scoring and low scoring smaller companies. This differentiation persists even after controlling for commonly recognized alpha factors such as value, liquidity, and momentum.

Because creating size-neutral portfolios of ESG leaders and laggards does not notably alter their return characteristics, we believe that implementing an CSP-influenced investment strategy using ESG scores in a small cap stock portfolio can be done in a straightforward manner by avoiding ESG laggards and focusing on ESG leaders.

## APPENDIX A

In order to calculate the factor returns for the small cap universe, we use 5 × 5 bivariate sorts on size vs. book to market, size vs. momentum, size vs. profitability, and size vs. investment. The average number of stocks in the bottom two size quintiles (ME1 and ME2) for our testing period is 2,323, while the larger three quintiles (ME3, ME4, and ME5) have 1,163 stocks on average. Note that this is very similar to the number of stocks in the Russell 2000 and 1000 Indexes respectively. Therefore we only extract the returns from the bottom two market capitalization quintiles to represent the Russell 2000 universe. For instance, we use ME1 LoBM (ME1 HiBM) data to represent the Small Growth (Value) portfolio and ME2 LoBM (ME2 HiBM) data to represent the Big Growth (Value) portfolio. Note that small and big refer to the market capitalization difference within the small

cap universe unlike the broader market versions used by Kenneth French's data library. We then calculate the HML factor as the average return on the two value portfolios minus the average return on the two growth portfolios,

$$HML = 1/2 \ (Small \ Value + Big \ Value) - 1/2 \ (Small \ Growth + Big \ Growth)$$

We calculate factor returns for up minus down factor (UMD), robust minus weak factor (RMW), and conservative minus aggressive factor (CMA) similarly. Finally for the small minus big factor (SMB) we calculate the following factors first,

$$SMB_{(B/M)} = 1/2 \ (Small \ Value + Small \ Growth) - 1/2 \ (Big \ Value + Big \ Growth)$$
$$SMB_{(MOM)} = 1/2 \ (Small \ Up + Small \ Down) - 1/2 \ (Big \ Up + Big \ Down)$$
$$SMB_{(OP)} = 1/2 \ (Small \ Robust + Small \ Weak) - 1/2 \ (Big \ Robust + Big \ Weak)$$
$$SMB_{(INV)} = 1/2 \ (Small \ Conservative + Small \ Aggressive)$$
$$- 1/2 \ (Big \ Conservative + Big \ Aggressive)$$

We then use $SMB_{(B/M)}$ in the Fama French three-factor model, the average of $SMB_{(B/M)}$ and $SMB_{(MOM)}$ in the Carhart four-factor model, and the average of $SMB_{(B/M)}$, $SMB_{(OP)}$, and $SMB_{(INV)}$ in the Fama French five-factor model as the small minus big factor.

## APPENDIX B

Consider the following cross-sectional Fama-Macbeth regressions

$$r_t = \alpha_t + X_t\beta_t + u_t \qquad t = 1, ..., T \tag{1}$$

where $t = 1, ..., T$ are the months in the testing period, $n_t$ is the number of stocks in the universe, $r_t$ and $u_t$ are the $n_t \times 1$ vectors of stock returns and error terms respectively. We include $K$ fundamental factors in the model in addition to MSCI ESG score, that is, $X_t = [m_t \ F_t]$ is the $n_t \times (K + 1)$ matrix of factor scores with $m_t$ being $n_t \times 1$ vectors of normalized MSCI ESG scores and $F_t$ being $n_t \times K$ matrix of normalized fundamental factor scores at time $t$. $\beta_t$ shows the $(K + 1) \times 1$ vector of factor premiums and $\alpha_t$ is the intercept term. We can rewrite this problem by adding the intercept term to the factor matrix, that is, $Z_t = [1_t \ m_t \ F_t]$ where $1_t$ is the $n_T \times 1$ vector of ones. The solution to this regression problem is given by ordinary least squares estimation, which yields $\hat{\beta}_t = (Z'_t Z_t^{-1})Z'_t r_t$. This solution can be thought of as the returns of portfolios described by $(Z'_t Z_t^{-1})Z_t$, which we will decompose as $[l_t \ w_t \ \Omega_t]$ where $l_t$ is the vector of weights of the equally weighted portfolio; $w_t$ is the vector of weights of the factor mimicking portfolio for the MSCI ESG factor; and $\Omega_t$ is the matrix of weights of the factor mimicking portfolios for the $K$ fundamental factors. Note that $w_t$ is also the solution to the following portfolio optimization problem;

$$\min \quad w_t w'_t \tag{2}$$

$$\text{s.t.} \quad w'_t m_t = 1 \tag{3}$$

$$w'_t F_t = 0_K \tag{4}$$

where $w_t = [w_1, ..., w_{nt}]$ is the vector of portfolio weights and $0_K$ is the vector of $K$ zeroes. In order to achieve our goal of creating a factor neutral leaders portfolio, we will restrict the feasible region of the above problem to the leaders basket of stocks only, which is given by

$$\mathcal{L}_t = \{i \in 1, \ldots, n_t : m_{it} \geq Q(0.8)\} \tag{5}$$

where $Q$ denotes the quantile function of $m_{it}$. Since we are selecting from the high ESG score basket that already has high positive exposure to ESG scores, we can relax the constraint 3, and add non-negativity constraints to achieve a long only portfolio. Finally we need a constraint for portfolio weights to add up to 1, $w'_t \mathbf{1}_t = 1$. The resulting formulation is given below.

$$\min \quad \boldsymbol{w}_t \boldsymbol{w}'_t \tag{6}$$

$$\text{s.t.} \quad \boldsymbol{w}'_t \mathbf{1}_t = 1 \tag{7}$$

$$\boldsymbol{w}'_t \boldsymbol{F}_t = 0_\kappa \tag{8}$$

$$w_{it} \geq 0 \qquad it \in \mathcal{L}_t \tag{9}$$

$$w_{it} = 0 \qquad it \notin \mathcal{L}_t \tag{10}$$

The optimization problem for the laggards portfolio can be defined similarly by using

$$\mathcal{L}_t = \{i \in 1, \ldots, n_t : m_{it} \leq Q(0.2)\} \tag{11}$$

## REFERENCES

Berg, F., J. F. Koelbel, and R. Rigobon. 2020. "Aggregate Confusion: The Divergence of ESG Ratings." MIT Sloan School Working Paper 5822-19. https://doi.org/10.2139/ssrn.3438533.

Bouten, L., C. H. Cho, G. Michelon, and R. W. Roberts. 2017. "CSR Performance Proxies in Large-Sample Studies: 'Umbrella Advocates,' Construct Clarity, and the 'Validity Police.'" Working paper. https://doi.org/10.2139/ssrn.3107182.

Burke, L., J. M. Logsdon, W. Mitchell, M. Reiner, and D. Vogel. 1986. "Corporate Community Involvement in the San Francisco Bay Area." *California Management Review* 28: 122–141.

Deixonne, F. 2019. "SRI, Data and Bias: Fund Managers' Bermuda Triangle." Accessed January 4, 2021. https://www.securities-services.societegenerale.com/en/insights/views/news/sri-data-and-bias-fund-managerse-bermuda-triangle/.

Emerson, J. W., and S. Meredith. 2011. "Nationalistic Judging Bias in the 2000 Olympic Diving Competition." *Math Horizons* 18 (3): 8–11.

Fama, E. F. "Foundations of Finance." Chap. 9: 326–329. New York: Basic Books, Inc. 1976.

French, K. R. (Data Library; accessed March 25, 2021). https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Gibson, R., P. Krueger, and P. S. Schmidt. 2020. "ESG Rating Disagreement and Stock Returns." Finance Working Paper N° 651/2020. https://doi.org/10.2139/ssrn.3433728.

Margolis, J. D., H. A. Elfenbein, and J. P. Walsh. 2007. "Does It Pay to Be Good? A Meta-Analysis of the Relationship between Corporate Social and Financial Performance." Working Paper. https://doi.org/10.2139/ssrn.1866371.

McCourt, T. C., and L. Vandegrift. 2020. *Environment, Social, and Governance Investing 101*. Analyst Report. Raymond James.

Orlitzky, M. 2001. "Does Firm Size Confound the Relationship between Corporate Social Performance and Firm Financial Performance?" *Journal of Business Ethics* 33 (2): 167–180.

Orlitzky, M., F. L. Schmidt, and S. L. Rynes. 2003. "Corporate Social and Financial Performance: A Meta-Analysis." *Organization Studies* 24: 403–441.

Zitzewitz, E. 2006. "Nationalism in Winter Sports Judging and Its Lessons for Organizational Decision Making." *Journal of Business Ethics* 15: 67–99.

*To order reprints of this article, please contact David Rowe at d.rowe@pageantmedia.com or 646-891-2157.*