Responsible Investing: The ESG-Efficient Frontier

Lasse Heje Pedersen, Shaun Fitzgibbons, and Lukasz Pomorski*

This draft: November 2019

Abstract

We propose a theory in which each stock’s environmental, social, and governance (ESG) score plays two roles: 1) providing information about firm fundamentals and 2) affecting investor preferences. The solution to the investor’s portfolio problem is characterized by an ESG-efficient frontier, showing the highest attainable Sharpe ratio for each ESG level. The corresponding portfolios satisfy four-fund separation. Equilibrium asset prices are determined by an ESG-adjusted capital asset pricing model, showing when ESG increases or lowers the required return. Combining several large data sets, we compute the empirical ESG-efficient frontier and show the costs and benefits of responsible investing. Finally, we test our theory’s predictions using commercial ESG measures, governance, sin stocks, and carbon emissions.

Keywords: portfolio choice, ESG, socially responsible investing, impact investing, sustainable investing, CSR
JEL codes: D62, G11, G12, G14, G23, G34, G4, M14, Q01, Q5

* Fitzgibbons and Pomorski are at AQR Capital Management, Two Greenwich Plaza, Greenwich, CT 06830. Pedersen is at AQR Capital Management, Copenhagen Business School, NYU, and CEPR. We are grateful for helpful comments from Antti Ilmanen, Ronen Israel, and Dan Villalon as well as seminar participants at Copenhagen Business School. We are very thankful to Alice Zhao for her excellent research assistance. AQR Capital Management is a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR.

Electronic copy available at: https://ssrn.com/abstract=3466417
1. Introduction

Asset owners and portfolio managers overseeing trillions of dollars seek to incorporate environmental, social, and governance (ESG) considerations into their investment process. However, investors have little guidance in how to incorporate ESG in portfolio choice, and worse, opinions differ dramatically across academics and practitioners about whether ESG will help or hurt their performance. Some argue that ESG considerations must necessarily lower expected returns (e.g., Hong and Kacperczyk, 2009) while ESG proponents believe that ESG investing raises returns (e.g., Edmans, 2011; Nagy, Kassam, and Lee, 2015; Morningstar, 2019).

To reconcile these opposing views, we develop a theory that shows both the potential costs and benefits of ESG-based investing. Our theory explains how the increasingly widespread adoption of ESG affects portfolio choice and equilibrium asset prices. Further, we estimate the magnitude of these effects empirically.

Our findings can be summarized as follows:

1) Theoretically, we show that an investor optimally chooses a portfolio on the ESG-efficient frontier.
2) The portfolios that span the frontier are all combinations of the risk-free asset, the tangency portfolio, the minimum-variance portfolio, and what we call the ESG-tangency portfolio (four-fund separation).
3) Equilibrium asset returns satisfy an ESG-adjusted capital asset pricing model, showing when higher ESG assets have, respectively, lower or higher equilibrium expected returns.
4) We estimate the costs and benefits of responsible investing via the empirical ESG-efficient frontier based on a measure of governance (“G”) and show how ESG screens can have surprising effects.
5) We test the theory’s equilibrium predictions using four ESG proxies, providing a rationale for why certain ESG measures predict returns positively (governance) while others negatively (non-sin stocks, a measure of “S”) or close to zero (low carbon emissions, an example of “E,” and commercial ESG measures).

To provide some intuition for these results, let us explain our model. We consider three types of investors. Type-U ("ESG-unaware") investors are unaware of ESG scores and simply seek to maximize their unconditional mean–variance utility. Type-A ("ESG-aware") investors also have mean–variance preferences, but they use assets’ ESG scores to update their views on risk and expected return. Lastly, type-M ("ESG-motivated") investors use ESG information and also have preferences for high ESG scores. In other words, M investors seek a portfolio with an optimal tradeoff between a high

---

1 For example, Bloomberg reports on February 8, 2019 that Europe alone has “some $12 trillion committed to sustainable investing.” The Global Sustainable Investment Review 2018 reports over $30 trillion invested with explicit ESG goals as of the beginning of 2018. The 2017/18 annual report of the Principles for Responsible Investments (PRI), a proponent of ESG supported by the United Nations, reports that its signatories manage close to $90 trillion in assets.

2 Edmans (2011) writes that “certain socially responsible investing (SRI) screens may improve investment returns”; Nagy et al. (2015) find that portfolios that incorporate ESG as an investment signal “outperformed the MSCI World Index over the sample period while also increasing their ESG profile.” Morningstar (2019) reports that “41 of the 56 Morningstar’s ESG indexes outperformed their non-ESG equivalents (73%) since inception”; an article in the Financial Times of September 7, 2017 states that “The outperformance of ESG strategies in beyond doubt.”
expected return, low risk, and high average ESG score. While trading off three characteristics may seem challenging, we show that the investor’s problem can be reduced to a tradeoff between ESG and the risk-adjusted return.

Specifically, for each level of ESG, we compute the highest attainable Sharpe ratio (SR). We denote this connection between ESG scores and the highest SR by the “ESG-SR frontier” as seen in Figure 1. To understand why the ESG-SR frontier is hump shaped, consider first the tangency portfolio known from the standard mean–variance frontier: The tangency portfolio has the highest SR among all portfolios, so its ESG score and SR define the peak in the ESG-SR frontier. Further, the ESG-SR frontier is hump shaped because restricting portfolios to have any ESG score other than that of the tangency portfolio must yield a lower maximum SR.

Type-A investors choose the portfolio with the highest SR, that is, the “tangency portfolio using ESG information” in Figure 1. Type-M investors have a preference for higher ESG, so they choose portfolios to the right of the tangency portfolio, on the ESG-efficient frontier. Choosing portfolios below or to the left of the efficient frontier is suboptimal because, in this case, the investor can improve one or both of the ESG score and the SR, without reducing the other. Nevertheless, type-U investors may choose a portfolio below the frontier, because they compute the tangency portfolio while ignoring the security information contained in ESG scores (they condition on a coarser information set). Type-M investors with a small preference for ESG choose a portfolio just to the right of peak with nearly the maximum SR (higher than the SR achieved by type-U in the example depicted in Figure 1), while type-M investors with strong preferences for ESG choose portfolios on the far right of the ESG-efficient frontier (possibly with lower Sharpe ratios than U investors).

![Figure 1. ESG Efficient Frontier](https://ssrn.com/abstract=3466417)
We also derive the equilibrium security prices and returns. In particular, we show that expected returns are given by an ESG-adjusted CAPM, as seen in Figure 2. When there are many type-U investors and when high ESG predicts high future profits, we show that high-ESG stocks deliver high expected returns. This is because high-ESG stocks are profitable, yet their prices are not bid up by type-U investors, leading to high future returns. In contrast, when the economy has many type-A investors, then these investors bid up the prices of high ESG stocks to exactly reflect their expected profits, thus eliminating the connection between ESG and expected returns. Further, if the economy has many type-M investors, then high-ESG stocks actually deliver low expected returns, because ESG-motivated investors are willing to accept a lower return for a higher ESG portfolio.

![Figure 2. ESG-CAPM](https://ssrn.com/abstract=3466417)

To illustrate the theory empirically and investigate its testable implications, we turn to a range of empirical proxies for ESG: 1) accruals in financial statements (Sloan, 1996) as a proxy for how aggressive a company is in its accounting choices, reflecting the company’s governance; 2) ESG scores produced by MSCI, a leading provider of ESG ratings; 3) carbon intensity, as a proxy for how green a company is; and 4) a sin stock indicator, defined as in Hong and Kacperczyk (2009).

We use the first proxy, based on accruals, to empirically estimate the ESG-efficient frontier. This frontier shows a responsible investor’s opportunity set and helps quantify the costs and benefits of using ESG in investing. Starting with the benefit of ESG information, we find that the maximum SR that incorporates this ESG proxy is about 12% higher than

---

3 High-ESG firms are more profitable if such firms benefit from being less wasteful, having more motivated employees, being better governed, or because their customers are willing to pay a higher price for their products. See also the literature on corporate social responsibility, e.g., Baron (2009), Benabou and Tirole (2010), and Hart and Zingales (2017).
the maximum SR that ignores such information (corresponding the vertical difference between the two tangency portfolios in Figure 1). Turning to the cost of ESG preferences, doubling the average ESG score relative to the level that maximizes the SR leads to a reduction in SR of only 3%.

We also study a common way of incorporating ESG into a portfolio: restricting the investment universe by removing the assets with the weakest ESG scores. We document a seemingly counterintuitive result that investors who screen out assets with the worst ESG characteristics may build optimal portfolios that have lower aggregate ESG scores than portfolios of investors who do not impose ESG-type restrictions. This happens because unconstrained investors may short poor-ESG assets to hedge out risks or to finance larger positions in high-ESG assets. Not surprisingly, limiting the breadth of the investment universe detracts from financial outcomes as well: the ESG-SR frontier for investors who screen out poor-ESG stocks is strictly dominated by the unconstrained frontier.

Finally, we carry out a series of theory-motivated empirical tests that help explain how the four ESG proxies we consider correlate with returns. Using our measure of governance based on accruals, we find that governance positively predicts future profitability. We also observe some increase in investor demand for stocks of this type, although not to the point of making them more expensive compared with other stocks. In fact, stocks with attractive governance trade at relatively cheaper Tobin’s Q, and we document evidence that they subsequently exhibit significant positive abnormal returns. In contrast, our results for sin stocks indicate that sin- and not-sin stocks have similar future profitability, but the latter exhibit stronger investor demand, to the point of trading at expensive valuations. This helps explain the “sin premium” described by Hong and Kacperczyk (2009), although the statistical significance is low. The sin premium parallels the finding of Baker, Bergstresser, Serafeim, and Wurgler (2018) that “green municipal bonds are issued at a premium to otherwise similar ordinary bonds.” Finally, we find little evidence that our two remaining proxies, MSCI ESG scores and company carbon intensity, correlate with either fundamentals or subsequent returns; we also find only weak evidence that investors adjust their portfolios based on these variables in our sample.

We contribute to the literature both theoretically and empirically. A growing theoretical literature on ESG follows Merton (1987) and assumes that ESG-sensitive investors will refuse to hold certain assets. For example, Heinkel, Kraus, and Zechner (2001) and Luo and Balvers (2017) show that in equilibrium, such market segmentation leads to higher expected returns to non-green companies.

In addition to allowing such segmentation, we explicitly model many assets characterized by ESG scores in addition to the standard risk–return characteristics. Based on this general setting, we derive several interesting properties of the
solution to the portfolio problem with parallels to the classic Markowitz solution, including the novel result that the ESG-SR frontier characterizes the solution, under certain conditions. Further, we show when ESG should predict returns positively or negatively in equilibrium.

Empirically, our research bridges the gap between papers that argue that ESG hurts performance and those that arrive at the opposite conclusion. The former group, based on the segmentation theories cited previously, is supported by empirical literature showing that “sin stocks” (alcohol, tobacco, gaming) generate positive abnormal returns (Hong and Kacperczyk, 2009). In contrast, another strand of the literature shows that stocks with good governance (the “G” in ESG) generate positive abnormal returns (Sloan, 1996; Gompers, Ishii, and Metrick, 2003), and so do stocks with higher employee satisfaction (part of the “S” of ESG) (Edmans, 2011). Our model and empirical results help explain these opposing findings. We submit that if ESG is a positive predictor of future firm profits, then ESG is also a positive predictor of returns as long as the value of ESG is not fully priced in the market. Further, the model predicts that this relationship may be weakened with ESG becoming a neutral predictor when most investors see the value in ESG, and even flips sign, with ESG becoming a negative predictor of returns, when investors are willing to accept lower returns for more responsible stocks. So, according to our model, the results of Hong and Kacperczyk (2009) arise because their measure of sin stocks (belonging to the industries related to alcohol, tobacco, and gaming) is associated with low investor demand, while the ESG measures of Gompers, Ishii, and Metrick (2003) and Edmans (2011) are related to higher firm profits in a way that the market has not fully appreciated.5

Our paper is also linked to the economic theories of discrimination, both the theory of taste-based discrimination pioneered by Becker (1957) and that of statistical discrimination due to Phelps (1972). Indeed, ESG scores play a dual role in our model because ESG affects investor preferences both directly (a kind of taste-based discrimination) and indirectly because ESG scores are informative of risk and expected returns (a form of statistical discrimination). In equilibrium, the interplay between these two dimensions allows for a variety of potential outcomes. This flexibility is important, because the empirical literature (discussed previously) suggests that the link between ESG and returns is not trivial. Indeed, certain ESG measures predict returns positively while others predict negatively, which highlights the need for a theoretical framework that allows for a similar flexibility in outcomes, with testable predictions of when each applies.

Puget (2014) and Friedman and Heinle (2016) who consider a single risky asset to study issues related to corporate engagement of responsible investors.

5 Bebchuk, Cohen, and Wang (2013) find that the return predictability associated with the governance indicator of Gompers, Ishii, and Metrick (2003) has disappeared, conjecturing an explanation based on investor learning. We find that the governance metric of Sloan (1996) based on accruals has continued to predict returns post publication.
2. Portfolio Choice with ESG: The ESG-Efficient Frontier

2.1. Model: Markowitz Meets Sustainability Goals

We consider an investor’s problem of choosing a portfolio of \( n \) risky assets and a risk-free security. The risk-free return is \( r^f \) and the risky assets have excess returns collected in the vector of random variables denoted by \( r = (r^1, \ldots, r^n)' \). The assets have an ESG scores given by \( s = (s^1, \ldots, s^n)' \).

We consider three types of investors. Type-U investors are uninterested or unaware of ESG scores. They take expected excess returns to be \( E(r) \) with risk given by the variance–covariance matrix, \( \text{var}(r) \). Type-A (ESG-aware) investors use ESG scores to update their views on risk and expected return. Specifically, they use assets’ expected excess return, \( \mu = E(r|s) \), conditional on the ESG information \( s \), and the conditional variance–covariance matrix of excess returns \( \Sigma = \text{var}(r|s) \).\(^6\) Lastly, type-M (ESG-motivated) investors use ESG information and also have preferences for high ESG scores. The portfolio problem for U and A investors has the standard Markowitz solution, so we focus here on the solution for type-M investors. Section 3 discusses equilibrium asset prices with all these investors.

Investor M starts with a wealth of \( W_0^M \) and chooses a portfolio of risky assets, \( x = (x^1, \ldots, x^n)' \). Specifically, \( x^i \) is the fraction of capital invested in security \( i \), or said differently, the investor buys \( x^i W_0^M \) dollars’ worth of security \( i \). The investor’s utility depends on her future wealth and the ESG characteristics of the portfolio. Given her portfolio choice, the investor’s future wealth is

\[
W = W_0^M \left( 1 + r^f + x'r \right)
\]  

(1)

The investor seeks to maximize her utility \( U \) over final wealth \( W \) and average ESG score, \( \bar{s} = \frac{x's}{x't} \) given the following extended mean–variance framework:

\[
U = E(W|s) - \frac{\hat{\gamma}}{2} \text{Var}(W|s) + W_0^M f(\bar{s})
\]  

(2)

Here, \( \hat{\gamma} \) is the absolute risk-aversion parameter and \( f: \mathbb{R} \to \mathbb{R} \cup \{-\infty\} \) is the ESG-preference function.\(^7\) The ESG preference function depends on the average ESG score among the risky asset positions (i.e., \( \bar{s} \) is the weighted sum of ESG scores, scaled by the total position in risky assets, \( x't) \), meaning that the investor gets no “ESG utility” from investing in the risk-free asset. We consider more general ESG preference functions in Section 2.4. The overall utility can be written as follows:

\(^6\) There is an active debate on whether ESG has an effect on valuations and, even more so, whether it is relevant to future risks or returns. For example, Flammer (2015) and Kruger (2015) provide supportive evidence for valuations/returns, while Dunn, Fitzgibbons, and Pomorski (2018), Ilhan, Sautner, and Vilkov (2018), and Hoepner et al. (2019) show that ESG correlates with risks.

\(^7\) Economists generally hesitate to add arguments to the utility function since this flexibility means that you can justify almost any outcome, but here, we simply formalize the stated intentions of investors who control trillions of dollars as discussed in the introduction. We allow that the ESG preference function takes the value \( -\infty \) to capture screens as discussed further in Section 2.3.
where \( \gamma = \bar{\gamma} W_0 M \) is the relative risk aversion. Hence, by dropping constant terms, we see that the utility maximization problem is

\[
\max_{x \in \mathcal{X}} \left( x' \mu - \frac{\gamma}{2} x' \Sigma x + f\left( \frac{x'S}{x'T} \right) \right)
\]

(3)

where the set of feasible portfolios is \( \mathcal{X} = \{ x \in \mathbb{R}^n | x'1 > 0 \} \), that is, all long-biased portfolios (generalized sets of allowed portfolios are discussed in Section 2.3). We consider portfolios that invest at least as much long as short because it is difficult to define the overall ESG characteristic for a portfolio that is short overall, but in principle, the framework can be applied more generally.

2.2. Solution: ESG-SR Frontier

Recall that in standard mean–variance analysis, the investor optimally combines the tangency portfolio with the risk-free security. The tangency portfolio is the portfolio that maximizes the Sharpe ratio, namely the expected excess return divided by the standard deviation of excess returns. To generalize this idea, we consider the maximum SR for each level of ESG score. The maximum SR that can be achieved with an ESG score of \( \bar{s} \) is denoted by \( SR(\bar{s}) \):

\[
SR(\bar{s}) = \max_{x \in \mathcal{X}} \left( \frac{x' \mu}{\sqrt{x' \Sigma x}} \right) = \max_{x \in \mathcal{X}} \left( \frac{x' \mu}{\sqrt{x' \Sigma x}} \right) \quad \text{s.t.} \quad x'1 = 1 \quad \text{and} \quad x'S = \bar{s}
\]

(4)

We will shortly use this definition of the highest Sharpe for each ESG level, but first we rewrite the utility maximization problem (3) as follows:

\[
\max_{\bar{s}} \left[ \max_{\sigma} \left( \max_{x \in \mathcal{X}} \left( x' \mu - \frac{\gamma}{2} \sigma^2 + f(\bar{s}) \right) \right) \right] \quad \text{(5)}
\]

This expression means that we can think of the investor’s problem as first choosing the best portfolio given a level of risk \( \sigma \) and an ESG score \( \bar{s} \) and then maximizing over \( \sigma \) and \( \bar{s} \). The former problem is solved by choosing the portfolio with the highest SR for the given ESG score (a more detailed proof is given in Appendix B), which yields:

\[
\max_{\bar{s}} \left[ \max_{\sigma} \left( SR(\bar{s}) \sigma - \frac{\gamma}{2} \sigma^2 + f(\bar{s}) \right) \right] \quad \text{(6)}
\]

The optimal level of risk is given by \( \sigma = SR(\bar{s})/\gamma \). Inserting this risk level and simplifying the expression shows the following result.
Proposition 1 (ESG-SR Tradeoff). The investor should choose her average ESG score \( \bar{s} \) to maximize the following function of the squared Sharpe ratio and the ESG preference function \( f \):

\[
\max_{\bar{s}} \left[ \frac{(SR(\bar{s}))^2}{2\gamma} + f(\bar{s}) \right]
\]

We see that the investor trades off the benefits of good risk-adjusted returns (that is, a high Sharpe ratio) against the benefits of high ESG scores. Interestingly, if the investor has a higher risk aversion \( \gamma \), then she puts less weight on SR relative to ESG, everything else equal. This is because a higher risk aversion means that the investor takes less risk and, therefore, earns fewer benefits of a high SR. (We note, however, that this is only true “everything else equal”; that is, one could get a different conclusion if comparing people who differ both in their risk aversion and ESG preferences).

We next characterize how the maximum Sharpe ratio depends on the ESG score. We use the notation \( c_{ab} = a' \Sigma^{-1} b \in \mathbb{R} \) for any vectors \( a, b \in \mathbb{R}^n \).

Proposition 2 (ESG-SR Frontier). The maximum Sharpe ratio, \( SR(\bar{s}) \), that can be achieved with an ESG score of \( \bar{s} \) is

\[
SR(\bar{s}) = \sqrt{c_{\mu\mu} - \frac{(c_{s\mu} - c_{1\mu})^2}{c_{ss} - 2\bar{s}c_{s1} + \bar{s}^2 c_{11}}} \tag{8}
\]

The maximum Sharpe ratio across all portfolios is \( SR(s^*) = \sqrt{c_{\mu\mu}} \), which is attained with an ESG score of \( s^* = c_{s\mu}/c_{1\mu} \). Increasing the ESG score locally around \( s^* \) leads to nearly the same Sharpe ratio, \( SR(s^* + \Delta) = SR(s^*) + o(\Delta) \), because the first-order effect is zero, \( \frac{dSR(s^*)}{ds} = 0 \).

We next consider the nature of the optimal portfolio weights for an ESG-aware investor.

Proposition 3 (Four-fund separation). Given an average ESG score \( \bar{s} \), the optimal portfolio is

\[
x = \frac{1}{\gamma} \Sigma^{-1} (\mu + \pi (s - 1\bar{s})) \tag{9}
\]

as long as \( x' 1 > 0 \), where

\[
\pi = \frac{c_{1\mu} - c_{s\mu}}{c_{ss} - 2\bar{s}c_{s1} + \bar{s}^2 c_{11}} \tag{10}
\]

The optimal portfolio is therefore a combination of the risk-free asset, the tangency portfolio, \( \Sigma^{-1} \mu \), the minimum-variance portfolio, \( \Sigma^{-1} 1 \), and the “ESG-tangency portfolio,” \( \Sigma^{-1}s \).

We see that the optimal portfolio looks the same as the standard Markowitz solution, except that the expected excess returns \( \mu \) have been adjusted. In other words, the optimal portfolio can be found as follows. First, the investor considers “ESG-adjusted” returns in the sense that each stock’s expected excess return is increased if its ESG score \( s_i \) is above the
desired average score \( \bar{s} \); otherwise it is lowered. The amount of adjustment depends on the scaling parameter \( \pi \), or the strength of the preference for ESG.\(^8\) Second, the optimal portfolio is found in the standard way. This further implies that all investors, regardless of their risk aversion and ESG preferences, should choose a combination of four portfolios (or “funds”): the risk-free asset, the standard tangency portfolio, the minimum variance portfolio, and the portfolio that we name the “ESG-tangency portfolio.” The ESG-tangency portfolio is the tangency portfolio if we replace the expected excess returns with the ESG scores.

### 2.3. Example: How Investors Choose Portfolios using the ESG-SR Frontier

Figure 3 illustrates how the ESG-motivated investor M chooses her portfolio using the ESG-Sharpe ratio frontier.\(^9\) First, for every ESG level, she finds the portfolio with the highest SR. One way to think about this step is that the investor computes a standard mean–variance frontier for all portfolios with this level of ESG. Then, the investor computes the maximum Sharpe ratio as the slope of the line that goes from the risk-free security to the tangency portfolio (again, based only on portfolios with this ESG level). The investor then collects all these Sharpe ratios and plots them against the ESG levels as seen in Figure 3.

Figure 3 also shows investor M’s indifference curves. These curves slope down because investor M likes high Sharpe ratios and also likes high ESG scores and can trade off one versus the other to remain indifferent about all portfolios on each indifference curve. Investor M’s utility is maximized at the point where her indifference curve is tangent to the ESG-SR frontier. This solution is not the global maximum of the Sharpe ratio, as the investor optimally chooses a higher level of ESG to satisfy her nonfinancial preference for ESG.

This solution contrasts with that of our ESG-aware investor A, depicted in Figure 4. Investor A also considers ESG information to build a better forecast of returns but does not have any direct (nonfinancial) preference for ESG. That is, he might tilt toward portfolios with high ESG (or, for that matter, with low ESG) only in as much as they help maximize the investment outcome. This means that the investor has horizontal indifference curves, illustrating that his preference depends only on the Sharpe ratio. We can also imagine that this investor considers the ESG-Sharpe ratio frontier but will always choose the portfolio with the highest possible Sharpe.

Finally, investor U solves a standard mean–variance optimization just like investor A, except that U computes potentially different estimates of risk and expected returns. We illustrate this more specifically when we estimate the empirical ESG-SR frontier in Section 4.2.

---

\(^8\) When \( \pi = 0 \), portfolio choice simplifies to the traditional mean–variance optimization.

\(^9\) Our illustration relies on a simple example with four uncorrelated assets, with respective means of 5%, 10%, 15%, and 20%, variances of 0.0625, 0.0625, 0.25, and 0.25, and ESG scores of 0.1, 0.5, 0.4, and 0.2.
2.4. Generalized ESG Preferences

Some investors use screens to help implement their ESG views. For example, an investor might screen out any stock with a low ESG score, for example, $s_i < 0.2$. The previous analysis naturally goes through for the subset of non-screened stocks. We can also incorporate such screens more directly by changing the set of allowed portfolios to be $X = \{x \in \mathbb{R}^n | x'1 > 0, \forall i x_i = 0 \text{ if } s_i < s^*\}$. Some investors prefer to exclude short positions, which can be captured by $X = \{x \in \mathbb{R}^n_+\}$, or to exclude both short positions and screened stocks $X = \{x \in \mathbb{R}^n_+ | \forall i x_i = 0 \text{ if } s_i < s^*\}$. Investors may achieve a better risk–return tradeoff if they allow shorting, and further, shorting low-ESG stocks could be consistent with ESG preferences.\(^{10}\) Hence, investors may require that their position in low-ESG stocks is zero or negative, that is, $X = \{x \in \mathbb{R}^n | x'1 > 0, \forall i x_i \leq 0 \text{ if } s_i < s^*\}$. For any of these restrictions, we can use the following result because all these portfolio sets are “cone shaped.” We say that $X$ is cone shaped if $x \in X$ implies that $ax \in X$ for all $a > 0$ (said differently, this means that $X$ only depends on $x/x'1$).

**Proposition 4 (ESG-SR Frontier with Screens).** The conclusion of Proposition 1 continues to hold for any cone-shaped $X$.\(^{11}\)

We can also consider even more general ESG utility functions of the form $e(x, s): X \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{−\infty\}$, where $X \subseteq \mathbb{R}^n$ is cone shaped set of allowed portfolios. We assume that the ESG utility function is homogeneous of degree zero with respect to portfolios, that is, $e(ax, s) = e(x, s)$ for any $a > 0$. This is a natural assumption because it means that the cash holding does not affect the ESG utility. For example, the portfolio $x = (0.2, 0.2)$ means that you put 20% of your assets in each risky asset and the rest, 60%, in cash, while the portfolio $2x = (0.4, 0.4)$ means that you put twice as much money in the same portfolio of risky assets, leaving only 20% in cash—and homogeneity means that you get the same ESG utility because the risky portfolio is the same. This homogeneity is what allows the investor to first focus on the optimal combination of the Sharpe ratio and portfolio-level ESG score, and then second decide on the amount of risk. In the absence of this homogeneity, the investor’s problem cannot be summarized as the ESG-SR frontier (which also shows that our frontier results are not trivial).

One interesting example is $e(x, s) = f \left( \frac{x's}{\sqrt{x'\Sigma x}} \right)$, where the investor cares about how much ESG she gets per unit of risk. This specification has the advantage that it also works for long–short portfolios with $x'1 = 0$ and it retains much of the tractability of the specification considered earlier.

The generalized ESG preference function can also capture screens by having $e(x, s) = -\infty$ for all portfolios where $x_i \neq 0$ for any security with $s_i < 0.2$. A screen can be seen as an extreme version of nonlinear preferences across the stocks’ ESG scores. In other words, an investor may not view a portfolio of three stocks with ESG scores of $(0.1, 0.8, 0.9)$

\(^{10}\) We note that, in the approach based on the average ESG score, the optimal portfolio may include short positions, and this approach gives the investor “credit” if the short positions have lower ESG scores than the long ones. Fitzgibbons, Palazzolo, and Pomorski (2018) argue that ESG-sensitive investors should indeed be willing to short low-ESG stocks.

\(^{11}\) Obviously, the definition (4) of the SR function must depend on the same set of allowed portfolios, $X$. 

11
the same as one with (0.6, 0.6, 0.6) even if they have the same average—because the former has one really low-ESG stock. Rather than capturing this idea with a screen, a less extreme (and still tractable) version would be $e(x, s) = e_1 \frac{x_{1S}}{x_{11}} - e_2 \frac{x' \text{diag} \left( \frac{1}{s_1}, \ldots, \frac{1}{s_n} \right)x}{(x'1)^2}$, where $e_1, e_2 \in \mathbb{R}$ are parameters. Here, the utility is more penalized if the investor has a stock with an ESG score close to zero. In any event, the investor can still think in terms of an ESG-SR frontier as seen from the next proposition.

**Proposition 5 (Generalized ESG-SR Frontier).** If the investor has generalized ESG preferences $e(x, s)$, then the investor’s problem is

$$\max_e \left[ \frac{(SR(e))^2}{2\gamma} + \tilde{\epsilon} \right]$$

where $SR(\tilde{\epsilon})$ is the maximum Sharpe ratio for a given level of ESG utility:

$$SR(\tilde{\epsilon}) = \max_{x \in X} s.t. \tilde{\epsilon} = e(x, s) \left( \frac{x'\mu}{\sqrt{x'\Sigma x}} \right)$$

Finally, we note that the theory also goes through if each security has a multi-dimensional ESG score (e.g., one score for environmental concerns, another for social, and a third for governance, with investors having preferences over such combinations).

Having characterized the solution to the ESG-aware portfolio problem in a variety of cases, we note that such a solution exists under certain conditions. Rather than going into such theoretical details, the empirical section will show the practical applicability of the framework.

### 3. Equilibrium Asset Pricing with ESG: ESG-Adjusted CAPM

Having solved the Markowitz problem with ESG investors, we next endogenously derive the security returns $r$ and security prices $p = (p_1, \ldots, p_n)'$. The exogenous variables are the ESG scores $s$, the risk-free rate $r_f$, the final payoffs of the assets $v = (v_1, \ldots, v_n)'$, and the shares outstanding of each stock, normalized to 1. We denote the total market payoff by $v^m = v_1 + \cdots + v_n$.

Recall that the economy has three types of investors. Type-U investors do not use ESG information at all. Specifically, they have no preference for ESG (i.e., their ESG preference function is $f_U \equiv 0$), and they ignore the informational value of ESG signals $s$, assuming that the best forecast of future payoffs is the unconditional mean $\hat{\mu} = E(v)$ and payoff risk is

---

12 A sufficient condition for existence is that the ESG preference function $f$ is continuous, we consider a compact space of ESG levels, $\tilde{s} \in [s_{\text{min}}, s_{\text{max}}]$, and for all such ESG levels, the portfolio $x$ in (9) satisfies $x'1 > 0$. In this case, for any $\tilde{s}$, an optimal portfolio is given in (9) with a resulting objective function (7) that is continuous in $\tilde{s}$, and any continuous function attains its maximum on a compact space.

Electronic copy available at: https://ssrn.com/abstract=3466417
taken to be $\Sigma = \text{var}(v)$. ESG-aware type-A investors also don’t enjoy ESG utility ($f_A \equiv 0$), but they exploit ESG to update their views on securities, using $\hat{\mu} = E(v|s)$ as the expected payoff and $\hat{\Sigma} = \text{var}(v|s)$ to capture payoff risk. Lastly, ESG-motivated type-M investors use ESG information and have a preference for a high average ESG score. Investors of type-U have wealth $W_0^U$ and similarly for types A and M, and the aggregate wealth is $W_0 = W_0^U + W_0^A + W_0^M$. Market clearing requires that the total demand for shares from all investors must equal the shares outstanding.

Let us now consider equilibrium implications of the model, starting with simplest cases in which all investors are of the same type. If all investors ignore ESG (i.e., all type-U) then we are back to a standard CAPM equilibrium. Indeed, all investors hold the tangency portfolio, that is, the portfolio that maximizes SR (relative to their information set), the tangency portfolio equals the market portfolio, and each security’s expected excess return is driven by its standard market beta, $\beta_i = \frac{\text{cov}(r_i, r^m)}{\text{var}(r^m)}$. What is new here is that a (small) investor who understands that ESG scores are informative can exploit this insight. To see this, we model the informational value of ESG scores as $E(v|s) = \hat{\mu} + \lambda (s - s^m)$, where $s^m = \sum_i w_i^m s_i$ is the weighted average ESG score of the market portfolio, $w_i^m = p_i / \sum_j p_j$ is the weight of the market portfolio in stock $i$, and the parameter $\lambda \in \mathbb{R}$ determines how informative ESG scores are for future profits. The following proposition characterizes the equilibrium.

**Proposition 6.** If all investors are uninterested in ESG, i.e., of type-U ($W_0^A = W_0^M = 0$), then any security $i$ has equilibrium price

$$p_i = \frac{\hat{\mu}_i - \frac{\lambda}{W_0} \sum_j w_j^{m} (s_j - s^m)}{1 + r_f}$$

(11)

Unconditional expected excess return obeys the standard unconditional CAPM:

$$E(r_i) = \beta_i E(r^m)$$

(12)

but conditional expected returns are given by

$$E(r_i|s) = \beta_i E(r^m) + \lambda \frac{s_i - s^m}{p_i}$$

(13)

This proposition provides several intuitive results. First, the price (11) of any firm’s equity is given by its expected cash-flow payoff ($\hat{\mu}_i$) less a risk premium ($\frac{\lambda}{W_0} \sum_j w_j^{m} (s_j - s^m)$), discounted by the risk-free rate. Second, expected excess returns (12) are driven by market betas from the perspective of an investor who ignores ESG scores. Third, from the perspective of an investor who uses ESG scores, we see from (13) that stocks returns have alphas relative to the CAPM that depend linearly on ESG. Specifically, if a high-ESG score is indicative of a high future profit, that is, if $\lambda > 0$, then we see that stocks with ESG scores above average have higher conditional expected returns than those with below-average ESG scores. This is in line with the empirical findings such as those of Gompers et al. (2003), who show that an ESG-type
metric (governance) earns CAPM alphas. Of course, market prices adjust when more investors are aware that this type of information may be relevant. At the extreme, all market participants may incorporate it into their decision—we consider this case next.

Suppose next that all investors use ESG signals, but without ESG preferences (i.e., all are ESG-aware of type-A). In this case, we get a conditional CAPM equilibrium and now investors can no longer profit from using the informational value of ESG scores because this information is already incorporated into prices. This theoretical prediction is in line with the empirical finding of Bebchuk et al. (2013), who argue that market participants have gradually learned about the usefulness of governance and have impounded it into prices. Consequently, they show that the measures from Gompers et al. (2003) do not predict abnormal returns out-of-sample.

Finally, suppose that all investors use ESG in their signals and in their identical ESG preferences (i.e., all type-M). Such ESG preferences change the equilibrium in an interesting way. To derive this equilibrium, we first note that prices and excess returns are related by

\[ r = \text{diag} \left( \frac{1}{p_i} \right) v - 1 - r^f \]

where \( \text{diag} \left( \frac{1}{p_i} \right) \) means the diagonal matrix with elements \( \left( \frac{1}{p_i}, \ldots, \frac{1}{p_n} \right) \). Clearly any investor wants to maximize the SR for the chosen ESG score. Further, in equilibrium, all investors must choose the market portfolio, which must therefore maximize for SR among all portfolios with an ESG equal to that of the market, \( s^m \). Based on Proposition 3, we see that any investor buys the following portfolio:

\[ x = \frac{1}{\gamma} \text{diag}(p_i)^{\Sigma^{-1}} \text{diag}(p_i) \left( \text{diag} \left( \frac{1}{p_i} \right) \bar{\mu} - 1 - r^f + \pi(s - 1s^m) \right) \]

The total wealth invested in each stock is \( W_0x \), where \( W_0 \) is the aggregate wealth, and the total dollar supply is \( p \) because shares outstanding are normalized to 1. Hence, the equilibrium condition is \( p = W_0x \), and we derive the equilibrium in Appendix B. While all investors hold the market portfolio in this equilibrium with only type-M investors (everyone cannot be more ESG friendly than the average), but nevertheless a security’s required return is affected by its ESG as well as its conditional market beta, \( \beta_i = \frac{\text{cov}(r_i, r^m|s)}{\text{var}(r^m|s)} \), as seen in the following proposition.

**Proposition 7 (ESG-CAPM).** If all investors are ESG-motivated of type-M \( W_0^U = W_0^A = 0 \), then any security \( i \) has equilibrium price

\[ p_i = \frac{\bar{\mu}_i + \lambda(s_i - s^m) - \frac{\gamma}{W} \text{cov}(r_i, r^m|s)}{1 + r^f - \pi(s_i - s^m)} \]

(14)

\[^{13}\] The model is also consistent with \( \lambda < 0 \), when ESG is in conflict with financial outcomes (e.g., when corporations engage in charity).
where $s^m$ is the ESG score of the market portfolio and the corresponding $\pi$ is given by (10). The equilibrium conditional expected excess return is given by

$$E(r_i|s) = \beta_i E(r^m|s) - \pi (s_i - s^m)$$

(15)

If all investors are ESG-aware of type-A ($W_0^U = W_0^M = 0$), the same conclusions hold with $\pi = 0$.

This proposition shows that equilibrium asset prices are rather different when all investors derive utility from ESG (type-M) relative to an economy dominated by investors who ignore ESG (as in Proposition 6). With such ESG-motivated investors, the price of any firm’s equity depends on its ESG score in two ways: First, the ESG score affects the expected cash flow as seen in the numerator of (14). Second, a higher ESG score lowers the discount rate used in the denominator, thus increasing the price. Turning to the implications for returns in (15), we see that the firm’s cost of capital is given by the standard conditional CAPM expression ($\beta_i E(r^m|s)$) adjusted for whether the ESG score is above or below that of the market. In other words, the firm’s cost of capital is lower if its ESG score is higher or, equivalently, the firm can issue shares at higher prices. This low cost of capital encourages high-ESG firms to make real investments because, using this low discount rate, more projects will have a positive net present value. While we do not explicitly model firm decisions to invest in ESG, this insight helps us understand why firms may choose to increase their corporate investment in ESG or why firms with a stronger ESG profile may be able to realize higher growth than firms with relatively weaker ESG.\(^{14}\)

If all types of investors exist, then several things can happen. If a security has a higher ESG score, then, everything else equal, its expected return can be higher or lower. First, a higher ESG score increases the demand for the stock from type-M investors, leading to a higher price and, therefore, a lower required return, as seen in Proposition 7. In contrast, companies with poor ESG scores that are down-weighted by type-M investors will have lower prices and higher cost of capital.

Second, the force that can increase the expected return is that the higher ESG may be a favorable signal of firm fundamentals, and if many type-U investors ignore this, this fundamental signal may not be fully reflected in the price, as seen in Proposition 6. Of course, it is an empirical question whether favorable ESG characteristics signal good profitability (e.g., good governance leading to a well-run company or a social company with happy productive employees) or low profitability (e.g., a company spending shareholders’ money on charities that employees and customers may not appreciate); that is, the sign of $\lambda$ is an empirical question. Further, it is an empirical question whether the force of Proposition 6 or 7 is stronger, that is, the extent to which ESG information is incorporated into prices and the extent to which ESG-investors’ demand pressure affects required returns.

\(^{14}\) See for example Albuquerque, Koskinen, and Zhang (2018) for a literature review relevant for this modeling approach.
Finally, we can consider the effect of an increasing adoption of ESG investing over time (i.e., an increasing fraction of ESG-motivated investors or a stronger ESG preference among these). A future increase in ESG investing will lead to higher prices for high-ESG stocks,[15] corresponding to a larger $\lambda$ in the model. If these flows are unexpected (or not fully captured in the price for other reasons), then the return of high-ESG stocks increases (as seen by increasing $\lambda$ in Proposition 6). If these flows are expected, however, then expected returns are not affected (as seen in Proposition 7).

3.1. Testable Predictions of the Theory

To summarize, the theory makes the following predictions:

I. The tradeoff between risk, expected returns, and ESG can be summarized by the ESG-SR frontier.

II. Using ESG information can increase the investor’s SR by improving the ESG-SR frontier.

III. Given the investor’s information set, investors with stronger ESG preferences (or higher risk aversion) should choose portfolios with higher ESG scores and (marginally) lower SR.

IV. Even investors with preferences for the average ESG score should optimally choose portfolios with positions (long or short) in almost any security (as opposed to standard models of taste-based discrimination that imply stricter segregation).

V. ESG investors should choose a combination of four portfolios (or “funds”), the risk-free asset, the standard tangency portfolio, the minimum-variance portfolio, and the ESG-tangency portfolio.

VI. A security with a higher ESG score should have
   i. a higher demand from ESG investors, which lowers the expected return;
   ii. different expected future profits, which can increase the expected return if the market underreacts to this predictability of fundamentals;
   iii. potentially stronger flows from investors, which can increase the price in the short term.

4. Empirical Results

We now estimate the ESG-efficient frontier to analyze the realized costs and benefits of ESG investing and test some of our theory’s predictions using a range of variables capturing various ESG characteristics.

4.1. ESG Measures and Data

ESG is a broad umbrella term, and consequently, we chose four different proxies, each motivated differently and possibly followed by different investor clienteles. Our goal is not a horse race between them, but rather a discussion of

[15] Here, the future payoff $v$ is exogenous, but it can be endogenized as the future equilibrium price in a multiperiod model. In such a model, an increase in ESG investing next time period corresponds to an increase of $\pi$ in equation (14), which leads to higher prices for high-ESG stocks.
how different elements of ESG may be priced in the market, and an illustration of how our theory guides empirical tests for investors who want to incorporate some ESG metric into their portfolios. Our four proxies for ESG are:

1) **Accruals (negated).** We begin with a measure of governance (the “G” pillar of ESG) that can be computed over a long sample period based on accounting information. Specifically, we look at each firm’s accruals over assets with a sample period spanning January 1963 through March 2019. We negate accruals so that higher values indicate better ESG. The idea, coming from the accounting literature, is that low accruals indicate that a firm is conservative in its accounting of profits (e.g., Sloan, 1996) and better governed companies tend to adopt more conservative accounting processes (e.g., Kim et al., 2012). Indeed, research shows companies that are subject to SEC enforcement actions tend to have abnormally high accruals prior to such actions (e.g., Richardson, Sloan, Soliman, and Tuna, 2006) and companies with high accruals also have a higher likelihood of earnings restatements (e.g., Richardson, Tuna, and Wu, 2002).

2) **MSCI ESG.** One of the most widely used ESG scores by institutional investors is computed by MSCI, and our sample for this variable is from January 2007 through March 2019. The MSCI score is a comprehensive assessment of each company’s ESG profile. We use the top-level ESG score that summarizes each company’s E, S, and G characteristics, on an industry-adjusted basis, as a numerical score from 0 (worst ESG) to 10 (best ESG).

3) **Carbon intensity, CO2 (negated).** As a measure of how “green” a company is (the E in ESG), we compute its carbon intensity (CO2), defined as the ratio of carbon emissions in tons over sales in millions of dollars. Carbon emissions can be measured in different ways, but we use the sum of “scope 1 carbon emissions” (a firm’s direct emissions, e.g., from the firm’s own fossil fuel usage) and “scope 2 carbon emissions” (indirect emissions from the use of electricity); we do not include “scope 3” (other indirect emissions) because these are rarely reported by companies and are at best noisily estimated and inconsistent across different data providers (e.g., Busch, Johnson, and Pioch, 2018). As with accruals, we negate the CO2 variable so that higher values indicate better ESG (less carbon intensive, “greener” companies). These data are obtained from Trucost and are available from January 2009 through March 2019.

4) **Non-sin stock.** Stocks in certain “sin” industries are shunned by some ESG-conscious investors, for example tobacco, gambling, or controversial weapons (related to the S in ESG). We consider a “non-sin stock” indicator, taking the value of 0 for sin stocks and the value of 1 otherwise, so that higher values indicate better ESG. Sin industries are defined as in Hong and Kacperczyk (2009), and this indicator is available for our longest sample, January 1963 through March 2019.

---

16 The MSCI website states that as of August 2018, “MSCI ESG Research is used by 46 of the top 50 asset managers and over 1,200 investors worldwide” (https://www.msci.com/esg-ratings, accessed July 7, 2019).
We merge these datasets with the XpressFeed database for stock returns and market values, the Compustat database to compute firm fundamentals, institutional holdings from 13f holdings reports (as aggregated by Thomson Reuters), signed order flow computed from intraday data,\(^\text{17}\) and the risk model of Barra US Equity (USE3L) that is used in the computation of the empirical EGS efficient frontier.

### 4.2. Empirical ESG-SR Frontier

We next estimate the ESG-SR frontier discussed in the model. We first focus on the governance proxy based on accruals, as discussed in Section 4.1, because it is perhaps the most likely to predict returns (which we confirm in Section 4.6) and hence may lead to more interesting frontiers. We empirically implement the model as follows. To focus on a liquid and realistic investment universe with high data coverage, we consider stocks in the S&P 500 Index using monthly data from January 1974 to March 2019. Using S&P 500 stocks makes the analysis conservative in the sense that we rule out that our results are driven by microcap stocks.

To compute the ESG-Sharpe ratio frontier, investors must first compute risk and expected returns. To compute risk (i.e., the variance–covariance matrix of the S&P 500 stocks), we assume that all investors use Barra’s US Equity risk model (Barra USE3L model), an industry standard for use in portfolio management.\(^\text{18}\) ESG-unaware investors and ESG-aware investors compute expected returns in different ways. U investors focus on the general equity risk premium and the traditional value factor, book-to-market, while A investors additionally use ESG information.

Specifically, to compute the annualized expected return of any stock \(i\) in any month \(t\), U investors do as follows:

\[
E_t^U(r_{i,t+1}) = \overline{MKT}_t + bm_{i,t} \overline{BM}_t
\]

where \(\overline{MKT}_t\) is the equity risk premium, \(bm_{i,t}\) is stock \(i\)'s cross-sectional book-to-market z-score (i.e., the stock's book-to-price ratio minus the cross-sectional mean, and dividing by the cross-sectional standard deviation), and \(\overline{BM}_t\) is the return premium of the factor-mimicking value factor. For each factor, the return premium at time \(t\) is its constant Sharpe ratio, multiplied by its volatility as estimated using the Barra model. Details on the estimation method are given in Appendix A.

Similarly, A investors (and M investors, but for brevity we focus on type-A here) compute the annualized expected return of stock \(i\) as

\[
E_t^A(r_{i,t+1}) = \overline{MKT}_t + bm_{i,t} \overline{BM}_t + s_{i,t} \overline{ESG}_t
\]

\(^{17}\) The variables related to signed order flow are defined as in Chordia et al. (2002) and Chordia and Subrahmanyam (2004) and are available between January 1993 and December 2012. We thank Tarun Chordia for kindly making these variables available to us.

\(^{18}\) Estimating the covariance matrix is not a contribution of this paper, so we use a third-party risk model for convenience. For more details about the risk model, please see Barra documentation, available, for example, at [http://www.alacra.com/alacra/help/barra_handbook_US.pdf](http://www.alacra.com/alacra/help/barra_handbook_US.pdf).
where \( s_{i,t} \) is the stock’s ESG score at time \( t \) and \( ESG \) is the return premium of the ESG factor, which is based on accruals. The ESG score \( s_{i,t} \) is computed as the cross-sectional (negated) accrual z-score. Because a stock’s ESG score \( s_i \) is normalized as a cross-sectional z-score, we get the intuitive interpretation that an ESG score of 0 means an average stock in terms of the ESG measure, a measure of 2 means that the typical stock has ESG characteristics two standard deviations better than the average stock, and so on. For a portfolio, the average ESG score is computed as in the theory section, \( \bar{s} = \frac{x's}{x'1}, \) which provides a similar intuition for long-only portfolios, but we note that long–short portfolios can in principle attain an unbounded range of ESG scores.

We are ready to compute the empirical ESG-SR frontier as seen in Figure 5. Figure 5 shows the frontier both from the perspective of U and A investors (solid and dashed lines, respectively), and we distinguish what we call the ex ante perceived frontiers (Panel A) and the realized frontiers (Panel B). Each investor’s perceived frontier is computed as follows: Each month, the investor computes risk and expected returns as defined previously and then derives the ESG-SR frontier and the corresponding frontier portfolios. Panel A simply shows the time-series average of these perceived frontiers. Panel B shows the realized Sharpe ratios of these portfolios.

We begin our discussion of Figure 5 with Panel A, which illustrates the economic tradeoffs perceived by investors U and A. Intuitively, having more information can be expected to improve portfolio outcomes: The globally optimal Sharpe ratio is about 12% higher for A than it is for U. The ESG-unaware investor U maximizes the Sharpe ratio for the ESG score of 0.25, meaning that a typical stock in her portfolio is close to average for this ESG measure. This near-neutrality to ESG is not surprising because the U investor only uses information on book-to-market ratios, and any exposure to accruals only happens incidentally through the weak correlation between book-to-market and accruals. Moreover, the frontier is relatively symmetric in the neighborhood of 0, meaning that this investor perceives the cost of targeting a positive ESG score to be similar to the cost of targeting a same-magnitude negative tilt on ESG. For example, targeting an ESG score two standard deviations higher than optimal (i.e., moving from 0.25 to 2.25) lowers the Sharpe ratio by about 9%; targeting an ESG score two standard deviations weaker than optimal (–1.75) degrades the Sharpe ratio by 7%.

In contrast, investor A’s frontier looks very different. First, we note that the frontier peaks at the ESG score of 2.25, meaning that for investor A maximizing the Sharpe ratio means targeting a portfolio with significantly higher ESG score than the market. Moreover, the frontier is clearly asymmetric, in a way that suggests that decreasing a portfolio’s ESG score will be meaningfully more costly to the Sharpe ratio than increasing it would be. For example, a two-standard deviation increase from the optimal point (2.25 to 4.25) reduces the Sharpe ratio by about 3%; the penalty for a similar move in the opposite direction (2.25 to 0.25) is three times as high, 9%.

This explains why the two frontiers intersect: Forcing a particularly negative ESG score is perceived as more costly by A than by U, eventually leading to strictly worse portfolio Sharpe ratios than what U perceives. The two curves cross at the ESG score of approximately zero, which is also intuitive; at this point, the optimal portfolio is essentially the same.
Because we model A’s expected returns as linear in the standardized accruals, forcing a portfolio to be neutral to accruals makes it neutral to the ESG factor.

Finally, Panel B of Figure 5 shows the realized Sharpe ratios of the portfolios that underlie the frontiers in Panel A. We see that A’s (ex post) realized frontier is quite similar to the ex ante perceived frontier, because the ESG score that drives the frontier is explicitly incorporated into A’s returns forecast, and because our model of ex ante risk and expected returns captures well the ex post realized returns.

In contrast, U’s realized frontier in Panel B in Figure 5 has a different shape than the perceived frontier in Panel A because U ignores that this ESG measure predicts returns. Investor U’s realized ESG-SR frontier looks fairly similar to that of investor A for ESG scores close to zero because their portfolios are more similar in that range, but U’s frontier is otherwise below because, for any ESG target, investor U chooses a portfolio with a suboptimal tradeoff between market exposure, value, and accruals.

4.3. Impact of Restrictions: Screening Out the Worst ESG Stocks

Our empirical application has so far allowed the investors to deploy their capital in unconstrained portfolios, going long and short any stock in the investment universe. It is interesting to also consider realistic constraints faced by many ESG-sensitive investors. Among such constraints, undoubtedly the most popular one is screening out stocks with the weakest ESG characteristics (i.e., removing such stocks from the investable universe). To investigate how such restrictions affect the investment opportunity set, Figure 6 shows how the ESG-SR frontier is affected by screens. In particular, Figure 6 shows three different frontiers: one for the unconstrained investor A (exactly as in Figure 5A), another obtained when the investor removes the 10% of stocks with the lowest ESG characteristics, and a third frontier with a 20% screen.

The first observation is perhaps the most obvious: constraints reduce a portfolio’s expected performance. Not surprisingly, the frontier with the 10% screen is strictly below the unconstrained one, and the frontier with a 20% screen is lower still. This means that, for any desired level of the ESG score, the maximum attainable Sharpe ratio will be lower in a screened universe than in the unrestricted one.

What is perhaps more interesting is the magnitude by which the Sharpe ratio decreases. To benchmark the reduction, a useful rule of thumb is that, under certain assumptions, the Sharpe ratio is approximately linear in the square root of investment breadth (e.g., Grinold and Kahn, 1995). This implies that a 10% (20%) reduction in breadth should lower the Sharpe ratio roughly by 5% (by 10%, respectively). Indeed, these are roughly the magnitudes of the decrease for ESG scores below about –0.5. The penalty is about half as small closer to the ESG score of 0, perhaps because around that value the optimal portfolio does not invest in extremely weak ESG stocks (or, presumably, in extremely strong ESG stocks). However, for the values of ESG score meaningfully above zero, the magnitude of the
penalty is sharply higher than what might be inferred from the square root of breadth rule of thumb. For example, removing the 20% of stocks with the lowest ESG reduces the Sharpe ratio by over 25% when the investor seeks to achieve high portfolio ESG scores, due in part to the benefits of shorting low-ESG stocks.

A related, and equally surprising, finding from Figure 6 is that the portfolio with the highest Sharpe ratio (the tangency portfolio) has a lower ESG score when the worst ESG stocks are removed. The unconstrained investor A optimizes the Sharpe ratio at the portfolio ESG score of 2.25, as discussed previously. After removing 10% of weakest ESG stocks, the Sharpe ratio is maximized at the ESG score of 1.5; after removing 20%, the optimum is an ESG score of 1.

This is a surprising result: Investors who exclude low-ESG assets from their investment universes may optimally build portfolios with lower ESG scores than do investors who allow for such low-ESG assets. The intuition behind this finding is that low-ESG assets are effectively funding sources, allowing the unconstrained investor to short them to build larger long positions in high-ESG securities. Moreover, low-ESG assets may be useful hedging instruments for high-ESG assets, and may help the investor improve the Sharpe ratio of the overall portfolio, potentially by increasing their investment in high-ESG securities. With screening the investor may optimally choose portfolios that do not take a large position in high-ESG assets.

4.4. Does ESG Predict Future Fundamentals?

A necessary condition for ESG-type information to generate positive abnormal returns is that it correlates with future fundamentals. To test for this possibility, we relate our ESG proxies to future fundamentals. We consider two measures of fundamentals in Table 1. In Panel A, we consider the “accounting rate of returns,” defined as the return on net operating assets as in Richardson et al. (2006); and in Panel B, gross profitability over assets defined as in Novy-Marx (2013) and computed as (revenue minus cost of goods sold) over total assets. In both panels, these firm fundamentals are measured 12 months after the ESG variables. For each of our four ESG proxies defined in Section 4.1, we present two specifications, one based on a pooled sample with month fixed effects and with standard errors clustered at the firm level, the other using the Fama–MacBeth procedure with Newey–West standard errors. We also control for firm beta, size, and book-to-market, although these control variables are not critical for our results.

Regressions (1) and (2) in Table 1, Panel A indicate that firms with good governance (based on low accruals) realize higher accounting rate of return over the subsequent fiscal year. The economic magnitude of the effect is relatively large. A one-standard-deviation increase in negated accruals predicts a corresponding increase of 0.02 in the accounting rate of returns, or 20% of its average level of 0.1. This finding opens up the possibility, which we confirm later, that accruals contain information about future fundamentals that may not be fully priced in by the market (similar to findings

---

19 ESG may lead to price increases even without a fundamentals channel if investor demand for ESG characteristics goes up. This is perhaps more likely over short periods in time and may not lead to a consistent return premium over a longer period.

20 Regressions without these controls are available from the authors on request.
of Richardson et al., 2006). The corresponding regressions in Panel B provide additional corroboration using the other measure of future firm fundamentals, gross profitability. Again, we see that a decrease in accruals predicts an increase in future profitability, but this time by a relatively smaller amount: A one-standard-deviation move of accruals is associated with a 0.006 move in gross profitability, or about 2% of its average level of 0.3.

The results for accruals are robust to a variety of controls. For example, there may be a variation in accruals across industries, but the addition of industry dummy variables to regression (1) does not change the coefficient (it slightly increases from 0.208 to 0.209, with a t-statistic of 22.6 vs. 23.3). Similarly, running the regressions without controls for firm size, book-to-market, or beta, or without date fixed effects, has little effect on the result. Lastly, we find a strong positive effect on accounting returns and on profitability even 24 or 36 months after we measure accruals. We conclude that there is strong evidence that accruals correlate with future profitability.

In contrast, we find mixed evidence that any of the other ESG proxies predict future firm fundamentals. MSCI ESG data and the non-sin stock indicator both correlate with higher profitability but are insignificantly related to the accounting rate of return. For profitability, they are relatively less economically important than accruals; for example, a one-standard-deviation improvement in MSCI ESG score has about a third of the impact of accruals. Additionally, we find that negated carbon intensity correlates with accounting returns but not with gross profitability (if anything, it correlates with lower profitability in the future).

Finally, while we do not attempt to find the “best” ESG metric and run a horse race between various proxies, we run a regression that includes all our ESG indicators as dependent variables. In these regressions, negated accruals remain positive predictors of future fundamentals, and we see some evidence that high carbon emissions companies have relatively lower future accounting returns and—unlike in specifications (5) and (6) in Table 1—also lower gross profitability.

4.5. Does ESG Predict Investor Demand?

As we explain in the theory section, correlation with future fundamentals is not enough in itself to determine whether an ESG variable should help or hurt returns. For the full picture, one also needs to analyze investor demand for ESG. In this section, we consider institutional ownership, trading activity, and signed order flow to capture investors’ interest in owning or purchasing a given stock.

Table 2, Panel A uses a similar empirical setup as Table 1 to relate institutional holdings in percentages (based on 13f data) to ESG metrics three months earlier (lag chosen to ensure that ESG variables are known before we observe institutional holdings) and our usual controls. Institutional investors (whose interest we measure using 13f filings) seem to incorporate ESG when forming their portfolios. Indeed, all our four ESG proxies correlate positively with institutional holdings. The economic effect of these variables is noticeable. For example, a one-standard-deviation change in accruals
is associated with a change in institutional holdings of 0.3%–1.3%, depending on the specification. A similar one-standard-deviation change in MSCI ESG (CO₂ intensity) scores indicates an ownership increase of 0.6%–0.8% (1.3%). A move from the worst to the best MSCI ESG score (score of 0 to score of 10) is associated with a change in institutional holdings of 3%–4%. A move from a sin stock to a non-sin stock implies a drop in holdings of 4%–6%.

Panels B and C in Table 2 consider measures of trading activity (logarithm of the number of trades) and signed order flow (the fraction of dollar volume that is attributable to buys). For brevity, we only report pooled regressions with date fixed effects; Fama–MacBeth evidence is similar and available on request. The results are perhaps most intuitive for accruals, where we see that both the number of trades and the fraction of buys increase when this ESG proxy improves. For the other three metrics, however, evidence is not as straightforward. The number of trades seems to decrease for stocks with low carbon intensity or for non-sin stocks. For the latter, at least we see an increase in the fraction of buys; for the former, if anything, the larger number of trades are relatively more likely to be seller-initiated.

4.6. Does ESG Predict Valuation and Future Returns?

The findings so far suggest that at least some ESG proxies (specifically, accruals) robustly correlate with future fundamentals. At the same time, we find evidence that investors tilt their portfolios toward stocks with more attractive accruals. As we show in the theory section, the interplay between the two may potentially lead to a return premium (or discount if the latter effect is overwhelming). The prediction is perhaps easier to make for the other ESG proxies, where we found no correlation to future fundamentals but some evidence of investor demand. This, if anything, would indicate that high-ESG stocks based on these measures could be expensive and have low future returns. To assess these predictions, we consider valuations (Tobin’s Q) and risk-adjusted returns in Tables 3 and 4.

Table 3 shows how ESG correlates with the logarithm of the price-to-book ratio. Because our interest here is how much the market is willing to pay for ESG characteristics, we analyze the relationship between contemporaneous valuation and ESG proxies. We control for market beta, but we naturally omit the previously used control variables that are related to valuation by construction (i.e., size and book-to-market).

Regression (1) in Table 3 suggests that accruals’ predictability of future fundamentals may not be priced by the market. We see low valuation ratios for stocks with high scores of this ESG measure (i.e., high values of negated accruals), suggesting that such stocks are cheap. Hence, we see that stocks with low accruals tend to have stronger forecasted profitability and yet are trading at relatively cheap valuations. This opens up the possibility that such stocks generate attractive returns, which is something we confirm below.

At the same time, regressions (2) and (3) in Table 3 suggest that stocks with high ESG scores measured based on MSCI ESG or CO₂ trade at high valuations. In other words, such stocks appear expensive, consistent with the relatively higher
demand from investors that we documented earlier. Lastly, regression (4) shows little relation between valuation and sin characteristics.

We next turn to return predictability in Table 4. We analyze return predictability of ESG in a standard portfolio framework, sorting stocks into ESG quintiles (in the case of sin/non-sin indicator, into two portfolios) each month and then forming a spread portfolio that goes long the best ESG stocks and goes short the worst ESG stocks, for each of our four ESG proxies. Table 4 presents the resulting performance, for both the equal-weighted and value-weighted portfolios and controlling for a variety of asset pricing models.

Of our four ESG proxies, accruals strongly correlate with subsequent returns. The economic magnitude of this effect is substantial, up to 7% a year for the equal-weighted and up to 3% a year for the value-weighted portfolio, even after controlling for the five Fama–French factors augmented with momentum.

For the next two ESG proxies, MSCI ESG and carbon intensity, we find no or weak evidence of abnormal returns. It seems that, over our sample period, less carbon intense companies relatively outperformed based on the point estimate, but this effect is only significant at the 10% level. For carbon intensity, we note that our results are relatively similar to those of Bolton and Kacperczyk (2019), who find a carbon premium in returns but show that it disappears in richer specifications, for example when they control for industry composition.

Finally, we find weak evidence for the sin premium documented in Hong and Kacperczyk (2009). Because we consider a spread portfolio long in non-sin stocks (higher ESG) and short sin stocks (lower ESG), a sin return premium would be reflected as a negative alpha estimate. We note that our results are weaker than those of Hong and Kacperczyk (2009), possibly because of differences in methodology (we compare sin stocks versus the whole universe of non-sin stocks rather than versus a matched sample) and in sample period. Nonetheless, our point estimates indicate a sin premium of up to 5% a year, although this estimate is so large and statistically significant only for one- and three-factor alphas of value-weighted returns. The sin premium becomes insignificant when we consider five-factor and six-factor models (with both equal- and value-weighted returns), consistent with findings of Blitz and Fabozzi (2017).

5. Conclusion: Ethical, Saintly, and Guiltless (ESG) Investing

Investors, especially institutions such as pension funds, increasingly incorporate ESG views in their portfolios. Said simply and slightly tongue-in-cheek, many investors want to own ethical companies in a saintly effort to promote good corporate behavior, while simultaneously hoping to do this in a guiltless way that doesn’t sacrifice the investor’s returns.

21 The last years in our sample are particularly difficult for sin stocks. In particular, tobacco companies posted historically weak results; for example, the MSCI World Tobacco index underperformed the cap-weighted benchmark in each of 2016, 2017, and 2018, by about 1%, 9%, and 28%, respectively.
Investors must realistically evaluate the costs and benefits of responsible investing, and we hope that our framework will be a useful way to conceptualize and quantify these costs and benefits. Indeed, we show that a responsible investor’s decision can be conceptualized by the ESG-efficient frontier, a graphical illustration of the investment opportunity set. The benefit of ESG information can be quantified as the resulting increase in the maximum Sharpe ratio (relative to a frontier based on only non-ESG information). The cost of ESG preferences can be quantified as the drop in Sharpe ratio when choosing a portfolio with better ESG characteristics than those of the portfolio with maximum Sharpe.

In addition to its practical appeal, the ESG-efficient frontier is based on a rigorous theoretical framework. We explicitly derive the frontier and the corresponding set of optimal portfolios. The optimal portfolios are spanned by four “funds,” one of which captures stocks’ ESG scores. This finding can be viewed as a theoretical foundation for what is called “ESG integration,” meaning that ESG characteristics are used directly in portfolio construction (rather than as screens).

Empirically, we find that when ESG is proxied for by a measure of governance based on accruals, the maximum SR is achieved for a relatively high level of ESG. Increasing the ESG level even further leads to only a small reduction in SR, implying that ethical goals may be achieved at a small cost. When we impose realistic constraints on the portfolio, we see a downward shift in the ESG-SR frontier. This is an expected outcome, because imposing constraints reduces the maximum Sharpe ratio that one can attain for any given ESG score. More surprisingly, screens that remove the lowest-ESG assets from the investment universe can lead investors who maximize their Sharpe ratio to choose a portfolio with lower ESG scores than those chosen by unconstrained investors who allow investments in low-ESG assets. This result highlights nuances in optimally incorporating ESG into portfolio construction and suggests improvements to traditional approaches based on simple screening.

Turning to equilibrium asset prices, we derive an ESG-adjusted CAPM, which helps describe market environments that make ESG predict returns positively or negatively. To our knowledge, our model is the first to explicitly model heterogeneity in how investors use ESG information. We allow for investors to have preferences over ESG and for the possibility that investors can find investment intelligence from ESG information. We argue that this is a realistic feature, because not only do we observe large AUM deployed with ESG in mind (e.g., the Global Sustainable Investment Review 2018), but ESG is increasingly discussed as a potential “alpha” signal, both in academic outlets (going back to at least Sloan, 1996, and Gompers et al., 2003) and in practitioner journals (e.g., Nagy, Kassam, and Lee, 2015). This heterogeneity results in a range of possible equilibria depending on the relative importance of each investor type, leading to a relation between ESG and expected returns that is positive, negative, or neutral.

We test the empirical predictions of the theory using a range of ESG proxies that reflect different aspects of our model and that may represent different clienteles of investors. Perhaps the most intriguing case is our accruals-based
proxy, which we show helps predict future fundamentals, and while it attracts relatively more investor demand, it still trades at relatively cheap valuations and generates positive returns. At the other end of the spectrum are sin stocks, which do not have meaningfully different fundamentals but have much weaker demand by institutional investors. We show that this demand likely affects valuations, leading to some sin premium in the returns. Finally, we find no consistent premium or penalty to the last two ESG proxies we consider, MSCI ESG data and carbon emissions.

In conclusion, we think that our model provides a useful framework for responsible investment that we hope will be useful both for future research on the costs and benefits of ESG investing and for ESG applications in investments practice.
References


Fitzgibbons, Shaun, Christopher Palazzolo, and Lukasz Pomorski. 2018. “ESG 2.0: Hit’em Where It Hurts.” IPE.


Appendix A: Estimating the Empirical ESG-Efficient Frontier

As discussed in Section 4.2, we model expected returns as linear functions of factor exposures. For instance, investor U estimates expected returns as

\[ E_t^U(r_{t+1}) = MKT_t + bm_{t,t} BM_t \]

where \( MKT_t \) is the equity risk premium, \( bm_{i,t} \) is stock \( i \)'s cross-sectional book-to-market z-score, and \( BM_t \) is the return premium of the factor-mimicking value factor, and similarly for investor A, who additionally includes an ESG factor. To show how we estimate these models, it is helpful to write them in a general way that captures either investor type. We first show how we model the vector of realized returns, \( r_{t+1} \), and then later we turn to the expected returns, \( E_t^j(r_{t+1}) \), for investor \( j \in \{U,A\} \). Realized returns follow a standard factor model:

\[ r_{t+1} = X_t F_{t+1} + \epsilon_{t+1} \]

where \( X_t \) is a matrix of all securities’ factor exposures, \( F_{t+1} \) is a vector of factor returns, and \( \epsilon_{t+1} \) are the idiosyncratic shocks. In particular, for investor U, \( X_t \) is an \( N \times 2 \) matrix where the first column is a vector of ones and the second contains the book-to-market z-scores. For investor A, \( X_t \) is an \( N \times 3 \) matrix where the first two columns are the same and the third column is a vector of ESG z-scores. Even though investors U and A use different factor models (i.e., different \( X_t \) and \( F_{t+1} \)), we use the same notation for simplicity.

The factor returns \( F_{t+1} \) are unobserved, but they can be estimated as follows. Each time period, we run a cross-sectional regression of stock returns on their characteristics and note that the regression coefficients are the factor returns. Specifically, we run a GLS regression each period of stock-level returns on stock-level characteristics, using the Barra risk model to obtain an estimate of the residual covariance matrix, \( \Sigma_t = \text{var}(\epsilon_{t+1}) \), which yields the following estimated factor returns:

\[ \hat{F}_{t+1} = (X_t^T \Sigma_t^{-1} X_t)^{-1} X_t^T \Sigma_t^{-1} r_{t+1} \]

Here, we can interpret \( \theta_t := (X_t^T \Sigma_t^{-1} X_t)^{-1} X_t^T \Sigma_t^{-1} \) as the factor-mimicking portfolio weights, i.e., \( \hat{F}_{t+1} = \theta_t r_{t+1} \).

Finally, we need to compute expected returns:

\[ E_t^j(r_{t+1}) = X_t E_t^j(F_{t+1}) \]

which means that we need to compute estimate expected factor returns, \( E_t^j(F_{t+1}) \). The simplest way to do this is to assume that \( E_t^j(F_{t+1}) \) is constant over time, and then estimate the factor premiums as the sample average of factor returns. This simple method does not work well empirically, however, because it leads, for example, to perceived and realized ESG-SR frontiers that differ significantly even for investor A. This problem arises because investors have an
incentive to choose extreme portfolios when the perceived risk is time-varying (i.e., sometimes very low) while the perceived expected return is constant.

A more realistic specification is to assume that each factor $k$ has a time-varying risk and a constant Sharpe ratio, $E_t^1(F_{t+1}^k) = \sigma_t^{F,k} SR_{F,k}$. The volatility of each factor, $\sigma_t^{F,k}$, can be computed based on the factor-mimicking portfolio weights and the overall risk model, $\sigma_t^{F,k} = \sqrt{\theta_t^k \Sigma_t (\theta_t^k)^T}$. Finally, we estimate $SR_{F,k}$ as the realized full-sample Sharpe ratio of the volatility-scaled factor returns, $\hat{F}_{t+1}/\sigma_t^{F,k}$.

**Appendix B: Proofs**

**Proof of Proposition 1.** Consider the problem of maximizing the return given a level of risk $\sigma$ and an ESG score $\bar{s}$:

\[
\max_x \left( x' \mu - \frac{\gamma}{2} \sigma^2 + f(\bar{s}) \right) \quad \text{st. } \frac{x' \bar{s}}{\chi_1} = 0, \quad \sigma^2 = x' \Sigma x
\]

Clearly, maximizing the expected return for given level of $\sigma$ and $\bar{s}$ is achieved by maximizing the Sharpe ratio for that $\sigma$ and $\bar{s}$. Further, the resulting Sharpe ratio is the same for all levels of $\sigma$. To see why, suppose that $x_1$ is the optimal solution for $(\sigma_1, \bar{s})$ and $x_2$ is the optimal solution for $(\sigma_2, \bar{s})$. We can scale $x_2$ as $x_1/\sigma_1 x_2$ to have a volatility of $\sigma_1$, and importantly, this scaled portfolio still has the same average ESG score, $\bar{s}$. Given $x_1$ has the highest expected return among such portfolios, we have

\[
SR(x_1) = \frac{x_1' \mu}{\sigma_1} \geq \frac{x_2' \mu}{\sigma_2} = \frac{x_2' \mu}{\sigma_1} = SR(x_2)
\]

The symmetric argument shows that the opposite inequality also holds, so $SR(x_1) = SR(x_2) = SR(\bar{s})$.

Let us solve the problem explicitly. If we rewrite the first constraint as $x' \bar{s} = 0$, where $\bar{s} = s - 1 \bar{s}$, and introduce Lagrange multipliers $\pi$ and $\theta$, then we see that the solution is characterized by the first-order condition

\[
0 = \mu + \pi \bar{s} - \theta \Sigma x
\]

meaning that the optimal portfolio is given by

\[
x = \frac{1}{\theta} \Sigma^{-1}(\mu + \pi \bar{s})
\]

Both constraints clearly bind and the first one yields

\[
0 = \frac{1}{\theta} \bar{s}' \Sigma^{-1}(\mu + \pi \bar{s})
\]
So, the first Lagrange multiplier is

\[
\pi = -\frac{s'\Sigma^{-1}\mu}{s'\Sigma^{-1}s} = -\frac{(s - 1s')\Sigma^{-1}\mu}{(s - 1s')\Sigma^{-1}(s - 1s')} = \frac{c_{1\mu}s - c_{s\mu}}{c_{ss} - 2c_{1s}s + c_{11}s^2}
\]

The second constraint yields

\[
\sigma^2 = \frac{1}{\theta^2} (\mu + \pi s)'\Sigma^{-1}(\mu + \pi s)
\]

Using the first constraint, we can simplify as

\[
\sigma^2 = \frac{1}{\theta^2} \mu^'\Sigma^{-1}(\mu + \pi s)
\]

implying that the second Lagrange multiplier is

\[
\theta = \frac{1}{\sigma} \sqrt{c_{ss} - \left(c_{s\mu} - c_{1\mu}s\right)^2}
\]

This shows explicitly that we can write the optimal portfolio as \(x = \sigma v\), where the vector \(v\) only depends on the exogenous parameters and \(s\), that is, not on \(\sigma\). ■

**Proof of Proposition 2.** The maximum Sharpe ratio for a given ESG score \(s\) is the Sharpe ratio of the optimal portfolio given in the proof of Proposition 1.

\[
SR(s) = \frac{\mu'x}{\sigma} = \frac{\mu^'\Sigma^{-1}(\mu + \pi s)}{\sigma \theta}
\]

Using the last two equations in the proof of Proposition 1, we see that

\[
SR(s) = \sigma \theta = \sqrt{\mu^'\Sigma^{-1}(\mu + \pi s)} = \sqrt{c_{ss} - \left(c_{s\mu} - c_{1\mu}s\right)^2}
\]

Clearly the maximum Sharpe ratio is attained by the tangency portfolio, which is proportional to \(\Sigma^{-1}\mu\). This portfolio has the ESG score and Sharpe stated in the proposition. This result can also be derived by differentiating the \(SR(s)\) with respect to \(s\) and considering the first- and second-order conditions (there are two solutions to the first-order condition). ■

**Proof of Proposition 3.** This result follows from the proof of Proposition 1. ■

**Proof of Propositions 4–5.** These results follow based on arguments analogous to those in the first part of the proof of Proposition 1. ■
Proof of Propositions 6–7. We start with proposition 7. The equilibrium condition with all investors of type M is

\[ p = \frac{w}{y} \text{diag}(p_i) \Sigma^{-1} \text{diag}(p_i) \left( \text{diag} \left( \frac{1}{p_i} \right) \bar{\mu} - 1 - r^f + \pi(s - s^m) \right) \]

This condition can be simplified by multiplying both sides by \( \text{diag} \left( \frac{1}{p_i} \right) \) as

\[ 1 = \frac{w}{y} \Sigma^{-1} \left( \bar{\mu} - \text{diag}(p_i) (1 + r^f - \pi(s - s^m)) \right) \]

Solving this equation for the vector of prices \( p \) yields the following result:

\[ p = \text{diag} \left( \frac{1}{1 + r^f - \pi(s_i - s^m)} \right) \left( \bar{\mu} - \frac{y}{w} \Sigma \bar{1} \right) \]

which proves equation (14) stated in the proposition. To translate this result to expected excess returns, we multiply both sides by \( \text{diag} \left( \frac{1}{p_i} \right) \)

\[ 1 = \text{diag} \left( \frac{1}{p_i} \right) \text{diag} \left( \frac{1}{1 + r^f - \pi(s_i - s^m)} \right) \left( \bar{\mu} - \frac{y}{w} \Sigma \bar{1} \right) \]

and rearrange to obtain

\[ \text{diag}(1 + r^f - \pi(s_i - s^m)) = \text{diag} \left( \frac{1}{p_i} \right) \left( \bar{\mu} - \frac{y}{w} \Sigma \bar{1} \right) \]

Now we see that the vector of expected excess returns \( \mu \) is given by

\[ \mu = \text{diag} \left( \frac{1}{p_i} \right) \bar{\mu} - 1 - r^f = \frac{y}{w} \text{diag} \left( \frac{1}{p_i} \right) \Sigma \bar{1} - \text{diag}(\pi(s_i - s^m)) \]

The expected excess return of the market portfolio \( \left( \frac{p_i}{1'p} \right) \) is given by

\[ \mu^m = 1' \text{diag} \left( \frac{p_i}{1'p} \right) \mu = \frac{y}{w(1'p)} 1' \Sigma \bar{1} - 1' \text{diag} \left( \frac{p_i}{1'p} \right) \text{diag}(\pi(s_i - s^m)) \]

That is,

\[ \mu^m = \frac{y(1'p)}{w} \text{var}(r^m | s) - \pi(s^m - s^m) = \frac{y(1'p)}{w} \text{var}(r^m | s) \]

where we use the definition of the ESG score of the market \( s^m = \frac{1}{1'p} p's \). The expected excess return of security \( i \) can be written as \( \mu_i = z_i' \mu \) using the i'th unit vector \( z_i = (0, \ldots, 0, 1, 0, \ldots, 0) \), that is,

\[ \mu_i = \frac{y}{w} z_i' \text{diag} \left( \frac{1}{p_i} \right) \Sigma \bar{1} - z_i' \text{diag}(\pi(s_i - s^m)) = \frac{y(1'p)}{w} \text{cov}(r_i, r^m | s) - \pi(s_i - s^m) \]
Combining with the expression above for $\mu_m$, we get $\mu_i = \bar{\mu}_i \mu^m - \pi(s_i - s^m)$. Finally, when all investors are of type-A and choose the tangency portfolio, we have $\pi = 0$, which is seen from Proposition 3 and the expression for $\pi$.

Turning to the proof of Proposition 6, similar calculations show that prices are given by

$$p = \frac{1}{1+r^f}(\bar{\mu} - \frac{\gamma}{W} \Sigma 1)$$

and returns by the unconditional CAPM. Conditional expected returns are given by

$$E(r_i|s) = \frac{E(v_i|s)}{p_i} - (1 + r^f) = \frac{\bar{\mu}_i + \lambda(s_i - s^m)}{p_i} - (1 + r^f)$$

Using the expression for the price, we see

$$E(r_i|s) = \frac{\gamma \text{cov}(v_i,v^m)}{p_i} + \lambda \text{var}(r^m) E(r^m) + \frac{\lambda (s_i - s^m)}{p_i}$$

where we use that $\frac{\text{cov}(v_i,v^m)}{p_i(1'p)} = \text{cov}(r_i, r^m)$ and $\frac{\gamma (1'p)}{W} = \frac{E(r^m)}{\text{var}(r^m)}$. ■
Figure 3. ESG-Efficient Frontier and Indifference Curves for an ESG-motivated Investor. This figure shows an example of an ESG-Sharpe ratio frontier for an ESG-motivated investor M (solid line). The investor’s utility increases in both the Sharpe ratio and the ESG score of her portfolio, yielding a tradeoff illustrated by the downward-sloping indifference curves (dashed lines).

![Graph showing the efficient frontier and indifference curves for an ESG-motivated investor.](image)

Figure 4. ESG-Efficient Frontier and Indifference Curves for an ESG-Aware Investor. This figure shows an ESG-Sharpe ratio frontier (solid line) and an ESG-aware investor’s indifference curves (dashed lines), which are horizontal because this type of investor does not derive direct utility from ESG.

![Graph showing the efficient frontier and indifference curves for an ESG-aware investor.](image)
Figure 5. Empirical ESG-Efficient Frontier. We estimate the ESG-Sharpe ratio frontier for S&P 500 stocks, with returns driven by valuation (measured by each stock’s book-to-market ratio) and ESG (measured by each stock’s accruals to assets ratio, a measure related to governance). The figure shows annualized maximum Sharpe ratios attainable for each level of ESG constraint. The ESG-unaware investor U (solid blue line) solely utilizes book-to-market to estimate expected returns; The ESG-aware investor A (dashed line) uses both book-to-market and a measure of governance (the “G” in ESG) based on accruals to estimate expected returns. Panel A presents the perceived frontier, built using the ex ante estimates from each investor. Panel B presents the realized frontier, constructed using the portfolios from Panel A and computing their ex post performance.

Panel A: Ex ante perceived ESG-Sharpe ratio frontiers

Panel B: Realized ESG-Sharpe ratio frontiers
Figure 6. The Impact of screening on the ESG-Sharpe ratio frontier. This figure shows an ESG-aware investor’s perceived ESG-Sharpe ratio frontier (solid blue line, the same as the solid line in Figure 4A) as well as two frontiers for an investor who only allows herself to use a screened investment universes: removing 10% of stocks with the lowest ESG scores (dashed green line), or removing 20% of stocks (dotted red line).
Table 1: Does ESG Predict Firm Profits? This table reports the regression of future profitability on current ESG scores, where profitability is measured 12 months into the future. Profitability is computed as the accounting return (return on net operating assets, RNOA) in Panel A and as gross profit over assets in Panel B. We consider four ESG metrics and three control variables (market beta, the logarithm of market capitalization, and the logarithm of the book-to-price ratio). The ESG metrics are a measure of governance: “Accruals (negated)”; the overall “MSCI ESG” score; “CO2 (negated)” as a measure of low carbon usage; and “Non-sin stock” (all signed so that higher values are better ESG). “Estimation method” is either a pooled regression with month fixed effects (“pooled”) or Fama–MacBeth (“FM”), as indicated. Robust t-statistics are in parentheses, which are clustered at the stock level in pooled regressions or adjusted using a Newey–West weighting scheme in Fama–MacBeth regressions.

Panel A: Predicting the Accounting Return, RNOA

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accruals (negated)</td>
<td>0.208***</td>
<td>0.193***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(23.26)</td>
<td>(28.64)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSCI ESG</td>
<td>0.0001</td>
<td>0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO2 (negated)</td>
<td></td>
<td></td>
<td>5.9e-06***</td>
<td>6.0e-06***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4.91)</td>
<td>(7.34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sin stock</td>
<td>-0.007</td>
<td>-0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.18)</td>
<td>(-1.29)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta</td>
<td>-0.060***</td>
<td>-0.062***</td>
<td>-0.052***</td>
<td>-0.040***</td>
<td>-0.068***</td>
<td>-0.068***</td>
<td>-0.064***</td>
<td>-0.067***</td>
</tr>
<tr>
<td></td>
<td>(-31.79)</td>
<td>(-19.43)</td>
<td>(-11.62)</td>
<td>(-4.40)</td>
<td>(-17.90)</td>
<td>(-10.24)</td>
<td>(-33.78)</td>
<td>(-20.69)</td>
</tr>
<tr>
<td>Ln market cap</td>
<td>0.014***</td>
<td>0.014***</td>
<td>0.008***</td>
<td>0.006***</td>
<td>0.011***</td>
<td>0.011***</td>
<td>0.015***</td>
<td>0.015***</td>
</tr>
<tr>
<td></td>
<td>(30.14)</td>
<td>(26.85)</td>
<td>(6.54)</td>
<td>(4.89)</td>
<td>(12.45)</td>
<td>(23.91)</td>
<td>(32.75)</td>
<td>(26.56)</td>
</tr>
<tr>
<td>Ln (B/P)</td>
<td>-0.028***</td>
<td>-0.028***</td>
<td>-0.026***</td>
<td>-0.038***</td>
<td>-0.014***</td>
<td>-0.015***</td>
<td>-0.027***</td>
<td>-0.028***</td>
</tr>
<tr>
<td></td>
<td>(-23.73)</td>
<td>(-22.11)</td>
<td>(-9.27)</td>
<td>(-11.94)</td>
<td>(-6.72)</td>
<td>(-6.98)</td>
<td>(-22.58)</td>
<td>(-21.97)</td>
</tr>
<tr>
<td>RNOA (t)</td>
<td>0.725***</td>
<td>0.720***</td>
<td>0.756***</td>
<td>0.734***</td>
<td>0.763***</td>
<td>0.765***</td>
<td>0.710***</td>
<td>0.707***</td>
</tr>
<tr>
<td></td>
<td>(169.65)</td>
<td>(128.80)</td>
<td>(63.53)</td>
<td>(61.25)</td>
<td>(88.59)</td>
<td>(97.48)</td>
<td>(167.53)</td>
<td>(118.87)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.019***</td>
<td>-0.009</td>
<td>0.002</td>
<td>0.001</td>
<td>0.020***</td>
<td>0.021**</td>
<td>-0.006</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(-6.59)</td>
<td>(-1.56)</td>
<td>(0.19)</td>
<td>(0.06)</td>
<td>(2.78)</td>
<td>(2.32)</td>
<td>(-1.05)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,354,499</td>
<td>1,354,499</td>
<td>116,130</td>
<td>116,130</td>
<td>239,440</td>
<td>239,440</td>
<td>1,374,620</td>
<td>1,374,620</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.636</td>
<td>0.635</td>
<td>0.723</td>
<td>0.727</td>
<td>0.708</td>
<td>0.712</td>
<td>0.631</td>
<td>0.631</td>
</tr>
<tr>
<td>Estimation method</td>
<td>Pooled</td>
<td>FM</td>
<td>Pooled</td>
<td>FM</td>
<td>Pooled</td>
<td>FM</td>
<td>Pooled</td>
<td>FM</td>
</tr>
</tbody>
</table>

Electronic copy available at: https://ssrn.com/abstract=3466417
### Panel B: Predicting profitability

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCRUALS (negated)</td>
<td>0.061***</td>
<td>0.070***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.66)</td>
<td>(14.46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSCI ESG</td>
<td></td>
<td>0.001**</td>
<td>0.001***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.49)</td>
<td>(3.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO2 (negated)</td>
<td></td>
<td></td>
<td></td>
<td>-2.42e-06</td>
<td>-0.000*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.96)</td>
<td>(-1.79)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sin stock</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.003</td>
<td>0.005***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.95)</td>
<td>(2.88)</td>
</tr>
<tr>
<td>Beta</td>
<td>-0.009***</td>
<td>-0.008***</td>
<td>-0.017***</td>
<td>-0.015***</td>
<td>-0.022***</td>
<td>-0.014**</td>
<td>-0.025**</td>
<td>-0.013***</td>
</tr>
<tr>
<td></td>
<td>(-5.15)</td>
<td>(-3.59)</td>
<td>(-6.98)</td>
<td>(-3.30)</td>
<td>(-4.89)</td>
<td>(-2.29)</td>
<td>(-2.38)</td>
<td>(-5.91)</td>
</tr>
<tr>
<td>Ln market cap</td>
<td>-0.001*</td>
<td>-0.001***</td>
<td>-0.001*</td>
<td>-0.001**</td>
<td>-0.005</td>
<td>-0.004**</td>
<td>-0.001</td>
<td>-0.002***</td>
</tr>
<tr>
<td></td>
<td>(-1.85)</td>
<td>(-4.13)</td>
<td>(-1.93)</td>
<td>(-2.24)</td>
<td>(-1.43)</td>
<td>(-2.37)</td>
<td>(-1.37)</td>
<td>(-3.83)</td>
</tr>
<tr>
<td>Ln (B/P)</td>
<td>-0.002**</td>
<td>-0.002***</td>
<td>-0.006***</td>
<td>-0.006***</td>
<td>-0.036</td>
<td>-0.038**</td>
<td>-0.012</td>
<td>-0.014***</td>
</tr>
<tr>
<td></td>
<td>(-2.21)</td>
<td>(-3.34)</td>
<td>(-4.83)</td>
<td>(-8.58)</td>
<td>(-1.32)</td>
<td>(-2.25)</td>
<td>(-1.53)</td>
<td>(-3.42)</td>
</tr>
<tr>
<td>GPOA (t)</td>
<td>0.960***</td>
<td>0.960***</td>
<td>0.954***</td>
<td>0.948***</td>
<td>1.026***</td>
<td>1.017***</td>
<td>0.978***</td>
<td>0.980***</td>
</tr>
<tr>
<td></td>
<td>(160.64)</td>
<td>(252.59)</td>
<td>(102.37)</td>
<td>(177.07)</td>
<td>(25.35)</td>
<td>(63.36)</td>
<td>(49.30)</td>
<td>(132.27)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.020***</td>
<td>0.023***</td>
<td>0.026***</td>
<td>0.028***</td>
<td>0.019***</td>
<td>0.010</td>
<td>0.023***</td>
<td>0.015***</td>
</tr>
<tr>
<td></td>
<td>(8.73)</td>
<td>(7.91)</td>
<td>(5.26)</td>
<td>(3.51)</td>
<td>(3.11)</td>
<td>(1.13)</td>
<td>(6.50)</td>
<td>(4.45)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,521,202</td>
<td>1,521,202</td>
<td>171,284</td>
<td>171,284</td>
<td>361,540</td>
<td>361,540</td>
<td>1,877,268</td>
<td>1,877,268</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.712</td>
<td>0.747</td>
<td>0.866</td>
<td>0.892</td>
<td>0.087</td>
<td>0.684</td>
<td>0.267</td>
<td>0.686</td>
</tr>
<tr>
<td>Estimation method</td>
<td>Pooled</td>
<td>FM</td>
<td>Pooled</td>
<td>FM</td>
<td>Pooled</td>
<td>FM</td>
<td>Pooled</td>
<td>FM</td>
</tr>
</tbody>
</table>

Electronic copy available at: https://ssrn.com/abstract=3466417
Table 2: Does ESG Predict Investor Demand? This table reports the regression of investor demand on measures of ESG. Investor demand is measured as institutional ownership (obtained from 13f reports, led by three months) in Panel A, trading activity in Panel B (log number of trades in the next month), and signed order flow (dollar buy volume over total dollar volume) in Panel C. The ESG proxies and control variables are as in Table 1. The estimation method is either a pooled regression with month fixed effects ("pooled") or Fama–MacBeth ("FM"), as indicated. Robust $t$-statistics are in parentheses, which are clustered at the stock level in pooled regressions or adjusted using a Newey–West weighting scheme in Fama–MacBeth regressions.

**Panel A: Predicting Institutional Ownership**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accruals (negated)</td>
<td>1.060</td>
<td>3.208***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.74)</td>
<td>(2.98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSCI ESG</td>
<td></td>
<td>0.343**</td>
<td>0.420***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.55)</td>
<td>(6.98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO2 (negated)</td>
<td></td>
<td></td>
<td></td>
<td>0.002***</td>
<td>0.002***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3.37)</td>
<td>(14.65)</td>
<td></td>
<td></td>
<td>(3.37)</td>
<td>(14.65)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sin stock</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.416*</td>
<td>6.010***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.66)</td>
<td>(7.22)</td>
</tr>
<tr>
<td>Beta</td>
<td>1.610***</td>
<td>3.038***</td>
<td>6.371***</td>
<td>5.512***</td>
<td>5.774***</td>
<td>5.912***</td>
<td>5.707***</td>
<td>6.919***</td>
</tr>
<tr>
<td>(3.37)</td>
<td>(11.91)</td>
<td>(7.05)</td>
<td>(11.27)</td>
<td>(8.50)</td>
<td>(21.96)</td>
<td>(14.16)</td>
<td>(20.83)</td>
<td>(64.95)</td>
</tr>
<tr>
<td>(53.67)</td>
<td>(85.18)</td>
<td>(3.32)</td>
<td>(-2.67)</td>
<td>(50.48)</td>
<td>(108.99)</td>
<td>(62.29)</td>
<td>(64.29)</td>
<td>(64.95)</td>
</tr>
<tr>
<td>Ln (B/P)</td>
<td>2.282***</td>
<td>1.931***</td>
<td>-1.136***</td>
<td>-1.642***</td>
<td>0.321</td>
<td>0.354***</td>
<td>1.749***</td>
<td>1.251***</td>
</tr>
<tr>
<td>(13.90)</td>
<td>(13.83)</td>
<td>(-3.86)</td>
<td>(-9.22)</td>
<td>(1.20)</td>
<td>(5.08)</td>
<td>(10.98)</td>
<td>(8.28)</td>
<td>(8.28)</td>
</tr>
<tr>
<td>(-3.00)</td>
<td>(-9.55)</td>
<td>(24.56)</td>
<td>(18.45)</td>
<td>(-6.77)</td>
<td>(-17.28)</td>
<td>(-5.57)</td>
<td>(-15.59)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>737,865</td>
<td>737,865</td>
<td>180,326</td>
<td>180,326</td>
<td>378,623</td>
<td>378,623</td>
<td>962,867</td>
<td>962,867</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.475</td>
<td>0.422</td>
<td>0.033</td>
<td>0.083</td>
<td>0.454</td>
<td>0.450</td>
<td>0.470</td>
<td>0.424</td>
</tr>
<tr>
<td>Estimation method</td>
<td>Pooled</td>
<td>FM</td>
<td>Pooled</td>
<td>FM</td>
<td>Pooled</td>
<td>FM</td>
<td>Pooled</td>
<td>FM</td>
</tr>
</tbody>
</table>
### Panel B: Predicting Number of Trades

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accruals (negated)</td>
<td>0.282***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.44)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSCI ESG</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO2 (negated)</td>
<td>-0.0001***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.46)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sin stock</td>
<td>-0.159***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta</td>
<td>0.940***</td>
<td>0.989***</td>
<td>1.382***</td>
<td>0.937***</td>
</tr>
<tr>
<td></td>
<td>(43.56)</td>
<td>(21.81)</td>
<td>(29.97)</td>
<td>(43.52)</td>
</tr>
<tr>
<td>Ln market cap</td>
<td>0.724***</td>
<td>0.642***</td>
<td>0.898***</td>
<td>0.709***</td>
</tr>
<tr>
<td></td>
<td>(108.31)</td>
<td>(37.60)</td>
<td>(67.04)</td>
<td>(111.70)</td>
</tr>
<tr>
<td>Ln (B/P)</td>
<td>0.085***</td>
<td>0.075***</td>
<td>0.003</td>
<td>0.063***</td>
</tr>
<tr>
<td></td>
<td>(9.80)</td>
<td>(4.12)</td>
<td>(0.16)</td>
<td>(7.21)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.178***</td>
<td>2.519***</td>
<td>-0.415***</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>(-3.05)</td>
<td>(13.37)</td>
<td>(-2.95)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>Observations</td>
<td>263,217</td>
<td>28,703</td>
<td>49,264</td>
<td>312,487</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.892</td>
<td>0.647</td>
<td>0.737</td>
<td>0.886</td>
</tr>
<tr>
<td>Estimation method</td>
<td>Pooled</td>
<td>Pooled</td>
<td>Pooled</td>
<td>Pooled</td>
</tr>
</tbody>
</table>

### Panel C: Predicting Signed Order Flow

<table>
<thead>
<tr>
<th></th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.767*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.95)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.015*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-4.07)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.545**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.16)</td>
</tr>
<tr>
<td>Observations</td>
<td>264,242</td>
<td>28,736</td>
<td>49,318</td>
<td>313,711</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.121</td>
<td>0.166</td>
<td>0.011</td>
<td>0.122</td>
</tr>
<tr>
<td>Estimation method</td>
<td>Pooled</td>
<td>Pooled</td>
<td>Pooled</td>
<td>Pooled</td>
</tr>
</tbody>
</table>

Electronic copy available at: https://ssrn.com/abstract=3466417
Table 3: ESG and Valuation. We regress each firm’s valuation ratio (the logarithm of price to book) on the contemporaneous ESG score, controlling for the market beta. The ESG proxies are as in Table 1. Robust t-statistics are in parentheses, clustered at the stock level in these pooled regressions.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ln(Price-to-Book)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accruals (negated)</td>
<td>-0.470***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-11.59)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSCI ESG</td>
<td></td>
<td>0.058***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(8.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO2 (negated)</td>
<td></td>
<td></td>
<td>0.0001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(7.25)</td>
<td></td>
</tr>
<tr>
<td>Non-sin stock</td>
<td></td>
<td></td>
<td></td>
<td>-0.060</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.78)</td>
</tr>
<tr>
<td>Beta</td>
<td>0.338***</td>
<td>-0.348***</td>
<td>-0.449***</td>
<td>0.402***</td>
</tr>
<tr>
<td></td>
<td>(21.13)</td>
<td>(-8.56)</td>
<td>(-16.39)</td>
<td>(28.49)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.514***</td>
<td>1.245***</td>
<td>1.391***</td>
<td>0.444***</td>
</tr>
<tr>
<td></td>
<td>(27.37)</td>
<td>(21.81)</td>
<td>(38.32)</td>
<td>(5.65)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,708,222</td>
<td>203,502</td>
<td>427,857</td>
<td>2,120,679</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.077</td>
<td>0.046</td>
<td>0.050</td>
<td>0.073</td>
</tr>
<tr>
<td>Estimation method</td>
<td>Pooled</td>
<td>Pooled</td>
<td>Pooled</td>
<td>Pooled</td>
</tr>
</tbody>
</table>
Table 4: Does ESG Predict Returns? This table reports the performance of high-ESG minus low-ESG portfolios. Specifically, each month, stocks are sorted into portfolios based on quintiles of their ESG score proxies, and we then compute the return over the following month of the quintile with the best ESG scores minus that with the lowest scores. Stocks are equal weighted in Panel A and value weighted in Panel B. The ESG proxies are as in Table 1. We report the portfolios’ excess return, one-factor CAPM alpha, three-factor alpha that also controls for the Fama–French (FF) factors related to size and value, five-factor alpha that further controls for the FF factors related to profitability and investment, and six-factor alpha that also controls for momentum. *-Statistics are reported in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Accruals (negated)</th>
<th>MSCI ESG</th>
<th>CO2 (negated)</th>
<th>Sin stock (negated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess return</td>
<td>0.08***</td>
<td>0.00</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(4.41)</td>
<td>(0.28)</td>
<td>(1.59)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>CAPM alpha</td>
<td>0.08***</td>
<td>0.01</td>
<td>0.07**</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(4.39)</td>
<td>(1.00)</td>
<td>(2.09)</td>
<td>(-0.66)</td>
</tr>
<tr>
<td>Three-factor (FF) alpha</td>
<td>0.07***</td>
<td>0.01</td>
<td>0.05</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(4.03)</td>
<td>(0.60)</td>
<td>(1.63)</td>
<td>(-0.20)</td>
</tr>
<tr>
<td>Five-factor (FF) alpha</td>
<td>0.09***</td>
<td>0.00</td>
<td>0.06*</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(4.91)</td>
<td>(0.22)</td>
<td>(1.92)</td>
<td>(0.65)</td>
</tr>
<tr>
<td>Six-factor (FF+Mom) alpha</td>
<td>0.09***</td>
<td>0.00</td>
<td>0.05*</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(4.76)</td>
<td>(0.22)</td>
<td>(1.73)</td>
<td>(0.27)</td>
</tr>
</tbody>
</table>

Panel B: Value-weighted returns

<table>
<thead>
<tr>
<th></th>
<th>Accruals (negated)</th>
<th>MSCI ESG</th>
<th>CO2 (negated)</th>
<th>Sin stock (negated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess return</td>
<td>0.03**</td>
<td>0.00</td>
<td>0.05*</td>
<td>-0.03*</td>
</tr>
<tr>
<td></td>
<td>(2.30)</td>
<td>(0.01)</td>
<td>(1.89)</td>
<td>(-1.92)</td>
</tr>
<tr>
<td>CAPM alpha</td>
<td>0.04***</td>
<td>0.01</td>
<td>0.04</td>
<td>-0.05***</td>
</tr>
<tr>
<td></td>
<td>(3.12)</td>
<td>(0.70)</td>
<td>(1.52)</td>
<td>(-2.95)</td>
</tr>
<tr>
<td>Three-factor (FF) alpha</td>
<td>0.03***</td>
<td>0.01</td>
<td>0.03</td>
<td>-0.05***</td>
</tr>
<tr>
<td></td>
<td>(2.64)</td>
<td>(0.45)</td>
<td>(1.14)</td>
<td>(-2.85)</td>
</tr>
<tr>
<td>Five-factor (FF) alpha</td>
<td>0.03***</td>
<td>-0.01</td>
<td>0.05*</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(2.76)</td>
<td>(-0.31)</td>
<td>(1.85)</td>
<td>(-0.47)</td>
</tr>
<tr>
<td>Six-factor (FF+Mom) alpha</td>
<td>0.03**</td>
<td>-0.01</td>
<td>0.04*</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(2.52)</td>
<td>(-0.32)</td>
<td>(1.72)</td>
<td>(-0.63)</td>
</tr>
</tbody>
</table>