8-2021

ESG and Expected Returns on Equities: The Case of Environmental Ratings

Christopher C. Geczy
geczy@wharton.upenn.edu

John B. Guerard Jr.

ESG and Expected Returns on Equities: The Case of Environmental Ratings

Abstract

Using long-standing models for expected returns of US equities, we show that firm environmental ratings interact with those forecasted returns and produce excess returns both unconditionally and conditionally. Well-known factor models subsume neither environmental-related return differentials nor expected return premia from those scores and models. In addition, combining information from both inputs—expected return models and economic, social, and governance (ESG) information—may provide an advantage in selecting investments. For financial fiduciaries, this notion shifts the conversation about ESG reflecting only constraints to one of an expanded information and possibly investment opportunity set.

Keywords

ESG, sustainability, expected returns, factor models, earnings forecasts, fiduciary responsibility

Disciplines

Economics

ESG and Expected Returns on Equities: The Case of Environmental Ratings

Christopher C. Geczy and John B. Guerard Jr.

August 2021

PRC WP2021-15
Pension Research Council Working Paper
Pension Research Council
The Wharton School, University of Pennsylvania
3620 Locust Walk, 3302 SH-DH
Philadelphia, PA 19104-6302
Tel.: 215.573.3414 Fax: 215.573.3418
Email: prc@wharton.upenn.edu
http://www.pensionresearchcouncil.org

The authors thank Alimu Abudu, Peter Cachion, Nancy Gao, Troy Wang, and Evan Xu for computational and outstanding research assistance. We also thank the MSCI KLD and Wharton WRDS. All findings, interpretations, and conclusions of this paper represent the views of the authors and not those of the Wharton School or the Pension Research Council.

ESG and Expected Returns on Equities:
The Case of Environmental Ratings

Christopher C. Geczy and John B. Guerard Jr.

Abstract

Using long-standing models for expected returns of US equities, we show that firm environmental ratings interact with those forecasted returns and produce excess returns both unconditionally and conditionally. Well-known factor models subsume neither environmental-related return differentials nor expected return premia from those scores and models. In addition, combining information from both inputs—expected return models and economic, social, and governance (ESG) information—may provide an advantage in selecting investments. For financial fiduciaries, this notion shifts the conversation about ESG reflecting only constraints to one of an expanded information and possibly investment opportunity set.

Keywords: ESG, sustainability, expected returns, factor models, earnings forecasts, fiduciary responsibility
JEL Codes: G11, G12, G17, Q56, C58

Christopher C. Geczy
The Wharton School, University of Pennsylvania
2258 SHDH, 3620 Locust Walk
Philadelphia, PA 19104
geczy@wharton.upenn.edu

John B. Guerard Jr.
Scientific Advisory Board
McKinley Capital Management, LLC
3800 Centerpoint Drive, Suite 1100
Anchorage, AK
and Affiliate Instructor, Computational Finance and Risk Management Program
University of Washington
Seattle, WA
jguerard@mckinleycapital.com

Today, environmental concerns dominate ESG criteria cited by investors as influencing portfolio decisions, measured both by numbers of investors and by the total amount of assets subject to environmental criteria. Until recently, other ESG criteria were dominant. The shift reflects a change in preferences or at least a heightened perception about the importance of climate change and related issues facing the environment, which might have an anthropogenic component.

For example, the US Social Investment Foundation *2020 Trends Report* (US SIF 2020) indicated that 'Environmental Considerations' was the leading ESG criterion by assets for money managers in 2020 with $13.45 trillion out of approximately $17 trillion aggregated across all investment vehicles, including separate accounts and undisclosed vehicles (US SIF 2020:21, Figure 2.4). The leading individual criterion is related to climate change. In addition, of the top 14 criteria listed, five fall in some way under the 'E' environmental umbrella. In contrast, in the 2007 Report (US SIF 2007), environmental issues were ranked sixth, preceded by issues related to tobacco, Sudan, the MacBride Principles, human rights, community relations, and alcohol production, distribution, and sales (US SIF 2007:17, Figure 3.4).

The challenge faced by pension fiduciaries is honoring the long-standing principle that their legal and ethical duties must focus on the financial betterment of beneficiaries, rather than on any other (perhaps private) benefit including sustainability, those related to the common good or the environment, or those related to social goals, if for any reason such consideration results in tradeoffs against risk-adjusted returns. For plans governed by ERISA, 1998 guidance from Robert Doyle, then Director of the Office of Regulations and Interpretations of the United States Department of Labor (US DOL), set a requirement of a side-by-side comparison of risk-adjusted returns consistent with the Sharpe ratio, whenever socially responsible investments are considered for a plan:

> In discharging investment duties…fiduciaries must, among other things, consider the role of the particular investment [in the] investment portfolio. Because every investment necessarily causes a plan to forgo other investment opportunities fiduciaries also must consider expected return on alternative investments with similar risks available to the plan…If [those] requirements are met, the selection of a "socially responsible" mutual fund as either a plan investment or a designated investment alternative…would not, in itself, be inconsistent with … fiduciary standards (Doyle 1998:2).

One of the challenges faced by those overseeing ERISA plans has been perceived changes in guidance from the US DOL. For instance, reversing a stance articulated during the previous administration, Obama Administration Labor Secretary Thomas Perez said in October 2015: 'The question is this: Can an ERISA plan invest in projects or companies that serve the common good, while still keeping at the forefront the fiduciary principle of investing prudently and for the exclusive benefit of retirees and workers? I believe we can'. He also said that the '2008 [Bush Administration] guidance gave cooties to impact investing' (Perez np 2015)

In turn, more recently, the Trump DOL articulated yet another shift in tone in a 2020 proposed rule:

> As ESG investing has increased, it has engendered important and substantial questions and inconsistencies, with numerous observers identifying a lack of precision and consistency in the ESG investment marketplace. There is no consensus about what constitutes a genuine ESG investment, and ESG rating systems are often vague and inconsistent, despite featuring prominently in marketing efforts (DOL Proposed Rule, 2020:8).

This message raises further concerns about the fiduciary setting in which ESG criteria are considered, either via positive or negative screening, activism, engagement activities, or in other ways.

We address the important question about whether environmental scores widely referenced and utilized contain information directly related to expected returns and long-standing models for their forecasts that have survived multiple-comparison tests and out-of-sample tests alike. The results bridge concerns trustees would naturally have when making the required side-by-side comparisons of investments or portfolios selected, so as to have certain ESG characteristics with those that do not. Specifically, they meet the requirement of no decline in expected returns holding risk constant, as set out by Robert Doyle in 1998, and attendant to the basic notion of financial fiduciary duty also in non-ERISA settings.

Recent surveys also indicate that asset managers and investors may reference ESG characteristics in ways that defy the traditional 'constrained opportunity set' interpretation of the incorporation of ESG characteristics in investment decisions. For instance, in a recent assessment of the reasons (113) institutional investors reported considering ESG factors (Table 1 below), the most cited reason was Risk (84%), followed by Client Demand (81%), Social or Environmental Impact (80%), Returns (73%), Mission (70%), Fiduciary Duty (64%), UN Sustainable Development Goals (46%), and Regulatory Compliance (31%) (US SIF 2020:28, Figure 2.13). One interpretation of this ordering is that ESG criteria contain information important for investment selection, apart from the typical non-pecuniary or purpose-related reasons mentioned for screening or portfolio tilts toward 'good actor' firms in equity portfolios. When information on public investments is not limited to formal filings or releases governed, say, by Generally Accepted Accounting Principles (GAAP), other sources of information that would otherwise be difficult or expensive to collect and assess including proprietary analysis-based ESG scores may be valuable in assessing the cross-section of public firms.

Table 1 here

The proprietary analysis-based ESG scoring used in this study is MSCI ESG KLD STATS, a 1991-2017 database of firm ESG ratings, today subsumed in MSCI ESG Ratings. The inception of the ratings system followed with the launch of the Domini 400 Social Index (today the MSCI KLD 400 Social Index), to rate companies whose stocks were in the index. We show that firm

ESG characteristics computed via normalized MSCI KLD environmental scores interact with forecasted expected returns of US equities estimated from long-standing models for expected return first articulated by Guerard and Stone (1992), Bloch et al. (1993), Guerard, Gultekin, and Stone (1997), Guerard (1997a and 1997b) and further developed by Guerard, Markowitz, and Xu (2014, 2015). We focus on environmental characteristics because, especially for public firms, the potential cost of achieving high ratings in this category may be high, and because climate change is a major issue among SRI / ESG investors. The importance of using long-established models of expected return must be underscored. As Markowitz and Xu (1994), Lo and MacKinlay (1987), and Harvey, Lin, and Zhu (2016) have pointed out, data mining biases in the absence of multiple comparison test controls can lead to poor out-of-sample results. By relying on long-standing models developed before the KLD data rose to prominence, we may avoid some of the biases inherent in typical analysis.

Specifically, we find that firms with high ESG (environmental) scores have excess returns over those with low scores unconditionally, and also conditional on expected return from models above with 'bagged' value, earnings, and momentum components articulated in the early 1990s. In addition, a battery of now-traditional risk factor models including the CAPM, the Fama-French (1992) three-factor model, the Carhart (1997) extension, and a five-factor model that augments the Carhart model with the Fama-French quality factor (Fama and French, 2015) subsume neither environmental score-related return differentials nor expected return premia from the long-standing models. For pension trustees, or for consultants and managers they hire, combining information from both inputs (expected return models and ESG criteria) might provide advantages in constructing equity portfolios. For those fiduciaries whose concerns center on risk and return considerations alone when selecting investments, our results suggest that incorporating non-GAAP information via earnings, price momentum, and ESG characteristics along with a collection of weighted value measures may collectively and individually add value rather than serve to induce the cost of a constraint on the investment universe.

A Brief SRI / ESG Environmental Screen Literature Review

The empirical evidence is mixed on whether SRI / ESG portfolios incorporating constraints related to positive or negative screens induces a cost in investment performance or whether it is associated with additional gains.[1] The *Journal of Investing* has been an active SRI / ESG outlet for more than 20 years, starting with Luck and Pilotte (1993), Kurtz and DiBartolomeo (1996), and many of the Moskowitz Prize winners for research in socially responsible investing, including the first winner, Guerard (1997a). In this analysis, we apply the earnings forecasting model used in Guerard (1997a, 1997b) and the subsequent larger composite models of expected returns, Guerard, Markowitz, and Xu (2014, 2015), to show that incorporating ESG Environmental (ENV) criteria may potentially enhance stockholder returns. Specifically, we find that in certain implementations, incorporating the KLD environmental criteria enhances portfolio returns.

The Moskowitz Prize, awarded annually since 1996 for research in socially responsible investing, has recognized the environmental research analyses of Russo and Fouts (1997), Dowell, Hart, and Yeung (2000), and Naaraayanan, Sachdeva, and Sharma (2020) among its winning studies. Russo and Fouts (1997) used the Franklin Research & Development Corporation (FSRDC) environmental ranking. These authors report that in 1991 and 1992, in a 243-firm regression model using the return on assets (ROA) as the dependent variable, that ROA was positively and statistically associated with firm growth, industry growth, firm size, advertising intensity, and the FSRDC environmental ranking.

Dowell, Hart, and Yeung (DHY, 2000) start their analysis with a universe of the S&P 500 Companies, operating in countries with per capita income below $8,000 (in 1985 USD, relatively lower-income countries) during 1994-1997. They restrict their modeling to manufacturing firms and use the Investor Responsible Research Center (IRRC) environmental rating. The resulting universe is 89 firms. The dependent variable is Tobin's Q, measuring the firm Market Value of equity relative to the replacement costs of tangible assets, defined as book value of inventory plus the net value of physical plant and equipment. DHY (2000) study three ENV standards:

ENV1 = Local ENV standard,

ENV2 = US ENV standard, and

ENV3 = Stringent ENV standard.

DHY (2000) report that 72 of the 89 firms never changed ENV strategies; 16 changed once; and one changed twice. Of these changes, 12 were positive and six negative changes.

The DHY regressions show that Tobin's Q is positively and statistically associated with research and development expenditures, advertising intensity, and the IRRC environmental ranking. The smallest coefficient for ED2 (Table 3, regression 3-d) indicates that firms adopting their own stringent global environmental standards have a Tobin's Q that is higher than those using US standards abroad.

Naaraayanan, Sachdeva, and Sharma (NSS, 2020) study the New York City Pension System (NYCPS) Board Accountability Project (BAP, announced in November 2014) to hold boards accountable to long-term shareholders and give pensioners a voice concerning board diversity, climate change risks, and employee treatment. In the Russell 3000 stock universe during the 2000-2013 time period, 62 of the 181 BAPs were environmentally based. NSS (2020) use the

Thomson-Reuters (Asset4) ENV score. The reported regression results indicated that the return on assets was positively and statistically associated with the Fossil Fuel index return, form size, the market-to-book value ratio, and profitability. The authors did not find a statistically significant coefficient on the Thomson-Reuters environmental rating. Overall, the authors report that targeted BAP firms effectively reduced real EPA-measured toxic releases by a statistically significant amount. The BAP-targeted firms reduced their Toxic Release Inventory (TRI) and the Greenhouse Gas Reporting Program (GGRP) levels by up to 50 percent.

Environmental scores

Before the development of the KLD dataset in the early 1990s, which went on to become an industry standard, a large volume of ESG-related research focused on various ways to estimate the effects of ESG on company performance (for examples, see Gordon and Buchholz 1978; Aupperle, Carrol, and Hatfield 1985; and Rosen, Sandler, and Shani 1991). To our knowledge, the first academic papers to validate and link KLD data with firm characteristics were by Ruf, Muralidhar, and Paul (1993) and Graves and Waddock (1994).

KLD ratings started in 1991, covered about 650 companies, and they were based on a -2 to +2 ratings system in nine categories, including negative screens. The current study utilizes the 2017 version of the database, in which rankings start December 1991, end December 2017, and contain binary values of 0 and 1 for strengths and weaknesses in seven categories and six controversy scores for more than 3,000 companies.

Over time, the KLD database has been enhanced resulting from acquisitions and other methodology changes. For example, in 2000, the human rights category was added (Galema, Plantinga, and Scholtens 2008); in 2002 governance was added (Statman and Glushkov 2009); and in 2010 KLD decided to rank companies only on issues relevant for their industry instead of all issues.

From the early KLD studies (Sharfman 1996) and continuing to Statman and Glushkov (2016), among others, there has been an ongoing discussion about the challenges of creating a unique overall KLD-based score. The simplest way that sums all strengths and subtracts all weaknesses incurs its own set of biases and imbalances driven by data structure rather than companies' ESG attributes. Dorfleitner et al. (2014) study the relation between ESG score performance and stock performance in various markets worldwide, reiterating global evidence of positive association between firm ESG ratings and subsequent returns; however, the bias remains. The earlier literature attempted to address the implicit bias arising from weighting each issue equally. For example, in order to avoid treating each ESG strength and weakness as equally important, Waddock and Graves (1997) rely on the issues weighting scheme developed by Ruf, Muralidhar, and Paul (1993). Because such weightings are highly subjective, they are no longer used in the more recent studies. For example, Employee Relations strengths are evaluated on ten individual variables, with a maximum score of 8, while Human Rights strengths are evaluated on three variables with a maximum score of 2. Hence, because of the uneven ranges, the raw score will be much more affected by the Employee Relations strengths vs. Human Rights strengths. The same issues affect the weights of strengths vs. concerns. An area having a larger number of evaluated metrics will receive a higher implied weighting in the overall raw calculations.

Another challenge arises because of the changing coverage of the KLD dataset over time. Specifically, as the number of strengths and weaknesses changes in each category, summing the raw strengths and weakness, as was the earlier practice, creates score dynamics that are influenced by the dataset construction, rather than by the company's changing ESG policies. Kempf and

Osthoff (2007) address this problem by normalizing the net scores within each of the six categories. In addition, they introduce a way to transform the weakness measure into the same direction as the strengths.

Manescu (2011) adds an additional refinement that normalizes strengths and weaknesses separately because the number of each sub-strength and sub-weakness are different and also vary across time.

Modifying the KLD Environmental Score Data

Inside each of (now) seven subcategories (Governance, Community, Diversity, Employee Relations, Environment, Human Rights and Product Safety), KLD provides binary ratings on multiple individual measures of strengths and concerns criteria. For each of the seven categories and for each company in each year, the Category Raw Net Score is the sum of category strengths minus sum of category weaknesses. The Total Raw Net Score is the sum of the strengths across all categories minus the sum of all the weaknesses across all categories. There is a Total Net Score only if both strength and weakness exist. If strengths or weaknesses are missing entirely, the Net Score is missing for that company in that year and is not included in calculations.

To avoid the challenges of combining sub-score ratings across ratings subcategories, identified in the literature, and to focus simply on the Environment subcategory, we do not aggregate across subcategories, and we focus on yearly constitution of the rankings. Specifically, we separately sum the number of Environmental strengths and weaknesses a given company has in a given year, rejecting zeroes as non-covered, and then consider the simple difference between the two by firm and by year. We then compute each firm's Environmental net score (ENV) as the difference between the number of Environmental strengths and weaknesses it as (again, without firms that have zeroes). Finally, portfolio formation stratifies firms by their ENV score into quintiles yearly denoting Lo ENV and Hi ENV as the bottom and top quintiles firms.

By focusing on the Environment subcategory alone, we avoid the issue of combining different possible numbers of strength and weakness indicators in the different subcategories. By computing strengths and weaknesses summations separately, and rejecting zeroes which indicate non-coverage, and only then computing a total ENV score (the number of strengths minus the number of weaknesses), we avoid distortion induced by non-coverage. Finally, by focusing on the yearly cross-section of firms and defining Lo and Hi ENV firms as the bottom and top quintiles, we effectively dynamically adjust for changes in the underlying structure of the KLD data in a manner that takes the position of traders operating on information known to them at the time of portfolio formation.

Composite Models for Expected Returns Modeling and Stock Selection

In a number of composite models for expected returns, we utilize modifications of the expected return models outlined by Bloch et al. (1993). These models synthesized cross-sectional relationships between (and among) documented anomalies. Graham and Dodd (1934), Williams (1938), Graham, Dodd, and Cottle (1962), Elton and Gruber (1972a, 1972b), Latane, Tuttle and Jones (1975), Jacobs and Levy (1988), Dimson (1988), and Levy (1999) tested and reported known anomalies, including the low PE or high earnings-to-price (EP), high book value-to-price, high cash flow-to-price, high sales-to-price, net current asset value, and the earnings forecasting models in.[2] In addition, the model synthesizes small size, earnings forecasts, revisions, recommendations, breadth, earnings surprises, and dividend yield variables identified (see Banz (1981), Dimson (1988), Jacobs and Levy (1988), and Ziemba and Schwartz (1993) as anomalies).

The resulting model from Bloch et al. (1993) is referenced below as equation (1) relating total realized returns, TR, to eight selected variables. We refer to this model as the composite model, REG8:

$$TR = w_0 + w_1 EP + w_2 BP + w_3 CP + w_4 SP + w_5 REP + w_6 RBP + w_7 RCP + w_8 RSP + e_t \quad (1)$$

where: EP = [earnings per share]/[price per share] = earnings-price ratio;

BP = [book value per share]/[price per share] = book-price ratio;

CP = [cash flow per share]/[price per share] = cash flow-price ratio;

SP = [net sales per share]/[price per share] = sales-price ratio;

REP = [current EP ratio]/[average EP ratio over the past five years];

RBP = [current BP ratio]/[average BP ratio over the past five years];

RCP = [current CP ratio]/[average CP ratio over the past five years]; and

RSP = [current SP ratio]/[average SP ratio over the past five years].

Given concerns about both outlier distortion and multicollinearity, Bloch et al. (1993) test the relative explanatory and predictive merits of alternative regression estimation procedures and find that controlling for both outliers and multicollinearity via robust regressions is important. Second, Bloch et al. (1993) quantify the survivor bias (including dead companies in the database) and find that it was not statistically significant in either Japan or the US for the period tested. Third, they investigate period-to-period portfolio revision and find that tighter turnover and rebalancing triggers led to higher portfolio returns for value-based strategies. Finally, Markowitz and Xu (1994) develop a test for data mining.[3] In addition to testing the hypothesis of data mining, the test can also be used to estimate and assess the expected differences between the best test model and the average of simulated policies.

Studies of the effectiveness of corporate earnings forecasting variables[4] reported in Cragg and Malkiel (1968), Elton, Gruber, and Gultekin (1981), DeBondt and Thaler (1989), Wheeler (1994), and Guerard and Stone (1992) were reprinted in Bruce and Epstein (1994).[5] Analysts' forecasts of earnings per share (EPS), EPS revision, and the direction of EPS forecast revisions were incorporated into the Institutional Broker Estimation Services (I/B/E/S) in-print database in July 1972. The I/B/E/S database has computer-readable data from January 1976, domestically, and January 1987 for non-US.[6] We refer the reader to Brown (1998) which contains about 570 abstracts of I/B/E/S studies.

Guerard, Gultekin, and Stone (1997) report that analysts' forecast variables enhanced portfolio returns over the long run. CTEF, a composite model of earnings consensus forecasts, revisions, and breadth, the agreement among analysts' revisions (all from I/B/E/S), was highly statistically significantly correlated with stock returns. Guerard and Mark (2003) reported that CTEF, and a nine-factor model, denoted REG9 composed of REG8 plus CTEF, was highly (statistically) significantly correlated with stock returns.

Guerard, Xu, and Gultekin (2012) and Guerard, Markowitz, and Xu (2014) added price momentum (PM), price at *t-1* divided by the price seven months ago, *t-7*, which we refer to as 7/1 momentum. This is different from but correlated with the PR1YR momentum definition using prior returns measured from *t-1* to *t-12* to classify momentum. They denoted the ten-factor stock selection model as United States Expected Returns (USER). They reported, among other results, that: (1) the EP variable had a larger average weight than the BP variable; (2) the relative PE, denoted RPE, the EP relative to its 60-month average had a higher average weight than the PE variable; and (3) the composite earnings forecast variable, CTEF, had a larger weight than the RPE variable. In fact, in the USER model, only the price momentum variable, PM, had a higher weight than the CTEF variable (and only by 1%, at that).[7]

In what follows, we employ the USER Model shown in equation (2), augmenting REG8:

$$TR_{t+1} = a_0 + a_1 EP_t + a_2 BP_t + a_3 CP_t + a_4 SP_t + a_5 REP_t + a_6 RBP_t + a_7 RCP_t + a_8 RSP_t + a_9 CTEF_t + a_{10} PM_t + e_t \quad (2)$$

where: EP = [earnings per share]/[price per share] = earnings-price ratio;

BP = [book value per share]/[price per share] = book-price ratio;

CP = [cash flow per share]/[price per share] = cash flow-price ratio;

SP = [net sales per share]/[price per share] = sales-price ratio;

REP = [current EP ratio]/[average EP ratio over the past five years];

RBP = [current BP ratio]/[average BP ratio over the past five years];

RCP = [current CP ratio]/[average CP ratio over the past five years];

RSP = [current SP ratio]/[average SP ratio over the past five years];

CTEF = consensus earnings-per-share I/B/E/S forecast (FEP), revisions and breadth (BR),

PM = Price Momentum; and

e = randomly distributed error term.

The Guerard, Markowitz, and Xu (2014) USER model test substantiated the Bloch et al. (1993) approach. In addition to the fundamental ten-factor **USER** model, we isolate the following subset models to isolate particular effects in the models: **EWC** sets a1 through a10 = 10 percent, allowing tests out-of-sample optimization of USER weights. **EVALUE** sets a5 through a10 equal to zero and a1 through a4 equal to 25 percent, producing, in effect, a ‘bagged’ value model that naively blends traditionally estimated valuation ratios. **MQ** sets a1 through a8 = 0, isolating the CTEF earnings variable and 7/1 price momentum which are equally weighted.

Model Estimation

For each security, we use monthly total stock returns and prices from CRSP files, earnings book value cash flow, net sales from quarterly COMPUSTAT files, and consensus earnings-per-share, forecast revisions, and breadth from I/B/E/S files. We construct the variables used in (3) for each month starting in January 1990. The USER model is estimated using weighted least squared latent root regression analysis of Bloch et al. (1993) to control for multicollinearity among signal regressors and to address outliers (analysis over the 60-month (five-year) moving window for each period to identify variables statistically significant at the 10 percent level). The model uses the normalized coefficients as weights over the past 12 months with Beaton-Tukey bisquare outlier adjustment. We use the statistically significant coefficients to estimate the next month's expected return rank, E_i, for each security. The USER estimation conditions are virtually identical to those described in Guerard, Gultekin, and Xu (2012) and Guerard, Markowitz, and Xu (2014, 2015).[8]

Empirical Results

Table 2 presents summary statistics for our sample, and Figures 1 and 2 plot the numbers of firms having the requisite input over time for the calculation of the overarching USER model value and for the earnings subcomponent, CTEF. The number of firms having full input variable values is closely followed by the number of firms having the earnings variable. The former ranges from about 1,000 firms in 1976, just a few years after the NASDAQ exchange went online joining NYSE and AMEX exchanges in the data, to about 1,500 firms in 1991, the beginning of the KLD data. The number of firms having CTEF information peaks in 1989 at approximately 2,750, while the number of firms having complete data for USER peaks at about 2,550. Both decline to just under 2,000 firms at the end of the sample period in December 2017. Figure 2 tracks the number of firms in the tails of the distribution defined in the Low and High Environmental ratings equal weighted portfolios. A large spike is seen in 2013 when the number of firms covered by KLD was expanded to a large number of firms in the left tail of the environmental score distribution.

Table 2 here

Figures 1 and 2 here

Tables 3a and 3b report the baseline performance of five models of expected return via three encompassing risk factor models from the academic literature estimated over the period March 1991 through December 2017. The five expected return models are subsumed in the USER model described above: USER, EWC, EValue, MQ and CTEF. The EWC model naively equal weights inputs while the USER model optimizes the weights (with most of the period of the estimations being out of sample with respect to the optimized weights). The EValue model incorporates an equal weighting of the bagged, scaled price ratios articulated above, while the MQ model isolates the remaining variables (non-GAAP) formed using consensus earnings, earnings breadth, and earnings depth.

Tables 3a and 3b here

In Table 3a, quintile sorts on the encompassing USER model produce annualized alpha ranges from 3.2 percent to -0.1 percent in a one-factor CAPM using RMRF as the market measure. We see a U-shape pattern in one-factor market betas and inverted U-shapes in Adjusted R2s. The Q5-Q1 portfolio delivers an alpha of 3.3 percent with nearly zero beta and adjusted R2. In the four-factor model of Carhart (1997), which embeds the Fama French three-factor model, we see a similar pattern, except that across the CTEF quintiles, the momentum exposures are quite strong. Starting at a value of -0.55 and declining to -0.04 from the fifth to first quintile, momentum and CTEF sorts are quite highly correlated. The Q5-Q1spread returns a strong negative loading on Carhart (1997) momentum but nothing on the remaining factor-mimicking portfolios.

In the five-factor model that adds quality, we see a similar pattern: USER quintiles produce strong patterns in momentum-mimicking portfolio loadings, while the Q5-Q1 spread produces a strong negative loading on momentum of -0.51 while the four-factor alpha is estimated to be 3.2 percent. As we will demonstrate below, the momentum loading derives from the 7/1 price momentum variable within the model, and is in that sense an expected result. However, the four-factor alpha remains statistically significant and economically meaningful, indicating that USER encompasses information independent of the extended factor model. Moreover, the EWC and EValue results suggest that the contributions inherent in the definition of the USER model also add value. Because this model and its weightings were developed in advance of this time period, we are less worried about data mining than this finding would otherwise suggest.

MQ, which is a 50/50 combination of CTEF and price momentum, produces high alphas, but only for one-factor CAPM (6.2%). When momentum is accounted for in the four-factor model, the spread alpha is 90 bps and the momentum loading is a large -0.80 for the Q5-Q1 spread. When quality is introduced in the five-factor model, we see that momentum loadings are robust to the additional variable while the factor loading pattern is quite strong for both momentum and quality, but especially momentum. The five-factor alpha remains at 1.8 percent. In all of the one-, four- and five-factor models, we see the usual strong spread alpha (4.6%) which shrinks significantly when momentum and other zero-investment portfolio returns are included in the model where momentum subsumes the 7/1 momentum effect of the expected returns model. Finally, corresponding to the patterns above, the earnings composite variable CTEF survives one-, four- and five-factor regressions with positive and significant pricing error (alpha).

The Interaction of Environmental Scores and Expected Return Models

Tables 4a and 4b report one-, three-, and five-factor model time series regressions in which, for each of the expected return models above (USER, EWC, EValue, MQ, and CTEF), returns from portfolios constructed based on their ENV and model numerical values are calculated. High and low ENV and model groups are defined yearly by the 30/40/30 criteria (see Fama and French (1992) and Carhart (1997)). Each year, the 30 percent of firms with the highest (low) normalized environmental scores are include in the high (low) ENV group. Independently firms are included in the high (low) USER groups yearly based on their raw USER scores. Firms that are in both high environmental score and high USER score groups are characterized as High ENV + High USER and so on. Firms are equal weighted within groups.

Tables 4a and 4b here

The results in Table 4a generally demonstrate that High ENV have excess returns (alpha) holding USER constant, and that firms that have high expected returns via the models along with high environmental scores produce the greatest pricing errors (alphas) in the various factor model estimations. For example, in the CAPM (RMRF) regression in Table 4a, the High ENV + High USER portfolio produces an annualized intercept (alpha) of about 3.6 percent while the Low ENV + Low USER portfolio produces a historical alpha of 0.1 percent. The (High+High) - (Low+Low) spread produces a nearly zero-beta return (alpha) of about 3.5 percent. Moreover, when isolating the USER effect by going long and short along the USER dimension and examining the resulting ENV differentials, High ENV firms produce an 80 bps excess returns while the Low ENV score portfolio produces an excess return of -0.40 percent, yielding an alpha of approximately 1.20 percent.

This results for the CAPM are robust across the additional factor models where interesting patterns in factor loadings emerge, indicating the interaction between ESG characteristics and traditional factor exposures. For example, while patterns in market betas remain intact in moving from the CAPM to the Carhart (1997) four-factor model, small size effects emerge and in particular value (HML) factor loadings are significantly larger for High ENV + High USER portfolios than for Low ENV + Low USER portfolios. Specifically, the Fama-French HML (value less growth) factor loading for the (High+High) - (Low+Low) spread is -0.35, the momentum spread loading estimate is 0.18, and the intercept (pricing error or alpha) is an annualized 3.5 percent. In other words, perhaps as expected, firms that have low environmental scores seem to have a more pronounced value exposure than those with high scores, corresponding to received wisdom that ESG stock portfolios are generally tilted toward growth and away from the asset-heavy traditional value sectors and firms. These results are borne out by the five-factor model extending the Carhart (1997) model with the Fama-French Quality spread. Interestingly, quality subsumes the momentum loading and to a lesser extent value.

Part of the story behind the negative quality loading (-0.37 at the point estimate) in the regression is clear from the differences between the loadings on the Quality-mimicking portfolio spread, first holding ENV constant and then holding USER constant. For instance, the differential between High USER and Low USER loadings for High ENV is -0.14 (= 0.22 - 0.36), and for Low ENV, it is -0.15. In other words, USER correlates negatively with Quality, a fact that is known from previous literature which identifies USER as measuring expected profitability rather than realized profitability, which on average mean-reverts relative to expectations. Holding USER constant, for High USER, the spread in the quality loading between High ENV and low ENV is estimated as -0.22 (= 0.22 – 0.44) and for Low USER is 0.36 – 0.59 = -0.23. In other words, overall, equity portfolios load positively on the Quality factor portfolio, but Low ENV tends to load more strongly than High ENV, suggesting once again that positive environmental characteristics are negatively associated with realized measures of profitability, which is obliquely measured with respect to momentum and value, as is also known. Nonetheless, the alpha for the (High+High) - (Low+Low) spread is strong at an estimated 6.4 percent annualized value.

The evidence construed across model sub-components reinforces the story for the aggregate USER model. For instance, for the naïvely equal-weighted EWC model as well as the value-bagged model (EVALUE), the MQ model incorporating only price momentum and CTEF, and for CTEF itself, the CAPM and Carhart models load essentially the same with nearly identical intercepts. However, in the five-factor model, we see an inversion of the Quality loadings. In other words, in the unoptimized EWC model (recall that the input weightings in the USER model were optimized long ago, essentially out of sample, while the EWC model treats all inputs the same and the others break out sub-components) as well as the others, it is quite clear that Quality loadings are higher for High EWC, High EVALUE, High MQ and High CTEF versus their Low counterparts. The information ratio optimization inherent in the definition of USER seems to weight components that invert the relationship. Nonetheless, the intercepts in all five-factor regressions remain economically and statistically significant for all models including for CTEF. Taken together, the results strongly suggest an interaction between ENV and various models for expected returns, which in turn indicates that when one is creating portfolios (ESG or not), one would do well to consider both sources of information and that ESG information in this important and currently very relevant case of environmental scores may be additive in creating portfolios.

Conclusion

Using long-standing models for expected returns of US equities, we show that firm environmental ratings interact with those forecasted returns and produce excess returns both unconditionally and conditionally. Now-traditional factor models subsume neither environmental-related return differentials nor expected return premia from those scores and models. In addition, combining information from both inputs (expected return models and ESG information) may provide an advantage in selecting investments, opening the question of why. We speculate that the traditional inputs into quantitative estimate and models, namely data from accounting filings made under GAAP, are limited; and that information from earnings forecasts, their breadth, and their depth combine with ESG information to augment the information set referenced in successful strategies. For financial fiduciaries, this notion shifts the conversation about ESG reflecting only constraints to one of an expanded information and possibly investment opportunity set.

One troubling facet of the 1998 US DOL Doyle guidance is its sole emphasis on risk-adjusted return rather than on portfolio characteristics. As Geczy, Stambaugh, and Levin (2005, 2021) point out, side-by-side comparisons of ESG and non-ESG investments may suggest that there is no expected risk-adjusted return difference between them, or even that ESG investments may outperform their non-ESG counterparts (or high- vs. low-scoring investments) and yet still lead to lower Sharpe ratios at the aggregate portfolio level. The critical feature for portfolios is not only whether a given high ESG-scoring investment outperforms a low-scoring one, but whether, in focusing or tilting toward firms with positive ESG characteristics, investment diversification is lost, especially if the 'tilt' ends up being Boolean (ESG in … everything else out). It is surprising that guidance has not emerged framing this important point with more fidelity. After all, while we have shown that ESG scores can provide important information to investors about expected return, we have not shown that portfolios formed from *only* high scoring ESG firms maximize Sharpe ratios. Answering this key question in the broad cross-section of equities is fertile ground for research.

Acknowledgements

The authors thank Alimu Abudu, Peter Cachion, Jamie Doran, Nancy Gao, Troy Wang, and Evan Xu for computational and outstanding research assistance. We also thank the MSCI KLD, Wharton Research and Data Services, and especially Olivia S. Mitchell, Sarah Kate Sanders and the Pension Research Council staff and members for comments.

References

Aupperle, K., A. Carroll and J. Hatfield (1985). 'An Empirical Examination of the Relationship between Corporate Social Responsibility and Profitability.' *The Academy of Management Journal*, 28(2): 446-463.

Ball, R. and E. Ghysels (2018). 'Automated Earnings Forecasts: Beat Analysts or Combine and Conquer?' *Management Science*, 64(10): 4936-4952.

Banz, R. (1981). 'The Relationship Between Return and Market Value of Common Stocks.' *Journal of Financial Economics*, 9(1): 3-18.

Bernhard, A. (1959). *The Evaluation of Common Stocks*. New York: Simon & Schuster.

Bloch, M., Guerard, Jr., J., Markowitz, H., Todd, P., and Xu, G. (1993). 'A Comparison of Some Aspects of the U.S. and Japanese Equity Markets.' *Japan and the World Economy*, 5(1): 3-26.

Brown, L. (1998). 'Managerial Behavior and the Bias in Analysts' Earnings Forecasts.' https://papers.ssrn.com/sol3/papers.cfm?abstract_id=113508.

Bruce, B. and C. Epstein, eds. (1994). *The Handbook of Corporate Earnings Analysis: Company Performance and Stock Market Valuation*. Tokyo: Toppan.

Carhart, M. (1997). 'On Persistence in Mutual Fund Performance.' *Journal of Finance*, 52(1): 57-82.

Chan L., Y. Hamao, and J. Lakonishok (1990). 'Fundamentals and Stock Returns in Japan.' Working Paper No. 45, Center on Japanese Economy and Business, Graduate School of Business, Columbia University.

Cragg, J. and B. Malkiel (1968). 'The Consensus and Accuracy of Some Predictions of the Growth of Corporate Earnings,' *Journal of Finance*, 23(1): 67-84.

De Bondt, W. and R. Thaler (1989). 'Anomalies A Mean-Reverting Walk Down Wall Street.' *Journal of Economic Perspectives*, 3(1): 189-202.

Dimson, E. (1988). *Stock Market Anomalies*. New York: Cambridge University Press.

Dorfleitner, G., S. Utz, and M. Wimmer (2014). 'Patience Pays Off – Financial Long-Term Benefits of Sustainable Management Decisions.' Working paper. https://papers.ssrn.com/abstractid=2533957.

Dowell, G., S. Hart, and B. Yeung (2000). 'Do Corporate Environmental Standards Create or Destroy Market Value?' *Management Science*, 46(8): 1059-1074.

Doyle, R (1998). *Advisory Opinion 1998-04A*. United States Department of Labor, Pension and Welfare Benefits Administration.

Dremen, D. (1999). *Contrarian Investment Strategies: The Next Generation*. New York: Simon & Schuster.

Dremen, D. (1979). *Contrarian Investment Strategy: The Psychology of Stock Market Success*. New York: Random House.

Elton, E., M. Gruber, and M. Gultekin (1981). 'Expectations and Share Prices.' *Management Science*, 27(9): 975-1095.

Elton, E. and M. Gruber (1972a). 'Earnings Estimates and the Accuracy of Expectational Data.' *Management Science*, 18(8): B-367-B-481.

Elton, E. and M. Gruber (1972b). 'The Economic Value of the Call Option.' *Journal of Finance*, 27(4): 891-901.

Fama, E. and K. French (1992). 'The Cross-Section of Expected Stock Returns.' *Journal of Finance*, 47(2): 427-465.

Fama, E. and K. French (1995). 'Size and the Book-to-Market Factors in Earnings and Returns.' *Journal of Finance*, 50(1): 131-155.

Fama, E. and K. French (2015). 'A Five-Factor Asset Pricing Model.' *Journal of Financial Economics*, 116(1): 1-22.

Galema, R., A. Plantinga, and B. Scholtens (2008). 'The Stocks at Stake: Return and Risk in Socially Responsible Investment.' *Journal of Banking and Finance*, 32(12): 2646-2654.

Geczy, C., R. Stambaugh, and D. Levin (2021). 'Investing in Socially Responsible Mutual Funds.' *The Review of Asset Pricing Studies,* 11(2): 309-351.

Geczy, C., J. Guerard, Jr., and M. Samonov (2020). 'Warning: SRI Need Not Kill Your Sharpe and Information Ratios—Forecasting of Earnings and Efficient SRI and ESG Portfolios.' *Journal of Investing,* 29(2): 110-127.

Gordon, J. and R. Buchholz (1978). 'Corporate Social Responsibility and Stock Market Performance.' *Academy of Management Journal*, 21(3): 479-486.

Graham, B. and D. Dodd (1934). *Security Analysis: Principles and Technique*. 1st ed. New York: McGraw-Hill.

Graham, B., D. Dodd, and S. Cottle (1962). *Security Analysis: Principles and Technique*, 4th. ed. New York: McGraw-Hill.

Graves, S., and S. Waddock (1994). 'Institutional Owners and Corporate Social Performance.' *Academy of Management Journal,* 37(4): 1034-1046.

Guerard Jr., J. and B. Stone (1992). 'Composite Forecasting of Annual Corporate Earnings.' In A. Chen, ed., *Research in Finance 10*. Greenwich, CT: JAI Press, pp. 205–230.

Guerard, Jr., J., M. Gultekin, and B. Stone (1997). 'The Role of Fundamental Data and Analysts' Earnings Breadth, Forecasts, and Revisions in the Creation of Efficient Portfolios.' In A. Chen, ed., *Research in Finance 15.* Greenwich, CT: JAI Press, pp. 69-91.

Guerard, Jr., J. (1997a). 'Is There a Cost to Being Socially Responsible in Investing?' *Journal of Forecasting,* 16, 475-490.

Guerard, Jr., J. (1997b). 'Is There a Cost to Being Socially Responsible in Investing?' *Journal of Investing,* 6, 11-18.

Guerard, Jr., J. and A. Mark (2020). 'Earnings Forecasts and Revisions, Price Momentum, and Fundamental Data: Further Explorations of Financial Anomalies.' In C. Lee, ed., *Handbook of Financial Econometrics*. Singapore: World Scientific Handbook in Financial Economics, pp. 1151-1209.

Guerard, Jr., J. and A. Mark (2003). 'The Optimization of Efficient Portfolios: The Case for an R and D Quadratic Term.' In A. Chen, Ed., *Research in Finance 20*. Bingley, UK: Emerald Group Publishing Limited, pp. 213-247

Guerard, Jr., J., H. Markowitz, and G. Xu (2014). 'The Role of Effective Corporate Decisions in the Creation of Efficient Portfolios.' *IBM Journal of Research and Development*, 58(4): 6:1-6:11.

Guerard, Jr., J., H. Markowitz, and G. Xu (2015). 'Earnings Forecasting in a Global Stock Selection Model and Efficient Portfolio Construction and Management.' *International Journal of Forecasting,* 31: 550–560.

Guerard, Jr., J., M. Takano, and Y. Yamane (1993). 'The Development of Efficient Portfolios in Japan with Particular Emphasis on Sales and Earnings Forecasting.' *Annals of Operations Research*, 45: 91–108.

Guerard, Jr., J., G. Xu, and M. Gultekin (2012). 'Investing with Momentum: the Past, Present, and Future.' *Journal of Investing*, 21(1): 68-80.

Haugen, R. and N. Baker (1996). 'Commonality in the Determinants of Expected Results.' *Journal of Financial Economics,* 41(3): 401-439.

Harvey, C., Y. Lin, and H. Zhu (2016). '…and the Cross-Section of Expected Returns.' *Review of Financial Studies*, 29(1): 5–68.

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

Kempf, A. and P. Osthoff (2007). 'The Effect of Socially Responsible Investing on Portfolio Performance.' *European Financial Management,* 13(5): 908-922.

Kurtz, L. and D. DiBartolomeo (1996). 'Socially Screened Portfolios: An Attribution Analysis of Relative Performance.' *Journal of Investing*, 5(3): 35-41.

Lakonishok, J., A. Shleifer and R. Vishny (1994). 'Contrarian Investment, Extrapolation and Risk.' *Journal of Finance*, 49(5): 1541-1578.

Latané, H., D. Tuttle, and C. Jones (1975). *Security Analysis and Portfolio Management*. New York: The Ronald Press Company.

Lo, A. and A. MacKinlay (1987). 'Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test.' National Bureau of Economic Research Working Paper No. w2168. https://papers.ssrn.com/abstractid=346975

Luck, C. and N. Pilotte (1993). 'Domini Social Index Performance.' *Journal of Investing,* 2(3): 60-62.

Manescu, C. (2011) 'Stock Returns in Relation to Environmental, Social and Governance Performance: Mispricing or Compensation for Risk?' *Sustainable Development,* 19(2): 95-118.

Markowitz, H. and G. Xu (1994). 'Data Mining Corrections.' J*ournal of Portfolio Management*, 21(1): 60-69.

Moskowitz, M. (1972). 'Choosing Socially Responsible Stocks.' *Business & Society Review,* 1: 71-75.

Moskowitz, M. (1997). 'Social Investing: The Moral Foundation.' *Journal of Investing,* 6(4): 9-11.

MSCI ESG Research Inc. (2015). MSCI ESG KLD STATS: 1991-2014 Data Sets. https://www.wiso.uni-hamburg.de/bibliothek/recherche/datenbanken/unternehmensdaten/msci-methodology-2014.pdf.

Naaraayanan, L., S. Sachdeva, and V. Sharma (2020). 'Real Effects of Environmental Activist Investing.' European Corporate Governance Institute – Finance Working Paper No. 743/2021. https://ssrn.com/abstract=3483692

Pastor, L., Stambaugh, R., and Taylor, L. (2020). 'Sustainable Investing in Equilibrium.' Forthcoming in *Journal of Financial Economics*.

Perez, T. (2015). *United States Secretary of Labor's Remarks at Oct. 22, 2015 news conference on Interpretive Bulletin 2015-01 Relating to the Fiduciary Standard under ERISA in Considering Economically Targeted Investments.* US Department of Labor, Employee Benefits Security Administration, Washington, DC.

Rosen, B., D. Sandler, and D. Shani (1991). 'Social Issues and Socially Responsible Investment Behavior: A Preliminary Empirical Investigation.' *The Journal of Consumer Affairs,* 25(2): 221-234.

Ruf, B., K. Muralidhar, and K. Paul (1993). 'Eight Dimensions of Corporate Social Performance Determination of Relative Importance Using the Analytic Hierarchy Process.' *Academy of Management Proceedings*, Vol. 1993(1): 326–330.

Russo, M. and P. Fouts (1997). 'A Resource-Based Perspective on Corporate Environmental Performance and Profitability.' *Academy of Management Journal,* 40(3): 534-559.

Sharfman, M. (1996). 'The Construct Validity of the Kinder, Lydenberg and Domini Social Performance Ratings Data.' *Journal of Business Ethics,* 15: 287-296.

Social Investment Forum (2007). *2007 Report on US Sustainable and Impact Investing Trends*. Washington, DC: Social Investment Forum.

Statman, M., and D. Glushkov (2009). 'The Wages of Social Responsibility.' *Financial Analysts Journal,* 65(4): 33-46.

United States Department of Labor (2020). Proposed Rule: 29 CFR Part 2550 RIN 1210-AB95, *Financial Factors in Selecting Plan Investments.* https://www.napa-net.org/sites/napa-net.org/files/financial-factors-in-selecting-plan-investments-proposed-rule.pdf

US Social Investment Foundation (2020). *Report on US Sustainable and Impact Investing Trends,* 13th ed. (2020). Washington, DC: US SIF Foundation.

Waddock, S. and S. Graves (1997). 'The Corporate Social Performance-Financial Performance Link.' *Strategic Management Journal,* 18(4): 303-319.

Wharton Research Data Services. *KLD on WRDS.* https://wrds-www.wharton.upenn.edu/documents/1154/KLD-on-WRDS.pdf.

Wheeler, L. (1994) ‘Changes in Consensus Earnings Estimates and their Impact on Stock Returns.’ In B. Bruce and C. Epstein, eds., *The Handbook of Corporate Earnings Analysis*. Chicago: Probus Publishing.

Williams, J. (1938). *The Theory of Investment Value*. Cambridge: Harvard University Press.

Wimmer, M. (2013). ‘ESG-Persistence in Socially Responsible Mutual Funds.’ *Journal of Management and Sustainability*, 3(1): 9-15.

Womack, K. (1996) ‘Do Brokerage Analysts' Recommendations Have Investment Value?’ *Journal of Finance*, 51(1): 137-167.

Ziemba, W. and S. Schwartz (1993). *Invest Japan*. Chicago: Probus Publishing.

Endnotes

[1] The first commonly recognized paper on corporate social performance was by Milton Moskowitz (1972): he introduced the concept of social responsibility as a factor in the investment decision process and studied a handful of companies deemed to be acting according to corporate social responsibility practices and policies. Moskowitz (1997) reaffirmed his support of socially responsible investment (SRI) shortly after he established an award, the Moskowitz Prize, recognizing outstanding quantitative research in socially responsible investing. The Moskowitz Prize has been awarded annually since 1996, when Guerard (1997a, 1997b) won for research reporting no statistically significant costs associated with SRI, and with others either confirming or disconfirming this finding. In contrast, Geczy, Stambaugh, and Levin (2005), who were Honorable Mention awardees of the 2003 Prize competition, provided a detailed analysis demonstrating conditions under which SRI/ESG mutual fund portfolios created certainty equivalent costs relative to non-SRI/ESG portfolios.

[2] The major papers on the combination of value ratios for the prediction of stock returns (including at least CP and/or SP) include those of Jacobs and Levy (1988), Chan, Hamao, and Lakonishok (1990), Fama and French (1992 and 1995), Bloch, Guerard, Markowitz, Todd, and Xu (1993). Lakonishok, Shleifer, and Vishny (1994). Haugen and Baker (1996) later produced highly-cited variable testing which confirmed that fundamental variables enhanced portfolio returns over the long-run. Our point in this brief survey of anomalies is to acknowledge that Jacobs and Levy (1988), Chan et al. (1991), Bloch et al. (1993), and Ziemba and Schwartz (1991) were correct in their Berkeley Program in Finance and Q-Group presentations of the early 1990s on the inefficiencies of stock markets.

[3] Bloch et al. (1993) wrote their manuscript in 1991. At the time of the original estimation of an eight-factor regression model, the international Institutional Estimation Brokerage Service (I/B/E/S) was only four years old, having started in 1987.It lacked sufficient data for model building and testing, making it difficult for models with earnings forecasts to pass the Markowitz and Xu (1994) Data Mining Corrections test.

[4] Expected earnings have been used as a proxy for company's future cash flow in many studies. For a detailed analysis of analysts' consensus forecasts and share prices, see Elton, Gruber, and Gultekin (1981).

[5] The Bruce and Epstein (1994) and Brown (1998) works contain much of the rich history of earnings forecasting and resulting excess returns. Elton, Gruber, and Gultekin (1981) developed I/B/E/S database and published initial research using it. The Elton et al. (1981) paper is one of the more influential analyses in earnings forecasting and security analysis. Guerard, Takano, and Yamane (1993) employed Toyo Keizai earnings forecasts in Japan because of the limitations of the non-US I/B/E/S database. The Toyo Keizai earnings forecasts enhanced portfolio returns by over 200 basis points annually. Analysts were aware of the return-enhancement of I/B/E/S forecasts in US stocks; see Guerard and Stone (1992), research sponsored by the Institute for Quantitative Research in Finance, the 'Q-Group,' circa 1985. Womack (1996), Guerard, Gultekin and Stone (1997), Guerard, Markowitz, and Xu (2015), and Ball and Ghysels (2018), are among the thousands of studies of analysts' forecasting efficiency and how analysts' forecasts enhance portfolio return.

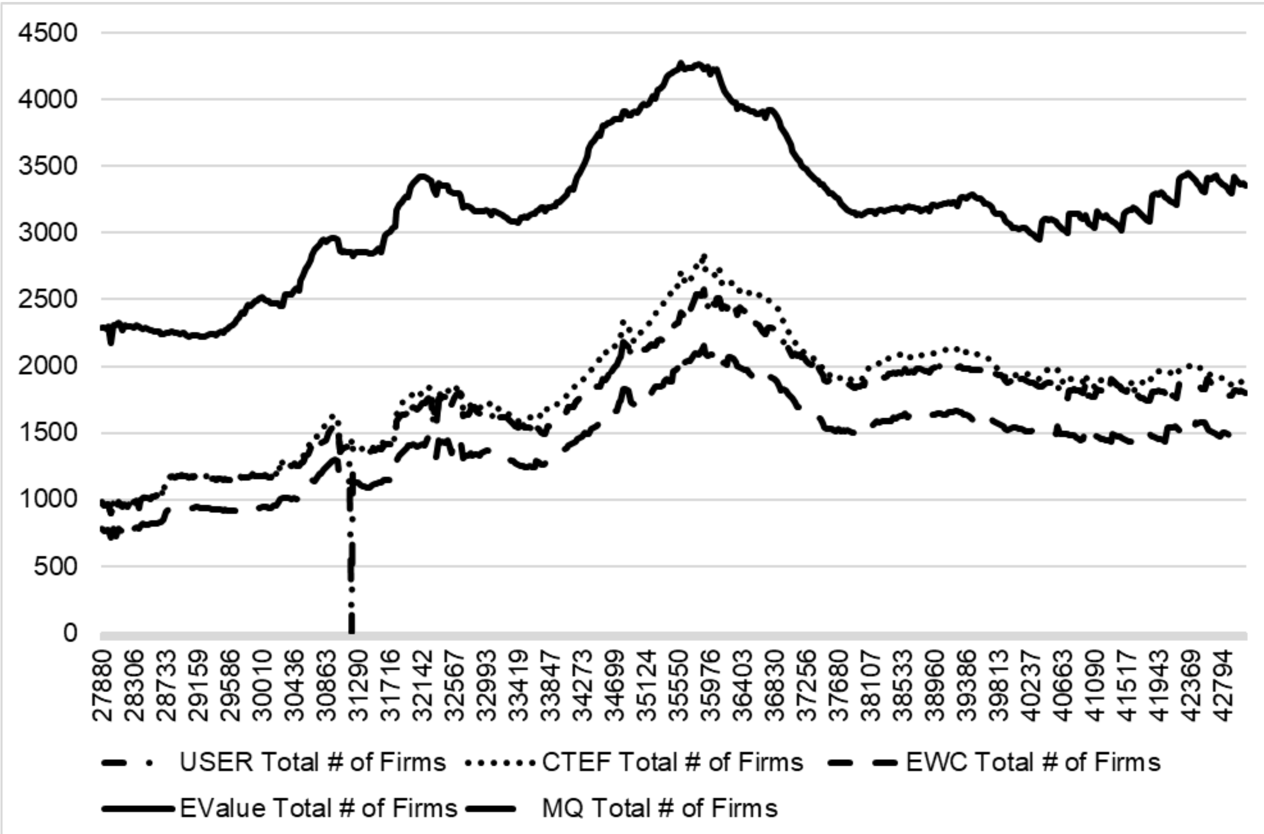
[6] The newly-created non-US I/B/E/S database did not have enough data in 1991-1994 to pass the Markowitz-Xu data mining test for its use in Japan.

---

[7] Wall Street practitioners have embraced the 'low PE' approach for well over 50 years.This is a form of the contrarian investment approach associated with Bernard (1959) and Dremen (1979, 1999). The authors believed in the low PE strategy, but not as an exclusive strategy. There is extensive literature on the impact of individual value ratios on the cross section of stock returns. We go beyond using just one or two of the standard value ratios (EP and BP) to include the cash-price ratio (CP) and/or the sales-price ratio (SP). The best writer on this topic, in our opinion, is Haugen (1995, 1999).

[8] Guerard and Mark (2020) reported monthly Axioma attribution statistics which, in the case of CTEF, indicates that the forecasted earnings acceleration variable loads on Medium-Term Momentum (0.257), Growth (0.151), and Value (0.469), and that Mean-variance CTEF and USER portfolios produced approximately 300-350 basis points of Specific Returns for the 20-year time period, 1996-2016. In US portfolios, equally-weighted 125 stock portfolios outperform Mean-variance (MV) 4 percent portfolios. In the Non-US and EAFE universes, Guerard and Mark (2020) reported that the CTEF ICs were higher than the USER ICs in their 10, 5, 3, and 1-year time sub-periods. The CTEF and USER produced approximately 400-500 basis points of Active Returns and about 250 basis points of Specific Returns. The Non-US portfolios offer more stock selection than US portfolios, with the addition of the REG8 plus CTEF (denoted REG9) and USER factors. The t-statistic on the risk stock selection effect in Non-US portfolios was maximized with ranked CTEF. The t-statistics on the risk stock selection effect were statistically significant for USER, although the t-statistic on the risk stock selection effect in the Non-US portfolios was only statically significant at the 10 percent level. Guerard and Mark (2020) reported that only ranked CTEF was statistically significant in the US, whereas globally, ranked CTEF and USER were statistically significant in Total Active Returns and Risk Stock Selection Returns.

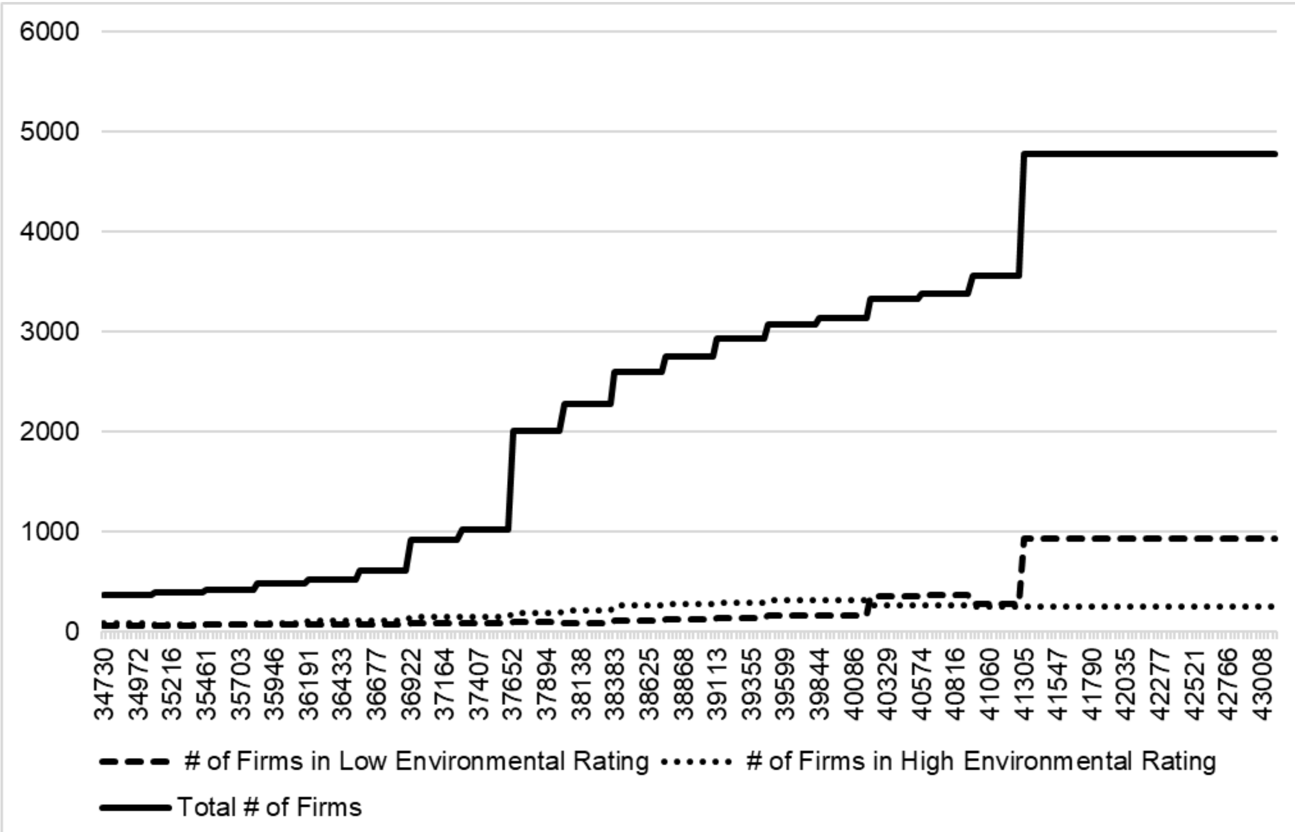
Figure 1. Total number of firms counted in model and selected sub-components



Note: The number of firms counted is the intersection of WRDS Compustat, CESP, and IBES databases.

Source: WRDS Compustat, CRSP, and IBES databases.

Figure 2. Total number of firms in each environmental rating category compared to sample total



Note: The number of firms counted is the intersection of WRDS Compustat, CESP, and IBES databases.

Source: WRDS Compustat, CRSP, and IBES databases.

Table 1. Reasons institutional investors report considering ESG factors, 2020 (US SIF Foundation)

Reason	**Number of Money Managers**	**% of Managers Responding**	**ESG Assets (in Billions)**
Risk	95	84%	$2,062
Client Demand	92	81%	$3,569
Social or Environmental Impact	90	80%	$3,476
Returns	82	73%	$2,355
Mission	79	70%	$2,445
Fiduciary Duty	72	64%	$3,557
UN Sustainable Development Goals	52	46%	$406
Regulatory Compliance	24	21%	$3,345
Total Responding	113		$3,621

Note: The table presents responses from 65 institutional investors who were asked to list the reasons they consider ESG factors among a series of possibilities. The table is ordered by the total ESG assets managed or held by these institutions. Assets and Numbers of Institutions may overlap since respondents could and did list more than one reason.

Source: Authors' adaptation of US SIF 2020 Trends Report, Figure 3.13.

Table 2. Summary statistics for the USER model and CTEF variable

		Annualized Arithmetic Return	Annualized Geometric Return	Annualized Volatility	# of Firms	Average Value
USER	Q5 (High)	15.8%	14.5%	20.5%	347	6.70%
	Q4	14.7%	13.7%	18.4%	347	0.03%
	Q3	14.6%	13.8%	17.7%	346	0.00%
	Q2	13.5%	12.5%	18.1%	347	0.00%
	Q1 (Low)	13.0%	11.1%	22.0%	347	-5.28%
	Q5-Q1	2.8%	2.3%	9.4%		
CTEF	Q5 (High)	15.8%	14.9%	18.8%	366	25.43%
	Q4	15.2%	14.3%	18.9%	365	13.42%
	Q3	15.1%	14.1%	19.0%	364	0.09%
	Q2	13.2%	11.7%	19.9%	365	-13.71%
	Q1 (Low)	11.1%	9.3%	20.9%	366	-24.54%
	Q5-Q1	4.7%	4.5%	6.5%		

Note: The table shows summary statistics for quintile portfolios formed on the ten-factor US Expected Return model (USER) of Guerard (1991, 1993) and Bloch et al. (1993) incorporating earnings yield, book-to-market, cashflow-to-price, and sales-to-price ratios, along with these ratios scaled by the average ratios over the previous five years as well as CTEF and price momentum. CTEF measures consensus earnings per share from I/B/E/S forecasts, revisions, and breadth. The sample ranges from May 1995 through December 2017. The table presents the average annual unconditional returns for the quintile groups, the average geometric return, the annualized standard deviation (volatility), the number of firms in each quintile, and the average of the USER model variable or CTEF, respectively.

Source: Authors' computations.

Table 3a. Baseline performance of five models of expected returns

Multifactor Models Regressions (March 1991 - December 2017)

	One-Factor Model (CAPM)			Fama-French/Carhart Four-Factor Model						Fama-French/Carhart Plus Quality Five-Factor Model						
USER	RMRF	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2
Quintile 5	1.24	3.2%	76.9%	**1.02**	**0.86**	0.07	-0.55	**3.6%**	90.8%	1.00	0.81	0.10	-0.54	-0.13	4.3%	90.9%
Quintile 4	1.13	2.2%	77.1%	**0.96**	**0.64**	**0.24**	-0.21	**2.0%**	94.3%	0.97	0.67	0.23	-0.22	0.08	1.5%	94.4%
Quintile 3	1.05	2.5%	74.6%	**0.91**	**0.63**	**0.28**	-0.22	**2.8%**	93.3%	0.92	0.65	0.27	-0.22	0.05	2.5%	93.3%
Quintile 2	1.10	1.6%	77.0%	**0.98**	**0.67**	**0.23**	-0.21	**2.5%**	94.2%	1.00	0.70	0.21	-0.22	0.09	2.0%	94.3%
Quintile 1	1.32	-0.1%	64.8%	**1.11**	**0.80**	**0.15**	-0.04	0.4%	93.9%	1.09	0.76	0.17	-0.03	-0.10	1.0%	94.0%
L/S (Q5-Q1)	-0.08	3.3%	0.2%	-0.09	0.06	-0.08	-0.51	3.2%	53.6%	-0.09	0.05	-0.07	-0.51	-0.03	3.3%	58.1%
EWC	RMRF	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2
Quintile 5	1.23	1.7%	75.1%	**1.05**	**0.72**	**0.22**	-0.31	2.9%	93.8%	1.05	0.72	0.22	-0.31	0.00	2.9%	93.8%
Quintile 4	1.17	1.4%	75.9%	**1.00**	**0.72**	**0.15**	-0.23	2.1%	93.8%	1.01	0.74	0.14	-0.23	0.05	1.9%	93.8%
Quintile 3	1.09	1.6%	74.5%	**0.94**	**0.68**	**0.22**	-0.23	2.1%	93.1%	0.94	0.68	0.21	-0.23	0.00	2.1%	93.1%
Quintile 2	1.15	1.2%	77.1%	**0.99**	**0.69**	**0.18**	-0.22	1.8%	94.4%	1.00	0.70	0.17	-0.22	0.02	1.7%	94.4%
Quintile 1	1.20	1.6%	73.5%	**1.02**	**0.79**	0.21	-0.24	2.2%	93.3%	1.00	0.76	0.23	-0.24	-0.09	2.7%	93.4%
L/S (Q5-Q1)	0.03	0.1%	3.2%	0.03	-0.07	0.01	-0.07	0.7%	9.8%	0.04	-0.04	-0.01	-0.07	0.09	0.2%	14.2%
EVALUE	RMRF	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2
Quintile 5	1.13	3.0%	67.0%	0.91	0.74	0.12	-0.36	5.0%	89.2%	0.89	0.70	0.14	-0.36	-0.13	4.7%	89.3%
Quintile 4	1.06	2.4%	69.4%	0.89	0.75	0.11	-0.18	3.0%	89.3%	0.87	0.71	0.13	-0.18	-0.09	3.5%	89.4%
Quintile 3	0.97	2.9%	69.1%	0.82	0.65	0.13	-0.19	3.6%	88.0%	0.80	0.62	0.14	-0.19	-0.08	4.0%	88.0%
Quintile 2	1.04	2.1%	70.7%	0.88	0.71	0.15	-0.19	2.6%	89.9%	0.87	0.68	0.16	-0.18	-0.07	3.0%	89.9%
Quintile 1	0.98	2.3%	69.5%	0.83	0.68	0.08	-0.17	2.9%	88.9%	0.80	0.63	0.11	-0.16	-0.14	3.7%	89.1%
L/S (Q5-Q1)	0.15	0.7%	18.2%	0.09	0.06	0.04	-0.19	2.0%	44.5%	0.09	0.07	0.03	-0.19	0.00	1.0%	55.4%

Note: The tables show performance and factor exposures of quintile portfolios formed on five expected return models: The ten-factor US Expected Return model (USER) of Guerard (1991, 1993) and Bloch et al. (1993) incorporates earnings yield, book-to-market, cashflow-to-price, and sales-to-price ratios, along with these ratios scaled by the average ratios over the previous five years as well as CTEF and price momentum. (CTEF measures consensus earnings per share from I/B/E/S forecasts, revisions and breadth, and PM is 7/1 price momentum). In addition, results are given for an equal-weighted model with the same characteristic variables (EWC), an equal-weighted naive value-based model using just the scaled price ratios above (EVALUE), and MQ, an equal-weighted model including CTEF and price momentum. Monthly Quantile portfolio returns or L/S zero investment portfolio returns are regressed on the one-factor US equity premium (RMRF) model (the CAPM), the Fama-French/Carhart four factor model, and a Fama-French/Carhart five-factor model that includes the Fama-French quality factor.

Source: Authors' computations.

Table 3b. Baseline performance of five models of expected returns, cont'd.

MQ	RMRF	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2
Quintile 5	1.13	3.5%	76.1%	1.04	0.86	0.02	-0.69	2.3%	90.4%	1.02	0.82	0.05	-0.69	-0.13	3.1%	90.5%
Quintile 4	1.14	2.8%	78.7%	0.99	0.73	0.18	-0.32	2.0%	94.1%	0.98	0.71	0.20	-0.32	-0.05	2.3%	94.1%
Quintile 3	1.05	1.6%	74.0%	0.91	0.66	0.24	-0.20	1.8%	92.3%	0.92	0.68	0.23	-0.20	0.05	1.6%	92.3%
Quintile 2	1.19	0.5%	73.9%	1.01	0.69	0.22	-0.13	2.5%	94.9%	1.02	0.70	0.21	-0.13	0.02	2.3%	94.9%
Quintile 1	1.38	-2.7%	61.4%	1.06	0.72	0.20	0.10	1.4%	93.8%	1.07	0.73	0.19	0.10	0.03	1.2%	93.7%
L/S (Q5-Q1)	-0.26	6.2%	6.0%	-0.02	0.14	-0.18	-0.80	0.9%	71.0%	-0.05	0.09	-0.14	-0.79	-0.16	1.8%	35.4%

CTEF	RMRF	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Intercept	Adj R^2	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2
Quintile 5	1.15	2.1%	79.4%	1.02	0.62	0.16	-0.14	2.3%	92.6%	1.04	0.65	0.14	-0.15	9.2%	1.7%	92.7%
Quintile 4	1.17	2.6%	77.7%	1.02	0.70	0.16	-0.19	3.0%	94.1%	1.01	0.70	0.16	-0.19	-0.5%	3.0%	94.1%
Quintile 3	1.12	2.9%	70.0%	0.94	0.77	0.12	-0.22	3.8%	89.8%	0.92	0.73	0.14	-0.22	-9.9%	4.4%	89.9%
Quintile 2	1.23	-0.6%	74.1%	1.04	0.80	0.17	-0.29	0.5%	94.4%	1.03	0.78	0.18	-0.28	-5.1%	0.8%	94.4%
Quintile 1	1.27	-2.5%	71.3%	1.04	0.83	0.13	-0.37	-0.5%	93.2%	1.02	0.80	0.15	-0.36	-8.6%	0.0%	93.2%
L/S (Q5-Q1)	-0.12	4.6%		-0.01	-0.21	0.03	0.22	2.8%		0.02	-0.14	-0.01	0.22	0.18	1.8%	

Note: The tables show performance and factor exposures of quintile portfolios formed on five expected return models: The ten-factor US Expected Return model (USER) of Guerard (1991, 1993) and Bloch et al. (1993) incorporating earnings yield, book-to-market, cashflow-to-price, and sales-to-price ratios, along with these ratios scaled by the average ratios over the previous five years as well as CTEF and price momentum. (CTEF measures consensus earnings per share from I/B/E/S forecasts, revisions, and breadth, and PM is 7/1 price momentum). In addition, results are given for an equal-weighted model with the same characteristic variables (EWC), an equal-weighted naive value-based model using just the scaled price ratios above (EVALUE), and MQ, an equal-weighted model including CTEF and price momentum. Monthly Quantile portfolio returns or L/S zero investment portfolio returns are regressed on the one-factor US equity premium (RMRF) model (the CAPM), the Fama-French/Carhart four-factor model, and a Fama-French/Carhart five-factor model that includes the Fama-French quality factor.

Source: Authors' computations.

Table 4a. The interaction of expected return models and ESG/KLD scores: The case of USER and Environmental scores

Multifactor Models Regression Parameters (March 1995 - December 2017)

	One-Factor Model (CAPM)					Fama-French/Carhart Four-Factor Model								Fama-French/Carhart Plus Quality Five Factor Model								
USER	RMRF	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2	TE	AR
High ENV + High USER	**0.95**	**3.6%**	**69.9%**	**9.4%**	**0.38**	**0.94**	**0.29**	**0.41**	**-0.07**	**4.6%**	**78.7%**	**8.0%**	**0.57**	**1.01**	**0.38**	**0.26**	**-0.19**	**0.22**	**1.8%**	**77.1%**	**9.5%**	**0.19**
High ENV + Low USER	1.04	2.8%	69.0%	10.5%	0.27	1.01	0.29	0.32	-0.16	4.2%	77.2%	9.1%	0.46	0.99	0.44	0.28	-0.10	0.36	0.9%	79.8%	8.2%	0.11
Low ENV + High User	1.11	-0.3%	64.0%	12.6%	-0.02	1.13	0.16	0.57	-0.10	0.4%	72.5%	10.8%	0.04	1.10	0.35	0.60	-0.29	0.44	-2.8%	80.5%	10.0%	-0.27
Low ENV + Low User	**1.05**	**0.1%**	**58.0%**	**13.5%**	**0.01**	**1.04**	**0.15**	**0.76**	**-0.24**	**1.0%**	**78.9%**	**9.6%**	**0.11**	**1.22**	**0.40**	**0.39**	**-0.15**	**0.59**	**-4.5%**	**76.9%**	**11.4%**	**-0.40**
High ENV: L/S USER	-0.08	0.8%	1.3%	9.8%	0.08	-0.07	0.00	0.09	0.10	0.4%	4.1%	9.6%	0.04	-0.02	0.07	0.02	0.09	0.14	-1.0%	3.6%	9.9%	-0.10
Low ENV: L/S USER	0.06	-0.4%	0.7%	8.6%	-0.05	0.09	0.00	-0.18	0.14	-0.7%	15.5%	7.9%	-0.08	0.12	0.05	-0.21	0.13	0.15	-1.8%	15.7%	8.2%	-0.22
(High+High) - (Low+Low)	-0.10	3.5%				-0.10	0.14	-0.35	0.18	3.5%				-0.21	-0.02	-0.13	-0.03	-0.37	6.4%			
EWC	RMRF	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2	TE	AR
High ENV + High EWC	**0.96**	**3.3%**	**67.0%**	**10.1%**	**0.33**	**0.91**	**0.32**	**0.44**	**-0.11**	**2.7%**	**77.2%**	**8.4%**	**0.32**	**1.00**	**0.48**	**0.33**	**-0.13**	**0.41**	**0.4%**	**79.7%**	**7.9%**	**0.05**
High ENV + Low EWC	1.03	3.0%	72.0%	9.7%	0.31	0.96	0.28	0.25	-0.14	3.3%	77.6%	8.7%	0.38	1.00	0.34	0.20	-0.15	0.17	2.3%	77.9%	8.6%	0.27
Low ENV + High EWC	1.13	0.0%	64.7%	12.5%	0.00	1.10	0.18	0.56	-0.18	-0.4%	75.7%	10.3%	-0.04	1.22	0.40	0.40	-0.21	0.59	-3.8%	79.3%	9.5%	-0.40
Low ENV + Low EWC	**1.04**	**-0.2%**	**58.3%**	**13.2%**	**-0.02**	**1.01**	**0.18**	**0.71**	**-0.21**	**-1.0%**	**76.5%**	**9.9%**	**-0.10**	**1.11**	**0.34**	**0.59**	**-0.23**	**0.44**	**-3.5%**	**78.5%**	**9.4%**	**-0.37**
High ENV: L/S EWC	-0.07	-2.0%	1.0%	9.6%	-0.21	-0.05	0.04	0.18	0.02	-2.8%	3.5%	9.4%	-0.30	0.00	0.14	0.12	0.01	0.24	-4.2%	6.1%	9.3%	-0.45
Low ENV: L/S EWC	0.09	-2.1%	2.6%	7.6%	-0.27	0.08	0.01	-0.16	0.02	-1.7%	7.0%	7.4%	-0.23	0.11	0.06	-0.20	0.01	0.15	-2.6%	8.4%	7.4%	-0.35
(High+High) - (Low+Low)	-0.08	3.5%				-0.10	0.14	-0.27	0.09	3.7%				-0.11	0.13	-0.26	0.10	-0.03	3.9%			
EVALUE	RMRF	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2	TE	AR
High ENV + High EVALU	**0.99**	**3.5%**	**64.5%**	**11.1%**	**0.32**	**0.94**	**0.31**	**0.45**	**-0.15**	**3.1%**	**74.9%**	**9.3%**	**0.34**	**1.01**	**0.45**	**0.35**	**-0.17**	**0.37**	**1.0%**	**76.6%**	**8.9%**	**0.12**
High ENV + Low EVALUI	1.03	2.9%	72.8%	9.5%	0.31	0.97	0.29	0.25	-0.13	3.1%	78.2%	8.5%	0.36	1.01	0.37	0.19	-0.14	0.21	1.9%	78.8%	8.4%	0.22
Low ENV + High EVALUI	1.15	0.5%	63.6%	13.1%	0.04	1.11	0.22	0.57	-0.20	0.1%	74.8%	10.8%	0.00	1.24	0.46	0.40	-0.23	0.63	-3.5%	78.6%	10.0%	-0.35
Low ENV + Low EVALU	**1.03**	**-0.3%**	**58.6%**	**13.1%**	**-0.03**	**1.01**	**0.17**	**0.69**	**-0.21**	**-1.1%**	**76.4%**	**9.8%**	**-0.11**	**1.10**	**0.33**	**0.58**	**-0.23**	**0.43**	**-3.5%**	**78.4%**	**9.4%**	**-0.38**
High ENV: L/S EVALUE	-0.04	-1.8%	0.0%	9.9%	-0.18	-0.03	0.03	0.19	-0.03	-2.2%	3.7%	9.7%	-0.23	0.00	0.09	0.15	-0.04	0.16	-3.1%	4.5%	9.6%	-0.33
Low ENV: L/S EVALUE	0.12	-1.6%	4.8%	7.9%	-0.20	0.10	0.06	-0.13	0.00	-1.1%	8.0%	7.7%	-0.15	0.14	0.13	-0.18	0.00	0.20	-2.3%	10.6%	7.6%	-0.30
(High+High) - (Low+Low)	-0.04	3.8%				-0.07	0.15	-0.24	0.06	4.2%				-0.08	0.12	-0.22	0.06	-0.06	4.6%			

Note: The tables show performance and factor exposures of portfolios formed using the ten-factor US Expected Return model (USER) of Guerard (1991, 1993) and Bloch et al. (1993) incorporating earnings yield, book-to-market, cashflow-to-price, and sales-to-price ratios, along with these ratios scaled by the average ratios over the previous five years as well as CTEF and price momentum. (CTEF measures consensus earnings per share from I/B/E/S forecasts, revisions, and breadth, and PM is 7/1 price momentum). The monthly returns of high (low) KLD Environmental score firms with high or low USER rankings or L/S zero investment portfolio returns are regressed on a one-factor US equity premium (RMRF) model (the CAPM), the Fama-French/Carhart four-factor model, and a Fama-French/Carhart five-factor model that includes the Fama-French quality factor. TE is the unbiased residual standard deviation. AR is the appraisal ratio (Information Ratio with unconstrained beta).

Source: Authors' computations

Table 4b. The interaction of expected return models and ESG/KLD scores: The case of USER and Environmental scores

MQ	RMRF	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2	TE	AR
High ENV + High MQ	**0.90**	**3.5%**	**69.8%**	**8.9%**	**0.40**	**0.90**	**0.27**	**0.35**	**0.02**	**2.2%**	**76.0%**	**7.9%**	**0.28**	**0.98**	**0.41**	**0.25**	**0.00**	**0.38**	**0.0%**	**78.5%**	**7.5%**	**0.00**
High ENV + Low MQ	1.09	2.9%	66.7%	11.6%	0.25	0.98	0.33	0.33	-0.28	4.0%	77.5%	9.5%	0.42	1.02	0.40	0.28	-0.29	0.20	2.8%	77.9%	9.4%	0.30
Low ENV + High MQ	1.06	0.6%	61.2%	12.7%	0.05	1.08	0.14	0.57	-0.04	-1.0%	69.5%	11.2%	-0.09	1.20	0.35	0.42	-0.06	0.56	-4.2%	73.0%	10.5%	-0.40
Low ENV + Low MQ	**1.12**	**-0.9%**	**55.5%**	**15.0%**	**-0.06**	**1.03**	**0.23**	**0.71**	**-0.36**	**-0.5%**	**76.8%**	**#####**	**-0.05**	**1.13**	**0.40**	**0.58**	**-0.38**	**0.48**	**-3.3%**	**78.9%**	**10.3%**	**-0.32**
High ENV: L/S MQ	-0.19	-1.7%	5.7%	11.3%	-0.15	-0.08	-0.05	0.01	0.29	-4.1%	22.3%	10.1%	-0.40	-0.04	0.02	-0.03	0.28	0.19	-5.1%	23.2%	10.0%	-0.51
Low ENV: L/S MQ	-0.06	-1.0%	0.1%	12.4%	-0.08	0.05	-0.08	-0.15	0.32	-3.0%	21.6%	10.9%	-0.27	0.07	-0.05	-0.17	0.31	0.09	-3.5%	21.6%	10.9%	-0.32
(High+High) - (Low+Low)	-0.22	4.5%				-0.14	0.04	-0.36	0.38	2.7%				-0.16	0.01	-0.33	0.38	-0.10	3.3%			

CTEF	RMRF	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Intercept	Adj R^2	TE	AR	RMRF	SMB	HML	MOM	Quality	Intercept	Adj R^2	TE	AR
High ENV + High CTEF	**0.95**	**2.3%**	**69.9%**	**9.3%**	**0.25**	**0.95**	**0.22**	**0.37**	**-0.01**	**1.1%**	**75.3%**	**8.4%**	**0.13**	**1.05**	**0.39**	**0.24**	**-0.03**	**0.47**	**-1.6%**	**78.8%**	**7.8%**	**-0.20**
High ENV + Low CTEF	1.05	4.2%	66.5%	11.2%	0.38	0.93	0.38	0.31	-0.26	5.2%	77.7%	9.1%	0.56	0.95	0.42	0.28	-0.26	0.12	4.5%	77.8%	9.1%	0.49
Low ENV + High CTEF	1.09	1.2%	60.8%	13.1%	0.09	1.07	0.15	0.56	-0.15	0.4%	70.9%	11.2%	0.04	1.19	0.37	0.40	-0.18	0.58	-2.9%	74.4%	10.5%	-0.27
Low ENV + Low CTEF	**1.09**	**-1.6%**	**58.6%**	**13.8%**	**-0.12**	**1.05**	**0.22**	**0.72**	**-0.24**	**-2.2%**	**76.8%**	**#####**	**-0.21**	**1.14**	**0.39**	**0.60**	**-0.26**	**0.46**	**-4.8%**	**78.9%**	**9.8%**	**-0.48**
High ENV: L/S CTEF	-0.10	-4.2%	1.4%	11.0%	-0.38	0.02	-0.16	0.05	0.24	-6.3%	14.6%	10.1%	-0.62	0.09	-0.03	-0.05	0.22	0.36	-8.3%	19.1%	9.9%	-0.85
Low ENV: L/S CTEF	0.00	0.4%	-0.4%	10.3%	0.04	0.02	-0.07	-0.17	0.08	0.3%	4.1%	10.0%	0.03	0.05	-0.02	-0.20	0.08	0.13	-0.4%	4.5%	10.0%	-0.04
(High+High) - (Low+Low)	-0.14	3.9%				-0.10	0.00	-0.35	0.23	3.3%				-0.10	0.00	-0.35	0.23	0.01	3.2%			

Note: The tables show performance and factor exposures of portfolios formed using the ten-factor US Expected Return model (USER) of Guerard (1991, 1993) and Bloch et al. (1993) incorporating earnings yield, book-to-market, cashflow-to-price, and sales-to-price ratios, along with these ratios scaled by the average ratios over the previous five years as well as CTEF and price momentum. (CTEF measures consensus earnings per share from I/B/E/S forecasts, revisions, and breadth, and PM is 7/1 price momentum). The monthly returns of high (low) KLD Environmental score firms with high or low USER rankings or L/S zero investment portfolio returns are regressed on a one-factor US equity premium (RMRF) model (the CAPM), the Fama-French/Carhart four-factor model, and a Fama-French/Carhart five-factor model that includes the Fama-French quality factor. TE is the unbiased residual standard deviation. AR is the appraisal ratio (Information Ratio with unconstrained beta).

Source: Authors' computations.

Appendix: KLD STATS[9]

MSCI KLD STATS ("KLD," STATISTICAL TOOL FOR ANALYZING TRENDS IN SOCIAL AND ENVIRONMENTAL PERFORMANCE) is a dataset with annual snapshots of the environmental, social, and governance performance of companies rated.

Strength and Concern (Positive and Negative Indicator) Ratings

KLD STATS covers indicators in seven major Qualitative Issue Areas including Community, Corporate Governance, Diversity, Employee Relations, Environment, Human Rights and Product. It presents a binary summary of positive and negative ESG ratings. In each case, if KLD assigned a rating in a particular issue (either positive or negative), this is indicated with a one in the corresponding cell. If the company did not have a strength or concern in that issue, this is indicated with a 0. KLD STATS data are organized by year. Each year, RiskMetrics takes a snapshot of its ratings and index membership to reflect the data at calendar year end. Each spreadsheet contains identifying information about the company, index membership, a listing of positive and negative ratings, involvement in controversial business issues, and total counts for each area. Additionally, the data provide summary count of all strengths and concerns the company received in a general category (either Qualitative Issue Area or Controversial Business Issue) in that year. The Environmental indicators are calculated separately but similarly to those in Geczy, Guerard, and Samonov (2020).

ENVIRONMENT (ENV-)

STRENGTHS

Beneficial Products and Services (ENV-str-A). The company derives substantial revenues from innovative remediation products, environmental services, or products that promote the efficient use of energy, or it has developed innovative products with environmental benefits.

Pollution Prevention (ENV-str-B). The company has notably strong pollution prevention programs including both emissions reductions and toxic-use reduction programs.

Recycling (ENV-str-C). The company is either a substantial user of recycled materials as raw materials in its manufacturing processes, or a major factor in the recycling industry.

Clean Energy (ENV-str-D). The company has taken significant measures to reduce its impact on climate change and air pollution through use of renewable energy and clean fuels, or through energy efficiency. The company has demonstrated a commitment to promoting climate-friendly policies and practices outside its own operations.

Communications (ENV-str-E). The company is a signatory to the CERES Principles, publishes a notably substantive environmental report, or has notably effective internal communications systems in place for environmental best practices. KLD began assigning strengths for this issue in 1996, and then incorporated the issue with the Corporate Governance: Transparency rating (CGOV-str-D), which was added in 2005.

Property, Plant, and Equipment (ENV-str-F). The company maintains its property, plant, and equipment with above average environmental performance for its industry. KLD has not assigned strengths for this issue since 1995.

Management Systems (ENV-str-G). The company has demonstrated a superior commitment to management systems through ISO 14001 certification and other voluntary programs.

CONCERNS

Hazardous Waste (ENV-con-A). The company's liabilities for hazardous waste sites exceed $50 million, or the company has recently paid substantial fines or civil penalties for waste management violations.

Regulatory Problems (ENV-con-B). The company has recently paid substantial fines or civil penalties for violations of air, water, or other environmental regulations, or it has a pattern of regulatory controversies under the Clean Air Act, Clean Water Act or other major environmental regulations.

Ozone Depleting Chemicals (ENV-con-C). The company is among the top manufacturers of ozone depleting chemicals such as HCFCs, methyl chloroform, methylene chloride, or bromines.

Substantial Emissions (ENV-con-D). The company's legal emissions of toxic chemicals (as defined by and reported to the EPA) from individual plants into the air and water are among the highest of the companies followed by KLD.

Agricultural Chemicals (ENV-con-E). The company is a substantial producer of agricultural chemicals, *i.e.,* pesticides or chemical fertilizers.

Climate Change (ENV-con-F). The company derives substantial revenues from the sale of coal or oil and its derivative fuel products, or the company derives substantial revenues indirectly from the combustion of coal or oil and its derivative fuel products. Such companies include electric utilities, transportation companies with fleets of vehicles, auto and truck manufacturers, and other transportation equipment companies.

[9] See Wharton Research Data Services, “KLD on WRDS.” https://wrds-www.wharton.upenn.edu/documents/1154/KLD-on-WRDS.pdf.