The Rise of Sustainable Investing: Does Employment of ESG Talents Predict Mutual Funds Activities?

Yutong Yan[†]

Department of Finance

The Chinese University of Hong Kong

yutongyan@cuhk.edu.hk

Abstract

In this proposal, I plan to examine whether hiring ESG-specialized professionals in mutual funds is related to the funds' portfolio choice, ESG rating, and subsequent funds' high ESG portfolio performance and fund flows. It will shed light on the role of ESG-specialized human capital in asset management business in the presence of fast growth in ESG investments. ESG-specialized professionals, who are knowledgeable about the nuances of climate change issues and sustainability, are expected to assist mutual funds to make strategic actions while navigating ESG-related risks and opportunities. To compile a comprehensive LinkedIn-based mutual fund employee database, I use large-scale web scraping and textual analysis techniques to analyze ESG employees' education backgrounds, past work experiences, expertise (e.g., science- vs management-oriented), etc. In addition, I merge the ESG employee information with the standard mutual fund databases (CRSP and Thomson Reuters) to create a Linkedin employee-fund database. I hypothesize that hiring ESG experts should be associated with the improvement in fund ESG rating and portfolio ESG score, better fund performance on high ESG stocks, and more fund flows.

Keywords: Human Capital; Big Data; ESG Preferences; Institutional Investors; Investment Decisions

[†]I also prepare slides, available at my website.

1 Introduction

According to a recent report by Bloomberg, Environmental, Social, and Governance (ESG) assets under management (AUM) could climb to more than a third of the projected \$140.5 trillion global total by 2025[1]. In response to the rapid growth of the ESG market, some institutions have broadened their perspectives and considered a firm's ESG achievement when making investment decisions. A recent international energy agency (IEA) report estimates that 14 million new jobs will be created by 2030 in the transition to ESG[2].

Despite the skyrocketing demand for ESG professionals, the role of human capital with specialization in ESG has not yet been explored in explaining mutual fund investments. Human capital component of financial services plays a vital role in the investment decision-making process. The standard human capital investment model (Becker, 1964) implies a positive relationship between human capital such as education and fund performance. Most existing works (Golec, 1996; Chevalier and Ellison, 1999; Gottesman and Morey, 2006) investigate the link between fund performance and fund managers human capital. There has been little attention paid to other fund employees in previous research. The skillsets of employees with ESG expertise are also crucial to the fund performance as per human capital theory. The two types of expertise provided by ESG-specialized professionals can be of substantial value to mutual funds: (1) analytical skills in corporate carbon neutrality, energy transition and low-carbon energy conversion systems; (2) technical skills in ESG investing, disclosure/reporting/standards, ratings assessment, risk management, and regulatory analysis.

Unfortunately, existing databases like Thomson Reuters only contain information about fund managers; information about fund employees cannot be obtained. Nevertheless, with the development of online employment and network building, LinkedIn[3], the largest professional networking site on the internet, is an ideal source of data. With

[1] Bloomberg Intelligence, "ESG assets may hit \$53 trillion by 2025, a third of global AUM," https://www.bloomberg.com/professional/blog/esg-assets-may-hit-53-trillion-by-2025-a-third-of-global-aum/.

[2] International Energy Agency, "Net zero by 2050: A Roadmap for the Global Energy Sector," https://www.iea.org/reports/net-zero-by-2050.

[3] LinkedIn photos have been used to analyze the facial characteristics of financial analysts (Peng et al., 2022).

my profound training in computer science, I am able to retrieve millions of employee profiles from LinkedIn. Profiles include information regarding employees' educational backgrounds and work experiences. Most professionals in the finance industry maintain detailed LinkedIn profiles with numerous textual descriptions in order to build up personal brands. It is therefore apparent that this database contains two distinct types of data: (1) structured data, such as majors and job titles, and (2) unstructured data, such as descriptions of school activities and work duties. Dictionary-based and deep learning-based methods can be applied accordingly to classify the data into two categories: ESG-related and others. I therefore construct two indicators of ESG-specialized professionals: ESG-related educational backgrounds (ESG_EDU) and ESG-related professional backgrounds (ESG_WORK) from the compiled LinkedIn database.

To begin, I examine whether the hiring of ESG-specialized professionals predicts mutual fund ESG ratings in next quarter, controlling for the effect of current ratings. These professionals may be able to facilitate ESG integration within the fund as well as conduct relevant research, such as regulatory frameworks and data. It is likely that these professionals can assist funds in tilting their holdings in favor of companies that are aligned with ESG investing objectives by advocating on their behalf. Therefore, the correlation between recruitment of ESG experts and next quarter's ESG ratings is expected to be positive.

Next, I intend to test if recruitment of ESG specialists predicts next quarter's mutual fund performance on high ESG stocks. In the event that my first hypothesis is supported by empirical evidence, then the natural question is whether hiring ESG specialists will lead to outperformance of funds' high ESG stock holdings? The second test uses the hiring of ESG-specialized professionals to predict fund high ESG excess returns in next quarter, controlling for the effect of current quarter high ESG excess returns. In light of diverse backgrounds of ESG professionals, which range from sustainability management to environmental sciences, they are believed to possess ESG stock-picking skills, such as distinguishing between greenwashing and green marketing and selecting better-performing companies with strong environmental and social values. Thus, I anticipate that funds with ESG experts will outperform others among high ESG stocks.

How do investors react to the hiring decisions of ESG-specialized professionals considering fund outperformance on high ESG stocks? It is widely accepted that the recruitment of these professionals implies the funds' investment preferences. Consequently, I intend to test whether ESG hiring can predict next quarter's fund flows, while controlling for current quarter fund flows and performance. Investors with preferences for sustainable investing also yearn for satisfactory returns. Accordingly, I assert that funds with these specialists will attract positive investment flows if empirical evidence from previous tests supports my hypotheses, as incorporating these specialists not only enhances the funds' ESG branding, but mitigates the performance-preference tradeoff.

This proposal is organized as follows. Section 2 describes my data sources and the steps of collecting LinkedIn data. Section 3 presents the textual analysis tools employed to construct *ESG_EDU* and *ESG_WORK* measure. Section 4 provides three tests capturing whether recruitment of ESG-specialized professionals can predict mutual fund ESG ratings, mutual fund performance on high ESG stocks, and mutual fund flows.

2 Data

2.1 Mutual Fund Characteristics

The proposed research project incorporates two mutual fund databases. The Center for Research on Security Prices (CRSP) Mutual Fund Database dates back to 1961. It provides information about the fund name, management company (also called fund family)[4], listing status, and assets under management. In addition, it provides detailed information about the fund family, address, website, etc., to which each fund belongs. The Thomson Mutual Fund Holdings Database, dating back to 1980, offers comprehensive information on the quarterly portfolio holdings of each fund. The two databases are linked using the Mutual Fund Links (MFLINKS) tables developed by (Wermers, 2000).

[4] Every management company has a unique management company code, provided by the CRSP mutual fund database.

2.2 LinkedIn Data Scraping

To verify the authenticity of work history, LinkedIn requires users to use their organizational email address. Therefore, the profile page of every employee listed on the company page can be easily accessed. It will be possible to collect the LinkedIn profile, as well as educational and professional information, of fund family employees. The workflow is as follows: the first thing to do is search for the fund family name on LinkedIn and arrive at the company page. There is an employee subpage on the company page that includes the name of each employee and a hyperlink to their LinkedIn profile. It is possible to encounter technical difficulties due to restrictions on the number of profiles per company that can be viewed per day on LinkedIn[5]. Considering the number of management companies is relatively small, it could be solved by using multiple proxy servers. As a result of website limitations, the data collection process is expected to take a few months. Job changes usually occur every few years, so this should not be a significant issue. When a user creates a LinkedIn account, the link to their profile remains unchanged. In order to resolve the issue of missing histories, I can maintain a database of all professionals who have worked for management companies and update it every year using the profile links saved in the database.

Data collection will be automated using Selenium[6], a tool for automating web interactions. In all cases, profiles will be saved in HTML format, which is a standard markup language used in web browsers to display documents. All the information will be extracted from the HTML language using regular expressions[7]. The user profile should include information about his educational background, such as the school and degree he has received, as well as his professional experience, such as the previous employers. If the starting and ending periods are missing, they can be inferred from other experiences on profile, for example, calculating their graduation date from their first job. In the profile, three categories have been defined: education, employment, and a personal statement.

[5] The current limit is 1000 profiles per day.
[6] The tool is available at https://www.selenium.dev/.
[7] Regular expressions are sequences of characters that specify a search pattern in text.

2.3 Dataset Potential Problems

Three potential issues arise with data collection. The first problem is caused by missing profiles. From 37 million in 2010 to 850 million in 2022, the LinkedIn user base has grown exponentially. Many professionals may not be interested in using LinkedIn or may not register before reporting all of their work histories, despite LinkedIn having a large user base today. Unfortunately, the problem of missing profiles cannot be solved. A reasonable assumption is that every management company is expected to have an equal number of missing profiles.

A profile history, including past modifications and deleted content, is not available at the present time. This second issue can be resolved by taking a snapshot once a year going forward.

In light of LinkedIn's growth since 2010, the ideal sampling period would be between 2010 and 2022. There were nearly 3000 fund families in total from 2010 to 2022. The number of employees varies from dozens to thousands depending on the size of the fund family. The results of (Guedj and Papastaikoudi, 2003) demonstrate that large fund families allocate resources unequally to each fund. Despite that, since information about employees can only be obtained for fund families, rather than individual funds, I assume that the effect of employee specialization is shared for all mutual funds within a fund family.

2.4 Educational Profile

Studies like (Chevalier and Ellison, 1999; Gottesman and Morey, 2006) investigate the relationship between manager education and mutual fund performance. There have been no studies that address the effect of employees' post-secondary degrees on green portfolio of mutual funds, investment performance, and mutual fund flows. Educational data from LinkedIn profile includes the user's past degrees and their major for each degree. Furthermore, users typically include a description of the purpose of the degree, for example, to educate environmental engineers. In addition, they also detail the courses and extracurricular activities offered by the degree.

The first step is to extract all post-secondary degrees and subjects related to them from users' profiles. To construct a dummy variable, the degrees are split into two groups: environmental and others, using dictionary-based methods. There have also been instances where a user majored in business but stated that he actively attended ESG activities, in which the user's degree can also be considered ESG-related. Textual descriptions will be classified into environmental and other categories using Natural Language Processing (NLP) algorithms in these cases. An example of educational profile with textual description can be found in Figure 1. If at least one employee with environmental educational background works at the fund family in quarter t, the dummy variable $ESG\_EDU_{i,t}$ for fund i is 1, otherwise 0.

2.5 Employment Profile

The research of (Kempf et al., 2017) indicates that experienced managers outperform in industries where they have gained on-the-job experience. No previous work has focused on employees' professional experience in the fund family. Currently, there are no studies that have examined the effect of employees' professional experience on the fund's environmental portfolio, returns on investments, or flows of mutual funds. With LinkedIn becoming the new medium for job search, mutual fund associates are more likely to provide a short description of each role followed by three to five bullet points of achievements. The profile description may also contain a lengthy statement containing several hundred words, in which the user states the central theme of their work and their goals for the future. As a result of the work history not being current or complete, the descriptions complement the situation. These textual descriptions, along with the job title, can be classified as ESG-related or not using Natural Language Processing (NLP) algorithms. A user is considered an ESG-specialized professional if any of the descriptions is classified as ESG-related by the algorithm. An example of employment profile with textual description can be found in Figure 2. Whenever at least one employee with ESG-relevant professional experience works at a fund family in quarter t, the dummy variable $ESG\_WORK_{i,t}$ for fund i will be 1, otherwise it will be 0.

2.6 ESG Mutual Fund

Annual Sustainable Funds U.S. Landscape Report by Morningstar provides an overview of ESG-focused mutual funds available to U.S. investors. (Raghunandan and Rajgopal, 2022) provides a list of ESG mutual funds and their ticker symbols[8]. The report includes nearly 500 ESG mutual funds from 2010 to the present. Morningstar has started offering an annual report since 2018. To determine when a fund first publicly announced its ESG or sustainable investing focus prior to 2018, the history of the fund family website can be found in the CRSP mutual fund summary data[9]. Thus, I define a mutual fund as an ESG fund when the fund family first lists ESG strategies on its website before 2018. A fund-level dummy variable $ESG\_Fund_{i,t+k}$ is 1 if a fund i has become a ESG fund in quarter t, for for $k \geq 0$, otherwise 0.

2.7 ESG Score

The MSCI ESG KLD database provides multiple dimensions across publicly traded companies[10]. There are more than 100 indicator variables from different angles each year since 1991[11]. This research study considers only environmental, social, and governance variables whose prefixes are ENV, PRO, and C_GOV, respectively, in the data. A yearly ESG score is computed based on the average of the environmental, social, and governance indicator variables. When missing values are present, they are filled with zero. An ESG score at the fund level is derived by matching the CRSP Mutual Fund holdings with MSCI stocks ESG scores[12]. As described by (El Ghoul and Karoui, 2017; Hwang et al., 2021), ESG scores of mutual funds are the value-weighted ESG scores of their portfolio holdings, denoted by $ESG\_Score_{i,t}$ for fund i in quarter t. Fund-level value-weighted environmental $E\_Score_{i,t}$, social $S\_Score_{i,t}$, and governance $G\_Score_{i,t}$

[8] Morningstar compiles the list using information contained in sources such as fund prospectuses and fund websites.

[9] Digital archives of the World Wide Web are available through the Wayback Machine. The website is available at https://archive.org/web/. All public websites' history can be traced on Wayback Machine.

[10] The indicator variables cover areas about Environment, Community, Corporate Governance, Diversity, Employee Relations, Environment, Human Rights, and Product issues.

[11] The full coverage of the Russell 3000 begun in year 2003.

[12] (El Ghoul and Karoui, 2017) has addressed the problem of an unrated significant portion of stocks.

scores are computed based on the variables with corresponding prefixes in the data.

2.8 Fund Flows

The annual fund flows are measured as:

$$Flows_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1} * (1 + r_{i,t})}{TNA_{i,t-1}},$$

where $TNA_{i,t}$ are total net assets for fund i in quarter t, and $r_{i,t}$ is the return of fund i in quarter t.

3 Natural Language Processing

Education and employment experiences from LinkedIn profiles are used to evaluate whether a fund employee is experienced in ESG. Profile information on LinkedIn can be divided into two types: 1) well-formatted information pertaining to degrees and majors, as well as the name of the organization. 2) unstructured text descriptions of personal interests and experiences. To examine the properties of different types of data, two different methods of textual analysis are considered for this research project.

3.1 Dictionary-based Textual Analysis

Despite recent progress in natural language processing (NLP), dictionary-based methods are still powerful for structured data without too many lexical variants. The preliminary data gathered from LinkedIn indicates that Blackrock has 30,000 employees with 72 degrees and 517 majors. Considering the naming convention for degrees and majors is similar, dictionary-based methods are an ideal starting point. The first step is to narrow down a seed dictionary of keywords/phrases relevant to ESG majors, from the structured data of majors and degrees. To facilitate the selection of seed words, Wordnet (Miller, 1995), an English database of synonyms is used. The dictionary is manually checked to ensure that there are no misleading words. Based on major and degree information, the ESG-major dictionary can categorize a LinkedIn user as either

ESG-educated or not.

3.2 Contextualized Text Classification

It has been shown that feature-based models, such as dictionary-based models, do not work well for personal statements with complex semantics (Peters et al., 2018). Examples in Table 1 have demonstrated the complexity of semantics. Recent research shows that pretrained deep learning models like BERT (Devlin et al., 2019) and its variants are more powerful than traditional textual analysis methods, primarily as a result of their ability to incorporate massive amounts of data. In this way, they are able to perform contextual modeling without relying on single words.

As pointed out by (Beltagy et al., 2019; Lee et al., 2020), if there is a distribution mismatch between training data and LinkedIn data, pretrained language models may have undesirable model behavior. In this regard, domain adaptation[13] (Ramponi and Plank, 2020) is of the utmost importance since LinkedIn data distribution is different from that of the general database. Domain adaptation techniques involve using knowledge gained from a training domain with vast data to improve the performance of a model in a target domain with insufficient data. It is necessary to further pretrain a model's parameters in order to fit the target domain. A pretrained BERT model is further trained on all textual personal statements collected from LinkedIn profiles in order to perform domain adaptation. The model can be referred to as LinkedIn-BERT, a BERT model specific to the LinkedIn domain. Consequently, LinkedIn-BERT is capable of performing classification tasks. The project is primarily concerned with identifying whether a LinkedIn user has ESG experience from their personal statements of their work experiences and school activities. In spite of LinkedIn-BERT's ability to solve the domain mismatch problem, the number of labelled data is still insufficient. In deep learning, patterns and features are automatically identified and extracted from large volumes of data without the involvement of humans (LeCun et al., 2015). This makes it an invaluable tool for analyzing large data sets. Recently, it has been demonstrated that the quality of the training dataset determines model performance. Researchers

[13]Figure 3 shows an animation of domain adaptation.

working on human-in-the-loop have proposed a few solutions to the problem of identifying what part of data to label iteratively. In a nutshell, the labeling process involves multiple iterations: create a small training set first, and then refine and expand it sequentially. Firstly, I apply a number of algorithmic techniques (Ratner et al., 2016) to filter high quality data containing seed words from the dictionary built in the last section. Using the heuristic-based training set, LinkedIn-BERT is able to classify the remaining unlabelled data. I then iteratively annotate examples with high uncertainty and add them to the training set to optimize the model's decision boundary. After a number of iterations, the training set has grown to a reasonable size and has become sufficiently diverse to meet the demands of the ESG classification task. The final step is to fine-tune LinkedIn-BERT based on the iteratively refined training set. A satisfactory level of accuracy should be achieved for the ESG classification task on personal statements. The goal of ESG classification task is to give every personal statement a label, ESG or non-ESG, by the fine-tuned LinkedIn-BERT.

4 Methods and Tests

4.1 Funds' Portfolio Choices and ESG Scores

In this subsection, I examine whether the recruitment of ESG-specialized professionals can predict the adjustment of ESG portfolios in mutual funds. In this test, I use a value-weighted ESG_Score at the fund-quarter level and estimate the following regression:

$$\begin{aligned} ESG\_Score_{i,t+1} &= \beta_0 + \beta_1 ESG\_Talents_{i,t} + \beta_2 ESG\_Fund_{i,t} + \beta_3 ESG\_Score_{i,t} \\ &+ Controls_{i,t} + FundFE + YearQtrFE + \epsilon_{i,t}, \end{aligned} \tag{1}$$

where i denotes fund and t denotes a quarter. The term $ESG\_Score_{i,t+1}$ is measured as the value-weighted ESG scores of fund i portfolio in next quarter. Value-weighted $ESG\_Score_{i,t}$ in quarter t is additionally controlled. To evaluate environmental, social, and corporate governance aspects of the fund ESG preference separately, $ESG\_Score_{i,t}$

can also be replaced accordingly with $E\_Score_{i,t}$, $S\_Score_{i,t}$, $G\_Score_{i,t}$. $ESG\_Fund_{i,t}$ that serves as a control variable is an indicator variable, equal to one when fund i has integrated ESG into investment objective at quarter t, otherwise zero.

$ESG\_Talents_{i,t}$ is an indicator variable for the fund's recruitment of ESG-specialized professionals, equal to 1 if either $ESG\_EDU_{i,t}$ or $ESG\_Work_{i,t}$ is equal to one and zero otherwise. To measure different effects of educational background and work experience, $ESG\_Talents_{i,t}$ can be substituted with $ESG\_EDU_{i,t}$ or $ESG\_Work_{i,t}$, respectively.

The fund-level control variables include the natural logarithm of total net assets $LogTNA$, fund expense ratio $ExpRatio$, standard deviation of fund daily returns over the previous month $FundVolatility$, annual fund turnover $Turnover$, and natural logarithm of the months since the first occurrence in the CRSP Mutual Fund database $LogAge$. Fund fixed effects and year-quarter fixed effects are included in all test specifications and standard errors are clustered at the fund level.

It is expected that the coefficient β_1 is positive and statistically significant. One interpretation is mutual funds recognize the contributions made by ESG-specialized professionals in decision making and subsequently adjust their investments towards companies with higher ESG scores, controlling the effect that the funds have publicized their ESG investment strategies.

4.2 Funds' High ESG Portfolio Performance

In this subsection, I examine the correlation between hiring ESG experts and fund performance on high ESG stocks. It is widely understood that ESG funds may underperform their counterparts due to investment constraints. However, whether funds with ESG experts outperform others on high ESG stock has not yet been studied.

As a first step, I categorize stocks with non-zero ESG scores into two ESG groups: Low and High. I choose 50^{th} as breakpoints. Zero and missing ESG scores are classified as Low ESG. The sorting is performed at every quarter t. This test only examines high ESG stocks in funds' portfolios. Funds with no high ESG stock holdings are discarded.

Two widely accepted performance metrics are used to measure fund performance:

quarterly raw return and risk-adjusted return on funds' ESG portfolios using (Carhart, 1997) model[14]:

$$r_{i,t} - rf_t = \alpha_i + \beta_{i1} RMRF_t + \beta_{i2} SMB_t + \beta_{i3} HML_t + \beta_{i4} UMD_t + \epsilon_{i,t},$$

where $r_{i,t}$ is the raw return for fund i's high ESG portfolio and quarter t, rf_t is risk-free rate during quarter t; $RMRF_t$ is the value weighted market return on all NYSE/ AMEX/NASDAQ firms in excess of the risk-free rate; SMB_t is the difference between small and big stock portfolios with the same weighted average book-to-market equity; HML_t refers to the difference between portfolios with high and low book-to-market equity ratios; the momentum factor UMD_t represents the average return on two portfolios with high prior returns minus the average return on two portfolios with low prior returns.

A portfolio of all stocks in high ESG group is also considered for benchmarking. Value-weighted raw return of the benchmark portfolio is denoted by $r_{ESG,t}$ in quarter t, and $\alpha_{ESG,t}$ is the excess return of the benchmark portfolio in quarter t.

A two-step Fama–MacBeth (Fama and MacBeth, 1973) regression approach with Newey–West standard errors (Gil-Bazo and Ruiz-Verdú, 2009) is employed to determine whether hiring of ESG professionals predicts fund performance on high ESG stocks:

$$\begin{aligned} \alpha_{i,t+1} - \alpha_{ESG,t+1} = \beta_0 + \beta_1 ESG\_Talents_{i,t} + \beta_2 ESG\_Fund_{i,t} + \beta_3(\alpha_{i,t} - \alpha_{ESG,t}) \\ + Controls_{i,t} + FundFE + YearQtrFE + \epsilon_{i,t}. \end{aligned} \tag{2}$$

where $\alpha_{i,t+1}$ is four-factor model excess return for fund i's ESG stock holdings in quarter $t+1$ and $\alpha_{ESG,t+1}$ is excess return of the benchmark portfolio in quarter $t+1$. Fund excess return subtracting ESG benchmark return, denoted as $\alpha_{i,t+1} - \alpha_{ESG,t+1}$, captures fund outperformance on high ESG stocks. $\alpha_{i,t} - \alpha_{ESG,t}$ is additionally controlled. In other words, I examine whether funds with ESG experts choose stocks that outperform others in light of high ESG stocks.

When hiring of ESG-specialized professionals is able to predict increasingly funds'

[14]Fama-French common risk factors and the risk-free rate are downloadable from Kenneth French's website.

high ESG portfolio performance by leveraging their expertise, the coefficient β_1 will be positive and statistically significant.

4.3 ESG Talents and Fund Flows

ESG investment strategies suggested by ESG specialists may contribute to performance persistence, thereby predicting positive fund flows. To further investigate this hypothesis, the following specification is tested:

$$\begin{aligned} Flows_{i,t+1} = \beta_0 + \beta_1 ESG\_Talents_{i,t} + \beta_2 ESG\_Fund_{i,t} + \beta_3 Flows_{i,t} + \beta_4 \alpha_{i,t} \\ + Controls_{i,t} + FundFE + YearQtrFE + \epsilon_{i,t}. \end{aligned} \tag{3}$$

where $Flows_{i,t+1}$ is the net flow of fund i in next quarter; excess return $\alpha_{i,t}$ and fund net flows $Flows_{i,t}$ in quarter t are also controlled. It is likely that the coefficient β_1 will be positive and statistically significant if hiring of ESG-specialized professionals can predict positive fund flows. I interpret the result as ESG-specialized professionals assist funds in converting into ESG funds, thus attracting investors with similar investment preferences.

4.4 Endogeneity

I acknowledge that there may be causality issues between the recruitment of ESG experts and the ESG ratings, performance, or flow of mutual funds. In the event that a fund anticipates positive inflows contributed by investors with ESG preferences in the upcoming quarters. The fund will recruit ESG-specialized professionals before receiving inflows in order to establish an ESG investment team. A use of shock-based methods is an effective method for testing the causal connection. Unfortunately, I have not yet been presented with such an opportunity. Consequently, I focus on the predictive power rather than the impact of ESG-specialized professionals recruitment.

References

Becker, Gary S (1964) *Human capital: A theoretical and empirical analysis, with special reference to education*: University of Chicago press.

Beltagy, Iz, Kyle Lo, and Arman Cohan (2019) "SciBERT: A Pretrained Language Model for Scientific Text," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3615–3620, Hong Kong, China: Association for Computational Linguistics, November, 10.18653/v1/D19-1371.

Carhart, Mark M (1997) "On persistence in mutual fund performance," *The Journal of finance*, 52 (1), 57–82.

Chevalier, Judith and Glenn Ellison (1999) "Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance," *The journal of finance*, 54 (3), 875–899.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2019) "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186, Minneapolis, Minnesota: Association for Computational Linguistics, June, 10.18653/v1/N19-1423.

El Ghoul, Sadok and Aymen Karoui (2017) "Does corporate social responsibility affect mutual fund performance and flows?" *Journal of Banking & Finance*, 77, 53–63.

Fama, Eugene F and James D MacBeth (1973) "Risk, return, and equilibrium: Empirical tests," *Journal of political economy*, 81 (3), 607–636.

Gil-Bazo, Javier and PABLO Ruiz-Verdú (2009) "The relation between price and performance in the mutual fund industry," *The Journal of Finance*, 64 (5), 2153–2183.

Golec, Joseph H (1996) "The effects of mutual fund managers' characteristics on their portfolio performance, risk and fees," *Financial Services Review*, 5 (2), 133–147.

Gottesman, Aron A and Matthew R Morey (2006) "Manager education and mutual fund performance," *Journal of empirical finance*, 13 (2), 145–182.

Guedj, Ilan and Jannette Papastaikoudi (2003) "Can mutual fund families affect the performance of their funds?" *Available at SSRN 467282*.

Hwang, Chuan Yang, Sheridan Titman, and Ying Wang (2021) "Investor Tastes, Corporate Behavior, and Stock Returns: An Analysis of Corporate Social Responsibility," *Management Science*.

Kempf, Elisabeth, Alberto Manconi, and Oliver G Spalt (2017) "Learning by doing: The value of experience and the origins of skill for mutual fund managers," *Available at SSRN 2124896*.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015) "Deep learning," *nature*, 521 (7553), 436–444.

Lee, Jinhyuk, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang (2020) "BioBERT: a pre-trained biomedical language representation model for biomedical text mining," *Bioinformatics*, 36 (4), 1234–1240.

Miller, George A (1995) "WordNet: a lexical database for English," *Communications of the ACM*, 38 (11), 39–41.

Peng, Lin, Siew Hong Teoh, Yakun Wang, and Jiawen Yan (2022) "Face Value: Trait Impressions, Performance Characteristics, and Market Outcomes for Financial Analysts," *Journal of Accounting Research*, 60 (2), 653–705.

Peters, Matthew E., Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer (2018) "Deep Contextualized Word Representations," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, 2227–2237, New Orleans, Louisiana: Association for Computational Linguistics, June, 10.18653/v1/N18-1202.

Raghunandan, Aneesh and Shiva Rajgopal (2022) "Do ESG funds make stakeholder-friendly investments?" *Review of Accounting Studies*, 27 (3), 822–863.

Ramponi, Alan and Barbara Plank (2020) "Neural Unsupervised Domain Adaptation in NLP—A Survey," in *Proceedings of the 28th International Conference on Computational Linguistics*, 6838–6855, Barcelona, Spain (Online): International Committee on Computational Linguistics, December, 10.18653/v1/2020.coling-main.603.

Ratner, Alexander J, Christopher M De Sa, Sen Wu, Daniel Selsam, and Christopher Ré (2016) "Data programming: Creating large training sets, quickly," *Advances in neural information processing systems*, 29.

Wermers, Russ (2000) "Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses," *The Journal of Finance*, 55 (4), 1655–1695.

Appendix

Table 1: This table shows the employment profile examples extracted from LinkedIn. The profile description has been preprocessed for readability. The first column is the textual profile description; the second column is ESG-related keywords; the third column is the ESG-related label given to the description. Keywords have been bolded in the description. To determine if a profile description is ESG-related, dictionary-based methods may not work well, as the third example shows here. While the description includes terms such as ESG, the person's past work experience is not relevant to ESG, depending on the context. Using deep learning based methods such as natural language processing algorithms can reduce false-positive cases induced by dictionary-based methods.

Profile Description	Keywords	Label (ESG-related)
Responsible for researching and evaluating economic risk of resource constraints and **climate change**. Collaborating with UNEP FI on ERISC (**environmental risk** integration into sovereign credit analysis) phase3 report. Focusing on stranded asset assessment and petroleum reserves projection.	climate change; environmental risk	YES
Identified companies for corporate engagement encouraging participation in the **Carbon** Disclosure Project's (CDP) water questionnaires	carbon	YES
Leader of BlackRock's Core Institutional Business group; 600+ client organization covering mid-market public and corporate pensions (DB and DC), foundations and endowments. Global multi-product asset management services including private equity and debt, liquid and illiquid alternatives, credit, hedge funds, real estate, infrastructure, smart beta, risk parity, and all public markets. Consultative work in asset allocation, risk budgeting, manager structure optimization, **ESG**, liquidity, OCIO and liability driven investing.	ESG	NO
Head of Responsible Investing, Global Fixed Income, Managing Director, is the Head of Responsible Investing, within Global Fixed Income at BlackRock. Ms. Schulten leads the coordination of **climate risk** evaluation and **sustainable investing** efforts within the fixed income division. She sits on the Global Fixed Income Executive Committee. Previous to BlackRock, Ms. Schulten spent 20 years as a sell side interest rate and options trader, most recently at Goldman Sachs.	climate risk; sustainable investing	YES

Figure 1: A sample educational profile is shown below. According to the profile, the individual studied sustainability during his master's program. Additionally, a brief description, circled in red, of his program and school activities is provided. The textual descriptions can be used to classify if this person is an ESG specialist.

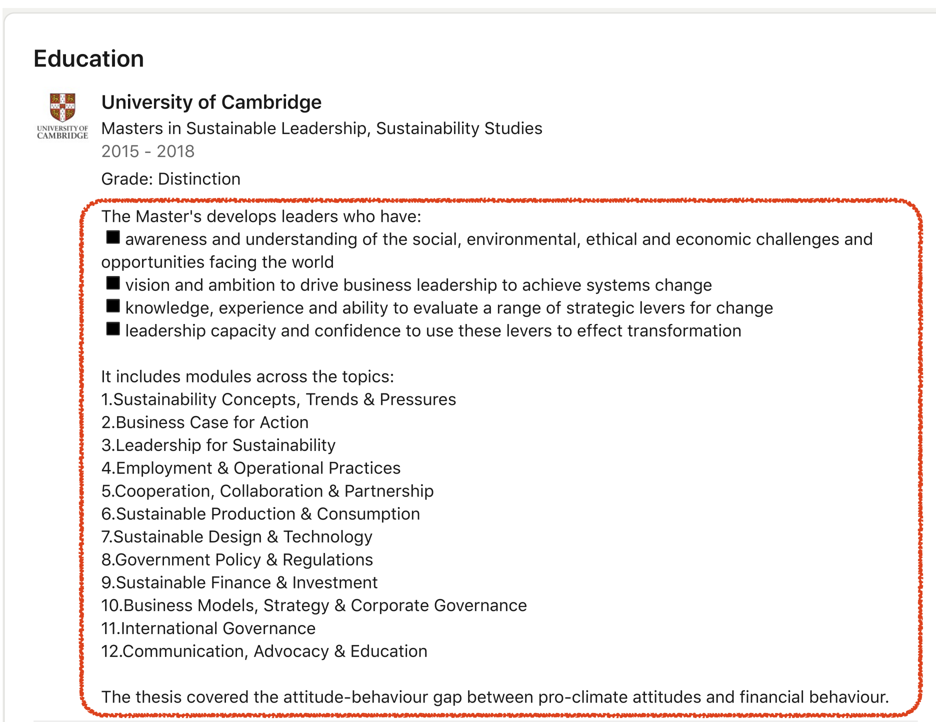

Figure 2: A sample employment profile is shown below. The name is colored in black for privacy. The descriptions, circled in red, of his extensive work experience, can be used to classify if this person is an ESG expert besides job title.

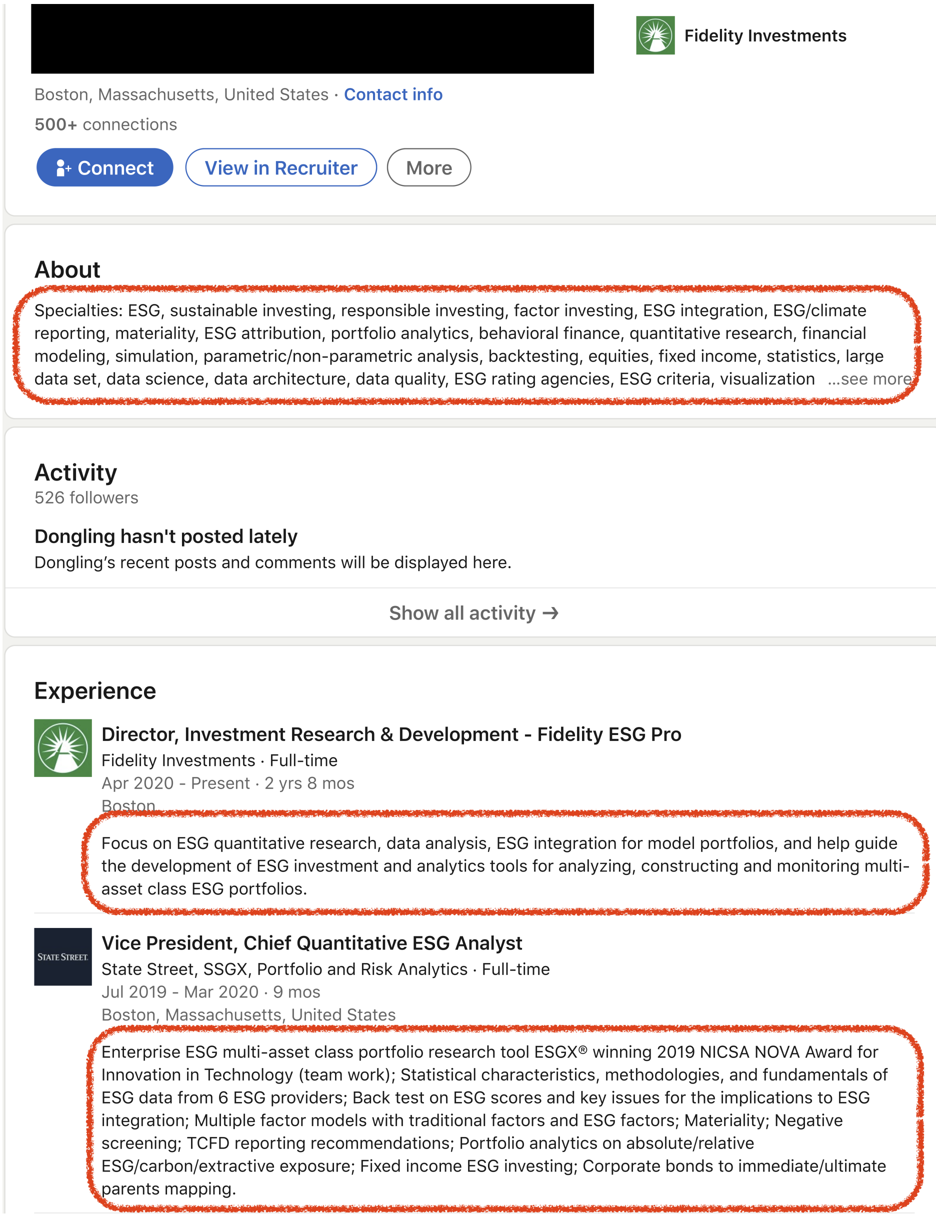

Figure 3: Domain adaptation involves uncovering latent factors common to both source and target domains. The feature space between domains is adapted to reduce marginal and conditional mismatches as a result of adaptation. Common domain adaptation techniques include feature alignment and classifier adaptation.