Jacob Bussell, Giancarlo Finis, Andrew Finn, Rocco Mathews, Isabella Valentino
MGT 411
3/24/25
Semester 2 Midterm Report

Abstract
This project aims to decompose the iShares ESG Aware MSCI USA ETF (ESGU) into distinct Environmental (E), Social (S), and Governance (G) factor portfolios to determine which factor is the primary driver of performance and risk within the ETF. Our hypothesis is that Governance (G) is the most significant determinant of risk-adjusted returns. Using data sourced from Bloomberg covering the years 2020-2024, we reconstructed the cap-weighting methodology employed by iShares and developed an alternative industry-based weighting scheme to facilitate comparative analysis. To test our hypothesis, we implemented a generalized Arbitrage Pricing Theory (APT) factor model to quantify the impact of each ESG component. We will further validate our findings using a Fama-MacBeth factor model to estimate the risk premium associated with each factor. Additionally, we will explore portfolio optimization and the integration of KPI derivatives to refine our analysis and provide deeper insights into the role of ESG factors in investment performance. This research offers a structured approach to isolating and evaluating ESG factors, delivering valuable implications for sustainable investment strategies and financial decision-making.

Background/Reading
The concept of ESG investing has evolved significantly since its origins in the 1970s, transitioning from "ethical investing" to "socially responsible investing" (SRI) and eventually to broader frameworks like "sustainable investing." Markowitz pioneering work posited that responsible corporate behavior could enhance financial performance, a view that contrasted with Friedman's assertion that ESG compliance incurs additional costs, thereby reducing returns. These conflicting perspectives gave rise to three hypotheses: a positive relationship between ESG and financial performance (e.g., Stakeholder Theory, Good Management Theory), a trade-off hypothesis, and a neutrality hypothesis. Over the decades, frameworks such as the UN Principles for Responsible Investment (PRI) have formalized ESG integration, while empirical studies like Friede, Busch, and Bassen's meta-analysis demonstrated consistent nonnegative impacts of ESG on financial outcomes. This historical evolution underscores the significance of ESG factor analysis and its alignment with this project's objectives of quantifying E, S, and G contributions to financial performance.

Naffa and Fain's study, titled *"Performance Measurement of ESG-themed Megatrend Investments in Global Equity Markets Using Pure Factor Portfolios Methodology,"* examines the construction of ESG-themed factor portfolios and their financial performance within global equity markets. The study focuses on nine ESG themes, including energy efficiency, water scarcity, and disruptive technologies, employing advanced methodologies like constrained weighted least squares (CWLS) and the generalized method of moments (GMM). Findings indicate that certain environmental megatrends and governance themes yield significant alphas under simple financial models, but these diminish under more comprehensive frameworks like the Fama-French five-factor model. The study highlights the potential of ESG investments to align with United Nations Sustainable Development Goals (SDGs) without sacrificing returns, although it emphasizes the need to manage transaction costs effectively. This research is directly relevant to the decomposition of ESGU into factor-specific portfolios, providing insights into robust methodological approaches for isolating and evaluating E, S, and G components.

The article *"Exploring the Nexus Between ESG Risk Variations and Investment Preferences"* investigates the relationship between ESG risk metrics and the performance of sustainable ETFs in U.S. and European markets during the COVID-19 pandemic. The study reveals that governance risks are consistently linked to ETF performance resilience, while environmental and social risks exhibit greater variability in their impact. This finding underscores the importance of the governance factor in mitigating financial volatility during crises. The study's emphasis on governance aligns closely with the hypothesis of this research project, which posits that the governance factor is the most significant driver of performance and risk within ESGU. Additionally, the insights into the fluctuating impacts of environmental and social factors suggest the necessity of dynamic and adaptive weighting strategies when constructing factor portfolios.

Friede, Busch, and Bassen's meta-analysis, *"ESG and Financial Performance: Aggregated Evidence from Over 2000 Empirical Studies,"* synthesizes a vast array of research to examine the relationship between ESG criteria and corporate financial performance (CFP). The authors found that approximately 90% of studies report a nonnegative correlation between ESG factors and CFP, with the majority showing a positive relationship. Governance factors, in particular, demonstrated strong associations with financial performance, especially in emerging markets and alternative asset classes like corporate bonds and green real estate. This comprehensive analysis provides empirical validation for the decomposition of ESG factors, supporting the methodological framework of this project. By highlighting governance as a critical driver of financial outcomes, the study strengthens the hypothesis that governance plays a key role in the performance of ESG-themed investments.

Data Collection

The foundation of our analysis relied on detailed data extracted from Bloomberg Terminals, which provided us with insights into the holdings of the iShares ESG Aware MSCI USA ETF (ESGU). Using Bloomberg's dataset, we identified and documented each company's overall ESG score as well as their individual Environmental (E), Social (S), and Governance (G) scores.

These scores, provided by MSCI, served as the cornerstone for constructing our factor portfolios and analyzing the contribution of each ESG factor to the ETF's performance and risk profile.

In addition to the ESG scores, we gathered the individual weights assigned to each company within the ESGU ETF. These weights were critical in determining the influence of each company's ESG performance on the overall fund. This allowed us to tie each company's scores directly to their contribution within the ETF, enabling more precise portfolio construction and replication. These individual E, S, and G weights from Bloomberg will be useful in constructing our industry weights for the portfolios. They will give us a concrete metric to see how heavily each asset is represented in its respective portfolios and how these can be applied to industry weight construction. This methodology will be further outlined later.

This methodology was applied to multiple years to gain an understanding of the progression of ESG scores and their contributions from 2020 to the present. Extracting this vast amount of data by hand proved a timely task but was indeed worth the effort. We now have data from 2020, 2022, 2023, and 2024 that we can use to examine trends over time. We currently have an issue with the holding data for 2021 not being complete, and the source which we attained it from(Blackrock website) is no longer available. The holdings data appears to only change marginally from year to year, so we were able to collect data for most of the stocks that are likely in 2021. We will need to use alternate data likely from Bloomberg to get the rest of the holdings and assign industry keys to each ticker. The more data, the more confident we feel in our evaluation methodology which we will dive deeper into shortly.

To provide context and better interpret the ESG scores, we normalized the E, S, and G scores by industry. We did this because the way Bloomberg scores companies on E, S, and G depends on what industry the company is in, and we want to be able to incorporate top ESG performers from all industries, not just industries Bloomberg scores higher. Doing this, as well as for applying industry weighting in our factor/portfolio construction, is why having the relevant industry each company is in is important. This is a big part of what we're missing from 2021 data and need to amend in the near future.

Factor Construction Experimentation

Factor Construction Methodology

To construct these individual factors is a difficult task, as there's little prior written information regarding *individual* E, S, and G factors. Given that we are trying to create an "E" factor portfolio(one example), we need to ensure the stocks in this portfolio have higher E scores relative to the stocks we're looking at, and lower relative scores in the other factors. This is to make sure each individual factor portfolio effectively represents the given factor more than the others. While this seems straightforward it's difficult to accomplish in a way that's mathematically robust.

We previously had been experimenting with the following methods to construct these factors.

Min-Max Scaling
The first method is the most straightforward, normalizing and min-max scaling the values in each score vector. After these adjustments, whichever score is biggest for each stock will go into that factor portfolio. Then, for each portfolio we take the tickers of the top *n* values to reduce the size of the portfolios.

The drawback of this is if a certain vector of scores has an outlier on the upside it can make the other scaled scores very low. This causes the selection of tickers for that factor to be very infrequent, only happening when the given ticker performs incredibly poorly in the other two factors. The selection of *n* is also relatively arbitrary.

Arbitrary Algorithm
The second method selection via a very simple decision based algorithm that we came up with. Say we are constructing factor portfolio E. For a given ticker of E-score in the top *n%* of E-scores *and* S and G-scores below their respective S and G means, the ticker is then added to the E factor portfolio. This ensures that all stocks in a given factor portfolio overachieve in that factor while underachieving in the others.

This method seems to work pretty well, with the primary drawback being that again the selection of *n* is somewhat arbitrary, but is becoming less arbitrary as we discover concerns with correlation between factors.

Linear Optimization
We defined a linear programming problem that uses binary decision variables to choose which stocks will maximize its respective E, S, or G score subject to defined constraints. For example, when constructing the E portfolio, we will choose the best 10-25 stocks that best maximize its respective E score while keeping its S and G scores below the overall ESG index's S and G mean values. This will specify the stocks that are most heavily weighed by E while fitting our desired constraints. Then, the same methodology will be followed for the S and G factors. The formulation of the linear program for the E portfolio will look like this (Cornuejols)

We have stocks S_i for $i = 1, ..., N$, $N = 275$. Each stock has three scores, E_i, S_i, G_i.

We can create our E-Portfolio by solving the following linear programming problem.

$$max \frac{1}{e^T k} k^T E$$

$$\begin{aligned} s.t. \qquad \frac{1}{e^T k} k^T S &\leq \frac{1}{N} e^T S, \\ \frac{1}{e^T k} k^T G &\leq \frac{1}{N} e^T G, \\ a \leq e^T k &\leq b, \\ k &\in \{0, 1\} \end{aligned}$$

We can then do this for the other two factor portfolios.

After reflecting a bit on our different methods we realized that our "arbitrary algorithm" we made effectively solves this LP problem, as it removes tickers that don't meet the constraints and then selects the top scores, effectively maximizing the mean. Therefore we will be using "both" of these methods as they are effectively the same. This is somewhat similar to the method used by Naffa and Fain, cited below, but is adjusted to our individual-scores use case.

We are running into some issues with this method of factor construction that weren't apparent in our previous experimentation. When we used this method on just the 2024 data we didn't have any issues with correlation between factors, but as we've collected earlier data this correlation seems to increase the further we go back. We believe that adding more companies to each factor will improve this, but need to implement it to be sure.

Portfolio/Factor Weighting Methodology

Industry Weighting

To stay consistent with our main goal of examining how scores are determinants of performance, we aimed to find a way to normalize the scores in context of each other. In doing so, we examined how the industries of each stock making up each portfolio is impacted by ESG. Through separating industries and examining these scores impacts to different industries, we created a weighting methodology that adjusts individual stock weights according to their industry's relative contribution to overall ESG performance. Starting with the previously defined portfolios focused on E, S, and G factors, we first grouped the constituent stocks by their respective industries using the industry classification key described in the Data Collection section. We then calculated the total weighted score for each industry by summing respective scores of each stock belonging to that particular sector. This provided a measure of each industry's overall ESG contribution within each respective portfolio.

Next, we found the proportional weight of each industry by dividing its total industry score by the sum of all industry scores represented in the portfolio. These fractions are the weights that would then be assigned each stock within its respective industry. This method makes industries with higher total ESG performance have a greater influence on the portfolio, while also ensuring that stocks in the same industry are given the same weight. By embedding industry context, our portfolios would be both performance sensitive and structurally aligned with sector contributions. This methodology was implemented the same way for each pillar E, S, and G for each year of data (2020, 2022-2024). This leaves us with 4 sets of industry weights for each asset that makes up the E, S, and G portfolios for those years.

This industry weighting methodology can be visualized step by step as so:

Let:

- i denote a stock
- k denote an industry
- $\mathcal{I}(i)$ be the industry of stock i
- S_i be the ESG-weighted score for stock i, defined as:

$$S_i = \text{Score}_i \times \text{Weight}_i$$

Steps

1. **Industry Scoring:** Compute the total ESG score per industry:

$$\text{IndustryScore}_k = \sum_{i \in k} S_i$$

2. **Normalization:** Compute the total ESG score across all industries:

$$T = \sum_{\text{industries } k} \text{IndustryScore}_k$$

3. **Weight Assignment:** Each stock i is assigned a weight:

$$w_i = \frac{\text{IndustryScore}_{\mathcal{I}(i)}}{T}$$

Cap Weighting

BlackRock’s iShares ESG Aware MSCI USA ETF (ESGU) employs a modified market capitalization weighting strategy designed to balance exposure to large-cap U.S. equities while integrating environmental, social, and governance (ESG) considerations. Rather than using a pure cap-weighted approach, ESGU weights its constituents based on their market capitalization as derived from its parent index, the MSCI USA Index, but introduces an ESG tilt through an optimization process. This process aims to maximize the portfolio's overall ESG score by overweighting companies with higher ESG ratings and underweighting or excluding those with poor ESG profiles or involvement in controversial business activities. Specific constraints are applied to ensure the ETF maintains a risk and return profile closely aligned with its benchmark. These include sector and country weight constraints (typically within ±5% of the parent index) and a tracking error constraint (e.g., 0.5% for ESGU) to limit deviation from the benchmark. Additionally, companies involved in controversial weapons, thermal coal, oil sands, civilian firearms, or with severe ESG controversies are excluded outright. This methodology allows ESGU to deliver a portfolio that closely mirrors the broad U.S. equity market while enhancing its ESG characteristics—achieving, according to BlackRock, an average ESG score uplift of 15% and a 42% reduction in weighted average carbon intensity compared to its parent index. In our project, we replicated this weighting framework to construct accurate factor portfolios and maintain consistency with how ESG considerations are operationalized within ESGU.

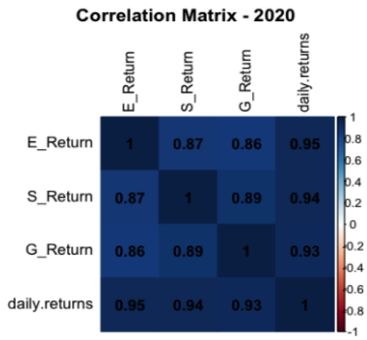
Comparison

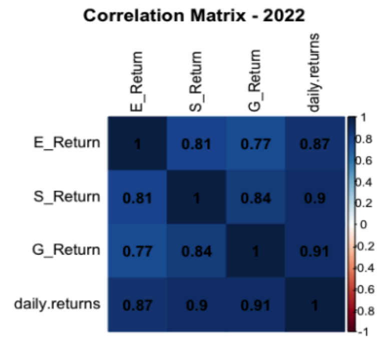
As depicted in the E Portfolio Composition graphic to the right, the industry weighting offers a diversified portfolio as opposed to the Cap Weighting's stark distribution. The Industry weighted portfolio composition aligns more with the team's research goals, as we feel that our portfolio should not be dictated by large market cap stocks such as META and AMZN, as this defeats the purpose of examining the intrinsic value of the scores themselves. By having large cap companies completely make up the portfolio, we lose the true question at hand: Which **score** is the biggest factor in performance? Not which large cap stock with a high score drives the portfolio. The industry weight methodology smoothes out this problem and lets the industries' contributions to ESG scores define the makeup of each portfolio.

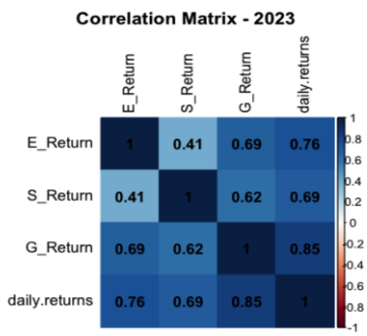
E Portfolio Comparison						
2020				2022		
Ticker	Industry_Weight	Cap_Weight		Ticker	Industry_Weight	Cap_Weight
CRM	14.25%	21.04%		EQIX	12.96%	5.50%
IBM	14.25%	11.56%		PEG	7.47%	2.91%
MCO	3.40%	5.68%		LULU	7.15%	3.90%
IFF	3.46%	1.21%		META	6.51%	30.46%
SBAC	3.46%	3.26%		J	5.23%	1.44%
D	13.40%	6.39%		HUBS	14.56%	1.34%
EL	7.99%	10.02%		CRM	14.56%	12.62%
MA	14.25%	37.06%		TWLO	14.56%	0.86%
PEG	13.40%	3.07%		VMW	14.56%	4.97%
PVH	12.13%	0.70%		JPM	2.45%	35.97%
2023				2024		
Ticker	Industry_Weight	Cap_Weight		Ticker	Industry_Weight	Cap_Weight
CARR	20.31%	2.49%		ECL	7.77%	1.50%
PEG	9.24%	1.55%		J	10.00%	0.38%
LULU	4.83%	3.29%		LULU	14.67%	1.07%
META	4.76%	46.14%		HUBS	9.06%	0.82%
HUBS	7.03%	1.48%		TGT	5.12%	1.41%
TXT	20.31%	0.80%		META	4.97%	34.47%
JPM	3.89%	24.94%		AMZN	14.67%	53.34%
GE	20.31%	7.05%		RCL	14.67%	1.41%
WDAY	7.03%	3.66%		WDAY	9.06%	1.52%
AMGN	2.29%	8.61%		GE	10.00%	4.09%

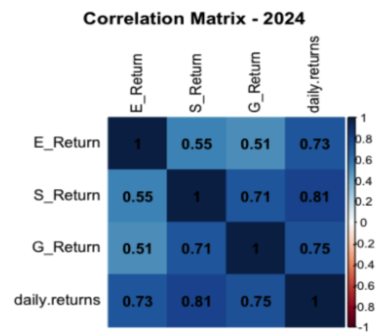
<u>Examining Hypothesis</u>

APT Model Results









The APT model was to act as a benchmark for our project, through regressing each factor portfolio's returns to the return of the MSCI ESG ETF, we could determine which factor was most closely correlated too, and in turn, made up most of the ETFs performance. The first test we conducted using the APT model was on the 2023 data, resulting in the G factor having the highest correlation, complimenting our initial hypothesis that the G factor led the others. However, after running the results for the 3 other years, (note 2021 data is missing due to data collection issues), we found that no one factor dominated the others. This, potentially a result of macroeconomic precedence over ESG, given the current and past 5 years domestic and global environment.

The high correlations given from the model led us to believe our methodologies regarding the portfolios may need to change. Firstly we used cap weighting just as the MSCI ESG ETF did, however we are going to rerun the regression using sector weighting which we believe should more fairly represent each company in the portfolio. Secondly the portfolios only consist of 10 stocks each, and due to data issues, 2 of the portfolios are missing one company. By increasing the number of stocks in each portfolio, correlations should decrease, especially in the factor portfolios that through industry weighting, potentially higher weight more volatile industries. This should in turn better represent the "real" correlations, and once we run the APT model for 2021, potentially lead to different results, proving or disproving our hypothesis.

Fama-MacBeth Regression Results

The method for the Fama-MacBeth two-stage regression can be seen below. This allows us to construct a “Risk Premium” for each factor that we’ve created. What these risk premiums represent isn’t the exposure of a given stock/portfolio to a factor, but rather given exposure to the factor, does that result in higher or lower returns on average for an investor.

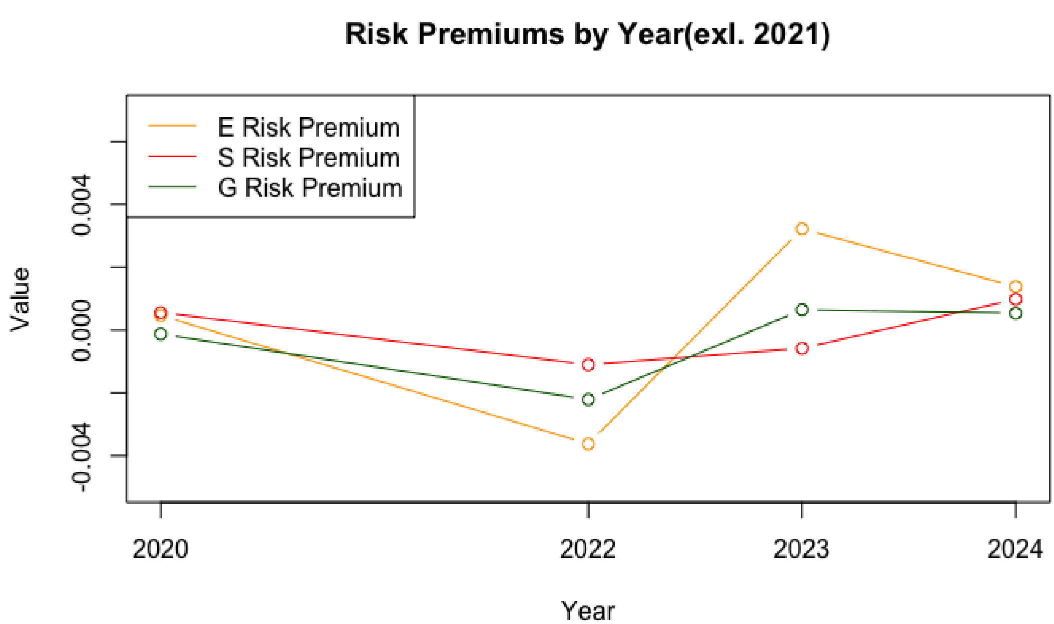
1. First regress each of n asset returns against m proposed risk factors to determine each asset's beta exposures.

$$
\begin{aligned}
R_{1,t} &= \alpha_1 + \beta_{1,F_1} F_{1,t} + \beta_{1,F_2} F_{2,t} + \cdots + \beta_{1,F_m} F_{m,t} + \epsilon_{1,t} \\
R_{2,t} &= \alpha_2 + \beta_{2,F_1} F_{1,t} + \beta_{2,F_2} F_{2,t} + \cdots + \beta_{2,F_m} F_{m,t} + \epsilon_{2,t} \\
&\vdots \\
R_{n,t} &= \alpha_n + \beta_{n,F_1} F_{1,t} + \beta_{n,F_2} F_{2,t} + \cdots + \beta_{n,F_m} F_{m,t} + \epsilon_{n,t}
\end{aligned}
\quad [1]
$$

2. Then regress all asset returns for each of T time periods against the previously estimated betas to determine the risk premium for each factor.

$$
\begin{aligned}
R_{i,1} &= \gamma_{1,0} + \gamma_{1,1} \hat{\beta}_{i,F_1} + \gamma_{1,2} \hat{\beta}_{i,F_2} + \cdots + \gamma_{1,m} \hat{\beta}_{i,F_m} + \epsilon_{i,1} \\
R_{i,2} &= \gamma_{2,0} + \gamma_{2,1} \hat{\beta}_{i,F_1} + \gamma_{2,2} \hat{\beta}_{i,F_2} + \cdots + \gamma_{2,m} \hat{\beta}_{i,F_m} + \epsilon_{i,2} \\
&\vdots \\
R_{i,T} &= \gamma_{T,0} + \gamma_{T,1} \hat{\beta}_{i,F_1} + \gamma_{T,2} \hat{\beta}_{i,F_2} + \cdots + \gamma_{T,m} \hat{\beta}_{i,F_m} + \epsilon_{i,T}
\end{aligned}
\quad [1]
$$

Eugene F. Fama and James D. MacBeth proposed this methodology in 1973, and it involves first running a regression of assets returns against factor returns to determine exposure of the assets to each factor, then running a cross-sectional regression to “price” the exposure to these different factors. The final values for the risk premiums are the means over T of the coefficients donated as gamma in the image above. Even though these risk premiums don't measure exposure to a given factor, it does allow us to see if a factor on average affects performance, and helps guide us towards answering our hypothesis.



We implemented this model using data from 2020-2024(excluding 2021), and got the results depicted above. These results show near-zero expected increase or decrease in returns as a result of exposure to these factors. Using language from the Fama/MacBeth paper, these factors don't appear to be "priced". It's important to repeat that while these values don't measure exposure, they do measure if exposure results in higher or lower average returns.

To verify these results we do need to examine our factor construction, the level of correlation for previous years is too high, and the small number of companies involved in each factor can affect the results seen in this model.

Next Steps

Evaluating Factor Construction

Given the issue we've uncovered with correlation between factors for older dates, we need to re-examine how we are constructing our factors/portfolios. A possible solution is that we used cap-weighting for both of the results above instead of industry weighting. The industry weighting more evenly spreads out the weight in the factor, which avoids the issue of say the E factor and the S factor each having a heavily weighted big-tech stock, which would cause them to be highly correlated. If this doesn't work we also believe that just increasing the number of companies in each factor should also minimize the correlation between the factors, but also lead to including lower companies in their respective E, S or G performance.

Further Research

Once we have completed answering our hypothesis, we can experiment with further research questions such as using linear programming to minimize various risk metrics given selected ESG constraints. We can experiment with a variety of constraints, as well as using returns from various times of volatility to see the effects of our optimization.

For each ETF, we will find the log returns of each asset over a given time period and designate changing decision variables between 0 and 1 as weights. Then, we will use these weights to determine which assets we will invest some fraction of wealth into to max/minimize different objective functions. This fraction of wealth will be determined by a solver (Excel, Python, R, etc.) based on these changing weights. The objective functions mentioned could include maximizing the expected return minus AVaR at a given alpha, or maximizing expected return minus semi deviation, etc.

Conclusion

After implementing our two models to assess the effects of E, S, and G factors on asset performance, our preliminary results show the factor the ETF is most exposed to varies, and that exposure to these factors doesn't have a significant effect on asset prices. This points away from our hypothesis that G is the factor with the most impact, but rather none of these factors are consistently impactful. To ensure these results are accurate we need to get results for 2021 and improve our factor construction.

Given time we look to experiment with various applications of our decomposition, such as optimizing portfolios with constraints on scores and factor hedging. To hone in on which strategy gives us the greatest values for said objective functions above, we will use the objective functions as a benchmark to decide which of the three portfolio construction methods performs best. Essentially we will use the three main construction methods to make three different sets of E, S, and G portfolios. Then, we will do the same asset weighting strategy mentioned above for each of the three different constructed sets. Finally, we will compare objective functions to see which portfolio construction strategy best maximizes expected returns minus VaR, or some other metric.

Works Cited

BlackRock. "IShares ESG Aware Equity Etfs." *iShares by BlackRock*, www.ishares.com/us/literature/product-brief/ishares-esg-aware-equity-brief.pdf. Accessed 25 Mar. 2025.

Friede, Gunnar, et al. "ESG and Financial Performance: Aggregated evidence from more than 2000 empirical studies." *Journal of Sustainable Finance & Investment*, vol. 5, no. 4, 2 Oct. 2015, pp. 210–233, https://doi.org/10.1080/20430795.2015.1118917.

G. Cornuejols, R. Tutuncu, Optimization Methods in Finance, Cambridge University Press.

Gudmundarson, Ragnar L., and Gareth W. Peters. "Assessing Portfolio Diversification via Two-Sample Graph Kernel Inference. A Case Study on the Influence of ESG Screening." *PLoS ONE*, vol. 19, no. 4, Apr. 2024, pp. 1–59. *EBSCOhost*, https://doi-org.stevens.idm.oclc.org/10.1371/journal.pone.0301804.

Landi, Giovanni Catello, et al. "Exploring the Nexus between ESG Risk Variations and Investment Preferences: Insights from Sustainable ETFs during the COVID-19 Era." *Socio-Economic Planning Sciences*, vol. 95, Oct. 2024, p. N.PAG. *EBSCOhost*, https://doi-org.stevens.idm.oclc.org/10.1016/j.seps.2024.102039.

Naffa, Helena, and Máté Fain. "Performance Measurement of ESG-Themed Megatrend Investments in Global Equity Markets Using Pure Factor Portfolios Methodology." *PLoS ONE*, vol. 15, no. 12, Dec. 2020, pp. 1–34. *EBSCOhost*, https://doi-org.stevens.idm.oclc.org/10.1371/journal.pone.0244225.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*(3), 607–636.