

Executive Summary

Introduction

Sustainable, socially responsible, and in particular, ESG based investing is undisputedly one of the biggest and most significant trends in finance for decades. Notwithstanding, both practitioners and academics are far from reaching an agreement on the financial performance of sustainable investments. The contradictory findings can, to some extent, be attributed to the fact that results are highly sensitive to various influences. In addition to the subjective nature of ESG ratings, studies are affected by ESG's dynamic character, its dependence on the market under study, and information overlaps with well-known return-driving factors. Parallel to the discussion on the relationship between sustainability and financial performance, a contentious debate is being held on the optimal integration of ESG criteria into the investment process. Whereas in the past the focus was on the exclusion of sin stocks, these days more sophisticated methods prevail.

This thesis examines an unprecedented long-term data-set derived from machine learning algorithms. It allows a prolonged view on ESG developments in the US. Alternative markets are examined over a shorter time horizon. Using characteristic and factor balanced portfolios, I investigate whether sustainability constitutes a profitable risk factor or characteristic. Risk factors are calculated to be cross-sectionally independent. Concluding, an optimized investment strategy for the integration of ESG criteria in the equity investment process is proposed. Results are derived for aggregated ESG scores as well as the sub-pillars Environment, Social, and Governance.

Methods

Whoever wants to conduct convincing research on sustainability in financial markets is confronted with the lack of long-term observations. In order to counteract this issue, I combine two data-sets to perform an extrapolation using machine learning algorithms. Five algorithms and two strategies for the derivation of the long-term data-set are tested for their performance. I use the best working algorithm, the Gradient Boosting Machine, to extrapolate both ESG as well as E, S, and G scores back to 1997. Gradient Boosting is a sequential ensemble technique, which starts with the most simple predictor - the mean of the variable to predict. The algorithm repeatedly fits trees to the residuals of the previously estimated model, gradually reaching the final model. To counteract overfitting, each intermediate model is scaled to have little impact on the final prediction. The subjectivity of