Erasmus University Rotterdam
Erasmus School of Economics
Master Thesis Econometrics & Management Science

Linking equity premiums to carbon emission costs through changes in the probability of default

Thomas Tichelman (539955)

Supervisor:	Pierluigi Vallarino
Second assessor:	Michel van der Wel
Company supervisor:	Jelle Thijssen
Date final version:	30th April 2024

Abstract

This paper performs an analysis of the effect of several carbon emissions tax scenarios on equity premiums, using a non-classical equity premium prediction model. The approach uses both real-world and risk-neutral probabilities of default as inputs for a novel econometric forecasting model based on the Merton framework to predict equity premiums. From an analysis of the predictive performance of this proposed Merton model in comparison to several benchmark models, we validated the use of the Merton framework in this context. This justified its use in the carbon tax scenario analysis for the overall set of stocks, as well as subsets based on company sectors and carbon emission segments. The analysis, conducted by adjusting the models inputted real-world probabilities of default based on carbon emission tax scenarios, reveals a negative relation between carbon emission taxes and equity premiums. In a scenario where the carbon emission tax increases by €86.42 per tonne of CO2, the sector equity premiums decrease between 1.41% and 12.10%, which would substantially affect valuations and investment decisions. The impact is largest for the sectors wholesale and retail trade; and electricity, gas, steam. The sensitivity of sector equity premiums to carbon emission taxes is dependent on carbon emissions but also on the way their financials are structured. The analysis of equity premiums within carbon tax scenarios reveals that the high emission segment, comprising companies categorized based on their emissions per unit of turnover, is the most affected. In the event of a €86.42 per tonne of CO2 tax increase, the high emission segment experiences a significant 10.04% decrease in equity premium. In contrast, the medium and low emission segments experience relatively minor decreases of 0.59% and 0.10%, respectively.

Contents

Chapter 1

Introduction

This paper examines the impact of different carbon emission tax scenarios on equity premiums through the use of a non-classical equity premium forecasting model. Utilizing historical financial statements and derivative instruments data as input for a novel econometric forecasting approach we forecast equity premiums. The approach was proposed and used by Kaserer and Berg (2008) to estimate the size of the (market) equity premium. This chapter explains the core concepts of the paper; gives the theoretical context and relevance; outlines the methodology used; and a brief overview of the content of the paper can also be found at the end of this chapter.

1.1 Equity premium

Estimating equity premiums is academically important, providing a framework to understand the risk-return relationship, shape investment strategies, and formulate financial models. van Ewijk, de Groot and Santing (2012), seek to review existing literature through a meta-analysis, aiming to understand how the size of the equity premium is influenced by its measurement methodology, its evolution over time, and its variation across different regions in the world. This paper assumes that the equity premium changes over time and defines it as the expected excess returns on equity. Throughout this paper, the expected excess returns are referred to as the equity premium. It is important to note that these expected excess returns may not necessarily be related to the market price of certain risks, which is a common interpretation in the literature. Several studies have suggested that it is possible to forecast the equity premium to a substantial degree by using various financial and economic indicators, such as dividend yields and book-to-market ratios (Fama & French, 1988; Kothari & Shanken, 1997).

Welch and Goyal (2007), conduct a comprehensive assessment of equity premium prediction up to 2005 using several classical predictive variables. They employ consistent methods, sample periods, and estimation frequencies for each variable, revealing substandard predictive power of various variables both in-sample and out-of-sample. They investigated classical predictive variables in the following classes: dividend-price ratio and dividend yield; earnings price ratio and dividend-earnings (payout) ratio; interest and inflation rates; book-to-market ratio; consumption, wealth, and income ratio; and aggregate net issuing activity. They argue for the instability and short-lived accuracy of the predictions based on their out-of-sample forecasts, especially when compared to the historical average benchmark.

Later studies examined if alternative predictors could offer more accurate and stable forecasts of the equity premia. One interesting class of predictors, which have been extensively investigated in the literature, are financial derivatives. Since these securities are inherently forward-looking, they may offer valuable insights into the future return distribution of underlying assets. Alexandridis, Apergis, Panopoulou and Voukelatos (2023), investigate the predictive power of twelve strategy benchmark indices quoted by the Chicago Board Options Exchange (CBOE). They find utilizing option-based predictors yields substantial economic benefits for a mean-variance investor when compared to the historical average benchmark. Similarly to equity options, credit default swaps (CDS) have also been used to predict equity premiums (Han & Zhou, 2011; Hilscher, Pollet & Wilson, 2015). This paper also uses CDS data as an instrument to extract risk-neutral probabilities of default which intern are used to predict equity premiums. More details about utilizing CDS data to predict equity premiums can be found in Section 1.2.

Next to financial derivatives, there is a vast list of other alternative predictors. Some approaches (see Baetje and Menkhoff 2016, among others) find that technical indicators such as moving-average rules, momentum rules, and volume rules yield stable out-of-sample predictions of the equity premium. Other papers aim to use data extracted from newspaper articles as predictors of equity premiums. For example, Adämmer and Schüssler (2020), find their method has substantial utility gains for mean–variance investors. Another predictor, explored by Wang, Pan, Liu and Wu (2019), is oil price increases. They find that price increases rather than price changes in oil hold predictive power over equity premiums. Moreover, they find that this performance is robust to the choice of sub-sample, which is an advantage over more traditional predictors.

Besides alternative predictors, there is a rich literature investigating different types of econometric models to improve equity premium forecasts. For instance, Rapach, Strauss and Zhou (2009), demonstrate that combining forecasts across models and variables leads to notable improvements in forecast accuracy. This stands in stark contrast to the limited effectiveness of individual forecasts, both from statistical and economic perspectives. Another approach is to place economic or statistical constraints on the equity premium forecasts. Li and Tsiakas (2017), for example, puts constraints on both. They find this results in an increase in certainty equivalent returns for a mean–variance investor when compared to the historical average benchmark. This paper makes a contribution to the equity premium forecasting literature by employing an existing econometric model that has not been previously utilized for predicting equity premiums. The model was first introduced by Kaserer and Berg (2008) and originally used to measure the size of the (market) equity premium but is now used for forecasting. The model is based on a structural framework and utilizes real-world and risk-neutral probabilities of default. This model is first introduced in Section 1.4 and is later discussed in more detail in Chapter 2.

1.2 Credit default swaps

This paper indirectly uses credit default swap (CDS) data to forecast the equity premium through the risk-neutral probability of default. In the current literature, there is mixed evidence for the predictive power of CDS data on equity performance. Han and Zhou (2011) find that the slope of the CDS curve (measured by the spread between the 5-year and 1-year CDS rates) serves as a predictor for future stock returns across various time-frames extending up to 6 months. They also find that the CDS slope primarily predicts returns for stocks that encounter substantial arbitrage costs. They argue two components play into these results: "Expectation hypothesis" and "Slow information diffusion".

First of all, the expectation hypothesis is the anticipation among investors that the firm's creditworthiness will decline in the forthcoming periods. This parallel resembles the expectation hypothesis governing the term structure of interest rates in a default-free context. In this analogy, a higher long-term rate in the present compared to the short-term rate implies an expectation of elevated short-term rates in the future.

Secondly, the case for slow information diffusion is supported by the following argument. The negative correlation Han and Zhou (2011) have observed between the CDS slope and future stock returns implies that the information conveyed by the CDS term structure isn't entirely reflected in current stock prices. If it were, they would anticipate that stocks associated with a high current CDS slope would yield higher expected returns, as they are at a greater risk of experiencing credit deterioration in the future. Nevertheless, Han and Zhou (2011) find that stocks with elevated CDS slopes tend to yield lower returns in subsequent periods. This phenomenon aligns with the notion that it takes time for information to flow from the CDS market to the equity market.

Contrary to the findings of Han and Zhou (2011), Hilscher et al. (2015) present evidence that, at daily and weekly intervals, equity returns lead credit protection returns, while the reverse is not observed. They claim this suggests informed traders focus more on the equity market than the credit default swap (CDS) market, which is in line with theories of market selection considering transaction costs. Additionally, they find that credit protection returns respond faster during salient news events, supporting explanations related to investor inattention.

The discrepancy in evidence regarding the predictive influence of the CDS market on the equity market may be due to the difference in the predictors used in each paper to determine the relation between the two markets. Han and Zhou (2011) focuses on the predictive power of the slope of the CDS curve on equity returns whereas Hilscher et al. (2015) investigate the predictive power of lagged credit protection returns on equity returns.

This paper indirectly uses CDS to predict equity premiums by first making a monotonic transformation of the CDS spreads to estimate the risk-neutral probability of default and subsequently using that as input for the equity premium estimate model, as further described in Section 2.4. The model used does not directly assess the predictive capability of the CDS market on the equity market. However, it can be reasoned that the model performs optimally when there is a flow of information from the CDS market to the equity market.

1.3 Carbon emission tax

Climate change risk can broadly be divided into two main risk categories. On the one hand, there are physical risks which include acute risks such as droughts, floods, hurricanes and wildfires but also chronic effects such as rising sea levels, the expansion of tropical pests and diseases and loss of biodiversity. See, for example, Attoh, de Bruin, Goosen, van Veldhoven and Ludwig (2022), which explore physical climate risk in Dutch real estate. On the other hand, there are transition risks which include public policy changes, disruptive innovation or investor and consumer attitudes regarding more environmentally friendly investments and products. This is exemplified by the paper of Ferentinos, Gibberd and Guin (2021), which quantifies the effect of a policy intervention addressing carbon emissions in the English housing market. This paper focuses on the transition aspect of climate risk. More specifically, it examines the impact of carbon emission taxes on equity premiums.

The relevance of carbon emission taxes is growing with governments worldwide intensifying their efforts to combat climate change. These effects could significantly impact the valuations of different companies, thereby influencing their stock prices. Several papers indicate an increase in the importance of climate risk in asset pricing since the Paris Agreement was signed on 12 December 2015 (Bua, Kapp, Ramella and Rognone (2022); Bolton and Kacperczyk (2021)). For example, Bolton and Kacperczyk (2021), find that the premium associated with emissions is larger during the 2016-2017 period when compared with the period 2005-2015. They argue that the Paris Agreement possibly heightened both the recognition of emission-related risks and the anticipation of regulatory measures aimed at decreasing carbon emissions. It is therefore reasonable to focus on a sample starting in 2016 until the present.

The link between environmental, social and governance (ESG) and corporate financial performance (CFP) is well researched. Friede, Busch and Bassen (2015), performed a second-order meta-analysis in which they aggregated over 2200 empirical studies. If we assume a close connection between ESG and carbon emissions, this paper implies that companies with lower carbon emissions financially outperform those with higher emissions[1]. This effect could be due to different factors. One reason is the fact that carbon emissions are an externality. Presently, there exists substantial uncertainty concerning the extent to which different firms will need to absorb the expenses associated with their carbon emissions in the future. Luo and Tang (2014) performed a study in Australia that provides evidence that the announcements of proposed carbon taxes have a generally adverse effect on shareholder wealth, as indicated by abnormal negative returns. The adverse influence differs among sectors, with the most pronounced impact observed within the materials, industrial, and financial sectors. They argue that the stock market's response to a proposed regulation depends on the alteration in the likelihood of the regulation's adoption and the monetary value of the anticipated effect of the regulation on shareholder wealth.

By assessing the implications of different taxation levels, this paper seeks to provide practical insights into potential changes in stock prices and corporate valuations. This, in turn, contributes

[1]They found that approximately 90% of these analyses find a non-negative relationship between ESG and CFP, while the remaining ones find a significant negative relationship. It is even more notable that the vast majority of these studies indicate a significant positive ESG-CFP relation.

to a better understanding of the practical outcomes of climate-related policies on the business landscape. This paper aims to assess the potential effects of carbon taxation on individual corporations, as well as on entire sectors and carbon emission segments.

1.4 Merton estimator

This paper uses a non-classical approach to estimating equity premiums and their sensitivity to carbon emission costs. The method links the equity premium to carbon emission costs through changes in the probability of default under different levels of carbon taxes. The structural model, proposed by Merton (1974), is traditionally used to calculate the real-world probability of default (PD) by modeling a company's equity as a call option on its assets and using input such as expected returns and volatility. Kaserer and Berg (2008), use this Merton framework to derive an estimator of the Sharpe ratio. The Sharpe ratio estimator relies mainly on the input real-world and risk-neutral probability of default.

Subsequently, we can combine the Sharpe ratio estimate with a volatility estimate to derive an estimator of the equity premium. This estimator offers several advantages over more traditional estimators. Initially, it introduces a new line of thought for estimating the equity premium that diverges from existing methods. Moreover, it relies exclusively on observable parameters, negating the necessity for calibrating dividend or earnings growth, as is customary in dividend/earnings discount models. Similarly, the calibration of asset values or default barriers, typically required in traditional applications of structural models, is circumvented. Lastly, according to Kaserer and Berg (2008), the approach displays robustness to model changes. They specifically examine the model of Duffie and Lando (2001), which is currently one of the best structural models in the literature, to show this robustness. The carbon tax scenarios are incorporated into this framework by adjusting the estimated real-world probabilities of default under several carbon tax scenarios.

1.5 Overview of the paper

The following chapters are structured as follows. The methodology is explained in Chapter 2. Chapter 3 discusses the different data sets used for the research. The data chapter is followed by Chapter 4, in which the results are presented. Finally, the conclusions drawn from the results are presented in Chapter 5.

Chapter 2

Methodology

In this chapter, the methods of the paper are discussed and rationalized. First, Section 2.1 outlines the theoretical framework and accompanying assumptions that were made in this paper. Following this, Section 2.2, outlines how we derive the real-world PDs for the different companies. Similarly, Section 2.3 explains how the risk-neutral PDs are calculated. After that, Section 2.4 highlights how real-world and risk-neutral PDs are used to estimate Sharpe ratios. This section includes two suggested adjustment factors to the Sharpe ratio estimate. Section 2.5 delves into several volatility estimates which are used in combination with the Sharpe ratio estimates to come up with equity premium estimates.

2.1 Theoretical framework

This paper uses a framework that resembles the Merton framework to predict equity premiums. However, it is important to note that there are some clear distinctions between the Merton framework proposed by Merton (1974) and the framework in which this paper is written. Our framework differs from the Merton framework in that the latter assumes a constant Sharpe ratio, whereas we do not make this assumption. This difference arises from the Merton framework's assumption of constant drift, volatility, and risk-neutral rate, whereas our approach involves a more dynamic interpretation. Specifically, we assume exogenous changes in market conditions govern the model parameters and, therefore, can change over time. One example of such an exogenous change in market conditions can be a change in the carbon emission tax rate. We assume that drift, volatility and risk-neutral rate parameters can change over time due to these changes in market conditions. The asset values V_t are modeled as a geometric Brownian motion with volatility σ_t and drift μ_t, i.e.,

$$dV_t = \mu_t V_t dt + \sigma_t V_t dB_t^{\mathbb{P}}, \tag{2.1}$$

where $B_t^{\mathbb{P}}$ denotes a standard Wiener process under the real-world $\mathbb{P}$-measure. Under the assumptions of no arbitrage opportunities and the existence of a risk-free asset, Girsanov's Theorem allows us to change from the real-world $\mathbb{P}$-measure to the risk-neutral $\mathbb{Q}$-measure utilizing the Radon-Nikodym derivative. In our case, this is relevant for relating the risk-neutral probability of default to the model parameters. Written under the $\mathbb{Q}$-measure the equation becomes,

$$dV_t = r_t V_t dt + \sigma_t V_t dB_t^{\mathbb{Q}}, \tag{2.2}$$

where $B_t^{\mathbb{Q}}$ denotes a standard Wiener process under the risk-neutral $\mathbb{Q}$-measure and r_t being the return of the risk free asset at time t. Lastly, the default barrier L, which in the Merton framework is defined as the limit below which the asset value triggers a default, is the only model parameter chosen to remain constant in our framework. The boundary L can be interpreted as a company's debt level. The company defaults when the value of the assets falls below the value of the debt.

In the Merton framework, there exists an equation relating the probabilities of default to the model parameters. Due to the dynamic nature of our model parameters, it is no longer possible to assume that this equation holds. However, in this paper, we assume that for short horizons up to one year, we can accurately approximate the relations between the probabilities of default with the standard Merton framework. This assumption can be interpreted as the fact that on short horizons the probability of large enough exogenous shocks changing the probabilities of default is negligible. We therefore assume that we can use the equation,

$$P^{def}(t,T) = \Phi\left(\frac{\ln\frac{L}{V_t} - \left(\mu_t - \frac{1}{2}\sigma_t^2\right)\cdot(T-t)}{\sigma_t\cdot\sqrt{T-t}}\right), \tag{2.3}$$

to approximate the real-world probability of default for maturities up to one year. Here, Φ represents the cumulative standard normal distribution function. Similarly, we assume that we can use the equation,

$$Q^{def}(t,T) = \Phi\left(\frac{\ln\frac{L}{V_t} - \left(r_t - \frac{1}{2}\sigma_t^2\right)\cdot(T-t)}{\sigma_t\cdot\sqrt{T-t}}\right), \tag{2.4}$$

to approximate the risk-neutral probability of default. Furthermore, we make an assumption specifically on the effect of a carbon emission tax rate increase. Tax costs are inferred to correlate with a company's earnings, posited on the premise that they likely scale proportionately with production volume, under the assumption of fixed marginal carbon emissions per product. These additional taxes, proportional to production figures, then decrease the drift μ_t of the company. Tax costs do not cause shocks as they are always proportional to earnings and therefore we assume they do not change the volatility. Lastly, we assume that the value of the assets of a company is not affected as a result of changes in the carbon emission tax rate. The aggregate of these assumptions leads the risk-neutral probabilities to be constant to a carbon emission tax shock.

Another feasible interpretation is to assume that the change in real-world probability is completely driven by the change in the asset value V_t. This assumes constant drift and volatility, and consequently, the Sharpe ratio remains unchanged. In reality, the solution would likely be somewhere in between these two solutions. Therefore our estimate of the change in the Sharpe ratio can be interpreted as an upper bound of the shock in our framework. Making different assumptions on the dynamics of the model parameters would empirically involve calculating the real-world and risk-neutral probabilities of default using estimated parameter values which defeats the purpose of using this framework to come up with a Sharpe ratio estimator. We were also not able to predict the changes in the risk-neutral probability of default using the changes in the real-world probability of default with a regression.

2.2 Real-world probability of default

The model that we use to estimate the probability of default (PD) is the corporate rating model (CRM) which was developed and calibrated by Zanders. This model utilizes historical balance sheets and profit and losses (PNL) as well as cash flow data to assess the likelihood of a company defaulting on its debt. The corporate rating model computes multiple sub-scores and sub-ratings, which are combined to come to a final credit risk rating (CRR) assigned to a specific company. To generate the sub-scores and sub-ratings, various types of scoring functions are employed, including linear functions and sigmoid functions, for instance. These sub-scores and sub-ratings are combined in a multi-step process to result in a CRR. The CRRs are mapped to a score range from "AA" to "C" and can be directly linked to PD values I.e., every score has a (midpoint) PD assigned to it. To make the PDs more sensitive to small changes in the CRR, we map the CRRs to PD values by linearly interpolating the midpoint PDs assigned to the scores. This allows us to directly link a CRR to a PD value.

The strategy to consider the potential impact of carbon emission taxes relies on Zanders' internal expertise. Essentially, the approach involves computing the carbon tax expenses incurred by individual companies and subsequently projecting the consequences on various other metrics, including net profit, tangible net worth, EBIT, EBITADA, and current assets. These adjusted metrics are then plugged into the CRM model to get the carbon tax-adjusted PDs.

2.3 Risk-neutral probability of default

The other component necessary for the estimation of the company Sharpe ratios is the risk-neutral probability of default. We compute the risk-neutral PD by utilizing CDS spreads along with an estimated recovery rate (RR) and time to maturity. The derivation of equation 2.5 can be found in Chan-Lau (2006) which base their derivation on the constant hazard model of Duffie (1999). The risk-neutral default probability implied by the CDS spread is approximately,

$$Q_i^{def}(t,T) = 1 - e^{\frac{-S_{i,t}\,(T-t)}{1-RR_{i,t}}} \tag{2.5}$$

where $S_{i,t}$ and $RR_{i,t}$ are the CDS spread and recovery rate of firm i at time t. The recovery rate is assumed to be constant and equal to 40% in line with market conventions and industry practices. $T-t$ denotes the time to maturity over which the risk-neutral probability of default is calculated. In this paper, it is always equal to 1 year. As mentioned in Section 2.1, we have made several assumptions regarding the response of parameters in the Merton model to changes in carbon emission tax rates. These assumptions lead to the inference that risk-neutral probabilities remain unchanged despite variations in carbon tax rates.

2.4 Sharpe ratio estimator

Once the real-world and risk-neutral probability of default estimations are obtained, the subsequent step involves calculating the company Sharpe ratios. To achieve this, we adopt the estimator developed by Kaserer and Berg (2008), which is rooted in the Merton framework. Our framework assumes that on short horizons we can use equations 2.3 and 2.4 which Kaserer and Berg (2008) show can be rewritten to come up with a Sharpe ratio estimate,

$$\widehat{SR}_{i,t} := \frac{\mu_{i,t} - r_t}{\sigma_{i,t}} = \frac{\Phi^{-1}(Q_i^{def}(t,T)) - \Phi^{-1}(P_i^{def}(t,T))}{\sqrt{T-t}}, \tag{2.6}$$

where $P_i^{def}(t,T)$ and $Q_i^{def}(t,T)$ are the real-world and risk neutral PDs respectively. These are derived as mentioned in the previous paragraphs. $\mu_{i,t}$ and $\sigma_{i,t}$ are the expected return and volatility for company i in period t, while r_t is the risk-free rate in that period. $\widehat{SR}_{i,t}$, is the estimated company Sharpe ratio of company i in period t.

2.4.1 Adjustment factor

The Merton framework models a company's equity as a call option on its assets. A default is modeled if the call option is out of the money at the time of maturity[2]. Therefore the assumption is made that a default can only occur at the time of maturity. The first-passage framework, proposed by Black and Cox (1976), removes this assumption making it more realistic for modeling default probabilities. Kaserer and Berg (2008) come up with an adjustment term such that the estimator fits into the more realistic first-passage framework. A drawback to this approach is that the asset volatility, company Sharpe ratio, maturity, rating grade and risk-free rate need to be chosen at realistic values. These values need to be estimated leading to the possibility of estimation error. This paper introduces a second approach to compute useful adjustment factors. The approach is a regression to find the optimal adjustment.

This paper suggests trying two approaches to estimating the adjustment factor. The first is using the plug-in approach proposed by Kaserer and Berg (2008) centered around the first-passage time framework. The second is a new method that is centered around a regression to find the optimal adjustment.

First-passage adjustment

In the context of the Merton framework, default is constrained to happen exclusively at the bond's maturity, a simplification that diverges from real-world dynamics. In this subsection, we explore a first-passage-time framework developed by Black and Cox (1976), in which default is allowed to occur before the maturity date. Making it a more realistic representation of the real world, which one may anticipate results in better estimates of the Sharpe ratio. The approach is similar to the approach Kaserer and Berg (2008) use to show their estimator is robust to a model change from the Merton framework to the first-passage-time framework.

[2] An option is considered out of the money when the current price of the underlying asset is not favorable for the option holder to exercise the option for a profit. In this case, the call option is out of the money when the stock price is below the strike price.

In the first-passage framework the equation for $P^{def,FP}(t,T)$ and $Q^{def,FP}(t,T)$ are slightly different than in the Merton framework. Consequently, equation 2.6 no longer holds and there exists no analytical formula for estimating the Sharpe ratio. The equation for the real-world probability of default, $P^{def,FP}(t,T))$, in the first-passage framework is,

$$P_i^{def,FP}(t,T)) = \Phi\left(\frac{b - m_t^p(T-t)}{\sigma_t\sqrt{T-t}}\right) + e^{\frac{2m_t^p b}{\sigma_t^2}}\Phi\left(\frac{b + m_t^p(T-t)}{\sigma_t\sqrt{T-t}}\right), \tag{2.7}$$

with $b = ln\left(\frac{L}{V_t}\right)$, $m_t^p = \mu_t - \frac{1}{2}\sigma_t^2$, $\sigma_t = \sigma_{V,t}$ according to Kaserer and Berg (2008).

Similarly the equation for the risk-neutral probability of default, $Q^{def,FP}(t,T)$, in this framework is,

$$Q_i^{def,FP}(t,T) = \Phi\left(\frac{b - m_t^q(T-t)}{\sigma_t\sqrt{T-t}}\right) + e^{\frac{2m_t^q b}{\sigma^2}}\Phi\left(\frac{b + m_t^q(T-t)}{\sigma_t\sqrt{T-t}}\right), \tag{2.8}$$

with $b = ln\left(\frac{L}{V_t}\right)$, $m_t^q = r_t - \frac{1}{2}\sigma_t^2$, $\sigma_t = \sigma_{V,t}$.

As there is no closed-form solution for the Sharpe ratio the adjustment factor needs to be estimated using a different approach. We use a similar methodology to Kaserer and Berg (2008), with one difference: we use estimated values to calculate the adjustment factor, whereas they use hypothetical values to show that the adjustment factor typically behaves well and remains close to 1. First, realistic values for asset volatility, company Sharpe ratio, maturity and real-world PD need to be chosen. The volatility and company Sharpe ratio are estimated using sample moments of the stock returns over the past 60 months. The maturity is chosen to align with the period over which the real-world PD is calculated. The real-world PD is taken from the CRM model mentioned in section 2.3. We also use a risk-free rate equal to the Euribor rate with the same maturity as the real-world PD. Subsequently, the distance to default barrier measure b is computed, using equation 2.7, we numerically search for the value of b for which the right-hand side is equal to the real-world PD selected value in the initial step. Using the estimated $\hat{b}$ and the parameters established in the initial step as inputs, the first-passage risk-neutral PD, $Q_i^{def,FP}$, is calculated according to equation 2.8. In the final step, we plug the real-world and risk-neutral PDs into our Merton Sharpe ratio estimator as presented in equation 2.6. This gives us a Sharpe ratio estimate when assuming the (incorrect) Merton framework while the underlying truth is the first-passage framework.

$$\widetilde{SR}_{i,t,M} = \frac{\Phi^{-1}(Q_i^{def,FP}(t,T)) - \Phi^{-1}(P_i^{def,FP}(t,T))}{\sqrt{T-t}} \tag{2.9}$$

To determine the first-passage adjustment factor, we compare the input Sharpe ratio with the (incorrect) Merton estimator. This comparison allows us to compute the appropriate adjustment factor for a company whose underlying properties match those estimated in the initial step of the methodology when utilizing a Merton estimator. The adjustment factor can be calculated as,

$$AF_{FP,i,t} = \frac{\widehat{SR}_{i,t,Input}}{\widetilde{SR}_{i,t,M}}, \tag{2.10}$$

With $\widetilde{SR}_{i,t,M}$ being the Sharpe ratio when using the inappropriate Merton estimator in the first-passage framework for firm i in period t, and $\widehat{SR}_{i,t,Input}$ representing the inputted Sharpe ratio calculated using sample moments for the same firm in the same period. The corresponding first-passage adjustment factor is denoted as $AF_{FP,i,t}$.

Regression adjustment

The second approach to obtain an adjustment factor is a regression to find the optimal adjustment. The regression should include lagged variables such as returns, asset volatility, company Sharpe ratio and real-world PD. A specific variable we would like to include is the (PD/ volatility) as these seem to have a multiplicative effect on each other for the first-passage adjustment factor in Kaserer and Berg (2008). These variables were selected based on the fact that the first-passage adjustment is dependent on the same variables and thus they theoretically have some explanatory power. The regression gives better insight into where the adjustment term comes from. The regression can be written as,

$$AF_i^* = SR_i^{Rel}/\widehat{SR}_{i,M} = \alpha + \sum_{j=1}^{k} \beta_j X_{i,j} + \epsilon_i, \tag{2.11}$$

With AF_i^* being the optimal adjustment factor.

SR_i^{Rel} is the realized Sharpe ratio defined as the annualized daily Sharpe ratio for observation i, which corresponds to a specific company in a specific year. The calculation is done by first deriving the daily realized Sharpe ratio from the sample moments of the daily excess returns, followed by multiplying this ratio by $\sqrt{252}$ to annualize it. The $X_{i,j}$'s are the covariates: realized returns, Sharpe ratios, historical volatilities, real-world PDs, and the ratio of real-world PD to historical volatility, incorporating lagged values up to 3 periods. We fit all possible regression equations including different combinations of these covariates and use the AIC to identify the best performing model. The β_j coefficients correspond to their respective variables. We estimate the coefficients, including α, using the ordinary least squares (OLS) estimator. In the end, the adjusted Sharpe ratios estimate is computed as follows,

$$\widehat{SR}_{i,Reg} = \widehat{AF}_{Reg}\widehat{SR}_{i,M} = \left(\hat{\alpha} + \sum_{j=1}^{k} \hat{\beta}_j X_{i,j}\right) \widehat{SR}_{i,M}, \tag{2.12}$$

With $\widehat{AF}_{Reg}$ being the regression adjustment. The parameters $\hat{\alpha}$ and $\hat{\beta}$ are the OLS estimators of the regression shown in equation 2.11. This approach to calculating the adjustment term has two main advantages. First of all this method is much more flexible as it does not assume a certain framework like the first-passage adjustment does. The second advantage is the fact that this method provides much better insights into what effect each variable has on the adjustment factor. This can in turn give insights into where our Merton framework leads to adequate results and where it does not.

2.5 Volatility estimate

This section discusses the potential volatility estimates and their advantages and disadvantages. An overview of these advantages and disadvantages can be found in Table 2.1.

Category	Model	Advantages	Disadvantages
Historical Volatility	Monthly HV	Smaller period adjustment Non-parametric	Limited amount of datapoints Backwards looking
	Daily HV	More datapoints Non-parametric	Need period adjustment Backwards looking
GARCH	Monthly GARCH	More dynamic	Need t+12 estimate parameter uncertainty
	Daily GARCH	More datapoints More dynamic	Need t+252 estimate parameter uncertainty

Table 2.1: Overview of proposed volatility estimation models along with their anticipated advantages and disadvantages.

First of all the Historical Volatilities (HV) offer several advantages and disadvantages. One advantage is that this is a non-parametric approach and, therefore, does not suffer from parameter uncertainty. A disadvantage is that the HV methods are backward-looking. Therefore, the estimates relate to historical volatility, which is not necessarily the same as the current or future volatility. For both the daily and the monthly historical volatility we go back for 12, 24 and 60 months to see which volatility works best for our application. A disadvantage of the monthly HV approach when compared to the daily approach is that there is a limited amount of data points to construct the historical volatility leading to a larger estimation error. For the yearly historical volatility, the amount of datapoints becomes so limited that this approach was not considered. The monthly HV approach offers an advantage over the daily HV method due to its closer alignment with the timeframe of the yearly volatility being estimated. The method therefore suffers less from inaccuracy introduced by the time adjustment.

The next category of models considered is the GARCH model. What sets these models apart from the HV models is the fact that they allow for volatility to change over time making them suffer less from the backward-looking nature of the approach. A yearly GARCH would suffer from the fact that there are very few data points in the dataset. This leads to very large estimation errors of the parameters, making this approach unusable and was therefore not considered. The advantage of this GARCH model is that we only need a 1-step-ahead forecast. On the other hand, we have the daily GARCH. This model can use lots of data points within the sample allowing for better calibration of the parameters. The downside of this approach is the fact that we need a 252-step ahead forecast. As the amount of steps increases GARCH models converge to the historical volatility estimate. At 252 steps we can expect the GARCH model to make almost the same prediction as the historical volatility estimate losing any potential gain from using the model. The monthly GARCH strikes a balance between these two approaches. There are more data points in the sample than for the yearly GARCH and we only need a 12-step ahead estimate which is much less than the 252 steps needed for the daily GARCH. Therefore this is the best prospect of the different GARCH models.

Finally, we would like to remark on the use of an implied volatility (IV) estimate. Based on the Black-Scholes framework and the observed option prices in the market the implied volatility can be calculated. The main advantage of this approach is that it is forward-looking due to the use of options. The maturity over which the IV is calculated could therefore be aligned with the period we are interested in instead of using historical data like the HV and GARCH approaches. The disadvantage of this method is the fact that it is under the risk-neutral $\mathbb{Q}$-measure instead of the real-world $\mathbb{P}$-measure. Therefore, theoretically, this estimate is not completely justified. In practice, however, we expect this estimate of the volatility under the $\mathbb{Q}$-measure to be a good proxy for real-world volatility. We were however unable to obtain a suitable dataset for the implied volatility. The only option available to us was to manually retrieve the implied volatility of each company individually. Therefore, the decision not to pursue this approach was taken based on practical considerations of time constraints and the limited expected improvement in the performance of implied volatility compared to alternative volatility estimation methods.

2.6 Equity premium

Once we have estimated a Sharpe ratio and a volatility we can combine these estimates to get an equity premium estimate. The equation for the equity premium is,

$$EP_{i,t} = \mu_i - r = \frac{\mu_{i,t} - r_t}{\sigma_{i,t}} \sigma_{i,t} = SR_{i,t}\ \sigma_{i,t}, \tag{2.13}$$

where $EP_{i,t}$ is the equity premium $\mu_{i,t}$ represents the expected returns for company i in period t. The risk-free rate is denoted as r_t.

As we try several options of volatility estimate as mentioned in Section 2.5 we get different estimators of the equity premium. We also use different adjustment factors leading to even more equity premium estimates. Irrespective of which Sharpe ratio and volatility estimate used the equation for the equity premium estimator is,

$$\widehat{EP}_{i,t} = \widehat{SR}_{i,t}\ \hat{\sigma}_{i,t} \tag{2.14}$$

Using $\widehat{SR}_{i,t}$ as the estimator from equation 2.6 and $\hat{\sigma}_{i,t}$ as the estimated volatility. Our estimation of the excess return corresponds directly to the predicted equity premium, thus yielding $\hat{R}_{i,t} = \widehat{EP}_{i,t}$.

2.7 Benchmarks and evaluation

To examine the performance of the proposed equity premium estimator, we compare its performance with four benchmarks. An overview of the benchmarks along with their advantages and disadvantages can be found in Table 2.2.

Benchmark	Advantages	Disadvantages
Zero	Non-parametric No variance	Inflexible Bias
CAPM	Strong theoretical foundation	Omitted risk drivers
FF3	Strong theoretical foundation	Omitted risk drivers
Hist. Avg.	Non-parametric Easy to compute	Does not account for changing market conditions

Table 2.2: The considered benchmark models along with their corresponding advantages and disadvantages.

The first and simplest is the zero benchmark, a non-parametric 'baseline model' that always estimates an equity premium of 0. This benchmark has an interesting bias-variance trade-off as it removes all forecast variance, but has the potential to introduce bias into its estimates due to its complete inflexibility. The next two benchmarks are the CAPM, developed by Sharpe (1964) and Lintner (1965), and its extension, the Fama and French 3-factor model, first introduced by Fama and French (1996). These models calculate the sensitivity of each stock to each of the risk factors in the factor model and assign an equity premium according to the sensitivity to the risk factors and their premia. The sensitivities and risk premia are calculated over the last 60 months. The CAPM model uses only the market factor, while the Fama and French 3-factor model includes the size and value factors as well as the market factor. The models have a strong theoretical foundation as they are derived from an econometric model with a few reasonable assumptions on market dynamics. A disadvantage of these models is that they likely do not account for all the possible risk drivers as they only use one or three risk drivers. The last benchmark is the historical average benchmark, which looks at the average excess returns over the last 60 months for a given stock. The historical benchmark estimates are the annualized average monthly excess returns over this period. The approach is non-parametric and easy to compute but is unable to account for changing market conditions.

To evaluate the performance of the proposed method we look at its ability to predict the excess returns as this is the only (noisy) estimate of the equity premium we have. By assessing the root mean squared errors (RMSE) and mean absolute error (MAE) of the proposed method and the benchmark methods we can conclude its performance. To be specific the equation used to calculate the RMSE is,

$$RMSE_m = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(R_i - \hat{R}_{m,i}\right)^2}, \tag{2.15}$$

where $RMSE_m$ is the root mean squared error of model m. The excess return R_i corresponds to a specific company and year. The prediction $\hat{R}_{m,i}$ is the equity premium (expected excess return) forecast by model m for observation i. Lastly, N refers to the amount of observations found in the sample for which the RMSE is calculated. Similarly, the MAE is calculated using the equation,

$$MAE_m = \frac{1}{N}\sum_{i=1}^{N}\left|R_i - \hat{R}_{m,i}\right|, \tag{2.16}$$

which averages over the absolute error of the predictions.

Both the RMSE and MAE are calculated for each model such that we can assess how much each model's performance is affected by outliers in their prediction errors. The RMSE is more sensitive to outliers, therefore we can compare the performance with respect to these two metrics to get insight into the amount and size of outliers of the estimation errors.

To test for significant outperformance of the benchmark with respect to the proposed equity premium forecast we use a paired t-test for the difference in mean of the squared errors and the absolute errors. This test is performed on the cross-section of stocks for each period in the sample. This approach allows to to tell which model is significantly better than the others at a given horizon. However, making overall statements over the entire sample period presents multiple testing issues that need to be controlled for. The reason we did not opt for a more conventional significance test predicting data with a temporal dimension is due to the extremely short sample period of our dataset[1]. Tests like the Diebold Mariano and Giacomini-White which rely on estimating autocovariance within the estimation period can therefore not be run. We chose to perform the paired t-test on the models in comparison with the historical average benchmark, given its widespread application in equity premium forecasting and its simplicity.

The paired t-test makes assumptions about the data. First of all, each of the paired measurements must be obtained from the same subject. This holds as we are measuring prediction errors of different models for the same stock. Something that is less clearly the case is that the measured differences between the pairs should be normally distributed. Using the differences between squared/absolute prediction error, rather than their actual values, leads to a more normally distributed outcome. This is an advantage compared with what would occur if an unpaired t-test is used. Taking the differences likely does not completely get rid of the non-normality. Lastly, the paired t-test assumes subjects are independent in the sense that the measurement of one subject should not affect the expectation for any other subject. This assumption is likely to not hold as it is very well known that stock returns are often highly correlated. If we measure that one stock performed very differently from the expectation other stocks will likely perform very differently from their expected returns as well. When using the paired t-test we have to keep in mind that not all the assumptions are met and should not rely on its results too heavily.

Testing the performance of the proposed model is necessary to validate its underlying framework, particularly for carbon tax sensitivity analyses. By assessing its effectiveness across different sectors and carbon emissions segments, we check the validity of the model in these subsets. This validation process strengthens our confidence in using the framework to explore the sensitivity of different subsets of stocks to carbon tax shocks.

[1] The prediction period contains only 5 periods, spanning from 2018 to 2022. More details about the final dataset can be found in Section 3.6

2.8 Carbon tax scenarios

We evaluate three distinct carbon tax scenarios derived from prices found in scenarios by the Network for Greening the Financial System (NGFS). To be clear we are investigating the effect of government behavior concerning carbon tax rates which is different from the social cost of carbon. The scenarios are published in the First NGFS Comprehensive Report (NGFS, 2019). The scenarios are also featured on the NGFS's website, providing not only insights into potential scenarios but also information about the utilized data. The NGFS partnered with an expert group of climate scientists and economists to design a sizable set of hypothetical scenarios which are categorized as "Orderly," "Disorderly," and "Hot house world" depending on their physical and transition risk profile. We chose one scenario from each category and below we provide the NGFS's definitions of these scenarios.

- **Orderly:** *Below 2°C*, gradually increases the stringency of climate policies, giving a 67% chance of limiting global warming to below 2°C.
- **Disorderly:** *Divergent Net Zero*, reaches net zero around 2050 but with higher costs due to divergent policies introduced across sectors leading to a quicker phase-out of oil use.
- **Hot house world:** *Nationally Determined Contributions (NDCs)*, includes all pledged targets even if not yet backed up by implemented effective policies.

In the analyses, we consider the effect that the predicted 2050 carbon tax rate would have on equity premiums if governments would choose to implement them immediately. Note we do not use any other information like inflation, GDP growth and technological change from these scenarios, we only use the carbon tax rates in the year 2050. The specific model employed was created by the NGFS and is referred to as REMIND-MAgPIE 3.0-4.4. The tax rates per tonne of CO2 from the model can be found in Table 2.3. The model determines the price per tonne of CO2 in 2010 US dollars (USD). The carbon tax sensitivity analyses are performed on 2021 data, therefore we need to convert the 2010 USD into 2021 USD and 2021 Euros. To get the 2021 USD amount we controlled for the inflation on the period from 2010 to 2021. To come up with the 2021 Euro values took into account the average EUR/USD rate during the year 2021. Both steps were taken using Bureau of Labor Statistics data. The price actual price of carbon emissions is estimated as the price of the emissions trading systems (ETS) in the EU. In July of 2021, this was equal to €55. This allows us to estimate the carbon shocks in the different scenarios.

Scenario	2010 USD	2021 USD	2021 EUR	Shock in EUR
Below 2°C	134.84	167.20	141.42	86.42
Divergent Net Zero	700.76	868.94	734.95	679.95
NDC	50.67	62.83	53.14	-1.86

Table 2.3: Overview of carbon price scenarios both in 2010 US dollars as provided by the NGFS as well as 2021 US dollars. The 2021 Euro value is also provided along with the carbon shock in Euros given a 2021 carbon price of 55 Euro.

2.9 Carbon tax sensitivity analyses

To assess the possible effect of several carbon emission tax scenarios we investigate what happens when using the carbon-adjusted probability of default instead of the regular probability of default. This results in different predicted equity premiums. We assess the size of this impact for different sectors and emission segments by comparing the unadjusted equity premium predictions to the carbon-adjusted ones. As mentioned previously the analysis is performed for three different carbon price scenarios which are based on predictions of future carbon prices by the Network for Greening the Financial System (NGFS). The carbon prices in the different scenarios can be found in Table 2.3.

Chapter 3

Data

This chapter outlines the five datasets utilized in this research. We give an overview of all the variables included in each dataset and indicate what they are used for. This chapter also offers insights into the data by presenting descriptive statistics and providing additional context regarding the data's frequency, volume, and sources. At the end of this chapter an overview of our final dataset after linking all data sources is provided.

3.1 Financial statements data

The financial statements are sourced from company balance sheets, PNLs and cash flow statements. The data used for this research originates from Bureau van Dijk and has been precleaned by Zanders as this dataset has been used to calibrate their CRM model. The full dataset we received from Zanders contains data on about 1.8 million companies. The sample comprises annual data from 2000 to 2022. As not all companies are present throughout the entire period, there are approximately 7.2 million data points. Please refer to Table B.1 in Appendix B for the exact number of observations each year.

Each data point has 67 different variables. From these 67 variables, we use 31 for the PD forecasts. A list of the variables used can be found in Table B.1 in Appendix B. These 31 variables are input for the CRM model. The CRM model uses them to generate a list of 15 risk drivers which have been listed in Table 3.1. The dataset was narrowed down to include only those companies that had data available for each year from 2012 to 2021. If a company met this requirement, all of its data points from 2012 to 2022 were selected, including 2022 where possible but not mandatory. This resulted in a list of about 90,000 companies with about 900,000 data points in total. This dataset is used as input for the CRM model which utilizes it to estimate real-world PDs for the individual firms. In the final dataset, not all these companies and periods are used as only data points with corresponding data points in other datasets are included in the final dataset.

Pillar	**Risk driver**
I - Operations	Turnover growth
	Gross profit margin
	Operating profit margin
	Return on sales
	Return on capital employed
II - Liquidity	Current ratio
	Debtor days
	Creditor days
III - Capital structure	Gearing
	Solvency
IV - Debt service	Total debt / EBITA
	Interest coverage ratio
V - Turnover & Category	Turnover
	Size
	Industry

Table 3.1: List of risk drivers used in the CRM model

3.2 CDS data

Credit default swaps (CDS) are popular over-the-counter credit derivatives. In a CDS, the protection buyer pays periodic premiums to the protection seller. If a credit event occurs, the protection seller covers losses by paying the difference between the nominal and current market value of a specified reference obligation to the protection buyer. However, practical complexities arise, including defining credit events, specifying reference obligations, and clarifying “market value”. In academic research, we often restrict our research to one type of CDS to prevent discrepancies between different CDS spreads in the same paper. This paper for example only focuses on CDS on “Senior” debt with one year to maturity.

Zanders is subscribed to the ICE database which provides them with current market data. Daily, they make comprehensive requests for various market data from the database and save the results in an internally managed database. The original dataset contains 3487 companies on a sample spanning from January 2014 to October 2023. For the analyses, we use the par mid spreads with a tenor of 1 year. The dataset retrieved from Zanders contains daily data but the data is transformed into yearly data. As a proxy for the yearly CDS, we use average CDS spreads over January of each year. This gives the CDS spreads time to adjust to the financial statement information usually disclosed at the end of December and reduces the effect of outliers on specific days. In the final dataset we only use the CDS spreads for the companies found in the other dataset. The CDS spreads are used to calculate the risk-neutral PDs. In the final dataset, the sample period starts in 2017 because, as of this year Zanders started to make large cross-market requests about CDS data. This dataset is what limits our final dataset to predictions of equity premiums to a sample starting in 2018.

3.3 Return data

The return data, sourced from the Eikon database, encompasses the daily stock prices of companies listed in the New York Stock Exchange (NYSE) and EURO 600 index, amounting to 2525 companies. Furthermore, it incorporates the Euribor 1-year rate. The sample period extends from 2013 to 2022, with a monthly frequency. This dataset serves various purposes: it enables the calculation of yearly excess returns, which provide a somewhat noisy estimate of the equity premiums targeted for forecasting; moreover, it is used for the generation of HV and GARCH volatility estimates, as well as the computation of realized company Sharpe ratios.

3.4 Fama and French data

For the application of the CAPM and Fama-French 3-factor benchmarks, we retrieved data from the French (2024) website, which contains monthly observations spanning from July 1926 to December 2023. The dataset includes monthly observations of the three components within the Fama-French 3-factor model: the market factor (MKT), size factor (SMB), and value factor (HML), alongside the monthly risk-free rate. The Fama-French factors are constructed using value-weight portfolios formed based on market capitalization and book-to-market value. The SMB (Small Minus Big) factor represents the difference in average returns between portfolios containing small market capitalization stocks and portfolios with large market capitalization stocks, while the HML (High Minus Low) factor captures the difference in average returns between portfolios with high book-to-market value stocks and portfolios with low book-to-market value stocks. The market factor is derived as the average return of all portfolios minus the risk-free rate. Finally, the monthly risk-free rate is determined as the one-month Treasury bill rate. All stock returns used for the factor generation are retrieved from NYSE, AMEX, or NASDAQ.

3.5 Carbon emission data

The carbon emission data utilized in this study is sourced from the Eikon database. These values, obtained for the year 2021, have been interpolated by Eikon themselves to reflect the real or estimated total carbon emissions within scopes 1 and 2. The dataset is used to evaluate the implications of our model for the sensitivity of equity premiums on carbon emission taxes. Carbon emissions are only retrieved for the companies for which it was possible to find financial statements, CDS and return data. This data is used in combination with the carbon emission price scenarios to calculate the cost shocks to the companies and inspect the possible impact of these shocks on the equity premiums. The companies in this dataset are also grouped into sectors and carbon emission segments such that we can inspect the aggregated impact of carbon emission tax.

3.6 Final dataset

The final dataset contains 179 companies, with names successfully matched across the three data sources and having the relevant data points to make at least one equity premium prediction.

This process was done using a Python package published by De Nederlandsche Bank (DNB)[1] designed for matching large lists of messy company names between datasets. It is important to note that not all data points and time periods are available for each company for which we were able to match their names. This section aims to provide insight into the composition of the companies included in the final dataset. The final dataset is focused on the data utilized for equity premium predictions, spanning the period from 2018 to 2022. Depending on the type of data this requires different sample periods. The return dataset extends from 2013 to 2022, the financial statements dataset covers the years 2015 to 2022, and the CDS spread data spans from 2017 to 2022.

Size class	Count	Percentage
Very large	161	90%
Large	16	9%
SME	2	1%

Table 3.2: Distribution of companies dataset across different size classes.

Table 3.2 presents the distribution of companies across different size classes in the final dataset. The majority of companies, comprising 90% of the dataset, are classified as very large enterprises with a turnover exceeding 100 million euros. About 9% of the companies fall into the category of large enterprises, with turnovers ranging between 10 million and 100 million euros. Merely 1% of the dataset comprises small and medium-sized enterprises (SMEs) with turnovers below 10 million euros.

Country	Count	Percentage
US	74	41%
FR	22	12%
GB	13	7%
IT	11	6%
NO	7	4%
NL	7	4%
Other	45	26%

Table 3.3: Distribution of companies across different countries.

The distribution of companies across different countries is outlined in Table 3.3. The United States (US) dominates the dataset, constituting 41% of the total count. Following the US, France (FR) accounts for 12% of the companies, while the United Kingdom (GB) and Italy (IT) represent 7% and 6% respectively. Norway (NO) and the Netherlands (NL) each contribute 4% of the dataset. The remaining 26% is distributed among other countries not individually listed.

Table 3.4 provides insights into the currencies companies report in. Nearly half of the companies, 49%, report in United States Dollars (USD), while 33% report in Euros (EUR). A smaller fraction, 6%, report in British Pounds (GBP). The remaining 12% of companies report in currencies not explicitly listed in the table.

[1] The "name_matching" Python package enables the efficient matching of company names from different datasets. It preprocesses names, applies cosine similarity to identify potential matches and employs fuzzy string matching algorithms. With a simple implementation, users can customize matching criteria based on their needs. The package can be found on GitHub (DNB, 2022).

Currency	Count	Percentage
USD	88	49%
EUR	59	33%
GBP	10	6%
Other	22	12%

Table 3.4: Distribution of currencies in which companies report their financial statements

Company sector	2018	2019	2020	2021	2022	Sales	CO2
Manufacturing	50	51	51	59	64	27.30	5.17
Information and communication	15	15	16	18	19	28.35	0.66
Electricity, gas, steam and air co	10	11	11	15	16	23.84	15.10
Wholesale and retail trade	11	11	11	13	15	4.10	0.78
Mining and quarrying	7	8	8	12	14	40.58	13.94
Other	30	32	33	44	45	13.50	1.57

Table 3.5: The number of equity premium predictions for each sector for each year. For each sector, the average sales, carbon emissions and CO2 sales ratio are reported. Note the sector averages have been created on 2021 and thus correspond to the prediction counts of 2022. Sales are reported in billions of Euros and carbon emissions are reported in millions of tones of CO2

The amount of companies for which we can make equity premium estimates per sector is reported annually in Table 3.5. The table shows the five most common sectors in the dataset and also provides the count of the companies that fall into all other sectors. In the final year of 2022, the manufacturing industry stands out with the highest number of observations, totaling 64 companies. In comparison, the sectors of information and communication; electricity, gas, steam, and air co; wholesale and retail trade, and mining and quarrying have observation counts ranging between 19 and 14 companies. Additionally, there are 45 companies spread across various other sectors. It also includes the average sales and carbon emissions of each sector for the year 2021.

Emission segment	2018	2019	2020	2021	2022	Sales	CO2
Low	46	47	47	55	56	27.88	0.29
Medium	34	35	36	47	53	25.62	2.48
High	38	39	40	51	55	15.13	12.47

Table 3.6: The number of observations for each carbon segment for each year. For each segment, the average sales, carbon emissions and carbon sales ratio are reported. Note the segment averages have been created on 2021. Sales are reported in billions of Euros and carbon emissions are reported in millions of tones of CO2

Table 3.6 also includes the count of companies for which equity premium estimates can be predicted, segmented by carbon emission intensity levels. The segmentation is performed using 2021 data and sorts the companies into segments based on their carbon emissions per unit of turnover. In 2022, approximately one-third of the data points fall into each emission segment, with counts per segment ranging between 53 and 56. Additionally, the table provides insights into the average sales and carbon emissions within each segment.

Chapter 4

Results

This chapter displays and discusses the most important results from the empirical analyses. First, the performance of the equity prediction framework is discussed in Section 4.1. This section analyses the validity of the proposed framework in different periods, sectors and carbon emission segments. Subsequently, the implications of this framework on the equity premium in several carbon tax scenarios are discussed in Section 4.2.

4.1 Equity premium prediction

This section presents and compares the performance of the proposed equity premium prediction models with the benchmarks. By observing the performance of the Merton framework relative to the benchmark models, we can get an indication of the validity of the framework. This is done first for the full cross-section of stocks for each period and for the full sample period. This is followed by an analysis of each sector and carbon emission segment for each period and for the full sample period. This gives us an indication of the validity of using the Merton framework on these subsets of the data, which is necessary for the carbon sensitivity analyses in Section 4.2.

We have chosen to present the equity premiums estimated using the 60-month historical volatility as the Merton estimates because this model version performs similarly to the rest and is likely the most stable. Using a different volatility estimate leads to similar results which can be found in Appendix C. This section presents the Merton and first-passage prediction performance. We opted not to utilize the regression adjustment model due to our inability to accurately predict the optimal adjustment factor. This limitation is reflected in the insignificance of all parameters in all the tried regressions. More details on the results of the regression analyses can be found in Section 4.1.5. This section includes the zero, historical average and CAPM benchmarks. Due to the similar performance of the CAPM and Fama-French 3-factor models we have opted to only report the less complicated CAPM model. The results of the Fama-French 3-factor model can also be found in Appendix C.

4.1.1 Evaluation metrics

We employ the root mean squared error (RMSE) to evaluate the models' performance. This metric is reported both annually and cumulatively over the period from 2018 to 2022 on the full cross-section of stocks. For the Merton model and historical average benchmark, this is also done on the sectors and carbon emission segments separately. Additionally, we also report the mean absolute error (MAE) both annually and cumulatively over the period from 2018 to 2022 on the full cross-section of stocks. This is done because the RMSE is more sensitive to outliers than the MAE which allows us to compare the models concerning the size and amount of outliers they have.

4.1.2 RMSE evaluation

The RMSEs of the proposed models and the benchmarks are reported in Table 4.1 along with the number of observations per year. The table also contains p-values calculated using a two-sided paired t-test on the difference in the mean of the squared errors between the historical average benchmark and other models. First of all, it is worth noting that, within each year, there appears to be a substantial co-movement in model performance across all the models. It is expected that the proposed Merton and first-passage models exhibit co-movement in their RMSE values, as they are computed in a highly similar manner and only differ by an adjustment factor. Notably, the benchmarks' performances also demonstrate a co-movement with the other models. This is likely attributed to certain years exhibiting higher variance or larger absolute mean in returns throughout the cross-section of stock, resulting in poorer prediction performance than other years across all models. In 2021 for example, the stock market experienced a period of very high returns which none of the models were able to predict resulting in poor performance across all models.

In three of the five years, as well as for the full sample, every benchmark outperforms the proposed Merton and first-passage models and the difference with the historical average is significant at a confidence level of 95% or higher. However, in 2018, both proposed models significantly outperform the historical average benchmark, with the Merton model also slightly outperforming the CAPM benchmark. Conversely, in 2020, both models outperform the three benchmarks, albeit not significantly in comparison to the historical average benchmark, which ranks as the poorest-performing model for that year. It is important to note that the reliability of the p-values is somewhat compromised due to certain assumptions not being fully met, as discussed in 2.7, particularly in the case of the full sample p-values. However, it becomes clear that the proposed models are outperformed by the benchmarks overall.

In terms of the comparison between the Merton model and the first-passage model, the Merton model outperformed the first-passage model in three of the five years, with the margins between the models usually very small. The Merton model also has a slightly lower (0.3327) RMSE in the full sample as compared to the RMSE of the first-passage model (0.3355). Even though the difference in performance does not seem that large it is still significant with the p-value being 0.00 when rounded to two decimal places[1]. The outperformance of the Merton model

[1] In line with Table 4.1, the p-value corresponds to a two-sided paired t-test on the difference in the mean of the squared errors of the Merton and first-passage models in the full sample.

RMSE		2018		2019		2020		2021		2022		Full Sample	
Framework	Merton	0.2692	0.00	0.3010	0.00	0.2546	0.11	0.4428	0.01	0.3280	0.03	0.3327	0.05
	First-passage	0.2756	0.00	0.2993	0.00	0.2513	0.09	0.4458	0.01	0.3353	0.01	0.3355	0.03
Benchmark	Zero	0.2445	0.00	0.2455	0.25	0.2905	0.27	0.4027	0.01	0.2494	0.00	0.2965	0.01
	CAPM	0.2712	0.00	0.2444	0.21	0.2883	0.24	0.3616	0.06	0.2504	0.00	0.2882	0.00
	Hist. Avg.	0.3353	-	0.2327	-	0.3017	-	0.3742	-	0.2860	-	0.3114	-
N		125		128		130		161		173		717	

Table 4.1: The yearly and overall RMSE of the Merton model and first-passage model estimates along with the zero, historical average and CAPM benchmarks. To the right of each RMSE a p-value is shown, this p-value corresponds to a two-sided paired t-test on the difference in the mean of the squared errors in the given period between the historical average benchmark and the other models. The amount of observations in each sample can be found at the bottom of the table.

is likely due to the additional noise introduced by the first-pass model. This noise stems from the need to estimate the volatility of the assets and the Sharpe ratio of the companies, which are calculated using sample moments in the returns and can thus act as sources of noise. This additional noise likely outweighs the benefit of using a more realistic model in our application.

4.1.3 Sectors and carbon emission segments RMSE evaluation

The objective of this paper is however not to come up with the most effective equity premium prediction model. We are interested in creating a framework that can accurately predict the sensitivity of the equity premium to changes in the carbon emission tax. The validity of the proposed framework is checked by analyzing the predictive performance of the models. As we are interested in calculating the sensitivity of different sectors and carbon emission segments we subsequently analyze the validity of the framework in these subsets.

For this, we present the RMSEs of the Merton and historical average models for every sector cumulatively over the period 2018 to 2022 in Table 4.2. The RMSEs of both the Merton and historical average models for the different sectors are reported annually in tables C.3 and C.4 in Appendix C, respectively. Similarly, Table 4.3 presents the results for the carbon emission segments, corresponding to the Merton and historical average benchmark models. The annual RMSEs of the Merton and historical average models for the carbon emission segments can be found in tables C.5 and C.6 in Appendix C, respectively.

By comparing the performance of the models across different sectors and carbon emission segments for several time periods, as well as comparing the performance between the two models, we can gain insight into the viability of the proposed framework for each subset of the data under consideration. This evaluation informs us whether the framework can be reasonably applied to infer the sensitivity of the equity premium to carbon taxes across different sectors and carbon emission segments.

As mentioned, Table 4.2 displays the RMSE values per sector for the Merton and historical average models, respectively. The performance of the Merton models varies across different sectors. However, this does not necessarily imply that the Merton framework is invalid in certain sectors. Instead, it could be attributed to the fact that in some samples the realized returns are absolutely larger and therefore more difficult to accurately predict.

The Merton model for instance has a full sample RMSE of 0.5427 in the mining and quarrying sector while the full sample RMSEs of the other sectors are all between 0.2446 and 0.3482. We however observe that the historical average benchmark has similarly poor performance for this sector when compared to the other sectors. This leads us to believe that the relatively large RMSE is caused by more extreme absolute returns as opposed to the Merton framework being invalid for this sector.

Company sector	Merton	Hist. Avg.
Manufacturing	0.3482	0.3176
Information and communication	0.2446	0.2429
Electricity, gas, steam and air co	0.2461	0.2756
Wholesale and retail trade	0.2841	0.3144
Mining and quarrying	0.5427	0.4678
Other	0.3097	0.2870

Table 4.2: The full sample RMSEs of the Merton and historical average models split out over the five most common sectors in our dataset. The RMSEs of all the other companies can be found at the bottom of the table.

When analyzing individual time periods in tables C.3 and C.4 in Appendix C, we see that either the Merton model or the historical average benchmark consistently outperforms the other across all or most sectors. This can be attributed to the different input data and methodological approaches used by the two models, resulting in predicted equity premiums having different means. Due to the substantial co-movement of stock prices across various sectors, the overall performance of the stock market in a given year significantly influences the performance of each model. Therefore, one approach may consistently outperform the other across all sectors during a specific period. This observation does not falsify the Merton framework in certain years. This is because the difference in performance by year between the Merton model and historical average benchmark is very much correlated leading us to conclude that this difference in performance is caused by unexpectedly large absolute returns and not due to the invalidity of the Merton framework.

Since there are no instances where the Merton estimator substantially underperforms compared to the historical average benchmark across periods or sectors, we cannot falsify the validity of the Merton framework using this metric. Therefore, the Merton framework is valid for our carbon sensitivity analyses for all sectors and each year.

Emission segment	Merton	Hist. Avg.
Low	0.2798	0.2714
Medium	0.2968	0.3080
High	0.4220	0.3678

Table 4.3: The RMSEs for the low, medium, and high emission segments of the Merton and historical average models over the full sample period. The classifications have been made based on emissions-to-sales ratios, splitting the data into three similarly sized categories.

Table 4.3 presents the RMSE values for the Merton model and the historical average benchmark, respectively, across the emission segments for the full sample. In the table, we once again observe that the performance of the Merton model and the historical average benchmark are remarkably similar. Specifically, the models' relative performance across the different emission segments remains highly consistent. Tables C.5 and C.6 , which can be found in Appendix C, show that for each period, the relative ranking of the model performance for each segment was identical for both the Merton model and the historical average benchmark. Additionally, if one of the models outperforms the other in one of the segments for each period, it always outperforms all three. The performance of the Merton model is relatively close to the historical average benchmark, and we do not observe specific periods or carbon emission segments for which the performance is much worse than the historical average. We therefore again cannot falsify the validity of the Merton framework using this metric. Therefore, the Merton framework is valid to use in our carbon sensitivity analyses for all carbon emission segments and in all years.

4.1.4 MAE evaluation

The last metric used to analyze the performance of the equity premium predictions is the mean absolute error (MAE). The MAE is less sensitive to outliers than the RMSE, so if a model's RMSE is substantially affected by a few outliers, the MAE will be less affected. The annual and full sample MAE of the proposed models and benchmarks can be found in Table 4.4.

MAE		2018		2019		2020		2021		2022		Full Sample	
Framework	Merton	0.2066	0.00	0.2269	0.00	0.1964	0.00	0.2827	0.00	0.2392	0.12	0.2334	0.32
	First-passage	0.2126	0.00	0.2242	0.00	0.1942	0.00	0.2849	0.00	0.2442	0.06	0.2353	0.20
Benchmark	Zero	0.1929	0.00	0.1951	0.03	0.2169	0.03	0.2573	0.03	0.1828	0.00	0.2097	0.00
	CAPM	0.2145	0.00	0.1919	0.05	0.2203	0.07	0.2272	0.21	0.1895	0.00	0.2083	0.00
	Hist. Avg.	0.2675	-	0.1769	-	0.2356	-	0.2369	-	0.2200	-	0.2272	-
N		125		128		132		161		173		717	

Table 4.4: The yearly and overall MAE of the Merton model and first-passage model estimates along with the zero, historical average and CAPM benchmarks. To the right of each RMSE a p-value is shown, this p-value corresponds to a two-sided paired t-test on the difference in the mean of the absolute errors in the given period between the historical average benchmark and the other models.

Comparing these results with those in Table 4.1, which displays root mean square error (RMSE), reveals similar insights into model performance. For each period the models have the same relative ranking for the MAE metric as for the RMSE, with the only exception being that the first-passage model now outperforms the CAPM in 2018. The difference between the two models is very small in both metrics.

There is a decrease in differences between models which could be attributed to MAE generally yielding smaller values than RMSE. We are not convinced this decrease in difference is due to outliers in the prediction errors as the relative performance seems to be conserved. This suggests that the previously noted subpar performance of equity premium estimators, as indicated by RMSE, is not primarily driven by outliers in predictions.

4.1.5 Regression adjustment

As described in section 2.4.1, we have also attempted to create an adjustment factor based on a regression to find a better performing adjustment factor that is also more interpretable. This section explains how we tried to create such an adjustment factor and presents the results of the best regression we found. As previously mentioned, the prediction results of the regression adjustment model are not presented, as we were unable to develop a model that performed satisfactorily. More details as to why the performance was deemed insufficient to include in the paper are also provided in this section.

First of all, to find the best performing regression to predict optimal adjustment factors we performed a grid search over a large number of different predictor combinations. We analyzed five predictors: realized returns, Sharpe ratios, historical volatilities, real-world PDs, and the ratio of real-world PD to historical volatility, incorporating lagged values up to 3 periods and exploring the inclusion of a constant term. We performed a grid search over the $32,767$ possible combinations of these variables and evaluated the models on the AIC score. The best performing model found using this procedure is a regression using only a constant and the real-world PD to volatility ratio. A summary of the regression results can be found in Table 4.5

Variable	Coef.	Std. Error	t	P > $\|t\|$
Constant	0.6780	0.477	1.422	0.156
PD_{t-1}/vol_{t-1}	-10.7477	8.462	-1.270	0.205

Table 4.5: OLS regression results for the top-performing model based on AIC score. This regression includes a constant term and the ratio of real-world PD to historical volatility. The table presents coefficients, standard errors, t-statistics, and corresponding two-sided p-values for each regressor. The number of observations used in the regression is 523. The R-squared of the regression is 0.003

From the regression results, we conclude that we are unable to find significant regressors to predict the optimal adjustment factor within our dataset. This can be concluded from the fact that the p-values for the constant and the ratio of real-world PD to historical volatility in the previous period are 0.156 and 0.205 respectively. These are considered to lack statistical significance even when employing the liberal significance threshold of 0.10. Insignificant coefficients in the regression indicate that the independent variables lack a significant impact on the dependent variable, rendering the model ineffective for prediction and its results difficult to interpret. The R-squared value of 0.003 also indicates our regression is unable to explain much of the variance in the optimal adjustment factor. Due to the regression's inability to accurately predict the optimal adjustment factor we have decided not to use the regression adjustment model in the final results of the paper.

One possible solution for the regression is to set all parameters to zero, resulting in a trivial regression and an adjustment factor that is always zero. This approach would yield the same predictions as the zero benchmark, which we have demonstrated to generally outperform the Merton model. This may explain why we did not find any significant parameters in our regression analyses.

4.2 Carbon tax scenario

After assessing the validity of the framework, we now examine the implications of this framework on the sensitivity of equity premiums to carbon tax rates. The analysis concentrates on predictions of 2022 equity premiums, investigating three distinct carbon tax scenarios and comparing them to the unadjusted case. This section first presents the carbon sensitivity analysis of several sectors followed by the same analysis for the carbon emission segments.

4.2.1 Sector analyses

The results of the carbon sensitivity analyses by sector are presented in Table 4.6. We opted to focus on the five most common sectors within our dataset: manufacturing; information and communication; electricity, gas, steam, and air co; wholesale and retail trade; and mining and quarrying. These sectors contain a considerable amount of observations, which ensures that the sample remains representative of each sector and accounts for noise in the data. The results for all other companies are aggregated under the title "Other" which can be found at the bottom of Table 4.6

Company sector	N	Unadjusted	Divergent net zero		Below 2		NDC	
Manufacturing	64	11.05%	2.33%	(-8.72%)	8.47%	(-2.58%)	11.13%	(+0.08%)
Information and communication	19	7.46%	1.22%	(-6.24%)	3.95%	(-3.50%)	7.65%	(+0.19%)
Electricity, gas, steam and air co	16	2.77%	-7.36%	(-10.13%)	-1.61%	(-4.37%)	2.89%	(+0.12%)
Wholesale and retail trade	15	12.74%	-1.34%	(-14.08%)	0.64%	(-12.10%)	14.87%	(+2.13%)
Mining and quarrying	14	26.96%	13.62%	(-13.34%)	23.24%	(-3.73%)	27.01%	(+0.05%)
Other	38	15.73%	11.65%	(-4.09%)	14.33%	(-1.41%)	15.85%	(+0.11%)

Table 4.6: The predicted equity premiums for the four most common sectors in 2022. Equity premiums for all other companies are listed in the bottom row. It includes the unadjusted equity premium alongside those in three carbon tax scenarios. Values in parentheses denote the difference between the unadjusted equity premiums and the respective scenario.

As per the findings presented in Table 4.6, our framework predicts that equity premiums will decrease with an increase in carbon tax rates, which is expected. This is evident from the fact that all sectors have decreased predicted average equity premiums in the "divergent net zero" and "below 2" scenarios, where the carbon tax rates are higher than the current tax rates. Additionally, it can be observed that a decrease in carbon tax, as is the case in the "nationally determined contributions" (NDC) scenario, leads to an increase in predicted equity premium.

Looking at the divergent net-zero scenario, different sectors show different sensitivities to the carbon tax shock, with shocks to the equity premium ranging from -4.09% to -14.08%. The wholesale and retail trade; and mining and quarrying sectors are affected the most with a negative equity premium shock of 14.08% and 13.34% respectively. However, the mining and quarrying sector is still expected to have the highest equity premium of all sectors, which could be due to the exceptional performance of the companies in our dataset. This large effect is to be expected for the mining and quarrying sector as these companies are likely to emit a lot of CO2 as shown in Table 3.5. The result of the wholesale and retail trade is less obvious as this sector emits much less CO2 per unit of sales. The large effect could be due to these companies often having very small profit margins, a relatively small cost shock can therefore have a large relative impact on the profit margin.

The corporate rating model is likely more sensitive to changes in the profit margin of wholesale and retail trade companies as this is a very important metric for such a company. From the calibrated parameters of the CRM, in combination with generally small profit margins, it can be observed that of the considered sectors the wholesale and retail trade generally has the highest sensitivity to changes in profit margin for every size class in most cases[2]. Remember that we assume that tax costs are fully absorbed by companies, with full price rigidity in effect. As a result, these costs have a direct impact on profit margins.

On the other hand, companies in the information and communications sector likely do not emit as much CO2 at shown in Table 3.5. This is in line with them being the least affected of the sectors analyzed, with only a negative shock of 6.24% to the equity premium in the divergent net-zero scenario. The only group of companies affected less by the tax scenario is the "Other" category which only experiences a negative shock of 4.33%. In the "divergent net zero" scenario, the worst-performing sector is the Electricity, gas, steam and air co sector. This is due to the sector performing relatively poorly in the unadjusted case with an equity premium of only 2.77% and receiving a moderately large negative shock of 10.13% resulting in a predicted equity premium of -7.36%.

When comparing the results in the "below 2" scenario to the "divergent net zero" scenario we observe that the shocks to the equity premiums have a smaller magnitude now ranging from 1.41% to 12.10%. This is to be expected as this scenario considers a much smaller tax shock. What is noteworthy to see is that for most sectors the shock to the equity premium is somewhere between a third and half of the shock in the "divergent net zero" scenario. The exception to this is the wholesale and retail trade sector which even for this smaller tax shock has a reduction in the predicted equity premium of 12.10%, which is not much less than the shock in the "divergent net zero" scenario of 14.08%. Again this could be due to the margins in this sector being very small and company financials therefore looking unfavorable at relatively small cost shocks if they are not passed onto the customers. This is one way variations in marginal sensitivities to carbon tax shocks can arise not just from the nominal tax amount, but also from differences in companies and sectors. Moreover, the non-linear nature of the PD model leads to varying sensitivities of company PDs to the same nominal shock. Different risk drivers and corresponding sub-scores can become more or less sensitive to the marginal of the input variables due to non-linear transformation and combination functions in the PD model. Consequently, the marginal sensitivity to carbon tax shocks can differ for the same company depending on the shock's magnitude. The lack of a straightforward economic interpretation is attributed to the complexity of the PD model.

Lastly, for the "nationally determined contributions" (NDC) scenario, we see small positive shocks to the equity premium ranging between 0.05% and 2.13%. This was to be expected as this scenario corresponds to a small decrease in carbon emission taxes. Remarkably, the mining and quarrying sector is affected the least in this scenario with a positive equity premium shock

[2]This is reasoned through observing the derivative of the profit margin sub-score in the CRM model. The derivative is most often higher for the wholesale and retail trade companies given the wholesale and retail trade company does not have a much larger profit margin than the company from a different sector. We reason this is often the case even though we have not tested this quantitatively.

of 0.05% whereas it had the second largest shock to its equity premium in the "divergent net zero" scenario of -13.34. This again highlights that the predicted equity premiums' marginal sensitivity to the tax shock can change considerably for different nominal shocks.

4.3 Carbon emission segment analyses

The effects of the tax scenarios on the carbon emission segments are shown in Table 4.7. These results give us a better understanding of the direct relationship between carbon emissions and the sensitivity of the equity premium to carbon emission taxes. The results are in line with what we would expect, the higher the emission-to-sales ratio the larger the effect of carbon taxes on the equity premium. In the divergent net-zero scenario, the higher emission segment is affected the most with a negative shock of 18.56%. This is substantially higher than the negative shock for the medium segment of 5.96%, which in turn is much higher than the negative shock for the low segment of only 0.84%.

Emission segment	N	Unadjusted	Divergent net zero		Below 2		NDC	
Low	56	12.99%	12.16%	(-0.84%)	12.89%	(-0.10%)	13.00%	(+0.00%)
Medium	53	10.64%	4.67%	(-5.96%)	10.04%	(-0.59%)	10.65%	(+0.01%)
High	54	14.70%	-3.87%	(-18.56%)	4.66%	(-10.04%)	15.56%	(+0.87%)

Table 4.7: The predicted equity premiums for the low, medium, and high emission segments. The classifications have been made based on emissions-to-sales ratios, splitting the data into three equally sized categories. It includes the unadjusted equity premium prediction alongside those in three carbon tax scenarios. Values in parentheses denote the difference between the unadjusted equity premiums and the respective scenario.

For the emission segments, the other two scenarios behave similarly. In the "below 2" scenario, the shocks to the carbon emissions are smaller than for the "divergent net zero" scenario. In this scenario, the negative shocks range from 10.04% for the high emission segment and 0.10% for the low emission segment. Similarly to the sector analyses we again observe small positive shocks to the equity premium in the "nationally determined contributions" scenario which is expected given this scenario considered a small decrease in the carbon tax rate.

Chapter 5

Conclusion

5.1 Equity premium prediction

The proposed Merton and first-passage equity premium prediction models do not outperform the benchmark models in forecasting equity premiums. The zero, historical average and CAPM benchmark slightly outperform the proposed models with respect to the RMSE and MAE for most years, as well as when aggregating over the full sample period. However, it is important to note that accurate equity premium predictions are not the main objective of this research. The aim of testing the predictive performance of the proposed Merton model is to assess the validity of using the Merton framework for our carbon sensitivity analyses. We conclude that the performance of the Merton model is sufficient to justify its use in the carbon sensitivity analyses.

The models' performances are highly correlated within each time period. For instance, in 2021, the stock market experienced a period of very high returns that none of the models could predict, resulting in poor performance across all models. This indicates that the model performance is largely driven by the overall movement of the stock market during a specific period. Due to the limited sample period, drawing definitive conclusions about the performance of various models is statistically difficult.

The Merton and first-passage models exhibit similar predictive performance. In the full sample, the Merton model has a slightly lower RMSE of 0.3327 compared to the first-passage model's RMSE of 0.3355. Although this difference is relatively small it is significant. This suggests that the additional noise introduced by the first-passage adjustment outweighs the benefits of using a more realistic model in our application. We recommend using the Merton model instead of the more complicated first-passage model as it provides no additional benefits.

To evaluate the effect of outliers on prediction model performance, we compare the performance measured by the RMSE and MAE. The MAE is less sensitive to outliers, so if a model performs better relative to the MAE than the RMSE, it may indicate that the model is affected by outliers in the prediction errors. We find that the models' relative performance is very similar when measuring performance with the RMSE and MAE. The models maintain the same relative ranking for the MAE metric as for the RMSE metric, with only one exception in one period. Based on these results, we conclude that the model performances are not majorly affected by a small number of large outliers in the prediction errors.

To assess the validity of the Merton framework across various sectors and emission segments, we calculated the RMSE separately for each part of the dataset, both annually and for the entire sample period. Our comparison of the Merton model with the historical average benchmark suggests that we cannot falsify the framework's validity for any of the periods, sectors, or carbon emission segments based on the equity premium prediction results. Based on these results, it is justifiable to explore the effects of the Merton framework on equity premiums across various sectors and carbon emission segments.

Lastly, the regression adjustment model was not included in the analyses of the equity premium predictions because none of the parameters in the attempted regressions are significant. Using these regression results would therefore make no sense and result in very poor performance of the model. An interpretation of only finding insignificant parameters is that they could all be equal to zero. This trivial regression would be the same as the zero benchmark, which we showed to outperform the Merton model.

5.2 Carbon tax scenario

The results of the carbon tax scenario analyses indicate that carbon tax rates have a negative relationship with the equity premium. Increased costs lead to larger probabilities of default, resulting in smaller equity premium predictions for all sectors and emission segments. Conversely, a tax reduction leads to higher equity premiums. The magnitude of these effects depends on the size of the tax shock but also company financials play a role. Different sectors and carbon emission segments are predicted to have different sensitivities to carbon tax shocks.

Our results show that different sectors have different sensitivities to carbon tax shocks. In the most extreme, divergent net-zero scenario, which corresponds to a carbon tax rate increase of €679.95 per tonne of CO2, the wholesale and retail trade sector, along with the mining and quarrying sector, experience the largest decrease in the equity premium. Specifically, these negative shocks amount to 14.08% and 13.34% respectively. The large effect on the mining and quarrying sector can be explained by the fact that these companies emit a large amount of CO2. The large impact observed in the wholesale and retail trade sector is explained by the fact that these companies have low profit margins, making it difficult to absorb costs that cannot be passed on to customers. It was also observed that probability of default of a company in the wholesale and retail trade sector is often more sensitive to changes in profit margin than other sectors. We conclude that the size of the shock to the equity premium depends on the level of carbon emissions, but also on the way in which companies' finances are structured. The negative shocks to the other sectors are all between 4.09% and 10.13%. The shocks to the equity premium are relatively large and different from each other, which would certainly affect valuations and investment decisions. We can conclude that a tax shock of €679.95 per tonne of CO2 would have a significant negative impact on almost all equity premiums, with different sensitivities across sectors.

In the "below 2" scenario, which corresponds to a carbon tax rate increase of €86.42 per tonne of CO2, the results of our sector analyses are very similar albeit less severe. The negative equity premium shocks range from 1.41% to 12.10% which are very substantial and would affect investment decisions. It is interesting to note that the wholesale and retail trade sector again experiences a very large negative shock of 12.10% to its equity premium, while the mining and quarrying sector is much less affected in this scenario with a negative shock of only 3.73%. We attribute this difference to the difference in profit margins between the sectors. Even this lower tax rate is likely to erode a large part of the profit margins of the wholesale and retail trade sector, while the mining and quarrying sector has large enough profit margins to absorb this additional cost. Finally, the "nationally determined contributions", which correspond to a reduction in the carbon tax of 1.86 per tonne of CO2, leads to small increases in equity premiums between 0.05% and 2.13%. Once again, the wholesale and retail trade sector is by far the most affected, followed by the information and communication sector with an increase of only 0.19%. This scenario would have minor effects on investment decisions.

The analysis of the carbon emission segments shows similar results as the sector analyses. From the results, we can conclude that carbon emissions are a very important driver of the effect on the equity premium in the carbon emission tax scenarios. The results show that the high emitting segment, in the "divergent net zero" scenario, could face a very large reduction in equity premium of 18.56% whereas the low emitting segment would only expect a 0.84% reduction in equity premium in the same scenario. The medium segment falls in between with a reduction of the equity premium of 5.96%. The shocks to the equity premium are slightly milder in the "below 2" scenario and even slightly positive in the "nationally determined contributions" scenario. We can conclude that carbon taxes could have a profound impact on equity premiums, with higher emitters being affected the most.

References

Adämmer, P. & Schüssler, R. A. (2020). Forecasting the equity premium: Mind the news! *Review of Finance*, *24*(6), 1313–1355. doi: 10.1093/rof/rfaa007

Alexandridis, A. K., Apergis, I., Panopoulou, E. & Voukelatos, N. (2023). Equity premium prediction: The role of information from the options market. *Journal of Financial Markets*, *64*, 100801. doi: 10.1016/j.finmar.2022.100801

Attoh, E. M., de Bruin, K., Goosen, H., van Veldhoven, F. & Ludwig, F. (2022). Making physical climate risk assessments relevant to the financial sector – lessons learned from real estate cases in the netherlands. *Climate Risk Management*, *37*, 100447. doi: 10.1016/j.crm.2022.100447

Baetje, F. & Menkhoff, L. (2016). Equity premium prediction: Are economic and technical indicators unstable? *International Journal of Forecasting*, *32*(4), 1193–1207. doi: 10.1016/j.ijforecast.2016.02.006

Black, F. & Cox, J. C. (1976, May). Valuing corporate securities: Some effects of bond indenture provisions. *The Journal of Finance*, *31*(2), 351–367. doi: 10.1111/j.1540-6261.1976.tb01891.x

Bolton, P. & Kacperczyk, M. (2021, Apr). Do investors care about carbon risk? *Journal of Financial Economics*. doi: 10.3386/w26968

Bua, G., Kapp, D., Ramella, F. & Rognone, L. (2022). Transition versus physical climate risk pricing in european financial markets: A text-based approach. *SSRN Electronic Journal*. doi: 10.2139/ssrn.4154034

Chan-Lau, J. A. (2006). Market-based estimation of default probabilities and its application to financial market surveillance. *IMF Working Papers*, *06*(104), 1. doi: 10.5089/9781451863642.001

DNB. (2022). *De nederlandsche bank name_matching.* Retrieved from `https://github.com/DeNederlandscheBank/name_matching`

Duffie, D. (1999). Credit swap valuation. *Financial Analysts Journal*, *55*(1), 73–87. doi: 10.2469/faj.v55.n1.2243

Duffie, D. & Lando, D. (2001, May). Term structures of credit spreads with incomplete accounting information. *Econometrica*, *69*(3), 633–664. doi: 10.1111/1468-0262.00208

Fama, E. F. & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, *22*(1), 3–25. doi: 10.1016/0304-405x(88)90020-7

Fama, E. F. & French, K. R. (1996, Mar). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, *51*(1), 55. doi: 10.2307/2329302

Ferentinos, K., Gibberd, A. & Guin, B. (2021). Climate policy and transition risk in the housing market. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3838700

French, K. R. (2024). *Current research returns.* Retrieved from `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

Friede, G., Busch, T. & Bassen, A. (2015). Esg and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance amp;amp; Investment*, *5*(4), 210–233. doi: 10.1080/20430795.2015.1118917

Han, B. & Zhou, Y. (2011). Term structure of credit default swap spreads and cross-section of stock returns. *SSRN Electronic Journal*. doi: 10.2139/ssrn.1735162

Hilscher, J., Pollet, J. M. & Wilson, M. (2015). Are credit default swaps a sideshow? evidence that information flows from equity to cds markets. *Journal of Financial and Quantitative Analysis*, *50*(3), 543–567. doi: 10.1017/s0022109015000228

Kaserer, C. & Berg, T. (2008). Linking credit risk premia to the equity premium. *SSRN Electronic Journal*. doi: 10.2139/ssrn.1103502

Kothari, S. & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics*, *44*(2), 169–203. doi: 10.1016/s0304-405x(97)00002-0

Li, J. & Tsiakas, I. (2017). Equity premium prediction: The role of economic and statistical constraints. *Journal of Financial Markets*, *36*, 56–75. doi: 10.1016/j.finmar.2016.09.001

Lintner, J. (1965, Feb). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*(1), 13. doi: 10.2307/1924119

Luo, L. & Tang, Q. (2014). Carbon tax, corporate carbon profile and financial return. *Pacific Accounting Review*, *26*(3), 351–373. doi: 10.1108/par-09-2012-0046

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, *29*(2), 449. doi: 10.2307/2978814

NGFS. (2019, Apr). A call for action: Climate change as a source of financial risk.

Rapach, D. E., Strauss, J. K. & Zhou, G. (2009). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies*, *23*(2), 821–862. doi: 10.1093/rfs/hhp063

Sharpe, W. F. (1964, Sep). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425. doi: 10.2307/2977928

van Ewijk, C., de Groot, H. L. & Santing, A. C. (2012). A meta-analysis of the equity premium. *Journal of Empirical Finance*, *19*(5), 819–830. doi: 10.1016/j.jempfin.2012.07.002

Wang, Y., Pan, Z., Liu, L. & Wu, C. (2019). Oil price increases and the predictability of equity premium. *Journal of Banking amp;amp; Finance*, *102*, 43–58. doi: 10.1016/j.jbankfin.2019.03.009

Welch, I. & Goyal, A. (2007). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, *21*(4), 1455–1508. doi: 10.1093/rfs/hhm014

Appendix A

Company contributions

A.1 Company and ESG efforts

Zanders was established in 1994 and has, in recent years, become a leading consultancy firm specializing in treasury management, risk management, and corporate finance. The company boasts a workforce of over 450 consultants, catering to a clientele of more than 500 clients. Its primary office is situated in Utrecht, the Netherlands, with additional offices in Belgium, Germany, Switzerland, Sweden, the United Kingdom, the United States, the United Arab Emirates and Japan.

This research was conducted within the Financial Institutions department, which provides financial risk advisory services to banks, insurers, and asset managers. Currently, most financial institutions are actively involved in activities such as investigating, measuring, or modeling climate change risks within their portfolios. Therefore within the Financial Institutions department, several employees collaborate within an Environmental, Social, and Governance (ESG) expert group, primarily focused on addressing climate change risks. This paper mainly leverages a carbon-adjusted probability of default (PD) model, that was developed by the ESG expert group. More on the specifics of this model can be found in Section 2.2.

Appendix B

Data

The data for this paper was sourced from four different databases. In Table B.1 an overview is given on the different data sources, mode of access and the data retrieved from each one along with their use cases.

Data type	Database	Mode of access	Used case
Financial statements and P&L	Bureau van Dijk	Zanders	Real world probability of default
Credit default swaps spreads	ICE	Zanders*	Risk neutral probability of default
Equity prices	Refinitiv Eikon	University (EUR)	Realized returns, SR and volatility estimates
Carbon emissions	Refinitiv Eikon	University (EUR)	Carbon adjustments
1 Year Treasury Rate	Refinitiv Eikon	University (EUR)	Realized Sharpe ratios
Fama and French	Kenneth R. French website	Public website	CAPM and Fama-french 3-factor models

Table B.1: Overview of the different databases used along with the mode of access, data retrieved, and their use case

The following sections discuss the process of retrieving each type of data as well as the decisions made as to which data was retrieved. Not all retrieved data was used in the paper since not every retrieved data point had a corresponding data point in the other databases.

B.1 Financial statements and P&L

The financial statements and P&L data were retrieved from Bureau van Dijk by Zanders. The dataset used was the dataset that was collected and used to calibrate their credit risk models and spanned 1,808,751 companies in a period from 2000 until 2022. Because there are a lot of missing observations in the dataset, there are only 7,209,974 data points. The data was provided as a parquet file which was loaded into SQL to process the data. A decision was reached by examining various subsets and assessing the number of available companies, striking a balance to include as many companies as possible while maintaining a sufficient number of observations per company. To ensure the reliability of the selection in this initial data screening, the decision was made to focus exclusively on companies with observations spanning each year from 2012 to 2021. In Table B.3 the observation counts for each year in the dataset are reported. It can be seen that the observation counts started to increase as of 2011. They peak in 2020 and starkly decrease in 2022. For the companies meeting this criteria, we selected the data points from 2012 until 2022 which means 2022 was optional. The resulting dataset contains 91,365 companies and 930,059 observations. Adding the constraint of having to include 2022 this would be decreased to

15,351 companies and 169,111. It was decided against creating a dataset in which the possibility of one or several missing years within the sample is permitted as the probability of default model for which this data is used works best if several lagged terms are available to estimate several growth rates.

The resulting set of data points still had several issues which had to be resolved. It was found that several companies had duplicate observations. The companies where this issue was observed were almost exclusively found to be from Sweden so we expect there to be an issue with Bureau van Dijk's data collection process for this county. It was considered unnecessary to delete these duplicate observations since they consistently contained identical data, provided that the company and year remained the same. The data is used by looking up rows based on company and year and therefore it does not make a difference to the final results. Given the nature of the data, which involves balance sheets, due to accounting standards certain balance sheet items must add up to other balance sheet items by construction. To give an example, Intangible fixed assets, Tangible fixed assets and Other fixed assets should add up to be equal to Fixed assets. This is often not exactly the case, however. When observing the discrepancies it was concluded these errors most likely stemmed from some numeric computational error. This conclusion was drawn as the error terms were almost exclusively exactly equal to powers of two or sums of only a few powers of two. The most common errors were 256, 512 and 767 which are equal to 2^9, 2^{10} and $2^9 + 2^{10}$ respectively. The number of powers of two never surpassed the number of aggregated balance sheets or P&L items. Therefore, we conclude that during the conversion process from the Bureau van Dijk database to the version we received, small rounding errors were introduced by a digital system. As most values in the dataset are in the hundreds of thousands and we expect the rounding mistakes to not have been made for smaller values we expect these data imperfections to not have any noticeable effect on our results.

Finally, it was decided to access the code on which the CRM model runs directly. We received two Python scripts and one SQL script along with the Zanders credit risk python package. The SQL script converted the Bureau van Dijk data such that it could be used in the CRM model. The script essentially added all lagged terms up to $t - 3$ in each row, resulting in a structure similar to the one found in the Postman API template. The first python script ran the SQL script such that all the data was ready for the CRM model. The second script ran the CRM model. This method worked and resulted in 'final scores' for each company in each year, these were then mapped to probabilities of default.

B.2 Credit default swap

Zanders is subscribed to the ICE database which provides them with current market data. Daily, they make comprehensive requests for various market data from the database and save the results in an internally managed database. The process of making daily requests and saving the received has been going on since 2013 giving an upper limit to how far back any data obtained from this source can go. However, not all data has been in the request set for the entire period since 2013. This routine process allows Zanders to accumulate a vast dataset of market information. Among the products requested, credit default swaps (CDS) hold particular significance for this research.

Variable		
Size	Other Shareholders Funds	Capital
Segment	Long Term Debt	Operating Expenses
Company Name	Other Non-Current Liabilities	Amortization
Intangible Assets	Provisions	Financial Revenues
Tangible Assets	Short Term Debt	Financial Expenses
Other Assets	Account Payable External	Interest Paid
Inventory	Account Payable Internal	Taxation
Account Receivable External	Other Current Liabilities	Extra Revenues
Account Receivable Internal	Turnover	Extra Expenses
Other Current Assets	Cost of Goods Sold	Dividends
Cash Equivalent		

Table B.2: The full list of variables included in the financial statement dataset that is used for calculating the real-world probabilities of default with the corporate rating model (CRM).

status year	observation count	status year	observation count
2000	15644	2012	413593
2001	21040	2013	554950
2002	26487	2014	508373
2003	33977	2015	470451
2004	42805	2016	541309
2005	53381	2017	581080
2006	119752	2018	591383
2007	150720	2019	628988
2008	187473	2020	726616
2009	185299	2021	595457
2010	117562	2022	344579
2011	299055		

Table B.3: Amount of observations per year in the BvD dataset

We requested Zanders to provide a list of companies in their dataset with data on one-year maturity CDS. This resulted in a list of 3,487 companies for which some data of interest was available. Subsequently, we refined our selection by cross-referencing these names with our selections from the Bureau van Dijk and Eikon databases to identify matching or similar entries. In the end, we requested all available one-year maturity CDS data for a list of 304 companies between 01-01-2014 and 23-11-2023 which resulted in 429,724 observations. For the research we only use the CDS data of January of each year, this choice is justified in Section 3.2, leaving us with 33,754 observations. It is important to note that companies may have multiple CDS, each differing in characteristics like “Seniority”, “RestructuringType”, and “Coupon”. To uniquely identify each CDS, we rely on the BBGCDSTicker, although it is worth mentioning that some BBGCDSTickers may be missing. It was chosen to always use a 'Senior' CDS which was available for each company in the dataset.

The dataset comprises 57 columns, with our primary focus being centered on the “ParSpreadMid” column. This column serves as our CDS spread, used for calculating the risk-neutral PD. For 15,483 of the 33,754 observations “ParSpreadMid” column had to be interpolated by taking the average of the “ParSpreadBid” and “ParSpreadOffer” columns, both of which are complete with no missing observations.

Appendix C

Results

Model		2018		2019		2020		2021		2022		Full sample	
Daily	HV 12	0.2629	0.00	0.2973	0.00	0.2556	0.11	0.5321	0.00	0.3066	0.17	0.3549	0.00
	HV 24	0.2699	0.00	0.2927	0.00	0.2540	0.10	0.4845	0.00	0.3756	0.00	0.3556	0.00
	HV 60	0.2700	0.00	0.2963	0.00	0.2539	0.10	0.4537	0.00	0.3321	0.02	0.3362	0.02
	GARCH	0.2694	0.00	0.3222	0.00	0.2540	0.10	0.4598	0.00	0.3297	0.02	0.3416	0.01
Monthly	HV 12	0.2606	0.00	0.3006	0.01	0.2564	0.14	0.5070	0.00	0.2994	0.35	0.3453	0.02
	HV 24	0.2699	0.00	0.2946	0.01	0.2525	0.10	0.4679	0.01	0.3663	0.00	0.3483	0.01
	HV 60	0.2692	0.00	0.3010	0.00	0.2546	0.11	0.4428	0.01	0.3280	0.03	0.3327	0.05
	GARCH	0.2669	0.00	0.3209	0.00	0.2566	0.15	0.4785	0.00	0.3338	0.01	0.3481	0.00
Benchmark	Zero	0.2445	0.00	0.2455	0.25	0.2905	0.27	0.4027	0.01	0.2494	0.00	0.2965	0.01
	FF3	0.2707	0.00	0.2424	0.31	0.2892	0.25	0.3707	0.59	0.2496	0.00	0.2904	0.00
	CAPM	0.2712	0.00	0.2444	0.21	0.2883	0.24	0.3616	0.06	0.2504	0.00	0.2882	0.00
	Hist. Avg.	0.3353	-	0.2327	-	0.3017	-	0.3742	-	0.2860	-	0.3114	-
N		125		128		130		161		173		717	

Table C.1: The annual and full sample RMSE for the different variants of our equity premium prediction model and the benchmark models. The number of observations in each period can be found at the bottom of the table. The two-sided p-values of a paired t-test on the squared errors compare the historical average benchmark to the other models.

Model		2018		2019		2020		2021		2022		Full sample	
Daily	HV 12	0.2016	0.00	0.2280	0.00	0.1990	0.01	0.3443	0.00	0.2251	0.64	0.2436	0.02
	HV 24	0.2067	0.00	0.2238	0.00	0.1974	0.01	0.3134	0.00	0.2715	0.00	0.2476	0.00
	HV 60	0.2084	0.00	0.2272	0.00	0.1968	0.00	0.2913	0.00	0.2416	0.08	0.2364	0.14
	GARCH	0.2080	0.00	0.2347	0.00	0.1968	0.00	0.2968	0.00	0.2408	0.09	0.2387	0.07
Monthly	HV 12	0.2012	0.00	0.2249	0.00	0.1982	0.01	0.3240	0.00	0.2199	0.99	0.2370	0.15
	HV 24	0.2058	0.00	0.2225	0.00	0.1962	0.01	0.2989	0.00	0.2615	0.01	0.2414	0.04
	HV 60	0.2066	0.00	0.2269	0.00	0.1964	0.00	0.2827	0.00	0.2392	0.12	0.2334	0.32
	GARCH	0.2045	0.00	0.2357	0.00	0.1975	0.01	0.3054	0.00	0.2442	0.05	0.2412	0.04
Benchmark	Zero	0.1929	0.00	0.1951	0.03	0.2169	0.03	0.2573	0.03	0.1828	0.00	0.2097	0.00
	FF3	0.2141	0.00	0.1900	0.09	0.2215	0.08	0.2324	0.56	0.1890	0.00	0.2272	0.00
	CAPM	0.2145	0.00	0.1919	0.05	0.2203	0.07	0.2272	0.21	0.1895	0.00	0.2092	0.00
	Hist. Avg.	0.2675	-	0.1769	-	0.2356	-	0.2369	-	0.2200	-	0.2083	-
N		125		128		130		161		173		717	

Table C.2: The annual and full sample MAE for the different variants of our equity premium prediction model and the benchmark models. The number of observations in each period can be found at the bottom of the table. The two-sided p-values of a paired t-test on the absolute errors compare the historical average benchmark to the other models.

RMSE Merton	2018	2019	2020	2021	2022	Full Sample
Manufacturing	0.2993	0.3292	0.2935	0.4523	0.3277	0.3482
Information and communication	0.2051	0.2151	0.1907	0.3085	0.2661	0.2446
Electricity, gas, steam and air co	0.3145	0.2864	0.2392	0.2099	0.1991	0.2461
Wholesale and retail trade	0.2704	0.2685	0.3035	0.2670	0.3038	0.2841
Mining and quarrying	0.1873	0.3531	0.1490	0.9829	0.3184	0.5427
Other	0.2427	0.2897	0.2229	0.3237	0.3930	0.3097

Table C.3: The RMSE of the Merton model for the five most common sectors for each year and the full sample. At the bottom of the table is the RMSE of the other observations can be found.

RMSE hist. avg.	2018	2019	2020	2021	2022	Full Sample
Manufacturing	0.3790	0.2439	0.2953	0.3621	0.2878	0.3176
Information and communication	0.3188	0.2172	0.1769	0.2270	0.2551	0.2429
Electricity, gas, steam and air co	0.3069	0.2247	0.2879	0.2506	0.2977	0.2756
Wholesale and retail trade	0.2881	0.2636	0.4201	0.2509	0.3259	0.3144
Mining and quarrying	0.3398	0.2119	0.1957	0.8589	0.1683	0.4678
Other	0.2855	0.2174	0.3361	0.2724	0.3058	0.2870

Table C.4: The RMSE of the historical average benchmark for the five most common sectors for each year and the full sample. At the bottom of the table is the RMSE of the other observations can be found.

RMSE Merton	2018	2019	2020	2021	2022	Full Sample
Low	0.2602	0.2720	0.2136	0.3046	0.3220	0.2798
Medium	0.2783	0.2212	0.2075	0.4123	0.2807	0.2968
High	0.2894	0.3983	0.3310	0.5866	0.3889	0.4220

Table C.5: The RMSE of the Merton model for the three carbon emission segments for each year and the full sample.

RMSE hist. avg.	2018	2019	2020	2021	2022	Full Sample
Low	0.2866	0.2146	0.3414	0.2092	0.2883	0.2714
Medium	0.3793	0.1661	0.2106	0.4091	0.2788	0.3080
High	0.3670	0.3081	0.3311	0.4830	0.3032	0.3678

Table C.6: The RMSE of the historical benchmark model for the three carbon emission segments for each year and the full sample.