University of St.Gallen

Machine Learning-Enhanced Pricing in the Secondary Market for Cat Bonds

Bachelor Thesis

Supervisor:

Prof. Dr. Alexander Braun

Institute for Insurance Economics

Author:

Niklas Leander Kampe

Weiherweg 3, 8610 Uster

Major in Economics (BVWL)

16-611-618

University of St. Gallen

May 22nd 2021

Abstract

The pricing of cat bonds represents a central research topic in the area of insurance-linked securities, which has repeatedly brought about new insights over the past few years. Within the scope of this thesis, cat bond pricing is dealt with regard to the optimization of the spread forecasting using machine learning approaches of Random Forests and Neural Networks. Therefore, benchmark models of OLS and Penalized Regressions are set up and applied, whereby the results define the main performance benchmark for the machine learning-based approaches to spread forecasting. The analysis is carried out using an extensive secondary cat bond market data set, which compromises the quarterly observations of virtually all cat bond tranches between the beginning of the 2000s and the end of 2020. The evaluation of the underlying pricing model and its determinants for the regression-based performance benchmark models as well as the machine learning-based optimization models uses existing econometric approaches of primary and secondary market analyzes, which are tested for significance based on the data set. On the one hand, the results of the main analysis show that no regression-based benchmark improvement in forecasting due to penalized regressions compared to OLS models can be accomplished. On the other hand, regarding the machine learning-based optimization, given the correct model specification and determination of hyperparameters, the random forest model achieves extraordinary and robust out-of-sample spread forecasting results. Significantly outperforming and robust out-of-sample results are also recorded by the neural network against the benchmark regression-based models, while, however, still clearly lagging behind the random forest performance. All in all, the results underline the strength and high potential of machine learning-based approaches to forecasting-related problem settings compared to multivariate regressions, in particular in the context of cat bond spreads.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AUS Australia

CEL Conditional Expected Loss

EL Expected Loss

EQ Earthquake

EUR Europe

HAC Heteroskedasticity and Autocorrelation Consistent

IG Investment Grade

ILS Insurance-Linked Securities

IS In-Sample

JP Japan

LA Latin America

LIBOR London Interbank Offered Rate

LM Langrange Multiplier

ML Machine Learning

NA North America

NIG Non-Investment Grade

NN Neural Network

OLS Ordinary Least Squares

OOS Out-of-Sample

PFL Probability of First Loss

RF Random Forest

RMS Risk Management Solutions

RMSE Root Mean Square Error

ROL Rate on Line

SPV Special Prupose Vehicle

TTM Term to Maturity

USAA United Services Automobile Association

USD US Dollar

VIF Variance Inflation Factor

1 Introduction

Since the beginning in 1997[1], catastrophe (cat) bonds have been an attractive alternative to traditional financial instruments who have enjoyed substantial growth in recent years (SwissRe, 2012, p.4). The unique structure that enables (re)insurers to transfer the financial risks from global natural disasters onto the financial market on the issuer side as well as the low correlation with other financial asset classes and attractive return opportunities on the investor side are the main drivers of such growth (Braun, 2014, p.1). The formal characteristics of cat bonds are special in that the payment of a regular coupon and the return of the nominal principal are possible objects to partial or complete losses due to the occurrence of predetermined types of natural disasters in specific regions as soon as a predefined trigger threshold is met.

The literature on cat bonds has increased significantly in interest and volume in recent years, especially with regard to econometric pricing approaches, whereby the focus lies on examining the spread, which defines the main determinant of the coupon and thus of the pricing of cat bonds. It turns out, however, that the primary market and therefore the spread at issuance with regard to its determinants represents the main research objective. In this context, the pricing approaches of Lane (2000), Lee and Yu (2003) as well as Lei et. al (2008) define the theoretical foundation for novel econometric approaches. On this occasion, the research paper by Braun (2014), who translated the underlying knowledge about the spread determinants under the addition of new determinants into a novel econometric model and concluded new empirical evidences about the cat bond spread at issuance. This work mainly constitutes the basis of new approaches that have been developed in the near past. In contrast, studies on the secondary market spread are still relatively rare, which is mainly due to the still not fully examined primary market approaches as well as the limited availability of secondary market data. One of the first studies under secondary market data was approached by Dieckmann (2010), who investigated the main drivers of the spread. To date, the research by Gürtler et. al (2012) represents the most comprehensive study of the secondary market spread, which also builds upon on the primary market econometric pricing model of Braun (2014) by further evaluating the determinants under secondary market characteristics.

Regression-based models are a widely used method in the area of financial instruments in order to be able to forecast individual parameters such as future prices as well as to be able to easily and understandably examine the individual determinants and their relationships with the dependent variable (see, e.g., Campbell and Thompson, 2008; Thornton and Valente, 2010). This also occurs in the context of econometric research approaches to cat bond spreads, both for the primary and for the secondary market (Braun, 2014, p.16; Gürtler et al., 2012, p.10). Such models are particularly suitable in that they provide a full picture of the relevant determinants on the basis of the coefficients and their statistical significance. With increasing complexity, but also with regard to the underlying assumptions, regression-based models repeatedly reach their limits and sometimes do not represent an optimal framework for forecasting-related problem settings, especially in the area of asset pricing.

[1] The first cat bond was issued in 1997 by Residential Re in order to protect the United Services Automobile Association (USAA) against the risk of major hurricanes in the United States (SwissRe, 2012, p.4). Since then, the total amount of issued cat bond risk capital sums up to around USD 135 billions (Evans, 2021c).

Machine Learning (ML) models represent an alternative and intensively used method to address these limits of regression approaches. In addition, research on forecasting financial assets' core metrics shows that ML methods produce increased performances, in particular in the context of cat bonds (Götze et al., 2020, p.24). However, one of the central challenges here is the comparatively small market, which means that the premise of extensive training data sets for ML models is not fully met. Nevertheless, the results of Götze et. al (2020, p.24) for the primary cat bond market show that, under comprehensive and careful model specification, optimized spread forecasting results can be achieved. Nevertheless, there is still no detailed study of ML models for spread forecasting based on secondary market data. This issue represents the main target of this thesis, which is dedicated to the performance comparison of the cat bond spread forecasting of benchmark regression models, i.e. Ordinary Least Squares (OLS) and Penalized Regressions, and ML models of Random Forests and Neural Networks.

The structure of the thesis is given as follows: The first step is to provide an introduction to cat bonds, to afterwards dive into the characteristics of the secondary market which together represent the general qualitative foundation. This is followed in the second step by evaluating the pricing determinants of cat bonds, in particular the determinants of the cat bond spread. This is done using cat bond-specific as well as macroeconomic and financial market determinants. The resulting underlying pricing is basically based on the results of Braun's econometric approach (2014) and Gürtler et. al's (2012) adaptation to the secondary market. This is followed by the explanation of the cat bond data set, which is then used for the data-based evaluation of the pricing model defined above. Finally, the main part of this thesis follows, which is dedicated to the application of ML-based models for forecasting the cat bond spread, i.e. Random Forests and Neural Networks. The performance evaluation is based on a comparison with various extended regression models, in particular OLS and Penalized Regressions. In order to satisfy the complex model specification explained above, comprehensive hyperparameter tuning is carried out using appropriate grid search procedures. Finally, to further evaluate the model specification and the forecasting results obtained, a robustness check is carried out.

The contribution of this work to the existing research around cat bond pricing is that it is still relatively low with regard to the secondary market, which is underlined by the recommendation of Braun (2014, p.28) of using a similar econometric approach on secondary time series data. Furthermore, the work by Götze et. al (2020) proves that there is great potential in the application of ML methods in spread forecasting and that it requires further practical application. The general usage of ML models in the area of the ILS market and especially cat bonds is still very low, whereby the results of this thesis offer a valuable contribution to this small research area and a reference benchmark for further studies. In addition, the connection of ML methods with secondary market studies of the spread represents a new application at the time of publication. Furthermore, the thesis comprises one of the most extensive cat bond data sets, which contains tranches from the beginning of the 2000s to the end of 2020. Thus, the existing results of the econometric models by Braun (2014) and Gürtler et. al (2012), under the introduction of additional, data-based determined variables, can be reassessed under more recent data. Ultimately, the results should provide both re(insurers) and investors with a practical foundation in understanding and determining the secondary market spread on the basis of ML models.

2 Cat Bond Pricing: Linear Modeling

This chapter is dedicated to an introduction to cat bonds and its secondary market structures as well as to the evaluation of the cat bond pricing determinants, which build the foundation to the definition of the underlying econometric pricing model. The evaluation is carried out based on existing research theories, taking into account the focus on the secondary market and the pricing optimization to be achieved. Afterwards, this model serves as the basis for the main analysis of this thesis, which is dedicated to the pricing optimization.

2.1 Cat Bonds at a Glance

Cat bonds represent a central component of Insurance-Linked Securities (ILS), which generally provide protection to (re)insurers, as well as governments and corporations. Since the mid 1990s, cat bonds have served as an attractive financial instrument for both investors and issuers. The basic structure shows the special features of cat bonds (see also Figure 1): The sponsor - re(insurer) or other institutions (this work henceforth uses the term reinsurer) - concludes a risk transfer contract (or catastrophe swaps) with an explicitly set up Special Purpose Vehicle (SPV). With the target of capitalization, the SPV issues notes - in this case cat bonds - to the financial market and its investors in the amount of the limit of the risk transfer contract whereby the raised capital is held in a collateral/trust account. Thus, the deposit is capped according to the reinsurance contract and the investment instrument for investors is timely limited by the predefined maturity. The last central component is defined by the trigger, which includes the circumstances under which the paid-in capital is withdrawn in full (or in part) from the trust account to cover the reinsurers' claims due to natural catastrophes. If the predefined trigger threshold is not reached, investors receive the paid-in capital back in full after maturity (equivalent to other bond instruments). In addition, the investors receive regular coupon payments over the term of the bond, which are paid for the associated investment risks and are relatively high in contrast to classic financial assets due to the possibility of a total or partial loss of the paid-in capital. Finally, the SPV has the opportunity to invest parts of the principal in return-stable investment opportunities in order to be able to cover the coupon payments to the investors. (SwissRe, 2012, p.3; Braun, 2014, p.3-5)

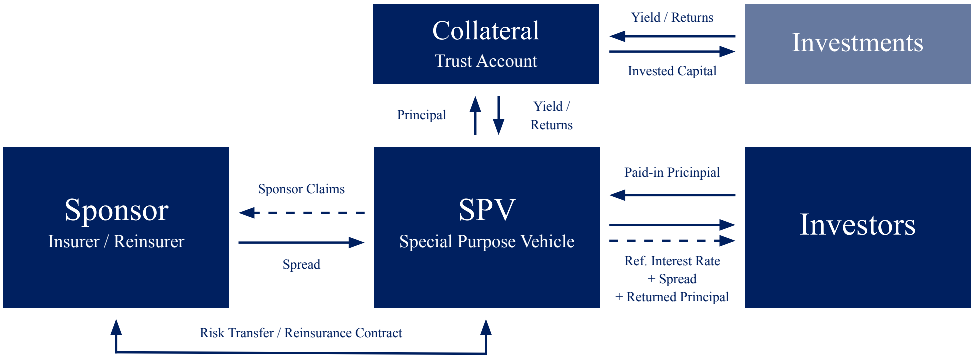

Figure 1: General Cat Bond Structure - Own illustration (Braun, 2014, p.5; SwissRe, 2012, p.3)

The attractiveness of cat bonds as financial instruments for sponsors and investors is further proved by the issuance volumes over the last two decades, which peaked in 2020 at around USD 16.4 billion with a total of USD 46.4 billion outstanding. In addition, a further USD 7.5 billion of cat bonds had already been issued in 2021 until the end of April, which corresponds to an outstanding volume of USD 48.28 billion at this point in time (Evans, 2021a). Neglecting the consequences of the global financial crisis in 2007, a constantly growing cat bond market can be identified, which is shown by Figure 2[2]:

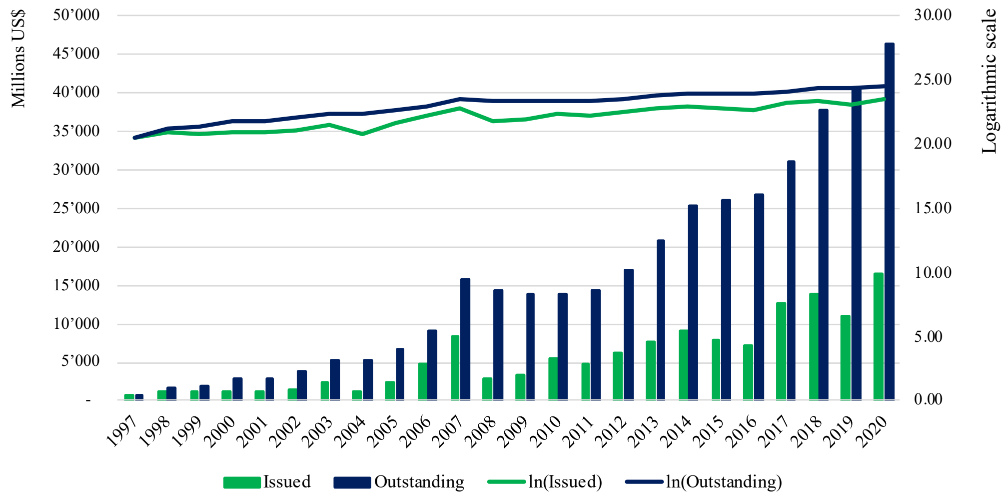

Figure 2: Cat Bond & ILS Risk Capital - Issued vs. Outstanding - Own illustration (Evans, 2021c)

The fundamental advantages of cat bonds as an alternative to classic financial assets from an investor's point of view is essentially characterized by the high coupon payments and thus high return opportunities (SwissRe, 2012, p.10). As of the end of April 2021, the coupon payments for more than half of outstanding cat bonds were between 4% and 12% (Evans, 2021b). Furthermore, cat bonds have a low correlation with market or credit risk, which primarily arises from the special characteristics and hence leads to high diversification opportunities compared to other financial instruments (Weistroffer, 2010, p.4). Lastly, cat bonds are characterized by a stable pricing structure and thus low volatility over the entire term, which is usually between three to five years (SwissRe, 2012, p.10). From the sponsors' point of view, the attractiveness is shown on the one hand by the transfer of the risks occurring from natural catastrophes to the financial market and thus also by the controllable management of counterparty credit risks related to natural catastrophes (SwissRe, 2012, p.9; Weistroffer, 2010, p.4). Additionally, such securitisation offers additional risk diversification for the sponsor itself and opens up a new funding source by monetizing embedded values from policies written (SwissRe, 2012, p.10). On the other hand, sponsors also benefit from the pricing stability over the term, which supports the relevant risk management (SwissRe, 2012, p.9).

[2]The bar chart component, which shows the absolute numbers of the issuance volumes and outstanding volumes, refers to the left y-axis and the line diagram, which shows their respective growth paths over time, refers to the right y-axis.

2.2 Secondary Cat Bond Market Characteristics

Compared to the primary market, which received the main focus in cat bond research, the secondary market has essential peculiarities that need to be understood in the context of the pricing analysis of this thesis. Cat bonds are generally traded on an active secondary market, with intermediary brokers bringing buyers and sellers together through indicative bid and ask spreads (Risk Management Solutions, 2012, p.8). The core group of investors is made up of dedicated ILS managers who hold around 70% of newly issued cat bonds, with the remainder being mostly held by institutional investors, money managers and pension funds (Risk Management Solutions, 2012, p.8).

While the predefined determinants of the cat bond at issuance offer a clear foundation for pricing in the primary market, investors in the secondary market are faced with changing determinants over time to maturity, e.g. EL[3] valuations, bond ratings, seasonal and cyclical fluctuations of natural catastrophes as well as the market environment. The adjustment of the cat bond-specific determinants over time is mainly due to the fact that the risk situations covered by the bond can change over the term, whereupon the determinants have to be adjusted and thus the pricing in the form of the spread is object to change accordingly. Seasonal fluctuations are based on the fact that the cat bond market is subject to a cyclical component that is driven by the probability of occurrence of natural disasters and the associated higher trigger probability. This ultimately leads to reoccurring increased probabilities of loss over the term for the investors. Cyclical components can be related specifically to the bond and its covered natural catastrophes as well as the whole reinsurance cycle itself (Risk Management Solutions, 2012, p.8). In practice, it has been shown that the occurrence of natural disasters lead to a "hard market" situation, which is associated with low supply and increasing reinsurance prices (Risk Management Solutions, 2012, p.8). During such conditions, the market is faced with cheaper cat bonds, which in turn offers sponsors attractive conditions for issuing new bonds and ultimately increases the market volume (Risk Management Solutions, 2012, p.8). On the other hand, there is the situation of a "soft market" in which the reinsurance prices fall fundamentally and the market volumes decrease/stagnate because sponsors fall back on traditional reinsurance due to high issuance costs (Risk Management Solutions, 2012, p.8). With regard to the pricing and the spread as the main determinant, it can be seen that secondary market spreads are often higher in times of hard markets than in times of soft markets, which seems intuitive with regard to the definition (Risk Management Solutions, 2012, p.8).

In addition, it can be assumed that macroeconomic and financial market conditions also have an impact on the secondary market next to cat bond- and reinsurance-specific circumstances, even despite the low correlation due to the abnormal structure of cat bonds compared to traditional financial assets. As Risk Management Solutions (RMS) (2012) state in their report, the pricing of ILS, and thus also of cat bonds, basically depends on reinsurance pricing, which in turn is strongly related to frequency and the probability of the occurrence of corresponding natural catastrophes, as stated above. With regard to the influence of financial markets, it can also

[3] The EL defines one of the main determinants of the risk associated with holding cat bonds and is calculated specifically for each bond by specialized risk modeling companies during the structuring process (Risk Management Solutions, 2012, p.8). A more detailed explanation will follow in chapter 2.3.1.

be seen that there is an indirect relationship, which became apparent, for example, during the financial crisis of 2008 and the associated low market liquidity, in which the demand for cat bonds as a source of liquidity for investors increased (Risk Management Solutions, 2012, p.8). Lastly, practice shows that the secondary market for cat bonds is exposed to high volatility on the one hand, but also receives increasing liquidity over time, which underlines the steadily growing attractiveness of cat bonds in the secondary market as well (Braun, 2014, p.2; SwissRe, 2012, p.16).

2.3 Cat Bond Pricing Determinants

The evaluation of the cat bond pricing determinants is closely aligned to the existing research from the primary and secondary market. The underlying research maily focuses on Braun (2014) for the primary market and Gürtler et. al (2012), who expanded their model to the secondary market on the same foundation. In this context, a distinction is made between cat bond-specific as well as macroeconomic and financial market determinants. With the addition of novel model extensions in terms of new variables, a data-based evaluation is carried out using an OLS regression in order to check the influence of the determinants on the spread as well as their intensity and statistical significance. Therefore, the results represent the basis of the benchmark regression models as well as the optimizing ML models for cat bond spread forecasts.

2.3.1 Cat Bond-specific Determinants

One of the most essential determinants for the cat bond spread is defined by the *Expected Loss (EL)*, which is broadly confirmed across existing research (see, e.g., Lane and Mahul, 2008; Galeotti et al., 2012; Braun, 2014). The EL represents the loss cost or average annual loss, i.e. the average value of losses over a full range of scenarios, which is determined by specialized modeling companies such as Applied Insurance Research Worldwide, Risk Management Solutions, Inc. or Eqecat, Inc (Gürtler et al., 2012, p.13; PartnerRe, 2015, p.2). The calculation is based on the Probability of First Loss (PFL) and the Conditional Expected Loss (CEL), according to the formal notation $EL = PFL * CEL$ (Braun, 2014, p.6). Since the EL thus represents the percentage loss of the nominal value within a year, it is already clear from the definition that there should be a strongly positive correlation with the cat bond spread, whereby the spread can be viewed as a risk compensation (Gürtler et al., 2012, p.12).

Another important determinant of the spread is the *Trigger Type*, which connects the payout of a cat bond with the covered catastrophe (Galeotti et al., 2012, p.18). A distinction must be made between five central trigger mechanisms: The Indemnity Trigger refers the actual loss amount of the sponsor, the Parametric Trigger is constructed using physical measures, the Index Trigger refers to a pre-specified index, the Modeled Loss Trigger is constructed under the usage of specific catastrophe modeling software and lastly the Hybrid/Multiple Trigger adapts combinations of different triggers (Galeotti et al., 2012, p.18). All triggers are constructed in order to define the payout in case of a covered catastrophe. Cummins and Weiss (2009) as well as Galeotti et. al (2012, p.18), next to other researchers, suppose that indemnity triggered cat bonds are related with higher spreads compared to cat bond prices with different trigger mechanisms. This is mainly due to a more extended verification process of losses after the occurrence of natural disasters as well as possible moral hazards by the sponsor based on information asymmetries (Braun, 2014, p.7; Gürtler et al., 2012, p.4). The trigger type is therefore evaluated as an influential determinant and should, if necessary, be defined as a dummy variable with reference to the indemnity trigger, which will be considered in more detail at a later stage.

One of the main characteristics of cat bonds and also determinants of the spread are the *Peril Types* and *Peril Regions*, and in particular their combination. They define which natural hazard risk is covered by the bond, in the form of a specified combination of a peril region (covered territory) and a peril type (reference peril) (Braun, 2014, p.5; Jaeger et al., 2010). In research, the regions North America, Europe, Japan and Rest of the World as well as the types Earthquake, Hurricane (including cyclones, typhoons and other tropical storms, depending on regional significance), Wind (including windstorms, winterstorms, thunderstorms and other storms to covered by hurricanes) and Others are often specified (see, e.g., Gürtler et al., 2012, p.14). In addition, multiple risk coverage in the form of multiple peril types and regions is also possible and used extensively in practice. This suggests that a higher number of covered peril combinations with a higher spread for investors due to increased risk of a triggering event. Empirical results, e.g. that of Berge (2005), underline the number of perils as a significant spread determining factor. It can therefore be recommended to consider the number of peril types and regions as valid spread determinants. This was analogously taken into account by Guertler et. al (2012, p.3) and confirmed high significance for the secondary cat bond market.

In addition to the determinants already discussed, classic components of cat bonds in comparison with other financial assets are also to be assumed as determinants of the spread, especially the *Volume*, the *Maturity* and the *Term to Maturity*. The latter shows its relevance especially in secondary market analyzes. Analogous to traditional bond instruments (Edwards et al., 2007), research on cat bond pricing shows that a higher volume is associated with a lower spread, which in turn is due to lower transaction costs and the associated higher fungibility and ultimately lower required yields to maturity (Braun, 2014, p.7). Even if the results of the primary market analysis by Braun (2014, p.28) show that there is only a slight influence, this can be more significant when using more extensive data or when analyzing under secondary market characteristics. With regard to maturity and the associated term to maturity for secondary market observations, one can assume that a liquidity premium will be added in the form of a higher spread as a result of longer maturities. Even if Diekmann (2010) and Guertler et. al (2012, p.31) could not find any empirical evidence, it is advisable to also take into account the maturity and the term to maturity equivalent to the volume in order to check the influence based on any underlying data or market differences.

As a further determinant, similar to classic financial assets, the *Bond Rating* is taken into account. The specialty here is that the rating is mainly influenced by the occurrence probability of a triggering event because cat bonds are fully collateralized and rating agencies are not specialized in cat bond characteristics (Cummins and Weiss, 2009; Galeotti et al., 2012, p.20).

Based on the explanations above, the determinant of the EL has a major influence on the rating[4]. Cat bond ratings are generally issued by the three leading rating agencies Standard & Poor's, Moody's and Fitch. The rational assumption that a lower rating leads to a higher spread also applies to cat bonds and has already been empirically proven in research (see, e.g., Gürtler et al., 2012, p.25; Galeotti et al., 2012, p.27). In practice, a distinction is often made between non-investment-grade and investment-grade ratings, on which dummy variables are defined (Braun, 2014, p.19). In addition, there are approaches to define further ordinal variables which quantify the exact ratings in more detail in order to allow an even more granular approach, which shows high significance (Gürtler et al., 2012, p.25). Such ordinal modeling is also suitable for this secondary market analysis and is derived in Appendix 1 on the basis of classic ratings by the agencies stated above.

A final determinant, which has received more attention especially in the last few years, is the *Seasonality* of the determinants of peril types and regions and their combination in particular. One of the reasons for this is that, by definition, it only affects secondary market analyzes which are clearly lagging behind primary market analyzes of cat bonds regarding research volume and depth. In general, it can be assumed that the seasonally increased probability of the occurrence of a triggering catastrophe in a predetermined region leads to a temporarily higher spread. So far, two central concepts have been used in practice, which on the one hand provide for the exclusion of observations influenced by seasonality (Gürtler et al., 2012, p.11-12) or on the other hand follow a corresponding modeling using dummy variables (Herrmann and Hibbeln, 2019, p.15). The latter results provide solid empirical evidence that such consideration in a conceptual econometric framework shows that seasonal fluctuations define a major driver of cat bond spreads (Herrmann and Hibbeln, 2019, p.30). Thus, seasonality is set as a further determinant, which should be modeled as a dynamic dummy variable in the main analysis of this thesis based on the cat bond data set.

2.3.2 Macroeconomic & Financial Market Determinants

In addition to the cat bond-specific determinants, existing research shows that the macroeconomic and financial market environment also has an influence on the cat bond spreads. The *S&P500 Index* return is defined as the benchmark determinant for the financial market, which is in line with the secondary market model by Gürtler et. al (2012, p.14). Despite the argument of low correlation between cat bonds and the financial market due to the special structuring characteristics, it has been shown that there is still a small degree of dependency. In addition, Carayannopoulos and Perez (2015, p.1) proved that the dependence is even larger in times of financial market crises. Therefore, all in all, the influence of the financial market on cat bond spreads should be tested for significance in the data-based evaluation at a later stage (see, e.g., Gürtler et al., 2012, p.37).

[4] Such an alleged correlation of two determinants can lead to the problem of multicollinearity within regression analyzes. For this reason, too, a check follows in the later part of this thesis in order to take any measures so that the regression results are not influenced.

Another determinant that also describes the influence of the financial market environment are *Corporate Credit Spreads*. They show whether there is an influence of corporate credit spreads on cat bond spreads. Cat bond-specific research mainly focused on the U.S. corporate spreads by Bank of America (BofA) Merrill Lynch for different maturities and different rating classes for granular analysis (Gürtler et al., 2012, p.14); 2020, p.15). Intuitively and rationally, a positive correlation between corporate credit spreads and cat bond spreads is assumed and should be included in the underlying pricing model.

Finally, the *Reinsurance Cycle* defines the last macroeconomic determinant. It describes the cyclical movements in the reinsurance industry and describes the relationship between cat bond spreads and those cycles. Early fundamental models, such as those by Lane and Mahul (2008), highlighted such cyclical dynamics of cat bond spreads. In practice, the Guy Carpenter Global Catastrophe Rate-on-Line (RoL) Index is often used as a benchmark, in which the rate on line is determined as the premium divided by the insured limit (Gürtler et al., 2012, p.14). Empirical results show that cat bond spreads underlie cyclical movements similar to the reinsurance cycle and thus there is a positive correlation (see, e.g., Gürtler et al., 2012, p.40; Braun, 2014, p.28). In addition, the results suggest that the use of the RoL index as a variable for reinsurance cycles is well suited.

2.3.3 Evaluation of the Underlying Pricing Model

After the central cat bond-specific as well as macroeconomic and financial market determinants have been elaborated, the evaluation of the linear pricing model follows in this section. It defines a central part of this thesis in that it serves as the foundation for the benchmark regression models as well as for the ML models for optimizing the spread forecasting. The model is particularly based on the primary market model by Braun (2014) and the adapted secondary market model by Guertler et. al (2012). In the first step, the regular coupon payment rates are defined on the basis of the structure of cat bonds in chapter 2.1 as follows:

$$CouponRate_{i,t} = ReferenceFloatingInterestRate_{i,t} + Spread_{i,t} \tag{1}$$

where the reference floating interest rate is often defined on the basis of LIBOR or any treasury rates. It follows from (1) that the spread represents the variable of interest which governs the pricing of a cat bond and that the spread is in the foreground in the context of the main analysis. As already emerged from the cat bond-specific determinants, the EL is an essential part of the model and defines the main driver of the spread. In addition to the models by Braun and Gürtler, this assumption is confirmed by a wide variety of research projects (see, e.g., Lane, 2000, p.274; Wang, 2000; Lane and Mahul, 2008, p.5; Galeotti et al., 2012, p.31). Thus, the cat bond spread is defined as follows:

$$S_{i,t}^{Cat} = EL_{i,t} + MarginalTerm_{i,t} \tag{2}$$

with $S_{i,t}^{Cat}$ [5] referring to the spread of cat bond $i = 1, ..., n$ at observation time t, $EL_{i,t}$ to the specific EL which can also vary over time and finally a bond- and time-varying marginal term[6]. Based on the previously evaluated pricing determinants, the marginal term can be formally noted as follows:

$$MarginalTerm_{i,t} = f(CatBondDeterminants_{-EL_{i,t}}, MacroeconomicDeterminants_t) \quad (3)$$

where f is defined as a function of cat bond-specific determinants excluding the EL as well as macroeconomic[7] determinants for cat bond i at time t. The form of the function f is assumed to be linear in the context of the underlying pricing model, which is in accordance with existing econometric approaches of Braun (2014, p.20) and Gürtler et. al (2012, p.19) for both the primary and the secondary cat bond market. The underlying spread model can thus be ultimately simplified as follows and generalized with regard to the elaborated determinants:

$$S_{i,t}^{Cat} = \beta' X_{i,t} + \gamma' Y_t + \epsilon_{i,t} \quad (4)$$

where $S_{i,t}^{Cat}$ defines the spread of cat bond i at time t, $X_{i,t}$ refers to the cat bond-specific determinants at time t with the vector of weights β', Y_t to the macroeconomic and financial market determinants with the vector of weights γ' and lastly $\epsilon_{i,t}$ defines the random error term. With regard to the weightings, the model forms a suitable basis for regression-based analyzes, whereby the weighting vectors can be quantified by the corresponding regression coefficients. Furthermore, the investigation of the statistical significance of the coefficients leads to the decision as to which extend cat bond-specific as well as macroeconomic and financial market determinants should be included in the model. This happens on the basis of a comprehensive cat bond data set in the following chapter. In comparison, ML approaches break up the strict linearity of this model and use alternative modeling approaches based on the same variables to determine the cat bond spread. This represents the focus of the main analysis of this thesis in conjunction with the benchmark regression-based linear models.

[5]The term $S_{i,t}^{Cat}$ is in line with the termination by Braun (2014), where Braun's spread refers to the primary market and the spread in this thesis also includes to the time-varying secondary market (see subset t).

[6]Other research papers sometimes describe the marginal term as a "risk premium" (see, e.g., Braun, 2014, p.4). However, since this notation can be misunderstood because a higher EL also refers to a risk premium for investors in a classic financial understanding, the term marginal term is used as part of this thesis.

[7]The term of macroeconomic determinants also includes financial market determinants and thus all determinants that have been derived in chapter 2.3.2.

3 Cat Bond Data Set

The analysis and evaluation of the underlying pricing model as well as the subsequent optimization through ML models is carried out on the basis of an extensive secondary market data set of cat bond tranches, which haven been provided by leading companies in the ILS sector. The initial data sets are first explained on a high level, whereupon the data is extensively prepared, merged and checked for regression-related data issues. Afterwards, the resulting final data set is described under descriptive statistics. Finally, the pricing model is evaluated to examine the spread on the basis of an OLS regression and is primarily tested for statistical significance. The ultimately resulting model defines the foundation of the ML models and the extended regression models for benchmark comparison.

3.1 Data Generating Process

The initial data sets of cat bond tranches were provided by AON, Swiss Re and Lane Financial. In addition, based on the evaluated macroeconomic and financial market spread determinants and the upon defined cat bond pricing model, the yearly returns of the Global Property Catastrophe Rate-On-Line Index by Guy Carpenter, the quarterly returns of the S&P500 Index from Standard & Poor's as well as the quarter-end and rating-specific US Corporate Credit Spreads[8] by Bank of America (BofA) Merrill Lynch are used for the underlying macroeconomic data. Table 1 includes an overview of the raw data sets with the core information on the number of tranches, the total number of observations, the earliest and most recent issue date and the earliest and most recent reference date of the secondary market observations, as well as information on whether the data record only contains primary or also secondary market data. With regard to the macroeconomic and financial market data, the focus only lies on the number of data observations as well as the earliest and most recent reference dates:

	Unique	Obs.	Issue (Min)	Issue (Max)	Ref (Min)	Ref (Max)	Prim.	Sec.
Cat Bond-specific Data								
AON	357	7906	2013-05-31	2020-12-22	2017-05-31	2020-12-31	Yes	Yes
Lane Financial	798	798	1996-10-01	2021-03-18	1996-10-01	2021-03-18	Yes	No
Swiss Re	795	7883	2003-06-19	2020-12-21	2001-06-29	2020-12-31	Yes	Yes
Macroeconomic Data								
Standard & Poor's	N/A	124	N/A	N/A	1990-03-31	2020-12-31	N/A	N/A
Guy Capenter	N/A	31	N/A	N/A	1990-12-31	2020-12-31	N/A	N/A
BofA Merrill Lynch	N/A	1649	N/A	N/A	1996-12-31	2020-12-31	N/A	N/A

Table 1: Descriptive Statistics - Raw Data Sets

As stated in Table 1, the issue and reference dates of the observations of the cat bond tranches are between October 1996 and December 2020 and virtually represent the entire universe of the cat bond market. In addition, it can be seen that the data set from Lane Financial only contains

[8]The US Corporate Credit Spreads are based on the quarter-end, rating-specific ICE BofA US Corporate Index Option-Adjusted Spread (OAS) Indices, whereby the spreads are calculated as the difference between a computed OAS index of all bonds in a given rating category and a spot treasury curve (FED of St. Louis, 2021).

primary market data which are used to map missing data which do not vary over time. Different ratios between the unique cat bond tranches and the total number of observations of the data sets from AON and Swiss Re are essentially based on different frequencies in data reporting, which will be uniformed in the preprocessing procedure. Ultimately, the AON and Swiss Re data sets define the foundation for the upcoming cleaning, mapping and merging process.

An extensive data cleaning process is carried out on the basis of the raw data sets. In the first step, the necessary time-fixed and time-varying variables are defined on the basis of the pricing model, which must be included in the final data set. Since both the AON and Swiss Re data sets do not contain all the necessary data, they are mapped with the help of Bloomberg, Thomson Reuters Eikon, the Deal Directory of the online portal Artemis and the primary market data set from Lane Financial based on a unique asset identifier. In the second step, after the missing data have been mapped and added, all cat bond tranches are removed, which do not include the cat bond-specific variables defined in the pricing model. Furthermore, the quarterly frequency of the observations is set, since this represents the lowest common frequency across data sets. Higher frequented observations are thus removed. The dates for the quarter end are defined as the fiscal year quarter ends in March, June, September and December. In addition, Artemis' Cat Bond Losses Directory is used to remove any cat bond tranches that were partly or fully triggered during maturity. As a result, 46 additional cat bond tranches are omitted. Ultimately, observations that have implausible data are also removed, which includes negative terms to maturity (TTM) and zero bond volumes. Hence, a total of 13 implausible observations are omitted. In the third step, the cleaned data sets from AON and Swiss Re are merged. This is done using a unique asset identifier, with the AON data set being the underlying data set, which is supplemented with additional observations and additional cat bond tranches from the Swiss Re data set. In the last step, the macroeconomic and financial market data are mapped for each observation. For the quarterly S&P 500 returns, the reference quarter end is used as the mapping identifier, whereby the return of the RoL Index uses the reference year. Ultimately, the US Corporate Credit Spreads are mapped by the reference quarter end as well as the aggregated bond rating. Unrated cat bond tranches are mapped using the BB rating.

In addition to the cat bond-specific as well as the macroeconomic and financial market data, it was evaluated in chapter 2.3.1 that the seasonality of the peril type and region combinations plays a central role in the secondary market analysis of cat bonds. It is shown that a seasonally increased risk of the occurrence of a covered peril type in a specific region leads to a higher risk premium in the form of a higher spread for investors. Thus, for all combinations of natural disasters, such as hurricanes (including cyclones, typhoons and tropical storms), wind storms and winter storms and the regions, such as North America (NA), Japan (JP), Europe (EUR), Australia (AUS) and Latin America (LA), it gets quarterly defined whether there is a pre-season, high season, after-season or no-season for the peril. Consequently, a dummy variable "Seasonality" is defined, which takes on the value of 1 if a cat bond covers a peril type in a region at the end of the reference quarter that is in the high season at this point in time. This also applies if a cat bond covers several peril types and regions. This definition of the dummy variable is in line with the research paper on seasonality of cat bonds by Herrmann and Hibbeln (2019, p.15-16). Since the occurrence of earthquakes is not subject to any seasonality, no seasons are assigned to them (Herrmann and Hibbeln, 2019, p.11).

After the cleaning, merging and mapping process, the final data set consists of 660 cat bonds and 6914 quarterly observations. Based on the evaluated pricing model in chapter 2.3.3, the most essential included variables of the final data set are defined as follows:

Variable	Description
Spread	Spread as of reference date in addition to reference floating interest rate (variable of interest)
ISIN	"International Securities Identification Number" - Unique identifier for cat bonds
Name	Name of the cat bond tranche
Issuance Date	Date when the cat bond was issued
Maturity Date	Date when the cat bond matures/expires
Term	Difference between maturity date and issuance date (in months)
TTM	Difference between maturity date and observation/reference date (in months)
Volume	Total amount of the accumulated, paid-in principals
EL	Modelled loss percentage of the nominal value within one year
Peril Type	Types of natural disasters covered by the cat bond
Peril Types	Number of peril types covered by the cat bond
Peril Region	Natural disaster affected geographical regions covered by the cat bond
Peril Regions	Number of geographical regions covered by the cat bond
Seasonality*	1 = any included peril combination that has high season as of reference date; 0 = else
Trigger Type	Type of trigger that leads to partial/full payout
Trigger Type*	1 = indemnity; 0 = else
Rating Aggregated	Uniformed bond rating (aggregated by S&P, Moody's and Fitch)
Rating	0 = No rating; 1 = B; 2 = BB; 3 = BBB; 4 = A; 5 = AA; 6 = AAA
Rating Grade	Investment, non-investment grade or non-rated grade based on aggregated rating
Ref Date	Date of observation/Reference date (quarterly)
S&P500 Return	Quarterly return of the S&P500 Index
RoL Index	Year-on-year percentage change of the Guy Carpenter Rate-on-Line Index
Corp. Spread	US corporate credit spreads of different rating classes as of reference date

* = Dummy Variable

Table 2: Variables Selection

3.2 Descriptive Statistics

After the final data set with 660 cat bonds and 6914 observations has been successfully processed, this section provides descriptive statistics of the pricing determinants resulting from the evaluated pricing model. Table 3 shows the cat bond-specific categorical variables. Here, the number of observations as well as the proportions of the total amount of observations are given. It should be noted that the peril types and peril regions do not aggregate to 100%, since the risk coverage of multiple types and regions is a main characteristic of cat bonds.

The descriptive statistics show that most cat bonds have an indemnity trigger (43.42%). Furthermore, the peril type of hurricanes, which also include typhoons, cyclones and tropical storms depending on the geographical region, represents the most secured risk with 67.04% and North America is the most covered region with 82.12%. In addition, a look at the more detailed listing of possible peril type and region combinations shows that hurricanes in North America are covered by almost 62.5% of all cat bonds, which underlines the previous thesis. Ultimately, it can be said that almost 52% of all cat bonds have a non-investment grade and are therefore subject to a BB or B rating and that almost 43.7% of all cat bonds are not rated[9].

[9]The definition of the aggregated ratings of AAA, AA, A, BBB, BB and B, based on the original ratings by S&P, Moody's and Fitch, can be found in Appendix 1.

	Count	Percentage
Trigger Type		
Indemnity	3002	43.42
Industry Loss Index	1809	26.16
Parametric	996	14.41
Mortality Index	313	4.53
Modelled Loss	295	4.27
Multi	499	7.22
Subtotal - Trigger Type	6914	100
Peril Type		
Hurricane	4635	67.04
Earthquake	4300	62.19
Wind	2597	37.56
Other	1605	23.21
Peril Region		
North America	5678	82.12
Europe	1857	26.86
Japan	1188	17.18
Other	913	13.21
Peril Combinations		
NA - Hurricane	4320	62.48
NA - EQ	3631	52.52
NA - Windstorm	1247	18.04
NA - Winterstorm	856	12.38
JP - EQ	796	11.51
JP - Typhoon	339	4.9
JP - Winterstorm	0	0
JP - Windstorm	0	0
EUR - Windstorm	1389	20.09
EUR - EQ	108	1.56
EUR - Winterstorm	31	0.45
AUS - Cyclone	225	3.25
AUS - EQ	211	3.05
LA - Hurricane	518	7.49
LA - EQ	481	6.96
Rating		
AAA	0	0
AA	11	0.16
A	22	0.32
BBB	259	3.75
Subsubtotal - IG	292	4.22
BB	2407	34.81
B	1194	17.27
Subsubtotal - NIG	3601	52.08
NR	3021	43.69
Subtotal - Rating	6914	100

Table 3: Descriptive Statistics - Categorical Variables

Table 4 describes the characteristics of the continuous variables, which contain both cardinal cat bond-specific and cardinal macroeconomic and financial market determinants. The number of observations as well as the statistical values of the mean, the standard deviation and the four quartiles are examined in order to get an impression of the distribution of the values. The units of measurement of the variables are given in brackets if they differ from absolute units.

The descriptive statistics show that the average spread is 6.75% and the average EL is 2.3%, which corresponds to an approximate ratio of 3. This spread-EL ratio is in line with existing research and underlines the stylized fact stated by Dieckmann (2010, p.1). In addition, the cat bonds have an average volume of almost USD 147 million, although there is a large difference between the smallest and largest observations. Meanwhile, the average term of a cat bond is just about 42 months, which is also in line with the stylized fact that cat bond terms are normally between three and five years. As indicated in the previous section, some cat bonds have multiple

	Obs	Mean	Std. Dev.	Min	Q25	Med	Q75	Max
Spread (%)	6914	6.75	4.55	0	3.75	5.89	9	39
EL (%)	6914	2.3	2.31	0	1	1.58	3	16.59
Volume ($mio)	660	146.9	119.68	1.69	70	117.5	200	1500
Term (mth)	660	42.07	11.16	5.95	36.2	37.32	48.3	84.63
TTM (mth)	6914	22.58	14.18	0	11.24	21.34	32.45	83.25
Peril Types	660	1.91	1.01	1	1	2	2	4
Peril Regions	660	1.4	0.69	1	1	1	2	4
—								
S&P500 Return (qr, %)	68	2.86	8.39	-21.94	0.28	4.21	7.03	20.54
Corp. Spread (%)	186	4.02	2.25	1.01	2.44	3.51	4.69	16.59
RoL Index (yr, %)	18	-0.71	10.27	-11.2	-8.8	-3.7	5	36.59

Table 4: Descriptive Statistics - Continuous Variables

peril types and regions. This can be seen in the average number of types and regions of 1.91 and 1.4 respectively. In line with previous research, the maximum number is limited to four types or regions at a time (see, e.g., Gürtler et al., 2012, p.17). With regard to the macroeconomic and financial market data, an average quarterly financial market return based on the benchmark index S&P 500 of 2.86% is shown. Furthermore, the average US corporate credit spread is just about 4% and thus 2.75% below the cat bond spread, which is in line with the assumed higher risk premium for investors associated to cat bonds. Ultimately, the RoL Index recorded an average annual return of just under -0.7%.

Next to the descriptive statistics, Table 5 shows the correlation matrix and thus the respective correlations between the dependent and explanatory variables in order to get a first impression of the explanatory content and its directions as well as of possible multicollinearity, which will be evaluated further in the next section. In addition, Appendix 2 shows the correlation heat map based on these values which displays the correlations dynamically.

	Spread	EL	Volume	Term	TTM	Peril Types	Peril Regions	Seasonality*	Trigger Type*	Rating	S&P500	Corp. Spread	RoL Index
Spread	1.00												
EL	0.75	1.00											
Volume	-0.23	-0.18	1.00										
Term	-0.33	-0.24	0.08	1.00									
TTM	-0.14	-0.08	0.05	0.47	1.00								
Peril Types	0.30	0.27	-0.12	0.01	0.03	1.00							
Peril Regions	0.14	0.2	-0.19	0.08	0.02	0.15	1.00						
Seasonality*	0.18	0.16	-0.06	-0.07	-0.05	0.31	0.13	1.00					
Trigger Type*	-0.17	-0.15	0.14	0.03	0.04	0.14	-0.44	-0.04	1.00				
Rating	-0.17	-0.42	-0.08	0.01	-0.05	-0.14	-0.15	-0.04	-0.16	1.00			
S&P500	0.01	0.03	0.03	0.03	0.01	0.03	-0.02	0.01	0.04	-0.06	1.00		
Corp. Spread	0.17	0.04	-0.1	-0.11	-0.1	0.02	0.06	0.04	-0.13	0.07	-0.36	1.00	
RoL Index	0.07	0.08	-0.06	-0.08	0	0	0.03	0	-0.05	-0.05	0.15	-0.23	1.00

* = Dummy Variable

Table 5: Correlation Matrix

The first thing that stands out is that the correlation between the dependent variable of the spread and the EL is relatively strong at 0.75 and has a positive direction. This underlines the accompanying hypothesis that the EL is the essential determinant of the spread and a higher

EL is associated with higher risk compensation for investors. The negative correlation between the volume, the term and the term to maturity (TTM), however, initially speaks against the assumption that a lower volume or a lower (remaining) term leads to a higher spread. However, this can change in the multivariate pricing model environment, as it has already been pointed out by Gürtler et. al (2014) accordingly. The same applies to the dummy variable of the trigger type, whereby it normally should be assumed that an indemnity trigger actually leads to higher spreads. The positive correlation of the spread with the number of peril types and regions underlines the previous assumption. The self-constructed dummy variable of seasonality also shows that a higher probability of occurrence of the covered risk type goes hand in hand with higher spreads. This confirms the seasonality thesis as well as the correct modeling of that. The negative influence of the rating corresponds to the rational assumption that lower-rated cat bonds generally lead to a spread as a risk premium measure for investors. With regard to the macroeconomic and financial market data, the observed results are expected, with a positive correlation of the S&P 500 return and the US corporate credit spread with the car bond spread. On this occasion, the low correlation values underline one of the main characteristics of cat bonds and generally of ILS, which describes a high degree of independence from the traditional financial markets and thus represents an alternative attraction for investors. Ultimately, the positive correlation of the ROL Index underscores the assumptions of the influence of reinsurance cycles on cat bond spreads, even if it is relatively low in comparison.

3.3 Testing for Regression-based Data Issues

After the final cat bond data set has been descriptively examined, the tests for regression-based data issues follow in this section, in preparation for the data-based OLS evaluation of the predefined pricing model. The focus here is set on the phenomena of multicollinearity and heteroscedasticity. Tests for omitted variables are left out, since the model and variable assumptions are already based on extensive research, i.e. on Braun (2014) and Gürtler et. al (2012). Lastly, outlying observations (outliers), in particular very high spreads, are not omitted from the data set as they represent natural variations in the cat bond market and do not result from any measurement or data pipeline errors. Furthermore, unnatural variations due to partial or full triggers haven already been removed during the cleaning process in chapter 3.1, whereby zero-coupon cat bond tranches also remain in the final data set.

The first test deals with multicollinearity. The main problem is that two or more of the predictors in a regression model are moderately or highly correlated, which in turn can lead to falsified standard errors and thus falsified significance levels of the explanatory variables (PennState University, 2021). This can mistakenly result in no statistical significance being assigned for individual determinants, or vice versa, which ultimately leads to incorrect regression results (PennState University, 2021). A central method for examining multicollinearity in a multivariate environment are the Variance Inflation Factors (VIF), which are calculated as follows (PennState University, 2021):

$$VIF_{\mathrm{i}} = \frac{1}{1 - R_i^2} \quad , \quad \mathrm{i} = i^{th}\text{-predictor of the multivariate model} \tag{5}$$

where R_i^2 defines the coefficient of determination obtained by the i^{th}-predictor (PennState University, 2021). As can already be seen from the notation, the predictor-specific VIF quantifies how much the variance is inflated. Table 6 summarizes the VIFs per predictor according to the evaluated pricing model and its determinants:

VIF Factor	Variable
1.55	EL
1.12	Volume
1.42	Term
1.30	TTM
1.29	Peril Types
1.45	Peril Regions
1.13	Seasonality*
1.52	Trigger Type*
1.42	Rating
1.16	S&P500 Return
1.25	Corp. Spread
1.09	RoL Index

* = Dummy Variable

Table 6: Multicollinearity - Variance Inflation Factor Results

By applying the general rule in research, that is $VIFs > 4$ warrant further investigation, while $VIFs > 10$ are signs of serious multicollinearity requiring model and variable corrections. A look at the results shows that all VIFs are within the tolerance for only slight multicollinearity and therefore the regression modeling does not require any further adjustments.

The second test is dedicated to heteroscedasticity. In the context of regression analysis, heteroscedasticity can be understood as non-constant error variance, which means that the remaining residual variability changes as a function of something that is not in the model, e.g. in time (Astivia and Zumbo, 2019, p.1). While existing heteroscedasticity does not affect the regression coefficients of the predictors themselves or the fit of the model, it causes a bias in the standard errors as well as in the test statistics of the coefficients and also leads to a degree of uncertainty in the form of influenced F-statistics (Astivia and Zumbo, 2019, p.2-3). The White Test and the Breusch-Pagan Test represent two central statistics in order to mathematically check for heteroscedasticity. Both statistics examine the relationship between the squared error terms and predictors using an additional regression (Astivia and Zumbo, 2019, p.5). Formally, they can be defined as follows: (Astivia and Zumbo, 2019, p.5-6):

$$\epsilon_i^2 = \alpha_0 + \alpha_1 X_{1i} + ... + \alpha_p X_{pi} + u_i \tag{6}$$

$$\epsilon_i^2 = \alpha_0 + \alpha_1 X_{1i} + ... + \alpha_p X_{pi} + \gamma_1 X_{1i}^2 + ... + \gamma_p X_{pi}^2 + \delta_1 (X_{1i} X_{2i} + ... + \delta_{2^p - 1}(X_{p-1i} X_{pi}) + v_i \tag{7}$$

where (6) refers to the Breusch-Pagan and (7) to the White Test, ϵ describes the error term and X the explanatory variables of the underlying multivariate regression model. As can already be seen from the formal notations, the White Test represents a special case of the Breusch-Pagan

Test by applying an additional term, which consists of a higher-order, non-linear functional form of the X-terms (Astivia and Zumbo, 2019, p.6). Ultimately, the null hypothesis for both the Breusch-Pagan and the White Test is defined as follows (Astivia and Zumbo, 2019, p.6):

$$H_0 : \alpha_1 = \alpha_2 = ... = \alpha_p = 0 \iff Homoscedasticity \tag{8}$$

Table 7 shows the test statistics and the corresponding confidence levels of the Breusch-Pagan and White Test for the cat bond data set in context if this thesis:

	White Test	Breusch-Pagan Test
LM Statistic	4673.3730	1202.4141
LM-Test p-value	0.0000	0.0000
F-Statistic	161.7636	121.0676
F-Test p-value	0.0000	0.0000

Table 7: Heteroscedasticity - White and Breusch-Pagan Test Results

As can be seen from the results of the two tests, for the LM as well as the F-Statistic of each test, the confidence intervals (p-values) are around zero. The rejection of the homoscedasticity null hypothesis takes place according to the definition when they fall below a predefined confidence interval, which is set to 0.05 in the context of this thesis. Thus, the homoscedasticity assumption can finally be rejected. As a consequence of this regression issue within the framework of the data-based evaluation of the pricing model, heteroscedasticity-consistent (robust) standard errors are used in order to prevent the accompanying consequences of heteroscedasticity.

3.4 Data-based Evaluation of the Pricing Model

After the underlying pricing model has been defined in chapter 2.3.3 on the basis of existing empirical research, this section is followed by a further evaluation based on the cat bond data set. On the one hand, the aim of this analysis is to check the statistical significance of the explanatory variables and, on the other hand, to check the overall quality of adaptation and the comparison with existing benchmarks that is subject to similarly defined pricing models. First, let's reconsider the underlying pricing model for the subsequent OLS regression:

$$S_{i,t}^{Cat} = \beta' * X_{i,t} + \gamma' * Y_t + \epsilon_{i,t} \tag{9}$$

where $S_{i,t}^{Cat}$ defines the spread of cat bond i at time t, $X_{i,t}$ refers to the cat bond-specific determinants at time t with the vector of weights β', Y_t to the macroeconomic and financial market determinants with the vector of weights γ' and lastly $\epsilon_{i,t}$ defines the random error term. Taking into account the determinants evaluated in chapter 2.3.1 and 2.3.2, the following detailed regression notation results:

$$
\begin{aligned}
S_{i,t}^{Cat} = \; & \alpha_{i,t} + \beta_1 EL_{i,t} + \beta_2 Vol_{i,t} + \beta_3 Term_i + \beta_4 TTM_{i,t} + \beta_5 PerilTypes_i \\
& + \beta_6 PerilRegions_i + \beta_7 Seasonality*_{i,t} + \beta_8 Trigger*_i + \beta_9 Rating_i \\
& + \gamma_1 S\&P500_t + \gamma_2 Corp.Spread_t + \gamma_3 RoL_t + \epsilon_{i,t}
\end{aligned} \tag{10}
$$

with α referring to the intercept, β_i to the cat bond-specific variable regression coefficients, γ_i to the macroeconomic and financial market variable regression coefficients and lastly ϵ_i to the random error term. Furthermore, variables with the addition * are constructed as dummy variables and are thus $\in \{0,1\}$, whereby detailed variable descriptions can be reconsidered at the end of chapter 3.1. In addition, the natural logarithm of the cat bond volume is used, since it is not the explicit monetary value but rather the magnitude which defines the object of interest within this analysis. Again, this is in line with the specification by Gürtler et. al (2012, p.14). Since the presence of heteroscedasticity was demonstrated in the previous section, the OLS regression is performed using heteroscedasticity and autocorrelation consistent (HAC) standard errors to ensure robust inference (Astivia and Zumbo, 2019, p.7-8). Thus the standard errors and p-values are based on the Newey-West HAC covariance matrix (Astivia and Zumbo, 2019, p.7-8). The first OLS regression based on the model specification in forumla (10) and the 660 cat bond tranches and their 6914 observations produces the following results in Table 8:

Dep. Variable:	Spread	**R-squared:**	0.640
Model:	OLS	**Adj. R-squared:**	0.640
Method:	Least Squares	**F-statistic:**	225.2
No. Observations:	6914	**Prob (F-statistic):**	0.00

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
Intercept	0.0580	0.005	10.928	0.000	0.048	0.068
EL	1.3612	0.068	20.113	0.000	1.229	1.494
log(Volume)	-0.0048	0.001	-6.762	0.000	-0.006	-0.003
Term	-0.0005	5.29e-05	-10.194	0.000	-0.001	-0.000
TTM	1.901e-05	2.59e-05	0.734	0.463	-3.18e-05	6.98e-05
Peril Types	0.0061	0.001	12.039	0.000	0.005	0.007
Peril Regions	-0.0028	0.001	-3.365	0.001	-0.004	-0.001
Seasonality*	0.0019	0.001	2.483	0.013	0.000	0.003
Trigger Type*	-0.0045	0.001	-3.479	0.001	-0.007	-0.002
Rating	0.0053	0.001	8.132	0.000	0.004	0.007
S&P500 Return	0.0276	0.004	6.770	0.000	0.020	0.036
Corp. Spread	0.2723	0.022	12.478	0.000	0.230	0.315
RoL Index	0.0131	0.005	2.912	0.004	0.004	0.022

* = Dummy Variable

Table 8: Multiple Linear Regression Results

The results show that all explanatory variables have a high statistical significance, except for the term to maturity (TTM), which with a p-value of 0.463 has a high probability of not having any influence on the spread[10]. Thus, on the one hand, it can be concluded that the underlying pricing model determination by Braun (2014) and Gürtler et. al (2012) offers a suitable framework for the data set of this thesis. On the other hand, due to the insignificance of the variable TTM, it is advisable to eliminate it from the model, which in turn can lead to changed regression results due to the changed multivariate specification. Consequently, a second OLS-based evaluation is

[10]The p-value ($P > |z|$) defines the probability that the specific regression coefficient is zero and therefore has no influence on the dependent variable. Within this thesis, the significance level is set to 0.05 which means that all p-values > 0.05 define insignificant variable coefficients.

carried out, based on the model specification defined along with it, but without the variable TTM, whose results are shown in Table 9:

Dep. Variable:	Spread	**R-squared:**	0.640
Model:	OLS	**Adj. R-squared:**	0.640
Method:	Least Squares	**F-statistic:**	245.1
No. Observations:	6914	**Prob (F-statistic):**	0.00

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
Intercept	0.0580	0.005	10.923	0.000	0.048	0.068
EL	1.3615	0.068	20.121	0.000	1.229	1.494
log(Volume)	-0.0048	0.001	-6.757	0.000	-0.006	-0.003
Term	-0.0005	5.03e-05	-10.498	0.000	-0.001	-0.000
Peril Types	0.0061	0.001	12.060	0.000	0.005	0.007
Peril Regions	-0.0028	0.001	-3.379	0.001	-0.004	-0.001
Seasonality*	0.0018	0.001	2.468	0.014	0.000	0.003
Trigger Type*	-0.0045	0.001	-3.479	0.001	-0.007	-0.002
Rating	0.0053	0.001	8.115	0.000	0.004	0.007
S&P500 Return	0.0275	0.004	6.745	0.000	0.020	0.036
Corp. Spread	0.2718	0.022	12.484	0.000	0.229	0.315
RoL Index	0.0132	0.005	2.911	0.004	0.004	0.022

* = Dummy Variable

Table 9: Adjusted Multiple Linear Regression Results

The OLS results of the adjusted regression model show that omitting the variable TTM leads to statistical significance across all explanatory variables. In addition, the adapted model shows the same goodness-of-fit in the form of the (adjusted) coefficient of determination R^2 of 0.64, which in turn supports the variable adjustments. The regression coefficients of some explanatory variables take on very small values, which, despite the regression of absolute values, which generally leads to low values, suggests very little influence. In this regard, it could be argued that, given a predetermined coefficient threshold, individual variables could also be removed due to their low influence. On the other hand, it can be argued that ML methods use those variables more efficiently due to their higher model complexity and are therefore still important. As a result, based on the second argument, these variables remain in the underlying pricing model despite their low coefficients in order to provide a comprehensive foundation for the ML optimization of spread forecasting. A similar approach is used by Götze et. al (2020, p.6) whose primary market results are also used as a comparison measure for the ML results of this thesis.

Finally, on the basis of the data-based evaluation of the determinants in the predefined pricing model, the following model can be defined as the basis for the following optimization problems:

$$S_{i,t}^{Cat} = \beta' X_{i,t} + \gamma' Y_t + \epsilon_{i,t} \tag{11}$$

$$\text{s.t.}$$

$$X_{i,t} \in \{EL_{i,t}, Vol_{i,t}, Term_i, PerilTypes_i, PerilRegions_i, Seasonality*_{i,t}, Trigger*_i, Rating_i\}$$

$$Y_t \in \{S\&P500_t, Corp.Spread_t, RoL_t\}$$

This pricing model defines the foundation of the main part of this thesis, which is object to the application of the benchmark regressions and the optimizing ML models in order to forecast the cat bond spread.

4 Enhanced Cat Bond Pricing: ML-based Modeling

After the underlying pricing model has been defined and evaluated based on the data set in the previous chapter, this section pursues the main analysis in form of the ML-based optimization of the spread forecasting. In the first step, the testable hypotheses and their testing criteria are defined, after which, in the second step, the central hyperparameters of the benchmark regression as well as the optimizing ML models are tuned. Afterwards, various types of regression-based models, in particular OLS and Penalized Regressions, are carried out to forecast the cat bond spreads as performance benchmarks. In addition, existing research on ML models for cat bond pricing is used as a further benchmark in order to additionally evaluate the results (i.e. Götze et al., 2020). Ultimately, complex random forest and neural network models are applied as optimizing spread forecasting approaches. All models are finally tested in a robustness check.

4.1 Rolling Samples & Comparison Measures

The application of the extended regression models as benchmarks as well as the ML models in the form of random forests and neural networks is based on rolling samples. These are defined as four consecutive years being used as the In-Sample (IS) and the following year as the Out-of-Sample (OOS). Looking back at the cat bond data set, this results in 13 in-samples and 13 out-of-samples. In addition, the models are validated using a robustness check, with the two consecutive years being used as the out-of-samples. As a result, there are 12 in-Samples and 12 out-of-samples for the robustness check. This sample structure is in line with Götze et. al (2020, p.21-22), which optimally allows a benchmark comparison of the comparative measures. The detailed list and the number of observations in the samples are stated in Table 10:

IS	No. Obs.	OOS	No. Obs.	OOS - Robustness	No. Obs.
2003-2007	894	2008	327	2008-2009	697
2004-2008	1202	2009	370	2009-2010	755
2005-2009	1516	2010	385	2010-2011	773
2006-2010	1785	2011	388	2011-2012	850
2007-2011	1899	2012	462	2012-2013	829
2008-2012	1932	2013	367	2013-2014	775
2009-2013	1972	2014	408	2014-2015	878
2010-2014	2010	2015	470	2015-2016	924
2011-2015	2095	2016	454	2016-2017	983
2012-2016	2161	2017	529	2017-2018	1122
2013-2017	2228	2018	593	2018-2019	1210
2014-2018	2454	2019	617	2019-2020	1267
2015-2019	2663	2020	650	N/A	N/A

Table 10: Summary Statistics - Rolling Samples

As can be seen from Table 10, a mostly strict growth can be observed in both the in-samples and the out-of-samples. This underlines the thesis in chapter 2.1 that cat bonds and ILS instruments in general are becoming increasingly attractive for investors and that the market has been subject to almost continuous growth since the beginning of the 2000s. In addition, the ratio of out-of-sample and in-sample observations is approximately 25% across all IS-OOS combinations, which represents a solid basis with regard to training the ML models as well as hyperparameter tuning as such ratios are often used in practice as well.

The next step is the definition of the performance and comparison procedure for the regression-based and the ML models. For this purpose, the models are set up according to the underlying determinants of the pricing model from formula (11) and optimized with regard to their hyper-parameters. The models are then used to estimate the quarterly spreads of the out-of-sample periods based on the pricing determinants from the in-samples. This already shows that the difference between the observed value in the OOS and the forecasted value in the OOS represents an important performance measure of a model. In order to determine this quantitatively and also to make it robust with regard to outliers, the key measure *Root Mean Square Error* ($RMSE$), which is often used in practice, is suitable and formally defined as follows:

$$RMSE = \sqrt{\frac{1}{N} * \sum_{i=1}^{N} (S_{i,t}^{Cat} - \hat{S}_{i,t}^{Cat})^2} \tag{12}$$

where N refers to the number of observations, $S_{i,t}^{Cat}$ to the observed spread in the data set and $\hat{S}_{i,t}^{Cat}$ to the forecasted spread by the respective model. In addition, a goodness-of-fit measure is recommended within the framework of models in which a certain random error term can be assumed and the degree of influence of the variables are also not always clear. One of the central goodness-of-fit measures is the coefficient of determination R^2. It quantitatively measures the percentage of variation in the dependent variable explained by variation in the explanatory variables (Figueiredo et al., 2011, p. 60). Thus, the R^2 can take a value between zero and one, with a value of one speaking of a perfect explanation of the spread by the explanatory pricing determinants by the corresponding model. Formally, the R^2 is defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (S_{i,t}^{Cat} - \hat{S}_{i,t}^{Cat})^2}{\sum_{i=1}^{N} (S_{i,t}^{Cat} - \bar{S}_{i,t}^{Cat})^2} \tag{13}$$

where the same variables compared to the RMSE are defined in the same way and additionally $\bar{S}_{i,t}^{Cat}$ refers to the mean observed spread in the data. Taking into account the rolling samples, it is therefore necessary to calculate the two key figures of the $RMSE$ and the R^2 on the basis of the corresponding regression and ML models for the OOS forecasts in order to determine the comparative performance and to test the hypotheses defined in the next chapter.

4.2 Optimization Problem and Hypotheses

ML methods such as the Random Forest and the Neural Network are suitable for optimizing forecasting performance compared to regression models. In context of this paper, the term optimization is defined according to the previous chapter in terms of a lower forecasting error $RMSE$ and a higher coefficient of determination R^2. In addition, however, it is necessary to define testable hypotheses, which are rejected or confirmed in the application of the benchmark regression-based as well as the ML models. The focus on the results in terms of $RMSE$ and

R^2 and their comparison across models serves as the foundation for answering these testable hypotheses, and hence also in regard to answer the main research question of this thesis.

4.2.1 Hypothesis 1: Penalized Regression Performance

In addition to the basic OLS regression, alternative regression models are repeatedly presented as optimization opportunities in order to improve the forecasting performance of linear models. *Penalized Regression* is one of these alternative regression applications. This thesis focuses on Ridge, Lasso and Elastic Net, with a detailed introduction to the models in the corresponding chapters of the model application. Penalized regressions have an advantage in that they constrain/shrink the parameter estimates (Miller, 2013, p.5). Here, the variance of the estimator is reduced, but at the same time the bias is increased, whereby the variance-bias-tradeoff[11] is considered to be a central challenge of penalized regressions, whereby the same applies in ML model specifications (Miller, 2013, p.5). While large regression coefficients are heuristically viewed as complex models, penalized regression methods make it possible to restrict such complexity through the shrinkage of the coefficient vector in an multivariate model environment and thus possibly achieve better forecasting results (Taylor, n.d., p.9). Taking into account the results of Götze et. al (2020, p.22), optimized performance could not be achieved by penalized regression methods as part of the primary market analysis. This can not only be due to the relatively small data set but also to the definition of the optimization interval of the shrinkage parameter (hyperparameter), which was limited relatively strongly in the vicinity of zero. Thus, these advanced regression models could still be suitable, taking into account a broader data set and a more extensive shrinkage parameter search interval. The following hypothesis can thus be set up in order to optimize the regression-based benchmark models:

H_1 : *Penalized Regressions, i.e. Lasso, Ridge and Elastic Net, perform better in spread forecasting than OLS regressions, in terms of a higher mean coefficient of determination* R^2 *and a lower mean forecast error RMSE over all rolling samples.*

4.2.2 Hypothesis 2: NN Performance

Neural Networks represent a central method in the area of ML, which became particularly popular in the area of forecasting and prediction problem settings. In the area of financial market issues, it is evident that, provided that the model specification is correct, neural networks can lead to better results compared to regression-based models (Ahangar et al., 2010, p.45). Even if the results of the primary market alanysis by Götze et. al (2020, p.22) show that forecasting performance is below that of regression-based models, under the circumstances of a larger data set and a more detailed model specification, it can be assumed that neural networks can achieve

[11] In machine learning it's often observed that techniques focusing on variance reduction in the estimator cause an increase in the estimation bias, where the bias is generally defined as the expected value of the difference between the estimator and the observed value. The balance between these thresholds define the heart of a successful ML model development. (Avati, 2020, p.2)

a higher performance. Thus, the following two hypotheses are defined with regard to the spread forecasting performance of neural networks:

$H_2(a)$: *Neural Networks perform better in cat bond spread forecasting than OLS regressions, in terms of a higher mean coefficient of determination* R^2 *and a lower mean forecasting error RMSE over all rolling samples.*

$H_2(b)$: *Neural Networks perform better in cat bond spread forecasting than penalized regressions, i.e. Lasso, Ridge and Elastic Net, in terms of a higher mean coefficient of determination* R^2 *and a lower mean forecasting error RMSE over all rolling samples.*

4.2.3 Hypothesis 3: RF Performance

Random Forest models represent a central application method in the area of ML and are generally recognized as a well-founded alternative to regressions (Ceh et al., 2018, p.1). While the considered regression models are limited by the linearity assumption, random forests allow a higher model complexity, in particular with regard to non-linearities (Schonlau and Zou, 2020, p.3-4). The detailed analysis by Mullainathan and Spiess (2017, p.89) shows that random forest models in particular have advantages in terms of performance compared to regression-based models under the premise of the correct model specification, even under moderate sample sizes and a limited number of covariates. As a foundation in the area of the cat bond market, the results of Götze et. al (2020) and Makariou et. al (2020) show a high performance of random forest models in forecasting primary market spreads of cat bonds. This allows the conclusion that such results can also be expected in the secondary market. Furthermore, the paper of Götze et. al (2020, p.22) shows that a significantly higher performance of the random forest model is to be expected not only compared to the regression-based models, but also with regard to the neural network. Thus, the following three hypothesis can be defined regarding the optimization of spread forecasting with a random forest model:

$H_3(a)$: *The Random Forest performs better in cat bond spread forecasting than OLS regressions, in terms of a higher mean coefficient of determination* R^2 *and a lower mean forecasting error RMSE over all rolling samples.*

$H_3(b)$: *The Random Forest performs better in cat bond spread forecasting than penalized regressions, i.e. Lasso Ridge and Elastic Net, in terms of a higher mean coefficient of determination* R^2 *and a lower mean forecasting error RMSE over all rolling samples.*

$H_3(c)$: *The Random Forest performs better in cat bond spread forecasting than neural networks, in terms of a higher mean coefficient of determination* R^2 *and a lower mean forecasting error RMSE over all rolling samples.*

4.2.4 Hypothesis 4: Robustness Check Performance

In the case of an existing degree of uncertainty with regard to the model specification, especially with regard to the determinants as well as the hyperparameters, the concept of robustness checks offers further tests of the observed forecasting results (Young, 2015, p.2). The concept of increased out-of-samples used in this thesis, which is explained in more detail in the corresponding chapter 4.6, aims to confirm the results achieved with less out-of-sample observations. In conclusion, the last two hypotheses can be defined as follows:

$H_4(a):$ *The observed performances across the OLS regression, Penalized Regressions, the Neural Network and the Random Forest, in terms of the mean coefficient of determination R^2 and the mean forecasting error RMSE, can be confirmed over increased out-of-samples.*

$H_4(b):$ *The observed performance rating between the OLS regression, Penalized Regressions, the Neural Network and the Random Forest, in terms of the mean coefficient of determination R^2 and the mean forecasting error RMSE, can be confirmed over increased out-of-samples.*

4.3 Hyperparameter Tuning Results

In the context of penalized regression models and in particular ML models, hyperparameters define parameters that play a central role in the model specification due to the special algorithmic structure. They act as a control element in the training process in order to subsequently achieve parameter-optimized results in the test process. One of the main tasks is to avoid overfitting, which would lead to very good in-sample but bad out-of-sample performances (Mullainathan and Spiess, 2017, p.88). However, there is the main problem that a good out-of-sample performance is to be achieved, but only the in-samples are fitted (Mullainathan and Spiess, 2017, p.92). Thus, inspired by Feurer and Hutter (2019, p.5), the formal notation for any models that are based on hyperparameters in the framework of in-sample and out-of-sample settings of this thesis can be defined as follows:

$$\lambda_{valid} = \underset{\lambda \in \Lambda}{\operatorname{argmin}} \; \mathbb{E}_{(D_{IS}, D_{OOS}) \sim \mathcal{D}}[f(\mathcal{L}, \mathcal{A}_\lambda, D_{IS}, D_{OOS})] \tag{14}$$

where $f(\mathcal{L}, \mathcal{A}_\lambda, D_{IS}, D_{OOLS})$ is the loss function, in terms of the out-of-sample forecasting errors, of the model algorithm $\mathcal{A}$ under the hyperparameter space Λ based on the fitting of the in-sample data (D_{IS}) within the data set $\mathcal{D}$. Ultimately, λ_{valid} defines the respective validated hyperparameter which minimizes the forecasting error of the respective model.

Three central concepts of the evaluation of the model hyperparameters, based on formula (14), are *Grid Search*, *Random Search* and *Manual Search*. While the concept of manual search compares the OOS results of the models based on a predetermined search interval for each hyperparameter manually with multiple runs, the concepts of grid search and random search pursue more systematic approaches. On the one hand, grid search follows a search algorithm that checks all combinations of a predefined hyperparameter set with regard to an evaluation criterion (Liashchynskyi and Liashchynskyi, 2019, p.3). In the context of forecasting problems, a variant of the absolute error is often used. Since continuous search intervals or intervals with many search parameters impair the calculation time and thus also the performance of the model, specific discrete boundaries are often set in order to limit the trade-off with regard to the calculation time (Liashchynskyi and Liashchynskyi, 2019, p.3). On the other hand, there is the concept of random search, which overwrites the predetermined set of hyperparameters with a random selection of those and evaluates the resulting hyperparameters in all combinations with regard to the forecasting performance (Liashchynskyi and Liashchynskyi, 2019, p.3). The latter happens analogously to the fitting procedure of the grid search. In contrast to the grid search, there is the possibility that random elements of a continuous space can also be selected (Liashchynskyi and Liashchynskyi, 2019, p.3).

In this thesis, the concept of grid search is used to optimize the hyperparameters of penalized regressions as well as of the random forest and the neural network across the in-samples. The evaluation criterion for the selection of hyperparameters in the grid search algorithm is represented by the $RMSE$ (see chapter 4.1), as this defines one of the overall criteria for optimization. Furthermore, a high correlation with the coefficient of determination R^2 is assumed. If the criterion of computational time had been a central condition for optimizing the spread forecasting, the random search concept could have been used, but this is not the case in this analysis. Ultimately, Table 11 shows the hyperparameters to be optimized for each model, the search intervals[12] and the corresponding validated hyperparameter values:

Model	Hyperparameter	Search Interval	Validated Value
Lasso	Shrinkage parameter	$[e^{-8}, ..., e^{0}]$, n = 200	e^{-8}
Ridge	Shrinkage parameter	$[e^{-8},...,e^{0}]$, n = 200	0.00511
Elastic Net	Shrinkage parameter	$[e^{-8},...,e^{0}]$, n = 200	$9.3293 * e^{-7}$
	Mixing parameter	[0, ..., 1] , n = 200	0
RF	Number of trees	[200, ..., 2000] , n = 10	200
	Max depth of the trees	[10, ..., 110] , n = 11	30
	Number of min samples to split internal node	[2, 5, 10]	5
	Number of min samples at leaf Node	[1, 2, 4]	2
	Number of Features during choice for best split	[Auto, Sqrt]	Sqrt
	Bootstrap samples used when building trees	[True, False]	True
NN	Number of hidden layers	[1, 2, 3]	3
	Number of neurons (each hidden layer)	[1, 11], [1, 11], [1, 11]	4, 7, 5
	Batch size	[10, ..., 300] , n = 7	100
	Number of epochs	[50, ..., 700] , n = 7	100
	Activation function	[Linear, Relu, Tanh, Sigmoid]	Tanh
	Optimizer function	[SGD, RMSprop Adam]	Adam

Table 11: Hyperparameter Tuning

As it can be seen from Table 11, the hyperparameter tuning of the penalized regressions (Lasso, Ridge, Elastic Net) focuses on the shrinkage parameter λ. According to the definition of the models, a hyperparameter value of 0 means that the models correspond to an OLS specification. The exponential definition from the lower limit ($e^{-8} \approx 0.00034$) up to the maximum value of 1 ($= e^{0}$), and thus a heavy tailed distribution towards zero (positive skewness), is in line with

[12]The variable n in the search interval column of Table 11 defines the number of equally distributed steps between the interval limits according to the discrete interval.

existing penalized regression results, which often validate shrinkage parameters close to zero in the context of cat bonds (see, e.g., Götze et al., 2020, p.20). In addition, the mixing parameter for the Elastic Net regression defines the weights of the linear combination of the Lasso and Ridge shrinkage terms included. Hence, a mixing parameter of 0 omits the Lasso term and a parameter of 1 omits the Ridge term analogously.

With regard to the random forest model specification, the tuning of five central hyperparameters is carried out. These are defined by the total number of decision trees in the entire model, the maximum depth for each decision tree, the minimum number of samples for a split of an internal node, the number of samples that must be at least contained in a leaf node, the maximum number of variables considered during the two-fold choice for the best split (defined as a function of the total variables) and finally the condition as to whether bootstrap samples are permitted in the decision trees or whether the entire input sample is used for all trees. The determination of the search intervals, i.e. that of the cardinal parameters, is in accordance with existing research work with similar data and problem setting characteristics.

The central hyperparameters of the neural network are the number of hidden layers[13], which, in addition to the input and output layers represent the fundamental structure of the model, the number of neurons per hidden layer, the batch size, which determines how many sub-samples per training session, the number of epochs, which in turn determines the number of training sessions including the entire training set, as well as the type of activation and optimization function. The search interval limits of the neurons in the hidden layers were defined according to the number of input variables (11) and the number of output variables (1). Furthermore, the search interval of the amount of hidden layers is in line with that of Götze et. al (2020, p.20), which allows a comparison of the model performance on the basis of the other hyperparameter determinations. Ultimately, the interval limits of the other cardinal hyperparameters are in line with existing research with similar data and problem-setting characteristics.

The more precise roles of the validated hyperparameters within the structure of the penalized regressions as well as the random forest model and the neural network are explained in chapter 4.4.1 and chapters 4.5.1 and 4.5.2 respectively, in context of the model applications.

4.4 Regression Model Application

Linear regression models are a fundamentally powerful tool for describing the relationships between a dependent and one or more explanatory variables and ultimately predicting new values on the basis of these. One of the central methods is the Ordinary Least Squares (OLS), which was already used as part of the data-based evaluation of the pricing model. In addition, the concept of Penalized Regression is a valid alternative to the OLS method, which adds a shrinkage term within the regression formula. Thus, penalized regression can produce optimized forecasting results, especially with regard to high variances in the regression coefficients due to existing multicollinearity (Gürtler et al., 2012, p.7-8). In the context of ILS instruments and especially cat bonds, existing econometric approaches show that regression-based spread

[13] The number of hidden layers is tuned by a manual search approach.

forecasting can deliver good results (see, e.g., Braun, 2014; Gürtler et al., 2012; Götze et al., 2020). The OLS and penalized regression models, i.e. Lasso, Ridge and Elastic Net, are thus defined as performance benchmark models to the later ML models and are applied to the cat bond data set in this chapter.

4.4.1 OLS, Ridge, Lasso & Elastic Net Regression

Ordinary Least Squares (OLS)[14] is a central method within the framework of multivariate linear regression models to estimate the corresponding regression coefficients of the explanatory variables. The least squares estimators are optimized in such a way that they minimize the sum of the squared residuals (Adedia et al., 2016, p.2). In the context of this thesis, the OLS estimate of the cat bond spread represents the fundamental forecasting approach on which the penalized regression methods are ultimately based. The OLS regression coefficients can thus be defined as follows (Ogutu et al., 2012, p.3):

$$\begin{aligned} \hat{\beta}_{\text{OLS}} = \; & \underset{\beta}{\operatorname{argmin}} ||y - X\beta||_2^2 \\ & \text{s.t.} \quad ||y - X\beta||_2^2 = \sum_{i=1}^{n} (y_i - x_i^T \beta)^2 \end{aligned} \tag{15}$$

where β represents the regression coefficients of the vector X of explanatory variables and y the dependent variable, which is to be estimated. As stated above, the best estimators $\hat{\beta}$ thus minimize the squared residuals between the observed value and the estimated dependent variable from the explanatory vector. As can already be seen from the definition of the OLS regression, possible issues lie in the high variances of individual regression coefficients, which flow one-to-one into the variable estimation and thus can influence the accuracy of the estimation (Götze et al., 2020, p.7-8). Such problems often result from complex models with large vectors of explanatory variables (Götze et al., 2020, p.7-8).

Penalized Regressions define an alternative approach to address precisely these limitations of the OLS approach. Penalized regression models, such as Lasso, Ridge and Elastic Net, include a corresponding constraint of the model parameters in order to counteract existing multicollinearity within the model and the resulting high variances in the regression coefficients (Götze et al., 2020, p.8; Ginestet, n.d., p.2). One of these methods is *Lasso* (Least Absolute Shrinkage and Selection Operator), which is dedicated to both shrinkage adjustment and variable selection (Kyung et al., 2010, p.371). Analogous to OLS, it pursues the goal of minimizing the squared sum of the residuals, but with the addition of the non-differentiable constraint expressed in terms of the L^1-norm of the coefficients (Kyung et al., 2010, p.371). Thus, the Lasso regression coefficient can be formally defined as follows (Ogutu et al., 2012, p.3):

[14] The concept of the general multivariate regression as well as the OLS is taken as given knowledge of the reader in the context of this thesis. Therefore, a brief introduction to the topic follows, after which the concept of penalized regression is examined in more detail.

$$\begin{aligned}\hat{\beta}_{\text{Lasso}} = & \underset{\beta}{\operatorname{argmin}} ||y - X\beta||_2^2 + \lambda||\beta||_1 \\ & \text{s.t.} \quad ||y - X\beta||_2^2 = \sum_{i=1}^{n}(y_i - x_i^T\beta)^2 \quad , \quad ||\beta||_1 = \sum_{j=1}^{p}|\beta_j|\end{aligned} \tag{16}$$

where the first term is interpreted analogously to the OLS definition and the additional shrinkage term as the sum of the absolute regression coefficients β_j weighted with the shrinkage parameter λ. Hence, the term $||\beta||_1$ describes the L^1-norm penalty on β (Ogutu et al., 2012, p.3). As part of the tuning of the parameter λ, the degree of shrinkage can be explicitly controlled. Ultimately, the formal notation shows that a shrinkage parameter λ of zero leads to a classic OLS model specification.

In addition to the Lasso approach, *Ridge* regression defines another alternative to OLS in the context of penalized regressions. Analogously, Ridge also builds on the concept of the OLS regression and adds a shrinkage term. Here, however, the shrinkage term is based on the constraint expressed in terms of the L^2-norm of the coefficients, which in turn is controlled by the parameter λ. Consequently, the regression coefficient can be formally defined in the context of the Ridge approach as follows (Ogutu et al., 2012, p.3):

$$\begin{aligned}\hat{\beta}_{\text{Ridge}} = & \underset{\beta}{\operatorname{argmin}} ||y - X\beta||_2^2 + \lambda||\beta||_2^2 \\ & \text{s.t.} \quad ||y - X\beta||_2^2 = \sum_{i=1}^{n}(y_i - x_i^T\beta)^2 \quad , \quad ||\beta||_2^2 = \sum_{j=1}^{p}\beta_j^2\end{aligned} \tag{17}$$

with the first term corresponding to the fundamental OLS estimator and the additional shrinkage term as the sum of the squared coefficients of each explanatory variable, whose influence is in turn determined by the shrinkage parameter λ. Similar to the Lasso concept, the term $||\beta||_2^2$ hence describes the L^2-norm penalty on β (Ogutu et al., 2012, p.3).

The last concept in the context of penalized regressions is the *Elastic Net* regression. This approach serves as an extension of the Lasso approach in order to be even more robust to very high multicollinearity (Ogutu et al., 2012, p.3). This is done using the linear combination of the shrinkage terms of the Lasso and Ridge approach. While the L^1 part is intended for the variable selection, the L^2 part motivates grouped selection of highly correlated and stabilizes the solutions with respect to random sampling (Ogutu et al., 2012, p.3). Hence, the elastic net estimator can be formally defined as follows (Ogutu et al., 2012, p.3):

$$\begin{aligned}\hat{\beta}_{\text{ElasticNet}} = & (1 + \frac{\lambda_2}{n}) * \{\underset{\beta}{\operatorname{argmin}} ||y - X\beta||_2^2 + \lambda_1 * ||\beta||_1 + \lambda_2 * ||\beta||_2^2\} \\ & \text{s.t.} \quad ||y - X\beta||_2^2 = \sum_{i=1}^{n}(y_i - x_i^T\beta)^2 \quad , \quad ||\beta||_1 = \sum_{j=1}^{p}|\beta_j| \ , \ ||\beta||_2^2 = \sum_{j=1}^{p}\beta_j^2\end{aligned} \tag{18}$$

where the first term corresponds to the OLS estimator and $||\beta||_1$ as well as $||\beta||_2^2$ describe the L^1-norm and the L^2-norm penalty on β. Regarding the determination of the two shrinkage parameters, one parameter (λ_1) and the ratio between λ_1 and λ_2 are optimized in practice, resulting in the corresponding optimized model specification reagrding the hyperparameters.

With a view to the OLS regression and the penalized regression models, the cat bond spread represents the dependent variable y and the data-based evaluated pricing determinants in chapter 3.4 represent the explanatory variables within the vector X. In addition, it can be seen from chapter 4.3 that the tuned shrinking parameter of the Lasso model with e^{-8} converges to zero and thus corresponds approximately to an OLS specification. Furthermore, the shrinking parameters of the Ridge and the Elastic Net regression are set to 0.00511 and $9.3293 * e^{-7}$ respectively, the latter as part of a ratio of the L^1 term of zero, which leads to the depreciation of the Elastic Net regression to a Ridge specification. Thus, looking at the parameters, it is already evident that the results of the OLS and penalized regression should be very close to each other and that no optimization can be achieved by the penalized regression. Nevertheless, Table 12 shows the results of the cat bond spread forecasting in the form of the performance evaluation criteria $RMSE$ and R^2 for all rolling samples as well as the aggregated statistics:

IS	OOS	Measure	OLS	Lasso	Ridge	Elastic Net
2003-2007	2008	RMSE	0.0240	0.0240	0.0240	0.0240
		R2	0.8306	0.8306	0.8293	0.8304
2004-2008	2009	RMSE	0.0336	0.0336	0.0334	0.0335
		R2	0.5861	0.5861	0.5879	0.5865
2005-2009	2010	RMSE	0.0290	0.0290	0.0289	0.0290
		R2	0.6360	0.6360	0.6366	0.6362
2006-2010	2011	RMSE	0.0253	0.0253	0.0253	0.0253
		R2	0.6731	0.6731	0.6738	0.6733
2007-2011	2012	RMSE	0.0214	0.0214	0.0214	0.0214
		R2	0.7242	0.7242	0.7247	0.7244
2008-2012	2013	RMSE	0.0216	0.0216	0.0216	0.0216
		R2	0.7374	0.7374	0.7373	0.7374
2009-2013	2014	RMSE	0.0245	0.0245	0.0245	0.0245
		R2	0.6699	0.6699	0.6682	0.6693
2010-2014	2015	RMSE	0.0219	0.0219	0.0219	0.0219
		R2	0.7274	0.7274	0.7253	0.7266
2011-2015	2016	RMSE	0.0356	0.0356	0.0354	0.0355
		R2	0.3910	0.3910	0.3912	0.3911
2012-2016	2017	RMSE	0.0279	0.0279	0.0279	0.0279
		R2	0.6989	0.6989	0.6974	0.6983
2013-2017	2018	RMSE	0.0244	0.0244	0.0243	0.0244
		R2	0.7562	0.7562	0.7572	0.7566
2014-2018	2019	RMSE	0.0230	0.0230	0.0229	0.0230
		R2	0.6280	0.6280	0.6286	0.6283
2015-2019	2020	RMSE	0.0269	0.0269	0.0269	0.0269
		R2	0.5399	0.5399	0.5401	0.5400
Mean		RMSE	0.0261	0.0261	0.0260	0.0261
		R2	0.6614	0.6614	0.6614	0.6614
Median		RMSE	0.0249	0.0249	0.0249	0.0249
		R2	0.6715	0.6715	0.6710	0.6713
Std. Dev.		RMSE	0.0040	0.0040	0.0039	0.0040
		R2	0.0996	0.0996	0.0992	0.0995

Table 12: OLS, Lasso, Ridge & Elastic Net Regression Results

Due to the validated hyperparameters of both Lasso, Ridge and Elastic Net, which leads to approximate OLS specifications of the penalized regressions, an almost identical performance can be determined. The results of the out-of-sample forecasts show across all rolling samples that a $RMSE_{mean}$ of approximately 0.0261 and a coefficient of determination R^2_{mean} of approximately 0.6614 are achieved. Furthermore, the $RMSE$ and the R^2 within the rolling sample of the OLS forecast are between 0.0290 and 0.0356 and between 0.8306 and 0.3910, respectively. With regard to the penalized regression models Lasso/Ridge/Elastic Net, the $RMSE$ and the R^2 within the rolling samples are between 0.0240/0.0240/0.0240 and 0.0356/0.0354/0.0355, respectively between 0.8306/0.8293/0.8304 and 0.3910/0.3912/0.3911. Thus, on the one hand, it can be said that the addition of a shrinkage term does not lead to statistically significantly better forecasting results, whereby the hypothesis H_1 can ultimately be rejected. On the other hand, based on the overall results in the aggregated statistics, it can be concluded that there is high potential for optimization in the modeling of spread forecasts.

4.4.2 Limitations of Regression Models

Since the results of the spread forecasting in the previous chapter show that OLS and penalized regression models still leave optimization potential with regard to the goodness-of-fit as well as the forecasting errors, the question of the limitations of regression models arises. According to the definition of multivariate regression models, which include OLS, Lasso, Ridge and Elastic Net, it already emerges that there is a strict assumption of linearity between the explanatory variables. However, it can be assumed that, depending on the problem setting, the cat bond spread cannot be modeled as a linear combination of the explanatory variables, but instead requires a higher model complexity (Kuhn and Johnson, 2013, p.109). In such cases, the linear regression cannot capture the relevant characteristics (Kuhn and Johnson, 2013, p.109). Another limitation of linear regression models is the recording of values that deviate from the existing trend, so-called outliers. With recourse to the goal of minimizing the squared residuals, it becomes clear that such outliers have exponentially larger residuals and have a corresponding influence on the regression coefficients, which can thus be biased due to outlying observations (Kuhn and Johnson, 2013, p.109). Alternative modeling approaches against these limitations are offered by ML models such as random forests or neural networks, which have shown great optimization results in past research compared to multivariate regressions and are applied analogously to the problem of this thesis in the next chapter.

4.5 ML Model Application

As can already be seen in the previous chapter, regression-based models have several limitations. In the context of this, ML models are available to break the linearity assumption and therefore allow a higher model complexity. The preceding regression results as well as existing research approaches in the area of cat bond spreads (see, e.g., Braun, 2014, p.27; Gürtler et al., 2012, p.22-27) show that using linear models, cat bond spreads can be explained approximately well and the most important determinants can be worked out, but there exists high optimization potential with regard to determination and forecasting errors. In this section, spread forecasting is carried out using a random forest model and a neural network, with the hyperparameters of the models having already been evaluated.

4.5.1 Random Forest

Random Forest models have grown to a popular approach to various problem settings in ML in recent years. In practice, the application focuses in particular on classification and regression problems, in which random forests are extensively used in the course of optimization approaches (Cutler et al., 2011, p.1). In the context of this model concept, the main characteristics are in particular that it doesn't depend on a linearity assumption, time-efficient training processes are possible and regression-based forecasting problem settings can be processed in a straight forward manner. Hence, these characteristics of random forests state a fitting foundation to the target of optimizing the spread forecasting (Cutler et al., 2011, p.1).

In general, random forest models are based on Breiman's "Bagging" theory, which generally follows the method of generating multiple versions of a predictor and using these to get an aggregated predictor (Breiman, 1996, p.123). The multiple versions are formed by making bootstrap replicates of the learning set and using these as new learning sets, whereby the method already produced significantly higher accuracy scores than linear regressions (Breiman, 1996, p.123). Random forest models go one step further and are defined by the structure that they represent a tree-based ensemble, with each tree depending on a set of random variables (Cutler et al., 2011, p.2). The input variables are formalized as the vector $(X_1, ..., X_p)^T$, and the response/output variable as Y, which have an unknown joint distribution $P_{XY}(X, Y)$ (Cutler et al., 2011, p.2). The general goal is analogously to determine the function $f((X_1, ..., X_p)^T)$, which forecasts Y. Hence, the expected value of the loss function $E_{XY}[L(Y, f(X))]$ should be minimized in order to maximize the accuracy of the forecasting (based on Cutler et al., 2011, p.2). The structure of the relevant model specification includes the construct of a decision tree. This has the main components of a root node, several interior nodes and several leaf nodes (Götze et al., 2020, p.10). The process of the decision logic starts in the root node with the input of the vector of input variables, whereupon the data is divided into several sub-samples in the interior node, while the division is based on specific yes-no questions (Götze et al., 2020, p.10). A common criterion for the split is the mean squared residual, given by $Q = \frac{1}{n} + \sum_{i=1}^{n}(y_i - \bar{y})^2$ with $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ as the average forecasting value up to this node (based on Cutler et al., 2011, p.3-4). Consequently, two candidate descendants Q_L and Q_R and their the sub-sample sizes n_L and n_R are defined, whereby the effective split is to minimize $Q_{split} = n_L Q_L + n_R Q_R$ (Cutler et al., 2011, p. 3-4). After all, the construction of the trees and their nodes is stopped when a predetermined minimum threshold of observations in a leaf node is reached (Götze et al., 2020, p.10). As Götze et. al (2020, p.10) show in their research, the choice of variables within the decision tree is an important factor regarding the output performance. The impurity decrease through the split of the data into subsets in a node defines an indicator for the variable importance, whereby impurity decrease can be measured by the variance of the dependent variable in regression-related trees (Götze et al., 2020, p.10). Finally, the basic structure of a random forest decision tree can be summarized as follows (assuming a two-tree-model):

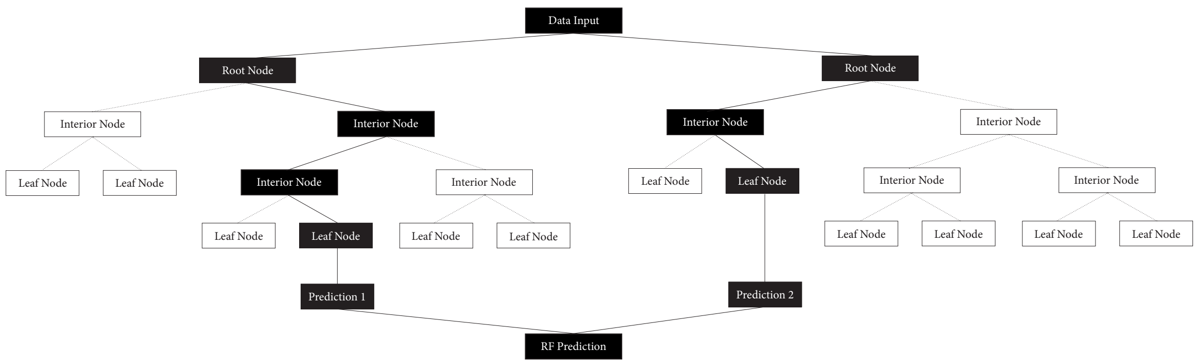

Figure 3: General Random Forest Structure - Own illustration

In the context of this thesis, the cat bond spread represents the output variable, which is forecasted by the random forest model prediction using the past four years' explaining variables to forecast the next year's spreads. With regard to the importance of the choice of input variables, as stated above, the model builds on the data-based evaluation of the pricing model, whereby eleven input variables, both cat bond-specific and macroeconomic, are defined. Looking back at chapter 4.3 of hyperparameter tuning, the number of trees in the entire decision tree is set to 200 ($N = 200$). In addition, the determination of the maximum depth of each tree shows that it is limited to 30 nodes. Thus, according to Figure 3, there are 200 individual trees with a maximum of 30 nodes each and thus a random forest total prediction in the form of the average of 200 individual predictions. Furthermore, the minimum number of samples must be determined for the split of an interior node on the one hand and to create a leaf node on the other hand. These are fixed at 30 and 5 respectively. In addition, it is important to determine the maximum number of variables to consider when looking for the best split, which are specified as a function of the total variables. The hyperparameter tuning results in the square root function ($sqrt$), from which it follows that $variables_{max}^{split} = \sqrt{variables_{total}}$. Finally, it is further determined whether bootstrap samples are used for each tree or the entire input data set. It can be seen that in combination with the other hyperparameters, the usage of bootstrap samples gives the best performance with regard to the evaluation criterion $RMSE$.

Now, that the hyperparameters have been explained in the context of the general random forest structure, the application of the model to the cat bond data set follows in order to forecast the cat bond spread. The last column in Table 13 shows the results for all rolling samples in relation to the performance criteria $RMSE$ and R^2. Furthermore, the mean forecasting results are given across all rolling samples in terms of $RMSE$ and R^2.

IS	OOS	Measure	OLS	Lasso	Ridge	Elastic Net	RF
2003-2007	2008	RMSE	0.0240	0.0240	0.0240	0.0240	0.0067
		R2	0.8306	0.8306	0.8293	0.8304	0.9869
2004-2008	2009	RMSE	0.0336	0.0336	0.0334	0.0335	0.0086
		R2	0.5861	0.5861	0.5879	0.5865	0.9735
2005-2009	2010	RMSE	0.0290	0.0290	0.0289	0.0290	0.0070
		R2	0.6360	0.6360	0.6366	0.6362	0.9796
2006-2010	2011	RMSE	0.0253	0.0253	0.0253	0.0253	0.0058
		R2	0.6731	0.6731	0.6738	0.6733	0.9822
2007-2011	2012	RMSE	0.0214	0.0214	0.0214	0.0214	0.0054
		R2	0.7242	0.7242	0.7247	0.7244	0.9845
2008-2012	2013	RMSE	0.0216	0.0216	0.0216	0.0216	0.0066
		R2	0.7374	0.7374	0.7373	0.7374	0.9818
2009-2013	2014	RMSE	0.0245	0.0245	0.0245	0.0245	0.0046
		R2	0.6699	0.6699	0.6682	0.6693	0.9890
2010-2014	2015	RMSE	0.0219	0.0219	0.0219	0.0219	0.0037
		R2	0.7274	0.7274	0.7253	0.7266	0.9900
2011-2015	2016	RMSE	0.0356	0.0356	0.0354	0.0355	0.0045
		R2	0.3910	0.3910	0.3912	0.3911	0.9823
2012-2016	2017	RMSE	0.0279	0.0279	0.0279	0.0279	0.0031
		R2	0.6989	0.6989	0.6974	0.6983	0.9909
2013-2017	2018	RMSE	0.0244	0.0244	0.0243	0.0244	0.0025
		R2	0.7562	0.7562	0.7572	0.7566	0.9951
2014-2018	2019	RMSE	0.0230	0.0230	0.0229	0.0230	0.0040
		R2	0.6280	0.6280	0.6286	0.6283	0.9879
2015-2019	2020	RMSE	0.0269	0.0269	0.0269	0.0269	0.0083
		R2	0.5399	0.5399	0.5401	0.5400	0.9584
Mean		RMSE	0.0261	0.0261	0.0260	0.0261	0.0054
		R2	0.6614	0.6614	0.6614	0.6614	0.9832
Median		RMSE	0.0249	0.0249	0.0249	0.0249	0.0054
		R2	0.6715	0.6715	0.6710	0.6713	0.9838
Std. Dev.		RMSE	0.0040	0.0040	0.0039	0.0040	0.0017
		R2	0.0996	0.0996	0.0992	0.0995	0.0083

Table 13: Random Forest Results

As can be seen from the results, the random forest model performs very well. In the individual rolling samples, the forecasting error $RMSE$ lies between 0.0025 and 0.0086, with a coefficient of determination R^2 between 0.9909 and 0.9584. It also shows that across all samples an almost perfect explanation of the variance of the cat bond spread is recorded by the eleven input variables ($R^2_{mean} = 0.9838$) and the mean forecasting error is only around 0.5% ($RMSE_{mean} = 0.0054$), which is an extraordinary result with regard to the distribution as well as the quartiles of the spreads contained in the cat bond data set (see chapter 3.2). In comparison to the OLS regression and the penalized regressions (Lasso, Ridge, Elastic Net), the latter converging to the OLS specification due to the hyperparameter evaluation, it can be seen that a lower mean $RMSE$ of approximately 0.02 and a higher mean R^2 of approximately 0.32 can be recorded. Thus, the hypotheses $H_3(a)$ and $H_3(b)$ from chapter 4.2.3 with regard to the optimized model performance of the random forest model compared to OLS and penalized regression can be clearly confirmed. The test of hypothesis $H_3(c)$ is carried out in the same way in the next chapter in the context of the neural network performance.

4.5.2 Neural Network

In addition to the random forest models, *Neural Networks* represent another central concept in the area of ML, which is mainly used in the area of complex applications such as computer vision and natural language processing (Gu et al., 2018, p.2242). However, there has also been an increased application to regression-based problems in recent years. Neural networks are generally part of the ML-family of "deep learning", which is based on the fact that their flexibility is due to the connection of several telescoping layers of nonlinear predictor interactions (Gu et al., 2018, p.2242). This degree of complexity also makes neural networks the most opaque and difficult to interpret models, which have a high degree of parameterization across the board of ML-models (Gu et al., 2018, p.2242).

The main characteristics can be described as follows: A neural network consists of an input layer, one or more hidden layers and an output layer (Götze et al., 2020, p.12). Each layer is assigned a number of neurons (= nodes). The input layer receives the number of n neurons, the hidden layers m neurons and the output layer p neurons, where the input neurons represent the input variables of the model and the output neuron(s) represent the corresponding output variable(s) which are object to the forecasting target (Gallo, 2015, p.179-180). Furthermore, each neuron receives a number of input signals x_i in addition to connection weights w_i, which together result in the activation value y (Gallo, 2015, p.180). These weights indicate the connection between an input neuron and a hidden neuron as well as a hidden neuron and an output neuron (Götze et al., 2020, p.13). A central approach to weight updates is the concept of Backpropagation, which carries out the update on the basis of the gradient descent method, whereby the partial derivatives of the sum of squared errors are used with respect to the weights (Götze et al., 2020, p.13). The algorithm first determines the output forecast based on the initial weightings, whereupon the forecast errors are calculated and returned to the hidden and the input layers (Götze et al., 2020, p.13). In this way, the errors are calculated in each layer on the basis of the existing weightings and updated using the partial derivatives and the gradient descent (Götze et al., 2020, p.13). This method represents a widely used approach in the specification of neural networks. In addition, a predetermined activation function f transforms the activation value into an output according to $f(y)$ (Gallo, 2015, p.180). All these central components of a neural network are determined using hyperparameter tuning (see chapter 4.3). The overall general structure of neural networks therefore results in the following picture (assuming a model with one hidden layer and multiple output variables):

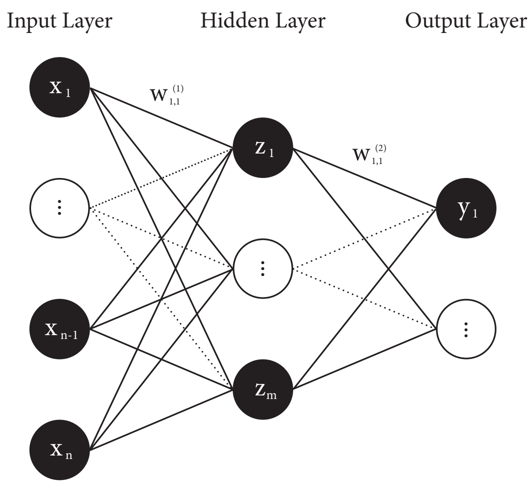

Figure 4: General Neural Network Structure - Own illustration (based on Götze et al., 2020, p.12)

In the context of this work, the cat bond spread represents the output variable, whereby the number of neurons in the output layer is consequently set equal to one ($p = 1$). With regard to the input variables, the model is based on the data-based evaluation of the pricing model (see chapter 3.4), which is ultimately based on eleven explaining variables ($n = 11$), eight cat bond-specific and three macroeconomic/financial market-related. The number of hidden layers and their number of neurons are objects of hyperparameter tuning. Table 11 in chapter 4.3 shows that within the framework of the model specification three hidden layers are used, to which in turn 4, 7 and 5 neurons are assigned. Furthermore, regarding the activation function σ, it emerges from the hyperparameter tuning that the *Hyperbolic Tangent Activation Function (tanh)*, defined by $\sigma(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ (Götze et al., 2020, p.13), performs best given the set of parameters. This function is generally based on the sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$, with activated neurons taking the values in the interval $[-1, 1]$ instead of $[0, 1]$ in the case of the sigmoid function (Götze et al., 2020, p.13). As already described in the general neural network strucutre, the weight updating process is take place under the backpropagation approach. In addition, it is important to evaluate the correct data-based optimizer function regarding the model learning rates. In the context of the model specification and data of this thesis, the *Adam* optimizer, which computes adaptive learning rates for different parameters from estimates of first and second moments, is determined (Kingma and Ba, 2015, p.1). Furthermore, the number of cycles through the full training data set can be determined, which is defined as the number of epochs (Deep AI, 2021). Within the context of this model, the number of epochs are set to 100, which in turn means that the full training data set will be run through the training process 100 times. Finally, the batch size has to be determined, which defines how many sub-samples of the in-sample training set are taken into account per training session and thus after how many observations in the training process the weights are getting updated. For example, if the training set has over 1000 observations and the batch size is set to 200 observations, five training sessions of 200 observations are run through. The results of the hyperparameter tuning show that, within the scope of this thesis, the batch size is determined as 100 and thus, in combination with the other validated hyperparameters, achieves the best results with regard to the evaluation criterion $RMSE$.

The calculation of the cat bond spread forecasting error defined as the $RMSE$ and the coefficient of determination R^2 using the predefined rolling samples takes place under the neural network specification just explained. The last column of Table 14 shows the results for each IS-OOS combination as well as mean, media and standard deviation values across all rolling samples.

IS	OOS	Measure	OLS	Lasso	Ridge	Elastic Net	RF	NN
2003-2007	2008	RMSE	0.0240	0.0240	0.0240	0.0240	0.0067	0.0241
		R2	0.8306	0.8306	0.8293	0.8304	0.9869	0.8162
2004-2008	2009	RMSE	0.0336	0.0336	0.0334	0.0335	0.0086	0.0288
		R2	0.5861	0.5861	0.5879	0.5865	0.9735	0.6660
2005-2009	2010	RMSE	0.0290	0.0290	0.0289	0.0290	0.0070	0.0270
		R2	0.6360	0.6360	0.6366	0.6362	0.9796	0.6839
2006-2010	2011	RMSE	0.0253	0.0253	0.0253	0.0253	0.0058	0.0236
		R2	0.6731	0.6731	0.6738	0.6733	0.9822	0.7258
2007-2011	2012	RMSE	0.0214	0.0214	0.0214	0.0214	0.0054	0.0245
		R2	0.7242	0.7242	0.7247	0.7244	0.9845	0.7247
2008-2012	2013	RMSE	0.0216	0.0216	0.0216	0.0216	0.0066	0.0253
		R2	0.7374	0.7374	0.7373	0.7374	0.9818	0.7381
2009-2013	2014	RMSE	0.0245	0.0245	0.0245	0.0245	0.0046	0.0240
		R2	0.6699	0.6699	0.6682	0.6693	0.9890	0.6803
2010-2014	2015	RMSE	0.0219	0.0219	0.0219	0.0219	0.0037	0.0188
		R2	0.7274	0.7274	0.7253	0.7266	0.9900	0.7159
2011-2015	2016	RMSE	0.0356	0.0356	0.0354	0.0355	0.0045	0.0203
		R2	0.3910	0.3910	0.3912	0.3911	0.9823	0.6572
2012-2016	2017	RMSE	0.0279	0.0279	0.0279	0.0279	0.0031	0.0189
		R2	0.6989	0.6989	0.6974	0.6983	0.9909	0.7966
2013-2017	2018	RMSE	0.0244	0.0244	0.0243	0.0244	0.0025	0.0174
		R2	0.7562	0.7562	0.7572	0.7566	0.9951	0.8673
2014-2018	2019	RMSE	0.0230	0.0230	0.0229	0.0230	0.0040	0.0196
		R2	0.6280	0.6280	0.6286	0.6283	0.9879	0.8102
2015-2019	2020	RMSE	0.0269	0.0269	0.0269	0.0269	0.0083	0.0210
		R2	0.5399	0.5399	0.5401	0.5400	0.9584	0.7284
Mean		RMSE	0.0261	0.0261	0.0260	0.0261	0.0054	0.0226
		R2	0.6614	0.6614	0.6614	0.6614	0.9832	0.7393
Median		RMSE	0.0249	0.0249	0.0249	0.0249	0.0054	0.0231
		R2	0.6715	0.6715	0.6710	0.6713	0.9838	0.7271
Std. Dev.		RMSE	0.0040	0.0040	0.0039	0.0040	0.0017	0.0031
		R2	0.0996	0.0996	0.0992	0.0995	0.0083	0.0580

Table 14: Neural Network Results

As can already be seen from the results, there is good forecasting performance across all rolling samples, which range from 0.0174 to 0.0288 for the $RMSE$ and from 0.8673 to 0.6660 for the coefficient of determination R^2. Furthermore, the hypothesis criteria show that a $RMSE_{mean}$ of 0.0226 and a R^2_{mean} of 0.7393 are achieved. Looking back to the results of the OLS regression and the penalized regressions (Lasso, Ridge, Elastic Net), a better performance can be identified with a lower $RMSE_{mean}$ of around 0.004 and an higher R^2_{mean} of around 0.08, whereby the hypotheses $H_2(a)$ and $H_2(b)$ are confirmed, which include a higher coefficient of determination and a lower forecasting error. The hypothesis $H_3(c)$, which describes a lower performance of neural networks compared to random forest models, can also be confirmed. A higher mean $RMSE$ of around 0.017 and a lower mean determination coefficient R^2 of around 0.16 of the neural network can be identified. In conclusion, with regard to neural networks, it can be said that they optimize the forecasting of cat bond spreads under the predetermined criteria and can thus be used as a valid instrument in the context of relevant analyzes. Nonetheless, the results of this thesis clearly show that although both ML models bring optimized results, the random forest model clearly outperforms all other models, and in particular the econometric approach of linear regressions based on Braun (Braun, 2014) and Gürtler et. al (Gürtler et al., 2012). Lastly, it should be mentioned that the neural network in this thesis compared to the one in the primary market analysis by Götze et. al (2020, p.22) outperforms any regression models, which might result from the more extensive data set as well as the more extensive specification of the model with regard to the hyperparameters. Thus, the application of neural networks over regression-based approaches for the cat bond spread forecasting is recommended compared to Götze et. al (2020).

4.6 Robustness Check

After the out-of-sample performances of all benchmark regression models and the optimizing ML models have been evaluated in the previous chapters, they must be subjected to a robustness check. The robustness check is primarily used to test model resilience in the form of changed, non-standard conditions (Micskei et al., 2012, p.1). Thus, it pursues the primary goal of identifying errors in the model specification and the associated falsified forecasting results (Micskei et al., 2012, p.1). In the context of this thesis, the check concentrates especially on testing the robustness against time-specific effects of the explanatory variables[15]. This is done by using enlarged out-of-samples of two years on the same basis of the in-samples of four years. Table 15 shows the results of the robustness check, which shows the forecasting performance for each model in the form of the $RMSE$ and the R^2 for the individual rolling samples as well as across all rolling samples in aggregated statistics.

IS	OOS	Measure	OLS	Lasso	Ridge	Elastic Net	RF	NN
2003-2007	2008-2009	RMSE	0.0240	0.0240	0.0240	0.0240	0.0067	0.0241
		R2	0.8306	0.8306	0.8293	0.8304	0.9869	0.8162
2004-2008	2009-2010	RMSE	0.0336	0.0336	0.0334	0.0335	0.0086	0.0288
		R2	0.5861	0.5861	0.5879	0.5865	0.9735	0.6660
2005-2009	2010-2011	RMSE	0.0290	0.0290	0.0289	0.0290	0.0070	0.0270
		R2	0.6360	0.6360	0.6366	0.6362	0.9796	0.6839
2006-2010	2011-2012	RMSE	0.0253	0.0253	0.0253	0.0253	0.0058	0.0236
		R2	0.6731	0.6731	0.6738	0.6733	0.9822	0.7258
2007-2011	2012-2013	RMSE	0.0214	0.0214	0.0214	0.0214	0.0054	0.0245
		R2	0.7242	0.7242	0.7247	0.7244	0.9845	0.7247
2008-2012	2013-2014	RMSE	0.0216	0.0216	0.0216	0.0216	0.0066	0.0253
		R2	0.7374	0.7374	0.7373	0.7374	0.9818	0.7381
2009-2013	2014-2015	RMSE	0.0245	0.0245	0.0245	0.0245	0.0046	0.0240
		R2	0.6699	0.6699	0.6682	0.6693	0.9890	0.6803
2010-2014	2015-2016	RMSE	0.0219	0.0219	0.0219	0.0219	0.0037	0.0188
		R2	0.7274	0.7274	0.7253	0.7266	0.9900	0.7159
2011-2015	2016-2017	RMSE	0.0356	0.0356	0.0354	0.0355	0.0045	0.0203
		R2	0.3910	0.3910	0.3912	0.3911	0.9823	0.6572
2012-2016	2017-2018	RMSE	0.0279	0.0279	0.0279	0.0279	0.0031	0.0189
		R2	0.6989	0.6989	0.6974	0.6983	0.9909	0.7966
2013-2017	2018-2019	RMSE	0.0244	0.0244	0.0243	0.0244	0.0025	0.0174
		R2	0.7562	0.7562	0.7572	0.7566	0.9951	0.8673
2014-2018	2019-2020	RMSE	0.0230	0.0230	0.0229	0.0230	0.0040	0.0196
		R2	0.6280	0.6280	0.6286	0.6283	0.9879	0.8102
Mean		RMSE	0.0260	0.0260	0.0260	0.0260	0.0052	0.0227
		R2	0.6716	0.6716	0.6715	0.6715	0.9853	0.7402
Median		RMSE	0.0245	0.0245	0.0245	0.0245	0.0052	0.0236
		R2	0.6731	0.6731	0.6738	0.6733	0.9853	0.7258
Std. Dev.		RMSE	0.0041	0.0041	0.0041	0.0041	0.0016	0.0032
		R2	0.0973	0.0973	0.0970	0.0972	0.0052	0.0600

Table 15: Robustness Check Results

[15] The definition of the rolling samples as part of the robustness check is in line with the approach of Götze et. al (2020, p.23), which further legitimizes the possibility of model comparison between primary market and secondary market spreads.

The results of the robustness check show that the random forest model and the neural network continue to dominate the performance of OLS and the penalized regressions. A look at the aggregated performance criteria shows that the OLS regression with a $RMSE_{mean}$ of 0.0260 and a R^2_{mean} of 0.6716 achieves almost the same results as the previous OOS results (0.0261 and 0.6614). The same applies to the penalized regressions Lasso/Ridge/Elastic Net, which in the robustness check show a $RMSE_{mean}$ of 0.0260/0.0260/ 0.0260 and a R^2_{mean} of 0.6716/0.6715/0.6715 compared to the previous results (0.0261/0.0260/ 0.0261 and 0.6614/ 0.6614/0.6614). The random forest model shows a $RMSE_{mean}$ of 0.0052 and a R^2_{mean} of 0.9853, which are also very close to the previous results of 0.0054 and 0.9832. Ultimately, the same scheme can be identified for the neural network, in which a $RMSE_{mean}$ of 0.0227 and a R^2_{mean} of 0.7402 compared to 0.0226 and 0.7393 are achieved in the robustness check. Based on the performance comparison, both hypothesis $H_4(a)$ and hypothesis $H_4(b)$, which test the models' performances and ratings across the enlarged out-of-samples, can be confirmed.

Finally, after all hypotheses have been tested within the application of the regression-based and ML models as well as within the robustness check, the overall results of the main analysis of this thesis can be summarized as follows:

	Test Criteria	Test Result
$\boldsymbol{H_1}$	$RMSE^{Penalized}_{mean,OOS} < RMSE^{OLS}_{mean,OOS}$, $R^{2\ Penalized}_{mean,OOS} > R^{2\ OLS}_{mean,OOS}$	rejected
$\boldsymbol{H_2(a)}$	$RMSE^{NN}_{mean,OOS} < RMSE^{OLS}_{mean,OOS}$, $R^{2\ NN}_{mean,OOS} > R^{2\ OLS}_{mean,OOS}$	confirmed
$\boldsymbol{H_2(b)}$	$RMSE^{NN}_{mean,OOS} < RMSE^{Penalized}_{mean,OOS}$, $R^{2\ NN}_{mean,OOS} > R^{2\ Penalized}_{mean,OOS}$	confirmed
$\boldsymbol{H_3(a)}$	$RMSE^{RF}_{mean,OOS} < RMSE^{OLS}_{mean,OOS}$, $R^{2\ RF}_{mean,OOS} > R^{2\ OLS}_{mean,OOS}$	confirmed
$\boldsymbol{H_3(b)}$	$RMSE^{RF}_{mean,OOS} < RMSE^{Penalized}_{mean,OOS}$, $R^{2\ RF}_{mean,OOS} > R^{2\ Penalized}_{mean,OOS}$	confirmed
$\boldsymbol{H_3(c)}$	$RMSE^{RF}_{mean,OOS} < RMSE^{NN}_{mean,OOS}$, $R^{2\ RF}_{mean,OOS} > R^{2\ NN}_{mean,OOS}$	confirmed
$\boldsymbol{H_4(a)}$	$RMSE^{OLS}_{mean,OOS_{adj}} \approx RMSE^{OLS}_{mean,OOS}$, $R^{2\ OLS}_{mean,OOS_{adj}} \approx R^{2\ OLS}_{mean,OOS}$ $RMSE^{Penalized}_{mean,OOS_{adj}} \approx RMSE^{Penalized}_{mean,OOS}$, $R^{2\ Penalized}_{mean,OOS_{adj}} \approx R^{2\ Penalized}_{mean,OOS}$ $RMSE^{NN}_{mean,OOS_{adj}} \approx RMSE^{NN}_{mean,OOS}$, $R^{2\ NN}_{mean,OOS_{adj}} \approx R^{2\ NN}_{mean,OOS}$ $RMSE^{RF}_{mean,OOS_{adj}} \approx RMSE^{RF}_{mean,OOS}$, $R^{2\ RF}_{mean,OOS_{adj}} \approx R^{2\ RF}_{mean,OOS}$	confirmed
$\boldsymbol{H_4(b)}$	$RMSE^{RF}_{mean,OOS_{adj}} < RMSE^{NN}_{mean,OOS_{adj}} < RMSE^{Penalized}_{mean,OOS_{adj}} \approx RMSE^{OLS}_{mean,OOS_{adj}}$ $R^{2\ RF}_{mean,OOS_{adj}} > R^{2\ NN}_{mean,OOS_{adj}} > R^{2\ Penalized}_{mean,OOS_{adj}} \approx R^{2\ OLS}_{mean,OOS_{adj}}$	confirmed

Table 16: Hypotheses Test Results

5 Conclusion and Outlook

The aim of this thesis is to optimize the spread forecasting within the context of cat bond pricing. For this purpose, an extensive cat bond secondary market data set represents the foundation on which the benchmark models of OLS, Lasso, Ridge and Elastic Net as well as the optimized ML models of a random forest and a neural network are applied, all under the premise of specifically tuned hyperparameters. On the one hand, the results show that the application of the penalized regressions Lasso, Ridge and Elastic Net, in order to enhance the benchmark regression-based models, does not lead to any improvement in spread forecasting compared to OLS. This results from the determination of the respective shrinkage parameters, which reduce the penalized regression models approximately to an OLS specification by almost fully omitting the shrinkage terms. Thus, both the OLS regression and the penalized regression offer the same benchmark results for the ML optimization. On the other hand, with regard to the neural network, it can be seen that a significantly better forecasting performance is achieved. However, the highest performance, which also comes on its own with almost perfect goodness-of-fit and very low forecasting errors over all rolling samples as well as in the aggregated statistics, is achieved by the random forest model. In addition, all models hold during the robustness check and, despite increased out-of-samples, show the approximately same forecasting performances. Furthermore, the essential cat bond pricing determinants based on Braun (2014) and Gürtler et. al (2012) can be largely confirmed, whereby the term to maturity was not statistically significant. Moreover, the additional determinant of seasonality underlines the seasonal components of cat bond spreads in the secondary market. Looking back to the ML optimized spread forecasting in the context of primary market data by Götze et. al (2020), it can be concluded that the performance of the ML models can be improved with a larger cat bond data set as well as the exact evaluation and tuning of the model hyperparameters in context of the model specification, with neural networks being especially sensitive to the model parameter determination. In conclusion, the secondary market analysis of this thesis shows that the application of ML models for the cat bond spread forecasting has enormous potential for optimization and represents a significantly better alternative to regression-based models.

With regard to the existing research on ILS instruments and especially cat bonds, this thesis provides several key findings. The data-based evaluation of the econometric pricing model complements an important addition in the still thin research field of the secondary market. Furthermore, the ML-based optimization of the spread forecasting under secondary market data represents a novel approach, whereby the previous primary market results by Götze et. al (2020) can be completed with regard to the entire cat bond market. Moreover, the results show that ML methods, in particular random forest models, represent a valid approach to optimize general asset pricing, also outside of the ILS market. This applies above all to financial market-related research areas in which the essential explanatory input variables are known, but the distribution is largely unknown or this changes with the prevailing market conditions.

List of References

Adedia, D., Adebanji, A., Labeodan, M., & Adeyemi, S. (2016). Ordinary least squares and robust estimators in linear regression: Impacts of outliers, error and response contaminations. https://www.sciencedomain.org/download/MTI3MTdAQHBm.pdf

Ahangar, R., Yahyazadehfar, M., & Pournaghshband, H. (2010). The comparison of methods artificial neural network with linear regression using specific variables for prediction stock price in tehran stock exchange. *(IJCSIS) International Journal of Computer Science and Information Security*. https://arxiv.org/pdf/1003.1457.pdf

Astivia, O. L. O., & Zumbo, B. D. (2019). Heteroskedasticity in multiple regression analysis: What it is, how to detect it and how to solve it with applications in r and spss. https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1331&context=pare

Avati, A. (2020). Bias-variance analysis: Theory and practice. http://cs229.stanford.edu/summer2020/BiasVarianceAnalysis.pdf

Berge, K. (2005). *Katastrophenanleihen: Anwendung, bewertung, gestaltungsempfehlungen* (Doctoral dissertation). Techn. Universität Dresden. Lohmar, Eul. https://epub.uni-regensburg.de/8751/

Braun, A. (2014). Pricing in the primary market for cat bonds: New empirical evidence. *Working Paper on Risk Management and Insurance No. 116*. https://www.ivw.unisg.ch/~/media/internet/content/dateien/instituteundcenters/ivw/wps/wp116.pdf

Breiman, L. (1996). Bagging predictors. https://link.springer.com/content/pdf/10.1007/BF00058655.pdf

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? https://dash.harvard.edu/handle/1/2622619

Carayannopoulos, P., & Perez, F. M. (2015). Diversification through catastrophe bonds: Lessons from the subprime financial crisis. *The Geneva Papers*. https://link.springer.com/content/pdf/10.1057/gpp.2014.14.pdf

Ceh, M., Kilibarda, M., Lisec, A., & Bajat, B. (2018). Estimating the performance of random forest versus multiple regression for predicting prices of the apartments. https://www.researchgate.net/publication/324916118_Estimating_the_Performance_of_Random_Forest_versus_Multiple_Regression_for_Predicting_Prices_of_the_Apartments

Cummins, D., & Weiss, M. (2009). Convergence of insurance and financial markets: Hybrid and securitized risk-transfer solutions. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6975.2009.01311.x

Cutler, A., Cutler, D., & Stevens, J. (2011). Random forests. *Machine Learning*. https://www.researchgate.net/publication/236952762_Random_Forests

Deep AI. (2021). *Epoch*. https://deepai.org/machine-learning-glossary-and-terms/epoch

Dieckmann, S. (2010). A consumption-based evaluation of the cat bond market. https://www.emerald.com/insight/content/doi/10.1108/S2514-465020190000007002/full/html

Edwards, A. K., Harris, L. E., & Piwowar, M. S. (2007). Corporate bond market transaction costs and transparency. *The Journal of Finance*, 1421–1451. http://www.jstor.org/stable/4622305

Evans, S. (2021a). *Catastrophe bond & ils market dashboard.* https://www.artemis.bm/dashboard/

Evans, S. (2021b). *Catastrophe bonds & ils outstanding by coupon pricing.* https://www.artemis.bm/dashboard/cat-bonds-ils-by-coupon-pricing/

Evans, S. (2021c). *Catastrophe bonds & ils risk capital issued & outstanding per year.* https://www.artemis.bm/dashboard/catastrophe-bonds-ils-issued-and-outstanding-by-year/

FED of St. Louis. (2021). *Ice bofa bbb us corporate index option-adjusted spread.* https://fred.stlouisfed.org/series/BAMLC0A4CBBB

Feurer, M., & Hutter, F. (2019). Chapter 1 - hyperparameter optimization. https://www.automl.org/wp-content/uploads/2019/05/AutoML_Book_Chapter1.pdf

Figueiredo, D., Silva, J., & Rocha, E. (2011). What is r2 all about? https://www.researchgate.net/publication/312451551_What_is_R2_all_about

Galeotti, M., Gürtler, M., & Winkelvos, C. (2012). Accuracy of premium calculation models for cat bonds—an empirical analysis. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6975.2012.01482.x

Gallo, C. (2015). Empirical asset pricing via machine learning. *Artificial Neural Networks: Tutorial.* https://www.researchgate.net/publication/261392616_Artificial_Neural_Networks_tutorial

Ginestet, C. (n.d.). Ma 575 linear models: Regularization: Ridge regression and lasso. http://math.bu.edu/people/cgineste/classes/ma575/p/w14_1.pdf

Götze, T., Gürtler, M., & Witowski, E. (2020). How to deal with small data sets in machine learning: An analysis on the cat bond market. https://www.researchgate.net/publication/342395001_Improving_CAT_bond_pricing_models_via_machine_learning

Gu, S., Kelly, B., & Xiu, D. (2018). Empirical asset pricing via machine learning. *The Review of Financial Studies.* https://dachxiu.chicagobooth.edu/download/ML.pdf

Gürtler, M., Hibbeln, M., & Winkelvos, C. (2012). The impact of the financial crisis and natural catastrophes on cat bonds. https://www.econstor.eu/obitstream/10419/64631/1/725704713.pdf

Herrmann, M., & Hibbeln, M. (2019). Seasonality in catastrophe bonds and market-implied catastrophe arrival frequencies. *The Journal of Risk and Insurance.* https://onlinelibrary.wiley.com/doi/full/10.1111/jori.12335

Jaeger, L., Müller, S., & Scherling, S. (2010). Insurance-linked securities: What drives their returns? *The Journal of Alternative Investments Fall 2010.* https://jai.pm-research.com/content/13/2/9

Kingma, D., & Ba, J. (2015). Adam: A method for stochastic optimization. https://arxiv.org/pdf/1412.6980.pdf

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling.* Springer.

Kyung, M., Gilly, J., Ghoshz, M., & Casellax, G. (2010). Penalized regression, standard errors, and bayesian lassos. *Bayesian Analysis.* https://www.researchgate.net/publication/228619969_Penalized_Regression_Standard_Errors_and_Bayesian_Lassos

Lane, M. (2000). Pricing risk transfer transactions. https://www.casact.org/sites/default/files/2021-02/2001-lane.pdf

Lane, M., & Mahul, O. (2008). Catastrophe risk pricing: An empirical analysis. https://www.researchgate.net/publication/23550633_Catastrophe_Risk_Pricing_An_Empirical_Analysis

Lee, J.-P., & Yu, M.-T. (2003). Pricing default-risky cat bonds with moral hazard and basis risk. https://www.researchgate.net/publication/228289849_Pricing_Default-Risky_Cat_Bonds_With_Moral_Hazard_and_Basis_Risk

Lei, D. T., Wang, J.-H., & Tzeng, L. Y. (2008). Explaining the spread premiums on catastrophe bonds. http://www.fin.ntu.edu.tw/~conference/conference2008/proceedings/proceeding/10/10-2(A39).pdf

Liashchynskyi, P., & Liashchynskyi, P. (2019). Grid search, random search, genetic algorithm: A big comparison for nas. https://arxiv.org/pdf/1912.06059.pdf

Makariou, D., Barrieu, P., & Chen, Y. (2020). A random forest based approach for predicting spreads in the primary catastrophe bond market. https://www.semanticscholar.org/paper/A-random-forest-based-approach-for-predicting-in-Makariou-Barrieu/2e3b5fa9ea0907af7babada0acfc4be1cede2ba9

Micskei, Z., Madeira, H., Avritzer, A., Majzik, I., Vieira, M., & Antunes, N. (2012). Robustness testing techniques and tools. https://www.researchgate.net/publication/278692960_Robustness_Testing_Techniques_and_Tools

Miller, J. W. (2013). Lecture 11: Penalized regression - statistical learning (bst 263). https://jwmi.github.io/SL/11-Penalized-regression.pdf

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*. https://arxiv.org/pdf/1003.1457.pdf

Ogutu, J. O., Schulz-Streeck, T., & Piepho, H.-P. (2012). Genomic selection using regularized linear regression models: Ridge regression, lasso, elastic net and their extensions. https://link.springer.com/content/pdf/10.1186/1753-6561-6-S2-S10.pdf

PartnerRe. (2015). The drivers of catastrophe bond pricing. https://www.partnerre.com/assets/uploads/docs/The-Drivers-of-Catastrophe-Bond-Pricing.pdf

PennState University. (2021). *Lesson 12: Multicollinearity & other regression pitfalls*. https://online.stat.psu.edu/stat501/book/export/html/981

Risk Management Solutions. (2012). Cat bonds demystified - rms guide to the asset class. https://forms2.rms.com/rs/729-DJX-565/images/cm_cat_bonds_demystified.pdf

Schonlau, M., & Zou, R. (2020). The random forest algorithm for statistical learning. https://journals.sagepub.com/doi/pdf/10.1177/1536867X20909688

SwissRe. (2012). What are insurance linked securities (ils), and why should they be considered? https://www.casact.org/community/affiliates/CANE/0912/Cat-Bond.pdf

Taylor, J. (n.d.). Statistics 203: Introduction to regression and analysis of variance - penalized models. https://statweb.stanford.edu/~jtaylo/courses/stats203/notes/penalized.pdf

Thornton, D. L., & Valente, G. (2010). Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective. https://www.semanticscholar.org/paper/Out-of-Sample-Predictions-of-Bond-Excess-Returns-An-Thornton-Valente/0a8ae9e78307e614d21e00153aab9b8d48a76c77

Wang, S. S. (2000). A class of distortion operators for pricing financial and insurance risks. https://www.jstor.org/stable/253675?seq=1

Weistroffer, C. (2010). Insurance-linked securities - a niche market expanding.

Young, C. (2015). Model uncertainty and robustness: A computational framework for multi-model analysis. https://web.stanford.edu/~cy10/public/mrobust/Model_Robustness.pdf

Appendix

Appendix 1: Rating Aggregation

	S&P	Moody's	Fitch	Aggregated	Ratings
IG	AAA	Aaa	AAA	AAA	6
	AA+	Aa1	AA+	AA	5
	AA	Aa2	AA	AA	5
	AA-	Aa3	AA-	AA	5
	A+	A1	A+	A	4
	A	A2	A	A	4
	A-	A3	A-	A	4
	BBB+	Baa1	BBB+	BBB	3
	BBB	Baa2	BBB	BBB	3
	BBB-	Baa3	BBB-	BBB	3
NIG	BB+	Ba1	BB+	BB	2
	BB	Ba2	BB	BB	2
	BB-	Ba3	BB-	BB	2
	B+	B1	B+	B	1
	B	B2	B	B	1
	B-	B3	B-	B	1
					1
NR	-	-	-	NR	0

Table 17: Aggregated Ratings - Overview

Appendix 2: Correlation Heatmap

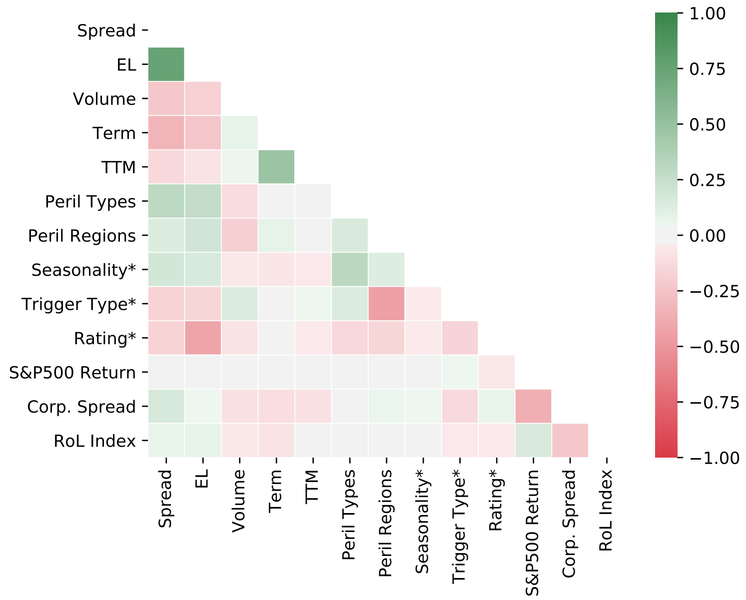

Figure 5: Correlation Heatmap - Own illustration

Declaration of Authorship

"I hereby declare

- that I have written this thesis without any help from others and without the use of documents and aids other than those stated above;
- that I have mentioned all the sources used and that I have cited them correctly according to established academic citation rules;
- that I have acquired any immaterial rights to materials I may have used such as images or graphs, or that I have produced such materials myself;
- that the topic or parts of it are not already the object of any work or examination of another course unless this has been explicitly agreed on with the faculty member in advance and is referred to in the thesis;
- that I will not pass on copies of this work to third parties or publish them without the University's written consent if a direct connection can be established with the University of St.Gallen or its faculty members;
- that I am aware that my work can be electronically checked for plagiarism and that I hereby grant the University of St.Gallen copyright in accordance with the Examination Regulations in so far as this is required for administrative action;
- that I am aware that the University will prosecute any infringement of this declaration of authorship and, in particular, the employment of a ghostwriter, and that any such infringement may result in disciplinary and criminal consequences which may result in my expulsion from the University or my being stripped of my degree."

Date and signature

May 22$^{\text{nd}}$ 2021, ...

By submitting this academic term paper, I confirm through my conclusive action that I am submitting the Declaration of Authorship, that I have read and understood it, and that it is true.

Declaration of Discretion

The undersigned

hereby undertakes and warrants to treat any information obtained by Lane Financial LLC, Swiss Re Group and Aon plc concerned in strict confidentiality. In particular, he shall only permit people other than the referees to inspect his written work with the express consent of all the parties that have provided information.

The undersigned hereby takes cognizance of the fact that the University of St. Gallen may check his work for any plagiarism with the help of a plagiarism software and that the undersigned shall have to notify Lane Financial LLC, Swiss Re Group and Aon plc surveyed accordingly.

Date and signature

May 22$^{\text{nd}}$ 2021, ...