Portfolio Diversification Through Catastrophe Bonds: An Empirical Analysis

Paul Luis Schröder[1]

Francesco Zaccaron[2]

May 2025

Master's Thesis

Department of Finance

Stockholm School of Economics

Supervisor: Tobias Sichert

Abstract

This thesis evaluates the diversification benefits of catastrophe (Cat) bonds in multi-asset portfolios using weekly data from March 2005 to March 2025. We analyze Cat bond correlations with traditional assets (equities, corporate bonds, treasuries) across different market regimes (including the Financial Crisis and COVID-19) and test if benchmark assets span Cat bonds using mean-variance spanning tests. We perform out-of-sample portfolio backtesting incorporating transaction costs to assess practical benefits. Results confirm Cat bonds offer a strong risk-return profile and low correlation with traditional assets, largely maintained during high-volatility periods. Spanning tests indicate significant diversification benefits, including variance reduction across market regimes and alpha generation primarily during normal market conditions. Out-of-sample tests show including Cat bonds generally improves point estimates of risk-adjusted returns and reduces drawdowns, even with transaction costs, although in formal statistical tests significance is mixed. We conclude Cat bonds are effective diversifiers, improving portfolio resilience across market conditions.

Keywords: catastrophe bonds, Insurance-Linked Securities (ILS), portfolio diversification, asset allocation, risk management, alternative investments, financial markets.

[1]42671@student.hhs.se

[2]42689@student.hhs.se

Contents

1 Introduction

In 2024, the catastrophe (Cat) bond market surpassed the milestone of over USD 50 billion in instruments outstanding, driven by the issuance of roughly USD 17.7 billion in new Cat bonds over the year and exhibiting annual returns in excess of 16% (Naik, 2024). Such rapid expansion and impressive returns naturally invite closer scrutiny of Cat bonds and their potential role within modern diversified investment portfolios. At the same time, the financial impacts of natural disasters continue to escalate globally and traditional reinsurance markets increasingly face limitations due to cyclical pricing behavior and capacity constraints. As a result, insurance-linked securities (ILS), and particularly Cat bonds, have emerged as crucial instruments, enabling insurers and reinsurers to transfer catastrophic risks directly to capital market participants (Morana and Sbrana, 2019).

Against this dual backdrop, this thesis aims to examine the role these instruments can play in modern diversified portfolios, while also considering their potential contribution to addressing some of the Sustainable Development Goals (SDGs).

Cat bonds represent a distinctive asset class primarily due to their theoretically low correlation with traditional financial assets, which has suggested over the years notable diversification potential (Braun, 2016; Demers-Bélanger and Lai, 2020; Drobetz et al., 2020; Haffar and Le Fur, 2022). These securities enable issuers (usually insurance and reinsurance firms, but also sovereigns and multilateral institutions) to transfer specific catastrophic risks to investors. Buyers are compensated through periodic coupon payments made up of the return on collateral reinvestment as well as the additional risk premium based on the specific catastrophe risk attached to the bond. If a specified catastrophic event occurs and meets the predefined trigger conditions, the invested amounts are used to cover - partly or fully - the resulting losses. This way, investors can expose themselves to specific physical risks theoretically uncorrelated with economic or financial cycles and events (Braun, 2016; Kish, 2016; Drobetz et al., 2020).

Despite widespread acknowledgement of cat bonds' diversification potential, opportunities for further research remain (Braun, 2016; Demers-Bélanger and Lai, 2020; Drobetz et al., 2020; Haffar and Le Fur, 2022). Specifically, this study aims to evaluate the performance and risk characteristics of Cat bonds both individually and within diversified portfolios, examining their behavior under various market conditions and during periods of stress, including the Global Financial Crisis (GFC), the 2017 hurricane season, the COVID-19 pandemic shock, and the recent inflationary environment. Our approach encompasses a comprehensive correlation analysis between Cat bonds and set of benchmark assets, formal mean-variance spanning tests and extensive out-of-sample portfolio testing.

Our analysis, drawing on data from 2005 to 2025 finds that Cat bonds retain low correlation with major asset classes even during market stress periods. While these correlations showed some time variation, they remained relatively stable even during crisis periods. Furthermore, asset spanning tests provided evidence that Cat bonds offer diversification benefits through both return enhancement and variance reduction. Our out-of-sample portfolio backtesting provided practical evidence of Cat bonds' diversification benefits, even when accounting for transaction costs. Across allocation strategies and rebalancing frequencies, portfolios including Cat bonds consistently outperformed their counterparts in terms of risk-adjusted metrics and downside protection.

Our evidence therefore suggests that Cat bonds can significantly enhance risk adjusted portfolio performance and resilience.

The remainder of this paper is organized as follows: Section 2 provides an overview of the history and structure of Cat bonds, and discusses their potential role regarding the SDGs. Section 3 reviews the relevant academic literature on the risk-return profile and diversification benefits of Cat bonds. Section 4 describes the data and presents the methodology employed in the empirical analysis. Section 5 reports the empirical results, and Section 6 contextualizes them within the existing academic discussion. Section 7 highlights the limitations of this study, and Section 8 draws conclusions from our findings.

2 Cat Bonds: History, Structure and SDG implications

This section will serve as a broad overview of the Insurance Linked Securities (ILS) market and more specifically the catastrophe bond market. It will provide a brief historical overview of both markets, describe the structure of typical Cat bond issue, and finish by looking at current Cat bond market trends and their broader implications for global SDG.

2.1 Cat bonds: History

Although reinsurance has been around for decades, research into hybrid financial reinsurance products is a more recent phenomenon (Cummins and Weiss, 2009). The unprecedented losses caused by Hurricane Andrew in 1992 (estimated at USD 15.5 billion in 1992 dollars), which caused eight insurance companies to go bankrupt starkly revealed the capacity constraints within the traditional insurance and reinsurance system (Cummins et al., 2002; Cummins and Trainar, 2009). These events highlighted the need for novel sources of risk-bearing capacity and triggered intensified efforts by academics and financial innovators to develop vehicles transferring catastrophe risk to capital markets. In this period, the first insurance futures and options were also being introduced by the Chicago Board of Trade (Cummins and Weiss, 2009; Mevorach, 2018; Polacek, 2018). Shortly after, the catastrophic insurance losses (estimated at USD 15.3 billion) brought by the 1994 Northridge earthquake reopened the debate on optimal risk transfer mechanisms and added to the industry stress. This event also notably saw insurance companies starting to refuse to cover earthquake risk for the first time, which in turn led to the creation of the California Earthquake Authority (CEA), in order to provide insurance in the State. The same year witnessed the first Cat bond issuance, a 5-year bond offered by Hannover Re, for USD 85 million (Zeller, 2008). But while the next few years saw some other small and private issuances - such as Hannover Re and St. Paul Re in in 1996 (Canabarro et al., 2000) - it wasn't until 1997 that the first major, widely recognized, tradable (i.e. SEC Rule 144A compliant) public Cat bond issuance took place (Cummins and Weiss, 2009; Kish, 2016): it was again the CEA which was looking for new forms of financing insurance risk and planned to use capital markets to raise USD 1 billion in the first large Cat bond offering. However, the deal fell through, delaying the first large public Cat bond offering to 1997, when USAA (an insurance company for military personnel and veterans) raised USD 500 million to cover hurricane risk, albeit with a maturity of just one year (Mevorach, 2018).

Since the inception of Cat bonds in the years following Hurricane Andrew, several key milestones and development phases shaped the Cat bond market. These include Hurricane Katrina in 2005, the global financial crisis in 2008, the following period of low interest rates fueling investor appetite, and finally the hurricane season of 2017 and the resiliency and growth of the market since (Mevorach, 2018; Polacek, 2018).

In the years leading to Hurricane Katrina, the Cat bond market was still rather small, with annual issuance volumes hovering around USD 1-2 billion between 1998 and 2001, and concentrated, with the two largest issuers making up the lion's share of total issuances. When the traditional reinsurance market hardened as prices soared and availability dwindled as a consequence of the USD 62 billion in insured losses due to Katrina, more capital flowed into the Cat bond market as it became an attractive alternative for sponsors seeking capacity and diversification. This led to record new issuances both in 2006 (USD 4.7 billion) and in 2007 (USD 8.3 billion) (see Figure 1) and the emergence progressively less concentrated market (Braun, 2016; Mevorach, 2018; Polacek, 2018).

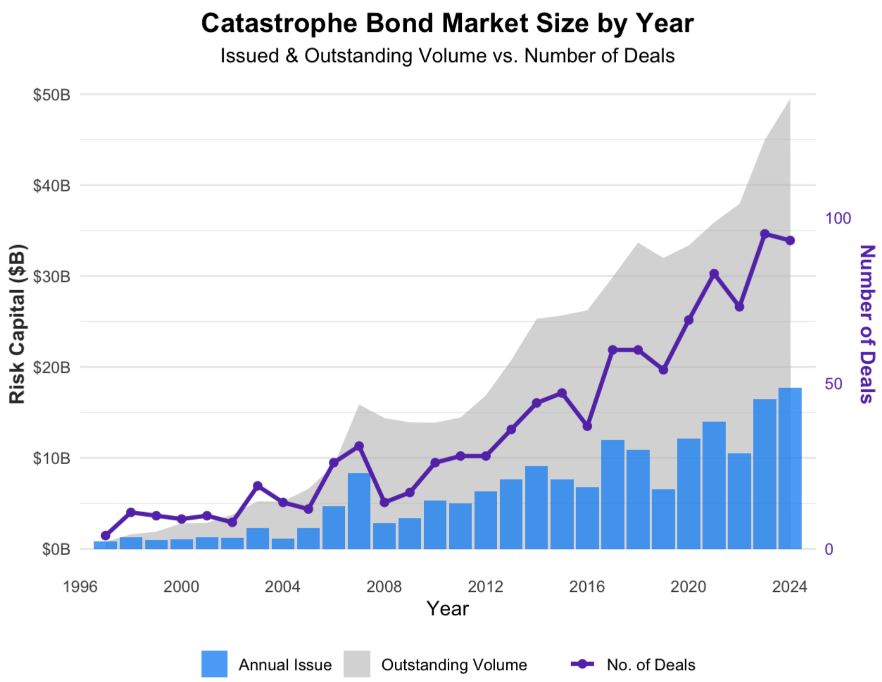


Figure 1: Catastrophe Bond Market Size (1997-2024). Source: Artemis.bm

The global financial crisis of 2008, however, caused new issuances to decline again as Lehman Brothers had served as a Total Return Swap (TRS) counterparty to collateral trusts in several major Cat bond issues (Clark et al., 2016). This event exposed structural vulnerabilities, especially concerning collateral investment practices and counterparty risk associated with structures like Total Return Swaps. Crucially, in the aftermath of the Lehman collapse, the market rapidly shifted towards increasingly safer collateral practices (Artemis.bm, 2013; Re, 2011), primarily by almost exclusively investing the proceeds directly into U.S. Treasury Money Market Funds (TMMFs) or short-term notes issued by highly-rated entities like the World Bank. This market-driven reform largely eliminated the collateral-related credit risk and was essential in rebuilding investor trust in the securitization process, leading to the development of more solid structures already by early 2009 (Braun, 2016; Polacek, 2018).

In the years following the 2008 financial crisis, the Cat bond market again saw significant growth (Figure 1). Driven by the low-interest-rate environment and the tight spreads on corporate debt, many new institutional investors were drawn to the relatively higher yields and Cat bonds' promise of uncorrelated risk exposure. This led to the value of outstanding Cat bonds more than doubling from 2010 to June 2017 (USD 23 billion), with multiple record new issuances in 2014 and 2017 (Drobetz et al., 2020; Polacek, 2018).

Lastly, the 2017 hurricane season represented a historic stress test on the Cat bond market, as Hurricanes Harvey, Irma, and Maria led to the triggering of loss events in several Cat bond tranches (estimated as up to USD 1.4 billion of loss exposure), and marking a particularly severe period for Cat bond investors (Polacek, 2018). However, despite the record losses in the second half of 2017, new Cat bond issuances in the first half of 2018 (USD 9.4 billion) did not slow down, highlighting the novel resilience of a now more established asset class (Polacek, 2018; Artemis.bm, 2018). Since then, the Cat bond market has continued to expand year after year, with record new issuances in 2020, 2021, 2023, and again setting a new annual record in 2024 (with volumes reaching USD 17.5 billion). More recently, the market saw a record first-quarter issuance in 2025 (around USD 7 billion), pushing the total outstanding market size beyond USD 52 billion (Figure 1).

2.2 Cat bonds: Structure

Cat bonds were developed to help issuing firms in transfering some of their exposure to catastrophic tail risk events (Braun, 2016; Drobetz et al., 2020; Kish, 2016). They are issued through the use of a bankruptcy-remote special purpose vehicle (SPV) to separate the legal and financial liabilities and credit risk of the issuer from the bonds themselves (Braun, 2016; Kish, 2016).

The typical structure of a Cat bond involves a few key parties and steps (Figure 2). The sponsor (typically an insurance or reinsurance firm, though government entities and corporations do also act as such), with the help of structuring agents or investment banks, establishes an SPV, usually in a jurisdiction offering favorable regulatory and tax treatment (common locations include Bermuda, the Cayman Islands, Ireland, and Singapore).

Next, the sponsor enters into a reinsurance contract with the SPV, and the SPV issues the Cat bond. The reinsurance contract defines the premium the sponsor pays the SPV (the spread, which represents the largest portion of the investor's return), the trigger mechanism specifying the exact conditions under which the reinsurance contract pays out, and the amounts that have to be paid if the trigger events occur - often defined by attachment (first loss) and exhaustion (total loss) points (Braun, 2016; Kish, 2016; Mevorach, 2018).

The terms of the Cat bond issued by the SPV mirrors the reinsurance contract between the sponsor and SPV. The principal raised from investors is deposited into a collateral account managed by the SPV and invested into highly-rated, liquid, short-term securities – predominantly US Treasury Money Market Funds or notes issued by supra-national entities such as the World Bank. Both the investment yield generated by this collateral and the risk premium paid by the sponsor (spread) are then used to pay the investor coupons (Braun, 2016; Kish, 2016).

Should a triggering event occur and meet the specified conditions, the SPV releases some or all of the funds from the collateral account to the sponsor, depending on the trigger structure and event severity relative to the points of attachment and exhaustion. In that scenario, the investor will bear losses - in part or in full - on their principal investment (Braun, 2016; Kish, 2016; Mevorach, 2018). Finally, a critical element in this regard is the bond's specific trigger mechanism. There are several primary types commonly used today, including indemnity triggers (based on the sponsor's actual losses), industry loss triggers (based on overall industry losses), parametric triggers (based on physical event characteristics like wind speed or earthquake magnitude), and modeled loss triggers (based on estimated losses from catastrophe models).

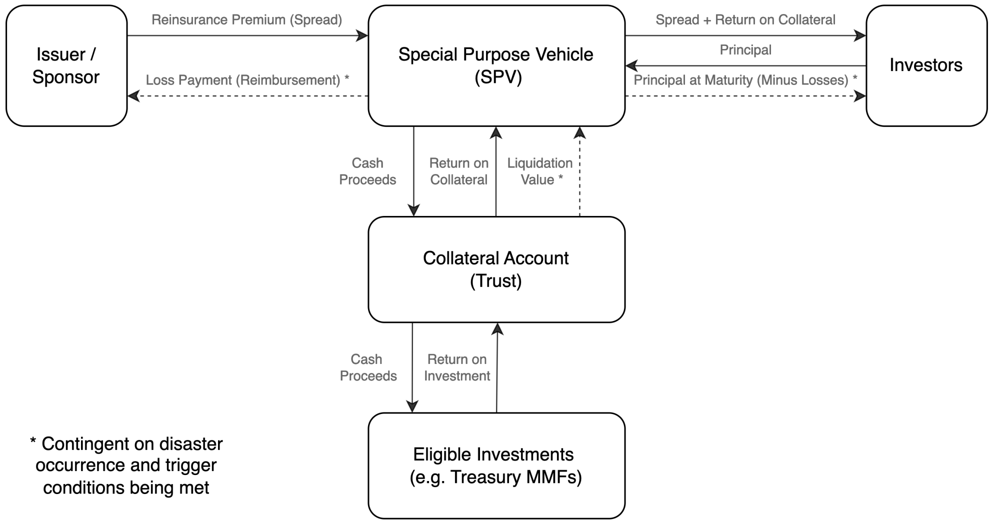


Figure 2: Cat bond structure. Source: Authors, based on Braun (2016), Kish (2016), Drobetz et al. (2020)

Cat bonds typically have maturities ranging from one to five years, with three-year being the most common term. They are issued in sizes of at least USD 100 million to cover the high transaction costs and usually pay quarterly coupons. Since these coupons are made up of the spread paid by the sponsor and the yield from the secure collateral account (typically tracking short-term rates like SOFR or T-bill yields), Cat bonds function as floating-rate instruments and are by design only exposed to the relevant natural risk they aim to cover (Kish, 2016; Mevorach, 2018). As such, the primary investment risk stems from the potential loss of some or all of the invested principal should a triggering event occur. Because of their inherent exposure to low-probability, high-impact catastrophic events, Cat bonds are typically rated as non-investment grade ('high yield') by rating agencies. Because of the SPV structure and the full collateralization (Lakdawalla and Zanjani, 2012), the rating solely reflects the modeled probability of default due to a triggering catastrophe event occurring, and is independent of the sponsor's creditworthiness. Accurate pricing thus necessitates sophisticated catastrophe risk modeling – which plays a crucial role given the sparse historical data for low-frequency, high-severity events. This modeling (performed by just a handful of highly specialized firms) quantifies the probability of attachment, probability of exhaustion, and the Expected Loss (EL), which are fundamental inputs for primary pricing, typically in fact expressed as a spread over EL (see Figure 3) (Braun, 2016; Kish, 2016;

Carayannopoulos et al., 2022).

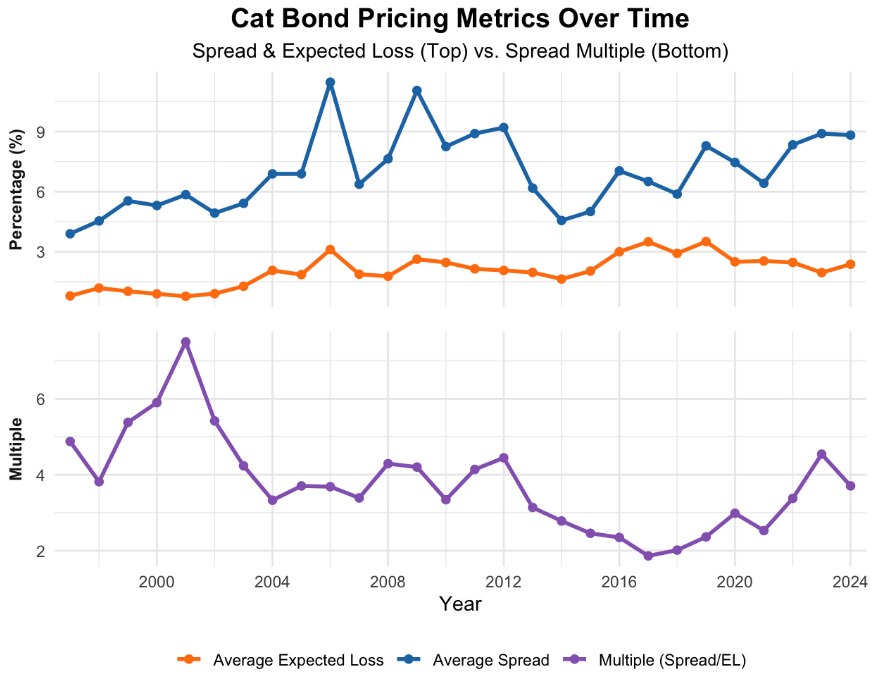


Figure 3: Catastrophe Bond Pricing Metrics: Average Spread, Average Expected Loss (EL), and Spread Multiple (1997-2024). Source: Artemis.bm

2.3 Cat bonds: SDG implications

Catastrophe bonds are increasingly discussed as potential instruments for advancing certain Sustainable Development Goals (SDGs) outlined by the United Nations, specifically SDG 11 (Sustainable Cities and Communities) and SDG 13 (Climate Action). However, their true role in sustainable development is still debated, with both potential opportunities and unresolved issues. In this sense, recent research on these instruments has focused not only on their financial characteristics, but also their suitability and effectiveness in addressing challenges tackled by the SDGs (Etzion et al., 2019; Motlagh et al., 2024; Phan and Schwartzman, 2024).

In relation to SDG 13, catastrophe bonds can enhance climate resilience by providing pre-arranged, rapidly disbursable liquidity immediately following major natural disasters, thus helping to efficiently bridge the gap between emergency relief measures and long-term reconstruction. This mechanism can be especially valuable for disaster-prone nations, and particularly sovereigns with existing fiscal constraints, offering a way to manage severe financial shocks without necessarily increasing sovereign debt, and as such significantly boost their overall resilience to natural disasters (Maran, 2024; Phan and Schwartzman, 2024). Yet, a significant debate persists regarding their effectiveness in incentivizing proactive climate risk management. In fact, one of the main points of criticism in this sense is that these instruments focus overwhelmingly on transferring financial risk after a catastrophe, rather than funding or encouraging ex-ante preventative measures like resilient infrastructure development or enhanced early warning systems (Morana and Sbrana, 2019; Motlagh et al., 2024).

Another critical component of SDG 13 involves accurately pricing climate risk to align financial incentives with sustainable climate adaptation. While the catastrophe bond market has driven advancements in catastrophe modeling, significant challenges persist. Evidence suggests that financial markets, including the Cat bond segment, struggle to consistently and accurately accurately price climate-related disaster risks, even with advancements in modeling techniques (Barnett et al., 2020; Jarzabkowski et al., 2015; Morana and Sbrana, 2019). Key difficulties include deep uncertainty surrounding future climate pathways, the complex non-linear effects of global warming on disaster frequency and intensity, existing data limitations, and the inherent issues in modeling unprecedented future conditions based on historical patterns (Morana and Sbrana, 2019; Eren et al., 2022). This broad set of issues highlights the ongoing limitations to the effectiveness of market signals in driving comprehensive climate risk management and adaptation.

Cat bonds can also contribute to SDG 11 (Sustainable Cities and Communities) by bolstering the financial resilience needed to protect communities against natural disasters. By providing efficient risk transfer capacity for sponsors (insurers, governments), they indirectly aim to reduce the economic and social shocks experienced by individuals and businesses, particularly in vulnerable regions (Etzion et al., 2019; Motlagh et al., 2024; Phan and Schwartzman, 2024). This potential is particularly relevant for developing economies often lacking deep traditional reinsurance markets, where multilateral institutions like the World Bank have actively facilitated access to cat bond financing (OECD, 2024). Although issuance is expanding, the market remains concentrated in developed nations (US, Europe, Japan) (Etzion et al., 2019; Motlagh et al., 2024; Polacek, 2018). Concerns about equity persist, however. High transaction costs, complex structuring requirements, and the need for sophisticated modeling capacity can create significant barriers for lower-income countries (Maran, 2024; OECD, 2024). Furthermore, the trigger design (especially parametric ones) can lead to a sort of basis risk, where significant real-world damage occurs but the bond doesn't pay out - as controversially seen in recent hurricane events impacting bonds for Mexico or Jamaica (Artemis.bm, 2017, 2024). Nevertheless, research suggests that OECD countries with advanced financial markets and high carbon emissions currently derive the most substantial benefits from Cat bonds, raising concerns regarding equitable distribution of these benefits (Maran, 2024). Lastly, an ongoing critique of the Cat bond market relates to investor returns. Historically, investors have achieved surplus returns as fewer Cat bonds have triggered than anticipated, raising concerns that surplus investor returns represent inefficient allocation of capital, where resources could instead have been deployed more directly toward resilience-building and proactive risk mitigation aligned with broader SDG objectives (Etzion et al., 2019).

In summary, while Cat bonds offer potential mechanisms relevant to SDGs 11 and 13, their overall effectiveness and alignment with sustainable development principles has not yet been fully explored. They can provide valuable post-disaster financial liquidity, but face significant limitations including a dominant focus on risk transfer over proactive risk reduction, challenges in accessibility and cost, the potential for basis risk to undermine protection for vulnerable sponsors, and difficulties in accurately pricing escalating climate risks. Addressing these challenges will be essential to fully leverage Cat bonds as tools for sustainable development.

3 Literature Review

This literature review builds upon the explanations laid out in the preceeding section by providing a historical overview of academic research into Cat bonds and to highlight the gap this thesis aims to fill within the existing literature.

Early foundational papers researching Cat bond pricing recognized that catastrophe risk was theoretically uncorrelated with economic cycles, suggesting Cat bonds as a potentially zero-beta investment offering portfolio diversification benefits. One of the earliest papers examining Cat bonds in more detail, by Litzenberger et al. (1996), formally assessed these securities, arguing their low correlation (derived from underlying "zero-beta" events) could improve portfolio efficiency. Other early academic pricing analyses, such as those by Canabarro et al. (2000); Lane (2000, 2002); Vaugirard (2003), used established bond pricing methods and the concept of expected excess returns (referring to the spread less the expected loss) to price Cat bonds with a probabilistic model. Others suggested different relationships to explain Cat bond spreads, such as a loglinear relationship between spread and expected loss (Major and Kreps, 2003; Mevorach, 2018) or more complex models, such as employing probability transforms (Wang, 2004). In the following years, as more Cat bonds were issued and data availability improved, pricing models began incorporating additional parameters. Some research identified cyclical patterns, including shifts in catastrophe risk pricing following major events and links to broader capital market or reinsurance cycles (Lane and Mahul, 2008).

As the Cat bond market grew following Hurricane Katrina and especially after the 2008 Global Financial Crisis (GFC), the interest in Cat bonds intensified. While some studies, like Galeotti et al. (2013), compared pricing models and potentially found limited evidence for cyclicality or market correlation in certain datasets, other research focusing specifically on the global financial crisis period reached different conclusions. Both Carayannopoulos and Perez (2015) and Gürtler et al. (2016) found a temporary but statistically significant increase in market correlation for Cat bonds during the global financial crisis. This time-varying correlation, primarily attributed to systemic factors (heightened counterparty risk concerns, widespread liquidity freezes forcing asset sales, and sharp shifts in investor risk aversion) appeared both relative to broader market indices and corporate credit spreads (Gürtler et al., 2016) as well as other asset classes

(Carayannopoulos and Perez, 2015). Crucially, however, while demonstrating cat bonds were not perfectly immune to systemic shocks, both papers still underscored their continued potential as diversifiers, as even during the high-point of the crisis the correlation remained relatively low compared to other asset classes, and risk-adjusted returns held up well during the crisis (Carayannopoulos and Perez, 2015; Gürtler et al., 2016). Finally, their evidence points to a return to low pre-crisis correlations following post-GFC structural reforms, and particularly the shift to safer collateral arrangements (Carayannopoulos and Perez, 2015).

Another major contribution to understanding primary market Cat bond pricing came from Braun (2016). Braun developed a comprehensive empirical model based on a deep dataset of bonds issued from 1997 to 2012. While confirming expected loss as the main driver, the model also identified significant pricing impacts from factors including trigger type, covered territory, sponsor prominence (an "issuer effect" possibly indicating market inefficiencies), the state of the reinsurance cycle, and broader market conditions like high-yield corporate bond spreads (Braun, 2016). Around the same time, Clark et al. (2016) was the first to test Cat bonds' out of sample portfolio performance, and provided the first evidence of superior risk-adjusted returns for strategies with Cat bonds, mainly driven by smaller standard deviations and idiosyncratic volatility.

Subsequent research explicitly tested Cat bonds' usefulness through multi-asset and multi-factor models. Drobetz et al. (2020) found Cat bonds acted as useful diversifiers against traditional asset classes and potentially as safe havens against extreme stock price drops (Drobetz et al., 2020). Other research employed various mean-variance (MV) spanning tests and different portfolio optimization strategies to test Cat bonds' diversification potential. Demers-Bélanger and Lai (2020), for instance, found evidence suggesting an increased MV efficiency frontier and an improved time-varying Sharpe ratio when including Cat bonds (Demers-Bélanger and Lai, 2020). However, drawing on Stochastic Dominance Efficiency (SDE) tests, which consider the entire return distribution rather than just mean and variance, their results were less conclusive and failed to reject the efficiency of portfolios without cat bonds, highlighting potential limitations of relying solely on MV analysis (Demers-Bélanger and Lai, 2020). Finally, Haffar and Le Fur (2022), analyzing data up to February 2020 using copula-GARCH methods and data on Cat bond fund performance, confirmed strong risk-adjusted performance and effective diversification benefits.

There are several gaps or areas for further investigation that this paper aims to address. Firstly, this paper seeks to test Cat bonds' portfolio usefulness using a broader and substantially more recent dataset that fully encompasses the COVID-19 pandemic period and the subsequent high-interest-rate, inflationary environment, as well as the intense 2022 Hurricane season. Secondly, while the existence and impact of transaction costs and liquidity premiums are acknowledged in the literature, their explicit incorporation into portfolio optimization models appears limited. This research intends to contribute by explicitly modeling estimated transaction costs in portfolio analysis to assess whether the diversification benefits persist given the substantial trading frictions currently inherent to Cat bond investing.

4 Methodology and Data

4.1 Data Acquisition, Processing, and Return Generation

4.1.1 Data

The benchmark investment universe comprises four asset class proxies broadly representing traditional portfolio allocation choices: US Equities (SPXTR, the S&P 500 Total Return Index), Global Equities (MSCI ex-US Total Return Index), US Investment Grade Corporate Bonds (FTSE US Broad IG Corporate Bond Total Return Index), and US Government Debt (ICE US Treasury 3-7 Year Total Return Index).

Our Cat bond proxy is the Swiss Re Global Cat Bond Total Return Index (SRGLTRR), which is the most comprehensive index and the most frequently cited in the academic literature (Cummins and Weiss, 2009; Braun, 2016), and which tracks the aggregate performance of all cat bonds issued under Rule 144A (which limits securities resale to qualified investors; see Swiss Re (2014)). Finally, our risk-free rate proxy is the S&P US Treasury Bill 0-3 Month TRI.

All indices are USD-denominated weekly datasets obtained from Refinitiv Eikon, and the sample period spans from March 2005 to March 2025.

Daily closing levels of the CBOE Volatility Index (VIX) are also sourced from CBOE's website.

Our t_0 estimates are manually compiled from various external data sources (see Appendix A.3.1).

4.1.2 Return Calculation

For each index series (P_t), both continuously compounded (log) returns and simple discrete returns were computed on a weekly basis. Simple discrete returns are calculated as:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{1}$$

And log returns as:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \tag{2}$$

Log returns are more useful for statistical modeling due to properties like time additivity and better distributional approximations (Tsay, 2010), while simple returns are employed to accurately calculate portfolio-level returns as weighted averages of individual asset returns (Campbell et al., 1997).

4.2 Descriptive Statistics and Time Series Diagnostics

A comprehensive diagnostic assessment was performed to understand the empirical properties of the asset returns series.

4.2.1 Distributional Properties

Standard descriptive statistics were computed for both log and simple returns, including mean, standard deviation, skewness, and excess kurtosis. Annualized mean and standard deviation were calculated as follows:

$$\mu_{log,annual} = \mu_{log,weekly} \times 52 \tag{3}$$

$$\sigma_{annual} = \sigma_{weekly} \times \sqrt{52} \tag{4}$$

Formal tests for departure from normality were conducted using the Jarque-Bera test (Jarque and Bera, 1980) and the Shapiro-Wilk test (Shapiro and Wilk, 1965), to check for the need for robust statistical techniques. Furthermore, standard risk metrics, including Maximum Drawdown, historical Value-at-Risk (VaR), and Expected Shortfall (ES) at a 95% confidence level were calculated to quantify downside risk and tail behavior (Artzner et al., 1999; Dowd, 2008).

The Maximum Drawdown (MaxDD) metric, particularly relevant for assessing tail risk for assets like Cat bonds, is calculated as:

$$\text{MaxDD} = \min_t \left(\frac{P_t - \max_{s \le t} P_s}{\max_{s \le t} P_s} \right) \tag{5}$$

where P_t represents the price or value index at time t.

4.2.2 Stationarity

Stationarity of the log return series, a fundamental assumption for many time series models, was evaluated using two complementary tests. The Augmented Dickey-Fuller (ADF) test examines the null hypothesis of a unit root (nonstationarity) against an alternative of stationarity (Dickey and Fuller, 1979), while the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test evaluates the null hypothesis of level (or trend) stationarity against the alternative of a unit root (Kwiatkowski et al., 1992). See the Appendix for the formalized hypotheses. Concordance between the tests (ADF rejection of H_0 and KPSS failure to reject H_0 at the 5% significance level) provides stronger evidence for stationarity.

4.2.3 Autocorrelation and ARCH Effects

We then investigated the temporal dependence structure of Cat bond log returns. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) were computed and plotted to identify potential ARMA structures. The Ljung-Box Q-test (Ljung and Box, 1978) was applied to the raw returns up to a lag of $\min(26, \text{floor}(N/5))$ to formally test the null hypothesis of no serial correlation:

$$Q(m) = n(n+2)\sum_{k=1}^{m}\frac{\hat{\rho}_k^2}{n-k} \tag{6}$$

where $\hat{\rho}_k$ is the sample autocorrelation at lag k, n is the number of observations, and m is the number of lags tested.

The same test was applied to the squared log returns to test for Autoregressive Conditional Heteroskedasticity (ARCH effects), indicating volatility clustering (Engle, 1982). Evidence of ARCH effects would point to the necessity to use GARCH-family models in subsequent analyses.

4.2.4 Structural Breaks

To assess parameter stability over the long sample period, particularly for the unique Cat Bond series, we tested for structural breaks in the mean log return level using the methodology of Bai and Perron (1998) and Bai and Perron (2003). This test endogenously determines the number and location of significant shifts in the intercept of a simple mean model:

$$y_t = x_t'\beta + z_t'\delta_j + u_t, \quad t = T_{j-1}+1, ..., T_j \tag{7}$$

where $j = 1, ..., m+1$ represents different regimes.

A minimum segment length (h) of 15% of the sample size was imposed to ensure stability and avoid detecting spurious breaks associated with short-term fluctuations.

4.3 Correlation Analysis

The analysis of the correlation between Cat bonds and traditional assets, and especially its stability across time and market conditions, is central to assessing their potential diversification benefits.

4.3.1 Unconditional Correlation

We compute three standard measures of unconditional association using the full sample. Pearson's product-moment correlation coefficient is used to quantify linear relationships. For potentially non-linear, monotonic associations, and as a robust alternative particularly with non-normally distributed data, rank-based correlation coefficients are also calculated, specifically Spearman's rho (ρ) and Kendall's tau (τ). Detailed definitions and formulas for the rank-based measures are provided in Appendix A.2.

4.3.2 Time-Varying Correlation Models

To understand whether asset correlations exhibit time-varying properties and state-dependency, we estimate conditional correlations using three distinct approaches:

Rolling Window Correlation Simple Pearson correlations are calculated over fixed-length rolling windows (of length $L = 26, 52, 104$ weeks). This provides a basic, model-free estimate of recent correlation.

EWMA We then employ an Exponentially Weighted Moving Average model whereby recent observations receive exponentially higher weight. The conditional covariance between assets i and j at time t is calculated recursively as:

$$\mathrm{Cov}_{i,j,t} = \lambda \cdot \mathrm{Cov}_{i,j,t-1} + (1 - \lambda) \cdot r'_{i,t-1} \cdot r'_{j,t-1} \tag{8}$$

Similarly, the conditional variances are updated as:

$$\mathrm{Var}_{i,t} = \lambda \cdot \mathrm{Var}_{i,t-1} + (1 - \lambda) \cdot (r'_{i,t-1})^2 \tag{9}$$

$$\mathrm{Var}_{j,t} = \lambda \cdot \mathrm{Var}_{j,t-1} + (1 - \lambda) \cdot (r'_{j,t-1})^2 \tag{10}$$

where $r'_{i,t-1}$ and $r'_{j,t-1}$ are the log returns at time $t - 1$ demeaned using the full-sample historical mean available at the time of calculation. We employ the standard λ value of 0.94 (Longerstaey and Spencer, 1996), which implies a relatively quick adaptation to changing market conditions. The conditional correlation is then derived as:

$$\mathrm{Corr}_{i,j,t} = \frac{\mathrm{Cov}_{i,j,t}}{\sqrt{\mathrm{Var}_{i,t} \cdot \mathrm{Var}_{j,t}}} \tag{11}$$

DCC-GARCH We estimate bivariate DCC-GARCH(1,1) models pairing Cat bonds with each benchmark asset (Engle, 2002). These models capture the time-varying conditional correlations while accounting for volatility clustering in individual return series. The implementation uses Student's t-distribution for the univariate GARCH models to accommodate fat tails in financial returns. See Appendix A.4 for the full model specification.

Standard model diagnostics, including Ljung-Box tests on the standardized residuals and squared standardized residuals, were performed to assess the adequacy of the fitted univariate GARCH models within the bivariate DCC framework.

4.3.3 Regime-Conditional Correlation

We analyze correlations separately within distinct market regimes to understand state-dependency. Regimes are defined based on:

Historical Events Predefined periods corresponding to major events, including the global financial crisis, the COVID-19 crash, as well as specific major hurricane seasons.

Table 1: Sample Period Subdivisions

Period Description	Start	End
Pre-GFC Period	Jan 2005	Aug 2007
Global Financial Crisis	Sep 2007	Mar 2009
Post-GFC Recovery	Apr 2009	Aug 2017
2017 Hurricane Season	Aug 2017	Oct 2017
Inter-Crisis Period	Nov 2017	Jan 2020
COVID-19 Impact	Feb 2020	Jun 2020
Post-COVID	Jul 2020	Sep 2022
Hurricane Ian Impact	Sep 2022	Dec 2022
Recent Period	Dec 2022	Mar 2025

Market Stress (VIX) We employ a proxy threshold-based model, whereby weekly observations are classified into *Normal* or *Crisis* regimes based on the VIX index weekly closing level. The threshold value of 25 is a commonly used heuristic as well as standard practice in the literature (Whaley, 2009; Bali and Hovakimian, 2009), and it approximately corresponds to the 85th percentile of the distribution. Values above this threshold typically correspond to periods characterized by significant market uncertainty and risk aversion.

$$\text{Regime}_t = \begin{cases} \text{Crisis}, & \text{if VIX}_t \geq 25 \\ \text{Normal}, & \text{if VIX}_t < 25 \end{cases} \tag{12}$$

4.3.4 Testing Correlation Differences

To statistically assess whether correlations differ significantly between regimes or periods, we utilize a non-parametric bootstrap approach. To compare two subperiods (e.g., Crisis vs. Normal VIX regime, or Pre-GFC vs. GFC), we repeatedly resample (R=1000) with replacement from the combined data corresponding to those two groups. Within each bootstrap sample, we calculate the Pearson correlation coefficient for each group (r_1^*, r_2^*) and compute their difference ($d^* = r_1^* - r_2^*$). The empirical distribution of these differences (d^*) across all resamples allows the construction of a 95% percentile confidence interval for the true difference in correlations. If this confidence interval excludes zero, we reject the null hypothesis of equal correlations between the two groups at the 5% significance level:

$$H_0 : \rho_{\text{period}_1} - \rho_{\text{period}_2} = 0 \tag{13}$$

This method is preferred over the traditional Fisher Z-test due to its robustness to the non-normality often encountered in financial returns and its validity even with smaller sample sizes (Efron and Tibshirani, 1994). Bootstrap confidence intervals are also generated for the within-period correlation estimates themselves.

4.4 Asset Spanning Tests

To formally test whether Cat bonds provide statistically significant diversification benefits relative to the benchmark asset portfolio, we employ mean-variance asset spanning tests. This involves estimating the regression:

$$r_{\text{catbonds},t} = \alpha + \sum_{i=1}^{N} \beta_i r_{\text{benchmark},i,t} + \epsilon_t \tag{14}$$

where r represents weekly log returns. Spanning implies that the benchmark assets can perfectly replicate the risk-return profile of the test asset. Following Kan and Zhou (2012), who highlighted the low power of the traditional Huberman and Kandel (1987) joint test, we focus on the individual hypotheses:

$$H_0 : \alpha = 0 \tag{15}$$

Testing this hypothesis assesses whether Cat bonds offer statistically significant abnormal returns (mean-variance improvement) after controlling for their systematic exposures to the benchmark assets. Rejection implies diversification gains via enhanced returns or alpha generation.

$$H_0 : \sum_{i=1}^{N} \beta_i = 1 \tag{16}$$

This tests whether the minimum-variance portfolio constructed solely from benchmark assets can achieve the same variance as a portfolio including the test asset for certain levels of expected return. More intuitively, it tests if the benchmark portfolio that best mimics the test asset's return requires full investment. Rejection implies the test asset helps shift the efficient frontier inwards (variance reduction) due to a distinct risk profile (DeRoon and Nijman, 2001).

These hypotheses are assessed using Wald tests performed on regression estimates. Given the evidence of heteroskedasticity (tested via Breusch-Pagan), the Wald tests employ robust standard errors computed using the HC3 variant of the HC covariance matrix estimator (for its superior performance in finite samples compared to HC0 or HC1 (MacKinnon and White, 1985; Zeileis, 2004)). The spanning tests are first conducted over the full sample period and then separately within the VIX-defined Normal and Crisis regimes to examine the stability of potential diversification benefits. A test for the equality of alpha across regimes is performed using an interaction model.

4.5 Out-of-Sample Portfolio Backtesting Framework

To evaluate the potential benefits of incorporating Cat bonds into investment portfolios, we conduct a comprehensive out-of-sample (OOS) backtesting simulation. This approach aims to assess strategy performance under conditions that mitigate in-sample overfitting and look-ahead bias, providing a more realistic evaluation than purely in-sample analyses (White, 2000; Hansen, 2005).

4.5.1 OOS Design and Estimation Windows

The backtesting procedure employs a differentiated rolling-origin framework. An initial in-sample period from March 2005 to March 2007 (2 years) is used exclusively for the first estimation cycle. Subsequently, at each rebalancing decision point t (start of weeks $i = 1, 1 + R, 1 + 2R, ...$, where R is the rebalancing frequency in weeks), the required inputs (expected returns, covariance matrix) are estimated using only data available up to the end of week $t - 1$. The portfolio weights determined at time t are then held constant and applied to the realized returns from week t through $t + R - 1$.

Covariance Estimation Window The covariance matrix ($\hat{\Sigma}_t$) is estimated using a fixed-length rolling window of $L = 104$ weeks (2 years) immediately preceding t. This window length is chosen as a common compromise between capturing recent changes in market dynamics and maintaining estimation stability (Fleming et al., 2001).

Mean Estimation Window Expected returns ($\hat{\mu}_t$) are estimated using an expanding window approach, incorporating all available data from the beginning of the sample up to $t - 1$. This is implemented via a blended estimation technique described below, where the sample component uses the full available history (n_{obs} increases over time). The use of an expanding window for mean estimation reflects the assumption that long-run average returns are more stable than short-term volatility and correlation structures.

4.5.2 Rebalancing Frequencies

The simulation is performed independently for three distinct rebalancing frequencies (R): Quarterly ($R = 13$ weeks), Monthly ($R = 4$ weeks), and Weekly ($R = 1$ week). Comparing these frequencies allows for an assessment of the practical trade-off between potentially capturing shorter-term signals (higher frequency) and incurring higher transaction costs and turnover (DeMiguel et al., 2009).

4.5.3 Input Parameter Estimation

Expected Returns We employ a blended estimation approach, shrinking the historical sample mean towards a predefined prior estimate. The blended estimate for asset i at time t is:

$$\hat{\mu}_{log,i,t} = \alpha_t \cdot \mu_{log,sample,i,t} + (1 - \alpha_t) \cdot \mu_{log,prior,i} \tag{17}$$

where $\mu_{log,sample,i,t}$ is the historical mean log return from the start of the data up to $t - 1$, $\mu_{log,prior,i}$ is the fixed prior mean log return for asset i (see Appendix

A.3.1), and $\alpha_t = \frac{n_{obs,t}}{n_{obs,t}+k}$ is the time-varying shrinkage intensity. $n_{obs,t}$ is the number of observations available at time t. The parameter $k = 208$ (equivalent to 4 years of weekly data) determines the weight given to the prior, consistent with Bayesian shrinkage principles aimed at reducing the impact of estimation error in sample means computed over short periods (Jorion, 1986; James and Stein, 1961). Expected simple returns for the Sharpe Ratio optimization are derived via $\exp(\hat{\mu}_{\log,t}) - 1$.

Covariance Matrix The procedure is as follows: we first calculate the L-week rolling sample covariance matrix of log returns (S). To enhance stability and ensure positive semi-definiteness, particularly when the number of variables is large relative to the number of observations, we apply a shrinkage estimation methodology derived from (Schäfer and Strimmer, 2005; Opgen-Rhein and Strimmer, 2007). This approach involves two separate shrinkage steps:

1. The sample variances (diagonal elements of S) are shrunk towards their median variance using an analytically determined shrinkage intensity λ^*_{var} (Opgen-Rhein and Strimmer, 2007).

2. The sample correlation matrix derived from S is shrunk towards the identity matrix using a separate, analytically determined shrinkage intensity λ^* (Schäfer and Strimmer, 2005).

The final shrinkage covariance estimate ($\hat{\Sigma}_{\text{Shrink}}$) is then reconstructed from the shrunken variances and the shrunken correlation matrix. This method avoids the need to specify a single, potentially suboptimal, target matrix F as in the original Ledoit-Wolf formulation (Ledoit and Wolf, 2004a,b) while retaining guaranteed positive definiteness and improved efficiency over the sample covariance matrix, especially for high-dimensional data (Schäfer and Strimmer, 2005).

4.5.4 Portfolio Optimization Strategies

We evaluate whether the addition of Cat bonds can improve portfolio performance by testing the out of sample performance, with and without Cat bonds, of four distinct allocation strategies, representing different investor objectives and sensitivities to input parameters:

Equal Weight (EQ) The naive $1/N$ benchmark, requiring no estimation or optimization, represents a simple, robust baseline. Weights are $w_i = 1/N$.

Minimum Variance (MV) Solves the quadratic program:

$$\begin{aligned}
\min_{w} \quad & \frac{1}{2} w' \hat{\Sigma}_t w \\
\text{subject to} \quad & w' \mathbf{1} = 1 \\
& w_i \geq 0, \quad \forall i \\
& w_{\text{cat}} \leq \text{cap} \quad \text{(if applicable)}
\end{aligned} \tag{18}$$

This strategy ignores expected returns entirely, focusing solely on minimizing portfolio variance based on the estimated covariance matrix. It is often considered more robust to estimation error than mean-variance efficient strategies (Jagannathan and Ma, 2003; Clarke et al., 2006).

Maximum Sharpe Ratio (SR) Solves the optimization problem:

$$\begin{aligned}
\max_{w} \quad & \frac{w' \hat{\mu}_{\text{excess},t}}{\sqrt{w' \hat{\Sigma}_t w}} \\
\text{subject to} \quad & w' \mathbf{1} = 1 \\
& w_i \geq 0, \quad \forall i \\
& w_{\text{cat}} \leq \text{cap} \quad \text{(if applicable)}
\end{aligned} \tag{19}$$

Here $\hat{\mu}_{\text{excess},t}$ are the estimated expected simple returns in excess of the risk-free rate at time t. This represents the tangency portfolio in standard mean-variance theory (Sharpe, 1966). Due to the objective function structure and its high sensitivity to mean estimates, it is solved using non-linear numerical optimization incorporating constraints via normalization and adjustment steps.

Maximum Diversification Ratio (DR) Aims to maximize portfolio diversification by maximizing the ratio of the weighted average of individual asset volatilities to the portfolio volatility (Choueifaty and Coignard, 2008):

$$\begin{aligned}
\max_{w} \quad & \frac{w' \sigma_{\text{diag},t}}{\sqrt{w' \hat{\Sigma}_t w}} \\
\text{subject to} \quad & w' \mathbf{1} = 1 \\
& w_i \geq 0, \quad \forall i \\
& w_{\text{cat}} \leq \text{cap} \quad \text{(if applicable)}
\end{aligned} \tag{20}$$

$\sigma_{\text{diag},t}$ is the vector of asset volatilities derived from the diagonal of $\hat{\Sigma}_t$. This approach seeks broad diversification across risk sources without relying on expected return estimates.

To handle potential numerical issues during optimization, a fallback mechanism is implemented: if an optimizer fails to find a valid portfolio satisfying all constraints at a given rebalancing date, the portfolio weights for that period are set to the Equal Weight (1/N) allocation.

4.5.5 Constraints

All optimized portfolios enforce non-negativity ($w_i \geq 0$) reflecting typical long-only mandates, and full investment ($\sum w_i = 1$). Variants imposing a $w_{\text{cat}} \leq 0.25$ constraint are implemented to assess the impact of limiting exposure to the alternative asset, representing a more realistic allocation option.

4.5.6 Transaction Costs

The impact of trading frictions is modeled using proportional transaction costs applied ex-post at each rebalancing date t. The cost is calculated as:

$$TC_t = \sum_{i=1}^{N} c_i \cdot |w_{i,t} - w_{i,t-R}^{+}| \tag{21}$$

where c_i is the asset-specific proportional cost rate, $w_{i,t}$ is the target weight for asset i determined at rebalance time t, and $w_{i,t-R}^{+}$ is the weight held *after* the previous rebalance at time $t-R$. This model captures the cost associated with portfolio turnover (Balduzzi, 1999; DeMiguel et al., 2009). These transaction cost figures are representative estimates, compiled from various studies and market analyses, and should be taken as rates approximately reflecting efficient institutional execution, with a particular emphasis on keeping the relative costs between asset classes consistent and realistic rather than providing exact individual numbers.

Table 2: Estimated One-Way Transaction Cost Estimates by Asset Class. Source: Authors' estimates.

Asset Class	Cost (bps)
US Treasuries	10
US Equities	15
Global Equities	25
US Corporate Bonds	30
Catastrophe Bonds	90

All tests are also run without transaction costs to provide a frictionless benchmark.

4.6 Performance Evaluation and Statistical Comparison

4.6.1 Performance Metrics

A suite of standard performance and risk metrics is calculated on the weekly OOS portfolio return series: Annualized Geometric Mean Return, Annualized Standard Deviation, Annualized Sharpe Ratio, Sortino Ratio, Maximum Drawdown, 95% historical Value-at-Risk (VaR), and 95% historical Expected Shortfall (ES). Portfolio Turnover, measuring trading activity, is also calculated and annualized. These metrics provide a comprehensive view of realized performance and risk (Bacon, 2012).

The Sortino ratio, which considers only downside volatility in assessing risk-adjusted returns, is calculated as:

$$\text{Sortino} = \frac{r_p - r_f}{\sigma_d} \tag{22}$$

where r_p is the portfolio return, r_f is the risk-free rate, and σ_d is the downside deviation defined as:

$$\sigma_d = \sqrt{\frac{1}{N}\sum_{t=1}^{N} \min(r_t - MAR, 0)^2} \tag{23}$$

with MAR (Minimum Acceptable Return) set to the risk-free rate.

4.6.2 Statistical Robustness and Comparison

To assess the reliability of observed performance differences, statistical tests and confidence intervals are employed as follows:

Portfolio Return Diagnostics The out of sample (OOS) return series for each strategy and frequency combination are examined for normality using Jarque-Bera and Anderson-Darling tests, and for ARCH effects using the Ljung-Box test on squared returns, to better understand their distributional properties.

Sharpe Ratio Difference Test The statistical significance of differences in risk-adjusted performance between key strategy pairs (e.g., with vs. without Cat bonds) is formally tested. The implementation uses the asymptotic test derived by Ledoit and Wolf (2008). This test operates on the weekly excess returns, evaluating the null hypothesis of equal weekly Sharpe Ratios, and its advantage lies in its robustness to the serial correlation and non-normality often observed in financial returns. The formal description of the test can be found in the Appendix A.5.

Bootstrap Confidence Intervals To quantify the estimation uncertainty surrounding the calculated performance metrics, 95% confidence intervals are generated using the stationary block (length $l = 15$) bootstrap applied to the weekly out of sample portfolio return series (Künsch, 1989; Politis and Romano, 1994). The non-parametric bootstrap approach employs 1000 resamples to generate a stable empirical distribution for each performance metric. Percentile confidence intervals are reported for the main metrics.

5 Results

This section presents the empirical findings derived from the tests outlined in Section 4. The analysis covers the set from March 2005 to March 2025. First, we examine the distributional properties (Section 5.1) and time series characteristics (Section 5.2) of Cat bonds compared to traditional asset classes. We then analyze correlation structures over the full sample as well as across different market regime subsamples (Section 5.3). Evidence from formal asset spanning tests is then presented (Section 5.4), and finally we report the results of the Out-of-Sample portfolio backtesting (Section 5.5).

5.1 Distributional Properties and Risk-Return Profile

Table 3 presents the key descriptive statistics for the weekly log returns of all assets. Cat Bonds exhibit an annualized mean return of 7.05%, higher than Global Equities (5.53%), US Corporate Bonds (4.13%), and US Treasuries (2.91%), but still lower than US Equities (9.77%). The annualized volatility of Cat Bonds (5.52%) is substantially lower than equity markets (Global Equities: 18.90%, US Equities: 17.68%), marginally lower than US Corporate Bonds (6.12%), and higher than US Treasuries (3.80%). This combination results in a superior Sharpe ratio for Cat bonds (0.988), exceeding all other assets in the sample (Table 3).

Table 3: Risk-Return Statistics of Weekly Log Returns

Asset	Ann. Mean	Ann. SD	Ann. SR	Max DD	ES (95%)
CatBonds	0.0705	0.0552	0.9880	0.0851	-0.0125
GlobalEquities	0.0553	0.1890	0.2082	0.5956	-0.0661
USCorpBonds	0.0413	0.0612	0.4145	0.1533	-0.0196
USEquities	0.0977	0.1768	0.4623	0.6041	-0.0597
USTreasuries	0.0291	0.0380	0.3463	0.0919	-0.0114

Table 4: Risk-Return Statistics (cont.)

Asset	Skewness	Exc. Kurtosis	MDD Date	Jarque-Bera p
CatBonds	-9.667	284.261	Sep 2017	<0.001
GlobalEquities	-1.524	12.614	Oct 2008	<0.001
USCorpBonds	-1.799	17.338	Mar 2020	<0.001
USEquities	-0.943	8.534	Oct 2008	<0.001
USTreasuries	-0.094	1.339	Jun 2008	<0.001

Table 4 summarizes further distributional statistics for the returns. Cat bonds exhibit extreme negative skewness (-9.67) and excess kurtosis (284.26), substantially exceeding the non-normality observed in other assets such as Global Equities (Skewness: -1.52, Exc. Kurtosis: 12.61). This pronounced negative skewness and leptokurtosis reflect the nature of Cat Bonds, characterized by generally stable returns punctuated by occasional large losses during catastrophe events.

The distinct risk profile of Cat bonds extends to other risk measures like maximum drawdown, which for Cat bonds occurred during September 2017, coinciding with Hurricanes Harvey, Irma, and Maria. This drawdown was notably less severe than the drawdowns experienced by Global Equities (59.56%) and US Equities (60.41%) during the Global Financial Crisis. Furthermore, the 95% Expected Shortfall for Cat bonds (-1.25%) also indicates more contained tail risk than equities (Global: -6.61%, US: -5.97%).

Finally, formal tests strongly reject normality for all asset returns. Both Shapiro-Wilk and Jarque-Bera tests yield $p < 0.001$ across all series, supporting the use of non-parametric and robust methods throughout the analysis.

5.2 Time Series Characteristics

Table 5 summarizes the results of stationarity tests on the log return series. The Augmented Dickey-Fuller (ADF) test rejects the null hypothesis of a unit root for all return series ($p = 0.01$), suggesting stationarity. The KPSS test fails to reject the null of stationarity for Cat bonds, Global Equities, US Corporate Bonds, and US Equities at the 5% significance level. However, for US Treasuries, the KPSS test statistic (0.502) exceeds the 5% critical value (0.463), failing to confirm stationarity for this series based on the KPSS test (Table 5).

Table 5: Stationarity Test Results (Log Returns)

Asset	ADF p-value	KPSS Statistic	KPSS Crit (5%)	Stationary
CatBonds	0.01	0.178	0.463	TRUE
GlobalEquities	0.01	0.038	0.463	TRUE
USCorpBonds	0.01	0.152	0.463	TRUE
USEquities	0.01	0.155	0.463	TRUE
USTreasuries	0.01	0.502	0.463	FALSE

Analysis of the Autocorrelation Function (ACF) for Cat Bond log returns reveals significant temporal dependence, with a negative spike at lag 1 (-0.15) and a positive spike at lag 2 (0.10). This pattern suggests possible ARMA structure in Cat bond returns, consistent with lower liquidity and potential staleness in Cat Bond pricing compared to more traditional, highly liquid assets.

Finally, the Bai & Perron test for structural breaks in the mean of Cat bond log returns detected no statistically significant structural breaks over the full sample period using a minimum segment length of 15% of the sample size. This suggests relative stability in the long-term mean return process.

5.3 Correlation Analysis

5.3.1 Unconditional Correlation

Table 6 presents standard unconditional correlation coefficients calculated over the full sample period. The Pearson correlation coefficient shows that Cat bonds exhibit low positive linear correlations with Global Equities (0.062), US Corporate Bonds (0.085), and US Equities (0.076), while the correlation with US Treasuries is essentially zero (0.002). Rank-based measures (Spearman: 0.039, Kendall: 0.027 for Global Equities), which are less sensitive to outliers and non-normality, confirm this low degree of association (Table 6). The consistently small correlation values across all three measures provide strong preliminary evidence for the diversification potential of Cat bonds.

Table 6: Unconditional Correlation Coefficients with Cat Bonds

Asset	Pearson	Spearman	Kendall
GlobalEquities	0.062	0.039	0.027
USCorpBonds	0.085	0.062	0.042
USEquities	0.076	0.030	0.020
USTreasuries	0.002	0.032	0.022

5.3.2 Time-Varying Correlation

The practical utility of Cat bonds as diversifiers depends critically on whether their low average correlation with benchmark assets remains stable across different market conditions. As such, dynamic correlation estimates are analyzed for evidence of instability or state-dependency. Table 7 summarizes statistics for 52-week rolling correlations and the mean DCC correlation. While the time variation is substantial (see e.g., dynamic correlation plots in Appendix B) and, for instance, the 52-week rolling correlation between Cat bonds and Global Equities ranges from -0.494 to 0.579, with a standard deviation of 0.171, the average correlations from these dynamic models (e.g., 52w Rolling Mean: 0.044, DCC Mean: 0.030 for Global Equities) remain low and broadly consistent with the unconditional estimates.

Table 7: Summary Statistics of Time-Varying Corr. Models with Cat Bonds

	Rolling Window (52-week)				DCC-GARCH
Asset	Mean	Min	Max	Std. Dev.	Mean
GlobalEquities	0.044	-0.494	0.579	0.171	0.030
USCorpBonds	0.099	-0.262	0.697	0.176	0.085
USEquities	0.042	-0.384	0.526	0.169	0.039
USTreasuries	0.038	-0.397	0.486	0.155	0.014

5.3.3 Regime-Conditional Correlation

To understand how correlations behave under different market conditions, we analyze them separately during *Normal* and *Crisis* (VIX $\geq$ 25) periods. The former accounts for approximately 83.5% of the total observations (871), while the latter regime accounts for approx. 16.5% of observations (172).

Table 8 highlights the performance differences between regimes. During Crisis periods, traditional assets experience significant returns deterioration: Global Equities and US Equities exhibit large negative annualized returns (-66.9% and -56.9%, respectively) with elevated volatility (34.0% and 31.9%). In contrast, Cat bonds maintain positive mean returns (2.8%) with relatively contained volatility (6.3%), resulting in a positive Sharpe ratio (0.45). Unsurprisingly, US Treasuries, as traditional safe-haven assets, also perform well during crises (return: 10.0%, Sharpe: 2.19).

Table 9 presents the average correlation with Cat Bonds within each regime and tests the significance of the difference using the non-parametric bootstrap described in Section 4.3.4. For all traditional asset classes, the point estimate for correlation with Cat bonds is higher during Crisis periods compared to Normal periods. However, this increase is only statistically significant at the 5% level for US Corporate Bonds (Difference = 0.283, 95% CI: [0.125, 0.569]). For Global Equities (Difference = 0.062, 95% CI: [-0.117, 0.378]), US Equities (Difference = 0.037, 95% CI: [-0.185, 0.290]), and US Treasuries (Difference = 0.130, 95% CI: [-0.019, 0.236]), the observed increases in correlation during crisis periods are not statistically distinguishable from zero (Table 9). This suggests Cat bonds largely

Table 8: Asset Performance Statistics by VIX Market Regime

Asset	Regime	Ann. Mean	Ann. Vol.	Sharpe Ratio	Mean (N-C)	Sharpe (N-C)
CatBonds	Crisis	0.028	0.063	0.445	–	–
	Normal	0.079	0.053	1.477	0.051	1.032
Global Equities	Crisis	−0.669	0.340	−1.969	–	–
	Normal	0.198	0.133	1.490	0.867	3.459
US Corp Bonds	Crisis	0.027	0.106	0.256	–	–
	Normal	0.044	0.047	0.929	0.017	0.673
US Equities	Crisis	−0.569	0.319	−1.781	–	–
	Normal	0.229	0.124	1.850	0.798	3.631
US Treasuries	Crisis	0.100	0.046	2.189	–	–
	Normal	0.015	0.036	0.421	−0.085	−1.768

retain their low correlation properties even during market turmoil, particularly relative to equities.

Table 9: Cat Bonds Correlation by VIX Regime (Boot. CI Test, Signif. at 5 %)

Asset	Crisis Corr.	Normal Corr.	Diff.	95% Boot CI	Sign.
GlobalEquities	0.095	0.032	0.062	[-0.117, 0.378]	No
USCorpBonds	0.274	-0.009	0.283	[0.125, 0.569]	Yes
USEquities	0.095	0.058	0.037	[-0.185, 0.290]	No
USTreasuries	0.108	-0.022	0.130	[-0.019, 0.236]	No

Detailed results comparing Cat bond correlations between specific historical periods with bootstrap CIs are available in Appendix B, Table 17.

5.4 Asset Spanning Tests

We test if benchmark assets can replicate the risk-return profile of Cat bonds using Wald tests with HC3 robust standard errors. The spanning regression alpha estimates (Table 10) indicate whether Cat bonds offer significant returns not explained by the benchmarks. The annualized geometric alpha is positive for the full sample (6.98%), and in both the Normal (7.63%) and Crisis (2.05%) regimes, indicating persistent diversification potential not captured by the benchmark assets.

Table 11 shows the results for the full sample period. Both individual hypotheses (H_0: $\alpha = 0$ and H_0: $\sum \beta_i = 1$) and the joint hypothesis are strongly rejected ($p < 0.001$). The rejection of H_0: $\alpha = 0$ (ChiSq = 27.02) indicates that

Table 10: Weekly vs. Ann. Geom. α Estimates & Regime Difference

Period	Weekly α	Ann. α
Full Sample	0.001297	0.0698
Crisis Regime	0.000390	0.0205
Normal Regime	0.001415	0.0763
H0: $\alpha_{\text{Crisis}} = \alpha_{\text{Normal}}$		
	ChiSq(1)	P-Value
(Fail to Reject)	1.45	0.229

Cat bonds offer significant abnormal returns (alpha) not explained by exposures to the benchmark assets over the entire period. The rejection of H_0: $\sum \beta_i = 1$ (Est: 0.021, ChiSq = 416.9) implies that the risk profile of Cat Bonds is distinct and cannot be mimicked by a portfolio of benchmark assets requiring full investment, suggesting variance reduction benefits. Together, these facts provide compelling evidence that Cat bonds expand the investment opportunity set (Table 11).

Table 11: Asset Spanning Tests - Full Sample

Hypothesis	ChiSq Stat	P-Value	Conclusion (5% level)
H0: $\alpha = 0$	27.020	<0.001	Reject H0
H0: $\sum \beta_i = 1$ (Est: 0.021)	416.908	<0.001	Reject H0
H0: Joint	425.146	<0.001	Reject H0

Tables 12 and 13 report results from spanning tests conducted separately for Crisis and Normal regimes, revealing differing sources of diversification benefits. In the Normal regime (Table 12), the joint hypothesis is strongly rejected ($p < 0.001$). Notably, the alpha coefficient is statistically significant (ChiSq = 29.43, p = <0.001), suggesting significant return enhancement relative to benchmarks during normal market conditions. The hypothesis $\sum \beta_i = 1$ (Est: 0.010) is also strongly rejected (ChiSq = 328.9, $p < 0.001$), indicating variance reduction benefits during these periods as well.

Table 12: Asset Spanning Tests - VIX Normal Regime

Hypothesis	ChiSq Stat	P-Value	Conclusion (5% level)
H0: $\alpha = 0$	29.433	<0.001	Reject H0
H0: $\sum \beta_i = 1$ (Est: 0.010)	328.931	<0.001	Reject H0
H0: Joint	367.663	<0.001	Reject H0

Conversely, during the Crisis regime (Table 13), while the joint hypothesis remains rejected ($p < 0.001$) and the variance reduction benefit persists (H_0:

$\sum \beta_i = 1$ rejected, Est: 0.145, ChiSq = 68.4, p< 0.001), the alpha coefficient is no longer statistically significant (ChiSq = 0.23, p = 0.631). This suggests that while Cat bonds still help reduce portfolio variance during crises, the evidence for statistically significant return enhancement (alpha) specifically during these high-stress periods is not supported by these results.

Table 13: Asset Spanning Tests - VIX Crisis Regime

Hypothesis	ChiSq Stat	P-Value	Conclusion (5% level)
H0: $\alpha = 0$	0.231	0.631	Fail to Reject H0
H0: $\sum \beta_i = 1$ (Est: 0.145)	68.401	$<$0.001	Reject H0
H0: Joint	91.164	$<$0.001	Reject H0

Finally, the $\sum \beta_i$ estimates (Full Sample: 0.021, Crisis: 0.145, Normal: 0.010) are extremely low, indicating minimal systematic risk relative to the benchmarks. While marginally higher in crises, the exposure remains minimal across all periods, highlighting strong diversification potential stemming from factors other than traditional market movements.

5.5 Portfolio Backtesting

To assess the practical implications of including Cat bonds, we conduct out-of-sample portfolio backtesting simulations using various allocation strategies (Equal Weight – EQ, Minimum Variance – MV, Maximum Sharpe Ratio – SR, Maximum Diversification Ratio – DR) and rebalancing frequencies. Table 14 presents key performance metrics for quarterly rebalanced portfolios, incorporating estimated transaction costs. Comprehensive results for all frequencies, with and without transaction costs, are available in Appendix B (Tables 20 and 21).

Across all strategy types, portfolios including Cat bonds generally exhibit improved risk-return profiles compared to their noCat counterparts, both with and without transaction costs. As shown in Table 14 for quarterly rebalancing with transaction costs, this improvement is often characterized by similar or higher annualized returns coupled with lower annualized volatility, leading to consistently higher point estimates for risk-adjusted performance measures (Sharpe and Sortino ratios) for the wCat variants. For instance, the SR wCat strategy achieves an annualized Sharpe ratio of 0.907 (wTC), compared to 0.466 for SR noCat. Even with a 25% allocation cap (SR wCat25), the portfolio achieves a Sharpe ratio of 0.728 (wTC), substantially outperforming the SR noCat portfolio based on these point estimates. The risk-return scatter plots (e.g., Figure 8a and Figure 8b in

Appendix B) visually confirm this general northwest shift when Cat bonds are included.

Table 14: Performance Metrics for Quarterly Rebalanced Portfolios (With Transaction Costs)

Portfolio	Ann. Return	Ann. Vol	Sharpe	Sortino	Max DD
EQ noCat	0.060	0.094	0.522	0.100	0.310
EQ wCat	0.063	0.077	0.661	0.127	0.251
SR noCat	0.045	0.071	0.466	0.091	0.182
SR wCat	0.059	0.050	0.907	0.165	0.103
SR wCat25	0.052	0.053	0.728	0.142	0.134
DR noCat	0.043	0.053	0.578	0.112	0.165
DR wCat	0.052	0.041	0.947	0.176	0.118
DR wCat25	0.049	0.043	0.846	0.162	0.139
MV noCat	0.032	0.042	0.455	0.087	0.150
MV wCat	0.048	0.042	0.815	0.144	0.099

Maximum drawdowns are also consistently reduced when Cat bonds are included (Table 14). For instance, comparing quarterly rebalanced strategies with transaction costs, the EW portfolio's maximum drawdown decreases from 31.0% without Cat Bonds to 25.1% with their inclusion. The reduction is even more pronounced for optimized strategies, with the SR wCat portfolio exhibiting a maximum drawdown of just 10.3% compared to 18.2% for SR noCat. This improved downside protection is a key practical benefit. Cumulative performance plots (Figures 9a and 9b in Appendix B) visually confirm these improvements.

Moreover, the choice of rebalancing frequency impacts net performance due to transaction costs. Table 15 compares the Sharpe ratios for the SR wCat strategy under different frequencies, both before and after transaction costs. Without costs, more frequent rebalancing yields marginally better performance. However, when transaction costs are considered, the relationship predictably inverts: quarterly rebalancing produces the highest net Sharpe ratio (0.930), illustrating the practical trade-off between capturing short-term signals and minimizing trading frictions. The same relationship exists for all optimized portfolio strategies.

To assess statistical reliability, we test for significant differences in Sharpe ratios using the Ledoit and Wolf (2008) methodology. The improvement in Sharpe ratio when adding Cat Bonds is no longer statistically significant at the conventional 5% level for any of the strategies tested under these conditions (p-values range from 0.123 to 0.611). However, when examining the results without transaction

Table 15: Impact of Rebalancing Frequency on SR wCat Sharpe Ratio

Rebalancing Frequency	SR (noTC)	SR (wTC)
Weekly	1.034	0.614
Monthly	1.023	0.878
Quarterly	0.992	0.907

costs (see Appendix Table 18), the improvement for the SR strategy becomes statistically significant (p=0.034; monthly rebalancing). While the point estimates consistently favor Cat bond inclusion (Table 14), formal statistical validation of the benefit is weak after accounting for costs.

Table 16: Tests for Sharpe Ratio Differences (Quarterly Rebalancing, wTC)

Strategy	Sharpe (noCat)	Sharpe (wCat)	Difference	p-value (LW Test)
SR	0.466	0.907	0.441	0.123
DR	0.578	0.947	0.369	0.185
MV	0.455	0.815	0.360	0.277
EQ	0.522	0.661	0.139	0.611

Analysis of portfolio composition over time reveals the significant impact of the inclusion of catastrophe bonds in optimized portfolios. When available without constraints, Cat bonds consistently command a substantial allocation share across optimization objectives. As illustrated for selected strategies (monthly rebalancing, wTC) in Appendix Figure 10, this weight for the SR wCat strategy often ranges between 25% and 50% (Appendix Figure 10a). The inclusion of Cat bonds primarily displaces allocations to riskier assets like global and US equities, and to some extent corporate bonds, particularly evident in the SR and DR strategies (compare plots (a) vs (b) and related DR plots in Appendix Figure 10). In addition, the allocation to Cat bonds often increases dynamically during periods of heightened market stress, acting in some instances as a potential safe haven or diversifier. This behaviour was particularly pronounced during the GFC (2008-2009), where the weight in the weekly-rebalanced SR wCat portfolio peaked. At the same time, US Treasuries tend to remain a core holding over the backtesting period, and especially when complementing the Cat bond allocation in risk-minimizing strategies. The MV wCat strategy, for example (Appendix Figure 10c), maintains substantial weights in both Cat Bonds and US Treasuries, while the MV noCat portfolio (Appendix Figure 10d), is essentially almost exclusively invested in US Treasuries and US Corporate Bonds. Even when a 25% allocation cap is imposed (e.g., DR wCat25), optimized portfolios frequently allocate the maximum permitted weight to Cat Bonds (see Appendix Figure 10e).

These differences in portfolio composition translate into distinct performance characteristics over the out of sample period, particularly evident in calendar year results. In fact, the resulting performance metrics suggest distinct advantages to Cat Bonds portfolio inclusion, especially during periods of market instability. Across the strategies examined (monthly rebalancing, wTC), portfolios including Cat bonds generally outperformed their noCat counterparts in the majority of the full years from 2007 through 2024; for the SR maximising strategy, SR wCat posted higher annual returns than SR noCat in 11 out of these 18 years. Selected performance metrics over calendar years are provided in Appendix Table 19. It should be noted that the relative outperformance was often stronger during years of significant market stress: for example, in 2008 SR wCat returned 1.88% versus -2.87% for SR noCat, or during the turbulent market year of 2022 more recently, SR wCat returned -7.41% versus -13.59% for SR noCat.

Another clear measure of the contribution of Cat bonds during periods of stress can be seen, for example, in drawdown measures during the initial COVID-19 market crash (2020). In this regard, wCat portfolios across all strategy types experienced substantially smaller intra-year maximum drawdowns compared to their noCat equivalents; for instance, the SR wCat portfolio's maximum drawdown in 2020 was only 4.15% versus 18.23% for SR noCat (see Appendix Table 19). This effective risk mitigation likely contributed to the faster recovery observed in portfolios values including Cat bonds, as depicted in cumulative performance charts (see Figures 9a and 9b in Appendix B for strategies with and without Cat bonds

However, this relative advantage is not universal across all market conditions. In years characterized by strong equity bull markets (e.g., 2017 where SR wCat returned 0.29% vs 5.65% for SR noCat, or 2024 where SR wCat returned 18.33% vs 23.65% for SR noCat), the noCat strategies occasionally posted higher annual returns, suggesting a potential trade-off where the inclusion of Cat bonds might dampen upside capture in certain environments (illustrated in Appendix Table 19). Overall, calendar year data suggests that including Cat bonds can improve portfolio resilience, primarily by reducing losses during market crises. Comprehensive performance details for all strategies can be found in Appendix B (Tables 20 and 21).

6 Discussion

This section contextualizes the empirical findings on Catastrophe bond diversification and performance within the existing academic literature.

6.1 Risk-Return Profile & Distributional Characteristics

Consistent with prior research (Cummins and Weiss, 2009; Sterge and Van Der Stichele, 2016; Haffar and Le Fur, 2022), Cat bonds exhibited a compelling historical risk-return profile during the 2005-2025 sample period examined in this study. The analysis finds Cat bonds achieved a high Sharpe ratio (0.988) with an annualized mean return of 7.05% (Table 3). This Sharpe ratio is notably lower than the ones reported by Sterge and Van Der Stichele (2016) for a period ending 2014, potentially reflecting both their observation that low historical losses within the period inflated past returns and the yield compression observed in the market since then. The significant non-normality identified in Cat bond returns (Skewness -9.67, Kurtosis 284.26, Table 4) is also consistent with prior studies, especially Demers-Bélanger and Lai (2020), whose stochastic dominance tests also indicated mean-variance analysis might not fully capture the impact of Cat bonds' distributional properties.

6.2 Correlation Dynamics & Crisis Performance

Regarding the traditional diversification argument for Cat bonds, the low average unconditional correlation found with traditional assets (e.g., Pearson correlation of 0.062 vs Global Equities and 0.076 vs US Equities, Table 6) confirms a well-established finding (Cummins and Weiss, 2009; Sterge and Van Der Stichele, 2016; Haffar and Le Fur, 2022). The results of our dynamic analysis present a slightly more nuanced picture compared to some earlier studies. Our study, which crucially examines time-varying correlations over the more recent 2005-2025 period, covers not only the GFC, but also various recently impactful hurricane seasons as well as the COVID-19 crisis. While GFC-focused analyses like Carayannopoulos and Perez (2015) documented significant correlation increases with both equities (to 0.252) and corporate bonds (to 0.475), attributing this partly to specific structural issues of that crisis (collateral orcounterparty risks), using a VIX threshold of 25 over the longer period finds that the statistically significant correlation increase during stress was confined to US Corporate Bonds (rising from -0.009 to 0.274). The correlation with US equities (0.058 Normal vs 0.095 Crisis) and Global Equities (0.032 Normal vs 0.095 Crisis) did not show a statistically significant change (Table 9).

This divergence might reflect post-GFC market adaptations (Carayannopoulos and Perez, 2015), methodological differences (VIX vs. event-based crisis definition), or the distinct nature of the global financial crisis compared to later stress periods included in this sample. Indeed, direct comparison between the global financial crisis and COVID-19 periods in this study showed no significant difference in correlation changes for equities or corporate bonds (Table 17). Our finding aligns partially with Drobetz et al. (2020), who found Cat bonds acted as a safe haven (implying low or negative correlation during extreme stress) only against stocks in the post-GFC era, and contrasts with Clark et al. (2016)'s reported GFC DCC mean correlations (e.g., 0.1104 vs SPX). Finally, it should be noted that some of the differences in significance might stem from the different tests used (see Subsection 7.3) and their differences in statistical power.

6.3 Mean-Variance Spanning

Mean-variance spanning tests strongly indicate that Cat bonds expand the investment frontier, consistent with Clark et al. (2016)'s and Demers-Bélanger and Lai (2020)'s spanning rejections. The joint hypothesis of spanning is rejected across the full sample and within both VIX-defined Normal and Crisis regimes ($p<0.001$, Tables 11-13). The sources of these benefits differ across regimes, however. Over the full sample, a statistically significant annualized alpha of 6.98% is identified (Table 10). Examining the regimes reveals this alpha is primarily driven by performance during Normal market conditions (Ann. alpha 7.63%, $p<0.001$, Table 12). Contrary to expectations and some previous findings suggesting strong crisis alpha, the alpha during VIX Crisis periods is positive (2.05% ann.) but not statistically significant ($p=0.631$, Table 13). Despite the lack of significant crisis alpha, the distinct risk profile, quantified by the very low sum of betas relative to benchmarks (Full Sample: 0.021, Crisis: 0.145, Normal: 0.010, Section 5.4), confirms their limited exposure to traditional market risks and implies variance reduction benefits in all periods (H0: $\sum \beta_i = 1$ rejected across all tests).

6.4 Out-of-Sample Portfolio Performance

The out-of-sample backtesting, which incorporates a simplified transaction cost model (see Section 5.5) - a distinctive feature of this analysis - provided evidence of practical portfolio benefits, supporting similar positive OOS findings by Clark et al. (2016) and Demers-Bélanger and Lai (2020). The inclusion of Cat bonds consistently improved point estimates of risk-adjusted metrics (Sharpe, Sortino) and reduced downside risk (volatility, MaxDD) across various strategies (e.g., Table 14 and comprehensive results in Table 20). For example, the net Sharpe ratio for the SR strategy (quarterly rebalancing, wTC) improved substantially from 0.466 to 0.907. However, formal statistical tests indicate that these improvements, while economically meaningful, do not appear statistically significant at conventional levels ($p>0.12$ for all quarterly wTC comparisons, Table 16). Interestingly, statistical significance is found for the SR strategy difference when rebalanced monthly without transaction costs ($p=0.034$, Table 18), suggesting the benefits are sensitive to costs and frequency assumptions.

7 Limitations

This study's findings should be interpreted considering the following limitations of the analysis, which also suggest possible avenues for future research.

7.1 Data

We should start by acknowledging the well-known limitations regarding the reliability as a proxy and investability of the Swiss Re Cat Bond Indices suite. In fact, the Global Cat Bond TRI employed in this analysis is not a directly investable index, and while Swiss Re publishes its methodology (Swiss Re, 2014), the individual bond data is not available. Furthermore, the index returns are pre-cost—and do not reflect the practical realities of portfolio management, including transaction costs (bid-ask spreads, market impact, especially for less liquid bonds), management fees, etc. While this is a significant issue, qualitative comparisons with publicly available returns of relevant Cat bond funds suggest the index captures general market trends, although precise replication remains infeasible due to the factors mentioned.

Furthermore, while our 2005-2025 sample period covers most of the asset class's history, its relative brevity (20 years)—and the absence of a 'worst-case' catastrophic stress test within this timeframe—raise questions about the robustness of long-term return expectations and relationship modeling, particularly given the anticipated impacts of climate change on event frequency and severity.

Finally, the choice of benchmarks—very US-centric and traditional—was almost entirely limited by data availability. In this sense, a more comprehensive future analysis should incorporate insight from the performance of various alternative, widely investable assets such as private equity, real estate, and commodities.

7.2 Model & Methodology

The dynamic correlation and covariance models employed (e.g., Rolling Window, EWMA, DCC-GARCH, Sample Covariance with shrinkage), while standard, inherently rely on specific structural assumptions and parameter choices, such as distributional forms, decay factors, or estimation window lengths, involving necessary approximations. Sensitivity analyses were conducted to assess the impact of varying key parameters and alternative model specifications at different stages (e.g., covariance estimation), confirming the general validity of the main findings; these detailed checks are omitted for brevity.

Beyond model specification, parameter estimation relies significantly on assumptions and approximations. This is particularly acute for the expected return inputs required by mean-variance optimization, which are notoriously difficult to forecast accurately. While we tried to minimize the error by the use of blended mean estimation, the initial parameter estimates (detailed in Section 4.5.3) are also clearly affected, and their reliability depends on the chosen lookback period and priors. For assets with limited history like Cat bonds, these initial inputs carry significant estimation risk and must be interpreted with appropriate caution.

Another meaningful issue lies in the treatment of transaction costs. Clearly, the applied ex-post proportional cost model significantly simplifies reality. It ignores variable bid-ask spreads, market impact (especially important in an illiquid market such as this), and other operational costs, as well as—unrealistically—keeping these costs outside of the optimization function. While trading frictions for benchmark assets are generally low and well-known, the same is not true for Cat Bonds: the assumed 90 bps one-way cost should be considered an educated but arbitrary estimate; actual costs can vary widely. Consequently, the reported net-of-cost returns should be seen as optimistic upper bounds on practically achievable performance. In this sense, an obvious direction for future work lies in incorporating more sophisticated and realistic transaction cost and liquidity models in the analysis.

Finally, our analysis only concerned itself with mean-variance investors. While some measures of tail risk are taken into account, neither higher-moment performance comparisons nor different utility and preference sets have been treated. A significant future deepening of the analysis of the benefits of Cat bonds inclusion must confront these different contexts.

7.3 Statistical Inference & Significance Testing

The selection of significance testing techniques prioritized robustness to the observed non-normality and heteroskedasticity of Cat bond returns. Consequently, a critical issue to consider is the reduced statistical power inherent when applying these kinds of tests—including the non-parametric bootstrap for comparing correlation coefficients and the Wald tests within regime-specific spanning regressions—to subsamples characterized by limited observations. This reduction affects the sensitivity of most tests used in this thesis and increases the risk of Type II errors. As such, a failure to detect statistically significant differences within some subperiods might stem from insufficient data rather than the true absence of an effect. Observed effects in these subsamples, even if not meeting conventional significance thresholds—e.g., the performance improvements for EQ, MV, and DR strategies shown in Section 5.5, Table 16—should not be automatically dismissed. Their economic magnitude might still hold practical relevance for allocation decisions.

8 Conclusion

We employed a comprehensive empirical framework to study the diversification benefits of Cat bonds within multi-asset portfolios over a two-decade sample period (2005-2025). Our analysis provides evidence that, contrary to some findings during crisis periods, Cat bonds offer substantial diversification benefits that remain relatively stable across varying market conditions.

Using multiple correlation methodologies, we showed that Cat bonds present consistently low associations with traditional asset classes. While these correlations showed some time variation, they remained relatively stable even during crisis periods, as represented both by VIX market stress regimes and specific subsample analysis. The correlation increase during market stress was statistically significant only for US Corporate Bonds, and broadly of limited magnitude. This suggests that while Cat bonds' diversification benefits exhibit some state-dependency, varying by asset class and market condition, there's no evidence of their disappearance in times of crisis.

Furthermore, asset spanning tests provided evidence that Cat bonds expand the investment opportunity set. Both individual and joint hypotheses were strongly rejected across the full sample, confirming Cat bonds offer diversification benefits through both return enhancement (alpha) and variance reduction (distinct risk profile). When examined across market regimes, the variance reduction benefits remained statistically significant in both Normal and Crisis periods. However, the significant return enhancement (alpha) was found primarily during Normal market conditions, and was not statistically significant during high-stress (VIX Crisis) periods. Finally, the consistently low sum of betas supports the idea of their limited systematic exposure to traditional market risks across all conditions.

The comprehensive out-of-sample portfolio backtesting provided practical evidence of Cat bonds' diversification benefits, even when accounting for transaction costs. Across allocation strategies and rebalancing frequencies, portfolios including Cat bonds consistently outperformed their counterparts in terms of risk-adjusted metrics and downside protection, both over the full period and during the great majority of individual years. However, these improvements were not uniform across all strategies and market conditions, with some allocation approaches benefiting more substantially than others.

Our findings contribute to the literature by demonstrating that while Cat bonds may experience some correlation instability during severe market stress, their overall diversification efficacy does not seem to diminish meaningfully. In conclusion, this study provides robust evidence supporting the inclusion of catastrophe bonds within institutional portfolios, with diversification benefits that remain resilient—though not entirely immune to—varying market conditions and implementation approaches.

A Supplementary Methodological Details

This section provides supplementary formulations referenced in the main text or relevant to results presented in the Appendix.

A.1 Time Series Diagnostics Tests

(ADF) The Augmented Dickey-Fuller test regression takes the form:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \varepsilon_t \tag{24}$$

where the null hypothesis is $\gamma = 0$ (unit root).

(KPSS) The Kwiatkowski–Phillips–Schmidt–Shin test decomposes the series as:

$$y_t = \xi t + r_t + \varepsilon_t, \quad \text{where } r_t = r_{t-1} + u_t \tag{25}$$

testing the null hypothesis that the variance of u_t is zero (stationarity).

A.2 Rank Correlation Coefficients

Spearman's Rank Correlation (ρ) Spearman's rho assesses the strength and direction of a monotonic relationship between two ranked variables. It essentially calculates the Pearson correlation coefficient on the ranks of the data rather than their raw values. If d_i is the difference between the ranks for each observation i, and n is the number of observations, it is calculated (for data without ties) as:

$$\rho = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{26}$$

Kendall's Rank Correlation (τ) Kendall's tau measures the ordinal association between two variables based on the number of concordant and discordant pairs of observations. Kendall's tau-a (which does not adjust for ties), is calculated by considering all pairs of observations (X_i, Y_i) and (X_j, Y_j) where $i < j$:

$$\tau = \frac{2}{n(n-1)} \sum_{i<j} \text{sgn}(X_i - X_j)\text{sgn}(Y_i - Y_j) \tag{27}$$

where $\text{sgn}(z)$ denotes the sign function, returning $+1$ if $z > 0$, -1 if $z < 0$, and 0 if $z = 0$. The product $\text{sgn}(X_i - X_j)\text{sgn}(Y_i - Y_j)$ is $+1$ for concordant pairs (pairs ranked in the same order by both X and Y), −1 for discordant pairs (pairs ranked in opposite orders), and 0 if there is a tie in either X or Y for the pair.

A.3 Backtesting - Input Estimation Details

The estimation of input parameters for portfolio optimization is performed as follows:

Covariance Matrix The procedure utilizes a shrinkage covariance estimator based on the methods developed by Schäfer and Strimmer (2005) and Opgen-Rhein and Strimmer (2007). Instead of shrinking the sample covariance matrix (S) towards a single target matrix F as in the original Ledoit-Wolf formulation (Ledoit and Wolf, 2004a), this approach applies shrinkage separately to variances and correlations:

1. **Variance Shrinkage:** The empirical variances s_{kk} are shrunk towards their median value (v_{median}) using an analytically determined optimal intensity $\hat{\lambda}^*_{var}$:
$$s^*_{kk} = \hat{\lambda}^*_{var} v_{\text{median}} + (1 - \hat{\lambda}^*_{var}) s_{kk} \tag{28}$$
where $\hat{\lambda}^*_{var} = \min(1, \frac{\sum_{k=1}^{p} \hat{Var}(s_{kk})}{\sum_{k=1}^{p}(s_{kk} - v_{\text{median}})^2})$. Details on estimating $\hat{Var}(s_{kk})$ and the rationale are found in Opgen-Rhein and Strimmer (2007).

2. **Correlation Shrinkage:** The empirical correlation coefficients r_{kl} ($k \neq l$) are shrunk towards zero (i.e., the target is the identity matrix) using a separate optimal intensity $\hat{\lambda}^*$:
$$r^*_{kl} = r_{kl} \cdot \max(0, 1 - \hat{\lambda}^*) \tag{29}$$
where $\hat{\lambda}^* = \min(1, \frac{\sum_{k \neq l} \hat{Var}(r_{kl})}{\sum_{k \neq l} r^2_{kl}})$. The estimation of $\hat{Var}(r_{kl})$ is detailed in Schäfer and Strimmer (2005).

The final shrinkage covariance estimate $\hat{\Sigma}_{\text{Shrink}}$ is then reconstructed using the shrunken variances s^*_{kk} on the diagonal and calculating off-diagonal elements as $s^*_{kl} = r^*_{kl}\sqrt{s^*_{kk}s^*_{ll}}$. This method guarantees a positive definite and well-conditioned matrix while being computationally efficient (Schäfer and Strimmer, 2005).

Expected Returns The blended expected log return estimate for asset i at time t is:
$$\hat{\mu}_{i,t} = \alpha_t \cdot \mu_{\text{sample},i,t} + (1 - \alpha_t) \cdot \mu_{\text{prior},i} \tag{30}$$

where $\alpha_t = \frac{n_{obs,t}}{n_{obs,t}+k}$ with $k = 208$.

A.3.1 Time-zero Estimates

Our $\mu_{\text{prior},i}$ estimates were compiled from various online sources: MSCI ex-US and SPX data for the 13 years prior to the start of the sample was sourced from Yahoo Finance, Curvo and cross-checked with available official data from S&P and MSCI. Data for corporate bonds and US Treasuries (15 years prior to the sample period) was respectively sourced from A. Damodaran's Historical Data Archive at NYU Stern and approximately reconstructed from historical FRED 3,5 and 7-year Treasury yields data, as well as comprehensively cross-checked and sanity-checked with performance snapshots from existing indices investing in comparable asset classes.

The computation of a historical estimate for Cat bonds was not as obvious, since the asset class was essentially born shortly before our sample period start. We only used 4 years of prior data sourced from the specialist website Lane Financial LLC's historical publications and cross-checked it with historical performance snapshots for Cat Bonds and ILS from Artemis since their inception in the 90s.

A.4 Dynamic Correlation Model

The DCC-GARCH(1,1) model follows Engle (2002), modeling the conditional covariance matrix H_t as:

$$H_t = D_t R_t D_t \tag{31}$$

$$R_t = (\text{diag}(Q_t))^{-1/2} Q_t (\text{diag}(Q_t))^{-1/2} \tag{32}$$

$$Q_t = (1 - a - b)\bar{Q} + a(\epsilon_{t-1}\epsilon'_{t-1}) + bQ_{t-1} \tag{33}$$

where D_t contains time-varying standard deviations from univariate GARCH models (specified with Student's t-distribution), R_t is the time-varying correlation matrix, Q_t is the conditional covariance matrix of standardized residuals ϵ_t, $\bar{Q}$ is the unconditional covariance matrix of ϵ_t, and a and b are the DCC parameters.

A.5 Test Statistic for Sharpe Ratio Difference

The test for the equality of two Sharpe Ratios (H_0: $SR_i = SR_k$) follows the asymptotic methodology described by Ledoit and Wolf (2008). The test statistic, Z, is constructed as:

$$Z = \frac{\hat{\Delta}}{s(\hat{\Delta})} \tag{34}$$

where $\hat{\Delta} = \hat{SR}_i - \hat{SR}_k = \frac{\hat{\mu}_{excess,simple,i}}{\hat{\sigma}_i} - \frac{\hat{\mu}_{excess,simple,k}}{\hat{\sigma}_k}$ is the difference between the sample Sharpe Ratios of the two strategies i and k. Here, $\hat{\mu}_{excess,simple}$ represents the sample mean simple excess return of a strategy over the test period. The term $s(\hat{\Delta})$ represents the standard error of this difference, robustly estimated to account for potential non-normality and serial correlation in the return series. It is calculated using the Delta method based on a Heteroskedasticity and Autocorrelation Consistent (HAC) estimator ($\hat{\Psi}$) of the asymptotic covariance matrix of the relevant return moments ($\hat{\mathbf{v}} = (\hat{\mu}_{excess,simple,i}, \hat{\mu}_{excess,simple,k}, \hat{\gamma}_i, \hat{\gamma}_k)'$, where γ denotes uncentered second moments):

$$s(\hat{\Delta}) = \sqrt{\frac{\nabla f(\hat{\mathbf{v}})' \hat{\Psi} \nabla f(\hat{\mathbf{v}})}{T}} \tag{35}$$

Here, T is the number of observations and $\nabla f(\hat{\mathbf{v}})$ is the gradient of the Sharpe ratio difference function evaluated at the sample estimates $\hat{\mathbf{v}}$. Under the null hypothesis, this Z statistic is asymptotically distributed as a standard normal variable.

B Additional Tables & Figures

This appendix provides supplementary tables and figures related to the empirical results presented in Section 5.

B.1 Distribution Analysis

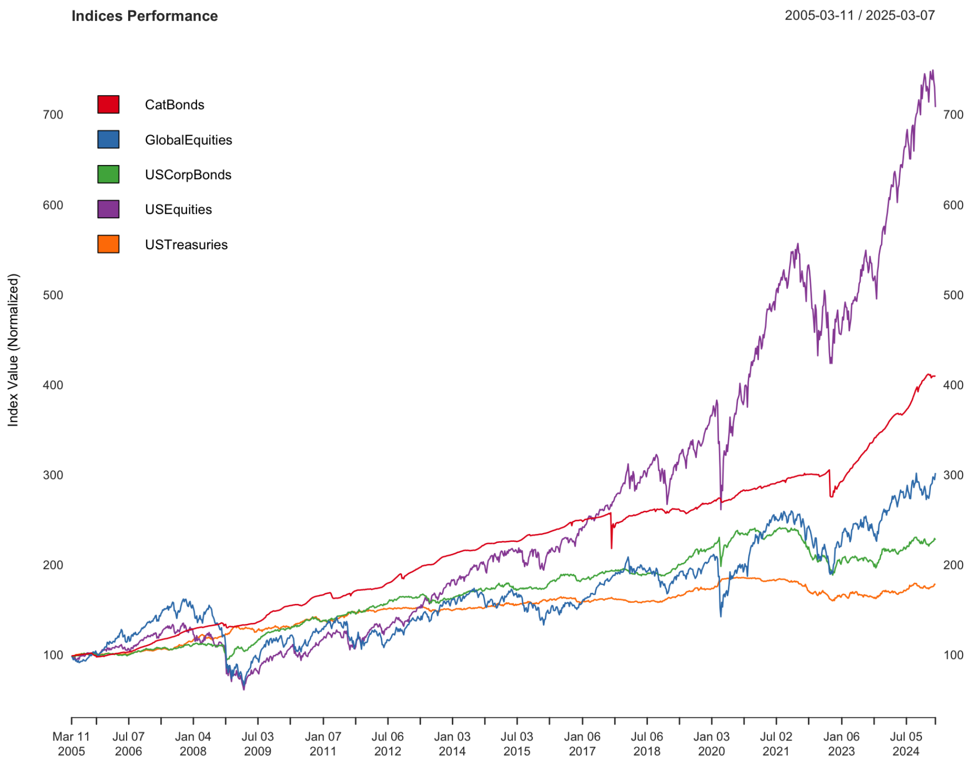


Figure 4: Sample Indices Performance

B.2 Correlation Analysis

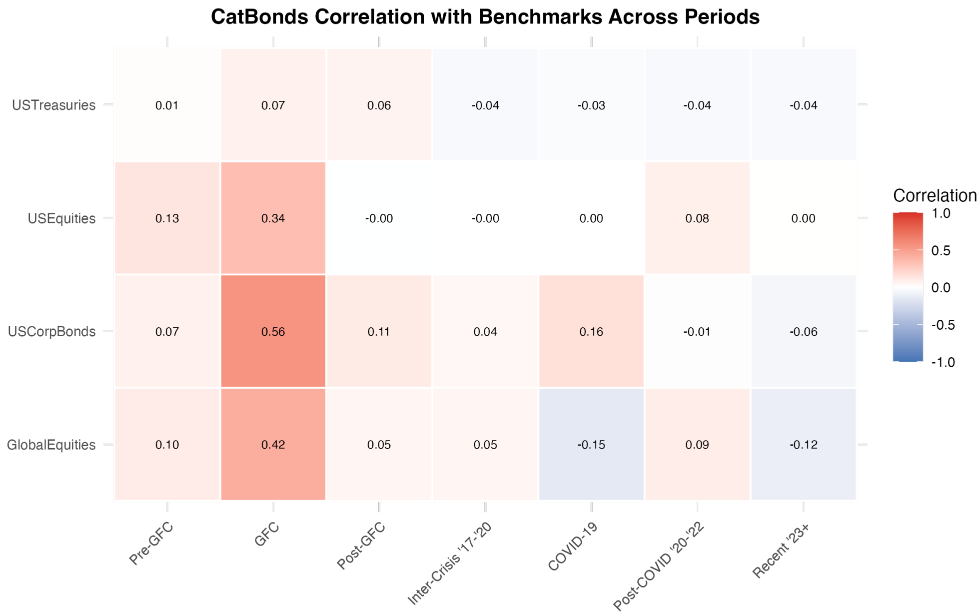


Figure 5: CatBonds Correlation with Benchmarks Across Periods

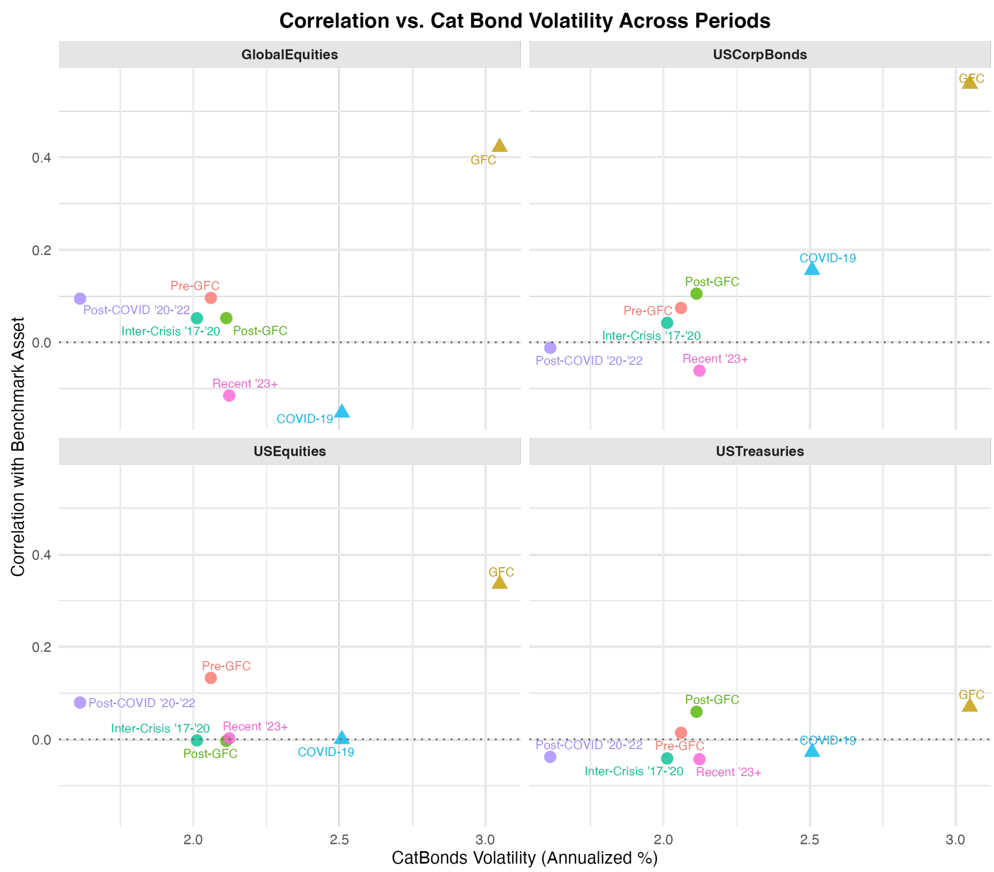
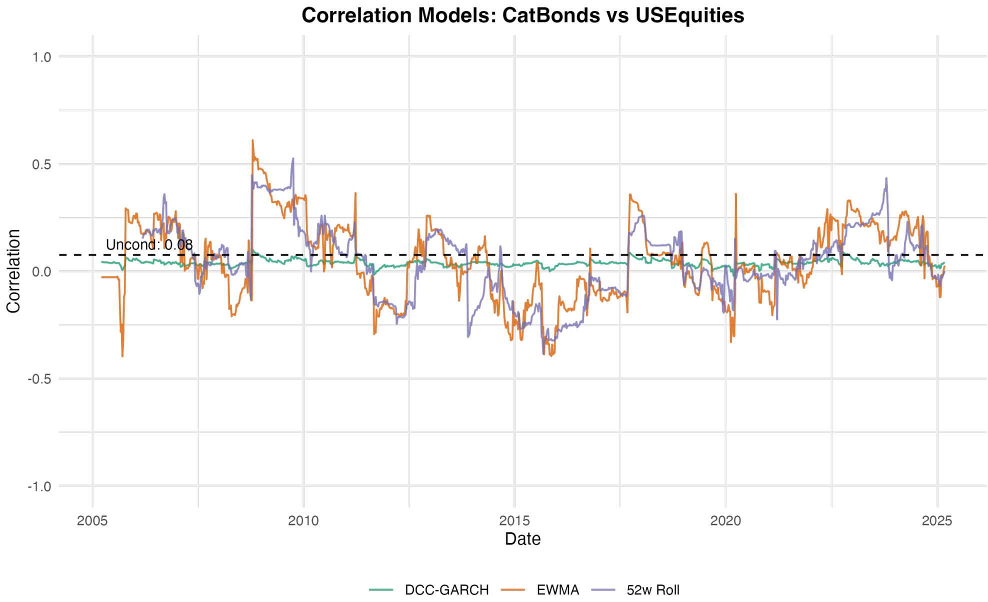
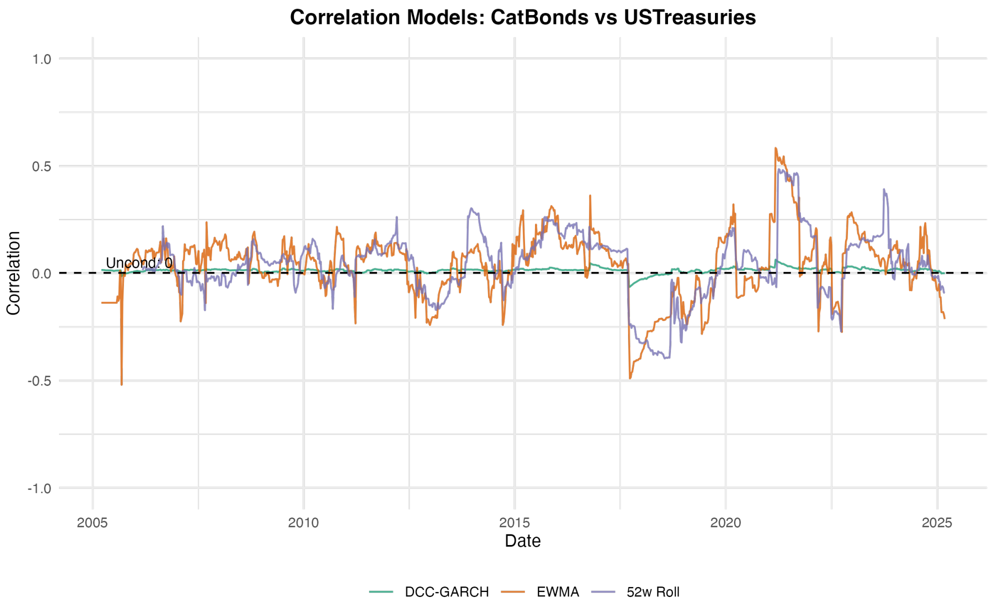


Figure 6: Correlation vs. Cat Bond Volatility Across Periods



(a) vs US Equities



(b) vs US Treasuries

Figure 7: Cat Bonds - Dynamic Correlation Estimates (Selected Models)

Table 17: Cat Bond Correlation Comparisons Across Economic Periods

Asset	Corr1	Corr2	Difference	Lower CI	Upper CI	Significance (5%)
Comparison: Pre-GFC vs GFC						
GlobalEquities	0.0960	0.4221	0.3260	-0.1817	0.6988	No
USCorpBonds	0.0743	0.5589	0.4846	-0.2401	0.8013	No
USEquities	0.1329	0.3361	0.2032	-0.3041	0.6216	No
USTreasuries	0.0142	0.0703	0.0561	-0.2148	0.2912	No
Comparison: Post-GFC vs Hurricane s. '17						
GlobalEquities	0.0523	-0.0530	-0.1053	—	—	Untested (Short Sample)
USCorpBonds	0.1052	-0.5510	-0.6562	—	—	Untested (Short Sample)
USEquities	-0.0033	0.7449	0.7482	—	—	Untested (Short Sample)
USTreasuries	0.0597	-0.7324	-0.7921	—	—	Untested (Short Sample)
Comparison: Inter-Crisis '17-'20 vs COVID-19						
GlobalEquities	0.0520	-0.1524	-0.2044	-0.7726	0.4113	No
USCorpBonds	0.0420	0.1561	0.1140	-0.7402	0.7978	No
USEquities	-0.0022	0.0002	0.0024	-0.6923	0.7489	No
USTreasuries	-0.0412	-0.0277	0.0136	-0.4973	0.6750	No

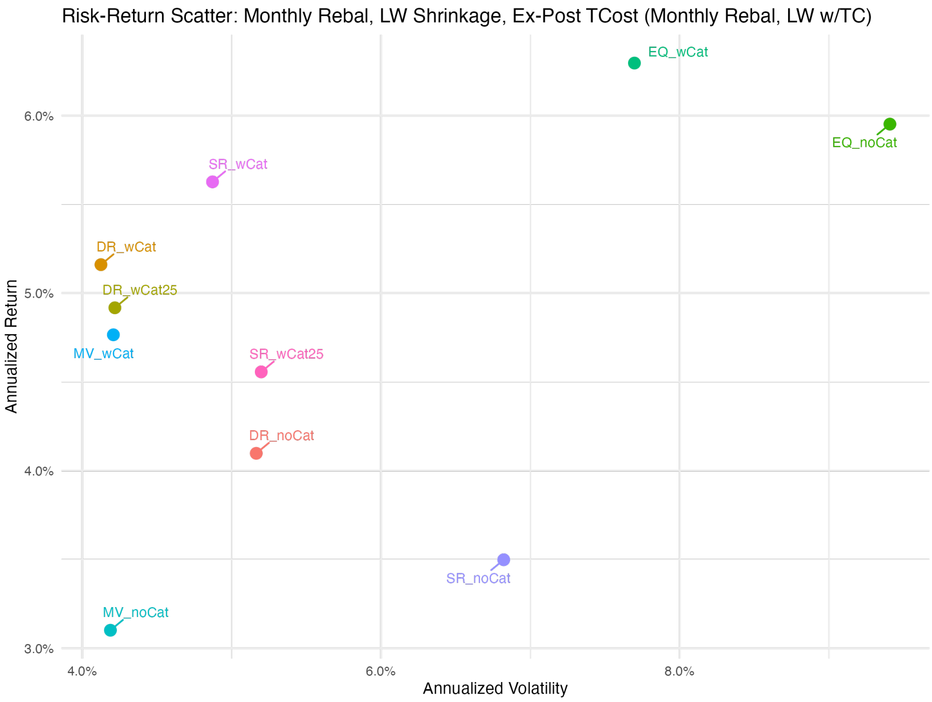
Continued on next page...

Table 17 – continued from previous page

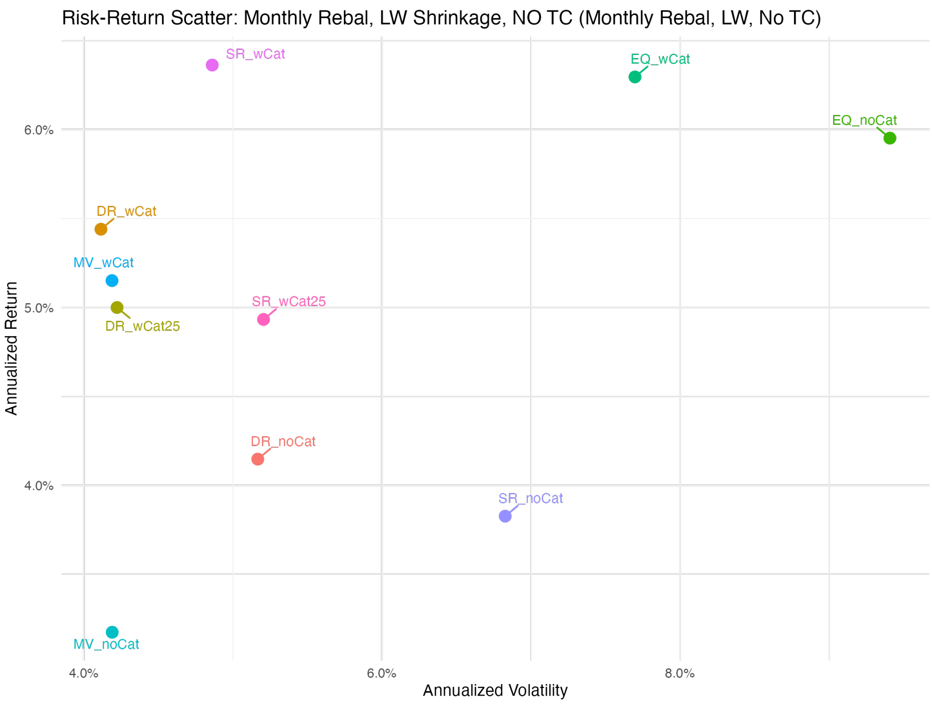
Asset	Corr1	Corr2	Difference	Lower CI	Upper CI	Significance (5%)
Comparison: Post-COVID '20-'22 vs Hurricane Ian '22						
GlobalEquities	0.0945	0.3314	0.2369	—	—	Untested (Short Sample)
USCorpBonds	-0.0120	0.5371	0.5491	—	—	Untested (Short Sample)
USEquities	0.0798	0.3756	0.2958	—	—	Untested (Short Sample)
USTreasuries	-0.0378	0.3068	0.3446	—	—	Untested (Short Sample)
Comparison: GFC vs COVID-19						
GlobalEquities	0.4221	-0.1524	-0.5745	-1.1326	0.2711	No
USCorpBonds	0.5589	0.1561	-0.4028	-1.2913	0.6162	No
USEquities	0.3361	0.0002	-0.3360	-1.1032	0.5516	No
USTreasuries	0.0703	-0.0277	-0.0980	-0.6153	0.5523	No
Comparison: Hurricane s. '17 vs Hurricane Ian '22						
GlobalEquities	-0.0530	0.3314	0.3844	—	—	Untested (Short Sample)
USCorpBonds	-0.5510	0.5371	1.0881	—	—	Untested (Short Sample)
USEquities	0.7449	0.3756	-0.3693	—	—	Untested (Short Sample)
USTreasuries	-0.7324	0.3068	1.0392	—	—	Untested (Short Sample)

Note: Corr1 and Corr2 represent correlations in the first and second periods of the comparison group, respectively. Difference is calculated as Corr2 - Corr1. CI denotes the 95% Confidence Interval for the difference, computed with a Bootstrap method where applicable. Significance is measured at the 5% level. Missing CI values (—) and 'Untested' status occur where tests could not be conducted due to small sample sizes.

B.3 Portfolio Backtesting Results

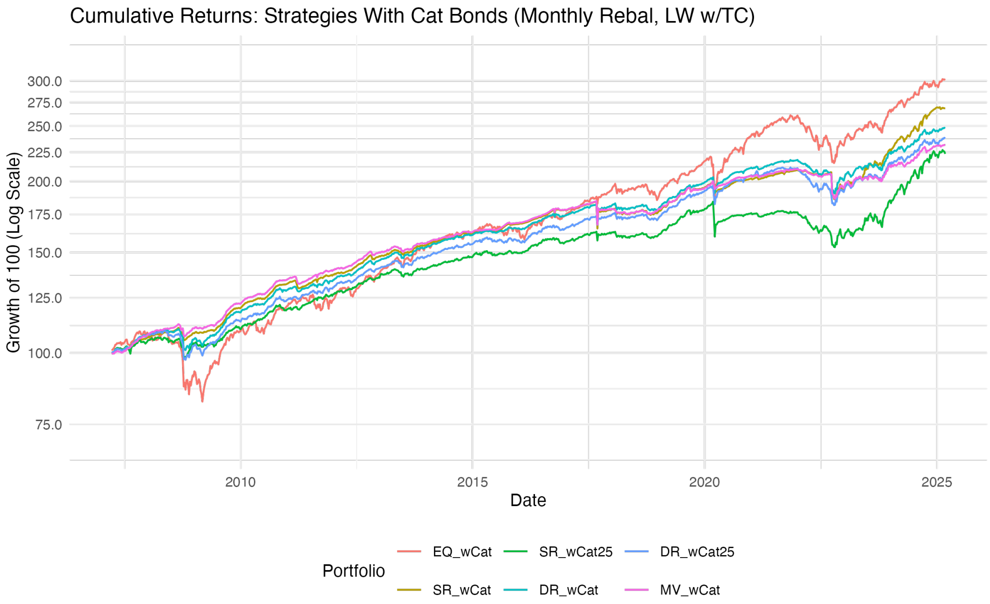


(a) Monthly Rebalancing with Transaction Costs

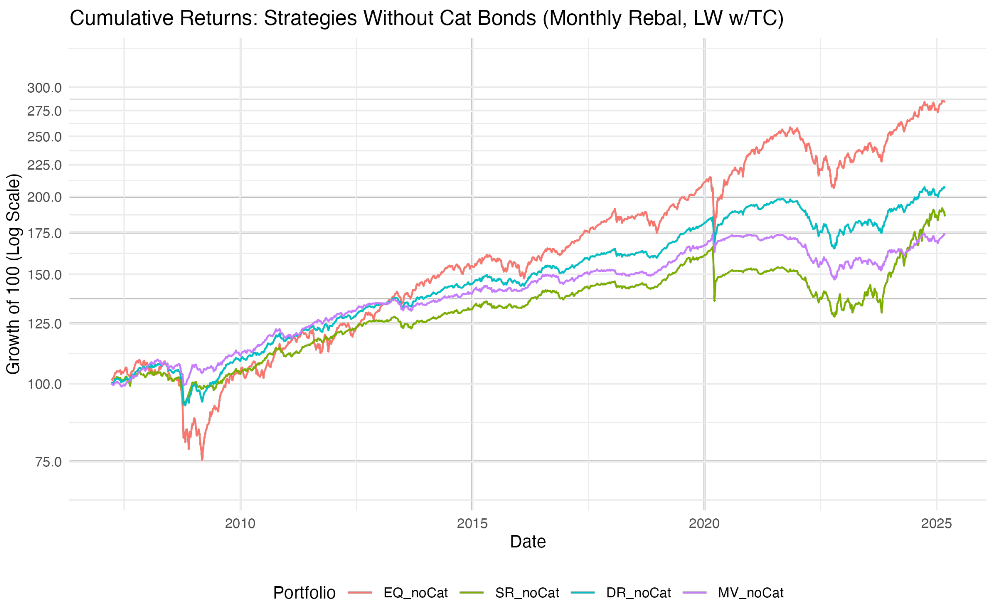


(b) Monthly Rebalancing without Transaction Costs

Figure 8: Risk-Return Scatterplots for Monthly Rebalanced Strategies

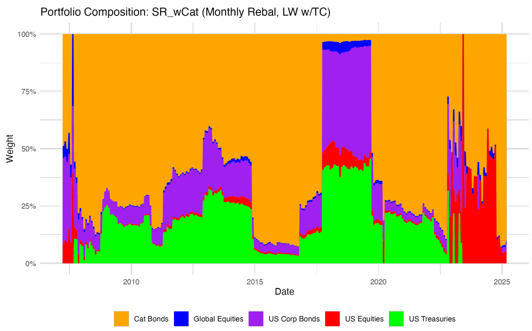


(a) Strategies With Cat Bonds

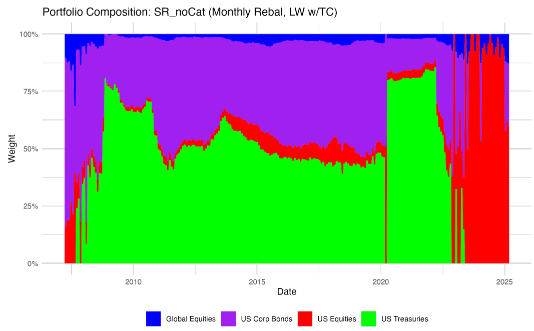


(b) Strategies Without Cat bonds

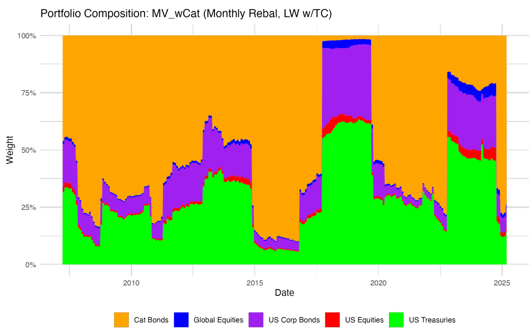
Figure 9: Cumulative Return Comparison: Monthly Rebalancing Strategies with and without Cat Bonds



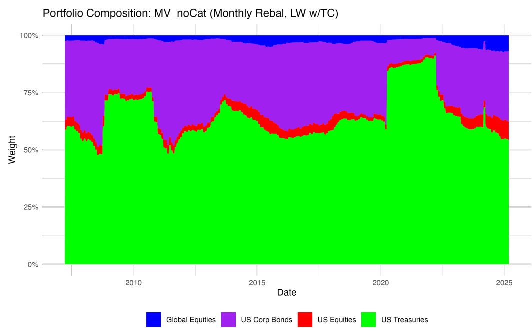
(a) SR wCat



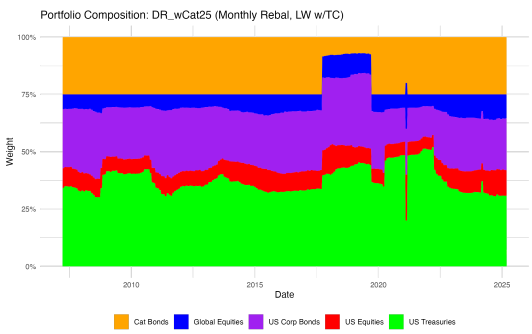
(b) SR noCat



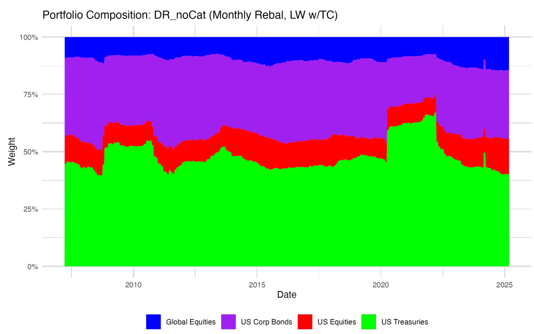
(c) MV wCat



(d) MV noCat



(e) DR wCat25



(f) DR noCat

Figure 10: Impact of Cat Bond Inclusion and Constraints on Portfolio Weights Across Strategies (SR, MV, DR). Selected plots (Monthly rebalancing, with Transaction Costs).

Table 18: Tests for Sharpe Ratio Differences (Monthly Rebalancing, No Transaction Costs)

Comparison	Sharpe1	Sharpe2	Difference	p-value	Result (5%)
EQ noCat vs wCat	0.522	0.661	0.139	0.611	No Sign. Diff.
SR noCat vs wCat	0.391	1.023	0.631	0.034	Significant
DR noCat vs wCat	0.557	0.988	0.431	0.139	No Sign. Diff.
MV noCat vs wCat	0.452	0.905	0.454	0.175	No Sign. Diff.

Table 19: Selected Calendar Year Performance Comparison

Year	Strategy	Portfolio	AnnRet	MaxDD	AnnRet Diff.
2008	SR	with Cat Bonds	0.019	0.049	0.048
		without Cat Bonds	−0.029	0.110	
	MV	with Cat Bonds	0.028	0.048	0.017
		without Cat Bonds	0.011	0.091	
	EQ	with Cat Bonds	−0.163	0.229	0.043
		without Cat Bonds	−0.206	0.278	
2017	SR	with Cat Bonds	0.003	0.105	−0.054
		without Cat Bonds	0.056	0.013	
	MV	with Cat Bonds	0.003	0.091	−0.042
		without Cat Bonds	0.045	0.013	
	EQ	with Cat Bonds	0.109	0.028	−0.023
		without Cat Bonds	0.132	0.008	
2020	SR	with Cat Bonds	0.038	0.041	0.085
		without Cat Bonds	−0.047	0.182	
	MV	with Cat Bonds	0.052	0.033	0.011
		without Cat Bonds	0.041	0.064	
	EQ	with Cat Bonds	0.104	0.151	−0.008
		without Cat Bonds	0.112	0.186	
2022	SR	with Cat Bonds	−0.074	0.111	0.062
		without Cat Bonds	−0.136	0.156	
	MV	with Cat Bonds	−0.075	0.113	0.031
		without Cat Bonds	−0.106	0.140	
	EQ	with Cat Bonds	−0.115	0.175	0.024
		without Cat Bonds	−0.138	0.199	
2024	SR	with Cat Bonds	0.183	0.020	−0.053
		without Cat Bonds	0.237	0.053	
	MV	with Cat Bonds	0.083	0.023	0.050
		without Cat Bonds	0.033	0.029	
	EQ	with Cat Bonds	0.105	0.027	0.017
		without Cat Bonds	0.088	0.035	

Note: Comparison of optimized portfolios with and without Cat Bonds for selected years. AnnRet = Annualized Return. MaxDD = Maximum Drawdown (as positive decimal). AnnRet Diff. = Annual Return Difference (with Cat Bonds - without Cat Bonds). Updated with new test data.

Table 20: Comprehensive Performance Metrics by Strategy and Rebalancing Frequency

Frequency	Portfolio	AnnRet	AnnVol	Sharpe	Sortino	MaxDD	VaR 95%	ES 95%	AnnTCost	AnnTurn
Weekly	EQ noCat	0.0595	0.0941	0.5218	0.1000	0.3098	-0.0176	-0.0309	0.0000	0.0000
	EQ wCat	0.0630	0.0770	0.6607	0.1275	0.2512	-0.0144	-0.0253	0.0000	0.0000
	SR noCat	0.0336	0.0589	0.3654	0.0728	0.1790	-0.0121	-0.0199	0.0099	2.4526
	SR wCat	0.0405	0.0449	0.6142	0.1004	0.1239	-0.0058	-0.0145	0.0188	2.2053
	SR wCat25	0.0395	0.0459	0.5796	0.1153	0.1357	-0.0086	-0.0154	0.0103	2.1869
	DR noCat	0.0402	0.0513	0.5377	0.1018	0.1686	-0.0103	-0.0164	0.0015	0.4245
	DR wCat	0.0480	0.0394	0.8757	0.1567	0.1261	-0.0062	-0.0126	0.0040	0.4021
	DR wCat25	0.0470	0.0414	0.8120	0.1535	0.1409	-0.0073	-0.0133	0.0018	0.3510
	MV noCat	0.0312	0.0412	0.4462	0.0854	0.1540	-0.0091	-0.0132	0.0010	0.2316
	MV wCat	0.0435	0.0383	0.7871	0.1279	0.1142	-0.0044	-0.0117	0.0045	0.4394
Monthly	EQ noCat	0.0595	0.0941	0.5218	0.1000	0.3098	-0.0176	-0.0309	0.0000	0.0000
	EQ wCat	0.0630	0.0770	0.6607	0.1275	0.2512	-0.0144	-0.0253	0.0000	0.0000
	SR noCat	0.0350	0.0682	0.3449	0.0645	0.2289	-0.0120	-0.0225	0.0032	0.7937
	SR wCat	0.0563	0.0487	0.8777	0.1564	0.1112	-0.0047	-0.0142	0.0069	0.7567
	SR wCat25	0.0456	0.0520	0.6293	0.1195	0.1680	-0.0086	-0.0169	0.0036	0.7368
	DR noCat	0.0410	0.0517	0.5478	0.1038	0.1686	-0.0103	-0.0164	0.0005	0.1103
	DR wCat	0.0516	0.0413	0.9206	0.1704	0.1263	-0.0061	-0.0128	0.0027	0.2586
	DR wCat25	0.0492	0.0422	0.8462	0.1615	0.1413	-0.0073	-0.0135	0.0008	0.1445
	MV noCat	0.0310	0.0419	0.4347	0.0832	0.1553	-0.0090	-0.0136	0.0007	0.1615
	MV wCat	0.0477	0.0421	0.8140	0.1426	0.1133	-0.0044	-0.0122	0.0037	0.3510
Quarterly	EQ noCat	0.0595	0.0941	0.5218	0.1000	0.3098	-0.0176	-0.0309	0.0000	0.0000
	EQ wCat	0.0630	0.0770	0.6607	0.1275	0.2512	-0.0144	-0.0253	0.0000	0.0000
	SR noCat	0.0445	0.0706	0.4659	0.0912	0.1823	-0.0112	-0.0219	0.0016	0.3913
	SR wCat	0.0593	0.0504	0.9067	0.1649	0.1033	-0.0050	-0.0145	0.0043	0.4297
	SR wCat25	0.0518	0.0532	0.7276	0.1424	0.1337	-0.0086	-0.0165	0.0015	0.3222
	DR noCat	0.0434	0.0531	0.5785	0.1120	0.1650	-0.0103	-0.0165	0.0005	0.1235
	DR wCat	0.0522	0.0407	0.9468	0.1762	0.1183	-0.0061	-0.0125	0.0020	0.1902
	DR wCat25	0.0494	0.0425	0.8457	0.1619	0.1388	-0.0073	-0.0135	0.0007	0.1120
	MV noCat	0.0321	0.0424	0.4553	0.0873	0.1502	-0.0090	-0.0136	0.0006	0.1300
	MV wCat	0.0475	0.0418	0.8152	0.1437	0.0989	-0.0048	-0.0120	0.0033	0.3105

Table 21: Comprehensive Performance Metrics by Strategy and Rebalancing Frequency (No Transaction Costs)

Frequency	Portfolio	AnnRet	AnnVol	Sharpe	Sortino	MaxDD	VaR 95%	ES 95%	AnnTCost	AnnTurn
Weekly	EQ noCat	0.0595	0.0941	0.5218	0.1000	0.3098	-0.0176	-0.0309	0.0000	0.0000
	EQ wCat	0.0630	0.0770	0.6607	0.1275	0.2512	-0.0144	-0.0253	0.0000	0.0000
	SR noCat	0.0438	0.0588	0.5346	0.1092	0.1662	-0.0118	-0.0195	0.0000	2.4526
	SR wCat	0.0602	0.0448	1.0336	0.1738	0.1141	-0.0045	-0.0135	0.0000	2.2053
	SR wCat25	0.0502	0.0459	0.8035	0.1650	0.1324	-0.0083	-0.0149	0.0000	2.1869
	DR noCat	0.0418	0.0513	0.5676	0.1077	0.1675	-0.0103	-0.0163	0.0000	0.4245
	DR wCat	0.0522	0.0393	0.9791	0.1768	0.1212	-0.0059	-0.0124	0.0000	0.4021
	DR wCat25	0.0488	0.0414	0.8544	0.1622	0.1401	-0.0072	-0.0132	0.0000	0.3510
	MV noCat	0.0322	0.0412	0.4706	0.0903	0.1524	-0.0091	-0.0132	0.0000	0.2316
	MV wCat	0.0482	0.0382	0.9065	0.1481	0.1051	-0.0041	-0.0115	0.0000	0.4394
Monthly	EQ noCat	0.0595	0.0941	0.5218	0.1000	0.3098	-0.0176	-0.0309	0.0000	0.0000
	EQ wCat	0.0630	0.0770	0.6607	0.1275	0.2512	-0.0144	-0.0253	0.0000	0.0000
	SR noCat	0.0383	0.0683	0.3912	0.0736	0.2250	-0.0119	-0.0224	0.0000	0.7937
	SR wCat	0.0636	0.0486	1.0226	0.1841	0.1047	-0.0046	-0.0136	0.0000	0.7567
	SR wCat25	0.0493	0.0521	0.6977	0.1335	0.1650	-0.0084	-0.0168	0.0000	0.7368
	DR noCat	0.0415	0.0517	0.5570	0.1056	0.1679	-0.0103	-0.0164	0.0000	0.1103
	DR wCat	0.0544	0.0411	0.9876	0.1836	0.1223	-0.0059	-0.0126	0.0000	0.2586
	DR wCat25	0.0500	0.0422	0.8642	0.1652	0.1407	-0.0073	-0.0135	0.0000	0.1445
	MV noCat	0.0317	0.0419	0.4515	0.0865	0.1533	-0.0090	-0.0136	0.0000	0.1615
	MV wCat	0.0515	0.0419	0.9054	0.1594	0.1056	-0.0043	-0.0118	0.0000	0.3510
Quarterly	EQ noCat	0.0595	0.0941	0.5218	0.1000	0.3098	-0.0176	-0.0309	0.0000	0.0000
	EQ wCat	0.0630	0.0770	0.6607	0.1275	0.2512	-0.0144	-0.0253	0.0000	0.0000
	SR noCat	0.0462	0.0707	0.4881	0.0958	0.1823	-0.0112	-0.0219	0.0000	0.3913
	SR wCat	0.0639	0.0504	0.9920	0.1812	0.1031	-0.0047	-0.0144	0.0000	0.4297
	SR wCat25	0.0534	0.0533	0.7556	0.1483	0.1325	-0.0086	-0.0165	0.0000	0.3222
	DR noCat	0.0439	0.0531	0.5881	0.1139	0.1643	-0.0103	-0.0165	0.0000	0.1235
	DR wCat	0.0543	0.0408	0.9946	0.1858	0.1177	-0.0061	-0.0125	0.0000	0.1902
	DR wCat25	0.0501	0.0425	0.8618	0.1651	0.1382	-0.0073	-0.0135	0.0000	0.1120
	MV noCat	0.0327	0.0424	0.4690	0.0901	0.1493	-0.0090	-0.0136	0.0000	0.1300
	MV wCat	0.0509	0.0417	0.8954	0.1585	0.0975	-0.0046	-0.0119	0.0000	0.3105

References

Artemis.bm (2013). The death of the total return swap in the cat bond market is almost upon us. https://www.artemis.bm/news/the-death-of-the-total-return-swap-in-the-cat-bond-market-is-almost-upon-us/.

Artemis.bm (2017). Jamaica questions CCRIF model after floods fail to trigger policy. https://www.artemis.bm/news/jamaica-questions-ccrif-model-after-floods-fail-to-trigger-policy/.

Artemis.bm (2018). Q1 2018 Catastrophe Bond & ILS Market Report. Technical report. https://www.artemis.bm/wp-content/uploads/2018/05/q1-2018-cat-bond-ils-market-report.pdf.

Artemis.bm (2024). NHC hurricane Otis report suggests Fonden Class D cat bond loss will be ˜50%. https://www.artemis.bm/news/nhc-hurricane-otis-report-suggests-fonden-class-d-cat-bond-loss-will-be-50/.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9(3):203–228.

Bacon, C. R. (2012). *Practical Portfolio Performance Measurement and Attribution*. Wiley, 1 edition.

Bai, J. and Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, 66(1):47–78.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1):1–22.

Balduzzi, P. (1999). Transaction costs and predictability: some utility cost calculations. *Journal of Financial Economics*, 52(1):47–78.

Bali, T. G. and Hovakimian, A. (2009). Volatility Spreads and Expected Stock Returns. *Management Science*, 55(11):1797–1812.

Barnett, M., Brock, W. A., and Hansen, L. P. (2020). Pricing Uncertainty Induced by Climate Change. *The Review of Financial Studies*, 33(3):1024–1066.

Braun, A. (2016). Pricing in the Primary Market for Cat Bonds: New Empirical Evidence. *Journal of Risk and Insurance*, 83(4):811–847.

Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton University Press, Princeton, N.J.

Canabarro, E., Finkemeier, M., Anderson, R. R., and Bendimerad, F. (2000). Analyzing Insurance-Linked Securities. *The Journal of Risk Finance*, 1(2):49–75.

Carayannopoulos, P., Kanj, O., and Perez, M. F. (2022). Pricing dynamics in the market for catastrophe bonds. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 47(1):172–202.

Carayannopoulos, P. and Perez, M. F. (2015). Diversification through Catastrophe Bonds: Lessons from the Subprime Financial Crisis. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 40(1):1–28.

Choueifaty, Y. and Coignard, Y. (2008). Toward Maximum Diversification. *The Journal of Portfolio Management*, 35(1):40–51.

Clark, S. P., Neale, F. R., and Dickson, M. (2016). Portfolio Diversification Effects of Catastrophe Bonds. *SSRN Electronic Journal*.

Clarke, R. G., Silva, H. d., and Thorley, S. (2006). Minimum-Variance Portfolios in the U.S. Equity Market. *The Journal of Portfolio Management*, 33(1):10–24.

Cummins, J., Doherty, N., and Lo, A. (2002). Can insurers pay for the "big one"? Measuring the capacity of the insurance market to respond to catastrophic losses. *Journal of Banking & Finance*, 26(2-3):557–583.

Cummins, J. D. and Trainar, P. (2009). Securitization, Insurance, and Reinsurance. *Journal of Risk and Insurance*, 76(3):463–492.

Cummins, J. D. and Weiss, M. A. (2009). Convergence of Insurance and Financial Markets: Hybrid and Securitized RiskTransfer Solutions. *Journal of Risk and Insurance*, 76(3):493–545.

Demers-Bélanger, K. and Lai, V. S. (2020). Diversification benefits of cat bonds: An indepth examination. *Financial Markets, Institutions & Instruments*, 29(5):165–228.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, 22(5):1915–1953.

DeRoon, F. A. and Nijman, T. E. (2001). Testing for mean-variance spanning: a survey. *Journal of Empirical Finance*, 8(2):111–155.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, 74(366):427–431.

Dowd, K. (2008). *Measuring market risk*. Wiley, Chichester, 2. ed. edition.

Drobetz, W., Schröder, H., and Tegtmeier, L. (2020). The role of catastrophe bonds in an international multi-asset portfolio: Diversifier, hedge, or safe haven? *Finance Research Letters*, 33:101198.

Efron, B. and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Chapman and Hall/CRC, New York.

Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics*, 20(3):339–350.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007.

Eren, E., Merten, F., and Verhoeven, N. (2022). Pricing of climate risks in financial markets: a summary of the literature. (130).

Etzion, D., Kypraios, E., and Forgues, B. (2019). Employing finance in pursuit of the sustainable development goals: The promise and perils of catastrophe bonds. *Academy of Management Discoveries*, 5(4):530–554.

Fleming, J., Kirby, C., and Ostdiek, B. (2001). The Economic Value of Volatility Timing. *The Journal of Finance*, 56(1):329–352.

Galeotti, M., Gürtler, M., and Winkelvos, C. (2013). Accuracy of Premium Calculation Models for CAT Bonds—An Empirical Analysis. *Journal of Risk and Insurance*, 80(2):401–421.

Gürtler, M., Hibbeln, M., and Winkelvos, C. (2016). The Impact of the Financial Crisis and Natural Catastrophes on CAT Bonds. *Journal of Risk and Insurance*, 83(3):579–612.

Haffar, A. and Le Fur, X. (2022). Dependence structure of CAT bonds and portfolio diversification: a copula-GARCH approach. *Journal of Asset Management*, 23(4):297–309.

Hansen, B. E. (2005). Challenges for Econometric Model Selection. *Econometric Theory*, 21(1):60–68.

Huberman, G. and Kandel, S. (1987). Mean-Variance Spanning. *The Journal of Finance*, 42(4):873–888.

Jagannathan, R. and Ma, T. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *The Journal of Finance*, 58(4):1651–1683.

James, W. and Stein, C. (1961). Estimation with Quadratic Loss. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 361–380. University of California Press.

Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255–259.

Jarzabkowski, P., Bednarek, R., and Spee, P. (2015). *Making a Market for Acts of God: The Practice of Risk Trading in the Global Reinsurance Industry*. Oxford University Press.

Jorion, P. (1986). Bayes-Stein Estimation for Portfolio Analysis. *The Journal of Financial and Quantitative Analysis*, 21(3):279–292.

Kan, R. and Zhou, G. (2012). Tests of Mean-Variance Spanning. *Annals of Economics & Finance*, 13(1):145–193.

Kish, R. J. (2016). Catastrophe (CAT) bonds: risk offsets with diversification and high returns. *Financial Services Review*, 25(3):303–329.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3):159–178.

Künsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics*, 17(3):1217–1241.

Lakdawalla, D. and Zanjani, G. (2012). Catastrophe Bonds, Reinsurance, and the Optimal Collateralization of Risk Transfer. *Journal of Risk and Insurance*, 79(2):449–476.

Lane, M., editor (2002). *Alternative risk strategies.* Risk Waters, London.

Lane, M. and Mahul, O. (2008). Catastrophe Risk Pricing: An Empirical Analysis. Technical Report 4765, Policy Research Working Paper.

Lane, M. N. (2000). Pricing Risk Transfer Transactions. *ASTIN Bulletin*, 30:259–293.

Ledoit, O. and Wolf, M. (2004a). Honey, I Shrunk the Sample Covariance Matrix. *The Journal of Portfolio Management*, 30(4):110–119.

Ledoit, O. and Wolf, M. (2004b). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

Litzenberger, R. H., Beaglehole, D. R., and Reynolds, C. E. (1996). Assessing Catastrophe Reinsurance-Linked Securities as a New Asset Class. *Journal of Portfolio Management*, 23(5):76–86.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Longerstaey, J. and Spencer, M. (1996). RiskMetrics Technical Document - 4th Edition. Technical report, J.P. Morgan/Reuters.

MacKinnon, J. G. and White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3):305–325.

Major, J. A. and Kreps, R. E. (2003). Catastrophe Risk Pricing in the Traditional Market. In Lane, M. N., editor, *Alternative Risk Strategies*, pages 201–222. Risk Books, London.

Maran, R. (2024). Do investors reward sovereign catastrophe bond issuance? Evidence from a panel of 26 disaster-prone countries. *Review of World Economics.*

Mevorach, C. (2018). Modelling Catastrophe Bond Pricing in the Primary Market - A Loglinear Approach. Technical report, University of Rochester.

Morana, C. and Sbrana, G. (2019). Climate change implications for the catastrophe bonds market: An empirical analysis. *Economic Modelling*, 81:274–294.

Motlagh, F., Hamideh, S., Gallagher, M., Yan, G., and Van De Lindt, J. W. (2024). Bonds for disaster resilience: A review of literature and practice. *International Journal of Disaster Risk Reduction*, 104:104318.

Naik, G. (2024). Cat-bond issuance pushes overall market to almost $50 billion. Bloomberg.com.

OECD (2024). Fostering Catastrophe Bond Markets in Asia and the Pacific. Technical report.

Opgen-Rhein, R. and Strimmer, K. (2007). Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach. *Statistical Applications in Genetics and Molecular Biology*, 6.

Phan, T. and Schwartzman, F. (2024). Climate defaults and financial adaptation. *European Economic Review*, 170:104866.

Polacek, A. (2018). Catastrophe bonds: A primer and retrospective. *Chicago Fed Letter*, (404).

Politis, D. N. and Romano, J. P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313.

Re, S. (2011). The fundamentals of insurance-linked securities. Technical report, Swiss Re.

Schäfer, J. and Strimmer, K. (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, 4.

Shapiro, S. S. and Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4):591–611.

Sharpe, W. F. (1966). Mutual Fund Performance. *The Journal of Business*, 39(1):119–138.

Sterge, A. J. and Van Der Stichele, B. (2016). Understanding Cat Bonds. *The Journal of Alternative Investments*, 19(1):27–40.

Swiss Re (2014). Cat Bond Indices Methodology. Technical report, Swiss Re.

Tsay, R. S. (2010). *Analysis of Financial Time Series*. Wiley Series in Probability and Statistics. Wiley, 3 edition.

Vaugirard, V. E. (2003). Pricing catastrophe bonds by an arbitrage approach. *The Quarterly Review of Economics and Finance*, 43(1):119–132.

Wang, S. (2004). CAT Bond Pricing Using Probability Transforms. *Geneva Papers: Etudes et Dossiers*, (278):19–29.

Whaley, R. E. (2009). Understanding the VIX. *The Journal of Portfolio Management*, 35(3):98–105.

White, H. (2000). A Reality Check for Data Snooping. *Econometrica*, 68(5):1097–1126.

Zeileis, A. (2004). Econometric Computing with HC and HAC Covariance Matrix Estimators. *Journal of Statistical Software*, 11(10).

Zeller, W. (2008). Securitization and Insurance – Characteristics of Hannover Re's Approach. *The Geneva Papers on Risk and Insurance - Issues and Practice*, 33(1):7–11.