POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

Executive Summary of the Thesis

On the interest rate risk: modelling, calibration and impact assessment for a simplified insurance portfolio.

Laurea Magistrale in Mathematical Engineering - Ingegneria Matematica

Author: Ludovica Grifa

Advisor: Prof. Daniele Marazzina

Co-advisor: Silvia Dell'Acqua

Academic year: 2023-2024

---

1. Introduction

This piece of work deals with the interest rate risk modelling in the insurance business. It compares the regulatory mandates to different stressing methodologies, measuring the implications these have on a simplified insurance portfolio. The impact assessment is carried out for some testing cases, that cover assets and liabilities portfolios characterized by different durations and biting scenarios (due to different duration gaps). The analyses performed blend economic intuition, linear algebra and statistical knowledge. They are based on a Python script, chosen over other languages for its versatility, assortment of libraries and for being free of charge.

2. Interest Rate Risk

The European Union has introduced a regulatory framework, called Solvency II (SII), aimed at stabilizing the insurance sector by setting robust capital requirements (SCR – Solvency Capital Requirements) to ensure the solvability of the insurance companies. The SCR is set by examining the extent to which the company's Own Funds (OF) decrease under adverse conditions. If not well measured and hedged, interest rate (IR) risk is one of the major causes for the SCR to increase. Indeed, the mismatch in timing between assets and liabilities creates what is known as the duration gap, a structural imbalance that intensifies the insurers' exposure to interest rates fluctuations. Most insurers are subject to a negative duration gap (liabilities longer than assets) and therefore exposed to drops in interest rates. When subject to a positive gap (possibly because of the hedging strategies), they are exposed to increases in interest rates, that can also trigger the "mass lapse scenario" (a massive withdrawal of liabilities from the policyholders, encouraged to look for higher yields in alternative financial instruments).

The SII framework allows the insurance companies to either adopt the Standard Formula (SF: a prescribed list of rules and methodologies) or the Internal Model (IM: a customized and self-defined model tailored to the company's specific risks) to calculate the SCR. When it comes to the IR risk, while the SF applies a "sort of" parallel shock to the interest rate curve, the IM usually handles changes in yields level, slope and curvatures. Beyond that, IM often integrate cross factors able to link changes in interest rates to a dynamic policyholder behaviour (response to these changes), allowing for a comprehensive modelling. The SF, originally designed to provide a simplified and consistent approach for calculating the SCR, has revealed several structural limitations as market conditions have evolved: relative changes, proportional to the base rate, fail at both measuring appropriate absolute movements when the base rate is close to zero and at treating negative rates as well as capturing the magnitude of increases in rates as the relative shock is applied uniformly across maturities.

In response to these challenges, EIOPA proposed in 2020 a new stressing methodology, validated through extensive backtesting against historical data, that blends additive and relative shocks ([1]). This methodology is based on a relative displaced approach, that temporarily shifts the base rates upward, enabling a relative stress to be always applied, and then shifts the result back by the same amount. A floor condition ensures that the stressed rates are not smaller than -1.25%. The concepts of the EIOPA's proposal have been accepted by the European Parliament and the Council in December 2023 (agreement reached on the amendments to the SII text regulation) and published in April 2024. Compared to the current, the alternative formulation usually provides bigger downward shocks than upward shocks and is still not able to capture the sharp increase in rates that occurred in December 2022.

3. Stressing Models

The calibration of interest rates shocks relies on the analysis of the historical behaviour of interest rates, complemented with a view of the current and future market environment, to cover scenarios that may not have yet occurred. Shocks, or stresses, can be defined and applied to various types of rates, such as zero, forward, or par swap. This piece of work considers zero rates, to allow for an immediate comparison to the SF calibration. Another key decision is to determine whether to measure shocks on the long-term from a market perspective (to hedge on an economic basis) or a regulatory perspective (to hedge the SII capital requirements) – this piece of work focuses on the former.

Zero market rates are derived from traded market swap par rates through a process known as "bootstrapping", based on the underlying assumption of linear interpolation of the missing (or illiquid) maturities. This study combines the analysis of three alternative stressing models together with the usage of three different time series of changes in rates, examining a total of nine cases.

The advantages and limitations of the three stressing models are as follows:

- Absolute Stresses: simple to implement and to understand, are on the one hand stable in low rates environments, avoiding large proportional changes, but on the other can produce unrealistic negative stressed rates, requiring adjustments like a flooring;
- Relative Stresses: capture the relationship between rate levels and volatility, ensuring scale sensitivity and reflecting the proportional changes typically observable in financial markets, but break for negative or near-zero yields and tend to underestimate shocks in low-yield environments;
- Relative Displaced Stresses: combine the proportional nature of relative stresses with the ability to handle low or negative yields and provide a flexible calibration through the displacement parameter but require an expert judgment and computational overhead.

The relative stress is discarded given its limitations, while the absolute and relative displaced are both further considered. It is worth noting that practitioners used to describe rates movement by adopting a simple relative approach till 2015-2019, when the rates became more and more negative. They then moved to the relative displaced representation, that worked well until 2023, when it was questioned by the sudden rates increase (about 300 bps in a year). Given the displacement levels adopted at that time (from 1 to 4%), the relative displaced approach was no longer able to produce such a big shock starting from low-negative yields and the absolute stress became to be considered.

The advantages and limitations of the three time series of rates changes are as follows:

- annual non-overlapping exhibits a low autocorrelation, but offers few data points;
- annual overlapping exhibits a high autocorrelation, providing an adequate number of data points;

- monthly non-overlapping combines the advantage of exhibiting a low autocorrelation together with the provision of an adequate number of data points, though requiring an additional complexity for a proper annualization of the values.

The autocorrelation is assessed by the means of the Durbin-Watson test and the Ljung-Box Q-test.

The annual non overlapping dataset is discarded not being suitable to perform a sensible distribution fitting exercise on its few data (see next Chapter), while the other two are both further considered. It is worth noting that the annual overlapping dataset provides bigger changes in rates than the monthly non overlapping one and its results are sensible to the choice of the starting month (i.e. a time series of changes from Jan to Jan produces different results to that covering the period June-June). Indeed, May, June and July have historically been months characterized by events that have significantly influenced the changes in zero rates, such the global financial crisis in 2009, the European sovereign debt crises in 2012 and the sudden rates spike in 2022 and 2023. Compared to the monthly non overlapping, the annual overlapping dataset also produces changes in rates that are much less symmetrical, sometimes being significantly left or right skewed.

4. Principal Component Analysis

4.1. Stress decomposition into PCs

A proper measurement and hedging of the IR risk must look at each relevant maturity of the liabilities cash flows. The stresses (and changes in zero rates) should be defined for each quantile and tenor of the yield curve, but this would require many risk factors (one per tenor), hard to be managed by the insurance companies in their processes. A handy solution is given by the usage of the Principal Components (PCs), that reduce the problem dimension and can be also associated to an economic meaning of the changes in yield curves. The covariance matrix of the original dataset is factorized following the eigenvalues equations, and the PCs are derived by projecting the original dataset on the eigenvectors, resulting in a transformed dataset, whose columns represent the PCs realizations.

The first three PCs (that correspond to a change in level, slope and twist of the base yield curve) are proven to explain about 99% of the total variance of the original datasets tested. Characterised by different peculiarities in the pairs (stress methodology, dataset), the PCs show common trait, with PC2 being that most similar to a normal distribution, PC1 showing the highest variance (i.e. wider range) and PC3 showing the lowest variance and highest skewness. The right skew of PC1 and PC3 distributions suggests the right tail (interest rates increases) is fatter/longer than the left one (interest rates decreases), meaning more extreme values on the higher end of the distribution. These come from the sharp rate increase that occurred in 2022-2023, that put into question the appropriateness of the relative displaced model, not capable of describing such big changes in short maturities.

4.2. PCs Distribution Fitting and Backtesting

A fitting exercise is carried out on the historical first three PCs to allow for a replication of the actual values and for a forecast of further tails. Five distributions are considered (Normal N; Student's t with different degrees of freedom, T-7, T-5, T-3; Exponential Generalized Beta of the Second Kind EGB2) to describe different behaviours and their optimal parameters found by both applying the Maximum Likelihood Estimation (MLE) and the Least Squares (LS) methods. The goodness of fit is assessed by visually considering the Probability Density Functions (PDFs) and Quantile to Quantile (Q-Q) plots, and by numerically comparing the log-likelihood (LL), Akaike and Bayesian Information Criterions (AIC, BIC) and the results of the statistical tests on the comparison of two distributions (Kolmogorov-Smirnov, KS, and Anderson-Darling, AD). The Root Mean Square Difference (RMSD), that measures the distance between the historical data and the corresponding percentiles of the fitted distributions, is also compared and a back-testing exercise performed to check the capability of the fitted distributions to replicate the historical data. The results obtained by using the MLE and LS estimation methodologies are very close and MLE is taken as reference.

Being characterized by a higher volatility and more extreme values, compared to the monthly non overlapping dataset, the annual overlapping one is more difficult to fit, resulting in higher RMSDs. For this reason, the monthly non overlapping dataset is chosen and, overall, the most preferable distributions are selected to be the T-3 for the absolute shock model and the EGB2 for the log relative displaced model. For the monthly dataset T-3 provides the best fitting for PC2 and PC3 and is able to capture the extreme PC1 down empirical realizations. For the annual dataset, EGB2 provides the best fit for PC2, an acceptable under-estimation on PC1 extreme realizations (visible from the Q-Q plots and the backtesting exercise) and is still able to capture the body of the distribution for PC3.
While the new SF formulation still fails at capturing the massive increase in zero rates observed between 31.12.21 and 31.12.22, the models defined via PCs are able to replicate it, though using high quantiles:

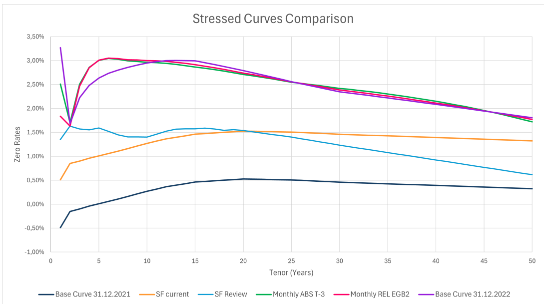


Figure 1: Stressed yield curves at 31.12.22 and 1 year annual historical movement

Quantiles					
			PC1	PC2	PC3
31.12.2021	up	Monthly ABS T-3	99.06%	99.35%	1.48%
		Monthly REL EGB2	99.97%	99.95%	6.40%

Table 1: Quantiles for PC1, PC2 and PC3.

The following chart compares the absolute stresses to the zero rates calculated on the 22YE yield curve, provided by the SF in both its current and new IR risk formulation, versus those defined by applying a combined stress scenario for both the absolute and log relative displaced (displacement 5%) models.
The PCs quantiles of the combined scenarios corresponds to those that produce a drop in own funds measured at the quantile 99.5% (see section 5).

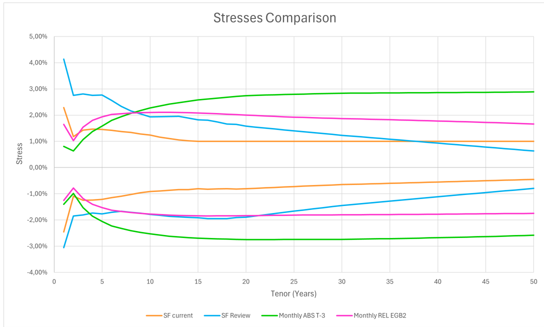


Figure 2: Stresses Comparison at 31.12.2022

The alternative SF IR shock (light blue line) provides a stronger stress than its current definition (orange line), but also a much higher stress than those defined by the combined PCs scenarios (green and purple lines) for very short tenors (T=2 and 3 years), whose variabilities are tiny, turning into small eigenvectors and small stresses. The monthly combined stress (green line), based on a T-3 distribution, is wider than the log-relative stress (pink line), based on the EGB2 distribution because of the fatter tails the T-3 is characterized by; both stresses are higher than the current and alternative EIOPA formulation.
The shock calibration shall be capable of replicating the historical changes observed in rates, leaving room to those that have not yet occurred: this feature is assessed by the means of a backtesting exercise, carried out on each PCs, one at a time. The historical zero rates movements given by that PC only are compared to those resulting from an application of the same PC's values corresponding to the 0.50% and 99.50% quantiles of its fitted distributions. An example is provided for PC1, fitted on the EGB2 distribution:

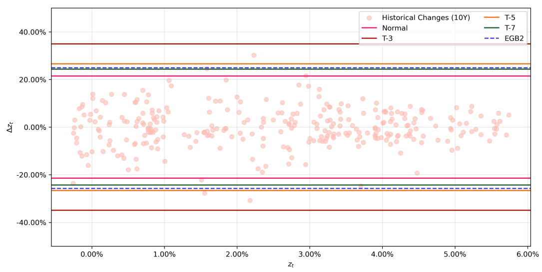


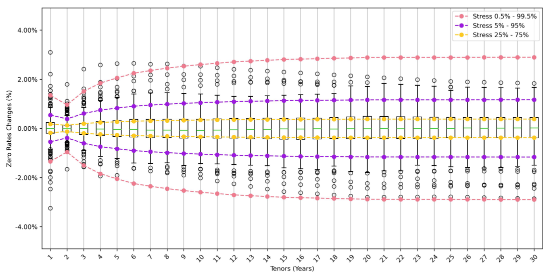
Figure 3: Backtesting for the Monthly Non-Overlapping Dataset $t = 10$years.

The number of breaches for two reference tenors is collected in the following table:

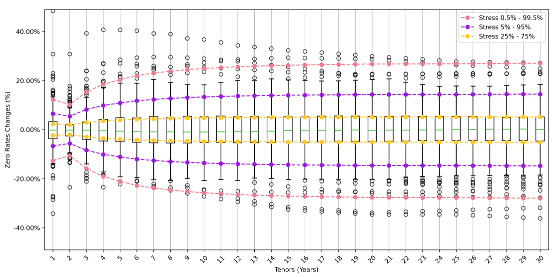
Breaches			
		Monthly Abs T-3	Monthly Rel EGB2
PC1	5Y	0	3
Breaches	10Y	0	3
PC2	5Y	0	4
Breaches	10Y	0	4
PC3	5Y	5	8
Breaches	10Y	5	8

Table 2: Breaches detected during backtesting.

The stress level chosen to assess the back-testing assumes to register a breach 1 every 200 years (therefore 2 breaches considering both stress sides) and the historical time series cover a period of 20 years. Based on this consideration, a number of breaches higher than 2 indicates that the fitted distribution is not even capable of correctly replicating the history on its extreme behaviours (tails). This rationale is sensible in a combined PC shock but becomes less prescriptive when stand-alone PC shocks are assessed. On the other hand, no breaches could raise red flags too, as the tails of the theoretical distribution may be too fat compared to the historical ones (i.e. too high stresses), as it happens for the T-3 distribution.
Finally, a visual comparison is provided to benchmark the empirical changes in rates (visualized though the usage of box-plots) to some relevant up and down PC1 stresses:



(a) Abs T-3



(b) Rel EGB2

Figure 4: Empirical Changes and Stresses.

In both cases the yellow lines (25%-75% quantiles of the theoretical distributions) well align with the same empirical quantiles (boxes), the purple lines (5%-95% quantiles) cross the outliers and the pink lines (0.5%-95% quantiles) show the same breaches already highlighted by the backtesting analysis, especially for the very short term.

5. Impact Assessment

The differences in calibrated stresses are then translated into differences in SCR (capital charges), by providing an impact assessment on the risk charge stemming from both the regulatory stresses and those calibrated in this piece of work. In the same exercise, the distribution of changes in Own Funds (OF) derived from the calibrations are compared to those derived from historical changes in rates, as an additional backtesting. Simplified assets and liabilities portfolios are combined to offer testing cases with different liabilities durations (short, medium, long) and different duration gaps (negative, positive), such that both the down and up scenarios are biting and and allowing for a holistic assessment of differences.
The IR capital charge, is derived:

- under the SII Standard Formula (both current and revised formulations) by calculating the change in OF in the two regulatory "down" (0.50%) and "up" (99.50%) scenarios;
- under the stress risk defined in the previous section, as the change in OF corresponding to the 0.50% / 99.50% quantiles of the change in OF distribution, derived by applying a large set (10.000 trials) of variations that cover changes in level, slope and curvature of the yield curve, by sampling independent 3 PCs values from their fitted distributions.

To prevent the sampling error and to ensure stable outcomes, the seed of the random number generator is fixed; to ease the compatibility between the absolute and relative shocks models (based on different distributions) the random values firstly sampled from a uniform distribution and then converted into PCs realisations by inverting their CDFs.
The following scatterplots chart the simulated asset, liabilities and own fund values against the PC1 ones, for both the negative (ndg) and

positive (pdg) duration gap cases of the insurance portfolio with medium duration. As expected, the values of assets and liabilities decreases as PC1 increases (higher discounting) and the speed pace is higher for higher durations, resulting into different interest rates exposures (i.e. biting scenarios) as reflected in the own funds movements:

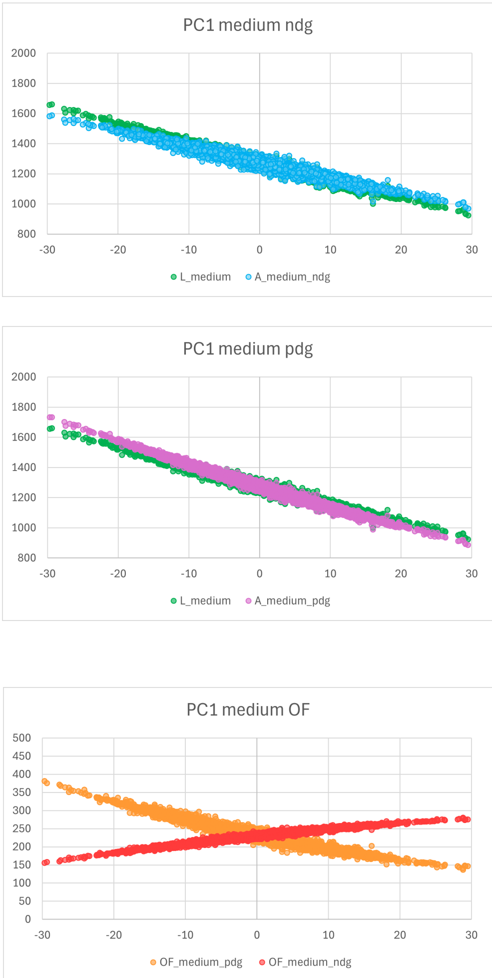


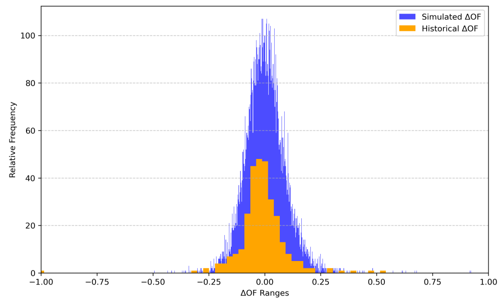
Figure 5: Movements of Assets, Liabilities and Own Funds as function of PC1 (Relative Displaced Shock Formulation).

The impact assessment confirms that the alternative proposal for the SF IR shocks is conservative compared to the current, providing a capital charge of about -17% vs -5% (IR down) and -14% vs -13% (IR up) for medium maturities. The only case where the current calibration provides a higher SCR is that with the upward shock biting and long durations. Indeed, in this example, all long term cashflows are subject to an alternative upward stress weaker than the current, whose values are floored to a minimum of 1% (please refer to Figure 2). The capital charges resulting from the PCs applications are higher than the regulatory ones (both current and alternative) and well compare to those of the internal models reported in the EIOPA's 22YE comparative study on market and credit risk modelling ([2]). The absolute stress formulation, based on PCs T-3 distributed, appears to be too conservative, due to the fat tails of the distributions, declaring the log-relative displaced (displacement=5%) shock formulation, whose PCs are fitted on EGB2 distributions, to be the most suitable choice.

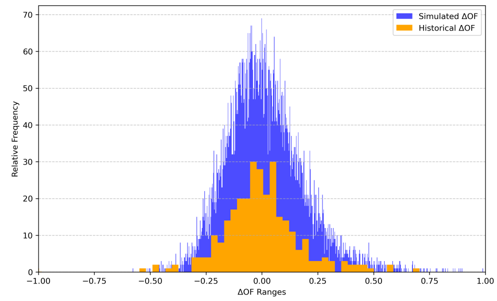
Capital Charges			
	Short D.	Medium D.	Long D.
SF Current 22YE Study	6%	9%	12%
SF Current Mean	6%	9%	12%
SF Review Mean	10%	16%	14%
IM Min	8%	14%	23%
IM Max	11%	18%	29%
EGB2 Mean	14%	25%	33%

Table 3: Comparison between Capital Charges.

It must be noted the simplified modelling of the insurance portfolios adopted in this piece of work replicates the same results reported by EIOPA in its comparative study, for the current SF IR shock definition, allowing for a comparison of the calibrations derived in this thesis to those adopted by the users of Internal Models.
The capital charges resulting from the EGB2 log-relative displaced shock formulation appear to be higher than (but still comparable to) those of the internal models. This is driven by the interest rate up capital charge, fitted to a time series that, differently from the results reported in the comparative study, embeds year 2023 too. Indeed, as mentioned, the models based on PCs decomposition quickly adapt to changes in the market environment, such the spike in zero rates observed in the last years.
The backtesting exercise is summarized by the following histograms, that compare the relative changes in OF distributions calculated by the means of historical and simulated PCs (EGB2 distribution) for the portfolio with medium (10 years) duration, in both cases with negative (ndg) and positive (pdg) duration gaps:



(a) Medium D. ndg



(b) Medium D. pdg

Figure 6: Comparison of Relative Changes in Own Funds Distributions.

It can be noted that the stresses simulated by the log-relative displaced models generate changes in OF (blue bars) capable of covering those that have been occurred in history (orange bars).

6. Conclusions

Under the SII framework, insurance and reinsurance companies are expected to hold a capital requirement that serves as a cushion to preserve their solvability under a combined stressed scenario, calibrated to occur once every 200 years. This corresponds to a Value at Risk (VaR) calculation with a 99.5% confidence level. Among the risks considered, the Interest Rate (IR) one plays a crucial role, expressing the mismatch between the durations of assets and liabilities — a concept commonly referred to as the duration gap.

This study explores the modelling of the IR risk by comparing the current and alternative Standard Formula (SF) settings to alternative stressing methodologies, usually adopted in the Internal Models (IM). While SF provides a standardized approach, IM adapt more rapidly to market conditions and capture a wider range of yield curve movements, usually defined as changes in level, slope and twist of zero rates by the means of the first three principal components. The main limitations of the current SF calibration are its inability to capture meaningful absolute movements when the base rates are close to zero, to handle negative rates and to capture historical magnitudes of rates increases. EIOPA made a big step forward by setting an alternative methodology that improves most of the current limitations, but however still fails at capturing the extreme rate increase that occurred in 2022.

The alternative stressing methodologies are calibrated to the historical behaviour of zero rates. Three models (absolute, relative displaced and relative) are explored in three different time series (annual non-overlapping, annual overlapping and monthly non-overlapping), for a total of nine cases under analysis. The relative model is discarded for its acknowledged limitations, as well as the annual non-overlapping dataset, due to its limited number of observations. The relative displaced model, chosen by EIOPA too for its new IR risk formulation, combines the proportional nature of changes in rates (a linear dependency is observable between the magnitude of the rates and the volatility of their changes) with the ability to handle low and negative yields, but may fail at capturing big jumps in short tenors. The annual overlapping dataset provides bigger changes in rates than the monthly non-overlapping one and its results are sensible to the choice of the starting month.

The Principal Component Analysis is adopted to reduce the dimensions of the problem (from 50 nodes, tenors of the yield curve, to 3 principal components, referring to changes in level, slope and twist of the zero rates). The resulting eigenvectors and empirical distributions of the PCs are extensively discussed: albeit being characterized by different peculiarities in the pairs (dataset, stress methodology) the PCs show common traits such as PC2 being that most like a Normal distribution, PC1 being that with highest variance and PC3 being that with highest skewness and lowest variance. A fitting exercise is carried out on the historical first three PCs to allow for a replication of the actual values and to forecast further tails. It establishes the best mathematical distributions that describe the empirical ones, discarding the an-

nual overlapping dataset for its poorer goodness of fit and choosing T-3 for the absolute shock formulation and EGB2 for the log-relative displaced model. Several statistical analyses and visual inspections are performed to come to the choice (Log-Likelihood, Akaike and Bayesian Information Criterion, Kolmogorov Smirnov (KS) and Anderson Darling (AD) tests, Root Mean Squared Distance and Q-Q plots).
Finally, the differences in stress calibrations are translated into differences in capital charges, by performing an impact assessment on several test cases that cover different portfolios durations and duration gaps. The log-relative displaced model, whose PCs are fitted on EBG2 distributions, seems to be the most suitable stressing model, with results comparable to those of the internal models reported in the EIOPA's 22YE comparative study on market and credit risk modelling. The backtesting exercise confirms the capability of the model to replicating the stresses historically occurred, when valued in terms of capital charges. It has to be noted that the definition of a simplified insurance portfolio by the means of its underlying assets and liabilities cash flows allows the analysis of several test cases and allows for a perfect replication of the capital charges EIOPA reports on its comparative study for the SF IR risk formulation; through stochastic simulations of the PCs values, sampled from their fitted distributions, it provides a comparison between the capital charges reported for the internal model firms participating in the study and those resulting from the analysis of this thesis.

7. Acknowledgements

I would like to express my sincere gratitude to my supervisor, Prof. Daniele Marazzina, for allowing me to pursue my master's thesis in this rich and dynamic area. My heartfelt thanks also go to my co-supervisor, Prof. Silvia Dell'Acqua, for her invaluable support and unwavering availability throughout this research. Her continuous feedback, push and constructive suggestions have been the foundation of this work.

References

[1] European Insurance and Occupational Pensions Authority (EIOPA). Opinion on the 2020 review of Solvency II, 2020.

[2] European Insurance and Occupational Pensions Authority (EIOPA). YE2022 Comparative Study on Market and Credit Risk Modelling, October 2023. Accessed: February 22, 2025.