Modeling Rate Relativity for Excess Casualty Insurance

Montana State University

5/6/2024

Project Sponsor:

Willis Towers Watson (WTW) – Insurance broking, risk management, consulting

Project Contact:

Jonathon Drummond – Head of Broking North America

Team Members:

Ethan Carney, Alexander Cimini, Quinlin Gregg, Nicholas Starring

Advisors:

Dr. Faraz Dadgostari
David Brower

Department of Agricultural Economics and Economics
Department of Mechanical and Industrial Engineering
Montana State University

Executive Summary
This project, conducted in collaboration with Willis Towers Watson (WTW), developed a predictive model for analyzing rate relativities and premiums across excess casualty insurance towers. Using a dataset of anonymized client towers, the team employed statistical models to model the relationships between layer limits, attachment points, and premiums. The model captured key non-linear trends in tower structures and achieved a strong predictive accuracy. The final deliverable was a broker and underwriter-friendly Excel tool with an R backend, enabling dynamic scenario testing and pricing analysis. This project highlights how these modern statistical methods can enhance pricing strategies in a dynamic excess casualty insurance market and provide a foundation for ongoing model refinement through expanded data sources and market integration.

Table of Contents

List of Figures

List of Tables

Introduction

Willis Towers Watson (WTW) is a global advisory, broking, and solutions company specializing in risk management, insurance broking, human capital consulting, and investment solutions. As the property and casualty insurance market evolves, understanding and pricing excess casualty towers has become increasingly complex due to shifts in loss trends, market capacity, and pricing volatility.

This project focuses on modeling and predicting rate relativities and premiums across excess casualty insurance towers. Using a dataset of 559 anonymized WTW client towers, we analyzed the structure and pricing patterns across primary, umbrella, and excess layers. Special attention was given to how variables such as layer limit, attachment point, and industry class influence relativity trends throughout the tower.
The primary objectives of this project are:

- Establish a data-driven understanding of how excess casualty towers are priced across different layers.
- Analyze market pricing patterns and layer-specific relativities to identify structural trends.
- Develop an Excel-based predictive model using statistical techniques such as Generalized Additive Models (GAM), quantile regression, and linear regression.
- Create a broker- and underwriter-friendly interface to enable scenario testing and pricing evaluations in real time.

The project scope includes a detailed analysis of historical tower structures, the construction of a flexible predictive model, and the development of an Excel tool for practical application. Broader pricing forecasting and operational implementation are out of scope, with the focus limited to enhancing understanding of current market pricing dynamics in excess casualty insurance.

Literature Review

Learning Sessions
As part of the initial phase of the capstone project, the team engaged in a series of structured learning sessions with industry professionals. These sessions served as the foundation for developing a comprehensive understanding of excess casualty insurance, including how towers are constructed and evaluated in practice. Each session was accompanied by an individual assignment designed to reinforce core concepts and introduce practical applications.

The first learning session was led by Mike Williams, North America Casualty Leader in Public Sector and Education. His presentation provided an introduction to the structure of excess casualty insurance, beginning with the composition of the primary layers—including Workers' Compensation, General Liability, Auto Liability, and the base umbrella policy. The session focused on how these foundational coverages interact with excess layers in a tower. The associated assignment involved analyzing sample quote data to extract key information such as limits, deductibles, premiums, surcharges, retention, and TRIA (Terrorism Risk Insurance Act) premiums. This exercise helped familiarize the team with the terminology and layout of quote sheets used in real-world underwriting.

The second set of learning sessions was delivered by Len Graziano, Strategy and Execution Leader, and Ned Kaufman, Excess Liability Industries Leader. These sessions dove deeper into the mechanics of tower construction and pricing strategy. The discussions focused on identifying key metrics within a tower structure, understanding market pricing dynamics, and evaluating how the structure affects total premium cost and efficiency. The first part of the assignment involved reviewing a sample "dummy" excess tower and identifying general observations, opportunities for restructuring, and ways to improve the clarity and utility of the financial exhibit in Excel. The second part consisted of reviewing a zip file of anonymized binder quotes for a client's excess program. Each team member was responsible for filtering through the quotes to build an individualized financial exhibit, incorporating components such as surplus lines taxes (SLT), estimating punitive damage wrap costs, and identifying more cost-effective restructuring options.

These learning sessions were critical in bridging academic knowledge with real-world industry practices. They informed both the conceptual understanding and technical modeling efforts used throughout the remainder of the project.

External Research and Industry Literature

In addition to practitioner insights, the project was supported by academic research on modern modeling techniques applicable to insurance pricing.

Hastie and Tibshirani's foundational work on General Additive Models provided theoretical underpinnings for the modeling approach. Their research emphasized the advantages of smoothing techniques in fitting flexible curves to data, which aligns closely with the project's need to model excess casualty layer relativities across a wide range of attachment points and limits (Hastie & Tibshirani, 1986).

Further, recent applications of GAMs in insurance modeling, such as the work by Henckaerts et al. (2017), demonstrated the practical advantages of additive models in claims reserving and risk forecasting. Their work showing how it can capture nonlinear effects in data informed the decision to use GAMs over more rigid linear models, ensuring greater adaptability to what the model may need to do.

All together, these academic resources provided a strong theoretical foundation for the statistical modeling choices made through the project.

Project Management

Project Management Methodology and Client Involvement

The team followed an iterative project management approach, emphasizing early learning, regular client feedback, and incremental development. Early phases focused on knowledge acquisition through structured assignments and learning sessions with WTW professionals. Once the project objectives and data were fully defined, the team transitioned into modeling and tool development, maintaining continuous collaboration with the client through weekly meetings.

Tools Used

To support the development and management of the project, the team utilized:

- Excel: Financial exhibit construction, scenario testing, model interface development.
- Python: Data cleaning, formatting, and automation tasks.
- R: Statistical modeling (GAMs, quantile regression, ANOVA testing) and data visualization.
- Google Drive: Document sharing and version control for group deliverables.
- GitHub: Code collaboration and model version control during the R and Python development phases.

These tools allowed the team to maintain flexibility and transparency while ensuring efficient collaboration across different technical tasks.

Project Scheduling, Milestones, and Deliverables

Table 1: Project Milestones Timeline

Phase	Timeline	Activities	Deliverables
Learning and Research	Weeks 1 - 9	Learning sessions with industry professionals	Assignment submissions, Knowledge base
Data Acquisition and Preparation	Weeks 10 - 11	Cleaning of the tower dataset & initial exploration	Cleaned and validated dataset
Model Development	Weeks 11 - 12	Statistical Modeling	Working predictive model
Tool Development	Weeks 12 - 13	Building Excel front	Excel Pricing tool

		end; Integrating R through VBA	
Final Testing and Presentation	Weeks 12 - 13	Scenario testing, final validations, and preparation of deliverables	Final Model, tool, documentation, client presentation, & poster presentation

Methods

Design Approach / Methods Used
The team employed a combination of statistical modeling, machine learning, and Excel-based simulation to analyze rate relativities across excess casualty insurance towers. The primary modeling technique was the Generalized Additive Model (GAM), a flexible, non-parametric method capable of modeling nonlinear relationships and interactions. Additional models used included quantile regression and linear regression.

The mathematical structure of a Generalized Additive Model (GAM) can be expressed as follows:

$$(1)\ g(E[y_i]) = \beta_0 + f_1(X_{i1}) + f_2(X_{i2}) + ... + f_p(X_{ip})$$

Where:
Y_i: is the response variable for the *i*-th observation
$E[Y_i]$: is the expected value of the response
$g(.)$: is the link function that connects the expected value of the response to the additive predictor
β_0 is the intercept
$f_j(X_{ij})$: are smooth functions that describe the potentially non-linear effects of the *j*-th predictor on the response
p: is the number of predictors

The second approach we used was a hybrid modeling approach to estimate lead ratios across excess casualty layers. First, we fit quantile regression models of the form:

$$(2)\ R = \frac{a}{M+b} + \varepsilon,\ R_i = \frac{ppm_i}{ppm_0}$$

where R is the lead ratio, M is the dollar midpoint of the layer, and a and b are regression coefficients.

To extend beyond observed layers, a scaling model was used to adjust predictions upward or downward in the tower. This took the form:

$$(3)\ log(S) = log(R/q(M)) + \varepsilon$$

where S is the scaling factor relative to the top known layer, M is the midpoint, and q(M) is the quantile prediction from eq (2).

Justification for Methods and Tools

These methods were selected to provide both flexibility in modeling and interpretability for the client. The GAM accounts for nonlinear effects and interactions between key variables like layer limit and attachment point. Excel provided a transparent interface that brokers and underwriters could easily use in their daily work. These tools matched both the complexity of the data and the usability goals of the final product.

Combining quantile predictions with scaled extrapolations provided a flexible, data-driven framework for forecasting lead ratios and quantifying uncertainty in unpriced layers. The quantile regression captured the full distribution of lead ratios at each layer, rather than just the mean. By then modeling how each quantile scaled with layer, we were able to generalize this structure to higher, unpriced layers using observed relationships in the data.

Data Collection

Data Needs Analysis

The project required granular, layered data on excess casualty programs, including attachment points, limits, and premiums. Categorical variables such as industry classes, region, and industry description would be helpful as well. Ideal supplementary data would have included lead layer limits for WC, AL, GL, and total client revenue to improve modeling precision.

Sources and Acquisition

All of the data was sourced from WTW in the form an an anonymized Excel file containing 559 excess casualty tower rows. Each tower included detailed information across primary, umbrella, and excess layers, along with attributes such as SIC codes, industry verticals, and geographic regions.

Data Collection / Cleaning Challenges

The primary challenge was ensuring data integrity, particularly in standardizing inputs across the dataset. Significant effort was required to clean and validate the data, a process done collaboratively with WTW team member Matthew Presifilippo. His involvement was crucial for confirming and correcting irregular entries.

Data Analysis

Techniques and Tools

The initial dataset consisted of 559 anonymized excess casualty towers provided by WTW. Each team member took a unique approach to cleaning the data, and the results were reconciled until a final dataset was agreed upon. Excel, Python, and RStudio were some of the tools utilized in the cleaning step. The GitHub repository was used for version control and cleansing iterations.

Data Cleaning

Before analysis could begin, the data required significant review and cleaning to ensure that only valid towers were included. The cleaning step was essential to ensure the accuracy and reliability of the subsequent analysis and modeling.

Towers were excluded from the analysis based on the following criteria:

- Missing or incomplete layer information (e.g., attachment points, limits, or premiums).
- Towers with structural inconsistencies, such as missing primary layers or illogical stacking of excess layers.
- Programs that could not be reasonably interpreted or aligned with standard market practices.

Final Data

The cleaned dataset consisted of 160 valid towers. Each tower was organized into multiple layers, with key fields including Client IDs, Umbrella Layers, Excess Layers, Industry Classifications, Price Per Million, and Rate Relativities.

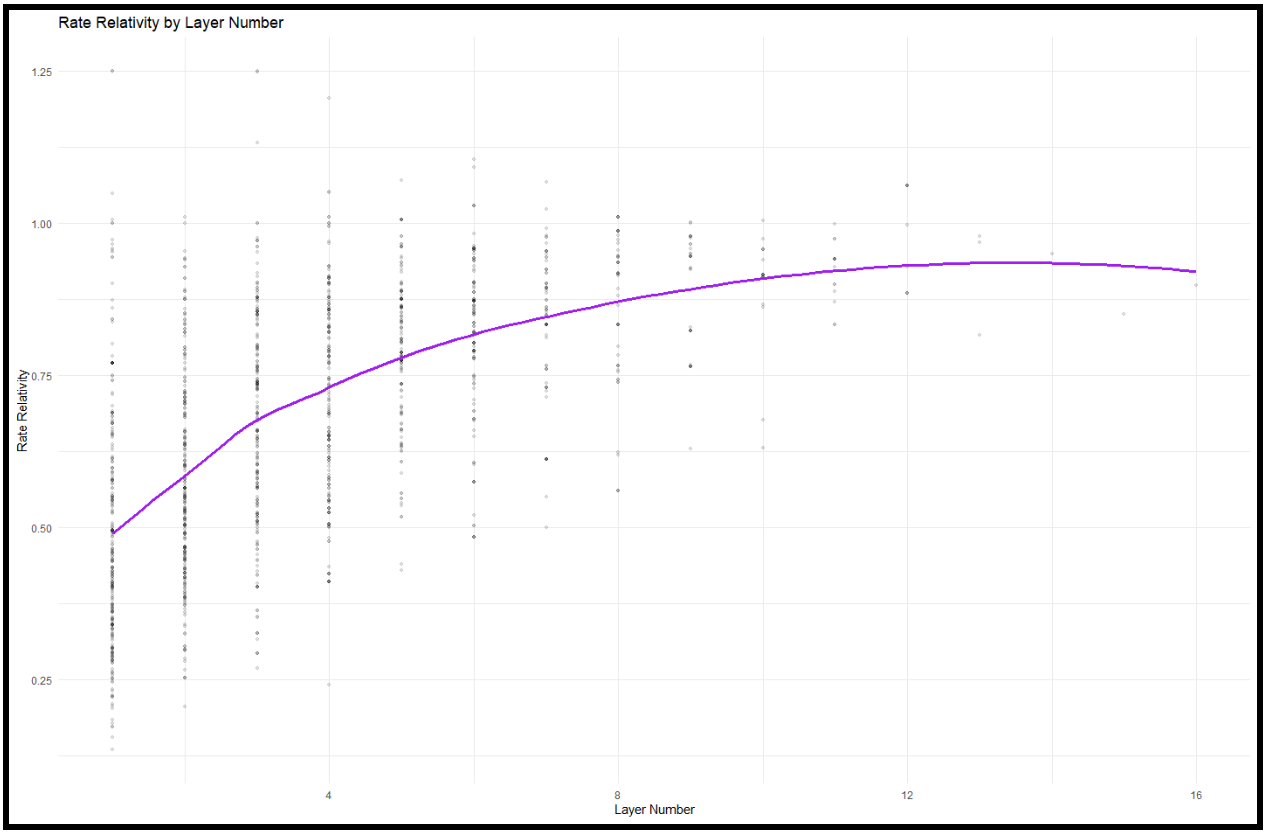


Figure 1: Rate Relativity by Layer Number

This figure shows the overall relationship between rate relativity and layer number across the full dataset. As expected, rate relativity generally increases as layer number rises, reflecting the added risk and uncertainty priced into higher excess layers. The curve flattens at higher layers, suggesting that beyond a certain point, additional increases in layer number result in diminishing adjustments to rate relativity. This pattern highlights common market behavior in structuring excess casualty towers.

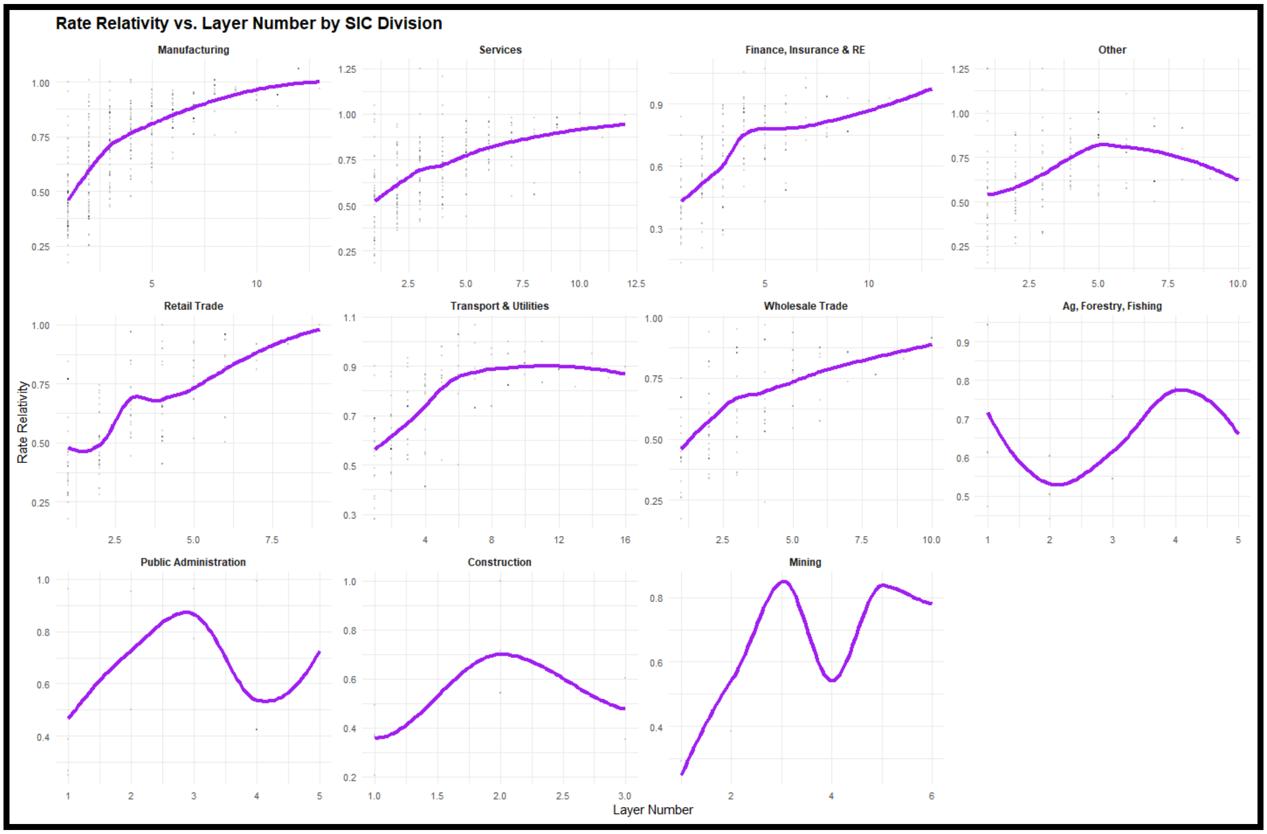


Figure 2: Rate Relativity vs. Layer Number by SIC Division

Figure 2 shows how rate relativities evolve as the layer number increases across different SIC industry divisions. In most industries, relativity generally rises with layer number, reflecting the decreased risk priced into higher layers of the tower. However, in divisions with fewer observations, the average curve tends to overfit the limited data, leading to noisier and less reliable trends.

Project Implementation

Proposed Solution

The proposed solution leverages current market data and statistical models to estimate rate relativities and premiums across different layers of an excess casualty tower. The model outputs allow brokers and underwriters to benchmark pricing against expected trends and identify potential structural inefficiencies within the tower design.

System Architecture & Design

The system was designed with a modular structure, separating the statistical modeling engine from the user-facing interface.

- The backend consists of an R-based predictive model optimized for rapid calculation and output formatting.
- The frontend is an Excel-based interface that enables users to input tower structures, run predictions, and view outputs seamlessly.
- Integration between R and Excel is handled through VBA scripting and controlled file exchanges (CSV format), ensuring accessibility without requiring users to directly interact with code.

Development Process
Development began with the construction of a baseline GAM model in R. Emphasis was placed on ensuring the model could compute predictions efficiently and export results in a format compatible with Excel.

The process flow is as follows (Figure 1):

1. The user enters the excess casualty tower structure into the Excel workbook.
2. Upon command (clicking a button), the input data is saved to a CSV file within the project folder.
3. VBA then initiates a shell command that triggers the R script.
4. The R script reads the input CSV, runs the predictions, and outputs a results CSV file.
5. Once processing is complete, Excel automatically imports the results and displays the predicted relativities within the worksheet.

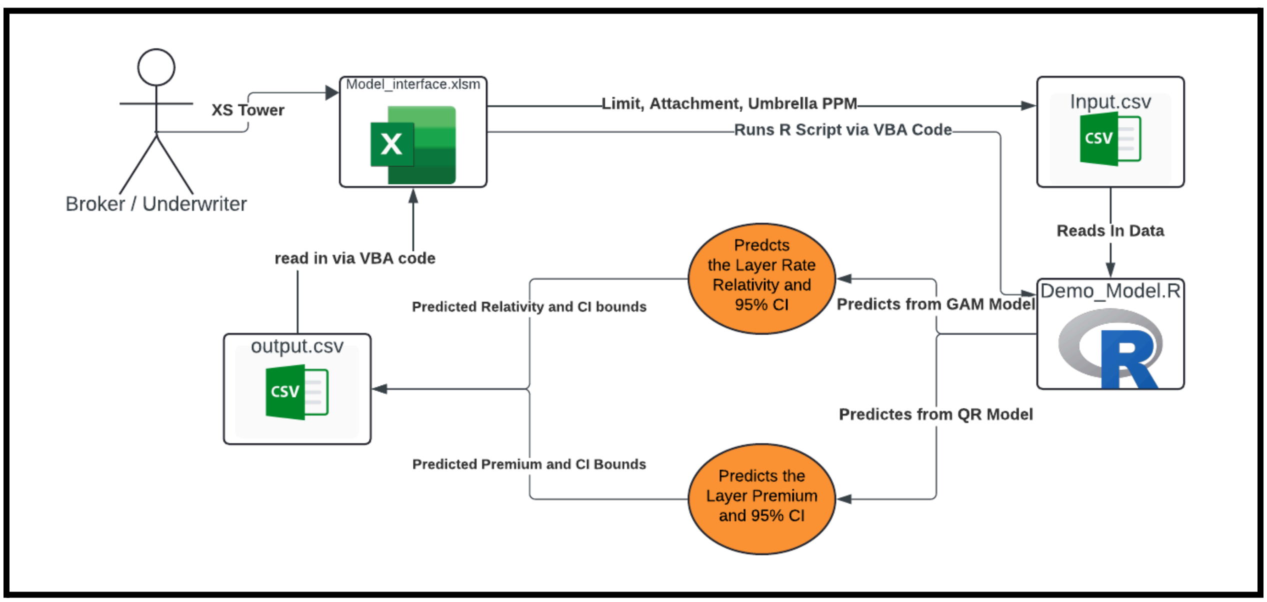


Figure 3: End-to-End Process Flow for Predictive Modeling Tool

This architecture allows users to run complex statistical predictions with minimal interaction and no technical barrier.

Implementation Plan

The model is designed for easy deployment within WTW broker workflows. Brokers and underwriters can utilize the Excel interface to quickly model client towers, benchmark current market pricing, and identify opportunities for restructuring.

Results and Discussion

Analysis of Results

The Generalized Additive Model (GAM) successfully captured key nonlinear relationships between layer limit, attachment point, and relativity. Across the training and testing datasets, the final model explained approximately 52% of the deviance and achieved a Root Mean Squared Error (RMSE) of 0.164 on the test set, indicating strong predictive performance.

The final equation from the GAM creation and analysis is as follows:

$$log(E[Relativity]) = \beta_0 + f_1(Layer\ Limit,\ Layer\ Attachment) + f_2(Umbrella\ PPM)$$

Where:

β_0: $Intercept$

f_1: $Smoothed\ Interaction\ surface\ for\ limit\ x\ Attachment,\ fit\ via\ tensor\ product\ splines$

f_2: $Smooth\ 1D\ Curve\ for\ Umbrella\ Price\ Per\ Million\ (PPM)$

During model development, numerous Analysis of Variance (ANOVA) tests were performed to evaluate different model specifications and assess the impact of adding various terms. Table 2 below highlights a selection of key comparisons to illustrate the model refinement process and demonstrate how the final structure was chosen.

Table 2: Highlighted ANOVA Tests

Model Specifications	Adj R-sq	Deviance Explained	Notes
Relativity ~ Layer Limit + Layer Attachment	-1.5	-236%	Linear terms only: poor fit
Relativity ~ s(Layer Limit) + s(Layer Attachment)	0.489	50.4%	Smooth terms only
Relativity ~ te(Layer Limit, Layer Attachment)	0.5	50.6%	Interaction via tensor product
Relativity ~ te(Layer Limit, Layer Attachment) + s(Umbrella PPM)	0.519	51.9%	Final selected model

Key insights from the model include:

- Relativity generally decreased as towers moved higher in attachment points and limits, aligning with market expectations (Figure 1).
- An inversion in relativity was observed for layers with limits above $50 million and attachments over $200 million, likely due to data sparsity in these regions (Figure 1).
- Umbrella Price Per Million (PPM) was identified as an important factor for refining relativity estimates in the lower layers (Figure 1).

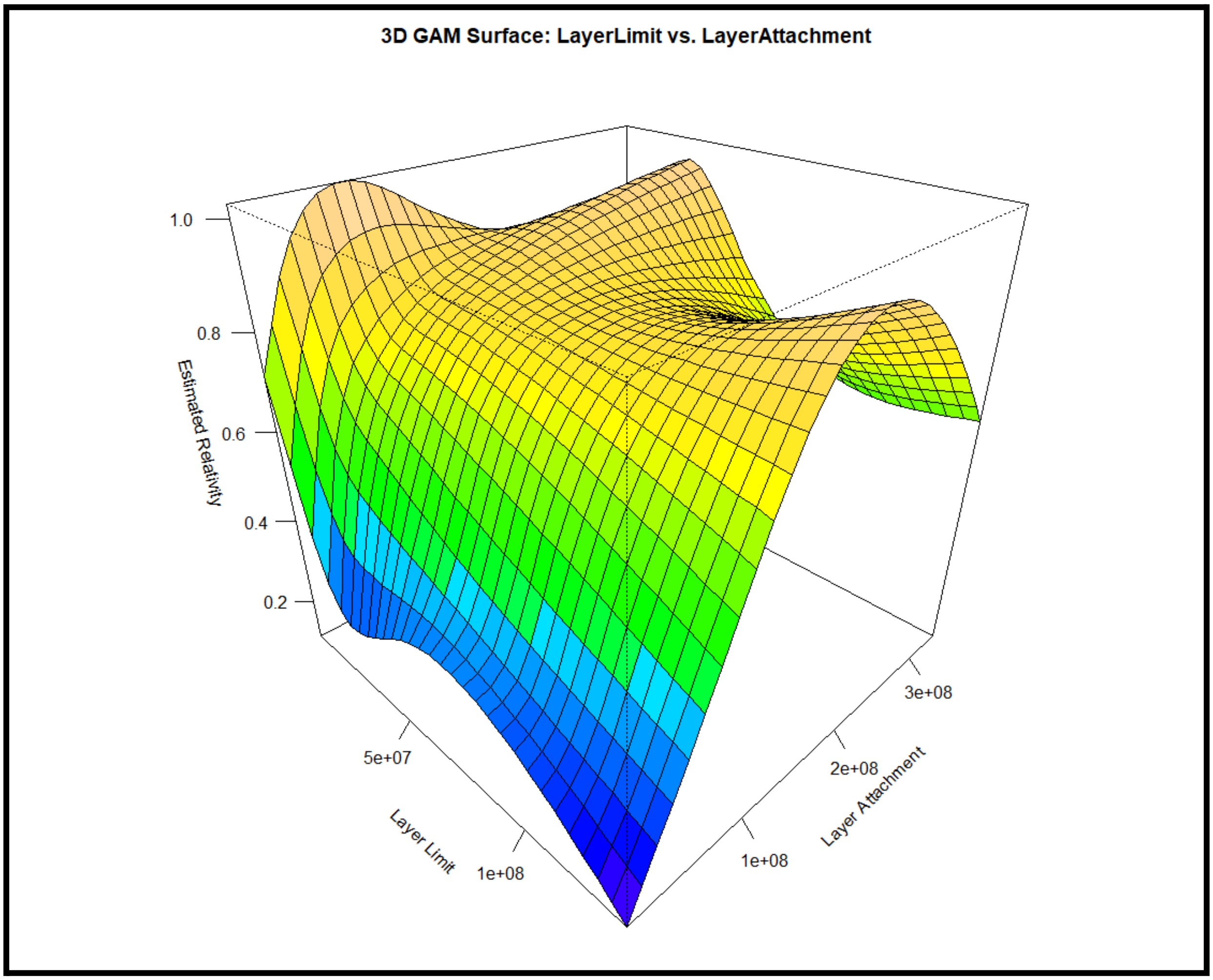
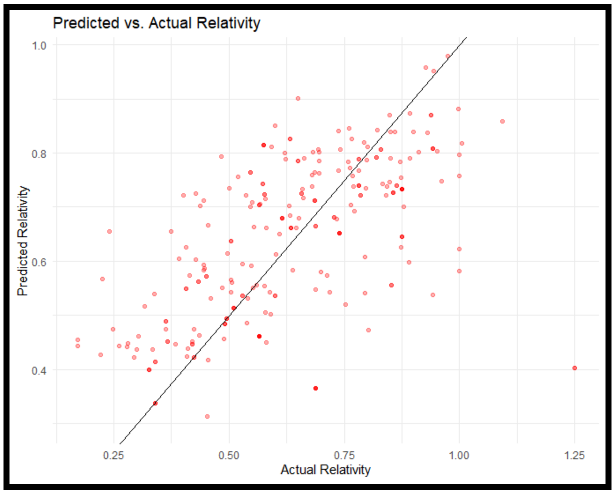


Figure 4: 3D GAM Surface between Layer Limit and Layer Attachment

Another form of model assessment was to analyze both the predictive vs. actual relativity plot and the residuals vs. predicted relativity plot.



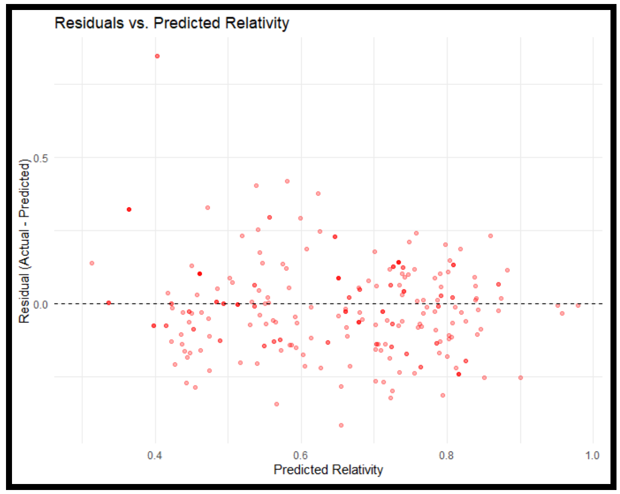


Figure 5: GAM Pedicted Relativity vs. Actual Relativity **Figure 6: GAM Residuals vs. Predicted Relativity**

Figure 2 compares the predicted relativity values from the model against the actual observed relativity in the test set. Ideally, the points would align closely along the 45-degree line, indicating near-perfect predictions. In our model, while some scatter is present, there is a strong positive trend along the line, suggesting that the model captures the general relationship between later structure and relativity well.

Figure 3 displays the residuals (the difference between actual and predicted values) against the predicted relativities. A well-performing model should show residuals randomly scattered around zero without obvious patterns. In our case, the residuals are relatively evenly distributed, indicating that the model does not suffer from significant bias across the range of predicted values. No strong curvature or heteroskedasticity (widening of residual spread) is visible, further supporting the model's validity.

The model's efficiency was evaluated using the Generalized Cross Validation (GCV) score and the scale estimate. A GCV of 0.0218 and a scale estimate of 0.0215 were achieved, indicating strong model performance and minimal overfitting. The close alignment between the GCV and the scale estimate suggests that the model generalizes well to unseen data without excessive complexity. These low values reflect the model's ability to balance flexibility and predictive accuracy, supporting its reliability in practical applications.

The quantile regression (QR) model explained 34.3% of the deviance, with an adjusted R^2 of 0.337, indicating moderate goodness of fit. The root mean squared error (RMSE) was 0.153, and the mean absolute percentage error (MAPE) was 36.0%. While the model generally performed well, prediction errors exhibited long tails, likely due to combining quantile and extrapolation-based regressions. Percent error increased with tower height—reflecting increased uncertainty in higher layers where insurers had less pricing visibility—but remained stable across layer limits. However, model behavior beyond $50 million limits remains untested, as such layers were rare (<7% of towers had more than 8 layers).

The Excel-based front-end, powered by VBA integration with R, allowed users to dynamically input tower characteristics and instantly retrieve modeled rate relativities, supporting practical pricing and restructuring analysis for brokers and underwriters.

Comparison to Project Objectives
The final deliverables closely aligned with the project's initial objectives:

- A working predictive model was developed that reflects how excess casualty towers are priced in current market conditions.
- A user-friendly Excel tool was created, designed specifically for brokers and underwriters to use without requiring programming knowledge.
- Scenario-testing capabilities were integrated, allowing users to model hypothetical tower structures and stress-test different market situations.

Risk Analysis
While the model performed well within the majority of the data range, some risks remain:

- Data sparsity at high attachment points and limits (> $250 million) can lead to increased prediction error.
- Ongoing maintenance is required to refresh the model as market conditions evolve.

Mitigation strategies include restricting usage to the validated range of the model and updating the dataset regularly to incorporate new client towers.

Economic Justification
Although a direct cost-benefit analysis was outside the project's formal scope, the developed model has the potential to save significant broker time, improve pricing accuracy, and increase client confidence through data-driven benchmarking. These improvements could indirectly lead to better market positioning and higher client retention.

Conclusion and Recommendations

Summary of Findings
This project successfully developed a flexible, data-driven model for analyzing and predicting rate relativities across excess casualty insurance towers. Using Generalized Additive Models (GAM), quantile regression, and linear regression, the team modeled how key factors, such as layer limit, attachment point, and umbrella pricing, affect pricing structure. The model captured industry trends effectively and provided valuable insights into the relationship between tower

structure and premium relativities. An Excel-based interface was built to make the model accessible and usable for brokers and underwriters in real-world pricing discussions.

The findings indicate that a data-driven approach can improve pricing accuracy, enhance client-facing deliverables, and enable more strategic structuring of towers, especially in a market facing increasing complexity and dynamic risk factors.

Final Deliverables

- GAM & QR-based predictive model: Built in R to estimate relativity for excess casualty layers.
- Excel user interface: Broker-friendly tool with VBA integration for dynamic scenario analysis.
- Scenario-testing capabilities: Ability to model different tower structures and compare modeled outputs against current market pricing.
- Process documentation and recommendations: Guidelines for future maintenance and enhancement of the model.

Detailed Recommendations

- Operationalize the Excel Tool: Deploy the Excel model internally to support real-time pricing discussions and tower structuring strategies with clients.
- Set Usage Boundaries: Limit model use to attachment points and limits well-represented in the dataset (e.g., below $250M) to ensure reliability.
- Embed Periodic Updates: Establish a schedule for refreshing the underlying data annually or semi-annually to maintain model relevance and accuracy.
- Integrate into Client Deliverables: Leverage the model to create more transparent client reports and benchmarking exhibits, enhancing the perceived value of WTW's broking services.

Suggestions for Future Work

We expect this unconventional modeling approach can be improved upon by WTW’s actuarial team(s) by implementing some of the recommendations above.

Reflections on the Project Experience and Learning Outcomes

This project provided valuable experience in blending academic modeling techniques with practical, broker-facing applications. The team gained hands-on exposure to the complexities of excess casualty insurance pricing, data cleaning challenges, model building, and stakeholder communication. Key takeaways include the importance of balancing statistical rigor with usability and the critical role of data structure and quality in successful model deployment. The project reinforced the real-world importance of flexibility, transparency, and collaboration in developing solutions for the insurance industry.

References

Posit, PBC. (2025). *RStudio (Version 2024.12.1+563)* [Computer software]. https://posit.co/products/open-source/rstudio/

Python Software Foundation. (2023). *Python 3.11.4* [Computer software]. https://www.python.org/downloads/release/python-3114/

Microsoft Corporation. (2025). *Microsoft Excel 2024* [Computer software]. https://www.microsoft.com/en-us/microsoft-365/excel

OpenAI. (2025). *ChatGPT* (April 29 version) [Large language model]. https://openai.com/chatgpt

Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science*, 1(3), 297–310. https://doi.org/10.1214/ss/1177013604

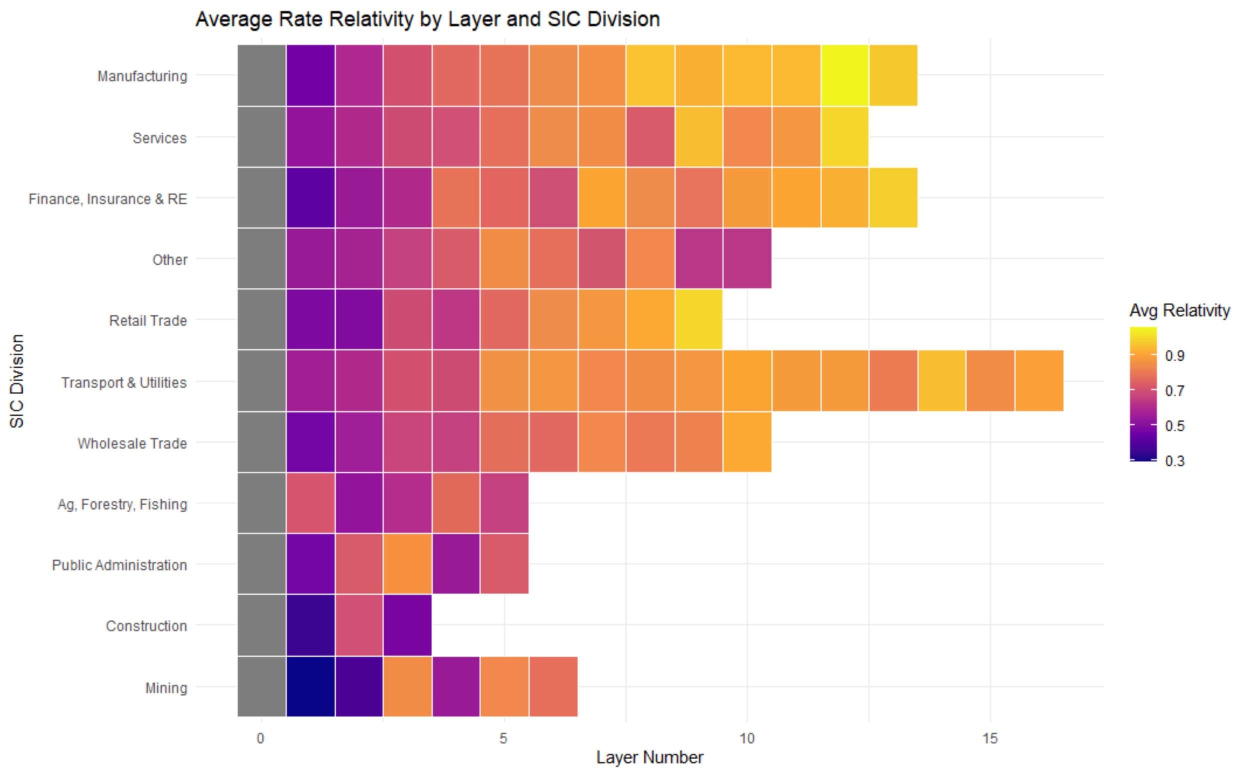
Henckaerts, Roel and Antonio, Katrien and Antonio, Katrien and Clijsters, Maxime and Roel, Verbelen, A Data Driven Binning Strategy for the Construction of Insurance Tariff Classes (May 12, 2017). Available at SSRN: https://ssrn.com/abstract=3052174 or http://dx.doi.org/10.2139/ssrn.3052174

Wood, S.N. (2017). Generalized Additive Models: An Introduction with R, Second Edition (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279

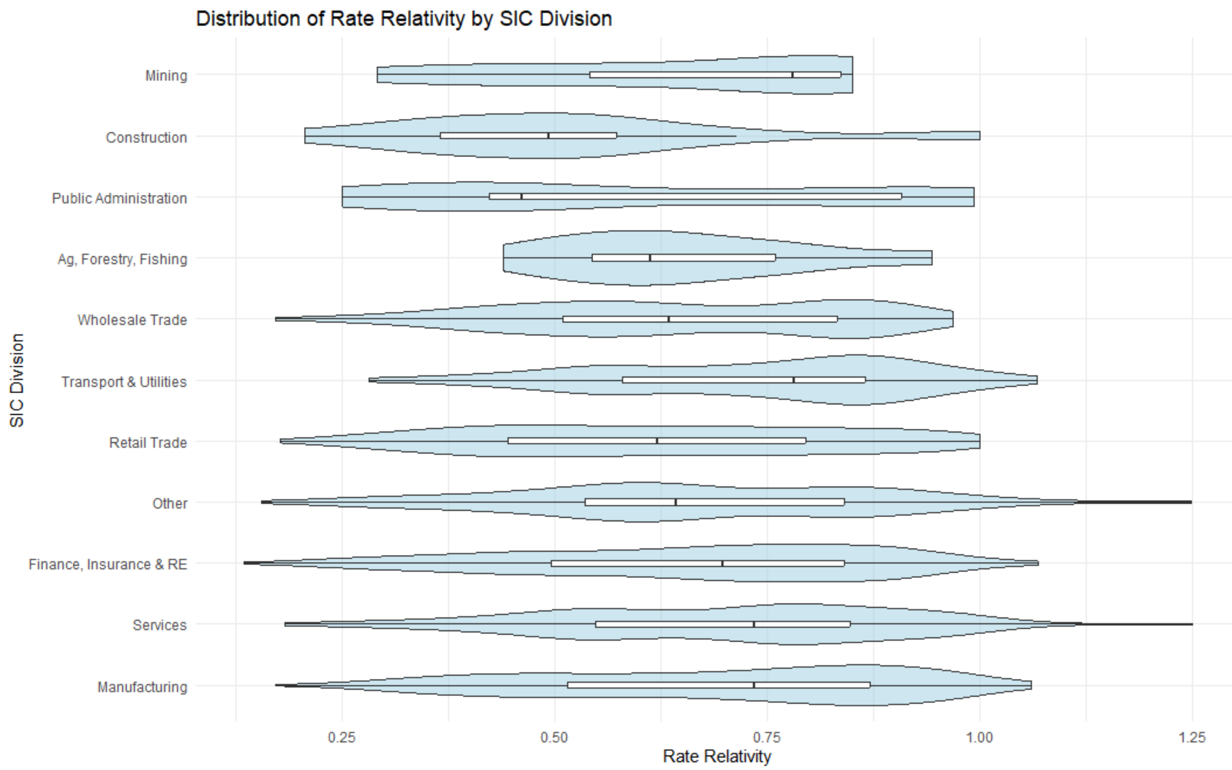
Appendices
Appendix A: R & Python Code

(Page intentionally left blank)
See attached R & Python files

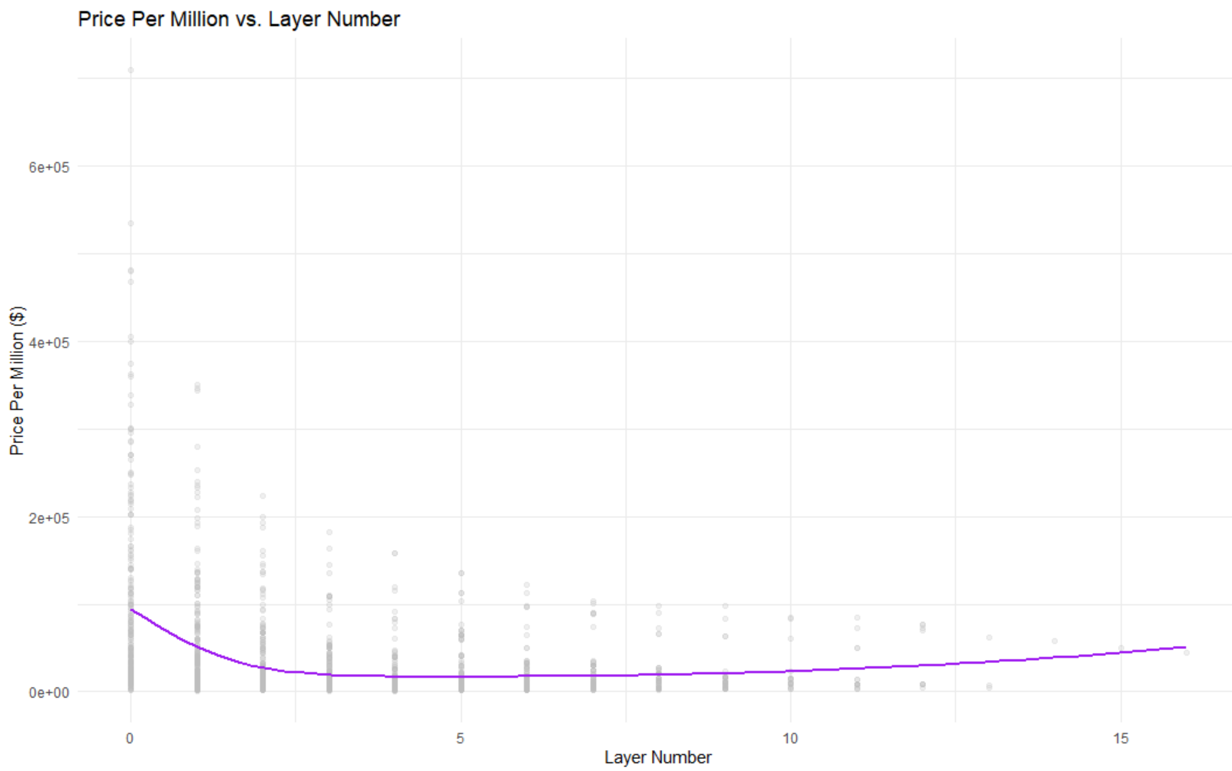
Appendix B: Heatmap of Average Rate Relativity by SIC Division



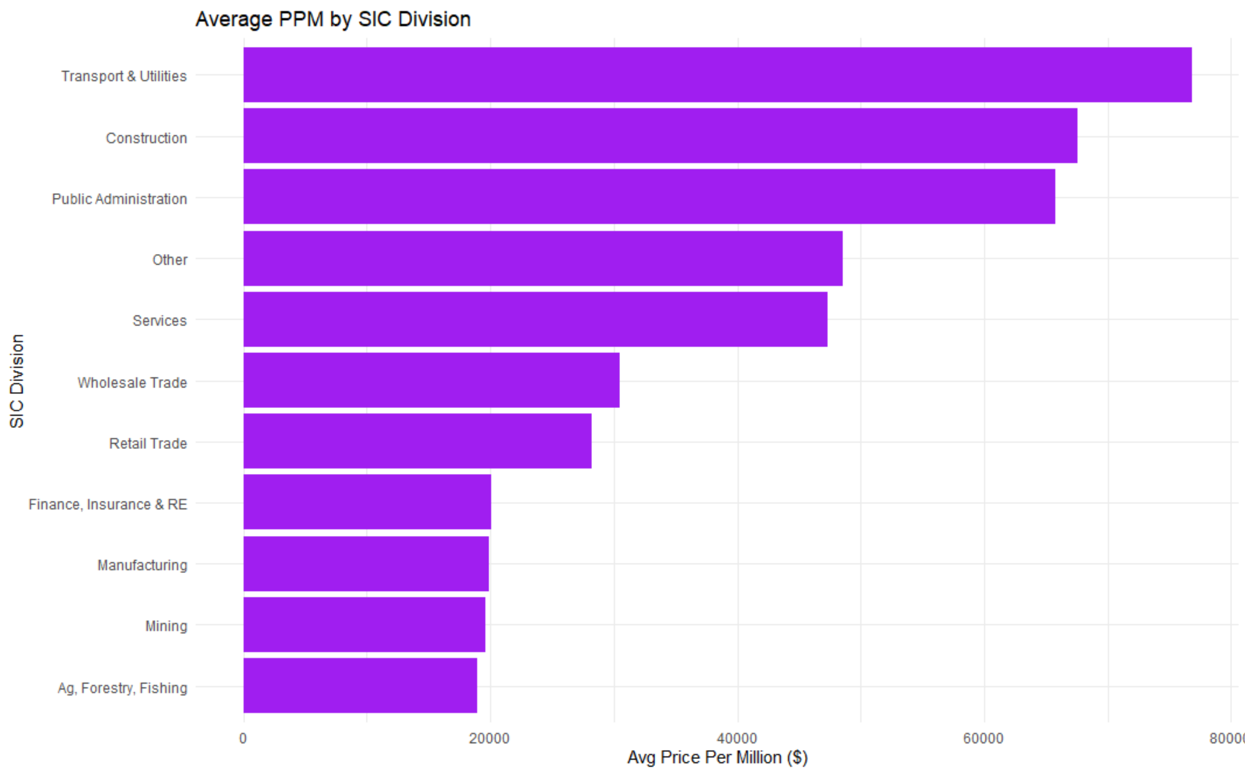
Appendix C: Distribution of Rate Relativities by SIC Division



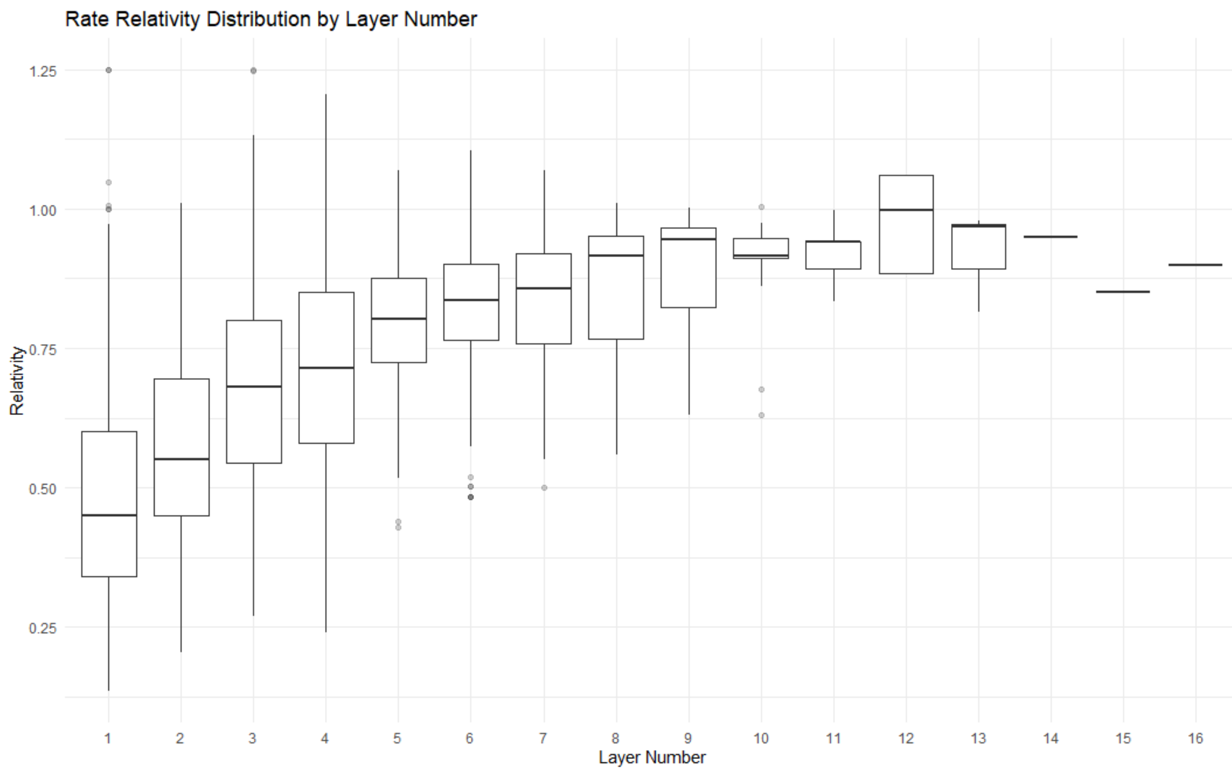
Appendix D: Layer PPM vs. Layer Number



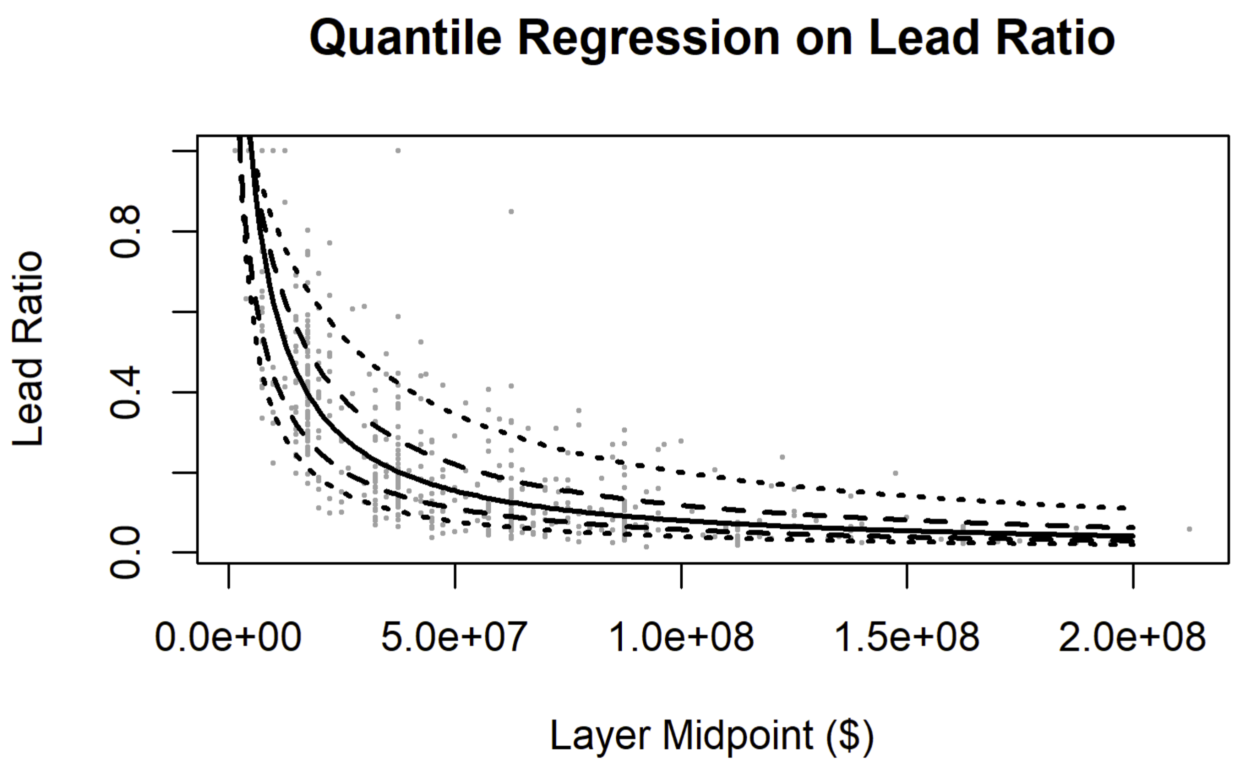
Appendix E: Average PPM by SIC Division



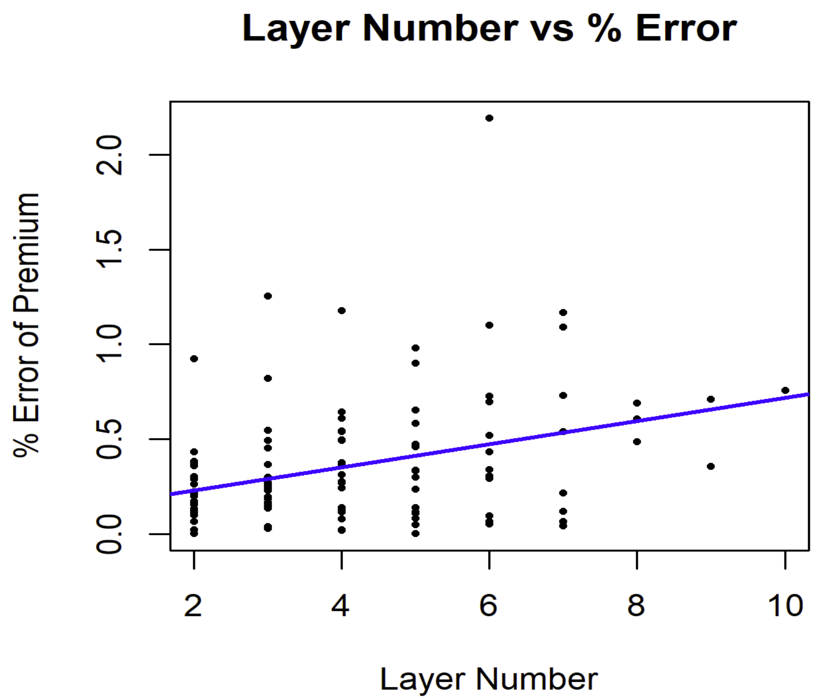
Appendix F: Rate Relativity Distribution by Layer Number



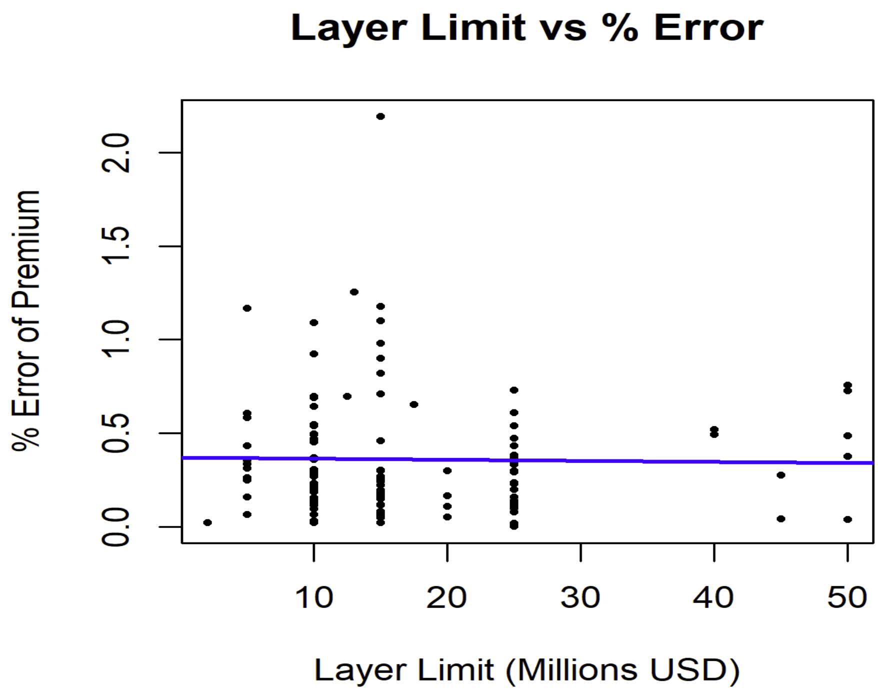
Appendix G: Quantile Regressions on Lead Ratio



Appendix H: Layer Number vs. Percent Error with Quantile Regression Approach



Appendix I: Layer Limit vs. Percent Error with Quantile Regression Approach



Appendix J: Team Member Time Budget Allowance

Team Member	Time Commitment
Ethan Carney	10-20 hours/week
Alexander Cimini	10-20 hours/week
Quinlin Gregg	10-20 hours/week
Nicholas Starring	10-20 hours/week
Total Time Commitment Per Week	40-80 hours/week

Acknowledgments

The team would like to sincerely thank all those who supported and guided us throughout the course of this project. We are especially grateful to Jon Drummond for sponsoring the project and collaborating with us throughout its development.

We are grateful to Ron Schuler and his actuarial team for their technical insight into the subtleties of modeling excess casualty insurance pricing.

We would also like to thank Mike Williams, Len Graziano, and Ned Kaufman for sharing their invaluable industry insights during the initial learning sessions, which laid the foundation for our understanding of excess casualty insurance.

Our appreciation extends to Matthew Prestifilippo for his collaboration, feedback, and assistance during the data preparation and model development phases. His technical guidance was instrumental in ensuring the project remained aligned with industry needs.

Finally, we would like to thank our academic advisors and mentors at Montana State University, particularly Dr. Faraz Dadgostari for his commitment to the Financial Engineering Capstone class, and Mr. David Brower for his guidance and for connecting us with capstone clients.

Without the support, guidance, and expertise of these individuals, the success of this project would not have been possible. Their contributions were critical in helping us develop a deeper understanding of the industry, refine our approach, and ultimately deliver a model and tool aligned with real-world needs.