

CLAIMS CORRELATION STUDY

December 2025

CLIENT

This research was conducted in collaboration with a Fortune 500 company.

ABOUT

This document summarizes the findings from research into the possible correlation between Moody's datapoints and a Fortune 500 company's claims events and loss sizes.

AUTHOR

Justin Whitehouse
Associate Director, Insurance

Victor Luo
Associate Director, Asset Management

Steven Morrison
Senior Director, Insurance

CONTACT US

Justin.whitehouse@moodys.com

Moody's and Fortune 500 company: Insurance claims correlation study

From late 2024 into early 2025, Moody's and a Fortune 500 company ("the Insurer") collaborated on a joint research project to identify company datapoints from across several data domains in the Moody's data estate that could help model claims risk.

At this stage in the research, the analysis has focused on key firmographic, financial, and credit risk data that has historically been difficult for the insurance industry to capture.

FINDINGS

We find significant trends and relationships when comparing the age, probability of default, and number of employees of insured companies with incurred claim events and loss sizes. By building predictive models for both claim events and total loss sizes, we quantified the contributions to prediction accuracy from these datapoints and others to help the Insurer understand which attributes provide the strongest contributions.

The evidence shows that there may be opportunities for other insurers in the industry to improve business performance by combining their own expertise with Moody's data and solutions.

Introduction

Over a period of six months from late 2024 into early 2025, Moody's and the Insurer collaborated on a joint research project to quantify the contributions from company datapoints in Moody's suite of data products to the prediction of claims risk above a baseline model.

The scope of the study included policy and claims data from lines of business across general liability, professional, and financial lines. These lines were written with two broad business groups: US small and medium-size privately owned enterprises and international businesses with a significant representation of large and public companies.

The risk factors covered included Moody's firmographic, financial statement, and credit risk data, as well as the Insurer's own policy data.

We conclude that a number of Moody's datapoints can provide significant contributions to models of both claim incidence and total incurred loss size above the baseline model.

We reach this conclusion from two approaches:

1. By performing quantitative analysis of the relationships between Moody's datapoints and the Insurer's claims data
2. By building gradient-boosted tree models to predict claim incidences and sizes

In section *Approach*, we describe the approach taken and assumptions made throughout the study. In section *The relationship between company data and claims*, we present results of the analysis of the relationships between observed claims with a range of the risk factor variables. In section *Using company data improves model predictions*, we present the marginal uplift found in machine learning model predictions gained from the sequential addition of more risk factor datapoints. Finally, in section *Moody's company data can give insurers an edge*, we recap our findings and discuss the possible implications for insurers that are considering incorporating an enhanced view of risk into their underwriting and business workflows.

Approach

Data

For the purpose of modeling and analysis, the Insurer provided Moody's with policy and claims data from two broad groupings of their business, international and US specialty, which included general liability, professional, and financial lines of business.

US SPECIALTY

US small and medium-size privately owned enterprises

INTERNATIONAL

International business, including a broad range of company sizes and types; includes significant representation of large and public companies, and companies from North America and Europe

Moody's and the Insurer worked together to construct two key modeling target variables from the dataset:

CLAIM INDICATOR	A binary variable that indicates whether any claim has been made against the policy
TOTAL INCURRED LOSS	A continuous, positive variable quantifying the total incurred loss against a policy given that at least one claim has been made against it

Moody's aligned the time series of the Insurer's claim events and loss sizes with a range of our datapoints:

- The age of the firm at the date of policy inception
- The industry classification code
- A range of financials including total assets, operating revenue, and number of employees
- The Moody's EDF-X probability of default and implied rating

We applied generic modeling assumptions across all lines of business and both policy groups:

- All monetary and financial values were converted to US dollars using exchange rates as of June 30, 2024.
- All numeric observations over a period in a window around a policy inception date were aggregated into the average and the change of the value over that period.

Analysis

The analysis proceeded in two phases: First, we performed univariate exploratory data analysis to examine the relationships between the dependent and independent variables. Following that we used the Extreme Gradient Boosting (XGBoost) model library to fit gradient-boosted tree models to both the claim indicator (using the classifier model) and the total loss given claim (using the regressor model with a gamma distribution objective).

The XGBoost modeling was performed in phases: First we built a baseline model using only policy information and no Moody's company data. Then, by sequentially adding data from the subsets for Moody's data above the baseline model existing data, we were able to quantify the marginal lift gained by incorporating each additional set of Moody's datapoints. The datapoints added in iteration of the model are as follows:

Model iteration	Datapoints	Source
0 — Baseline	Policy data including premiums	Insurer
1 — Add industry	<i>As above</i> + industry code	Moody's
2 — Add age	<i>As above</i> + firm age	Moody's
3 — Add financials	<i>As above</i> + range of financial metrics	Moody's
4 — Add probability of default	<i>As above</i> + probability of default and implied rating	Moody's

The relationship between company data and claims

Claim indicator

We plot the distribution of premium, firm age, probability of default (PD), and number of employees separately for policies with and without claim indicators. For both the international and US specialty portfolios we see significant differences in the characteristics of these distributions between instances where claims have and have not been made.

It is a fundamental feature of insurance that higher premiums reflect an assessment of greater risk. This is evidenced by the distributions of premiums from policies where claims have been made, peaking at larger premium values than those where no claims have been made.

The distribution of firm ages associated with claim events is also shifted to the right relative to those not associated with claim events. This may simply reflect that a longer period of operation provides more opportunities for incidents to occur.

There appears to be little difference in the PD distributions between policies with claims and those without. This may reflect the fact that the claims data contains a mixture of policy types, some of which may be less sensitive to creditworthiness.

In addition, the distribution of employee numbers is skewed toward higher values for policies where claims are made. The skew is particularly striking in the US specialty portfolio. Workers' compensation is not included in this analysis, but having more employees may be an indicator of more opportunities for incidences of third-party damage or injury to occur, or for accusations of malpractice to arise.

Any hypothesis about causality is speculative at this stage, but the evidence suggests there is some correlation insurers could explore further.

Figure 1: The distributions of key risk factors separated by whether they are associated with claim events or not (international portfolio)

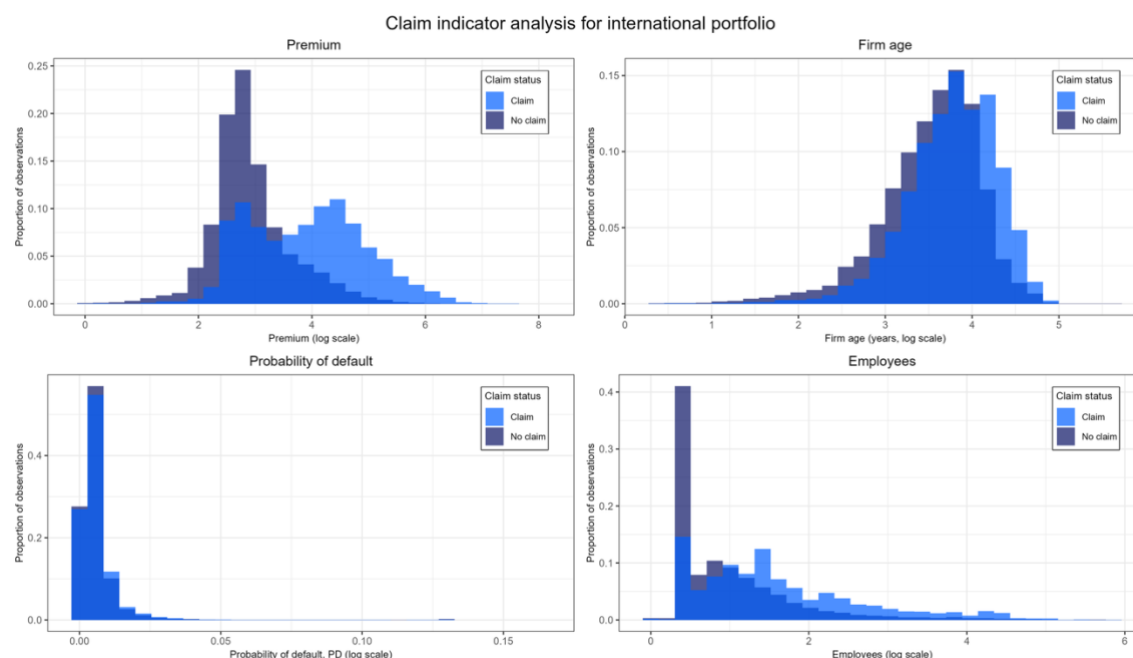
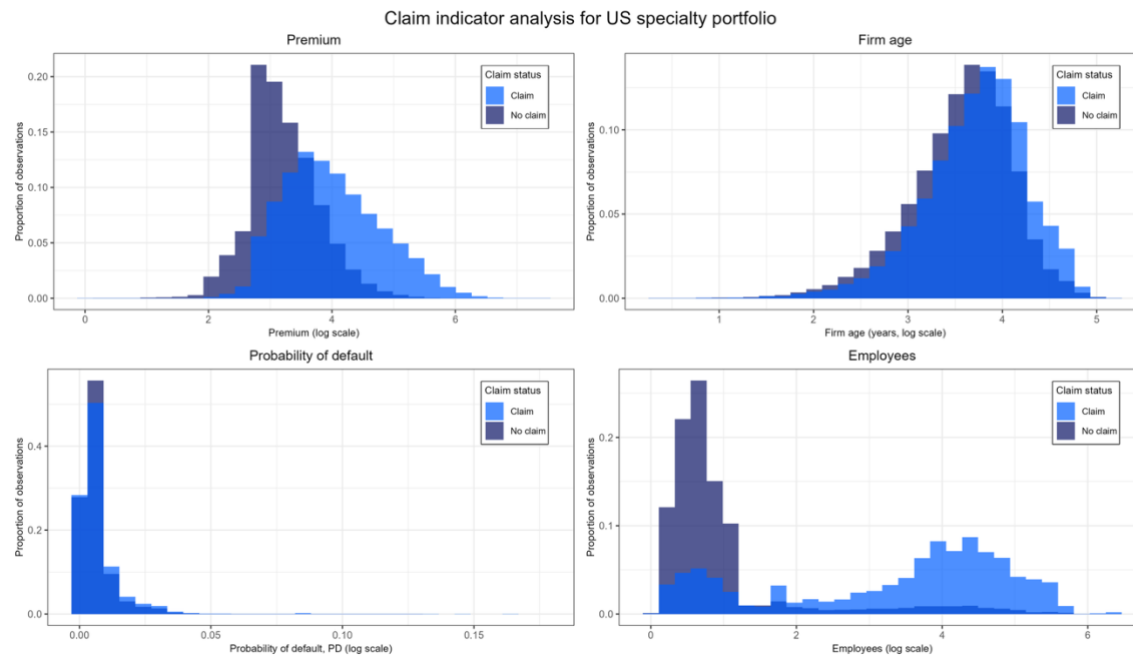


Figure 2: The distributions of key risk factors separated by whether they are associated with claim events or not (US specialty portfolio)



Total loss

We also performed a locally estimated scatterplot smoothing (LOESS)¹ regression of the mean values of ordered buckets of key risk factors against the total loss incurred given claim event.

The results clearly show that the (log) total loss incurred increases with (log) premium. The results also show an increasing relationship between firm age and loss incurred. The total loss may scale with premium by construction, but for both premium and firm age the increasing trend may also reflect that these variables are expected to correlate with company size, which in turn correlates with total incurred loss size.

The effect of size may also be present in the relationship between PD and loss size. At the lower end of the PD scale, we see the total loss size increasing as PD *decreases*. Larger companies tend to have lower PDs than smaller counterparts, so this observation could be another correlation with the company size.

This observation only explains the lower end of the PD scale; what happens in the mid-to-high PD range is less clear. One might expect that companies with higher PDs are under more stress and thus make more claims, reflecting greater total incurred losses. This would present in an upward trend in total loss incurred for higher PDs. Considering this expectation alongside the correlation with size, we might then expect to see a convex function or “smile” shape in the relationship between total incurred loss and PD.

In both portfolios, there is an increase in total loss as the PD decreases in the lower PD range, although for US specialty the trend reverses again at the very lowest PD range. At the high PD range, there is an increase in the loss with increasing PD in the US specialty portfolio. The same trend is arguably present but weaker in the international portfolio. The US specialty portfolio also exhibits a humped feature in the mid-PD range, which is so far unexplained.

The final variable considered is the number of employees. This may be expected to scale with company size, and the international portfolio data does show the same upward trend as seen with premium and

¹ Alternatively known as locally weighted scatterplot smoothing or LOWESS.

firm age. For US specialty, the relationship shows a “hump” in the midrange of the total number of employees and is otherwise arguably flat.

As stated earlier, we are presenting correlation in these charts, not causation. We have speculated that the size of the company is an important factor and hypothesized the existence of a “smile” in the relationship between PD and loss incurred.

Figure 3: Log total incurred loss in the international portfolio against selected risk factors

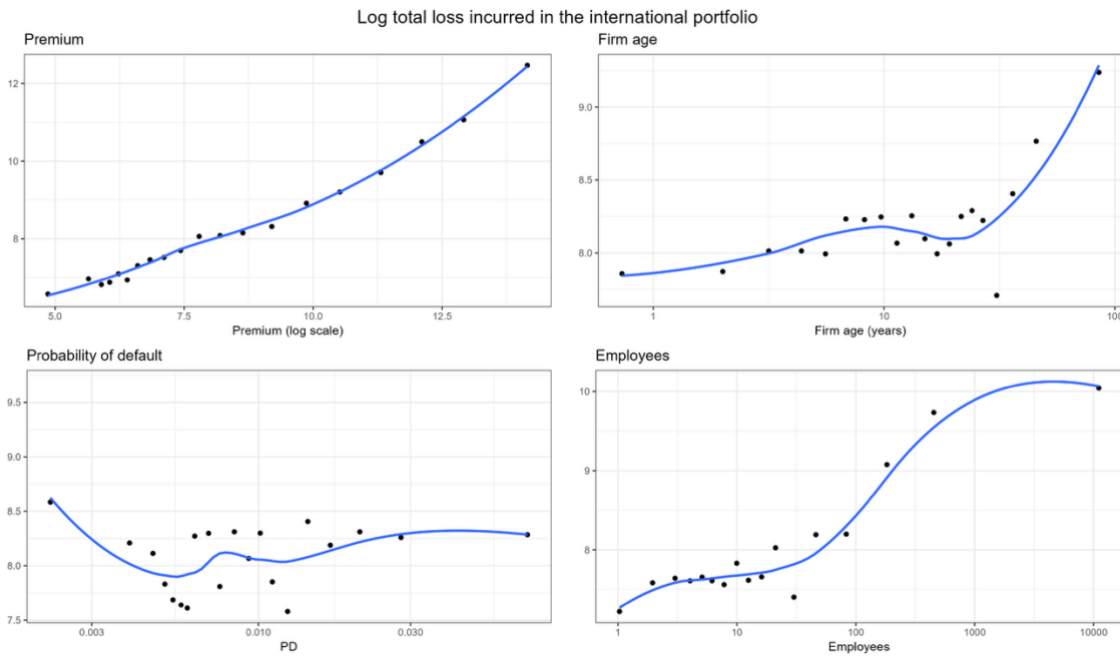
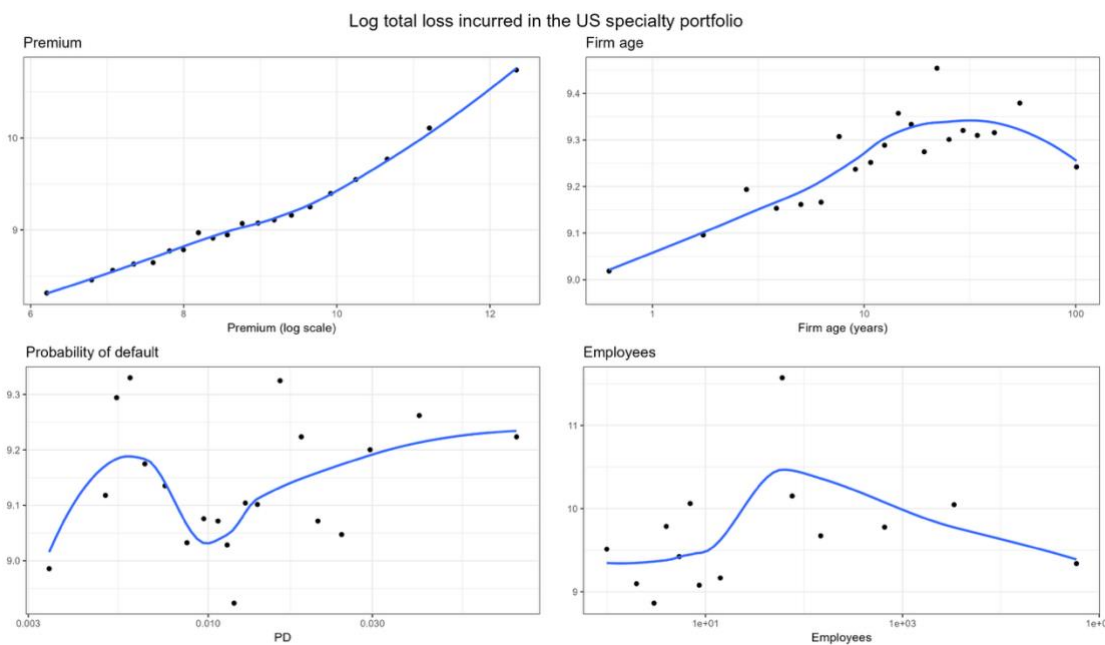


Figure 4: Log total incurred loss in the US specialty portfolio against selected risk factors



Using company data improves model predictions

The assessment of model performance presented here has been produced by analyzing a “holdout” dataset — a subset of the original full dataset that has been held aside from the process of building and parameterizing the model and thus has had no influence on the calibration of the final model parameters nor the optimization of the model hyperparameters.

By validating the model performance against the holdout data we can be confident that measured improvements in performance are not a consequence of overfitting and reflect genuine model improvement.

Claim indicator

By running the classifier model on the claim indicator we find a clear trend of increasing model accuracy with the addition of more datapoints. This is evidenced by the receiver operating characteristic (ROC) plots for each model with incremental additions on Moody’s data (Figure 5) and by the waterfall chart (Figure 6) that articulates the marginal increase in accuracy score associated with each model ROC.

The baseline model using only the Insurer’s existing policy data already provides a lot of predictive power, but it’s clear that incorporating additional company data from Moody’s can help achieve significant gains in accuracy.

There is a similar picture for the US specialty business but because this business is mostly composed of small and medium US-based entities that do not routinely publish financial statements, the coverage of financials data is very low. As such, the response to financials does not show an increase in accuracy.

Figure 5: ROC curves plot of the claim indicator model for the international portfolio

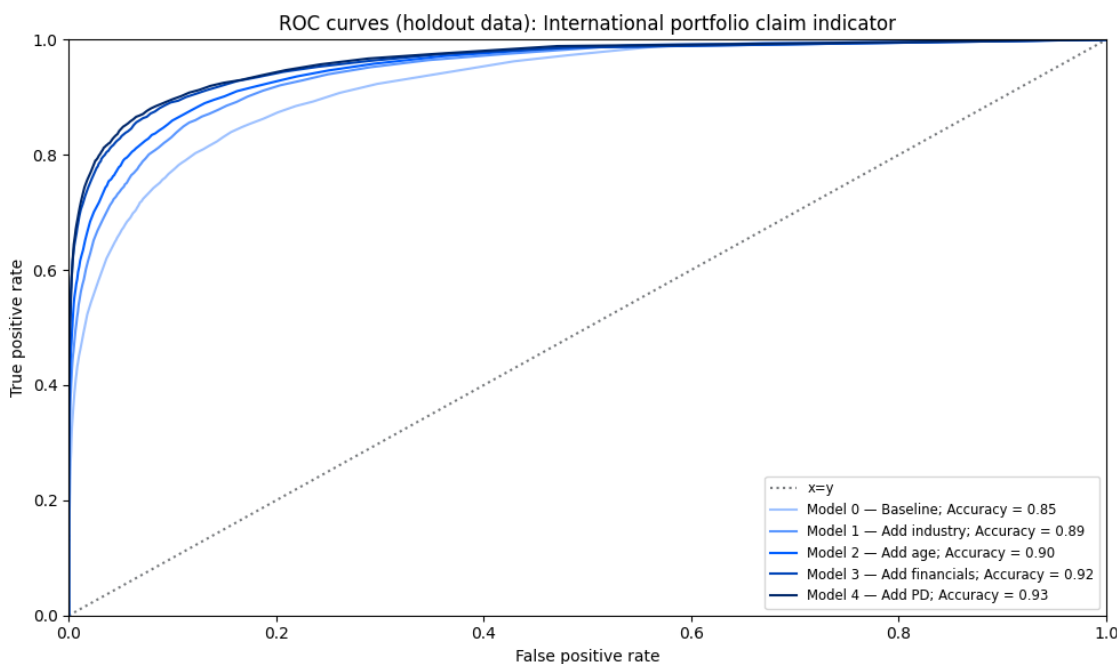


Figure 6: Waterfall chart presenting the incremental gains in accuracy score for the international portfolio claim indicator model from adding additional Moody's data to the model

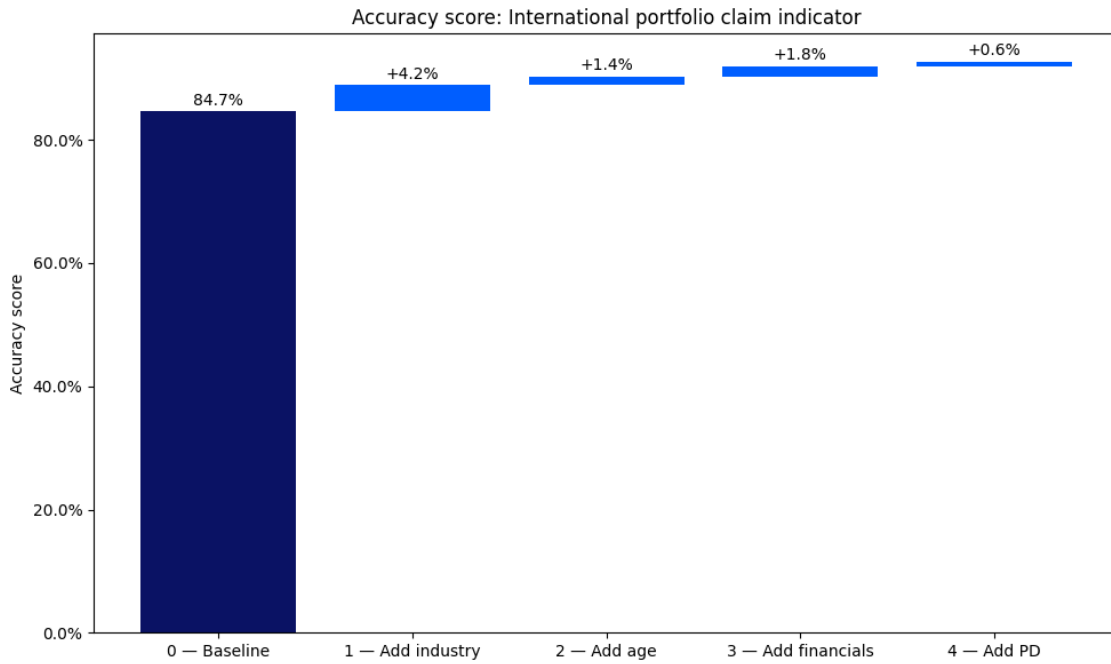


Figure 7: ROC curves plot of the claim indicator model for the US specialty portfolio

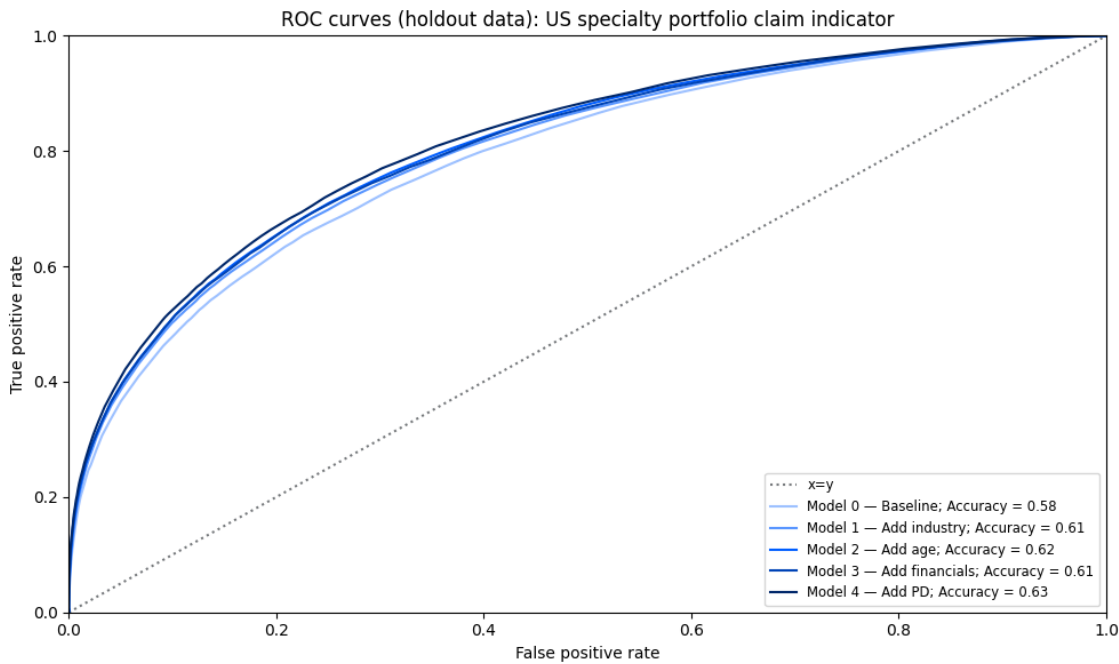
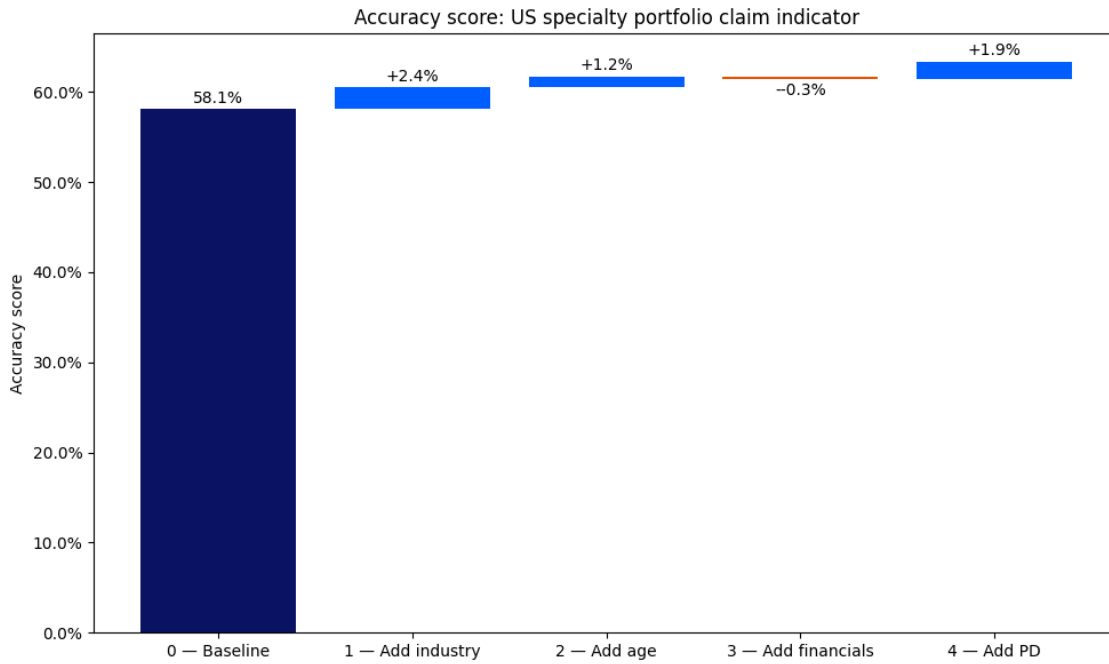


Figure 8: Waterfall chart presenting the incremental gains in accuracy score for the US specialty portfolio claim indicator model from adding additional Moody's data to the model



Total loss

Lorenz curves and Gini coefficients

To evaluate the relationship between Moody's company data and the total loss, we take a similar approach to the claim indicator model but instead quantify the predictive uplift across the models by using Lorenz curves and associated Gini coefficients. For the total incurred loss, we see a similar general trend to that seen for the claim indicator where the accuracy (as measured by the normalized Gini coefficient) increases with the addition of more Moody's company data. The accuracy decreases with the addition of company age data for both the international and US specialty businesses.

Figure 9: Lorenz curves plot of the loss amount model for the international portfolio

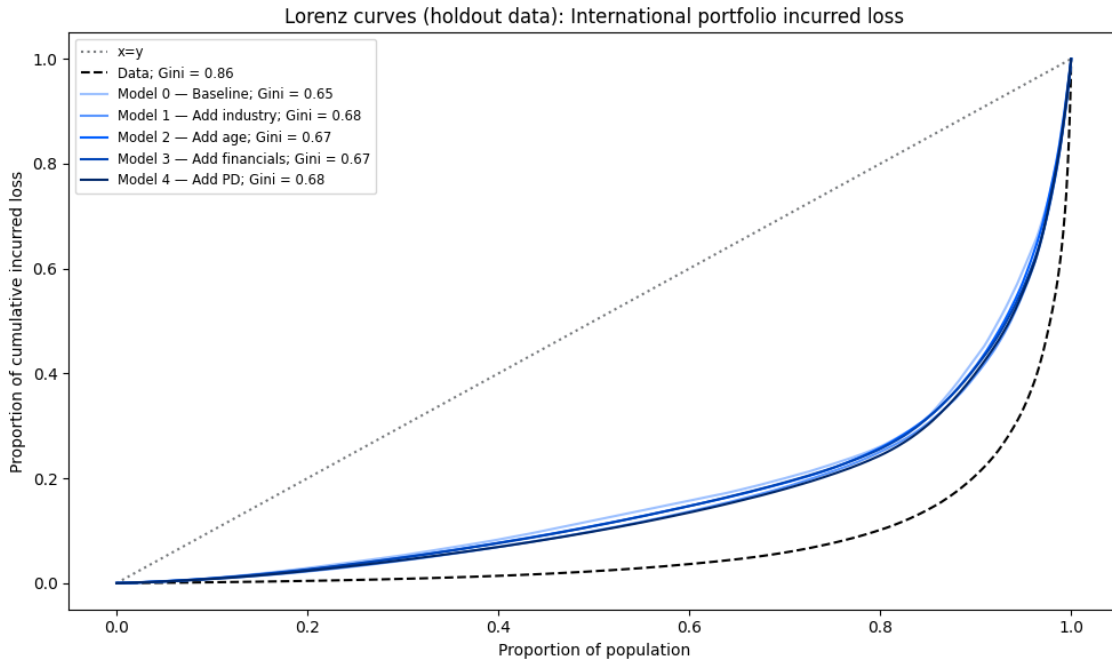


Figure 10: Waterfall chart presenting the incremental gains in normalized Gini coefficient for the international portfolio claim indicator model from adding additional Moody's data to the model

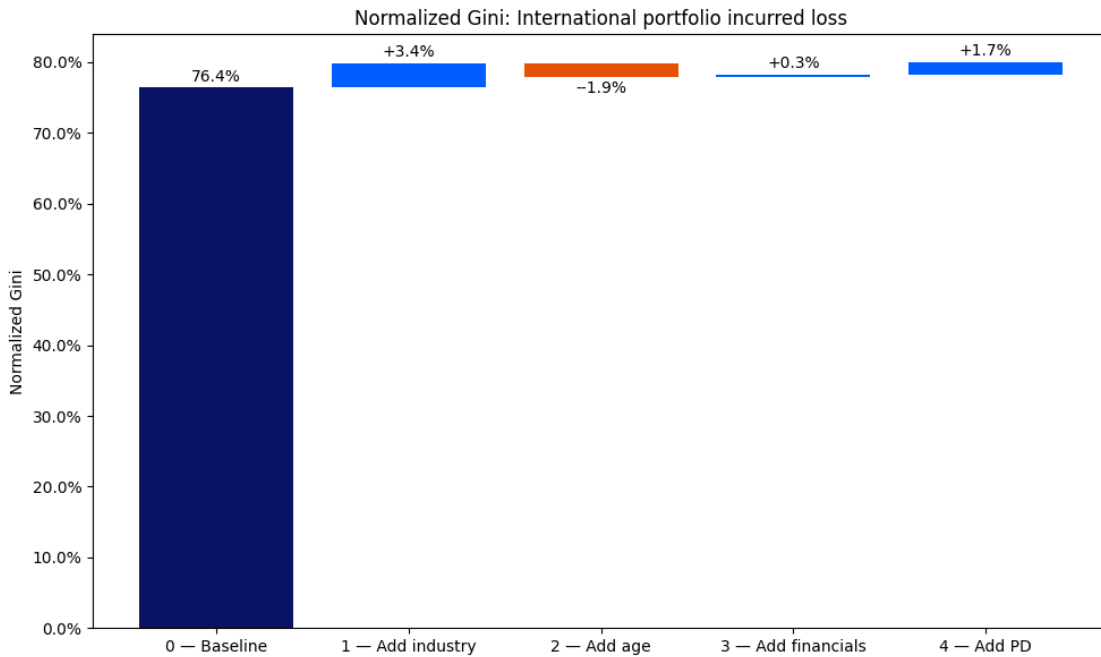


Figure 11: Lorenz curves plot of the loss amount model for the US specialty portfolio

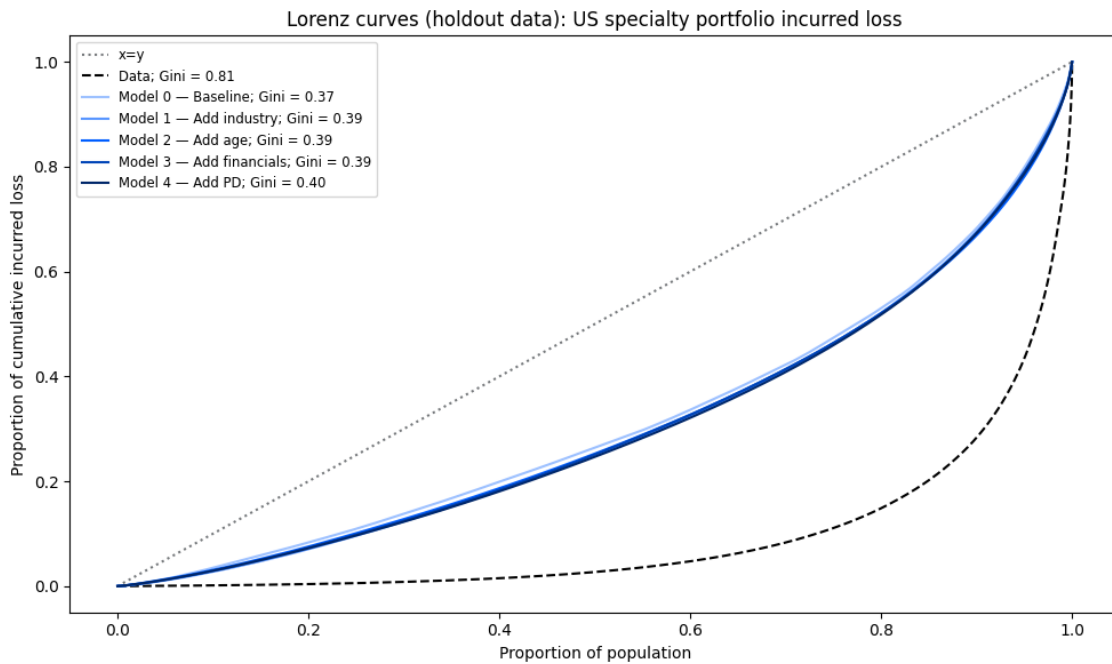
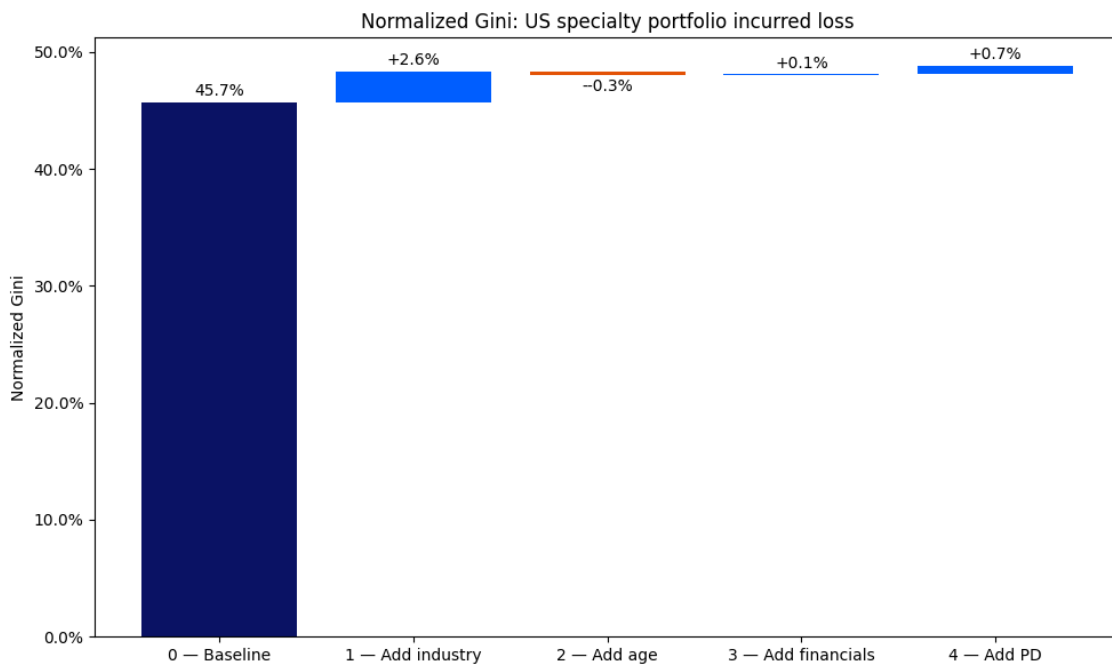


Figure 12: Waterfall chart presenting the incremental gains in normalized Gini coefficient for the US specialty portfolio claim indicator model from adding additional Moody's data to the model



Single lift charts

We can measure the lift on the holdout data provided by Model 4, which incorporates all available company data. When we order the predicted total losses and compute the mean predicted loss alongside the mean actual loss in each of 10 decile bins, we see that the model provides significant lift and tracks very well against the actual loss sizes in both the international and US specialty businesses.

Figure 13: A single lift plot shows that the lift provided by Model 4 — add PD tracks well against the holdout data for the international portfolio

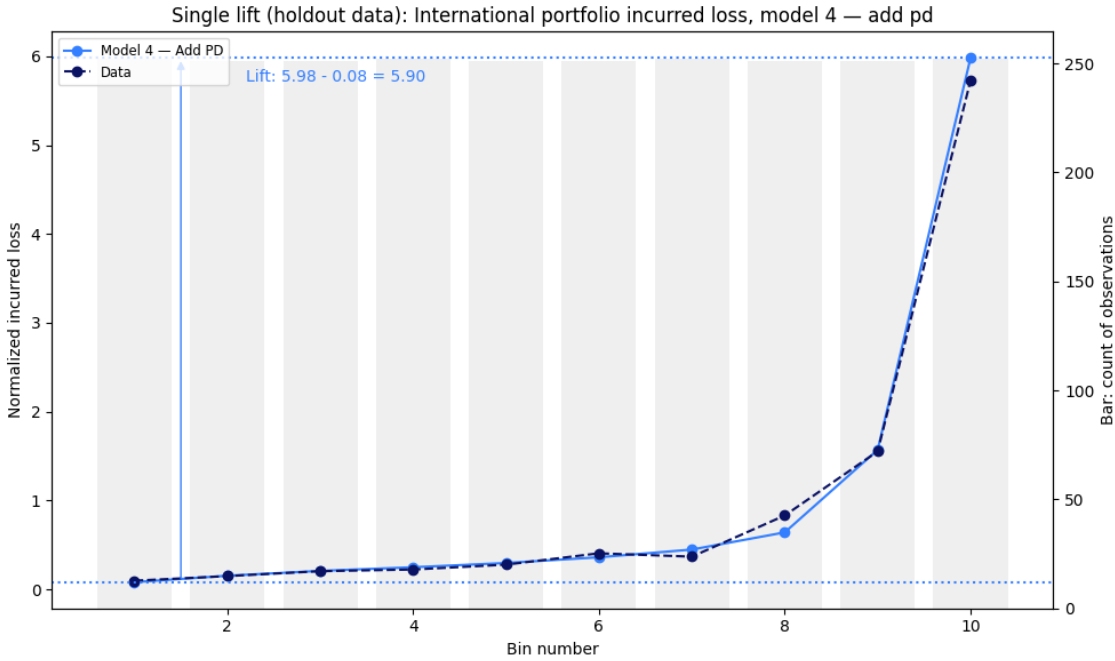
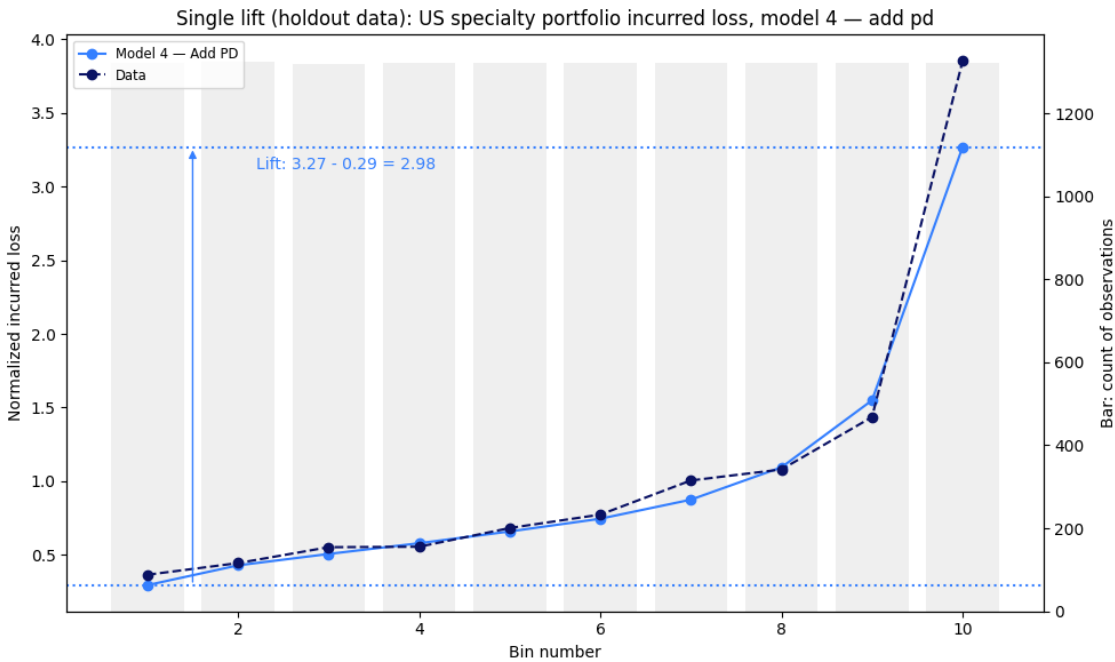


Figure 14: A single lift plot shows that the lift provided by Model 4 — add PD tracks well against the holdout data for the US specialty portfolio



Double lift charts

We also compare Model 4 with the baseline model, which uses only the Insurer's policy data as a predictor for total loss. In these "double lift charts," we compare the two models by ordering the observations by the ratio of the total loss prediction sizes in each model. The predictions from the two models diverge most at the left and right extremes of the chart. On the left side, Model 0 (the baseline) predicts much larger losses than Model 4 (all company data), and on the right side, Model 4 predicts much larger losses than Model 0. The charts show that for both sets of business and *at both extremes* the average loss predicted from Model 4 aligns more closely with the data than the average loss predicted from Model 0. This tells us that in places where the models disagree in their predictions, Model 4 is more accurate than Model 0.

Figure 15: A double lift plot shows that the lift provided by Model 4 — add PD tracks better against the holdout data compared with Model 0 — baseline for the international portfolio

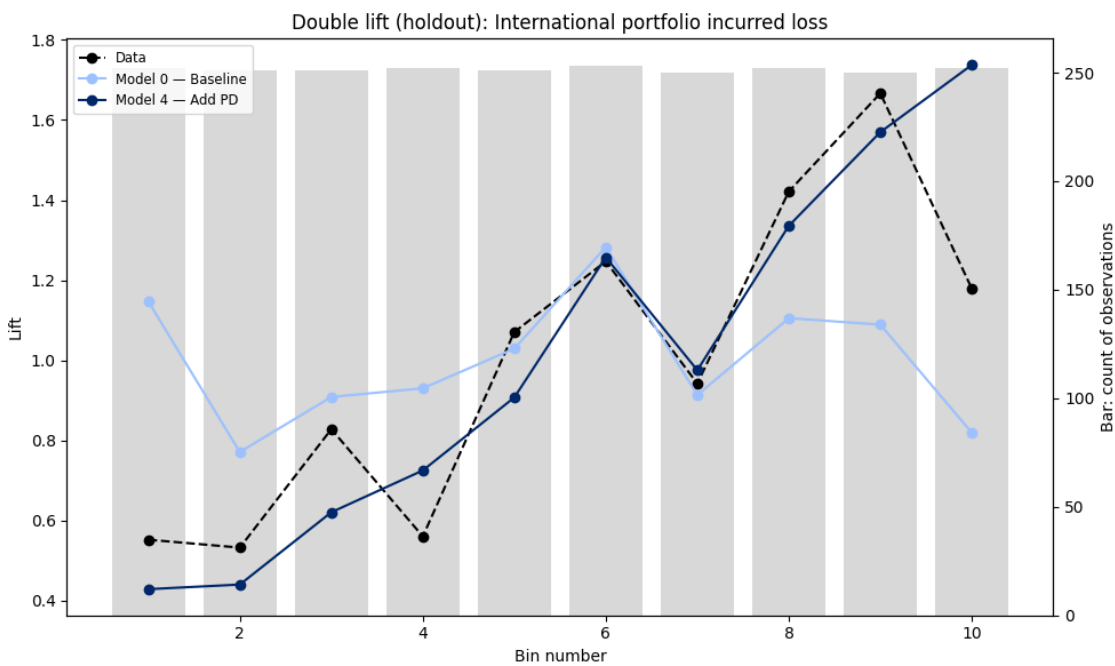
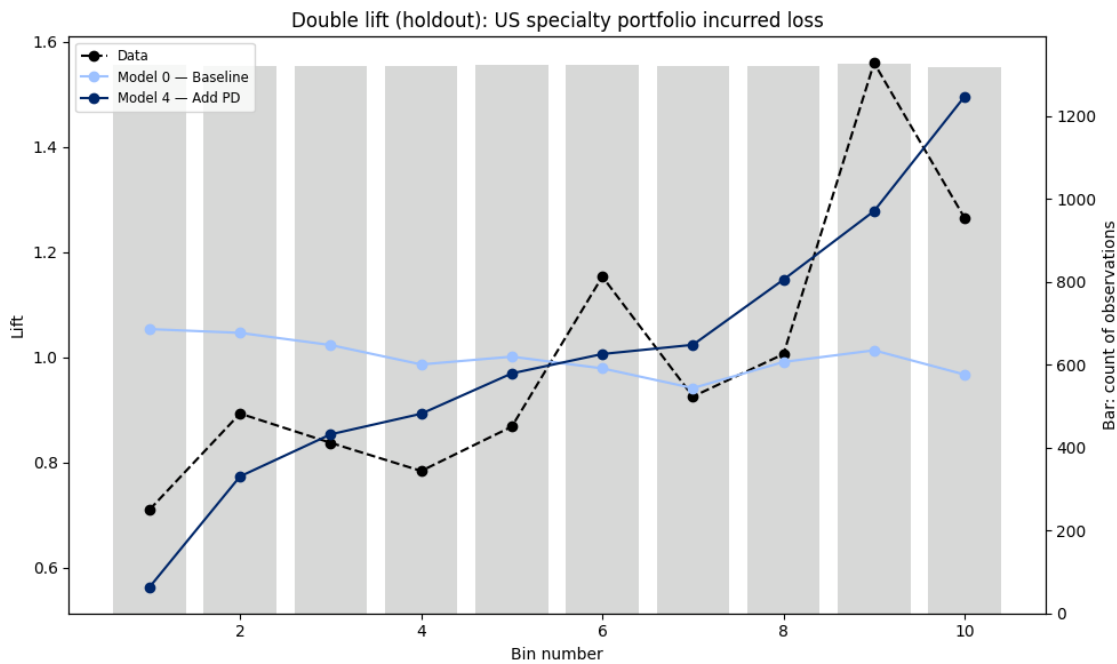


Figure 16: A double lift plot shows that the lift provided by Model 4 — add PD tracks better against the holdout data compared with Model 0 — baseline for the US specialty portfolio



Moody's company data can give insurers an edge

The results of this study demonstrate that insurers may be able to find significant contributions to predictive lift in their claims models by incorporating Moody's company data.

Features such as a company's financial performance and creditworthiness may reasonably be expected to correlate with the incidence of claims against financial lines of business; however, they may also be indicators of a company's overall operational health and risk of claims against other lines, such as professional, general liability and beyond. A company's size, age, and industrial sector are also likely indicative of claim risk, so being able to reliably capture even these simple features of a company can be beneficial.

© 2025 Moody's Corporation, Moody's Investors Service, Inc., Moody's Analytics, Inc. and/or their licensors and affiliates (collectively, "MOODY'S"). All rights reserved.

CREDIT RATINGS ISSUED BY MOODY'S CREDIT RATINGS AFFILIATES ARE THEIR CURRENT OPINIONS OF THE RELATIVE FUTURE CREDIT RISK OF ENTITIES, CREDIT COMMITMENTS, OR DEBT OR DEBT-LIKE SECURITIES, AND MATERIALS, PRODUCTS, SERVICES AND INFORMATION PUBLISHED OR OTHERWISE MADE AVAILABLE BY MOODY'S (COLLECTIVELY, "MATERIALS") MAY INCLUDE SUCH CURRENT OPINIONS. MOODY'S DEFINES CREDIT RISK AS THE RISK THAT AN ENTITY MAY NOT MEET ITS CONTRACTUAL FINANCIAL OBLIGATIONS AS THEY COME DUE AND ANY ESTIMATED FINANCIAL LOSS IN THE EVENT OF DEFAULT OR IMPAIRMENT. SEE APPLICABLE MOODY'S RATING SYMBOLS AND DEFINITIONS PUBLICATION FOR INFORMATION ON THE TYPES OF CONTRACTUAL FINANCIAL OBLIGATIONS ADDRESSED BY MOODY'S CREDIT RATINGS. CREDIT RATINGS DO NOT ADDRESS ANY OTHER RISK, INCLUDING BUT NOT LIMITED TO: LIQUIDITY RISK, MARKET VALUE RISK, OR PRICE VOLATILITY. CREDIT RATINGS, NON-CREDIT ASSESSMENTS ("ASSESSMENTS"), AND OTHER OPINIONS INCLUDED IN MOODY'S MATERIALS ARE NOT STATEMENTS OF CURRENT OR HISTORICAL FACT. MOODY'S MATERIALS MAY ALSO INCLUDE QUANTITATIVE MODEL-BASED ESTIMATES OF CREDIT RISK AND RELATED OPINIONS OR COMMENTARY PUBLISHED BY MOODY'S ANALYTICS, INC. AND/OR ITS AFFILIATES. MOODY'S CREDIT RATINGS, ASSESSMENTS, OTHER OPINIONS AND MATERIALS DO NOT CONSTITUTE OR PROVIDE INVESTMENT OR FINANCIAL ADVICE, AND MOODY'S CREDIT RATINGS, ASSESSMENTS, OTHER OPINIONS AND MATERIALS ARE NOT AND DO NOT PROVIDE RECOMMENDATIONS TO PURCHASE, SELL, OR HOLD PARTICULAR SECURITIES. MOODY'S CREDIT RATINGS, ASSESSMENTS, OTHER OPINIONS AND MATERIALS DO NOT COMMENT ON THE SUITABILITY OF AN INVESTMENT FOR ANY PARTICULAR INVESTOR. MOODY'S ISSUES ITS CREDIT RATINGS, ASSESSMENTS AND OTHER OPINIONS AND PUBLISHES OR OTHERWISE MAKES AVAILABLE ITS MATERIALS WITH THE EXPECTATION AND UNDERSTANDING THAT EACH INVESTOR WILL, WITH DUE CARE, MAKE ITS OWN STUDY AND EVALUATION OF EACH SECURITY THAT IS UNDER CONSIDERATION FOR PURCHASE, HOLDING, OR SALE.

MOODY'S CREDIT RATINGS, ASSESSMENTS, OTHER OPINIONS, AND MATERIALS ARE NOT INTENDED FOR USE BY RETAIL INVESTORS AND IT WOULD BE RECKLESS AND INAPPROPRIATE FOR RETAIL INVESTORS TO USE MOODY'S CREDIT RATINGS, ASSESSMENTS, OTHER OPINIONS OR MATERIALS WHEN MAKING AN INVESTMENT DECISION. IF IN DOUBT YOU SHOULD CONTACT YOUR FINANCIAL OR OTHER PROFESSIONAL ADVISER.

ALL INFORMATION CONTAINED HEREIN IS PROTECTED BY LAW, INCLUDING BUT NOT LIMITED TO, COPYRIGHT LAW, AND NONE OF SUCH INFORMATION MAY BE COPIED OR OTHERWISE REPRODUCED, REPACKAGED, FURTHER TRANSMITTED, TRANSFERRED, DISSEMINATED, REDISTRIBUTED OR RESOLD, OR STORED FOR SUBSEQUENT USE FOR ANY SUCH PURPOSE, IN WHOLE OR IN PART, IN ANY FORM OR MANNER OR BY ANY MEANS WHATSOEVER, BY ANY PERSON WITHOUT MOODY'S PRIOR WRITTEN CONSENT. FOR CLARITY, NO INFORMATION CONTAINED HEREIN MAY BE USED TO DEVELOP, IMPROVE, TRAIN OR RETRAIN ANY SOFTWARE PROGRAM OR DATABASE, INCLUDING, BUT NOT LIMITED TO, FOR ANY ARTIFICIAL INTELLIGENCE, MACHINE LEARNING OR NATURAL LANGUAGE PROCESSING SOFTWARE, ALGORITHM, METHODOLOGY AND/OR MODEL.

MOODY'S CREDIT RATINGS, ASSESSMENTS, OTHER OPINIONS AND MATERIALS ARE NOT INTENDED FOR USE BY ANY PERSON AS A BENCHMARK AS THAT TERM IS DEFINED FOR REGULATORY PURPOSES AND MUST NOT BE USED IN ANY WAY THAT COULD RESULT IN THEM BEING CONSIDERED A BENCHMARK.

All information contained herein is obtained by MOODY'S from sources believed by it to be accurate and reliable. Because of the possibility of human or mechanical error as well as other factors, however, all information contained herein is provided "AS IS" without warranty of any kind. MOODY'S adopts all necessary measures so that the information it uses in assigning a credit rating is of sufficient quality and from sources MOODY'S considers to be reliable including, when appropriate, independent third-party sources. However, MOODY'S is not an auditor and cannot in every instance independently verify or validate information received in the credit rating process or in preparing its Materials.

To the extent permitted by law, MOODY'S and its directors, officers, employees, agents, representatives, licensors and suppliers disclaim liability to any person or entity for any indirect, special, consequential, or incidental losses or damages whatsoever arising from or in connection with the information contained herein or the use of or inability to use any such information, even if MOODY'S or any of its directors, officers, employees, agents, representatives, licensors or suppliers is advised in advance of the possibility of such losses or damages, including but not limited to: (a) any loss of present or prospective profits or (b) any loss or damage arising where the relevant financial instrument is not the subject of a particular credit rating assigned by MOODY'S.

To the extent permitted by law, MOODY'S and its directors, officers, employees, agents, representatives, licensors and suppliers disclaim liability for any direct or compensatory losses or damages caused to any person or entity, including but not limited to by any negligence (but excluding fraud, willful misconduct or any other type of liability that, for the avoidance of doubt, by law cannot be excluded) on the part of, or any contingency within or beyond the control of, MOODY'S or any of its directors, officers, employees, agents, representatives, licensors or suppliers, arising from or in connection with the information contained herein or the use of or inability to use any such information.

NO WARRANTY, EXPRESS OR IMPLIED, AS TO THE ACCURACY, TIMELINESS, COMPLETENESS, MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OF ANY CREDIT RATING, ASSESSMENT, OTHER OPINION OR INFORMATION IS GIVEN OR MADE BY MOODY'S IN ANY FORM OR MANNER WHATSOEVER.

Moody's Investors Service, Inc., a wholly-owned credit rating agency subsidiary of Moody's Corporation ("MCO"), hereby discloses that most issuers of debt securities (including corporate and municipal bonds, debentures, notes and commercial paper) and preferred stock rated by Moody's Investors Service, Inc. have, prior to assignment of any credit rating, agreed to pay Moody's Investors Service, Inc. for credit ratings opinions and services rendered by it. MCO and all MCO entities that issue ratings under the "Moody's Ratings" brand name ("Moody's Ratings"), also maintain policies and procedures to address the independence of Moody's Ratings' credit ratings and credit rating processes. Information regarding certain affiliations that may exist between directors of MCO and rated entities, and between entities who hold credit ratings from Moody's Investors Service, Inc. and have also publicly reported to the SEC an ownership interest in MCO of more than 5%, is posted annually at ir.moody.com under the heading "Investor Relations — Corporate Governance — Charter and Governance Documents - Director and Shareholder Affiliation Policy."

Moody's SF Japan K.K., Moody's Local AR Agente de Clasificación de Riesgo S.A., Moody's Local BR Agência de Classificação de Risco LTDA, Moody's Local MX S.A. de C.V., I.C.V., Moody's Local PE Clasificadora de Riesgo S.A., and Moody's Local PA Clasificadora de Riesgo S.A. (collectively, the "Moody's Non-NRSRO CRAs") are all indirectly wholly-owned credit rating agency subsidiaries of MCO. None of the Moody's Non-NRSRO CRAs is a Nationally Recognized Statistical Rating Organization.

Additional terms for Australia only: Any publication into Australia of this document is pursuant to the Australian Financial Services License of MOODY'S affiliate, Moody's Investors Service Pty Limited ABN 61 003 399 657AFSL 336969 and/or Moody's Analytics Australia Pty Ltd ABN 94 105 136 972 AFSL 383569 (as applicable). This document is intended to be provided only to "wholesale clients" within the meaning of section 761G of the Corporations Act 2001. By continuing to access this document from within Australia, you represent to MOODY'S that you are, or are accessing the document as a representative of, a "wholesale client" and that neither you nor the entity you represent will directly or indirectly disseminate this document or its contents to "retail clients" within the meaning of section 761G of the Corporations Act 2001. MOODY'S credit rating is an opinion as to the creditworthiness of a debt obligation of the issuer, not on the equity securities of the issuer or any form of security that is available to retail investors.

Additional terms for India only: Moody's credit ratings, Assessments, other opinions and Materials are not intended to be and shall not be relied upon or used by any users located in India in relation to securities listed or proposed to be listed on Indian stock exchanges.

Additional terms with respect to Second Party Opinions and Net Zero Assessments (as defined in Moody's Ratings Rating Symbols and Definitions): Please note that neither a Second Party Opinion ("SPO") nor a Net Zero Assessment ("NZA") is a "credit rating". The issuance of SPOs and NZAs is not a regulated activity in many jurisdictions, including Singapore. JAPAN: In Japan, development and provision of SPOs and NZAs fall under the category of "Ancillary Businesses", not "Credit Rating Business", and are not subject to the regulations applicable to "Credit Rating Business" under the Financial Instruments and Exchange Act of Japan and its relevant regulation. PRC: Any SPO: (1) does not constitute a PRC Green Bond Assessment as defined under any relevant PRC laws or regulations; (2) cannot be included in any registration statement, offering circular, prospectus or any other documents submitted to the PRC regulatory authorities or otherwise used to satisfy any PRC regulatory disclosure requirement; and (3) cannot be used within the PRC for any regulatory purpose or for any other purpose which is not permitted under relevant PRC laws or regulations. For the purposes of this disclaimer, "PRC" refers to the mainland of the People's Republic of China, excluding Hong Kong, Macau and Taiwan.