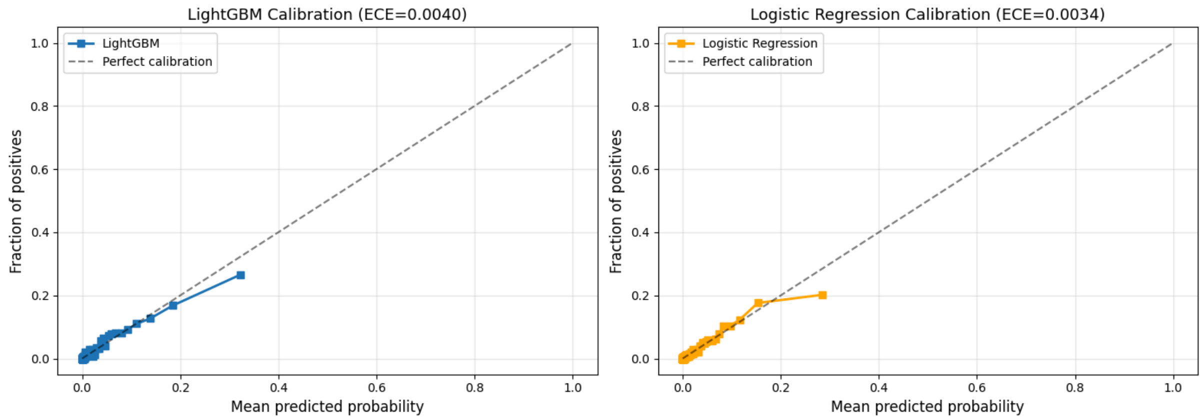
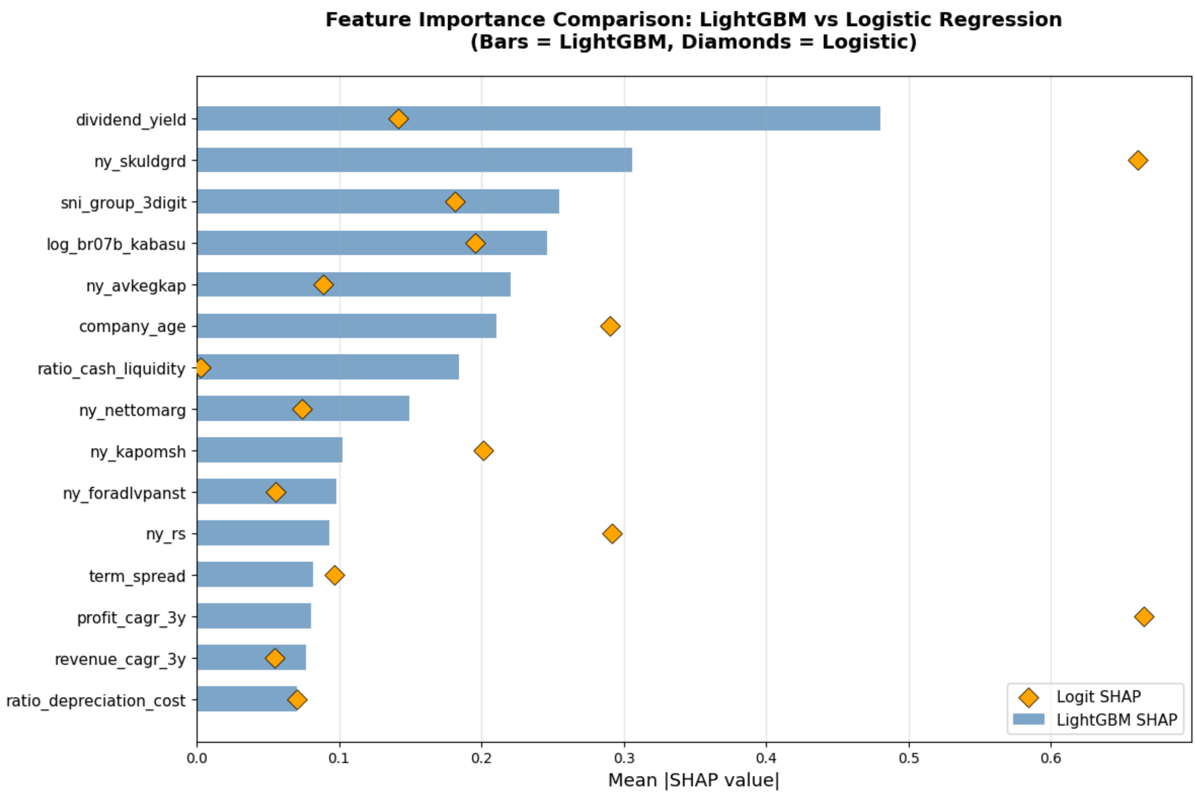
TABLE 1: MODEL PERFORMANCE COMPARISON

Model	AUC	PR-AUC	Brier Score	ECE
LightGBM	0.896632	0.164758	0.015836	0.003960
Logistic Regression	0.874605	0.128417	0.016276	0.003367
Δ (LightGBM − Logit)	0.022027	0.036341	−0.000440	0.000593



From Table 1 it is evident that LightGBM delivers higher discriminative power through its AUC (0.8966) and PR-AUC (0.1648) exceeding the logistic model by roughly 2 respectively 3.6 percentage points. This indicates tree boosting capturing nonlinear structure that is absent in the linear baseline. The Brier scores confirm that the LightGBM probabilities remain slightly closer to observed defaults and calibration remains strong for both models with an ECE below 0.01 but the logistic regression retains a marginal edge here.

The calibration curves show how the numbers translate visually. LightGBM’s curve lies close to the diagonal through most of the values but bends slightly below in the highest quantiles, indicating a mild tendency to overpredict risk for the most distressed firms. Similarly, logistic regression tracks the diagonal well in the central bins but flattens toward the extremes, reflecting its limited flexibility to capture highly nonlinear tails.



The figure compares feature importance between LightGBM and logistic regression using mean absolute SHAP values. Each bar/diamond pair shows how much a predictor contributes on average to the model's log-odds (the quantity models operate on before converting scores into probabilities). The features are ordered according to LightGBM's top drivers, which produces the clean descending pattern among the bars. Large bars indicate that LightGBM relies heavily on a predictor, whereas the position of the diamond reflects the logistic model's corresponding emphasis. Because SHAP values are additive and comparable across models, the plot highlights not only where the two approaches agree but also where LightGBM captures additional nonlinear or interaction-based patterns that the linear model cannot express.

The feature-importance comparison shows substantial divergence between the two models. LightGBM and logistic regression rely on different predictors to substantially different degrees and their rankings agree only partially. Dividend_yield is by far the most influential variable for LightGBM, yet the logistic model assigns it far less weight which signals how much the boosted trees extract predictive value from payout behaviour. LightGBM places ny_avkegkap and ratio_cash_liquidity among its top drivers, whereas their logistic SHAP values disagree.

The logistic model gives its highest emphasis to ny_skuldgrd and profit_cagr_3y. While ny_skuldgrd is also important for LightGBM appearing as its second-strongest predictor, the magnitude is substantially smaller. Overall, LightGBM never assigns mean SHAP values above 0.5, whereas logistic regression produces two such high-magnitude predictors. This contrast reflects how LightGBM distributes influence broadly across many features, often through nonlinear patterns whereas the logistic model places more concentrated weight on specific linear effects.

FIGURE 3A: LightGBM Feature Importance and Impact Direction
LightGBM - SHAP Summary (Beeswarm)

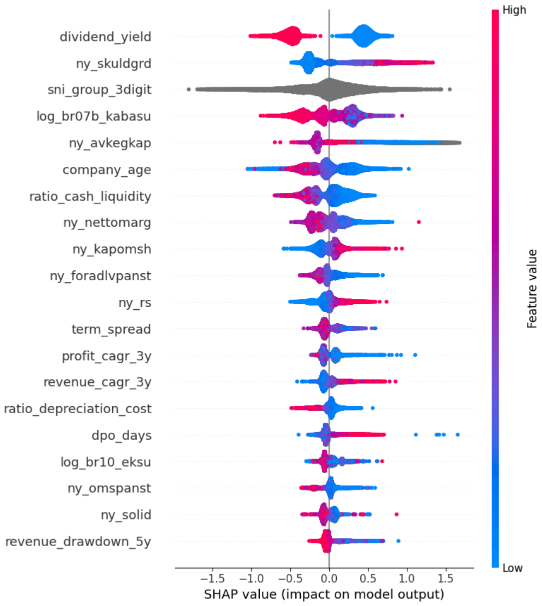
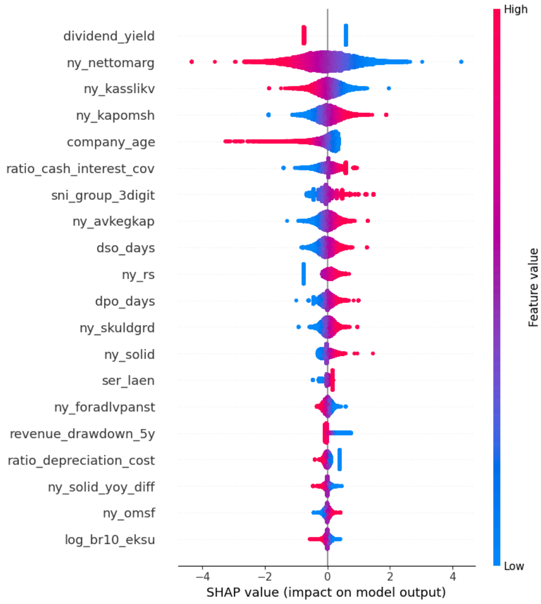


FIGURE 3B: Logistic Regression Feature Importance and Impact Direction
Logistic Regression - SHAP Summary (Beeswarm)



Overall, kanske onödigt mycket: The SHAP summary plots visualize how each predictor affects the model's output by displaying the full distribution of SHAP values across all observations. Each dot represents a firm-year, with the x-axis showing how much that observation shifts the predicted default risk (in log-odds), and the color indicating whether the feature value is low (blue) or high (red). Features appearing toward the top are those with the largest average impact on the model's predictions. Because SHAP is model-agnostic and locally additive, the plots make it possible to see both the strength and direction of each predictor's effect. While also illustrating how these effects are expressed within each model, linearly in the logistic regression and potentially in more complex forms in LightGBM.

Specifikt:
En övergripande skillnad är att lightgbm är mycket spräcklig. Mer detaljer i de enskilda fallen.
logistic ger inte de spräckliga. mer basic och generaliserat.

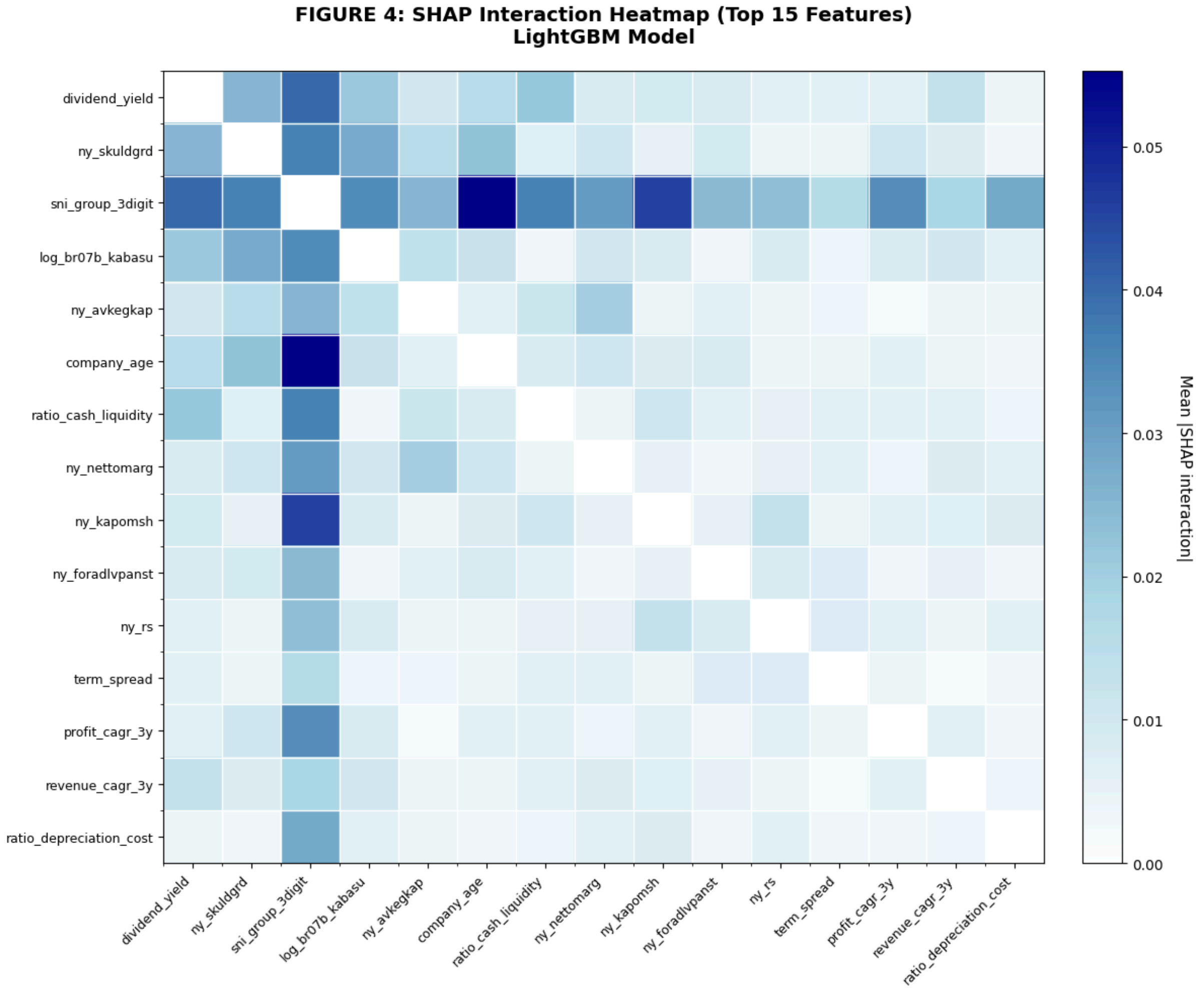
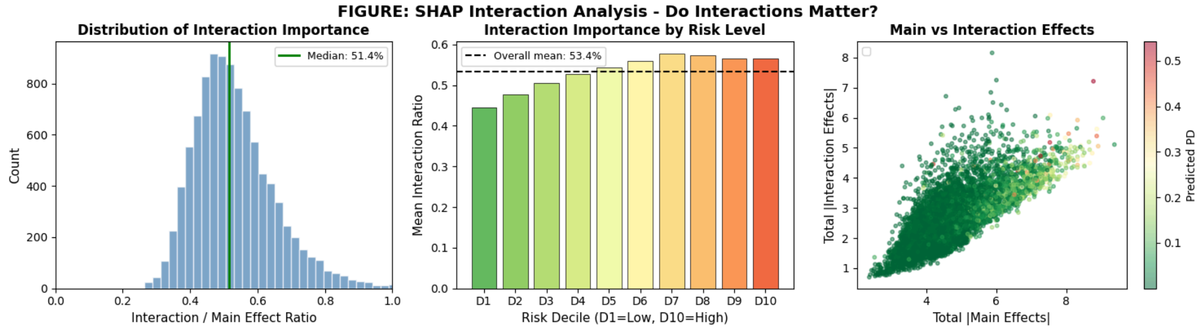
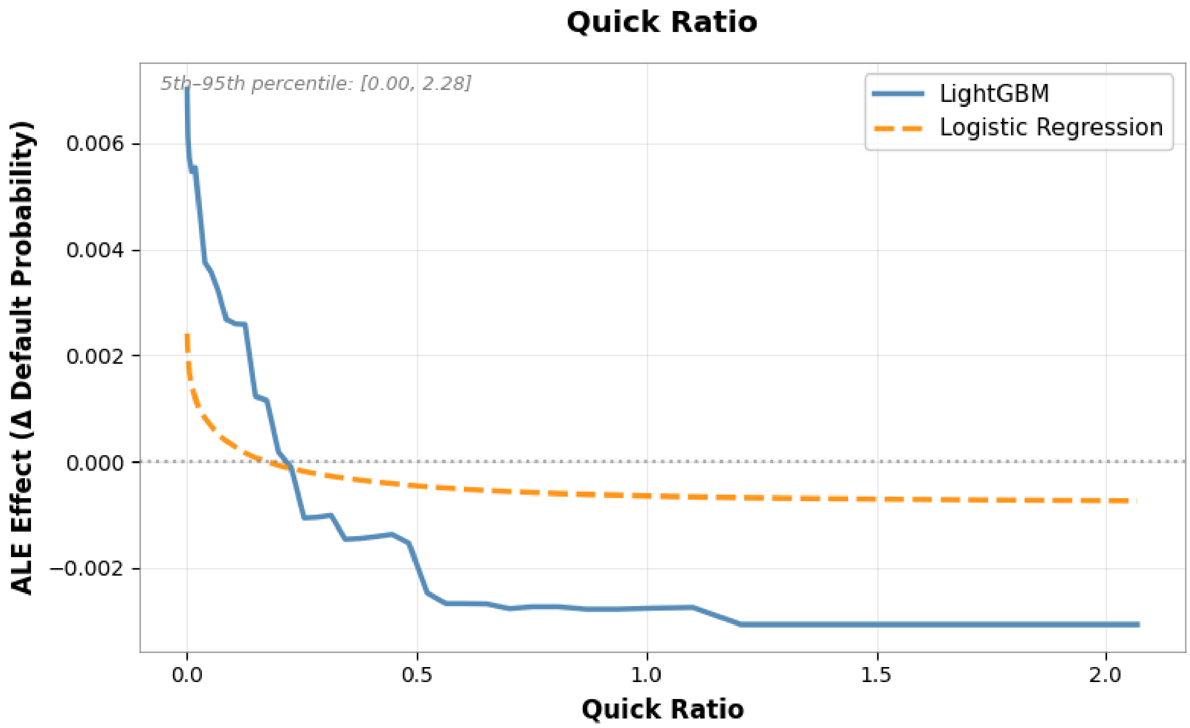


Figure 4 presents a SHAP interaction heatmap for the top 15 LightGBM features, showing how much each feature pair jointly influences the predicted log-odds beyond their individual main effects. Most cells remain very light, indicating that interaction effects are generally weak. The main exception is the row corresponding to sni_group_3digit (industry code), which shows darker cells. Two feature pairs stand out in particular, sni_group_3digit with company_age and sni_group_3digit with ny_kapomsh (capital turnover). This indicates that industry code in interaction with firm maturity and capital efficiency influence default risk.

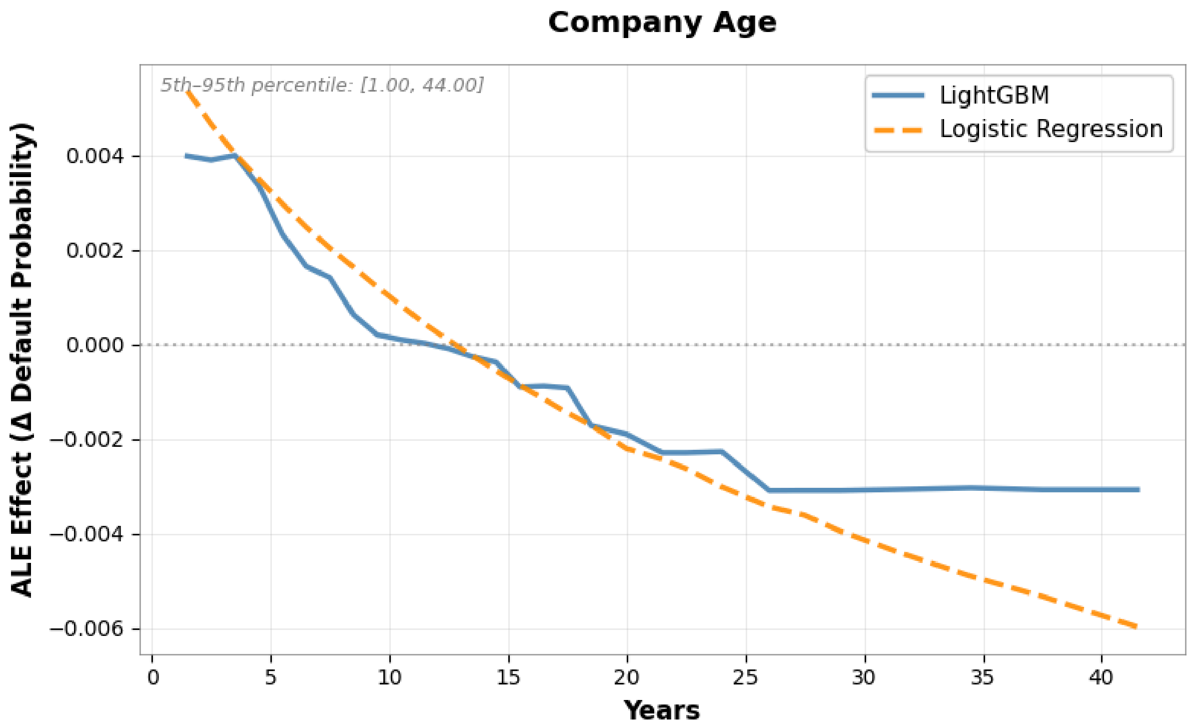


These figures show that the model is not purely driven by main effects, interaction effects are far from minor. The left histogram demonstrates this clearly, the median interaction-to-main-effect ratio is 51.4%, meaning that for half of the firms, interaction contributions are at least half as large as their individual SHAP effects. In the middle panel (overall mean 53.4%), we see that interaction importance rises slightly with risk level, but it remains consistently high across all deciles. Interactions become somewhat more influential in high-risk firms, yet the increase is modest. The right-hand scatterplot reinforces this pattern, showing that higher predicted default risk is associated with larger main effects and larger interaction effects alike.

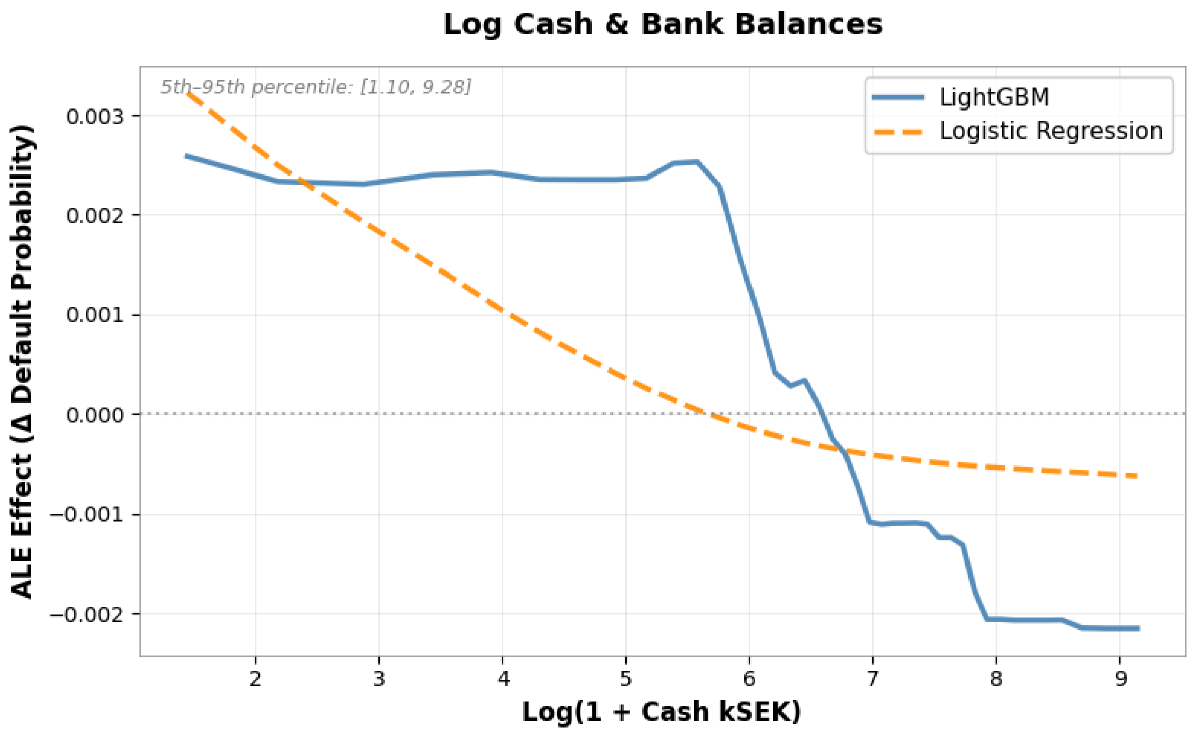
Taken together and in contrast to Figure 4, which highlights only where the strongest pairwise interactions occur, this figure quantifies that even though only a few feature pairs stand out, interaction effects as a whole are on average about half as large as the main effects. This makes interaction structure essential for interpreting the model.



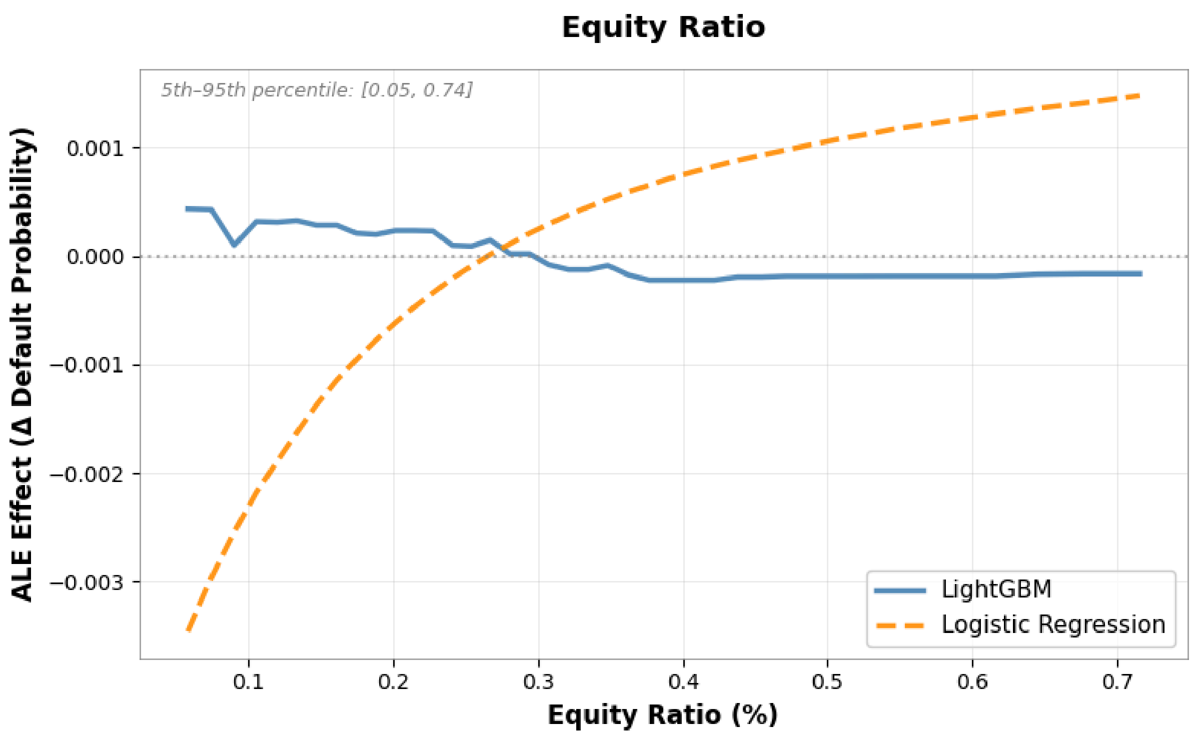
LightGBM's ALE curve declines sharply until a quick ratio of roughly 0.5. Below about 0.3, the effect rises steeply into positive ALE effects, the lower the liquidity the larger predicted default risk. Above the 0.5 threshold, the curve flattens and stabilizes. The logistic regression curve levels off much earlier and never drops as quickly and once it flattens it stays very close to zero. Both models therefore use the quick ratio primarily to identify firms with very constrained liquidity and neither provides additional risk reduction once the quick ratio exceeds approximately 0.5. LightGBM exhibits a clear threshold, while the logistic model shows the same pattern but in a less detailed way. In both curves, the lines cross the zero ALE level at around a quick ratio of 0.25.



The ALE plot for company age shows a broadly consistent pattern across the two models. Younger firms carry higher default risk and both curves fall noticeably until around 13 years where the effect turns negative which aligns well with financial intuition. LightGBM displays a clear threshold where the curve flattens after roughly 25 years. This indicates that additional firm age does not meaningfully reduce risk. By contrast, the logistic regression does not show a comparable flattening and continues to decline at a steady rate. Logistic regression treats age as a continuous factor, whereas LightGBM captures a more realistic diminishing effect where the marginal benefit of age eventually stops.

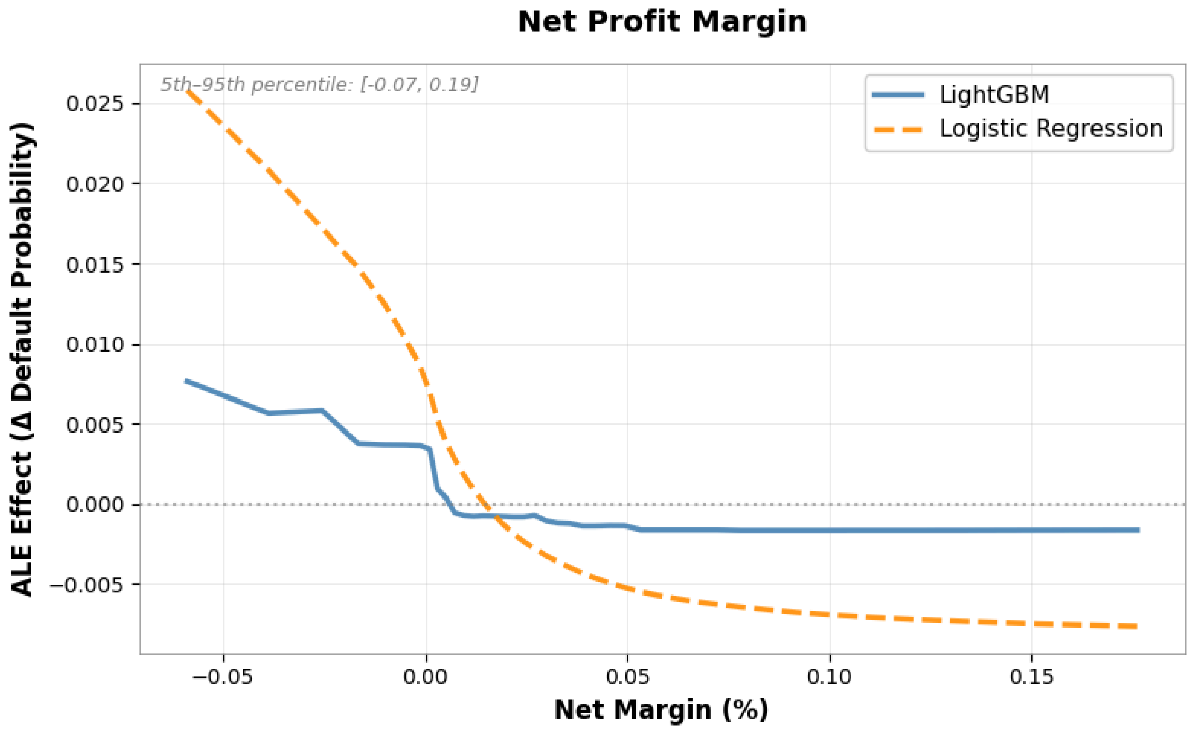


LightGBM shows a clear threshold at around 6 on the x-axis, which corresponds to roughly 400,000 SEK in cash and bank holdings. Once this level is reached, the ALE curve drops sharply indicating a substantial reduction in default risk when firms hold a buffer above this point. Logistic regression, by contrast, exhibits no such threshold. It follows a smooth pattern where more cash simply lowers risk in a gradual and linear way. Both models agree on the overall direction of the relationship, higher cash balances reduce default risk but LightGBM suggests that this effect only becomes important once liquidity exceeds a certain threshold-level.



KONSTIG
Big difference
lightgbm är väldigt plan jämfört med benchmark.
for lightgbm, Detta betyder att LightGBM knappt använder equity ratio som en viktig riskfaktor.
possible that the tree already uses other capital structure variables?

logistic
motsatt den finansiella logiken, visar begränsningen med logistic?
den logiskista modellen måste ge varje variabel en enda lutning.
Logit-modellen misstolkar equity ratio och skapar ett felriktat samband där högre soliditet kopplas till högre konkursrisk, medan LightGBM i stort sett avfärdar variabeln som irrelevant.
Kan förklaras av multikollinearitet? att den riktiga effekten mäts i en annan variabel?



As long as the firm is profitable, even slightly there is very little remaining risk for LightGBM to remove. Once profitability is above zero, the LightGBM curve remains almost completely flat. This shows that higher margins do not significantly lower the risk of default, what matters is simply that the margin is not negative. The logistic model shows much higher risk when the net margin is negative. Around the point where the margin turns positive, its ALE effect switches sign, indicating that the default risk starts to decrease rather than increase as it does in the negative region. LightGBM treats profitability primarily as a warning signal when it is below zero, whereas logistic regression interprets it as a continuously protective factor even where LightGBM shows no further reduction in risk.