Evaluating the Contribution of Open Banking Data to Credit Scoring Performance in the Spanish and French Markets

Abstract

The integration of open banking data into credit scoring models has garnered increasing attention from financial institutions as they seek to enhance risk assessment capabilities. This interest is driven by the potential of open banking data to provide a more granular and dynamic representation of borrowers' financial behavior, supplementing traditional sources of data. The ability to access transaction-level data enables lenders to infer patterns for dimensions such as income stability, discretionary spending, and financial resilience, which may be particularly valuable for assessing creditworthiness in populations with limited credit histories. As financial institutions seek to leverage these new data streams, there is the question of determining the extent to which open banking data contributes to model performance, particularly in markets where other sources of information are already well-developed.

This study examines the incremental predictive value of open banking data in credit risk assessment within the Spanish and French markets. To quantify its contribution, we develop two credit scoring models for each market: one that integrates open banking data and another that excludes this data source. Both models are constructed using an identical customer pool, ensuring that any observed performance differences can be attributed solely to the presence or absence of open banking data. The model performance was assessed using the rank-ordering GINI coefficient. Contrary to the prevailing assumption that open banking data significantly enhances credit scoring models, our findings indicate that substantial predictive accuracy can be achieved even in its absence. The analysis of the Spanish and French acquisition models reveals that while incorporating open banking data results in a measurable increase in GINI, the observed uplift remains modest.

Furthermore, our results highlight that a well-designed credit scoring model leveraging robust feature engineering and mobile app usage data can achieve comparable predictive performance even in the absence of open banking data. These findings challenge the notion that open banking data is indispensable for effective credit risk assessment in thin file populations and underscore the need for further research into its optimal application. Additionally, investigating the heterogeneity of open banking data impact across various borrower segments may offer deeper insights into its role in credit risk modeling.

Authors & Affiliations

Mr Francisco António Teixeira Mendonça[1], Mr Francisco Javier Ocáriz Gallego[2]

[1]Revolut Group Holdings Ltd, Lisbon, Portugal. [2]Revolut Group Holdings Ltd, Madrid, Spain