Predicting Financial Collapse Using Economic, Social, and Environmental Indicators

Author:
Naeem Almohtaseb, Viswa Sushanth Karuturi

1. Introduction

Financial collapses are critical events with far-reaching consequences for economies and societies. While traditional economic indicators like GDP growth and unemployment rates are commonly used to predict financial instability, the role of social and environmental indicators remains underexplored. This study investigates whether a combination of economic, social, and environmental indicators can improve the prediction of financial collapse. For this study, financial collapse is defined as a systemic banking crisis characterized by widespread bank failures, government interventions, or sovereign defaults, as classified by the IMF Systemic Banking Crises Database.

Our research question is: ***Can we predict the likelihood of a country experiencing financial collapse using a multilinear regression model with economic, social, and environmental indicators?*** We hypothesize that social and environmental indicators will provide additional predictive power beyond traditional economic metrics, offering early warning signals for financial collapse. This research is significant for policymakers and economists, as it may lead to more robust early warning systems for financial instability.

---

2. Methods

2.1 Data Sources

Data will be sourced from reputable online databases:

- **Economic Indicators**: GDP growth rate, unemployment rate, public debt-to-GDP ratio, and interest rates from the World Bank and IMF.
- **Financial Indicators**: Non-Performing Loan (NPL) ratios from the IMF Financial Soundness Indicators.

- **Social Indicators**: Divorce rates from UNECE, gambling/lottery data from the World Lottery Association, and crime rates from UNODC.
- **Environmental Indicators**: Natural calamities data from EM-DAT.
- **Political Stability**: Scores from the World Bank Worldwide Governance Indicators.
- **Target Variable**: Financial collapse indicator (binary) from the IMF Systemic Banking Crises Database.

2.2 Data Collection Process

Datasets will be downloaded and merged into a single CSV file using country-year as the primary key. Missing data will be handled through imputation (mean value for the region) or exclusion if missingness exceeds 20%. Variables will be standardized for comparability.

2.3 Variables Involved

- **Quantitative Predictors**: GDP growth, unemployment, public debt, interest rates, NPL ratio, divorce rates, gambling/lottery volume, crime rates, and natural calamities frequency.
- **Categorical Predictors**: Political stability score (ordinal).
- **Target Variable**: Financial collapse (binary).

2.4 Potential Sources of Bias

- **Selection Bias**: Countries with >20% missing data will be excluded.
- **Measurement Bias**: Single reputable sources will be used for subjective indicators.
- **Temporal Bias**: Data will be aligned with consistent reporting standards.

2.5 Addressing Collinearity

Collinearity between variables will be addressed by:

1. Calculating **Variance Inflation Factor (VIF)** and removing variables as needed.
2. Using **stepwise selection methods** (e.g., backward elimination based on AIC) to refine the model.

2.6 Reproducibility

All datasets are publicly available, and the merging process is documented in the R code provided in the appendix. Variable definitions and units of measurement will be clearly stated in the data dictionary. The code includes detailed comments to ensure reproducibility, and all code and datasets will be made available through a public GitHub repository to ensure transparency and ease of replication.