Midterm Project:

Behavior of World Bank Loan Repayment by Country Income Level

William Simpson

Dept. of Computer Science & Electrical Engineering

University of Maryland, Baltimore County

DATA 601: Introduction to Data Science

Prof. Mehmet Sarica

March 3, 2021

Introduction

As one of the world's leading international financial institutions, the World Bank Group disburses over $43 billion a year to some of the poorest countries to aid their development (World Bank, 2020). The movement of funds from World Bank donors merits analysis to ensure loan efficacy, financial responsibility, and transparency. Of particular interest is the repayment of World Bank loans and the factors that impact how recipient countries pay back the World Bank's investment in their development. Understanding this is crucial to effective and efficient allocation of World Bank resources.

This project aims to investigate country behavior in repaying development loans by examining historical loan statements from the World Bank's International Bank for Reconstruction and Development (IBRD) (World Bank, 2021). Utilizing Python, I specifically look for a potential correlation between country income level and loan repayment behavior. As the metric of a country's income level, my analysis uses Gross National Income (GNI) per capita, which measures wealth by calculating the cumulative income earned by people and businesses that reside within a nation's borders and dividing by population (Silver et al, 2021). In the context of this project, GNI is considered superior to the more common Gross Domestic Product (GDP) measurement because GNI more accurately captures wealth in nations that receive substantial development aid such as IBRD loans. To capture loan repayment behavior, I examine two metrics: 1) time to first installment of loan repayment, and 2) quantity of loan money sold to a 3rd party as opposed to paid back to the original lender (i.e., the World Bank) over the life of the loan.

Overview of Data

The primary analyzed dataset can be found in the World Bank's Data Catalog at the following link (https://finances.worldbank.org/Loans-and-Credits/IBRD-Statement-Of-Loans-Historical-Data/zucq-nrc3) (World Bank, 2021). The file size is approximately 369MB and the data consist of 1.005 million rows and 33 columns. After data processing, 5973 rows and 23 columns remained.

Data Processing

Prior to analysis, duplicate records and irrelevant dimensions of the data were eliminated from the dataset to isolate loan statements that had been successfully repaid and exclude those that had been cancelled or were still being disbursed. Values missing from the dataset were found to correspond to global-level, not country-level data and were thus dropped. To ensure the ease and accuracy of analysis, money values were converted to millions of dollars and dates were converted from simple text objects to datetime objects.

Following this, I merged a secondary World Bank dataset including countries categorized by income level (Prydz and Wadhwa, 2019) measured in GNI per capita, with the primary data. This dataset was selected due to its high level of granularity; for the period between 1989 and 2018, each country's GNI per capita is recorded along with the cutoff values between high-, upper-middle-, lower-middle-, and low-income brackets for that given year. This enables my analysis to account for fluctuation in a country's income level from year to year as well as the general global increase in GNI over time. For dates prior to 1989, the nearest recorded value in the secondary dataset was imputed. For dates after 2018, current income level categorization by the World Bank was used.

Analysis 1: Time to First Loan Repayment

Analysts at the World Bank should be aware of trends in how long it takes for a recipient country to begin repaying its loan once it obtains its first disbursement for several reasons. First, this information may supplement or inform other analyses on economic risk of default. Second, this data would be useful in forecasting future repayment behavior for a given country. The time elapsed from the date a loan was approved to the date of the first repayment is used in this analysis as a measure of speed of repayment. Grouping the loan statements in the dataset by country income level, and calculating the average time to first repayment, shows that middle income countries (upper-middle-income (UMC) and lower-middle-income (LMC)) tend to take longer to make the first repayment on their loans compared to high-income countries (HIC). Specifically, middle income countries require approximately an additional two years compared to HIC. This difference is illustrated in *Figure 1* and was found to be statistically significant ($p < 0.001$) using a one-way analysis of variance (ANOVA) test.

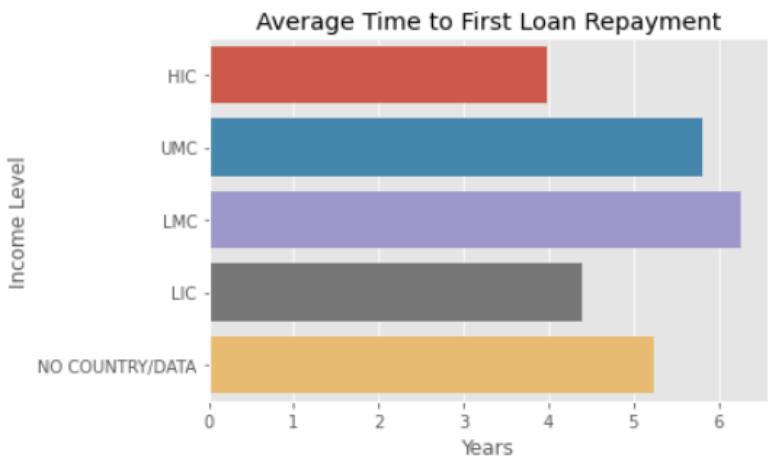


Figure 1.

Average Time to First Loan Repayment. High-income country (HIC), Upper-middle-income country (UMC), Lower-middle-income country (LMC), Low-income country (LIC). NO COUNTRY/DATA indicates no data existed in dataset or country no longer exists. Source: Author

One might posit that the time to first loan repayment would increase as GNI decreases, however, its notable that low-income countries (LIC) do not follow this pattern. LIC take approximately the same amount of time as high-income countries (4 years) to make their first repayment. One possible explanation for this would be that IBRD loans are designed to target

middle-income countries, not low-income countries. That is the role of the World Bank's International Development Association (IDA) loans. The data analysis is consistent with the fact that only *credit-worthy* low-income countries, meaning those that are not likely to default on loans and are able to pay them back, are being approved for IBRD loans. Therefore, it is reasonable to conclude that the low-income countries in this dataset of IBRD loans are intentionally selected for their ability to make repayment sooner and more reliably. When considering LIC are treated according to separate criteria, this analysis then shows a negative correlation between income level and speed of loan repayment for other country categories. Future research should focus on the underlying cause of this correlation. It is possible that higher income countries have more available funds with which to repay the IBRD, though subsequent analysis must also consider the effect of prepaid loans and the type of project for which the loan is earmarked.

Analysis 2: Loan Money Sold to 3rd Party

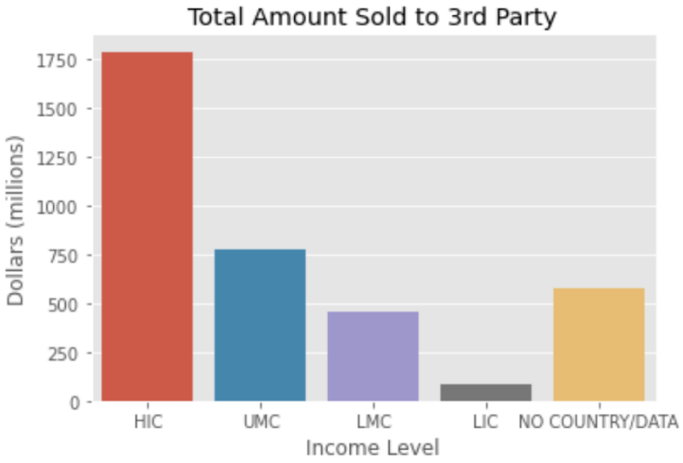
Another aspect of IBRD loan repayment enumerated in the primary dataset is whether the borrowing country sells a portion of their loan to a 3rd party. A country may choose to do so as a strategy in managing its risk of default on a given loan. Given that the average probability of high-income countries undergoing an external debt default crisis is less than half that of middle- and low-income countries, this analysis aims to test whether middle- and low-income countries sell more loan money to 3rd parties than high-income countries (Qian, Reinhart, & Rogoff, 2010).

As a first step, I calculated the total amount of loan money sold to a 3rd party across all loan statements in the dataset grouped by income level (*Figure 2*).

Figure 2.

Total amount of money sold to a 3rd party by country income level (1947 onward).

Source: Author



The results of this visualization, however, do not support the claim that middle- and low-income countries sell more to 3rd parties. To explain this, we can investigate how the amount of money sold to a 3rd party has changed over time. A plot of the yearly average portion of the original principal amount of the loan uncovers the existence of some notable outliers (*Figure 3*).

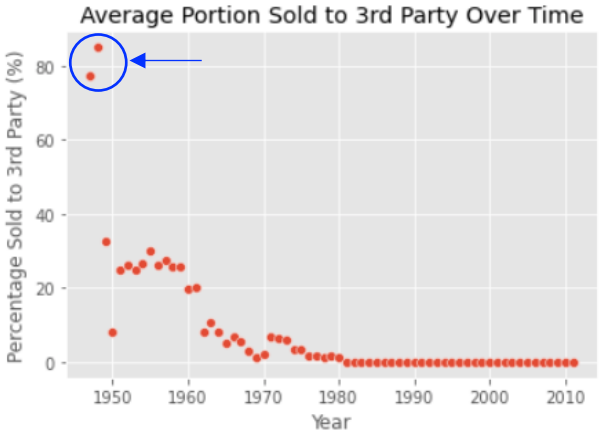


Figure 3.

On average, the portion of the original principal amount of the loan sold to a 3rd party has decreased. The first two datapoints circled in blue do not fit into the larger trend.

Source: Author

The first few data points (1947-1949) are likely obscuring any underlying trends in *Figure 2*. If we isolate these outliers (*Figure 4a*) it becomes clear that these datapoints represent almost exclusively HIC that received IBRD loans during the years immediately following WWII. While IBRD loans typically target middle-income countries, the World Bank's primary task during this period was to finance the post-WWII reconstruction of Western Europe, even though many of these countries, like France, were classified as HIC.

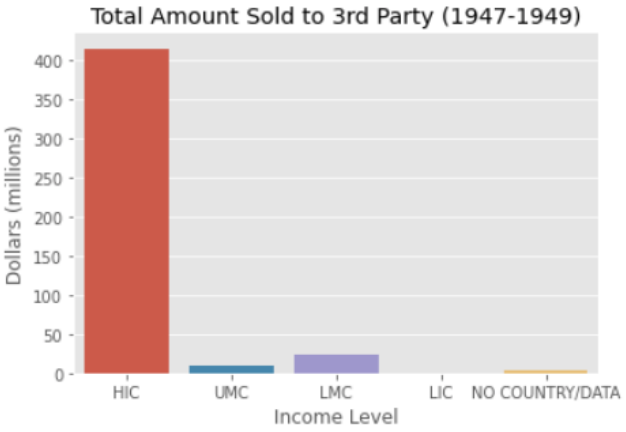


Figure 4a. ($p<0.001$)

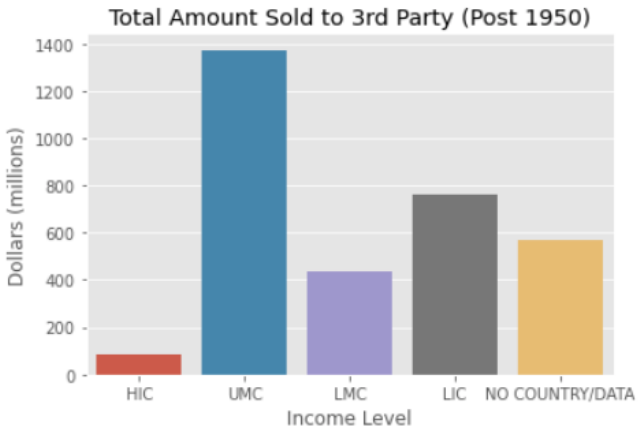


Figure 4b. ($p<0.001$) *Source: Author*

Furthermore, the loans during this time were substantial (e.g., $250 million), which likely inflates the amount of money sold to 3rd parties by high income countries in *Figure 2*. Post-1950, data (*Figure 4b*) reflect the expectation that middle- and low-income countries engage in the strategy of selling part of their original loan to a 3rd party more often than high-income countries. To support this conclusion, I found on average that the percentage of the initial loan sold to a 3rd party is highest (greater than 17.5%) for UMC compared to HIC (around 5%). Differences were determined to be statistically significant using ANOVA ($p<0.001$). Finally, further analysis of this characteristic of loan repayment might inspect the fact that *Figure 3* also depicts a notable drop in the sale of IBRD loans to 3rd parties after 1980.

Conclusions

Using historical data from the World Bank's Data Catalog, I analyzed two aspects of how countries repay IBRD loans, including the speed of repayment and the amount sold to 3rd parties. HIC were found to begin repaying loans an average of 2 years earlier than middle- and low-income countries, while UMC sold a significantly larger portion of the original principal loan amount compared to HIC. These findings lend themselves to additional analysis into the driving factors that cause them and can help inform future World Bank loan decisions and criteria.

References:

World Bank. 2020. World Bank Annual Report 2020. Washington, DC: World Bank. doi: 10.1596/978- 1-4648-1619-2 . License: Creative Commons Attribution–NonCommercial–NoDerivatives 3.0 IGO (CC BY-NC-ND 3.0 IGO).

Silver, C., Attkisson, A., Chen, J., Kagan, J., & Woolsey, B. (2021, April 27). *Gross National Income (GNI) Definition*. Investopedia. https://www.investopedia.com/terms/g/gross-national-income-gni.asp

World Bank. (2021, March 31). *IBRD Statement Of Loans—Historical Data | World Bank Group Finances*. World Bank Group Finances. https://finances.worldbank.org/Loans-and-Credits/IBRD-Statement-Of-Loans-Historical-Data/zucq-nrc3

Prydz, E. B., & Wadhwa, D. (2019, September 9). *Classifying countries by income*. The World Bank. https://datatopics.worldbank.org/world-development-indicators/stories/the-classification-of-countries-by-income.html

Qian, R., Reinhart, C., & Rogoff, K. (2010). *On Graduation from Default, Inflation and Banking Crisis: Elusive or Illusion?* (No. w16168; p. w16168). National Bureau of Economic Research. https://doi.org/10.3386/w16168