Investigation, Simulation and Optimisation of the Car Financing Market

Sopra Banking - Group 4

Adam Morris, Ellie Ryan, Conrad O'Driscoll, Matthew Shipley, Max Williams and Antony Stavrou

EMAT30005 Mathematical and Data Modelling

Department of Engineering Mathematics, University of Bristol

June 23, 2023

1 Introduction

Last year, the United States reached over \$1.4 Trillion in outstanding auto-loan debt, its highest level of all time [1]. The industry for car loans has been on a path of exponential growth since the lowering of central reserve interest rates in late 2008 [19]. However, loan delinquencies and defaults have been a historically significant cause in the bankruptcy of many lending establishments, and even entire economic recessions. Therefore, there exists a large incentive to be able to responsibly manage a portfolio of loans as to minimize the risk of defaults, whilst also benefiting from the profitable nature of asset backed lending. The aim of this project is to produce a model that can simulate different lender strategy scenarios, as well as produce the revenue data seen from these market outcomes.

In order to manage the risk of these portfolios, lenders must have an effective gauge on the risk level of each additional loan they issue. The processes by which they do this is to assess the risk of each prospective borrower before they can gain access to credit. By only accepting loans within an uncertainty tolerance that suits their risk appetite, they can ensure that the liability of the entire portfolio is minimised.

With such a large volume of business, with 83.67% of new cars being purchased with some type of financing, lenders are not often pressured to compete on price, nor compromise their standards on risk appetite in order to attract business [5]. In practice, different types of lenders service different tiers of customers. This leads to their behaviour within the market being flexible in pricing but not heavily varied, in terms of the customers they target. The bigger banks will stand by the target market of low risk customers however smaller banks that have a higher cost of credit will often have to utilise higher risk strategies to deliver a sustainable profit. There is an inherent imbalance in market shares across the banking sector with a tendency for competitors to follow the 'broken stick' distribution [11] [3]. Smaller lenders must therefore resort to alternative strategies to increase profits. One such example is outlined in the paper by Giovanni Dell'Ariccia and Robert Marquez, where lenders utilise informational structures of loans to modify their strategies [12].

With the average US household disposable income reaching \$3,856 in 2021, the budget with which they can service their debt is shrinking [17] [9]. Because of this, borrowers are partially restricted in their behaviour, within the market. More specifically, they always go for the cheapest offer they are presented with. The cheapest price they are offered with is a direct result of their credit worthiness and financial position. Some papers have investigated the effect of borrowing constraints on the behaviour of the customers especially those of lower income backgrounds [7].

Borrowers have a wide variety of options for whom they can try to obtain auto finance from, but the majority of automotive finance providers typically fall into three categories. The first of these are the manufacturers themselves who lend through their own captive finance company for example Toyota and Ford. Secondly, there are independent financial companies, some of which are wholly owned subsidiaries of major banks, such as Ally Financial or Wells Fargo. The third of these is a contract hire and leasing company that specialise in the provision of leasing agreements for a range of customers and vehicles for example, Lex Autolease or BT Fleet. Despite being of different origins, these types of finance providers typically follow the same process when reviewing a loan application. These lenders use the requested information provided by the applicant and a report is compiled by a third-party company, Experian Credit Score as a primary example, to gauge the credit worthiness of the applicant.

2 Processes in Applying for Car Financing

A prospective borrower in the process of buying a car has some key decisions to make, before proceeding to submit a finance application. The borrower needs to decide which car they want to purchase, with respect to their budget. Then the borrower should decide over how long they wish to repay the loan, referred to as the term. These periods typically vary from 1 year to 4 years. With these decisions made, a borrower is ready to make an application for their finance. They will make this application either directly to a lender or via a third party, who applies to multiple lenders at once. The applicant will often need to provide extensive information, as well as additional documents, as a form of verifying the information they have provided.

2.1 Verification process

In order to ensure the application contains valid information and the supporting documents are of genuine origin, the lender will process the application through a verification process.

This process is an increasingly important step in any loan application. The introduction of international standards for 'Anti Money Laundering' and 'Know Your Client' means that tight regulation has trickled down to the consumer level [10]. The additional verification is useful, not only as a required anti-fraud measure but, in increasing the certainty in the purported applicant information which relates to their credit worthiness as well. When banks receive loan applications, they have the freedom to decide how strict the verification process should be, that is how many documents they require to be verified. Due to larger banks by market share typically possessing a smaller appetite for risk, they typically only lend to the safest of borrowers who are deemed the most credit worthy through the information they provide in the application and the verification process. The disadvantage of this strict risk appetite is the increased verification time frame for these lenders. The upside; they can be more competitive in their interest rates offered to attract more customers due to their more accurate depiction of customer risk.

2.2 Loan Agreement Generation

A third party uses the information from the verification process to evaluate the applicant's credit tier using an existing model. The credit tier derived from the model gives the lender an indication as to whether a borrower would be a suitable loan applicant. In the event that the borrower is deemed worthy enough to lend to, the lender will compile a finance offer in the form of a credit agreement. This credit agreement includes the terms upon which the loan is offered and specifies at least the terms for the following properties of the loan: the credit amount, the duration, the type and the amount of interest rate. The loan agreement will then be presented to the prospective borrower and the borrower will then assess its affordability and make a decision whether to take on the loan or not.

3 Model Construction

3.1 Assumptions

In this section, the main assumptions taken throughout this approach are presented. These assumptions serve to reduce complexity while preserving as much realism as possible.

These assumptions are as follows; assume the economic environment remains the same during the simulation as well as the personal circumstances of borrowers resulting in a an unchanging default probability through a loan's lifetime. Secondly, we assume that the only loans available are fixed rate and so the monthly payment amount is constant for the duration of the loan. Borrowers are assumed to have no financial history and will not be recycled since they will not take out two auto loans within 10 years [4]. We also assume that the borrowed amount is equivalent to 100% of the cars purchase price such that borrowers do not have the option to pay a deposit on the loan or make pre-payments. We will also assume that each lender will rank borrowers credit worthiness from a third-party company such that all lenders have the same default probabilities associated with the same borrowers. Finally, there is no delay in the time between a borrower defaulting on their loan and the lender recovering the present value of the collateral for the loan; the current value of the car.

3.2 Car Financing with One Lender and One Borrower

This model explores the fundamental interactions on a one to one basis. It is assumed that this model includes a lender that provides unlimited funding, and will provide a loan offer regardless of the borrowers' credit risk. The objective of this model is to explore the loan repayments cycle of a single borrower to a single lender.

The process begins by calculating the required monthly payments. This is achieved using the PMT formula in equation (1), where A is the monthly contributions based on constant payments, P is the principle value of the asset (in the same currency as A), r is the interest rate per month and n is the duration in months.

$$A = P\frac{r(1+r)^n}{(r+1)^n - 1} \tag{1}$$

Each month, the borrower makes the appropriate payment to the lender over the course of the loan. The process ends when the loan has either been fully re-payed, or interrupted due to defaulting. If defaulting occurs, the loan is stopped, and the lender collects 80% of the current asset value as 80% is a standard recovery rate across the industry [4]. The value of the asset collected C, at a given month n, for an asset of principal value P is calculated using equation (2). The values for depreciation rate r_d are also included [6].

$$C = P\left(1 - r_d\right)^n \begin{cases} r_d = \frac{25}{12} & \text{for } n < 12 \\ r_d = \frac{78}{125} & \text{for } n > 12 \end{cases} \tag{2}$$

3.3 Introducing Multiple Borrowers

To make the first model more realistic, multiple, more representative buyers, that a lender can process are introduced. The verification process is modelled using a set of three documents which are assigned with a fixed probability relating to the chance that the borrower submits the application with the document correctly. If any of the three documents are submitted incorrectly the buyer is rejected.

In this model, each borrower is generated with a list of features that in practice the lender would receive from the documents provided through the verification process. Credit tiers and rates of defaulting are assigned to each borrower using a logistic regression model trained on a pre-existing data-set, including if a borrower deafulted or not [18].

To determine the features in which the credit tiers should be assigned upon, Serrano-Cinca's 'Determinants of default in P2P lending' highlights annual income and current housing situation to be important factors [20]. The data-set was cleaned to include these features as well as marital status, family members, age, and length of employment. Features such as bike ownership and mobile phone number were deemed to be irrelevant in determining the credit tier of a borrower and rows with missing features were removed to decrease the bias, hence increase the accuracy of the model. This cleaned data set was used to train a logistic regression model to assign probabilities to the binary outcome of a borrower defaulting. Such models prove to be a popular approach in credit risk modelling, as shown in Zahia's and Achchab's 'Modeling car loan prepayment using supervised machine learning' [22].

To train and test the accuracy of the model which is to be used by all lenders, as in assumptions section 3.1, 16,086 data points were split between the training set, using 80%, and the test set, using 20% of the data points. Using the test set to evaluate the model, an accuracy of 84.5% was recorded in classifying whether a borrower defaulted or not.

To generate the features of each borrower, a log normal distribution, using the mean and standard deviation from US data, was used to generate the income [9]. The house status attribute determines whether borrowers do or do not own at least one house, using the data from the home ownership rates in the US in 2005 [13]. Marital status follows a similar method, with a weighted distribution, using four different statuses; married, divorced, single and widowed [15]. The number of family members uses data that distributes the household size in the US, between 0 to 6 members [14]. To produce an age value for each borrower we used further data distributions from 2005 US economic data to assert weights to sets of ages, between 18 to 64 [8]. Length of employment is a feature that had to be combined with each set of ages. This implies that, for each set of ages, it's allocated a number of employed days based on the data from Employee Tenure in 2022 [2].

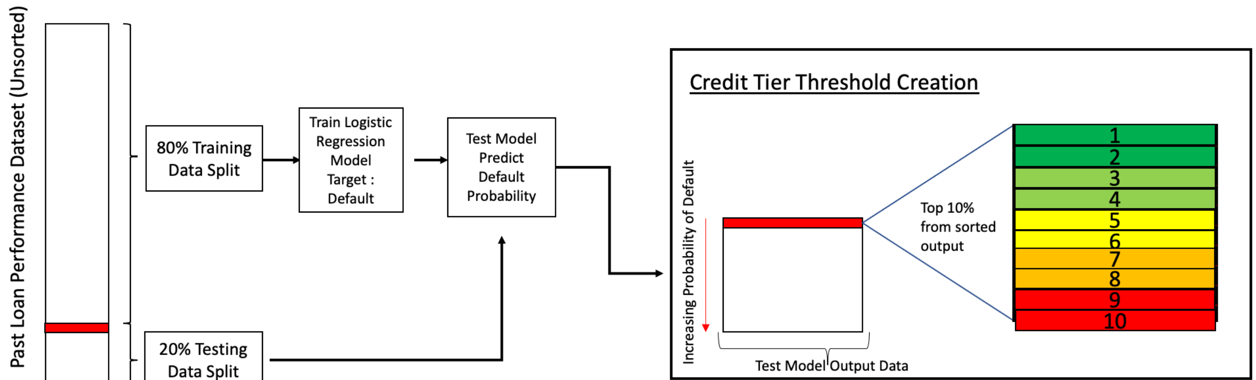

Figure 1: Defining credit tiers to be used when classifying generated borrowers

As in figure 1, to rank each generated borrower, a credit tier system is introduced, where tier 1 corresponds to the borrowers with the lowest probability of default and tier 10 the highest probability of default. The logistic regression model is used to not only calculate the probability of default for generated borrowers, but also determine the probability of default at the upper bounds of the credit tier thresholds. Borrowers with a default rate greater than 10% are deemed too risky for any bank as it would not be profitable over time, therefore do not pass through the tier assigning system [4].

To define the upper bound default probabilities for each tier, the model is used to predict the probability of default for each item of test set which has a default rate of below 10%. The outputs of the default rate probabilities are then sorted from lowest to highest (safest borrower to riskiest borrower). The program then takes and splits this sorted selection into 10 equally sized tiers and the predicted default probability for the riskiest borrower in each tier defines the upper bound in that credit tier. These upper bounds of default probabilities are then used to assign each of the generated borrowers a credit tier shown in table 1, where x is the default probability of a generated borrower.

Credit Tier	Default Probability
1	$0.00\% \leq x \leq 1.46\%$
2	$1.46\% < x \leq 1.98\%$
3	$1.98\% < x \leq 2.32\%$
4	$2.32\% < x \leq 2.67\%$
5	$2.67\% < x \leq 3.07\%$
6	$3.07\% < x \leq 3.45\%$
7	$3.45\% < x \leq 3.78\%$
8	$3.78\% < x \leq 4.11\%$
9	$4.11\% < x \leq 4.52\%$
10	$4.52\% < x \leq 10.00\%$

Table 1: Probabilities of default defining each credit tier

3.3.1 Results

Before developing the final stages of the simulation, it's vital to evaluate the current stages of the model, and what inferences can be made from such results. Figure 2 gives an insight into a typical collateral timeline for a large bank. Collecting collateral is an important aspect of a lenders portfolio, used primarily to offset any losses from a defaulted loan. As is demonstrated, the bank in question comfortably maintains a sufficient gap between losses and collected collateral, virtually eliminating the associated risk.

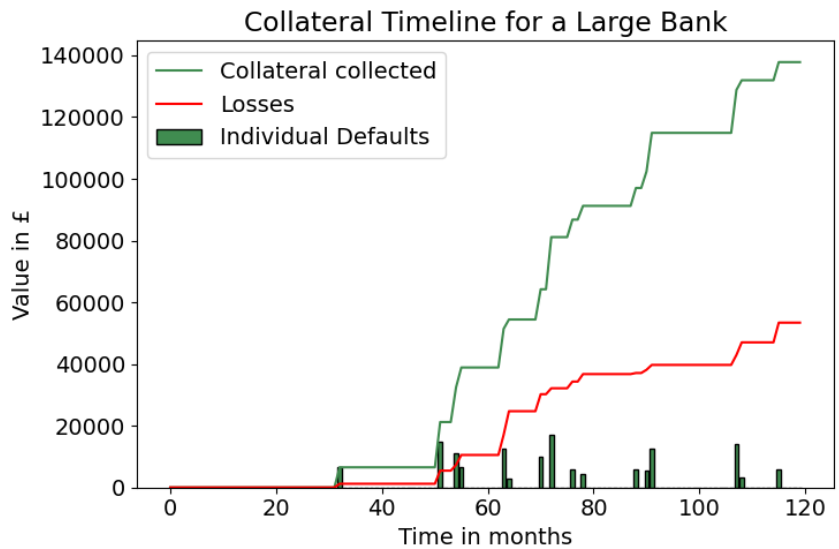

Figure 2: Collateral collected by a large bank, over a 10 year simulation. (Loss = loan remainder after default - collateral collected)

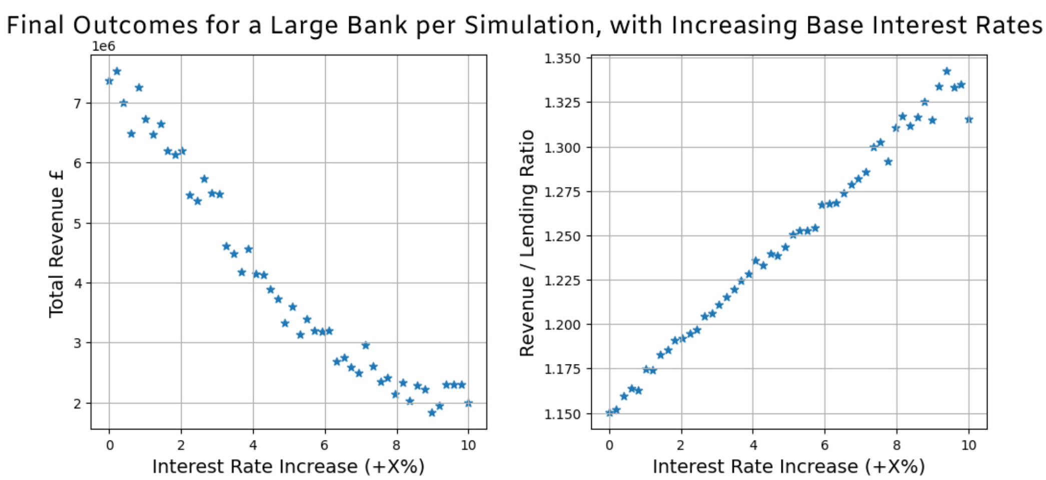

Figure 3: 50 sequential simulations each with rising interest rates for the generated lender.

In Figure 3, both the revenue to lending ratio and the total revenue have been plotted for 50 separate instances of the simulation for a large bank. Each iteration increases the possible offered base rates of interest by 0.2%, from an initial set of 10 discrete values ranging from 5-14%. The first iteration therefore has a prime rate of 5%, and worst of 14%, with the last having a prime of 15%, and worst of 24%.

As can be seen, lower interest rates produce higher absolute values of revenue, due to the greater affordability of borrowers, leading to larger distributions of customers. However, this in turn decreases the returns per person, hence a lower revenue to lending ratio is shown. Higher interest rates illustrate the inverse effect, suggesting the existence of a set of optimal values that maximises each function, with the lowest negative effect on one another. These values may be investigated by recreating these result over continuous ranges of interest rates. Due to the computational requirements needed to achieve sufficient accuracy, this has been avoided due to the scope and time-frame of this project. This result suggests further models should use a range of interest rates, for a large bank, beginning at around 9%, increasing to around 16%.

Although a single lender market is not a realistic scenario, this model highlights the importance of offering appropriate ranges of interest rates to borrowers. Additionally, lenders must decide what targets they wish to achieve within their portfolios; more profit per person or overall profit. Both objectives provide unique implications for future development. Maximising numbers of customers may prove more difficult in competitive environments and increase default risk, however maximising interest rates will allow undercutting from other lenders.

4 Simulating a Car Finance Market

This final model represents a fuller image of a real car-loan market. Multiple lenders all compete, in a common environment, in an attempt to maximise their own profit.

To create this competitive environment, 7 individually assigned lenders have been constructed. Within this, there are 3 categories; small (numbers 4-7), medium (numbers 2-3) and large (number 1) which have a market share corresponding to the broken-stick distribution as mentioned in section 1.

This allows for the simulation to follow real-world scenarios, where certain banks have greater lending capabilities and control. Each is prescribed a maximum lending capacity proportional to their market share. Furthermore, larger banks offer lower interest rates for borrowers with higher credit scores, and avoid those with the worst default percentages. Smaller banks are forced to take on higher-risk clients, therefore increase interest rates in an attempt to mitigate losses from defaulting. Specific market shares, interest rates and credit bands each bank has access to can be found in section A - Further Materials, table 2.

As mentioned in 2.1, the verification time delay is dependent on the size of the bank. To model this, a delay value, from a normal distribution around a given mean ranging from 2 days up to one month, is assigned to each bank depending on its size. A borrower has a prescribed patience tolerance, which is the maximum time they are willing to wait for their loan, again from a normal distribution within this range.

When presented with offers from all banks in the simulation, borrowers initially reject any lender with a verification delay greater than their patience tolerance. If any offers remain, the loan with the cheapest monthly payments is accepted, allowing borrowers some control over the process as well as leading lenders to change their strategies to out-compete their competition.

4.1 Market Strategies

In order to obtain a deeper understanding of the competitive nature of this market, it is critical to investigate and analyse the most popular and powerful strategies available to lenders. The chosen strategies under investigation are undercutting, collusion, dynamic pricing and reduced verification delays. Each scenario is run with identical seed values, ensuring that outcomes are a direct result of used strategies, and not due to different generations of borrowers and randomness.

4.1.1 Undercutting

Undercutting is a principle idea in a competitive client-based environment. It serves as a tactic to attract a larger market capitalisation by appealing to more consumers through significantly reducing prices. This however reduces the possible returns per customer.

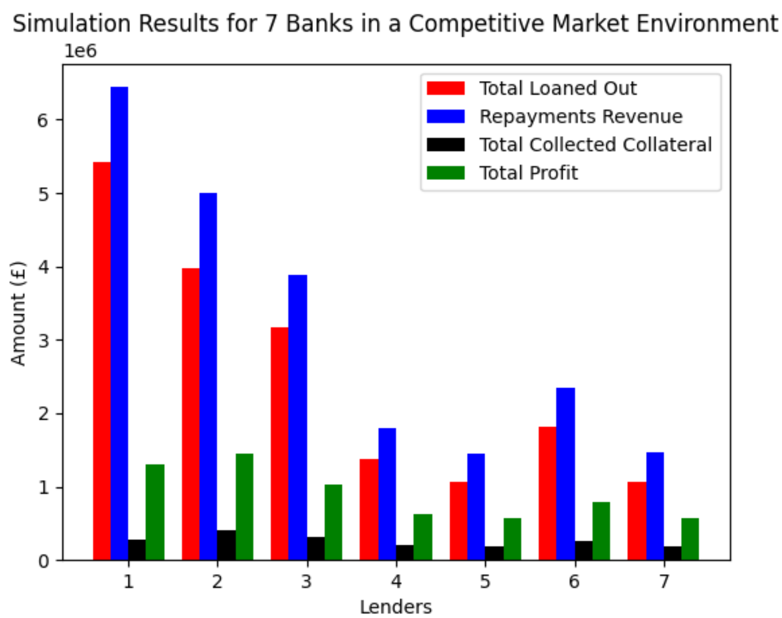

Figure 4: Banks 3 and 6 undercutting within their size bracket (Profit = Repayments Revenue + Collected Collateral - Total Loaned Out)

To analyse this strategy in practice, banks will only undercut those within the same size category. In figure 4, bank 3 has undercut bank 2 in the medium size band, and bank 6 has undercut banks 4, 5 and 7 in the small size band. The results show this strategy has clearly led to bank 6 out competing the other banks, even with a smaller market share than its band competitors. However, for the medium banks, bank 3 does not surpass bank 2. Though at first this may seem contradictory to the outlined hypothesis, there are additional factors explaining this result. When considering the bank sizes, it becomes clear that bank 3 reached its lending threshold, thereby rendering all other customers, for whom the medium banks were competing, to pick bank number 2.

4.1.2 Dynamic Pricing

In this strategy, banks in the same size band actively undercut each other in terms of interest rates offered, until they reach an equilibrium point. This equilibrium point is defined as the base rate percentages, which banks only slightly deviate from when competing with each other. This is a key and noteworthy idea, as banks cannot continue to undercut each other as this would result in all lenders ceasing to make profits. This allows banks to capitalize on a potential influx of customers by slightly reducing their rates to gain a competitive advantage.

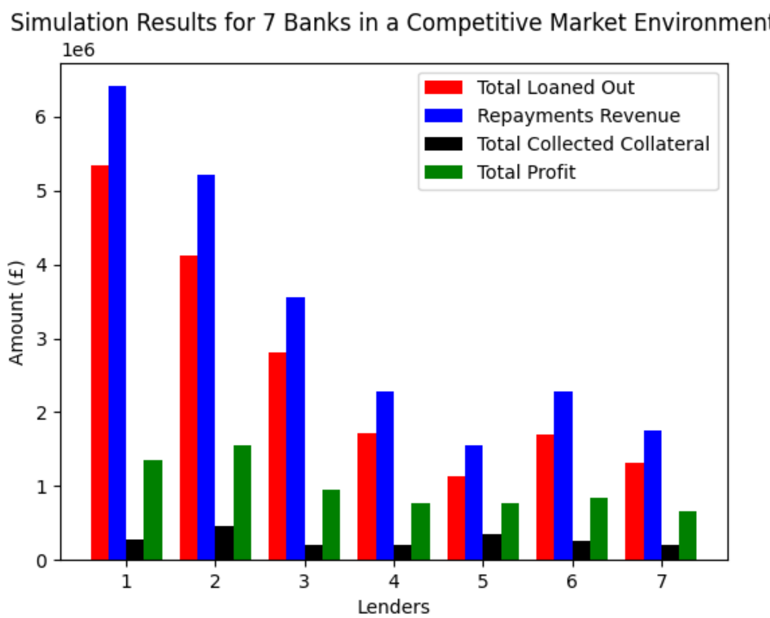

Figure 5: Dynamic Pricing Strategy across all Banks

In figure 5, the smaller banks are seen to have a more even distribution of market presence than seen in section 4.1.1. Lenders 4 and 6 have performed slightly better in this bracket, indicating an effective use of this strategy. Another cause of this result could be put down to the shorter time delays for the smaller banks as mentioned in section 4. This strategy appears to be less effective for individual banks, especially number 6. However, the collective market share has increased for small banks, allowing them to maintain a better position against larger banks. Eliminating competitors, within the same band, enables the remaining banks to absorb their competitors' clients. Given time, a bank may then be able to propel itself into the next band of relatively larger banks. This may seem very promising, although, the existence of more strategies, hinders this progression.

4.1.3 Collusion

An alternative to competitive strategies is to use co-operation with similar banks, also known as collusion. This is an agreement for a subsection of the market, called cliques, to work together to fix their prices and split borrowers evenly amongst them in order to ensure future business.

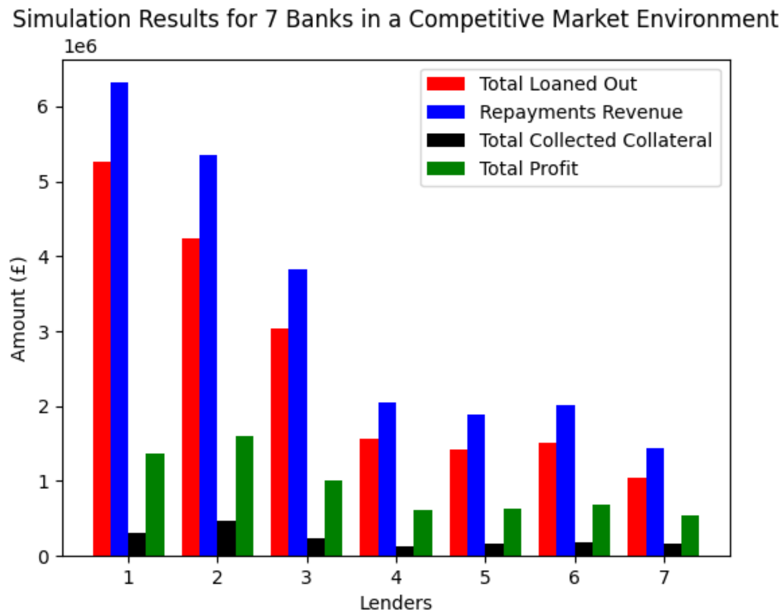

Figure 6: Banks 4-7 employing collusion strategy

In figure 6, banks 4 - 7 employ this collusion strategy. Profits for banks 4 - 6 are nearly identical with bank 7 slightly lower, a further uniformity of the market distribution within this band. Lender 7's decrease can be attributed to having the lowest lending capacity, therefore is limited on the number of possible offers they can accept at any one time. This strategy can therefore be used to hinder large banks from creating a monopoly. However, used incorrectly, a set of large banks could cause exactly this by gaining overarching control of the entire market. It is crucial therefore that government regulations limit these illegal activities, as was seen with the tech giant Apple in 2015 [21].

4.1.4 Reducing Verification Time Delay

Conducting a thorough background check of a potential borrower provides greater certainty of loan completion. Consequently, a very short background check increases uncertainty, leading to an inflated amount of risk being taken on by a lender. With smaller banks the most likely to take on such risks, a simple simulation was run, where smaller banks began accepting applicants wishing to expedite their acceptance.

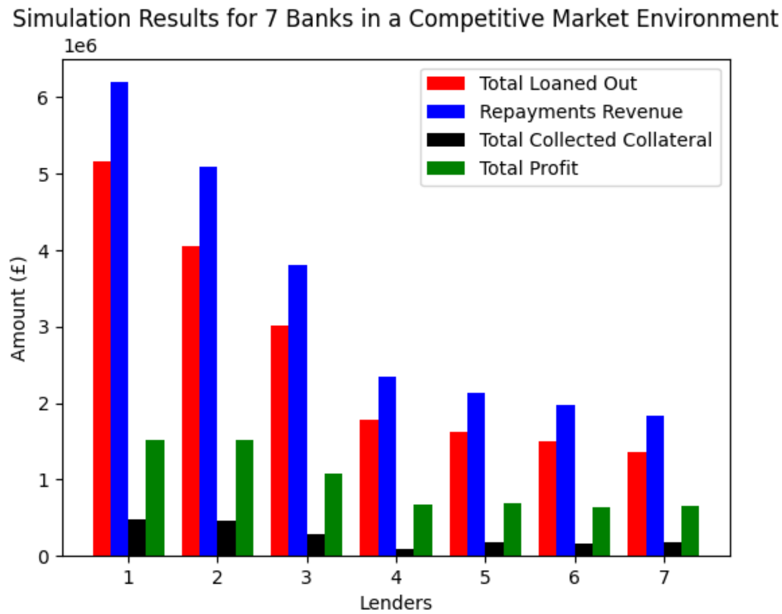

Figure 7: Banks 4 - 7 reducing verification time delays

Evidently, there seems to be minimal evidence of a correlation between impatient customers and the default rate. This seemingly insignificant change, proves to be quite critical, as there is a recorded growth across the board in revenue and profits, for the small bank tier. Smaller banks in this case are benefiting from those with higher credit tiers, but value the time-frame of receiving the loan over the monthly cost of the loan. It is therefore of interest to all banks, not only small, to reduce verification time delays where possible, as this will greatly benefit both parties.

5 Evaluation

In training our model, a reliable data set was difficult to acquire, as predicting the credit score of borrowers requires accurate real-world data that is not contained in the public domain. As the project is centred around credit risk, possessing accurate default rates is crucial. Although the chosen data set was considered the most appropriate, the credit tier system could be improved by utilising a more representative data set.

In practice, access to vast amounts of borrowers' data, on both current and previous loan transactions, is one of the reasons that large banks have a competitive advantage over smaller banks. Banks with more data to train their default prediction models allow a more accurate value for the actual default probabilities of borrowers. This report, therefore, is aimed toward banks to provide insight into these market dynamics. More specifically, focus lies heavily on small and medium band banks, as larger banks generally do not need to employ these additional strategies, beyond their no-risk customer policy.

An area of interest to take this report further would be to consider the borrowers previous loans and other current financial commitments, such as mortgages, as these can be vital indications of the risk of default. To advance the simulation, the length of time could be increased to allow recycling of borrowers into the loan process. This would lead to a replication of the financial history that banks have in the real world. Once an American citizen opens an account with an institution providing access to credit instruments, a record of their ability to pay back on any debt is formed to increase the accuracy of the borrowers credit score.

Another aspect to consider would be down payments and pre-payments. In practice, a borrower will often be required to contribute an initial portion of the retail value, most commonly a sum of 20% of the retail price from their own funds towards the car, often referred to as a deposit or down-payment. In this model this process was disregarded by assuming the borrower will take out a full loan on a lower priced asset. Ignoring down-payment, and assuming a smaller asset value, leads to a reduction in collateral collected, within the model. The effect is most predominant at the beginning of the loan. If a loan defaults immediately, a lender ideally loses nothing. If the down-payment is set at 20% and the vehicle is sold at 80% of its value as mentioned in section 3.2, the bank collects all of its money back [16].

6 Conclusion

In this report, a model was created to simulate a competitive car financing market that can handle a wide array of unique borrowers and lenders. Four separate strategies were investigated to provide insight into possible optimisations both parties can make during this process. Analysing the final results reveals that large banks are unaffected by the performance of smaller banks, due to their greater outreach to consumers. Contrary to this, small banks can benefit greatly from market strategies, particularly undercutting and reducing verification delays. A common target that all banks should adopt, however, is maximisation of profit rather than minimisation of risk. Although uncertainty was not directly considered, risk appetite is implied from the value of collateral collected, hence providing rough estimations of default rates. What is seen is that the large majority of this risk is alleviated through the credit scoring, verification and collateral collection methods. More accurate data sets would further confirm this, due to an improved classification of borrowers. Lenders will therefore benefit the most from catering to the largest amount of borrowers possible. Borrowers may capitalise on this by improving their credit scores as much as possible, allowing them to take advantage of lower interest rates. An exact algorithm for borrower decisions is difficult to prescribe as it is heavily reliant on their individual financial situation.

References

[1] Auto loan debt reaches a record-high $1.43 trillion. https://www.experian.com/blogs/ask-experian/research/auto-loan-debt-study/. Accessed: 2022-11-10.

[2] Employee tenure in 2022 - bureau of labor statistics. https://www.bls.gov/news.release/pdf/tenure.pdf. Accessed: 2022-11-7.

[3] Largest banks in the united states as of december 2021, by market capitalization. https://www.statista.com/statistics/431751/leading-banks-usa-by-market-cap/. Accessed: 2022-11-7.

[4] Meeting with sopra banking representative. "Private Communication".

[5] Share of used and new vehicles with financing in the united states from 2017 to 2022. https://www.statista.com/statistics/453000/share-of-new-vehicles-with-financing-usa/.

[6] Car depreciation calculator - good calculators. https://goodcalculators.com/car-depreciation-calculator/, 2019. Accessed: 2022-11-7.

[7] Orazio P Attanasio, Pinelopi Koujianou Goldberg, and Ekaterini Kyriazidou. Credit constraints in the market for consumer durables: Evidence from micro data on car loans. *International Economic Review*, 49(2):401–436, 2008.

[8] US Census Bureau. Age and sex composition in the united states: 2005. https://www.census.gov/data/tables/2005/demo//age-and-sex/2005-age-sex-composition.html, Dec 2021.

[9] US Census Bureau. United states household income report. https://www.census.gov/library/publications/2022/demo/p60-276.html, Dec 2021. Accessed: 2022-11-7.

[10] Ting-Hsuan Chen. Do you know your customer? bank risk assessment based on machine learning. *Applied Soft Computing*, 86:105779, January 2020.

[11] Joseph De Vita. Niche separation and the broken-stick model. *The American Naturalist*, 114(2):171–178, 1979.

[12] Giovanni Dell'Ariccia and Robert Marquez. Lending booms and lending standards. *The Journal of Finance*, 61(5):2511–2546, 2006.

[13] Statista Research Department. United states homeownership rate 2021. ref:https://www.statista.com/statistics/184902/homeownership-rate-in-the-us-since-2003/, Oct 2022. Accessed: 2022-11-7.

[14] Erin Duffin. Distribution of united states households by size 1970-2021. https://www.statista.com/statistics/242189/disitribution-of-households-in-the-us-by-

household-size, Sep 2022. Accessed: 2022-11-7.

[15] Erin Duffin. Marital status of the united states population 2021. https://www.statista.com/statistics/242030/marital-status-of-the-us-population-by-sex/, Sep 2022. Accessed: 2022-11-7.

[16] Jennifer Nelson. How much down payment is needed for a car? https://www.investopedia.com/how-much-down-payment-is-needed-for-a-car-5184453, Mar 2022.

[17] Bureau of Labor Statistics. Average annual household expenditure in the united states. https://www.statista.com/statistics/237203/average-expenditures-of-united-states-households/, Sep 2022. Accessed: 2022-11-7.

[18] Paolocons. Loan default classifier with keras (baseline). https://www.kaggle.com/code/paolocons/loan-default-classifier-with-keras-baseline/data, Nov 2021. Accessed: 2022-11-7.

[19] Anthony Sanders. The subprime crisis and its role in the financial crisis. *Journal of Housing Economics*, 17(4):254–261, December 2008.

[20] Carlos Serrano-Cinca, Begoña Gutiérrez-Nieto, and Luz López-Palacios. Determinants of default in p2p lending. *PloS one*, 10(10):e0139427, 2015.

[21] Julie Young. Collusion. https://www.investopedia.com/terms/c/collusion.asp, Sep 2021.

[22] Sara Zahi and Boujemâa Achchab. Modeling car loan prepayment using supervised machine learning. *Procedia Computer Science*, 170:1128–1133, 2020.

A Further materials

Bank size	Bank number	Market Share $M\%$	Interest Rate $r\%$	Credit Bands Available
Large	1	30	$9 \leq r \leq 12$	≤ 4
Medium	2,3	$5 \leq M \leq 30$	$11 \leq r \leq 16$	≤ 6
Small	4,5,6,7	≤ 5	$13 \leq r \leq 20$	≤ 10

Table 2: Table showing values for the different bank bands - values are based on information provided by Sopra Banking representative [4].

This section outlines some background information that is not directly relevant to our model but provides additional context.

An asset is defined as a resource with economic value that an individual, corporation, or country owns or controls with the expectation that it will provide a future benefit. In the context of borrowing against an asset, the asset is the owned resource that is used as collateral for the loan. In the example of car loans, the lender's asset is the loan itself whereas the borrower considers the car to be their asset. New and Used cars are one of the fastest depreciating commodities, with new cars expected to depreciate between 20% and 35% within the first year of ownership and used cars at 15% to 20% [6]. Fortunately, for lenders, it is difficult for borrowers to make fully financially informed decisions regarding value depreciation at the time of purchase. This is because it is only truly calculable retrospectively and as described by Stefan Lessmann, the sellers of used cars possess informational advantages over market research agencies and individuals, which enable them to forecast resale prices more accurately. [15]