Improving the Price Optimization Model for Nomis Solutions and E-Car's Data.

In Partial Fulfillment of the Requirements for

Machine Learning 1

Submitted By:

Apeles,

Dela Cruz,

Hwang,

Nigam,

Palomaria,

Vargas

Date: 4 August 2020

Abstract

Nomis Solutions has closed an initial deal with e-Car, an online auto loan company, to provide price optimization solutions and suggest new rates that will help maximize e-Car's profits. A dataset containing over 200,000 records which includes the outcome of loan applications from 2002 to 2004 was provided to Nomis Solutions by e-Car and was used in this study. This study implements a price optimization methodology which considers demand, price sensitivity, and risk. Nomis Solutions can present to e-Car this methodology that achieves over a 90% increase in e-Car's revenue and potentially close a larger deal with e-Car.

Introduction

Nomis Solutions has been tasked by e-Car's CEO, an online auto loan company, to help e-Car optimize its current loan rates and maximize its profits. This current offer of e-Car to Nomis Solutions was just an initial deal to prepare a business case that will be used to convince e-Car's board of directors and potentially, close a much larger deal. A dataset containing loan application outcomes has been provided to Nomis Solutions by e-Car.

The aim of this study is to provide a price optimization methodology that Nomis Solutions can present to e-Car's board of directors.

Data Description

The dataset contains 208,085 records of loan application outcomes from July 01, 2002 up to November 16, 2004. The dataset contains the following detailed information:

- Tier -- e-Car segmentation based on FICO scores
- FICO -- FICO score
- Approved Date -- Approval date of loan
- Term -- Number of months for loan payment
- Amount -- Loan amount in USD
- Previous Rate -- Previous annual percentage rate approved for the applicant
- Car Type -- Type of financing applied (New Car, Used Car, Refinancing)
- Competition Rate -- Published Rate of Competitor
- Rate -- Annual percentage rate offered by e-Car
- Outcome -- Indicator if customer signed the loan
- Cost of Funds -- Cost of Funding Loan for e-Car
- Partner -- e-Car categorization based on its funding partners

Assumptions

1. One price rating scheme will be implemented using the 3 years of data.
2. No negotiation can be done after rate has been provided by e-Car.
3. All "lost" customers deferred to competition rates.
4. Customers make decisions based on rates only.
5. When costlier rates are preferred by the customer, it is assumed that the competitor was not visible to the customer.
6. The existing segmentation of e-car will be used in price optimization.

Workflow

> Exploratory Data Analysis

The dataset will be explored first to check for trends that may affect succeeding analysis.

> Searching for Potential Pricing Errors

Since this study is about price optimization, the existing pricing technique will be scrutinized for potential pricing errorss.

> Segmentation

To make the analysis more manageable, segmentation will be done before price optimization.

> Price Optimization

Maximizing profit is the main goal of this study and this will be done through price optimization.

Preliminaries

Importing Libraries

Viewing the data

Out[3]:

	Tier	FICO	Approve Date	Term	Amount	Previous Rate	Car Type	Competition rate	Outcome	Rate	Cost of Funds	Partner Bin
33641	1	771	2004-04-24	60	18,650.000		N	4.190	1	3.990	1.100	1
60710	1	733	2002-10-29	60	36,000.000		N	4.990	0	4.890	1.790	1
91171	2	714	2003-04-21	48	26,389.000		N	4.490	0	4.350	1.320	3
41387	1	734	2004-08-31	36	18,000.000	7.450	R	4.990	1	4.990	1.670	1
33397	3	716	2004-04-21	66	10,640.100		U	5.290	1	6.690	1.100	1

The DataFrame above displays the first five rows of the dataset.

Out[4]:
```
Tier                          int64
FICO                          int64
Approve Date         datetime64[ns]
Term                          int64
Amount                      float64
Previous Rate                object
Car  Type                    object
Competition rate            float64
Outcome                       int64
Rate                        float64
Cost of Funds               float64
Partner Bin                   int64
dtype: object
```

The dataset contains integer (int64), float (float64), data (datetime64[ns]), and string (object) values.

Out[5]:

	Tier	FICO	Term	Amount	Competition rate	Outcome	Rate	Cost of Funds	Partner Bin
count	208,085.000	208,085.000	208,085.000	208,085.000	208,085.000	208,085.000	208,085.000	208,085.000	208,085.000
mean	1.927	726.731	56.809	26,009.516	4.807	0.220	5.623	1.329	2.030
std	1.051	44.784	11.204	11,108.799	0.586	0.414	1.547	0.278	0.911
min	1.000	587.000	36.000	5.000	2.990	0.000	2.450	1.020	1.000
25%	1.000	692.000	48.000	17,800.000	4.390	0.000	4.490	1.110	1.000
50%	2.000	726.000	60.000	25,000.000	4.790	0.000	5.090	1.262	2.000
75%	3.000	762.000	60.000	33,000.000	5.190	0.000	6.390	1.419	3.000
max	4.000	854.000	72.000	100,000.000	6.450	1.000	15.530	2.127	3.000

The DataFrame above displays the mean, standard deviation, minimum value, 25th percentile, 50th percentile, 75th percentile, and maximum values of each numerical variable.

```
# of Empty cells =  160867
```

Out[6]:
```
Tier                0
FICO                0
Approve Date        0
Term                0
Amount              0
Previous Rate       8
Car  Type           0
Competition rate    0
Outcome             0
Rate                0
Cost of Funds       0
Partner Bin         0
dtype: int64
```

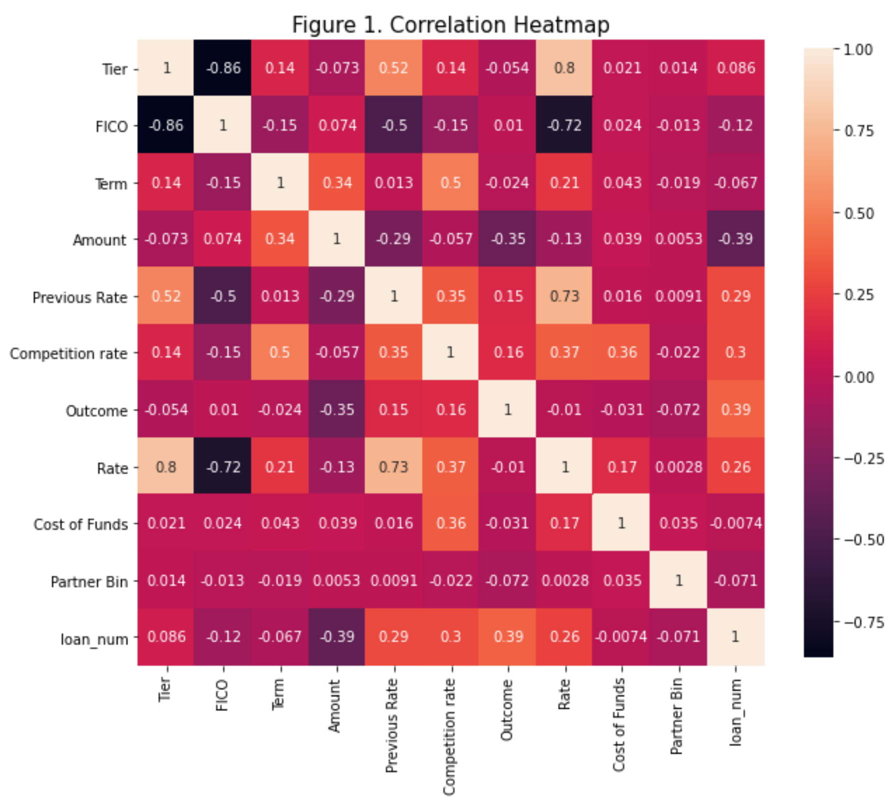
The dataset contains 160,867 empty cells (blank strings) and 8 missing values in the Previous Rate column.

Out[7]:

	Tier	FICO	Approve Date	Term	Amount	Previous Rate	Car Type	Competition rate	Outcome	Rate	Cost of Funds	Partner Bin
185252	1	734	2004-07-08	60	45,000.000	4.490	N	4.490	0	4.490	1.363	3
44247	1	766	2004-10-18	48	13,629.020	18.420	R	5.490	1	5.490	1.910	2
195316	1	793	2004-09-07	36	17,500.000	4.390	U	4.390	0	4.390	1.738	3
49490	4	667	2002-07-27	60	30,000.000	7.990	N	5.490	0	7.990	1.810	3
132610	3	818	2003-09-22	72	34,325.000	4.950	N	4.990	0	4.950	1.120	3

The empty cells (blank strings) and missing values in the Previous Rate column were filled using the current rates.

Exploratory Data Analysis

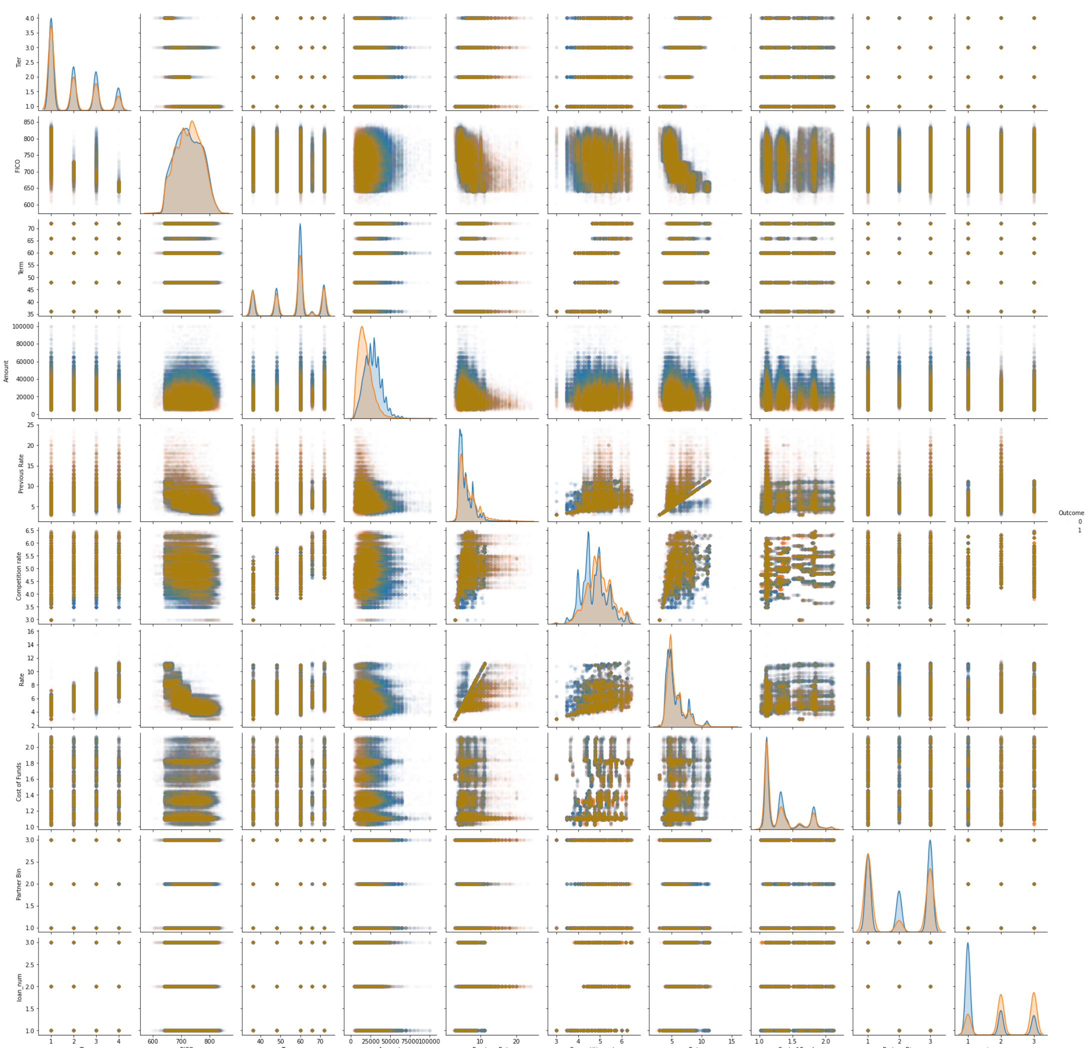


As expected, there are strong negative correlations with the Tier-FICO and FICO-Rate pairs since the tier segmentation is based on FICO scores and rates are based on risk which is indicated by the FICO scores.

Since tier segmentation and the rates are based on FICO scores, Tier and Rate have a strong positive correlation.

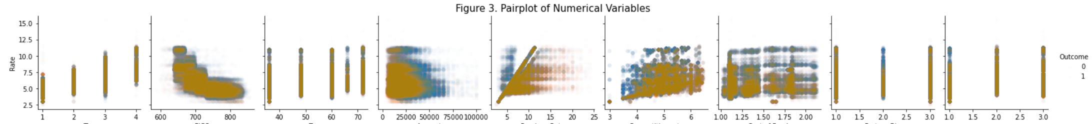
Also since empty cells of Previous Rate were filled with Rate values, they also have a strong positive correlation.

Figure 2. Pairplot of Numerical Variables



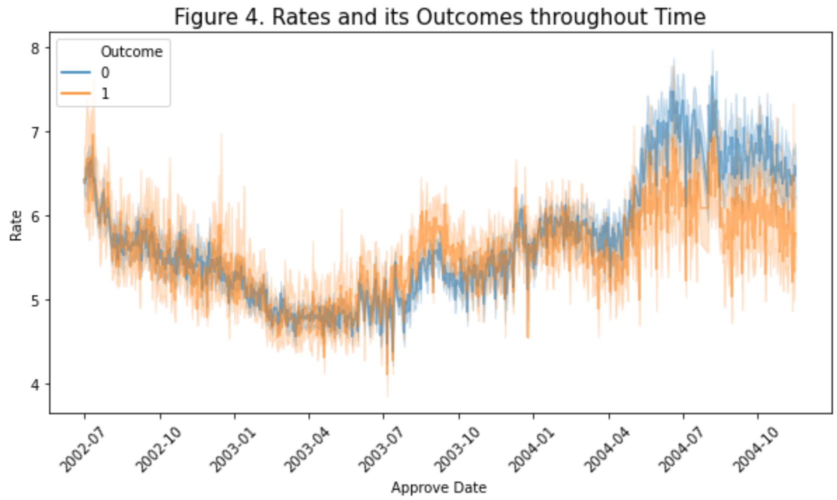
The plot above provides the pairwise relationship for each numeric feature pair.

Figure 3. Pairplot of Numerical Variables



Since Rate is the main concern of the study, its pairwise relationship plot with the other numeric variables was isolated.

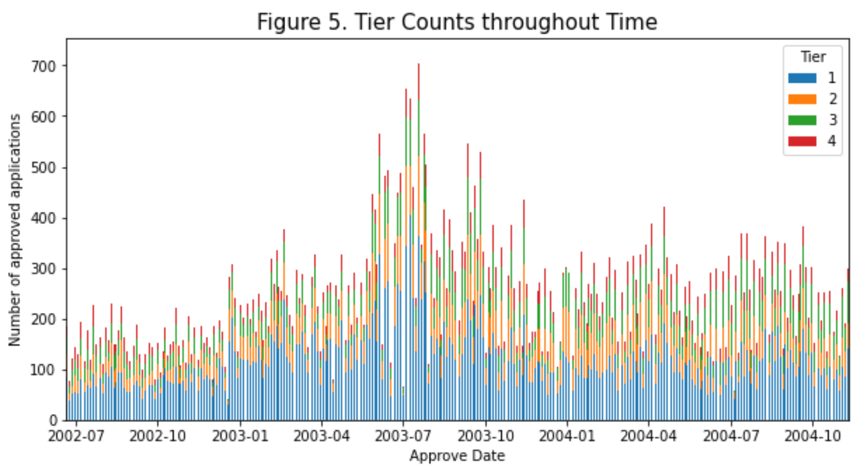
Based on the Rate and Amount pairplot, higher loan amounts have a tendency of not being accepted by customers.



From around July 2002 to April 2003, the rates with outcome 1 (signed by applicant) and 0 (not signed by applicant) were mixed together.

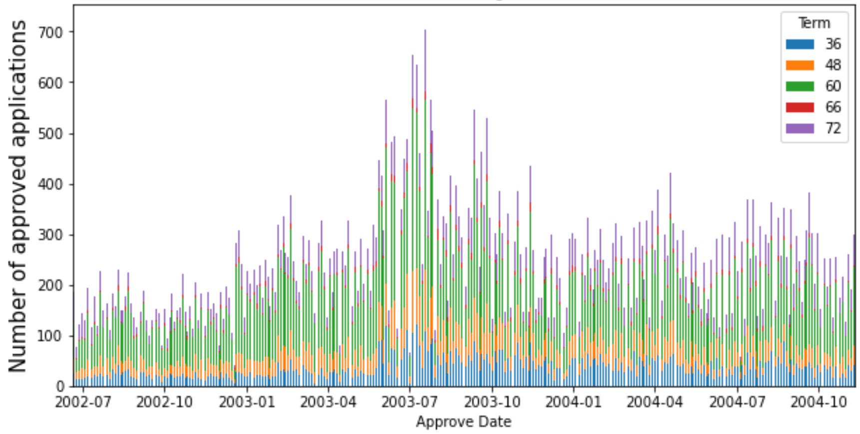
Surprisingly, from around May 2003 to January 2004, the higher rates had an outcome of 1 while the lower rates had an outcome of 0.

As expected, higher rates got an outcome of 0 while lower rates got an outcome of 1, from around February 2004 onwards.



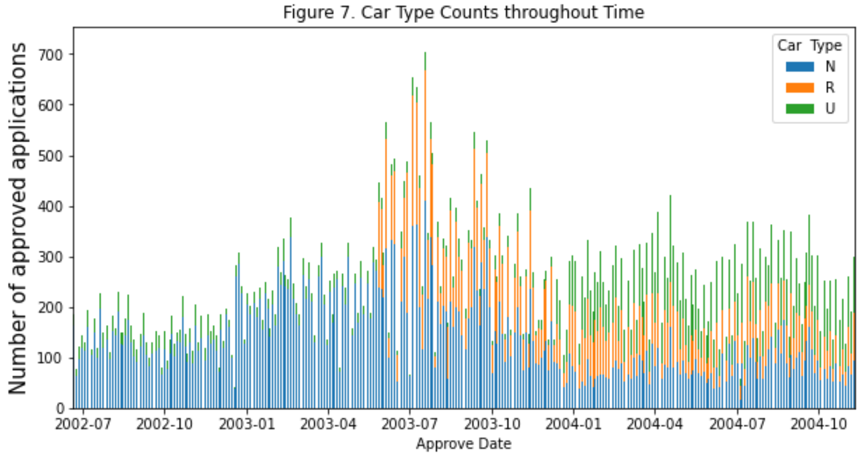
Throughout 2002 to 2004, the loan applications were mostly composed of Tier 1 loans. Since FICO and Tier are negatively correlated, this means that e-Car is approving more less risky loan applications.

Term Counts throughout Time



In the span of three years, the loan applications were mostly 60-month terms.

Figure 7. Car Type Counts throughout Time

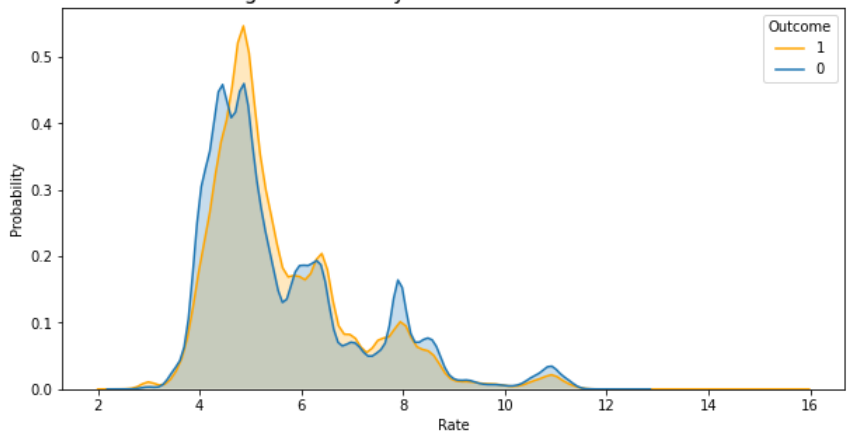


Throughout 2002 to 2004, the loan applications were mostly for new car financing. Refinancing loans surfaced at around 2003. This could suggest that e-Car has been more approving in a wider variety of loan types during that time. This could also explain the peak in number of approved applications on 2003.

Pricing Errors

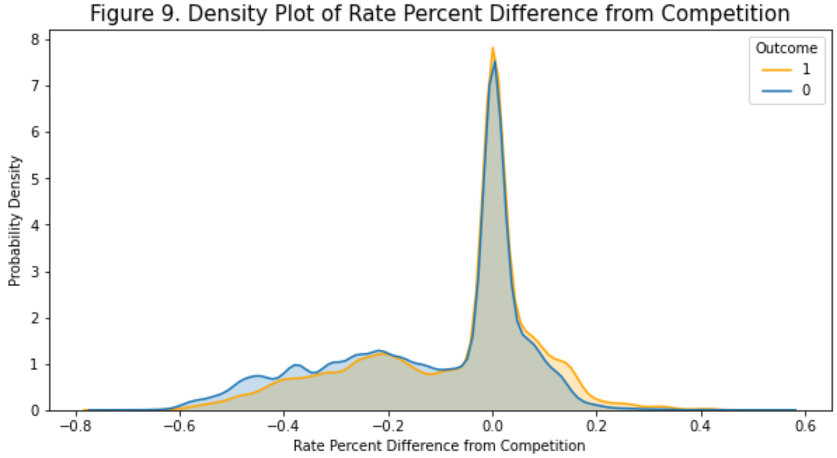
- How will you show that the current pricing technique contains "pricing errors"?
- How do we know that there are "pricing errors"? What are we trying to maximize in this project?

Figure 8. Density Plot of Outcomes 1 and 0



Outcomes 1 and 0 generally have the same rates.

Figure 9. Density Plot of Rate Percent Difference from Competition




```
When e-Cars is not cheaper than competition rate, there is a 75.36 probability that e-Car will lose
the customer.
When e-Car is more expensive than competition rate, there is a 19.11 probability that e-Car will win
the customer.
```

The probability computations above suggest that customers may prefer cheaper or costlier rates because of other factors. It will be assumed that these other factors cannot be controlled. Only the rates can be controlled.

Since the current pricing technique already involves segmentation, the problem lies in the rates being offered to each segment which is where there price optimization will come in.

Another potential problem in the pricing technique is the disregard for the loan-to-value ratio. Loan to value is one of the key risk factors that lenders consider when qualifying borrowers for a loan. The risk of default is always at the forefront of lending decisions, and the likelihood of a lender absorbing a loss increases as the amount of equity decreases.

It is also desired to maximize the net interest profit. That is the interest gained from the borrower after subtracting the cost of funds to be paid to the partner.

Segmentation

- How can we make the analysis more manageable? Should we segment the customers?

Pricing segment s_ijk is defined as a unique combination of product characteristics, customer characteristics and channels [1]. It is imperative to classify the data into different pricing segments due to following reasons:

- Pricing errors are identified for a specific pricing segment.
- The optimum rates vary across different segments.
- The bid-response function of a customer varies across different pricing segments.

In this study, a segment s_ijk is given by a unique combination of loan type (i), tier (j) and loan term (k). Loan type is a classification whether the loan is being applied for a new car, used or refinanced one. In this study, loan type is given by nominal classification: $k \in [1, 2, 3] \rightarrow \text{loan\_type} \in [\text{New}, \text{Refinanced}, \text{Used}]$.

A tier is an ordinal classification based on FICO score that was assigned by the company. A tier represents the risk band of the customer. Tiers are numbered based on increasing risk: 1 being the least risky, and 4 as the riskiest. In this study, a tier is given by $j \in [1, 4]$.

There are five (5) types of loan terms: $i \in [1, 5] \rightarrow \text{loan\_term} \in [36, 48, 60, 66, 72]$ months.

To illustrate, the segments are represented below as a matrix of segments s_{ijk}. Thus, the total number of segments is $n_{ijk} = n_i n_j n_k = 5 \times 4 \times 3 = 60$.

Type (i)	Tier (j)	36 mo	48 mo	60 mo	66 mo	70 mo
1-New	FICO=1	s_{111}	s_{112}	s_{113}	s_{114}	s_{115}
1-New	FICO=2	s_{121}	s_{122}	s_{123}	s_{124}	s_{125}
1-New	FICO=3	s_{131}	s_{132}	s_{133}	s_{134}	s_{135}
1-New	FICO=4	s_{141}	s_{142}	s_{143}	s_{144}	s_{145}
2-Refinanced	FICO=1	s_{211}	s_{212}	s_{213}	s_{214}	s_{215}
2-Refinanced	FICO=2	s_{221}	s_{222}	s_{223}	s_{224}	s_{225}
2-Refinanced	FICO=3	s_{231}	s_{232}	s_{233}	s_{234}	s_{235}
2-Refinanced	FICO=4	s_{241}	s_{242}	s_{243}	s_{244}	s_{245}
3-Used	FICO=1	s_{311}	s_{312}	s_{313}	s_{314}	s_{315}
3-Used	FICO=2	s_{321}	s_{322}	s_{323}	s_{324}	s_{325}
3-Used	FICO=3	s_{331}	s_{332}	s_{333}	s_{334}	s_{335}
3-Used	FICO=4	s_{341}	s_{342}	s_{343}	s_{344}	s_{345}

The mean and standard deviation of interest rate, and the total number of loans offered for each segment is shown below. The data was aggregated based on the outcome: 1 corresponds to loan accepted by the customer, 0 as lost customer.

Figure 10.0 shows the percentage of loans accepted in each segment. Based on the table, segments that belong under loan*type **1-New** ($s*{1jk}
$)have significantly lower percentage of accepted loans compared to segments under loan_t ype **2-Refinanced**($ s*{2jk}*
$)and **3-Used**($ s{3jk}$).

Segment 141 has the lowest percentage of loans accepted (<1%) for segments under loan_type 1-New (Figure N.n). Segment 231 has the lowest percentage of loans accepted (25%) for segments under loan_type 2-Refinanced. While, segment 341 has the lowest perentage accepted (26%) for segments under loan_type 3-Used. In general, the percentage of loans accepted increases from segment 1jk to segment 3jk (Figure 10.0).

Figure 11.0 shows the percentage of loans accepted in segments for each Tier ($s_{i1k}, s_{i2k}, s_{i3k}, s_{i4k}$). Based on the mean percentage, the average loan acceptance rate decreases from segment s*{i1k} to segment* s{i4k}. For each Tier segment, those segments that belong under loan_type **1-new** (i.e. s_1jk) have significantly lower percentage of loans accepted than those segments under loan_type **2-Refinanced** and **3-Used**.

Figure 12.0 shows the percentage of loans accepted in segments for each loan*term ($s*{ij1}, s*{ij2}, s*{ij3}, s*{ij4}, s*{ij5}
$). Based on the mean percentage, the average loan acceptance rate increases from segment$ s*{ij1} to segment* s{ij5}
$. For each loan_t erm segment, those segments that belong under loan_t ype **1-new** (i.e.$ s_1jk$) have significantly lower

percentage of loans accepted than those segments under loan_type **2-Refinanced** and **3-Used**. This is consistent with the previous observations.

Out[19]:

			mean					std					count				
		Term	36	48	60	66	72	36	48	60	66	72	36	48	60	66	72
Car Type	Tier	Outcome															
N	**1**	**0**	4.005	4.387	4.368	5.143	5.149	0.434	0.343	0.333	0.377	0.398	7479	8308	31873	1083	7333
		1	3.796	4.363	4.360	5.086	5.109	0.505	0.348	0.356	0.288	0.382	1220	538	2793	52	1079
	2	**0**	4.974	4.984	5.126	5.570	5.706	0.661	0.669	0.764	0.693	0.732	2157	2837	12406	565	4473
		1	4.848	4.954	4.956	5.415	5.553	0.808	0.546	0.617	0.531	0.620	71	111	772	38	586
	3	**0**	5.727	5.690	5.969	6.451	6.566	0.961	0.850	1.133	0.875	0.991	1568	2270	10259	668	4748
		1	5.842	5.568	5.724	6.404	6.324	1.176	0.683	0.940	0.914	0.768	33	55	575	52	654
	4	**0**	7.905	7.889	8.164	8.387	8.557	0.860	0.750	1.132	1.005	1.134	873	1319	5756	488	3431
		1	7.457	7.796	7.946	8.468	8.370	0.871	1.289	1.222	1.078	0.985	6	19	223	13	275
R	**1**	**0**	4.618	4.934	4.928	5.480	5.386	0.272	0.314	0.324	0.502	0.595	3812	2841	4205	157	1259
		1	4.642	5.008	4.989	5.476	5.366	0.322	0.435	0.381	0.610	0.644	3107	2076	3091	153	1147
	2	**0**	6.152	6.223	6.183	6.319	6.440	0.677	0.740	0.770	0.790	0.949	2164	1752	2776	113	933
		1	6.197	6.236	6.155	6.375	6.302	0.692	0.720	0.773	0.895	0.921	846	794	1396	131	732
	3	**0**	6.988	7.051	7.043	7.233	7.248	1.198	1.225	1.231	1.237	1.418	1464	1282	2646	111	891
		1	7.209	7.154	6.952	7.085	6.947	1.227	1.139	1.188	1.202	1.259	469	574	1174	90	511
	4	**0**	8.124	8.063	8.157	8.249	8.538	1.329	1.423	1.390	1.208	1.584	664	644	1104	97	415
		1	8.273	8.195	8.264	7.904	8.153	1.187	1.133	1.251	1.370	1.481	229	323	674	60	303
U	**1**	**0**	4.342	4.845	4.822	5.610	5.622	0.331	0.291	0.279	0.292	0.322	1365	1349	3801	169	967
		1	4.265	4.665	4.653	5.476	5.433	0.328	0.358	0.375	0.368	0.371	1509	1430	4398	186	1481
	2	**0**	6.133	6.191	6.258	6.781	6.845	0.628	0.569	0.551	0.653	0.609	635	852	2995	123	1083
		1	5.600	5.558	5.306	5.894	6.014	0.830	0.883	0.806	0.727	0.842	384	601	2074	107	965
	3	**0**	7.219	7.244	7.325	7.870	7.837	1.182	1.183	1.227	1.147	1.210	615	806	2931	173	1379
		1	6.393	6.264	6.153	6.696	6.882	1.152	1.200	1.178	0.831	1.039	232	471	1905	127	837
	4	**0**	9.516	9.580	9.559	9.766	10.003	1.231	1.230	1.236	1.251	1.274	389	539	1689	181	1033
		1	8.974	8.681	8.582	8.552	8.766	1.392	1.283	1.298	1.153	1.244	139	268	986	79	563

Out[20]:

	Tier	FICO	Approve Date	Term	Amount	Previous Rate	Car Type	Competition rate	Outcome	Rate	Cost of Funds	Partner Bin	segment
197424	1	799	2004-09-20	60	35,000.000	4.790	U	4.790	0	4.790	1.828	3	313
159323	3	682	2004-02-16	60	26,382.000	6.190	U	4.650	0	6.190	1.094	1	333
9240	2	705	2003-05-05	72	41,295.200	4.890	U	5.590	1	4.890	1.310	3	325
142893	1	753	2003-11-09	60	16,600.000	6.900	R	4.950	0	4.950	1.120	3	213
58679	1	736	2002-10-12	60	25,000.000	4.890	N	4.990	0	4.890	1.802	3	113

Out[21]: 60

Figure 10. Loan Acceptance Rate at Different Loan Types

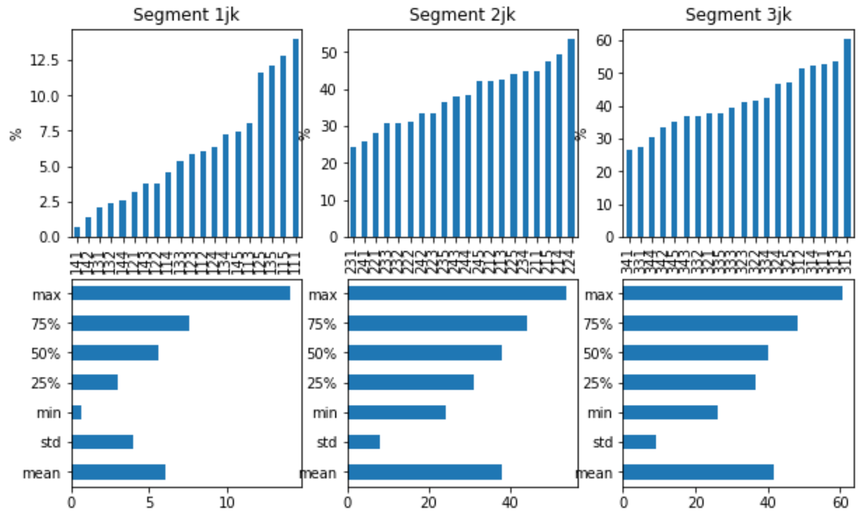


Figure 11. Loan Acceptance Rate at Different Tiers

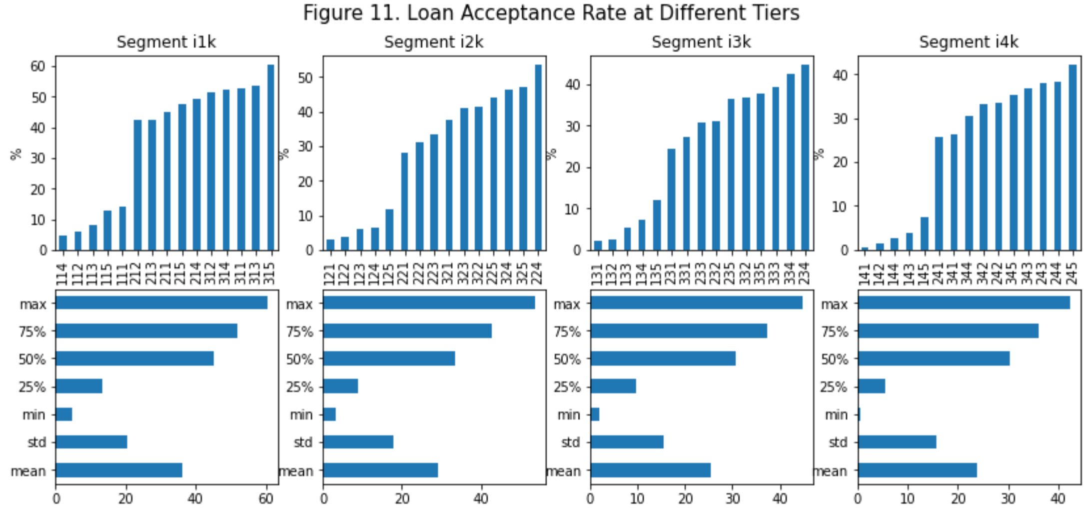
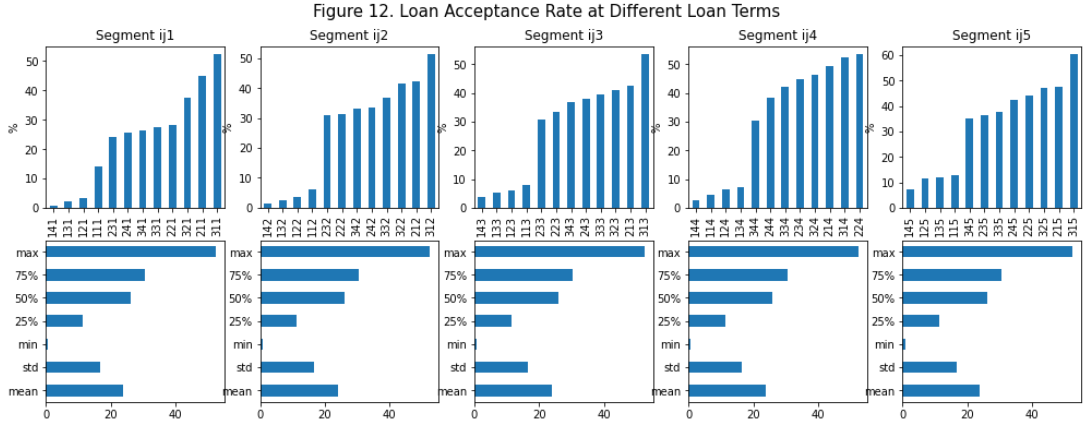


Figure 12. Loan Acceptance Rate at Different Loan Terms



Price Optimization by Segment

- How should we price the loans?
- Do you think you can recommend the right price to quote?
- Has there been a mis-pricing of APR quotes?
- Can we build a systematic approach that can scale with the number of segments given certain customer characteristics?

The primary objective of price optimization is to determine the optimum price for each segment using the data that were previously categorized. The objective function is to maximize the profit given the factors for each segment such as the current rate, previous rate, competitive rate, demand, and the risk involved.

In this study, an optimum price is defined as the interest rate that will maximize the overall profit. Overpricing a prospective loan decreases the probability that a customer will accept the loan, but it increases the profitability if the customer does accept. In contrast, underpricing a prospective loan increases the probability that a customer will accept the loan, but it may compromise profitability. Thus, the optimum loan prices that maximize the over-all profit can be found for each segment.

Objective Function

The objective function [1] is given by

$$\max_r TR(r) = \sum_{i}^{N=60} D_i F_i(r_i) PVNII(P_i, r_i, n_i)$$

where,

D_i = total demand per segment,
F_i = probability of accepting the loan per segment,
$PVNII(P_i, r_i, n_i)$ = Present value of net interest income

subject to the constraint:

$$r_i \geq 0$$

The present value of the net interest income is given by

$$\begin{aligned} PVNII(P, r, n) &= Pn(r - r_c) - (1 - s_n) \times Pr \sum_{i=1}^{n} \frac{1 - s_i}{1 - s_n} \\ &= Pn(r - r_c) - PRD \times LGD \end{aligned}$$

where,

r_c = cost of capital / funds
s_i = probability that a borrow will make payment i
PRD = probability of default
LGD = loss given default

Non-Linear Optimization

Contrained minimization method COBYLA (Constrained Optimization By Linear Approximation) was used to solve the objective function. The algorithm fits for optimization with non-linear inequality and equality constraints [2]. It minimizes the objective function $F(X)$ subject to M inequality constraints of the form $g(X) \geq 0$ on X, where X is a vector of N dimensions.

In this study, r is a vector of interest rates of $N = 60$ dimensions, where each dimension represents a unique pricing segment. The current interest rates were used as initial values. The only constraint imposed was $r \geq 0$.

Modeling the Bid-Response Function

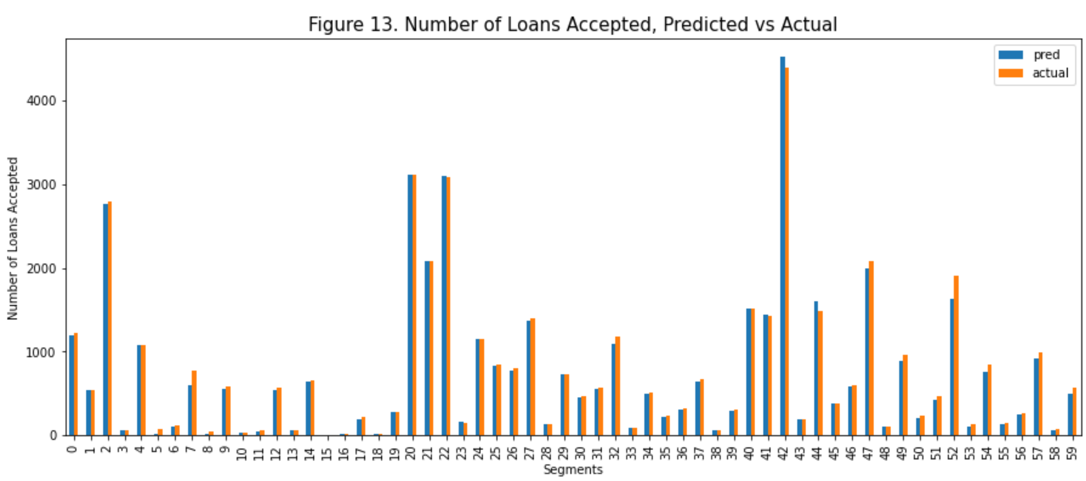
A bid response function gives the probability of winning for every possible price response. Low prices gives higher probability of winning, while high prices decreases the chance. A customer bid response can modeled using a logit function or power function [3]. The logit model applies even without knowledge on the competitor's pricing. In this study, the bid-response function, $F(r_i)$ was modeled using a bid-repsonse function of the form:

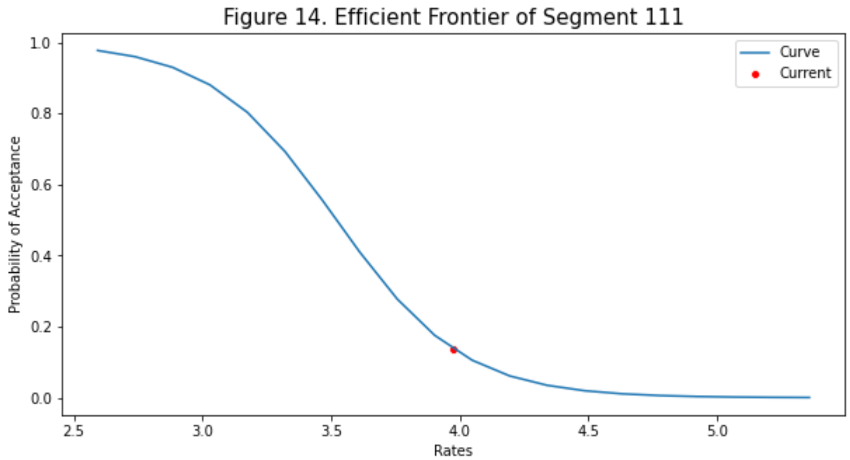
$$F(r_i) = \frac{1}{1 + e^{a + bX(r_i)}}$$

where,
$X(r_i)$ = is a measure of competitiveness of the offered APR r_i [1] and defined by,

$$X(r_i) = (\text{Previous}_{APR} - \text{Offered}_{APR}) + (\text{Competitor}_{APR} - \text{Offered}_{APR})$$

$F(r_i)$ was modeled by minimmizing SSE (Sum of Squared Residuals) for each segment.





Risk Free Loans

For a risk-free loan, $PRD = 0$. Thus,

$$PVNII(P, r, n) = Pn(r - r_c)$$

The objective function is then given by,

$$\max_r TR(r) = \sum_i^{N=60} D_i F_i(r_i) P_i n_i(r_i - r_{c_i})$$

The results are given below:

```
Optimization conducted successfully =  True
Total Revenue = 498974363191
Optimized Rates =
```

Out[31]:

Car Type	Tier	36	48	60	66	72
N	1	3.071	2.589	2.563	16.730	3.211
	2	3.885	2.653	3.016	4.406	3.493
	3	7.312	2.801	3.414	7.099	4.261
	4	8.237	4.935	4.048	7.816	7.716
R	1	6.297	4.818	5.323	4.723	5.487
	2	4.288	4.857	4.982	6.556	5.614
	3	5.355	6.428	5.080	5.733	5.684
	4	6.821	6.812	7.355	7.502	6.629
U	1	3.745	4.191	4.301	4.923	5.096
	2	4.121	4.923	5.084	5.752	5.723
	3	4.995	5.664	5.532	6.372	5.994
	4	5.183	7.241	7.201	7.438	7.760

Low-Risk Loans

All loans have a risk factor, that is the risk of default at a certain time period within the contract. This is modelled by the (Probability of Deafult)*(Loss Given Default).

Since there is no data available to estimate the associated risk for each segment, the above equation was modified and the risk probabilities were simulated using the products of equally-spaced scalars. Although the actual risk values may differ per segment, a logical risk was assigned with the highest probability assigned to the segment with the longest loan*term (i.e. $s*{ij5} $)andhighestFICOscore(i.e.Tier4segments$,s_{i4k}$). For the low-risk loan, the maximum risk assigned is 0.2, and the average risk across all segments is 0.06.

The objective function is then given by,

$$\max_r TR(r) = \sum_i^{N=60} D_i F_i(r_i) \left[P_i n_i (r_i - r_{c_i}) - P_i r_i {k_i}^2 \right]$$

where k_i is the computed array of risk probabilities, with the values given below:

Risks:

Out[32]:

Car Type	Tier	36	48	60	66	72
N	1	0.000	0.005	0.010	0.015	0.020
	2	0.001	0.021	0.040	0.060	0.080
	3	0.001	0.036	0.071	0.105	0.140
	4	0.002	0.052	0.101	0.150	0.200
R	1	0.000	0.005	0.010	0.015	0.020
	2	0.001	0.021	0.040	0.060	0.080
	3	0.001	0.036	0.071	0.105	0.140
	4	0.002	0.052	0.101	0.150	0.200
U	1	0.000	0.005	0.010	0.015	0.020
	2	0.001	0.021	0.040	0.060	0.080
	3	0.001	0.036	0.071	0.105	0.140
	4	0.002	0.052	0.101	0.150	0.200

```
Optimization conducted successfully =  True
Total Revenue = 492785397050
Optimized Rates =
```

Out[34]:

Car Type	Tier	36	48	60	66	72
N	1	2.940	2.823	2.743	14.570	3.336
	2	3.884	2.773	2.979	4.424	3.479
	3	7.038	3.349	3.177	7.212	4.095
	4	8.450	4.991	4.111	9.463	7.752
R	1	7.193	4.604	5.063	4.955	5.789
	2	5.429	4.550	5.057	5.861	5.619
	3	5.257	6.954	4.850	6.523	5.793
	4	7.189	6.968	7.141	8.193	7.371
U	1	3.780	4.238	4.297	4.884	5.064
	2	5.044	4.813	5.129	5.059	5.616
	3	4.767	5.342	5.537	5.553	5.948
	4	6.044	7.873	7.125	7.078	7.617

High-Risk Loans

The case when the firm is operating with a higher risk was considered. The risk probabilities were still simulated using the products of equally-spaced scalars, with the highest probability assigned to the segment with the longest loan_term and highest FICO score. For this case, the maximum risk assigned is **0.4**, and the average risk acrosss all segment is **0.15**, which is **twice** than the low-risk loan. Figure N.n shows the density plots of the computed risk probabilities for both low and high risk loans. The probabilities for the high-risk loan are more spread through a wider range of higher values.

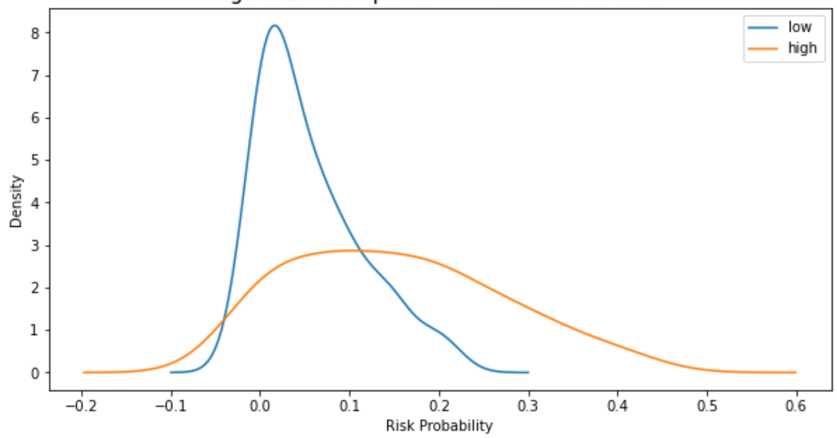
The objective function is the same as the low-risk loan, except for the newly-computed values of k_i. The results of optimization are given below.

```
Optimization conducted successfully =  True
Total Revenue = 497707282696
Optimized Rates =
```

Out[35]:

Car Type	Tier	36	48	60	66	72
N	1	3.115	2.623	2.609	16.654	3.232
	2	3.824	2.490	3.070	4.470	3.350
	3	9.921	4.325	3.374	7.618	3.932
	4	8.774	5.034	4.042	8.524	7.109
R	1	6.095	5.033	5.453	4.265	5.500
	2	5.905	3.985	5.341	6.717	5.721
	3	4.936	6.899	5.022	6.581	5.792
	4	5.714	6.443	6.288	5.868	7.051
U	1	3.971	4.041	4.252	4.866	5.185
	2	3.888	4.806	5.180	5.271	5.685
	3	4.884	5.303	5.496	5.803	6.040
	4	6.494	7.297	7.149	7.683	7.619

Figure 15. Comparison of Risk Probabilities

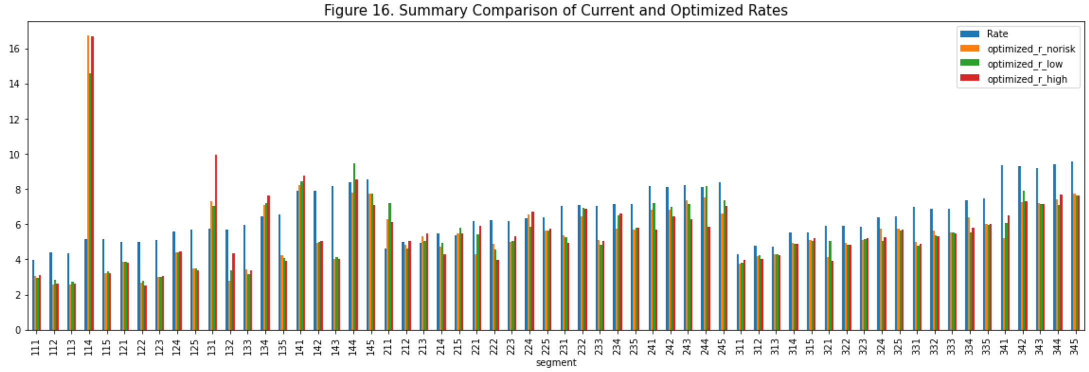


Comparison of Results

Figure 16 shows the comparison of the current and optimized rates for segments. In most segments, the optimum rate is slightly below the current rate. This means that in those segments, the firm can still increase profitability by offering lower APR to increase the probability of winning.

In segments where the optimum rate is higher than the current rate, offering a slightly higher APR will increase the profit, and the corresponding incremental reduction in the probability of winning is being offset by higher profitability.

Figure 16. Summary Comparison of Current and Optimized Rates



Comparison of Profit

The profit equation was used to calculate the present and potential revenue given the current and optimum APRs, respectively. The amounts computed are based on the simulated risk probabilities and fitted customer bid-response function which might differ in actuality for a specific segment.

The highest profit obtained using the present rate was **256,430,864,188.35**, under the risk-free condition. The calculated profit for the low and high-risk cases are just 0.01% lower than the profit calculated in the risk-free case.

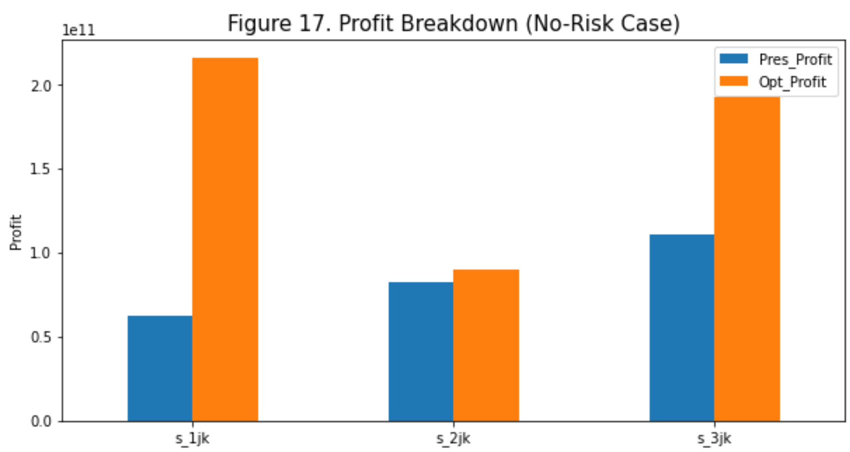
The potential profit calculated using the optimum rates are given in the second row of the table. For all three conditions, the potential profit obtained using the optimum rates is twice as large as the current profit. The highest potential profit is obtained in the risk-free condition, amounting to **496,840,870,620.74**. The potential profit obtained under low and high-risk conditions are just **~1%** lower than the amount computed under the risk-free condition.

Out[38]:

	Risk Free	Low Risk	High Risk
Present Rates	256,430,864,687.577	256,403,054,795.960	256,241,318,403.145
Optimized Rates	498,974,363,191.395	492,785,397,050.117	497,707,282,696.248

Breakdown of Profit

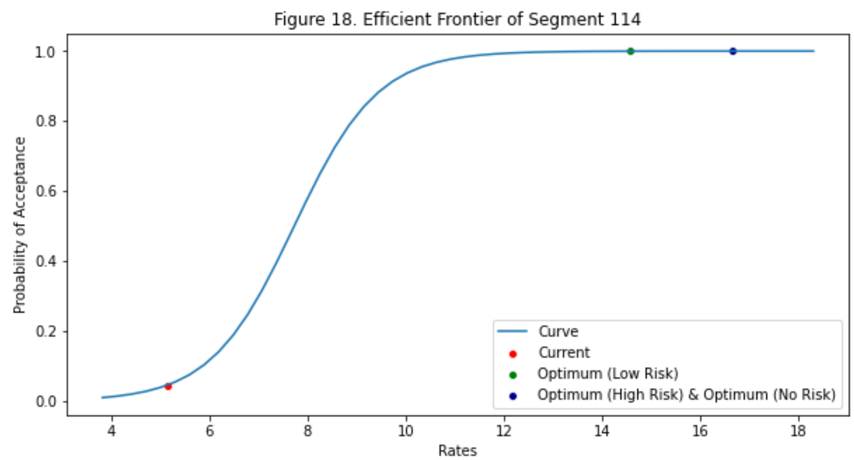
Under No-Risk case, segment $1jk$ (loan_type 1-New) will have the largest increase in profit (300% increase) if APR will be offered at the optimum rates. This is because of the fact that this segment has the lowest loan acceptance rate of only 6%, and thus, offering the optimum APR which is lower than the current rate for most segments under $s1jk$, will increase the probability of winning the customer. This only means that lowering the current APR by 1-2% would be offset largely by the profit gained from having more loans accepted. For segments $2jk$ (loan-type 2-Refinanced), the optimum profit is almost equal to the present estimate. As obsered, the optimum interest rates calculated under $s2jk$ are close to the present rates. This means that some of the present rates under these segments are near optimum.



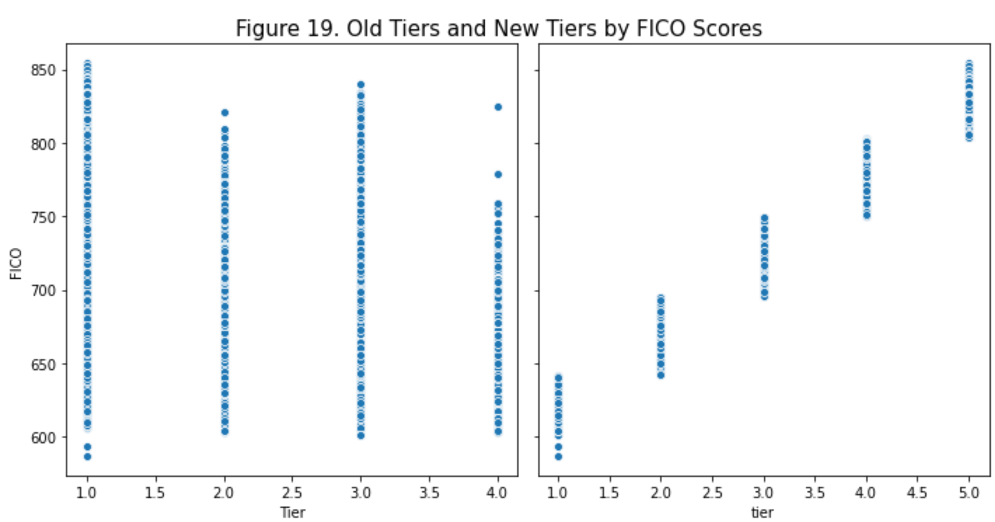
Recommendation

The authors would like to recommend to include the actual data of default payments for the estimation of the actual risk probabilities. The simulated risk probabilities assumes that the probability of default payment linearly increases with FICO score and loan term. However, actual risks may not be linear and vary across different loan type. The optimum rates should account for the risk probabilities as it affects the overall profit. Higher APR also increases the loss due to default payment, and thus pulls-down the profit.

The optimum rate calculated for the Segment 114 is very high at 14% compared to its present rate at 5%. This may be occuring due one of two reasons. The first is that the curve fitting of the given data showing that a higher rate would indeed have a higher probability of winning, or that this curve fitting is erroraneous. In this segment and similar ones, the authors would like to recommend that the behaviour be rechecked and either the current rate or competitive rate for that segment be used in the meantime.



The model can run for a higher number of segments as well with modifications in certain areas. For example, it was noticed that the Tiers assigned to the customers have overlapping FICO scores. To solve this, new tiers were made by splitting the range into 5 equal bins. This increased the number of segments from 60 to 75. While making these additional segments however, the number of points in each segment should be more than 1 but ideally, the more data points, the better the curve fit for predicting the behavior.



```
Optimization conducted successfully =  True
Total Revenue = 476870856217
Optimized Rates =
```

Out[43]:

Car Type	Tier	36	48	60	66	72
N	1	8.768	6.385	9.339	7.486	5.988
	2	6.492	7.523	3.431	6.756	4.273
	3	3.230	2.724	2.869	3.453	3.449
	4	3.075	2.472	2.728	13.255	3.266
	5	3.108	4.901	3.109	8.702	4.380
R	1	6.414	8.221	6.746	6.744	7.067
	2	5.182	5.633	5.743	5.974	6.184
	3	5.111	4.670	4.802	6.208	5.517
	4	6.991	5.277	4.743	6.813	4.385
	5	5.559	4.097	3.905	7.818	3.816
U	1	8.640	8.527	7.894	10.754	6.785
	2	5.621	6.095	5.875	6.502	6.545
	3	4.221	4.541	4.590	5.380	5.279
	4	4.003	4.177	4.205	4.956	5.227
	5	2.221	3.701	4.241	4.589	5.168

It can be seen that adding more segments may not always be a good option. More trials need to be conducted to identify the optimal number of segments to study.

Out[44]:

	60 Segments	75 Segments
Present Rates	256,430,864,687.577	259,713,544,417.498
Optimized Rates	498,974,363,191.395	476,870,856,217.055

Conclusion

In this study, the optimum rates for different risk types were calculated and the potential profit were estimated based on the optimum rates. Based on the estimate, the firm can almost double its current profit by offering the optimum APRs for each segment. In most cases, the optimum APR is lower than the current rate. This means that the firm can earn more profit by increasing its probability of winning, especially for segment $1jk$ (i.e. loan_type **1-New**), where the average loan acceptance rate is only **6%**, which is very low compared to acceptance rate of segments under loan_type **2-Refinanced** and **3-Used** which is around 40%. The breakdown of profit also showed that optimizing the APRs in this segment will increase its profit by **300%**.

Based on the profit breakdown analysis, some of the present interest rates under segment s_{2jk} (loan_type **2-Refinanced**) are already near the optimum. Thus, revising the APRs in these segments must be done graduallly so that current profit under these segments is maintained or improved.

References

[1] Phillips, Robert. (2013). Optimizing prices for consumer credit. Journal of Revenue & Pricing Management. Vol 12. DOI: 10.1057/rpm.2013.9.
[2] Villate, Jaime. (October 14, 2014). Maxima Manual: 37. Cobyla. Retrieved from http://maxima.sourceforge.net/docs/manual/de/maxima_37.html (http://maxima.sourceforge.net/docs/manual/de/maxima_37.html)
[3] Agrawal, V & Ferguson, M. (June 2007). Bid-Response Models for Customized Pricing. Retrieved from https://pdfs.semanticscholar.org/240f/8ebc74c9fdb0c5a0ff0f424d32391b556b19.pdf (https://pdfs.semanticscholar.org/240f/8ebc74c9fdb0c5a0ff0f424d32391b556b19.pdf)