DSC5211C QUANTITATIVE RISK MANAGEMENT

Semester 2, AY 2018/19

Lending Strategies
Risk Aversion Project

Group 12

Name

Cho Zin Tun

Toh Pei Xuan

Zhang Qian Ru

1 April 2019

1. Introduction

The rise of fintech has led to the introduction of various new business models including peer-to-peer (P2P) lending. P2P is a non-traditional form of lending, which is in contrast with the traditional one, where financial institutions offer loans to individuals or small and medium businesses (SME). Instead, P2P lending is a method that enables individuals to borrow and lend money by bypassing an official financial institution (Kagan, 2019). It promises lenders higher return rate on capital invested, compared to a saving account or other investments, as well as lower interest rate, higher level of access for borrowers. Its simplification on the lending process has attracted a large audience and P2P has gained massive popularity to the disadvantage of large financial institutions, that had a monopoly power when it came to lending, after 2008 financial crisis (Gupta, 2018) (FinTechNews, 2018).

1.1. Motivation of problem

P2P lending was pioneered by firms like Lending Club in the United States. The concept then spread to China which was then unregulated. In 2011, this industry in China was far bigger than the rest of the world combined, with outstanding loans of 1.49 trillion yuan, according to data tracker p2p001.com, managed by the Shenzhen Qiancheng Internet Finance Research Institute (Li, 2019) (Shu Zhang, 2018). Many P2P platforms lend to customers whom the commercial banks regard as risky. And when many of such customers could not repay their loans due to effects of an economic slowdown, it led to liquidity crises and many of the P2P platforms in China crashed. In order to tackle such situations, the China government had issued regulations to provide better protection to individual investors, aiming to reduce or avoid such risk spilling over into the overall financial system (Taylor, 2018).

All in all, for lending platforms to comply to regulations and to stay in business, it is crucial for them to study the possible risk that could be incurred from the borrowers when they fail in repaying their loans. Moreover, by doing so, they could also provide a recommendation system for investors to minimise the investment cost. It could be an additional selling point for the platform to attract more investors. Similarly, for investors especially risk-adverse investors, extra precaution can be taken to minimise cost (and risk) by examining the historical data of the borrowers.

1.2. Description of Problem

The objective of this project is to maximize the profit (equivalent to minimizing the cost) of the investors by recommending the category of loans and the corresponding number of units to invest in. This is done by analysing historical data of borrower default rate as well as return rate.

Borrowers were clustered into different categories (based on historical credit grade and loan amount) to simplify the problem as well as provide an overview of the problem. In each of these categories, stochastic modelling of the probabilities of borrowers defaulting and return rate of borrowers who default was done based on historical data. Hence, given the capital (amount that investor wants to invest) and demand of loans in various categories (loans available on the platform at that time), a risk aversion optimization model was built to recommend the number of loans to make for each category. This is done by optimising (minimising) the risk relating to borrowers defaulting and the amount of repayments made by these defaulters.

1.3. Literature Review

The ability to compare random outcomes based on the decision makers' risk preferences is crucial to modelling decision-making problems under uncertainty. Risk aversion modelling is widely used in many industries to either maximize the profit or minimize the cost of the process. Lin Fan et al. (2012) implemented a risk aversion model to examine the changing incentives of investment in different technologies for energy economics field where policies on using technologies are likely to change in favour to reduce greenhouse gas emissions (Lin Fan, 2012). Carrison (2008) built a two-stage stochastic program considering risk aversion to address the problem faced by an electricity retailer who searches to determine the forward contracting portfolio and the selling price for its clients (Carrion, 2008). Also, Cameron et al. developed a model for interdependent demand for health insurance and health care under uncertainty to highlight on the issue of insurance-induced distortions in the demand for health care services (A. C. Cameron, 1988).

2. Data

2.1. Data Source

Loan data was obtained from the website of Lending Club (Club, n.d.) which is one of the world's largest peer to peer lending platform (Lending Club, n.d.) The original raw data had 145 attributes, and contains information of each individual loan, including loan amount, loan status and borrower's information. Six columns were extracted for this study, and three new columns were derived for the calculation of parameters. See

Derived	Total return% at end of policy	total_pymnt / loan_amnt
Derived	Category by loan amount	5 categories based on loan_amnt with bin size of 7800; (Bin size = (max-min)/5)

Derived	Default or not	1 (default group) if "loan_status is "default", "charged off" and "late(31-120days)"; 0 (non-default group) otherwise

Table *1* for description of the information extracted and derived.

Analysis based on completed loans are more representative of the borrower's behaviour in general, hence in this study, analysis was done based on completed loans. If loans were still current, it would be difficult to classify if the borrower defaulted for such loans.

Also, to ensure that parameters derived were more representative of population parameters, year 2015 data with a 36-month term loan data was used. For such loans, majority, if not all loans would have been completed as of today and behaviour of all borrowers (whether the borrower default or not) would become known.

Source	Columns	Description
Original	loan_status	Current status of the loan (Current, Charged off, Late, Fully Paid, Default, Grace Period)
Original	loan_amnt	Loan amount that was provided to the borrower
Original	Grade	Assigned Loan grade (A, B, C, D, E, F,G)
Original	Term	36-months or 60-months term
Original	Issue_d	the month when the loan was issued
Original	total_pymnt	Total payment received up to record day
Derived	Total return% at end of policy	total_pymnt / loan_amnt
Derived	Category by loan amount	5 categories based on loan_amnt with bin size of 7800; (Bin size = (max-min)/5)
Derived	Default or not	1 (default group) if "loan_status is "default", "charged off" and "late(31-120days)"; 0 (non-default group) otherwise

Table 1: Description of columns in the data that was used

3. Mathematical Model

As mentioned in section 1.2, the objective of this project aims to determine how much an investor should invest in the various loan categories by using historical data available to maximize the profit (minimize the cost).

The problem is broken down by categorising loans based on credit grade and loan amount. There is a total of seven credit grades and five groups of loan amounts (total of 35 categories). Credit Grade is represented by ***i*** and has grades A, B, C, D, E, F and G; while Loan Amount is categorised into groups and is represented by $\boldsymbol{j}$ as L1 to L5. At each of the loan amount group, it is assumed that the behaviour of the borrower is the same (regardless of amount in that range) and that the loan amount of each category is the average amount loaned to the group (represented as $\boldsymbol{LA_j}$).

Cost of an investor in this project is based on the total loan amount minus the revenue. Revenue is the amount that the investor/company would receive when repayments are made. This is split into two parts based on whether the borrower defaulted or not. For those who did not default, revenue includes the original loan amount plus interests, where it is assumed that the interest returned ($\boldsymbol{IR_{ij}}$) is the same for all borrowers in the same category. For those who default, revenue refers to the amount returned before the borrower defaulted and it is assumed that this amount follows a normal distribution with some mean and standard deviation for each category represented by $\boldsymbol{Return\_Default\_mean_{ij}}$ and $\boldsymbol{Return\_Default\_sd_{ij}}$ respectively. It is also assumed that borrowers default with a certain probability for each category and is represented by $\boldsymbol{P(Default)_{ij}}$.

The decision variable ($\boldsymbol{K_{ij}}$) is used to determine the number of loans that should be made at each category based on some demand input (the amount of demand at each category, represented by $\boldsymbol{N_{ij}}$ which is a positive integer) and capital (resources) constraint.

The model is written as a minimisation of the Conditional value-at-risk measure (cost).

Problem:

$$\boldsymbol{Min}\ [\boldsymbol{VaR} + \frac{\sum_s \boldsymbol{Z_s}}{(\boldsymbol{1-\beta}) \times \boldsymbol{Number\ of\ scenarios}}]$$

s.t.

$$\boldsymbol{Cost(s)} = \sum_{i,j} \{\boldsymbol{LA_j} - [(\boldsymbol{1} - (\boldsymbol{Default})_{i,j,s}) \times \boldsymbol{LA_j} \times (\boldsymbol{1} + \boldsymbol{IR_{i,j}}) + (\boldsymbol{Default})_{i,j,s} \times \boldsymbol{LA_j} \times \boldsymbol{DefaultReturn}\%_{i,j,s}]\} \times \boldsymbol{K_{i,j}};$$

$$\boldsymbol{Z_s} \geq \boldsymbol{Cost(s)} - \boldsymbol{VaR} \qquad \boldsymbol{for\ all\ s};$$

$$\sum_{i,j} (\boldsymbol{LA_j} \times \boldsymbol{K_{i,j}}) \leq \boldsymbol{capital};$$

$$\boldsymbol{0} \leq \boldsymbol{K_{i,j}} \leq \boldsymbol{N_{i,j}} \qquad \boldsymbol{for\ all\ i\ and\ j};$$

$$\boldsymbol{K_{i,j}} \in \boldsymbol{Z^+} \qquad \boldsymbol{for\ all\ i\ and\ j.}$$

$(\boldsymbol{Default})_{i,j,s}$ are randomly generated value of a scenario s, using a Bernoulli distribution with probability of ***P(Default)**$_{ij}$* to determine if a borrower default (1) or otherwise (0).

$\boldsymbol{DefaultReturn}\%_{i,j,s}$ are randomly generated value of a scenario **s** using a normal distribution with mean ***Return_Default_mean**$_{ij}$* and standard deviation of ***Return_Default_sd**$_{ij}$* to determine the return percentage of a defaulter. Any randomly generated values that are negative are being set as 0.

It is assumed that whether individuals default or not and how much they returned is independent of the behaviour of others.

4. Model Parameterization

Using the data in section 2, several parameters were assumed for each of the 35 categories of borrowers. This includes the following:-

- Average loan amount of each of the 5 loan categories - obtained by averaging the "loan_amnt" column for each of the loan category (refer to **Table 2**);
- Overall 3-year interest that was paid back as a percentage of the loan amount for non-defaulters – obtained by averaging the "Total return% at end of policy" of non-defaulters for each of the credit and loan categories minus 1 (refer to **Table 3**); return rates increase from grade A to G but had different trend for loan categories;

- Overall probability of borrowers defaulting at each category – obtained by dividing the count of 1s in "Default or not" column over the total number of counts in "Default or not" column at each credit and loan category (refer to **Table 4**); default rates increase from grade A to G but had different trend for loan categories;
- Mean of return percentage by defaulters – obtained by averaging the "Total return% at end of policy" of defaulters for each of the credit and loan categories (refer to **Table 5**); no significant pattern across categories; and
- Standard deviation (sd) of return percentage by defaulters – obtained by deriving the sd of "Total return% at end of policy" of defaulters for each of the credit and loan categories (refer to **Table 6**); sd increase from grade A to G but had different trend for loan categories.

Loan Category	L1	L2	L3	L4	L5
Average Amount	5,018	10,473	17,546	24,990	32,970

Table 2: Loan amount of each loan category

	L1	L2	L3	L4	L5
A	9.05%	9.13%	9.14%	9.17%	10.15%
B	13.05%	13.20%	13.22%	13.18%	13.23%
C	17.03%	17.12%	17.14%	17.18%	17.56%
D	21.16%	21.38%	20.92%	20.88%	21.17%
E	24.71%	24.53%	23.58%	23.61%	23.22%
F	29.52%	28.14%	30.07%	26.59%	30.86%
G	31.99%	37.01%	27.33%	34.03%	36.63%

Table 3: Overall 3-year interest of non-defaulters

	L1	L2	L3	L4	L5
A	0.059	0.055	0.055	0.049	0.041
B	0.114	0.125	0.126	0.111	0.093
C	0.182	0.204	0.210	0.195	0.171
D	0.239	0.278	0.284	0.272	0.258
E	0.305	0.339	0.358	0.322	0.340
F	0.382	0.457	0.547	0.426	0.385
G	0.400	0.506	0.522	0.500	0.500

Table 4: Probability of default

	L1	L2	L3	L4	L5
A	0.643	0.644	0.643	0.639	0.648
B	0.644	0.644	0.643	0.651	0.625
C	0.627	0.632	0.624	0.628	0.631
D	0.619	0.619	0.614	0.604	0.588
E	0.609	0.608	0.579	0.599	0.587
F	0.634	0.625	0.559	0.567	0.550
G	0.516	0.545	0.551	0.601	0.611

Table 5: Mean of return% of defaulters

	L1	L2	L3	L4	L5
A	0.259	0.256	0.256	0.241	0.249
B	0.270	0.262	0.261	0.260	0.246
C	0.282	0.272	0.272	0.274	0.284
D	0.294	0.294	0.280	0.293	0.282
E	0.306	0.296	0.291	0.309	0.279
F	0.326	0.328	0.323	0.338	0.284
G	0.300	0.294	0.280	0.534	0.382

Table 6: SD of return% of defaulters

5. Analysis of the optimal policy

The model was run based on the parameters in section 4 and input values of demand of 10 persons per category and Capital of $1,000,000.

Various number of scenarios were tried and results are given in Table 7. The change in CVaR between 10,000 scenarios and 100,000 scenarios is much smaller than the change in CVaR between 100 and 1,000 scenarios. This shows that the model stabilises when the number of scenarios increased. This is due to the stochastic part of the model stabilising to the expected values of the distribution. 10,000 scenarios were used for the rest of the analysis.

No. of Scenarios	100	1,000	10,000	100,000
CVaR	-33,831.072	-11,392.329	-10,622.910	-10,507.184

Table 7: Results when beta is 0.8 with varying number of scenarios

Results were also tabulated for various Beta.

<table>
<tr><th>Beta</th><th>Policy</th><th>CVaR</th></tr>
<tr><td>0</td><td><table><tr><th></th><th>L1</th><th>L5</th></tr><tr><td>A</td><td>1.000</td><td>10.000</td></tr><tr><td>B</td><td>1.000</td><td>10.000</td></tr><tr><td>C</td><td></td><td>10.000</td></tr></table></td><td>-84,033.25</td></tr>
<tr><td>0.7</td><td><table><tr><th></th><th>L1</th><th>L2</th><th>L3</th><th>L4</th><th>L5</th></tr><tr><td>A</td><td>10.000</td><td>8.000</td><td>4.000</td><td>5.000</td><td>10.000</td></tr><tr><td>B</td><td>8.000</td><td>4.000</td><td>2.000</td><td>2.000</td><td>2.000</td></tr><tr><td>C</td><td>5.000</td><td>2.000</td><td>1.000</td><td>1.000</td><td></td></tr><tr><td>D</td><td>2.000</td><td></td><td></td><td></td><td></td></tr><tr><td>E</td><td>1.000</td><td></td><td></td><td></td><td></td></tr></table></td><td>-26,131.20</td></tr>
<tr><td>0.8</td><td><table><tr><th></th><th>L1</th><th>L2</th><th>L3</th><th>L4</th><th>L5</th></tr><tr><td>A</td><td>10.000</td><td>8.000</td><td>4.000</td><td>3.000</td><td>10.000</td></tr><tr><td>B</td><td>9.000</td><td>3.000</td><td>2.000</td><td>2.000</td><td>1.000</td></tr><tr><td>C</td><td>5.000</td><td>2.000</td><td>1.000</td><td>1.000</td><td>1.000</td></tr><tr><td>D</td><td>4.000</td><td>1.000</td><td>1.000</td><td></td><td></td></tr><tr><td>E</td><td>2.000</td><td>1.000</td><td></td><td></td><td></td></tr><tr><td>F</td><td>1.000</td><td></td><td></td><td></td><td></td></tr></table></td><td>-10,622.91</td></tr>
<tr><td>0.9</td><td>(ALL 0.000)</td><td>0</td></tr>
<tr><td>0.99999</td><td>(ALL 0.000)</td><td>0</td></tr>
</table>

Table 8: Results of various risk profile (by varying beta)

Table 8 shows values of CVaR when beta varies. When beta is 0, the CVaR value obtained refers to the expected cost (risk-neutral). The negative CVaR value implies that a profit of about $84K is expected when the corresponding recommended policy is implemented. Recommended policy is interpreted based on the number of loans to lend out at each credit category *i* and loan category *j*. For example, at the risk neutral position, it means to make loans to all 10 people in loan category 5 with credit grade A, B and C, and another 1 loan each to loan category 1 and credit grade A and B. While at the worst-case scenario of beta being close to 1, the model recommends not to take any risk by making any loans, and hence the worst-case scenario would result in a profit of 0. In general, expected profits decrease (CVaR increase) when beta increases (more risk averse). At beta of 0.8 and 0.7, the policy recommended is such that at 0.8 (less risk), the loans are spread across more categories as compared to when beta is 0.7. Risk is reduced by diversification of risk by lending to more groups of people.

6. Sensitivity Analysis

Sensitivity was performed for this model by changing various parameter values namely the capital, demand of loans and standard deviation of return percentage by those who default. There was no change to the policy for beta more than 0.9 for all cases.

Capital	Beta=0	Beta=0.7	Beta=0.8
0.5mil	L1 L5 A 1.000 B 10.000 C 5.000	L2 L3 L4 L5 A 1.000 1.000 5.000 10.000	L5 A 10.000
	-42,079.74	-16,332.87	-4,334.85
2mil	L1 L2 L3 L4 L5 A 1.000 10.000 10.000 B 10.000 10.000 10.000 10.000 10.000 C 10.000 3.000 10.000 D 10.000	L1 L2 L3 L4 L5 A 10.000 10.000 10.000 10.000 10.000 B 10.000 10.000 7.000 6.000 5.000 C 10.000 6.000 4.000 3.000 2.000 D 10.000 2.000 1.000 1.000 E 6.000 1.000 F 3.000	L1 L2 L3 L4 L5 A 10.000 10.000 9.000 8.000 10.000 B 10.000 8.000 6.000 4.000 3.000 C 10.000 5.000 3.000 2.000 2.000 D 9.000 2.000 1.000 1.000 1.000 E 6.000 2.000 F 3.000
	-156,824.49	-42,911.573	-15,648.00

Table 9: Results from changing capital at various beta values

Referring to Table 9, policies changed when capital changed. In general, when capital (supply) increases, more loans (demand) can be made and CVaR would hence drop (increase profits). With increased capital, a larger beta continued to recommend making loans in more categories. However, this was not the case when capital decreased and may be due to capital not being able to cover the increased cost when more loans are made (increased risk).

Demand	Beta=0	Beta=0.7	Beta=0.8
5	L1 L2 L3 L4 L5 A 1.000 5.000 5.000 B 5.000 5.000 5.000 5.000 5.000 C 5.000 1.000 5.000 D 5.000 E 1.000 F 1.000	L1 L2 L3 L4 L5 A 5.000 5.000 5.000 5.000 5.000 B 5.000 5.000 3.000 3.000 2.000 C 5.000 3.000 2.000 1.000 1.000 D 5.000 2.000 1.000 1.000 E 3.000 1.000 F 2.000	L1 L2 L3 L4 L5 A 5.000 5.000 5.000 5.000 5.000 B 5.000 5.000 3.000 3.000 2.000 C 5.000 3.000 1.000 1.000 1.000 D 5.000 1.000 1.000 1.000 E 3.000 1.000 F 2.000
	-78,234.948	-20,992.939	-7,177.107
20	L1 L5 A 1.000 B 1.000 20.000 C 10.000	L1 L2 L3 L4 L5 A 6.000 5.000 3.000 2.000 20.000 B 4.000 2.000 1.000 1.000 1.000 C 2.000 1.000 D 2.000 E 1.000	L1 L2 L3 L4 L5 A 9.000 5.000 3.000 2.000 15.000 B 6.000 3.000 2.000 1.000 1.000 C 4.000 1.000 1.000 1.000 1.000 D 3.000 1.000 E 2.000 F 1.000
	-84,222.325	-33,651.442	-10,867.929

Table 10: Results from changing demand at various beta values

Table 10 shows results of changing demand at various beta values. In general, when demand increase, CVaR decreases (increased profits) as more loans at certain categories could be increased. For example, the risk neutral profile increased the units loaned to credit grade B, loan category 5, while the risk averse (0.8) increase units for credit grade A with loan category 1 and 5. This showed that different risk profile had different categories that were preferred.

SD (default return%)	Beta=0	Beta=0.7	Beta=0.8
halved	L1 L5 A 1.000 10.000 B 1.000 10.000 C 10.000	L1 L2 L3 L4 L5 A 10.000 8.000 4.000 4.000 10.000 B 9.000 4.000 3.000 2.000 2.000 C 6.000 2.000 1.000 D 4.000 1.000 E 2.000	L1 L2 L3 L4 L5 A 10.000 8.000 5.000 4.000 6.000 B 10.000 4.000 3.000 3.000 2.000 C 6.000 2.000 1.000 1.000 1.000 D 5.000 2.000 E 3.000 F 1.000
	-83,992.300	-29,209.705	-15,536.174
doubled	L1 L5 A 1.000 10.000 B 1.000 10.000 C 10.000	L1 L2 L3 L4 L5 A 10.000 8.000 3.000 6.000 10.000 B 6.000 3.000 2.000 1.000 1.000 C 4.000 2.000 1.000 1.000 1.000 D 3.000 1.000 E 3.000 1.000 F 2.000	L5 A 10.000
	-87,237.430	-21,522.095	-2,910.978

Table 11: Results from changing sd of return% by defaulters, at various beta values

Table 11 shows results of changing standard deviation (sd) of return by defaulters at various beta values. In general, when standard deviation increase, there is no constant pattern for CVaR. For the risk neutral, there is no change to policy but CVaR decreases slightly. This is likely due to changes to the realised return rate and hence affected the overall profits. For beta at 0.7 and 0.8, the CVaR increases (profits drop) when sd increases. This may be due to the model recommending less risky choices to offset the increased risk due to increase in sd.

Credit grade G remained to be not recommended under all above scenarios which showed that the risk is too high at this category as default rate is the highest.

7. Conclusion

This study analysed historical loan data and had achieved the objective of minimising the cost by giving recommendation on loan categories and investment amount. Analysis was done for different risk profiles (including worst case, risk neutral, and different level of risk aversion) and varying various parameter values.

Overall, this model does not recommend to make any loans if the investor is more risk averse and shows that this industry has high risks. This corresponds to section 1.2 and the current situation where many P2P companies closed down. In particular, the model had showed that credit grade G in general has too high risk such that no policies under various situations were recommended for that category.

The model could be further improved from various aspects. One way is to increase the granularity of the groupings from five loan categories and seven credit grades to take into account sub grades and more loan categories. With enough computation power, this problem could even be solved at an individual basis. The other way is to increase the dimensions of the model. Currently only loan amount and credit grade is considered. Further work could be done to include other attributes which would impact default and return rate, like employment status and salary, to extend this model into a multi-dimension table for the grouping of categories.

Lastly, this model made assumptions about the distribution of default rate and return distribution. More work could be done to further analyse these distributions to increase the model accuracy.

References

A. C. Cameron, P. K. (January, 1988). A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia. *The Review of Economic Studies, 55*(1), 85–106.

Carrion, M. (20-24 July, 2008). Risk-aversion modeling: Medium-term electricity retailer decision-making. *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century.* IEEE. Retrieved from 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century.

Club, L. (n.d.). *Lending Club Offical Website* . Retrieved from https://www.lendingclub.com/info/download-data.action

FinTechNews. (Sep, 2018). *P2P Lending Faces Challenges Ahead* . Retrieved from FinTechNews: http://fintechnews.sg/23957/lending/p2p-lending-faces-challenges-ahead/

Gupta, N. G. (May, 2018). *The History of Peer-To-Peer (P2P) Lending*. Retrieved from ZuuOnline: https://zuuonline.sg/investment/crowdfunding/the-history-of-peer-to-peer-p2p-lending/

Kagan, J. (Feb, 2019). *Peer-To-Peer Lending (P2P)*. Retrieved from Investopeida: https://www.investopedia.com/terms/p/peer-to-peer-lending.asp

Lending Club. (n.d.). Retrieved from Wikepedia: https://en.wikipedia.org/wiki/Lending_Club

Li, O. (April, 2019). *Another P2P Lending Crisis in China, 220,000 Investors Losing 14.5 Billion Yuan*. Retrieved from EPOCH TIMES: https://www.theepochtimes.com/another-p2p-lending-crisis-in-china-220000-investors-losing-14-5-billion-yuan_2862296.html

Lin Fan, C. S. (2012). Electricity capacity investment under risk aversion: A case study of coal, gas, and concentrated solar power. *Energy Economics*(34), 54–61.

Shu Zhang, E. G. (12 August, 2018). *Beijing struggles to defuse anger over China's P2P lending crisis*. Retrieved from Reuters: https://www.reuters.com/article/us-china-lenders-p2p-insight/beijing-struggles-to-defuse-anger-over-chinas-p2p-lending-crisis-idUSKBN1KX077

Taylor, A. (12 12, 2018). *What's Going On With China's P2P Lending Crisis?* Retrieved from lending-times: https://lending-times.com/2018/12/12/whats-going-on-with-chinas-p2p-lending-crisis/

Appendix: Code and Data

DATA:

Full dataset from: https://www.lendingclub.com/info/download-data.action

Relevant fields from: Final Result Loan.xlsx

GAMS CODE:

```
$title LendingClub Problem

set
 i  credit grade     / A, B, C, D, E, F, G/
 j  loan category    /L1*L5/
 s  scenarios        /S1*S10000/;

Scalar beta      prob. of not being in the tail              /0.8/;

Table N(i,j)          Number of persons in each category - integer
         L1        L2        L3        L4        L5
A        10        10        10        10        10
B        10        10        10        10        10
C        10        10        10        10        10
D        10        10        10        10        10
E        10        10        10        10        10
F        10        10        10        10        10
G        10        10        10        10        10;

parameters LA(j) loan_amount
/L1=5018
L2=10473
L3=17546
L4=24990
L5=32970/;

Table intr(i,j)           interest rate each category_did not default
         L1           L2           L3           L4           L5
A        0.0905       0.0913       0.0914       0.0917       0.1015
B        0.1305       0.1320       0.1322       0.1318       0.1323
C        0.1703       0.1712       0.1714       0.1718       0.1756
D        0.2116       0.2138       0.2092       0.2088       0.2117
E        0.2471       0.2453       0.2358       0.2361       0.2322
F        0.2952       0.2814       0.3007       0.2659       0.3086
G        0.3199       0.3701       0.2733       0.3403       0.3663;


Table PD(i,j)          probability of default
         L1            L2            L3            L4            L5
A        0.05906       0.05519       0.05486       0.04869       0.04073
B        0.11411       0.12463       0.12611       0.11085       0.09326
C        0.18181       0.20421       0.20996       0.19475       0.17072
D        0.23909       0.27811       0.28393       0.27158       0.25844
E        0.30539       0.33930       0.35764       0.32247       0.33952
F        0.38170       0.45714       0.54688       0.42647       0.38462
G        0.40000       0.50588       0.52174       0.50000       0.50000;

Table MDR(i,j)          mean of return rate for those who default
         L1           L2           L3           L4           L5
A        0.6429       0.6445       0.6427       0.6387       0.6476
B        0.6436       0.6443       0.6432       0.6514       0.6254
C        0.6266       0.6320       0.6238       0.6285       0.6308
D        0.6194       0.6193       0.6138       0.6036       0.5883
E        0.6089       0.6082       0.5790       0.5991       0.5871
F        0.6339       0.6253       0.5594       0.5668       0.5496
G        0.5161       0.5446       0.5505       0.6015       0.6112;

Table SDDR(i,j)          sd of return rate for those who default
         L1           L2           L3           L4           L5
A        0.2591       0.2556       0.2563       0.2407       0.2492
B        0.2696       0.2624       0.2613       0.2604       0.2463
C        0.2824       0.2723       0.2717       0.2741       0.2837
D        0.2941       0.2936       0.2796       0.2930       0.2821
E        0.3064       0.2958       0.2911       0.3090       0.2786
F        0.3260       0.3280       0.3234       0.3383       0.2837
G        0.3001       0.2936       0.2796       0.5337       0.3817;

Scalar capital      capital available for borrowing                /1000000/;

parameters  default(i,j,s)      default status
            returnd(i,j,s)      default_percentage returned;
default(i,j,s) = randbinomial(1,(PD(i,j)));
returnd(i,j,s) = normal(MDR(i,j),SDDR(i,j));
returnd(i,j,s) = ifthen(returnd(i,j,s)<0, 0, returnd(i,j,s));

variables
        ppl(i,j) 'number of people for credit grade i, loan amount j'
        cost    'cost minus revenue'
        var    value at Risk
        cvar   conditional value at risk
        z(s)   tail profit;

Integer variables ppl;
positive variable z;
equations   cdef   costdefinition
            cc     capital constraint
            nK(i,j)   number of ppl for each category
            tails(s) calculate the value of the tail
            cvar_eq objective function  ;

cdef..      cost =e= sum((i,j,s), (-default(i,j,s)*returnd(i,j,s)*LA(j)-(1-default(i,j,s))*LA(j)*(1+intr(i,j)))*ppl(i,j))/card(s)+sum((i,j), LA(j)*ppl(i,j));

cc..        sum((i,j), LA(j)*ppl(i,j)) =l= capital;
nK(i,j)..       ppl(i,j) =l= N(i,j);
tails(s)..      z(s)=g=sum((i,j), (-default(i,j,s)*returnd(i,j,s)*LA(j)-(1-default(i,j,s))*LA(j)*(1+intr(i,j)))*ppl(i,j))+sum((i,j), LA(j)*ppl(i,j))-var;
cvar_eq..       cvar =e= var+1/((1-beta)*card(s))*sum(s,z(s));

models chance model        /all/

solve chance minimizing cvar using MIP;

display ppl.l, cost.l,var.l,cvar.l;
```