Working Paper, Step 6, Group B

# Evaluation/Deployment

## Introduction

In this final step, we transition from technical analysis to evaluating the business implications of our findings. This step synthesizes our findings, addresses the original questions posed, and evaluates the practical significance of our results. We explore how our model's performance compares to traditional methods, particularly the grade-based approach, and how it meets or exceeds the company's 2% return benchmark. The key objective of this final phase is to demonstrate the value and the business impact of our approach to "GreatYields" and to provide actionable recommendations based on our findings.

## Project overview

Over the past ten weeks, our project has systematically addressed the investment challenges faced by "GreatYields" through a series of well-defined phases. We started with business understanding, where we identified the project's goals and potential issues. In the data preparation step, we cleaned and explored the data to assess its quality and relevance. Following this, we moved on to model building, parameter fine-tuning, and evaluations. Eventually, our chosen model involved designing a two-stage model for loan classification- above and below 2% and numerical yield prediction. In the last step, we focused on testing the model's effectiveness through rigorous evaluation techniques to be able to address the customer's inquiries and evaluate our model from a strategic business perspective for our customer.

## "Model Reliability and Accuracy"

Before addressing the questions, it is imperative to establish our model's reliability and accuracy to ensure well-informed investment decisions.
In Appendix A, we present a graph plotting the weighted yield against cumulative investments for our model's predicted returns and the expected returns, calculated using a detailed yield formula described in Question 1. We have chosen to use weighted yield as our primary metric because it assigns greater significance to larger loan investments and ensures that the overall yield more accurately reflects the performance of substantial investments.
The graph shows our model's predictions align well with actual outcomes. Initially, the model may underestimate the yield, but it becomes more accurate with more data points (loans). Appendix A also includes a table displaying predicted returns, actual expected returns, the number of loans in the portfolio, and the standard deviation (SD) of expected returns. This close alignment underscores the model's high accuracy, which improves with larger investment amounts. The SD column reflects the volatility and risk of each loan portfolio, further discussed in Question 5.
This validation highlights the reliability of our model in predicting returns, forming a solid foundation for strategic investment decisions.

**Before answering your questions, it's important to note that we chose to address them in a different order. This approach better represents the chronological insights we gained throughout the project.**

**Question 1: What are the expected realized returns for the different loan grades? How are the returns distributed for each grade?**

To determine the expected realized returns, we calculated the yield for each loan grade using a formula that incorporates the time value of money and adjusts for the annual return.

$$\frac{p-f}{f} \cdot \frac{12}{m}$$

- Cash flow data is missing
    - p: full amount recovered from the loan
    - f: full amount invested in a loan
    - m: the actual duration of the loan

Working Paper, Step 6, Group B

Our analysis revealed that riskier loan grades tend to have lower returns and higher standard deviations as shown in Appendix B, indicating greater risk(Elaborate in question 5). The returns are distributed with a left tail, suggesting potential downside risks.

- Expected Realized Returns:
  - Higher-grade loans (A-C) tend to show more predictable and concentrated distributions of returns.
  - Riskier grades (D-G) exhibit wider interquartile ranges, indicating greater dispersion of returns.
  - Riskier loans (grades E-G) have an even split (50-50) between loans with returns above and below 2%.
- Distribution of Returns:
  - The returns for higher-grade loans are more predictable with less variability.
  - Lower-grade loans show higher standard deviations, indicating greater risk and variability in returns.
  - The distribution of returns for riskier grades (D-G) tends to have a left tail, suggesting potential downside risks.

Based on our findings, we will use grade A as the baseline benchmark for loans due to its more predictable and concentrated returns.

Potential pitfalls:
1.Loan grade transition- Loans might change risk grades over time, and not accounting for these transitions can affect the accuracy of return distributions.
2. Historical Data Reliance- Relying solely on historical data assumes past performance predicts future results, which might not hold true.
3. Disregarding investment size – consider weighted expected returns.
4.Changing economic conditions- Changing economic conditions can impact loan performance, making historical returns less indicative of future returns.

**Question 5: What is the risk level entailed in such investment (as measured by the volatility)?**

We assessed the risk level of our investment portfolio by analyzing the standard deviation of returns for each loan grade. Standard deviation measures the dispersion of return rates around the mean, offering insight into the investment's volatility. The risk level associated with each loan grade is captured by the weighted standard deviation of returns. As we move from lower to higher risk grades, the weighted standard deviation increases, indicating greater volatility in returns.
To ensure our risk assessment accurately reflects the overall portfolio performance, we utilized weighted standard deviation. This method accounts for the size of each investment, giving more importance to larger investments and providing a clearer picture of overall risk.

Methodology:

1. Calculate Standard Deviation for Each Loan Grade: We computed the standard deviation of returns for each loan grade by analyzing historical return data to determine the variability of returns within each grade.
2. Weight the Standard Deviations: We applied weights based on the size of the investments in each loan grade. The weighted standard deviation formula considers the proportion of the total investment represented by each loan grade, ensuring larger investments have a greater impact on the overall risk assessment.
3. Analyze Impact of Weighting: We examined how weighting affected the variability of yields, helping us understand the contribution of different loan grades to the portfolio's overall risk as shown in Appendix C.

Our analysis demonstrates that weighted standard deviation is an effective measure of volatility for assessing the risk level of loan investments. By accounting for the size of each investment, we provide a more accurate representation of overall portfolio performance. For a detailed explanation of our methodology and findings, please refer to Appendix C.

Working Paper, Step 6, Group B

Potential pitfalls:
1.Inaccurate loan data - Errors in the recorded loan amounts, payments, or durations can lead to incorrect risk assessments.
2.Assumptions in weighting -weighting Bias: The method used to assign weights to different loans might introduce bias. static weights- using static weights that don't adjust for changing market conditions or loan performance can lead to outdated risk assessments.
3.Risk concentration - Failing to account for the diversification of the loan portfolio can lead to an overestimation or underestimation of risk. A concentrated portfolio might be riskier than a diversified one.

**Question 4: What "average" returns can "GreatYields" expect from investing in peer lending loans? Keep in mind that ultimately, the goal is to maximize returns**

Based on the analysis presented and the methodology outlined in Appendix D, the graph below depicts the weighted annual returns for various investment amounts. To ensure optimal investment choices, we sorted the loans in descending order based on their predicted annual yield. This prioritization allows us to identify the loans with the highest expected returns. Next, we calculated the cumulative investment amount by progressively adding the loan amounts according to the sorted order.

It can be inferred from the graph that "GreatYields" consistently surpasses the benchmark return of 2% currently achieved by the company. The graph reveals that by utilizing our two-stage model, "GreatYields" has the potential to achieve an average annual return of 5.19% considering the maximum investment amount of $832,678,750.00.

Additionally, the attached table in Appendix D provides a more accessible way to view the expected returns for different investment amounts. It includes the number of loans invested in each portfolio and the associated risk, represented by the standard deviation of the expected returns. This detailed breakdown helps to clearly understand the potential returns and risks associated with various levels of investment.

Potential pitfalls:
1.The 2019 data used as the foundation for the model may not accurately represent the current market conditions.
2. Assumptions and limitations of the model: The analysis is based on particular assumptions and a two-stage model, which might have inherent constraints or fail to encompass all pertinent factors. Reports B-E elaborate on the possible drawbacks of the model.

**Question 3: If the data are indeed informative, what increased performance can be expected, compared to a baseline of simply selecting loans based on their ratings (grades)?**

As illustrated in the graph in Appendix E, our model demonstrates significantly higher weighted yields across various cumulative investment amounts compared to the Grade A benchmark. This robust performance highlights the advantage of utilizing comprehensive loan data over simple grade-based selection criteria.

Unlike our model's results that ensure optimal investment choices by sorting the loans in descending order based on their predicted annual yield, the Grade A loans were chosen for each cumulative investment randomly. As a result, the expected return for Grade A loans may vary significantly. This further emphasizes the advantage of our data-driven model, which systematically identifies the best investment opportunities to maximize returns while minimizing risk.

Furthermore, as shown in the table comparison in Appendix E, our model surpasses the standard approach not only in terms of expected return but also by reducing the standard deviation, which stands as an alternative measure of volatility. This reduction in standard deviation is particularly crucial in the realm of investment strategy, as it provides a more secure foundation for making investment decisions. By lowering the risk associated with the portfolio, our model ensures that the returns are not only higher but also more stable and reliable.

Working Paper, Step 6, Group B

**Question 2: Are the available loan data informative, thus can help selecting loans to invest in (i.e., can the data help choose loans better than random selection or selection by simple criteria, e.g., loan grade)?**

By utilizing this rich dataset, we have successfully developed a sophisticated model capable of constructing an investment portfolio that consistently delivers higher returns than the Grade A benchmark model. Our model's effectiveness is not limited to selecting loans with positive returns; it excels in identifying the best loans to invest in, achieving high yields while maintaining a low-risk profile. This is evident from the model's ability to avoid loans that are prone to charge-offs, thereby enhancing the stability and reliability of the investment portfolio.
The superior performance of our model, both in terms of higher returns and lower risk, highlights the crucial role of leveraging comprehensive loan data over simple grade-based selection criteria.

Potential Pitfalls:
1.Assumptions and Simplifications: The model and analysis depend on specific assumptions and simplifications that might not adequately reflect the intricacies of real-world investment situations.
2.Overfitting: There is a risk that the model may have been overfitted to the particular dataset used, potentially leading to decreased performance when applied to new or unseen data.
3.Data Quality and Completeness: The accuracy of the model depends on the quality and completeness of the loan data used. Incomplete or inaccurate data could lead to suboptimal investment decisions.

**Recommendations:**

1. Implement the predictive model to enhance loan selection and maximize returns.
2. Diversification of investments- recommend diversifying the investment portfolio by including a mix of loan grades to balance the risk and return. This approach can help mitigate the risk associated with high-grade loans while still capturing potential high returns from lower-grade loans.
3. Conduct further analysis with updated and more comprehensive datasets.
4. Maintain a constant monitoring and adjustment of the model to reflect changing market conditions and external factors.

Refer to the appendices for detailed methodologies, analyses, and graphical representations supporting these findings and recommendations.

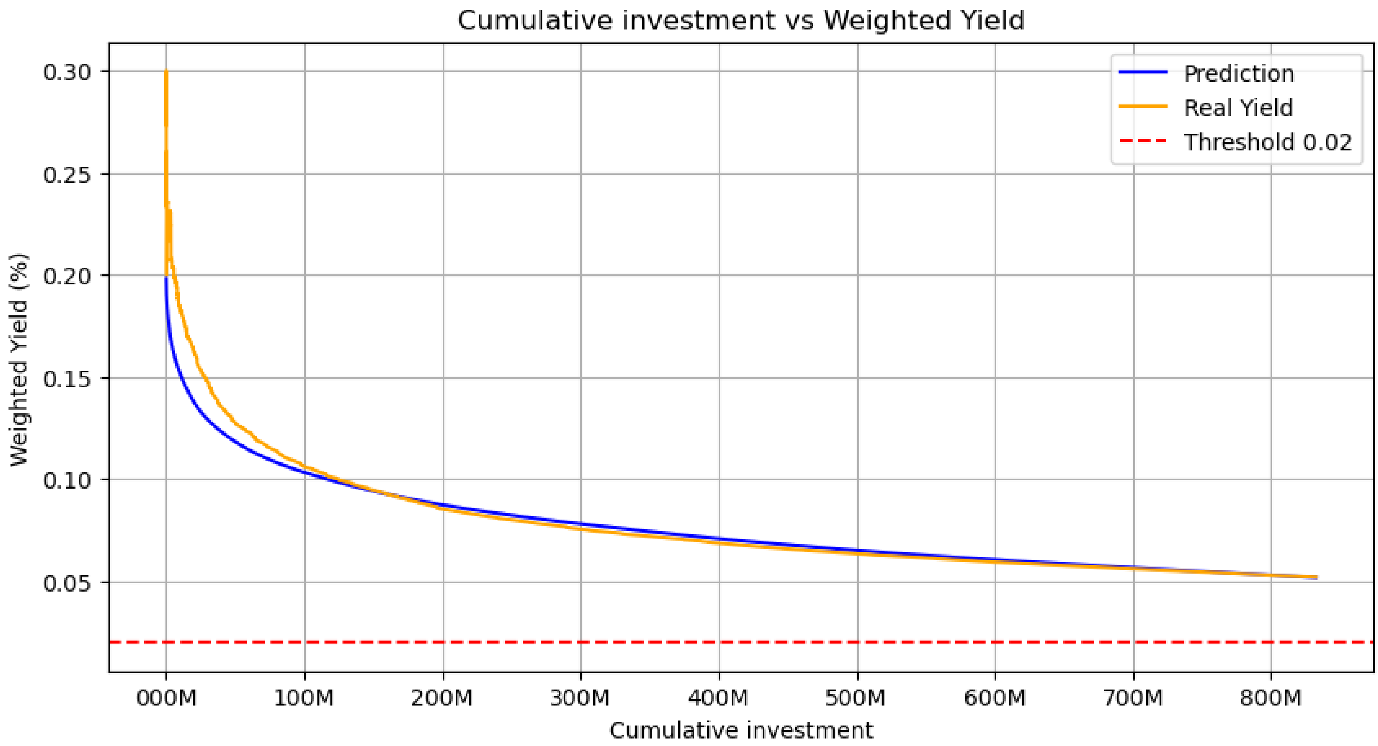
**Conclusion:**

This final step synthesizes the entirety of our efforts, addressing the business questions identified at the project's outset and responding to Walter's inquiries. Our analysis confirms the significant value of the available loan data, enabling the development of a model that outperforms traditional grade-based methods in predicting realized returns. The data's depth allows for more informed investment decisions, resulting in higher returns than the company's benchmark.

The business significance of these findings is substantial: by leveraging our model, the company can optimize its investment strategies, achieve better financial outcomes, and enhance overall portfolio performance. Practically, we recommend integrating our model into the investment decision-making process, continuously updating it with new data to maintain its accuracy and relevance. Additionally, further analysis and refinement could yield even greater insights, ensuring the company stays ahead in a competitive market.

# Appendices

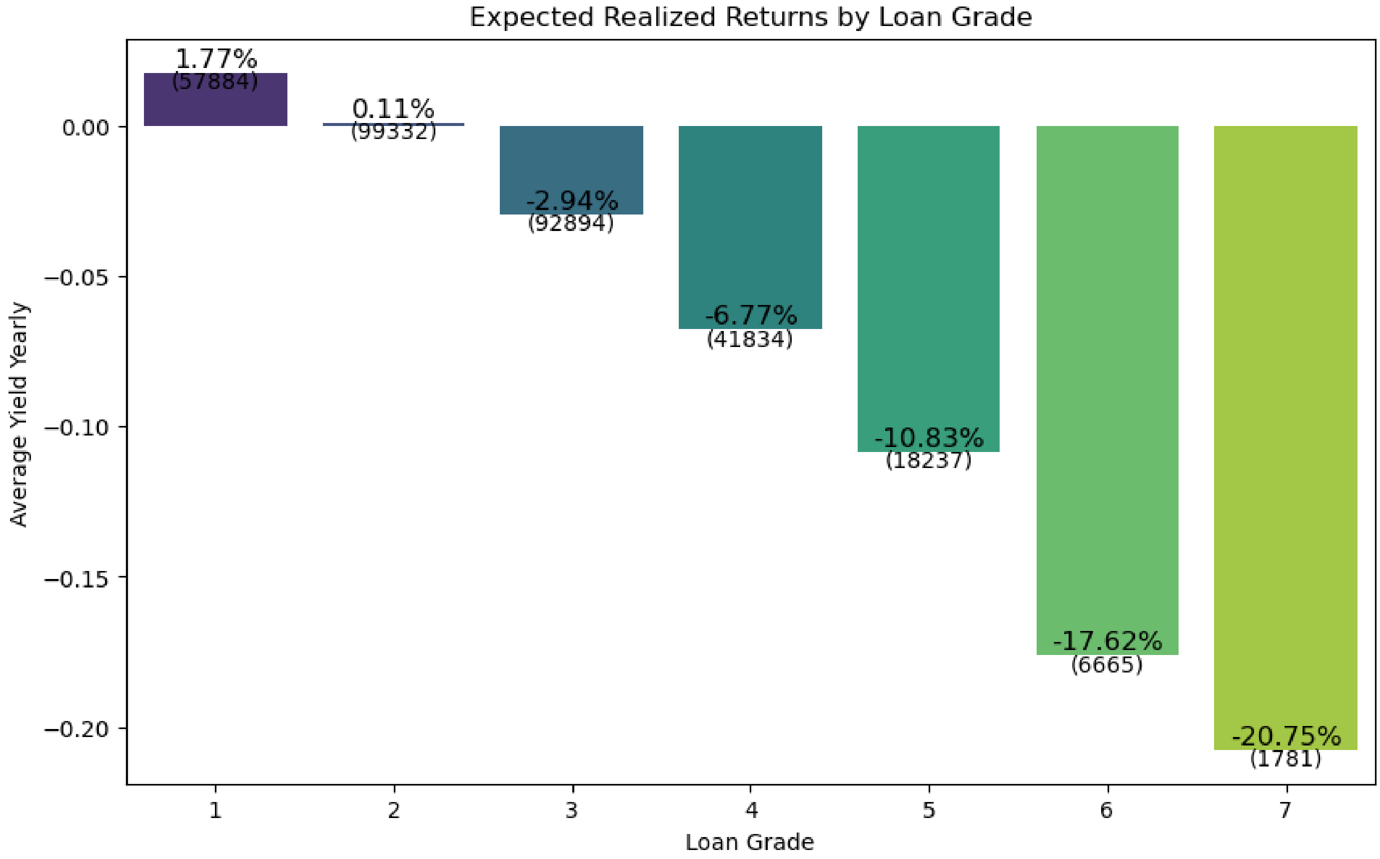
## Appendix A:




| Investment Amount | Predicted Return (%) | Expected Return (%) | Number of Loans | SD of Expected Return (%) |
|---|---|---|---|---|
| $100,000,000.00 | 10.35% | 10.60% | 7826 | 0.107% |
| $200,000,000.00 | 8.76% | 8.53% | 15902 | 0.102% |
| $300,000,000.00 | 7.82% | 7.54% | 23437 | 0.096% |
| $400,000,000.00 | 7.10% | 6.86% | 31055 | 0.094% |
| $500,000,000.00 | 6.52% | 6.34% | 37985 | 0.090% |
| $600,000,000.00 | 6.07% | 5.92% | 44600 | 0.088% |
| $700,000,000.00 | 5.68% | 5.60% | 51151 | 0.085% |
| $800,000,000.00 | 5.32% | 5.29% | 57473 | 0.083% |
| $832,678,750.00 | 5.20% | 5.19% | 59378 | 0.083% |

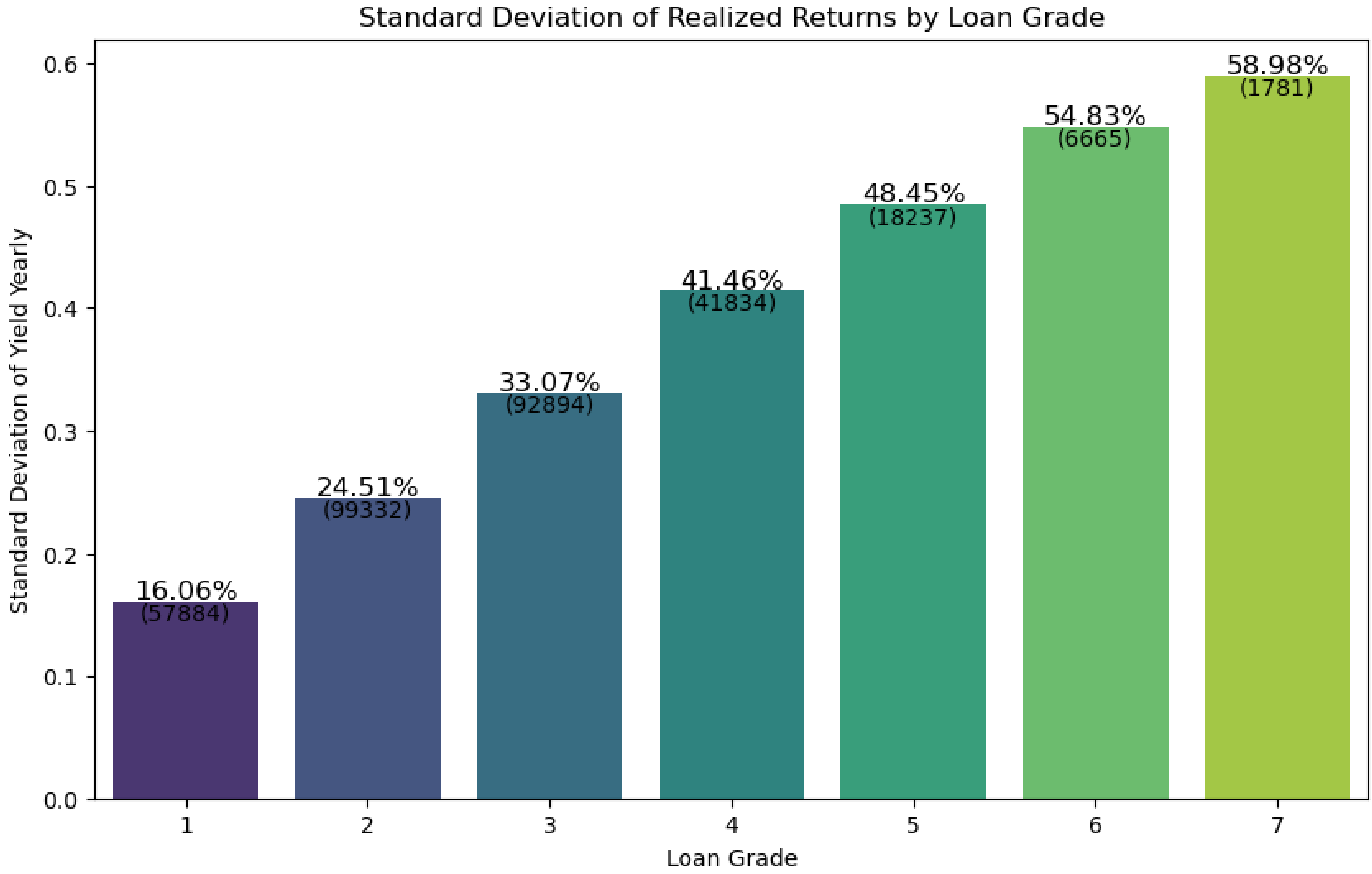
Working Paper, Step 6, Group B

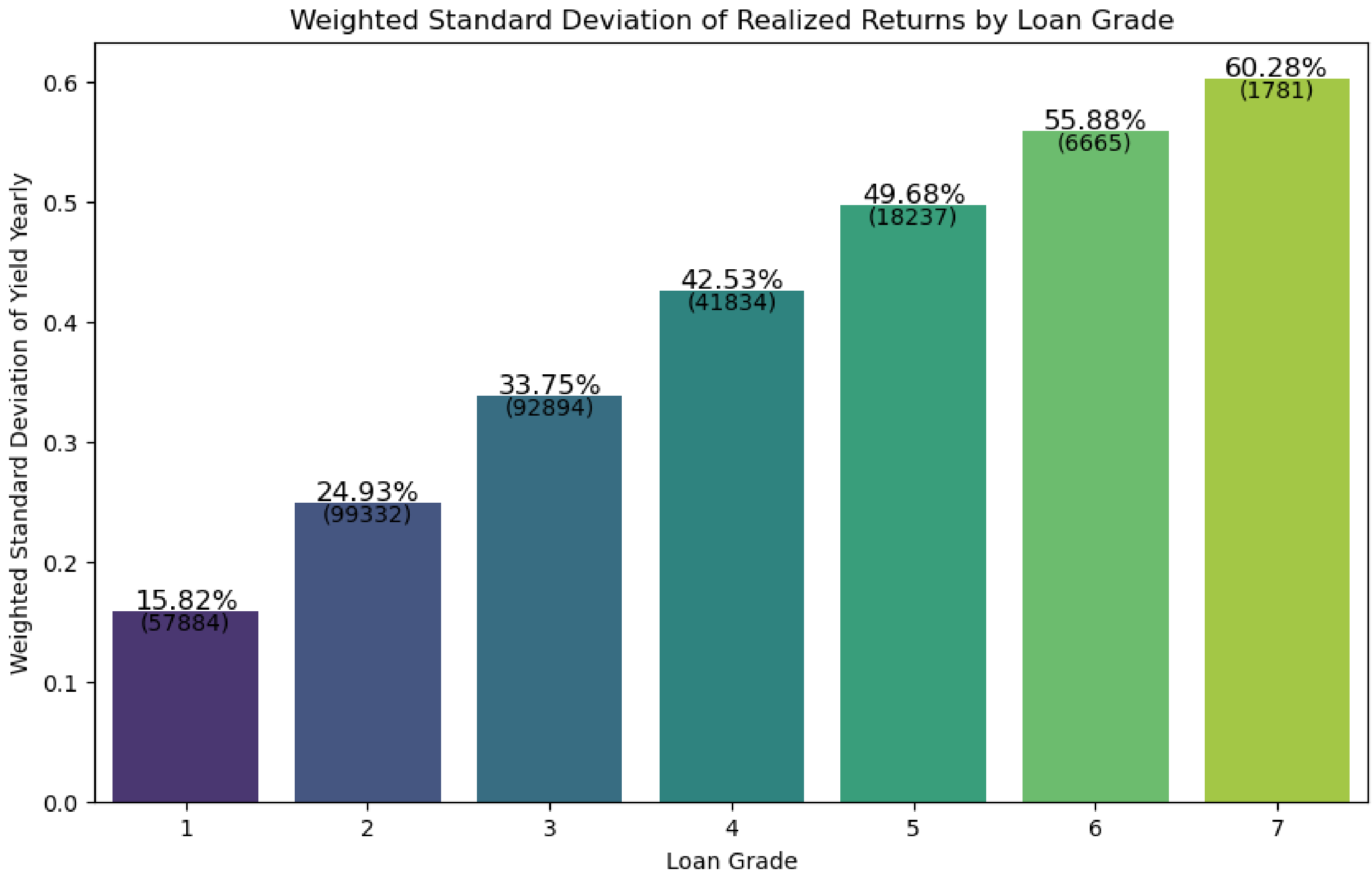
**Appendix B:**



Working Paper, Step 6, Group B

**Appendix C:**



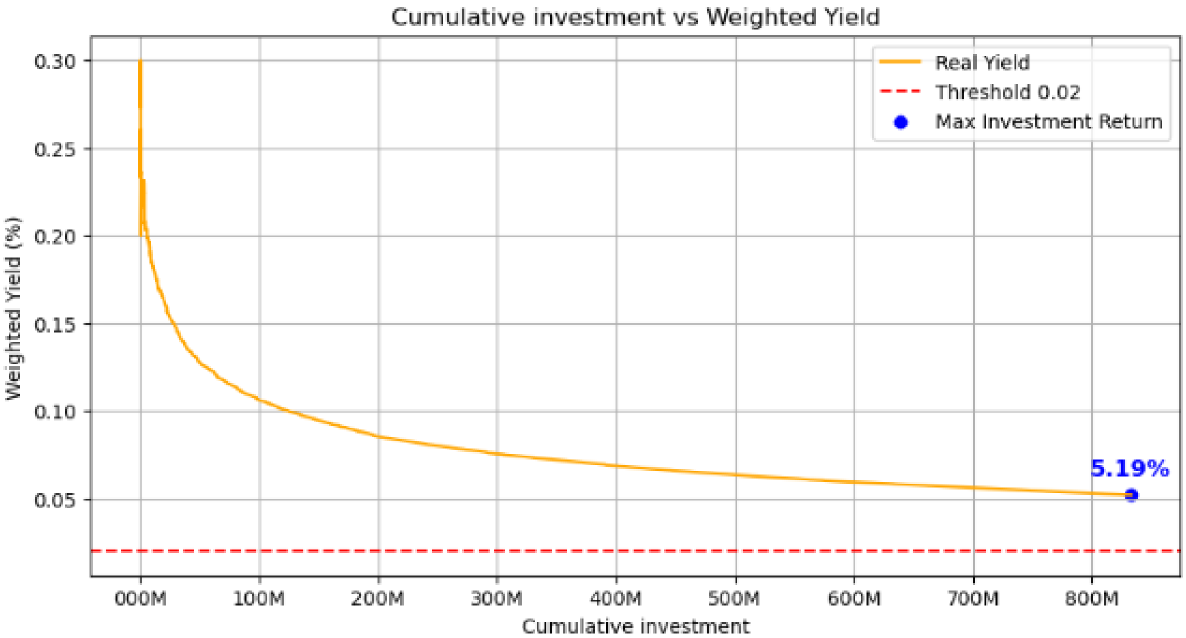



Working Paper, Step 6, Group B

| Investment Amount | Predicted Return (%) | Expected Return (%) | Number of Loans | SD of Expected Return (%) |
|---|---|---|---|---|
| $100,000,000.00 | 10.35% | 10.60% | 7826 | 0.107% |
| $200,000,000.00 | 8.76% | 8.53% | 15902 | 0.102% |
| $300,000,000.00 | 7.82% | 7.54% | 23437 | 0.096% |
| $400,000,000.00 | 7.10% | 6.86% | 31055 | 0.094% |
| $500,000,000.00 | 6.52% | 6.34% | 37985 | 0.090% |
| $600,000,000.00 | 6.07% | 5.92% | 44600 | 0.088% |
| $700,000,000.00 | 5.68% | 5.60% | 51151 | 0.085% |
| $800,000,000.00 | 5.32% | 5.29% | 57473 | 0.083% |
| $832,678,750.00 | 5.20% | 5.19% | 59378 | 0.083% |

Working Paper, Step 6, Group B

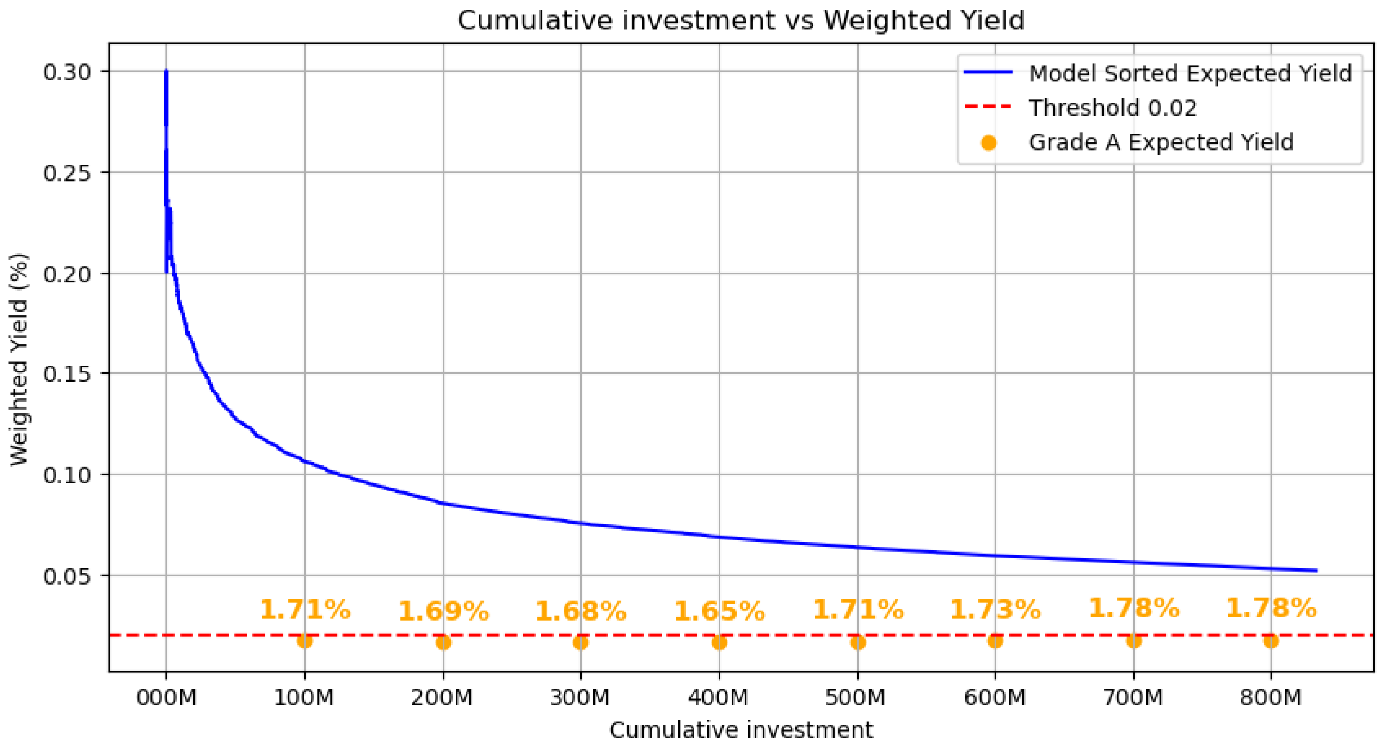
**Appendix D:**




| Investment Amount | Expected Return (%) | Number of Loans | SD of Expected Return (%) |
|---|---|---|---|
| $100,000,000.00 | 10.68% | 7826 | 0.107% |
| $200,000,000.00 | 8.53% | 15902 | 0.102% |
| $300,000,000.00 | 7.54% | 23437 | 0.096% |
| $400,000,000.00 | 6.86% | 31055 | 0.094% |
| $500,000,000.00 | 6.34% | 37985 | 0.090% |
| $600,000,000.00 | 5.92% | 44600 | 0.088% |
| $700,000,000.00 | 5.60% | 51151 | 0.085% |
| $800,000,000.00 | 5.29% | 57473 | 0.083% |
| $832,678,750.00 | 5.19% | 59378 | 0.083% |

Working Paper, Step 6, Group B

**Appendix E:**




| Investment Amount | Model Expected Return (%) | Model SD (%) | Grade A Expected Return (%) | Grade A SD (%) |
|---|---|---|---|---|
| $100,000,000.00 | 10.60% | 0.107% | 1.68% | 0.163% |
| $200,000,000.00 | 8.53% | 0.102% | 1.78% | 0.158% |
| $300,000,000.00 | 7.54% | 0.096% | 1.81% | 0.157% |
| $400,000,000.00 | 6.86% | 0.094% | 1.82% | 0.157% |
| $500,000,000.00 | 6.34% | 0.090% | 1.76% | 0.159% |
| $600,000,000.00 | 5.92% | 0.088% | 1.72% | 0.160% |
| $700,000,000.00 | 5.68% | 0.085% | 1.74% | 0.161% |
| $800,000,000.00 | 5.29% | 0.083% | 1.77% | 0.161% |
| $832,678,750.00 | 5.19% | 0.083% | 1.79% | 0.161% |