**Introduction and Survey of Relevant Literature**

Lending Club is the pioneer of peer to peer lending, a new investment option with nearly $16 billion in loans funded and 887,440 loans from the company’s inception in 2007 during the global financial crisis as defined by NBER (2010) to the end of 2015. Investors have been enticed by Lending Club’s promised returns even despite the macroeconomic situation in 2007; however, investors ultimately must assess the worthwhileness of investing in Lending Club. Asset pricing theory can be used to understand the value of claims to uncertain payments (Cochrane 2000), thus Lending Club assets can be assessed within that broader context of other forms of investment options.

Williams (1938) originally emphasized the creation of a portfolio that maximizes the discounted value of future returns, so ultimately examining Lending Club’s expected returns would be sufficient; however, Markowitz rejects William’s method and instead emphasizes expected returns as desirable and variance of returns as undesirable, incorporating another a key element in assessing a portfolio. The mean-variance frontier stems from Markowitz’s use of expected returns and variance to construct a portfolio that simultaneously maximizes the expected returns while minimizing the variance within some attainable set (1952) --- an efficient portfolio along the efficient mean-variance frontier line cannot do any better in terms of expected returns for risk. For a portfolio to yield higher expected returns, there must be more risk associated with it, so there is a trade-off aspect for the mean-variance frontier. Only two frontier returns, as a result, are necessary in constructing the mean-variance frontier (Cochrane 2000); two assets that lie on the frontier are the Risk Free Rate, which is the 3-Month Treasury Bill, and S&P 500 (Zha 2016). Examining Lending Club within the context of the mean-variance frontier is a case study utilizing the theoretical model for empirical analysis to assess the portfolio and highlight any discrepancies within the model.

Lending Club classifies borrowers using credit score, debt-to-income ratio, credit history, and income into grades with “A” being the highest and “G” being the lowest. Interest rates for borrowers are based off of the grades with 7.26% for borrowers with “A” status ranging to a maximum of 25.72% for borrowers with “G” status; however, despite the stated interest rates, loans have unique characteristics of monthly compounding interest, early payments, late payments, and defaults. This paper examines Lending Club loans in aggregate with respect to grade as a portfolio and contextualizes the portfolio with the mean-variance frontier. Furthermore, the paper claims that under empirical analysis, portfolios containing a single grade will be along the mean-variance frontier’s efficient line

The subsequent sections of the paper are organized to examine the evidence surrounding the claim that Lending Club portfolios with respect to grade are along the efficient line in the mean-variance frontier. The Data and Methodology section will detail the sources of data, provide descriptions of the data, and highlight any manipulation with the data. The Results section will present the empirical evidence from the data and construct the mean-variance frontier with Lending Club portfolios with respect to grade placed in the graph. Finally, the empirical evidence will be used in evaluating the claim of Lending Club portfolios being efficient.

**Data and Methodology**

The key data for the empirical analysis conducted is through Lending Club, utilizing the 2007-2011 and 2012-2013 data sets with 39,786 observations and 188,123 observations respectively. The data year ranges mark the starting years, and the final year chosen for analysis is 2015. The variables selected for analysis are as follows: Grade, Loan Status, Loan Term, Interest Rate (Annual), Funded Amount, and Total Payment with Issued date used for identifying separating loans that finished before 2015; these variables can be summarized below in Table 1 with the averages representing loans completed before 2015.

*Table 1: Lending Club Data Summary*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Grade** | **Total Loans** | **Completed Loans (by 2015)** | **Avg Loan Term (By 2015)** | **Avg Int Rate (By 2015)** | **Avg Funded Amount (By 2015)** | **Avg Total Payments (By 2015)** |
| **A** | **38661** | **20663** | **36.30199** | **7.466445** | **9805.885** | **10484.33** |
| **B** | **74640** | **26997** | **36.76542** | **11.69818** | **10636.35** | **11686.69** |
| **C** | **58099** | **16274** | **36.91877** | **14.54043** | **10694.94** | **11883.74** |
| **D** | **33206** | **8878** | **37.51115** | **17.02485** | **12551.97** | **14150.7** |
| **E** | **15100** | **2186** | **41.74199** | **17.02485** | **15269.84** | **17509.88** |
| **F** | **6760** | **489** | **44.58896** | **19.51211** | **15087.32** | **17211.53** |
| **G** | **1443** | **149** | **47.11409** | **21.14678** | **18996.81** | **22472.7** |
| **Overall** | **227909** | **75636** | **36.97414** | **12.04088** | **10826.09** | **11915.1** |

Time period for the empirical analysis is a limiting factor because Lending Club continues to grow in quantity of loans; however, any data used from 2016 completion dates and onwards would be at risk for a significant amount of uncertainty due to their incomplete status. Another clear limitation with respect to Lending Club’s data is the quantity of loans that are not completed by 2015 that started from 2013 and earlier: 75,636 out of 227,909 or approximately 33% of loans were completed by the time of analysis. Another complication appears when handling annualized interest rate for loans: loans have the nuances of late payments, defaults, and early payments. All are aspects that effect the annual interest rate to be different from before, and Table 2 highlights the significance of the issue within the data. As a note, classifications for Charged Off loans includ loan statuses of “Late (31-120 days)”, “In Grace Period”, and “Default,” which add to 37 loans. The total amount of charged off loans is 9,794.

*Table 2: Interest Rate Breakdown*

|  |  |  |  |
| --- | --- | --- | --- |
| **Grade** | **Full Interest Paid** | **Percent Partial Interest Rate Paid** | **Percent Charged Off** |
| **A** | **91.49688** | **8.50312** | **6.606011** |
| **B** | **80.45709** | **19.54291** | **12.07912** |
| **C** | **72.40383** | **27.59617** | **16.707632** |
| **D** | **67.45889** | **32.54111** | **20.207254** |
| **E** | **62.76304** | **37.23696** | **22.186642** |
| **F** | **56.64622** | **43.35378** | **25.97137** |
| **G** | **49.66443** | **50.33557** | **28.85906** |
| **Overall** | **79.4886** | **20.5114** | **12.94886** |

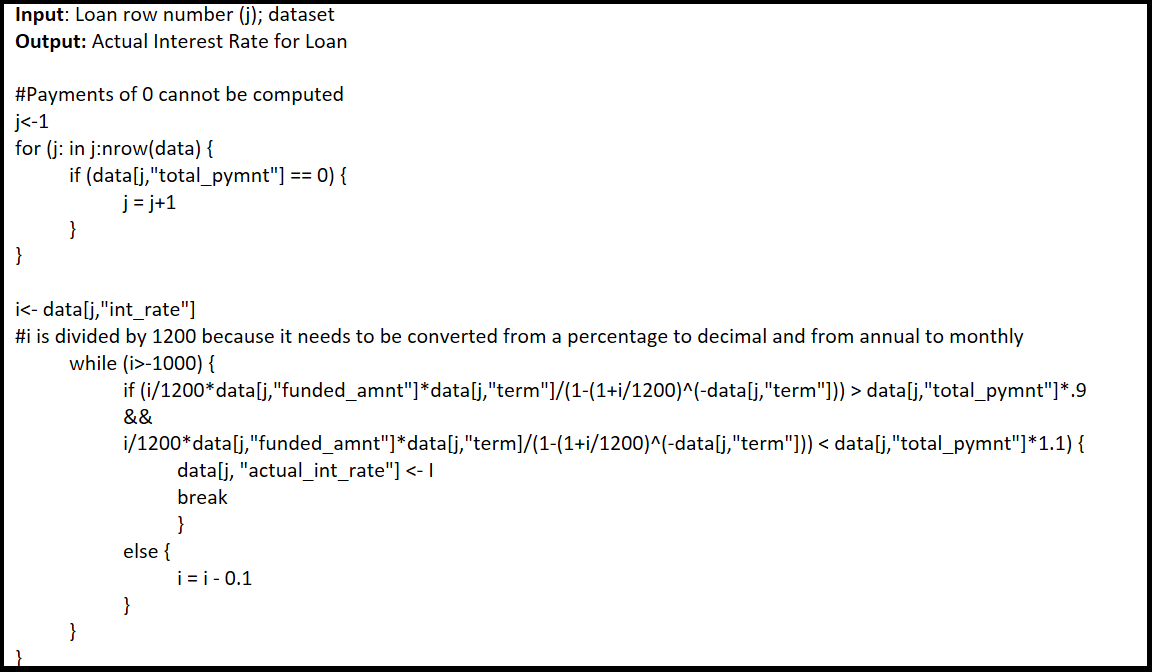
With approximately 20.5% of loans representing partial interest rates paid, the annualized interest rate provided needs to recomputed to accurately reflect annualized interest rates.

In order to calculate the actual annualized interest rate, the following formula is used with “r” as the monthly interest rate in decimal form, “P” as the funded amount on the loan, “N” as the total number of months for the loan, and total payment as “Total Loan Cost.”:

*Total Loan Cost = (r x P x N) / (1 -*

Given that Lending Club’s annualized interest rate did not reflect the nuances of the asset, an algorithm is constructed using total payments, term, and funded amount for the loan to derive an annualized interest rate. Figure 1 demonstrates the code used for the algorithm which used formula (1) (Calcunation 2016) as the basis for its construction. Key components of the algorithm are that total payments of “0” are ignored, theoretical total payments are calculated and checked to see whether they match the actual total payment within a degree of 10%, and computation begins with the actual interest rates and decrements by 0.1%.

*Figure 1: Algorithm for Determining Actual Interest Rate*



The data contains 21 loans with total payments of 0; these loans’ actual annualized interest rates are set to -250% from the lowest based off of the lowest computed interest rate of -213%. Table 3 provides a summary of the new actual interest rates computed through the algorithm.

*Table 3: Actual Interest Rate Descriptive Statistics*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Grade** | **Min Actual Int Rate** | **1st Quintile Actual Int Rate** | **Median Int Rate** | **3rd Quintile Actual Int Rate** | **Max Actual Int Rate** | **Expected Actual Int Rate** | **Standard Deviation Actual Int Rate** |
| **A** | **-209.88** | **6.54** | **7.51** | **7.9** | **9.63** | **4.61** | **13.78332** |
| **B** | **-250** | **10** | **11.14** | **12.12** | **16.39** | **6.175** | **18.90166** |
| **C** | **-250** | **12.355** | **13.99** | **15.31** | **18.61** | **6.207** | **23.13609** |
| **D** | **-250** | **12.69** | **15** | **17.77** | **21.87** | **6.603** | **25.67282** |
| **E** | **-250** | **11.35** | **16.69** | **19.03** | **22.47** | **6.771** | **26.37926** |
| **F** | **-139.59** | **8.21** | **18.17** | **19.04** | **23.76** | **6.631** | **25.13945** |
| **G** | **-105.28** | **4.44** | **19.66** | **20.9** | **24.89** | **24.89** | **27.26007** |
| **Overall** | **-250** | **7.51** | **10.78** | **13.71** | **24.89** | **5.824** | **19.96158** |

Other data used for analysis include time series of yearly averages for the US Consumer Price Index, US 3-Month Treasury Bill, and S&P 500 from 2007-2015 that can be found from the FRED. The CPI data computes US inflation rates during the time period of Lending Club using formula (2) (Mankiw 2009) to provide a baseline of real returns instead nominal returns. CPI data is incomplete for 2016, so averages from Jan. to March were used as an approximation for 2016. S&P 500 lagged values are found using the following year, and similarly to CPI, S&P 500 data is incomplete for 2016, so averages from Jan. to March were used as an approximation for 2016 . A summary of data for CPI, 3-Month Treasury Bills, and S&P 500 can be found in Table 4.

*Rate of Inflation = (𝐶𝑃*

*Table 4: Values for CPI, 3 Month T-Bill, and S&P 500*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Year** | **CPI1** | **CPI2** | **Inflation** | **3 Month T-Bill** | **S&P 500 Value** | **S&P 500 Lagged Value** | **S&P 500 Returns** |
| **2007** | **207.344** | **215.254** | **3.8149163** | **4.36** | **1477.19** | **1220.04** | **-17.40805** |
| **2008** | **215.254** | **214.565** | **-0.320009** | **1.37** | **1220.04** | **948.05** | **-22.29353** |
| **2009** | **214.565** | **218.076** | **1.636334** | **0.15** | **948.05** | **1139.97** | **20.243658** |
| **2010** | **218.076** | **224.923** | **3.1397311** | **0.14** | **1139.97** | **1267.64** | **11.199418** |
| **2011** | **224.923** | **229.596** | **2.0775999** | **0.05** | **1267.64** | **1379.35** | **8.8812439** |
| **2012** | **229.596** | **232.964** | **1.4669245** | **0.09** | **1379.35** | **1643.8** | **19.172074** |
| **2013** | **232.964** | **236.715** | **1.61012** | **0.06** | **1643.8** | **1931.38** | **17.494829** |
| **2014** | **236.715** | **236.995** | **0.1182857** | **0.03** | **1931.38** | **2061.07** | **6.714888** |
| **2015** | **236.995** | **237.911** | **0.386506** | **0.05** | **2061.07** | **1948.32** | **5.47046** |
| **2016** | **237.911** | **-** | **-** | **-** | **1948.32** | **-** | **-** |

Furthermore, inflation varies for Lending Club loans because of different issued dates, so inflation is averaged for the term of the loan as a way to accommodate the unique components of the each individual loan. All portfolios assets’ real returns are computed using formula (3) (Mankiw 2009).

*Real Return = (Nominal Return – Inflation)* ***(3)***

**Results**

*Table 5: Portfolio Assets Expected*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Asset** | **Average Stated Int Rate** | **Nominal Expected Actual Int Rate** | **Average Inflation** | **Real Expected Actual Int Rate** | **Standard Deviation of Actual Int Rate** |
| **A** | **7.466445** | **4.610442** | **1.469532** | **3.14091** | **13.71869** |
| **B** | **11.69818** | **6.175417** | **1.385317** | **4.7901** | **18.9229** |
| **C** | **14.54043** | **6.206902** | **1.412324** | **4.794577** | **23.11102** |
| **D** | **17.02485** | **6.603112** | **1.438259** | **5.164853** | **25.83946** |
| **E** | **17.02485** | **6.770709** | **1.588469** | **5.181224** | **26.20359** |
| **F** | **19.51211** | **6.631329** | **1.699662** | **4.931667** | **26.14608** |
| **G** | **21.14678** | **5.66557** | **1.742679** | **3.922891** | **28.43761** |
| **Rf** | **-** | **0.7** | **1.547815** | **-0.84781** | **0\*** |
| **S&P 500** | **-** | **4.273918** | **1.547815** | **2.726104** | **15.678514** |

*\*Assigned a standard deviation of 0 because it is the Risk Free Rate*

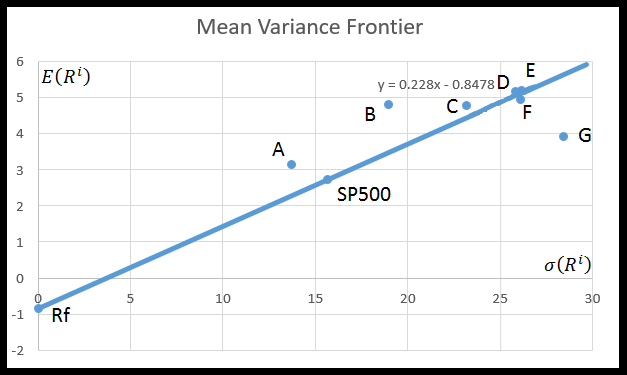
Table 5 summarizes the interest rates or returns as well as the standard deviations and for each portfolio with respect to grade, the Risk Free Rate, and S&P 500. The nominal expected actual interest rate column is derived from the algorithm in Figure 2, and the values within the column provide a nuanced view of interest rates by accounting for all relevant factors for the unique aspects of a portfolio constructed out of loans. Inflation differs when examining different starting years, so the average inflation over the term of the loan were aggregated for use in calculating the real expected actual interest rate. The real returns and standard deviation are relevant in computing the mean-variance frontier positions of the various portfolios of Lending Club and the efficient frontier line slope using Risk Free Rate and S&P 500.

Construction of the mean-variance frontier utilizes S&P 500’s average lagged real returns, which is an efficient portfolio, and the 3-Month Treasury Bill average returns, which is the default Risk Free Rate because for no risk a certain amount of returns can be expected. Using formula (4) (Cochrane 2000), the mean-variance frontier can be constructed with , as S&P 500 returns, and as the 3-month Treasury bill returns, and “m” representing the stochastic discount factor, “” as the correlation between stochastic discount factor and returns.

Using the values derived from Table 5 for Risk Free Rate and S&P 500 for expected returns and standard deviation, formula (4), and a value of -1 for “”, the efficienet line is constructed. The efficient portfolio line for the mean-variance frontier can be represented by the equation: =0.228-0.8478, and any portfolio that is efficient should be along the line.

When portfolios of Lending Club loans with respect to grade are also plotted within the mean-variance frontier graph, Figure 2 is produced.

*Figure 2: Mean-Variance Frontier 2007-2015*



The mean-variance frontier graph’s glaring discrepancy is that multiple portfolios of “A”, “B”, and “C” are completely outside the scope of the mean-variance frontier efficient line, and the fundamental issue is that the theoretical model asserts there can be no portfolios that give a higher expected return for the same amount of standard deviation than the efficient line. In addition, the results highlight several portfolios of “D”, “E”, and “F” approximately lying along the mean-variance frontier, and “G” as the only portfolio outside the scope of the mean-variance frontier. Another general trend is that portfolios consisting of grades that have borrowers with stronger characteristics have higher expected values for a given standard deviation. Overall, the mean-variance frontier construction in Figure 2 provides the empirical evidence in support of the claim that Lending Club portfolios are along the mean-variance frontier.

**Development of Key Arguments**

The portfolio of Lending Club’s aggregated loans with respect to grade in the empirical analysis produces several key supporting evidence for the portfolios being “efficient.” Literature dictates that anything outside of the attainable set cannot exist because there is no way to obtain more expected returns for the same amount of risk, and the existence of Lending Club’s portfolio outside the characterization of the mean-variance frontier’s efficient line constructed from the years of 2007 to 2015 is noteworthy. Reconciliation with the theoretical model of the mean-variance frontier can occur through one of two ways. One approach to reconcile an efficient frontier of which all portfolios must fall along or within is through reclassification of S&P 500 as a “non-efficient” portfolio, meaning it no longer falls along the mean-variance frontier. A reassessment of what is considered an efficient portfolio would logically include the portfolios of grade “A” and “B” along the mean-variance frontier efficient line. With S&P 500 formerly considered an efficient portfolio, construction of the mean-variance frontier would be at a max of the line produced by S&P 500 and the Risk Free Rate, so in constructing a new mean-variance frontier, the best known expected returns and lowest standard deviations must be used, thus including “A”, and “B”. Until another portfolio is found that satisfies both cases of maximizing expected returns and minimizing standard deviations, the positions of “A” and “B” would remain as efficient. “C,” while above the mean-variance frontier, offers a lower expected return than B while also possessing a higher standard deviation, so “C” can be considered a portfolio with idiosyncratic risk, thus “C” is not along the efficient line of the mean-variance frontier. While performance exceeding the S&P 500’s expected returns from 2007-2015 with lower standard deviation certainly highlights the strength of Lending Club portfolios of “A”, “B”, and “C”, a single time series from 2007-2015 is likely inappropriate for reconsideration of S&P 500 as an efficient portfolio. A more realistic approach is to take into account the macroeconomic factors occurring during the time period.

2007-2008 in the global macroeconomic context was a period known as the “Great Recession,” and the data from S&P 500 is a clear indication of such. The nominal returns for 2007 and 2008 from Table 4 were -17.4% and -22.3% respectively, and those values are significant enough to skew the expected returns for S&P 500 to levels that are significantly lower than the overall average. Furthermore, S&P 500 did not return back to 2007 values of 1477.19 until sometime in 2013 as shown in Table 4. The Great Recession serves as a powerful explanatory factor to keep S&P 500 as an efficient portfolio because the event ultimately skewed the expected returns; however, the overarching macroeconomic issue was not limited to an index fund such as S&P 500. Investments that are not considered hedges like gold will significantly drop in returns in a period of a recession. A reasonable assumption is that loans are not an investment that can be considered a hedge because ultimately repaying back a loan requires some sort of income which typically comes from a job. During a macroeconomic crisis, jobs are significantly more valuable and less in quantity, so paying back a portion of a loan monthly is much more difficult if an individual is unemployed. Recognizing that the portfolios with grades “A” and “B” fared better overall than S&P 500 despite facing the same macroeconomic conditions as S&P 500 strengthens the argument that the portfolios are along the efficient mean-variance frontier line. Data does not exist for a world without the Great Recession, but logically, the portfolios are less likely to default because without the overarching macroeconomic issues, individuals would be more likely to hold a job and borrow loans that can be repaid back. The expected returns and standard deviations for the portfolios would be different than what current data suggests, but the expected returns would be higher and standard deviations lower. Ultimately, “A” and “B” would have a position along the efficient mean-variance frontier line because the current expected returns and standard deviations are better than an efficient portfolio, so the readjusted increased expected returns and decreased standard deviation would fall in line with the efficient portfolio when a global macroeconomic issue does not occur.

The original claim argued that all Lending Club portfolios would be along the mean-variance frontier; however, the portfolios that could be argued being along the efficient line for the mean-variance frontier are “A” and “B”. The other portfolios of “C”, “D”, “E”, “F”, and “G” would likely not maintain a portfolio along the efficient mean-variance frontier line whether because of idiosyncratic risk or because of the information in Figure 2 that suggests that even in comparison to a skewed S&P 500, the portfolios do not exceed the efficient line.

**Conclusion**

The paper examined Lending Club loans as an aggregate with respect to each grade as a portfolio to determine the positions along the mean-variance frontier. Empirical analysis presents portfolios “A” and “B” along the mean-variance frontier, partially supporting the claim that Lending Club portfolios are within the mean-variance frontier; “C”, “D”, “E”, “F”, and “G” were all found to not be along the mean-variance frontier. Investors that are deciding on particular investment choices can utilize this information to invest in “A” and “B” portfolios and be reassured that the investment choice is optimal even when accounting for the existence of the portfolios outside mean-variance frontier in Figure 2 and broader macroeconomic conditions. A future direction for the research regarding contextualization of Lending Club portfolios along the mean-variance frontier can involve examining whether the portfolios are a hedging asset. If Lending Club is not a hedging asset, then the research conducted in this paper would be reinforced regarding the position of “A” and “B” as part of the mean-variance frontier. Another future direction for research is to perform a case study to examine Lending Club portfolios created using Lending Club’s algorithmic generation of notes to invest in. With algorithmic generation through Lending Club, is it worthwhile for the investor to use the algorithm as a source to invest in. Lending Club is a new, rich resource to investigate economic questions.

**List of Tables**

Table 1: Lending Club Data Summary.....………………………………………………………...3

Table 2: Interest Rate Breakdown.……...………………………………………………………...4

Table 3: Actual Interest Rate Descriptive Statistics………………………………………………5

Table 4: Values for CPI, 3 Month T-Bill, and S&P 500………………………………………….6

Table 5: Portfolio Assets Expected.…….………………………………………………………...7

**List of Figures**

Figure 1: Algorithm for Determining Actual Interest Rate……………………………………….5

Figure 2: Mean-Variance Frontier 2007-2015……………………………………………………8

**References**

Cochrane, John H. (2000) *Asset Pricing*, Princeton University Press, Princeton, NJ. 8,26-28

Mankiw, Gregory. (2009) *Principles of Macroeconomics*, South-Western Cengage Learning, Mason, OH. 219, 359.

Markowitz, Harry. (1952) Portfolio Selection. *Journal of Finance*, **7**: 77-91.

Williams, Burr. (1938) *The Theory of Investment Values*, 6.

Zha, Tao. (2016) *Asset Pricing*. ECON 410W. Lecture.

(2016). Inflation FAQ. *US Inflation Calculator*

(2015). Lending Club Statistics. *Lending Club*

(2016). Total Loan Cost Calculator. *CalcuNation*

(2010). US Business Cycle Expansion and Contractions. *NBER*.