CITADEL | Securities

correlation one

Testing the Efficient Market Hypothesis in Peer-to-Peer Loan Markets

Summer Invitational Data Open 2022

Ryan Lucas[1], George Chiriac[2], Varun Varanasi[3], Benny Sun[4]

[1]*Massachusetts Institute of Technology*, [2]*Princeton University*,
[3]*Yale University*, [4]*Duke University*

1 Introduction

Founded in 2006, LendingClub pioneered the peer-to-peer lending space by providing a platform for investors and potential borrowers to connect. Individuals on the platform make decisions on which loans to fund on the basis of the debtor's credit history and loan specifications.

As seen in LendingClub's public reports [1], the majority of capital on the platform originates from institutional investors. Banks and managed accounts do not lend money to individuals, but rather invest in multiple loans at once. LendingClub's reports classify loans by grade group, and present the performance of grade groups as a whole.

In this context, from an institutional investor's perspective, the platform is similar to a bond market, with baskets of loans classified by grade acting as assets. We measure the value of a basket of loans by the returns it brings. With this framework, the goal of our project is to study the efficiency of LendingClub's loan market to draw conclusions about the following:

1. **Market Stability**: Market efficiency is an integral component of stable markets which are characterized by their resilience to technological and demand shocks. Stable markets are also known to promote healthy fiscal environments in which economic decisions can be made without fear of unpredictable volatility and mis-valued assets.
2. **Market Fairness**: Efficient markets can be characterized as fair markets since under conditions of market efficiency, each participant is judged fairly on the basis of their creditworthiness. Therefore, an efficient market prevents discriminatory lending practices.

Our analysis of market efficiency is based on the Efficient Market Hypothesis (EMH). Published in 1970, Eugene Fama's Noble Prize winning work classifies three classes of market efficiency [2]:

1. **Weak Form**: The simplest of all three, the weak form of the EMH states that future asset prices cannot be predicted based on past performance.
2. **Semi-Strong Form**: The semi-strong of EMH states that asset prices spontaneously reflect all publicly available information.
3. **Strong Form**: The strong form of the EMH states that asset prices account for all public and private information. This final form of the hypothesis has the most lenient model assumptions, but is also the most difficult to test.

Our study touches upon each form of the EMH to answer the following questions:

1. Are the annual returns from LendingClub loans fully reflective of available return data?
2. Can publicly available information, such as macroeconomic indicators, be used to predict the returns on LendingClub loans?
3. Is there evidence for insider trading or other leakage of private information on the LendingClub platform?

2 Executive Summary

Using separate tests, we examine whether the LendingClub loan market is efficient under a weak, semi-strong and strong definition of EMH. We do so by incorporating data on loan principals, interest rates, terms, default rates and recouperations to construct a set of historical return-related features for all loans on the LendingClub platform.

As a test for the weak hypothesis, we analyse whether there are differences between cross-sectional returns across loan grade categories A-D. Using a group-means difference test, we find evidence for disparity in the expected returns between grade categories, which is suggestive of overvaluation for baskets of high grade loans (for example, A grade). Our group-means difference test makes several statistical assumptions such as homogeneity of variance and independent identically distributed returns. Violations of the former assumption can be caused by omitted variables, while the latter is a natural drawback of cross-sectional analysis.

To relax these assumptions, we adopt a time-series return forecasting approach for testing the semi-strong form of EMH. Here, we use publicly available macroeconomic indicators together with past return data to investigate the predictability of loan returns on the LendingClub platform. In this time series setting, we find that autoregressive return data alone is insignificant for predicting future returns.

However, we find evidence that US treasury yield data and unemployment data can be used to predict the returns for both A and D grade loans, which lie at the tail-ends of the economic spectrum for the loans studied. This means that current returns do not reflect all public information, which is indicative of LendingClub's failure to consistently assign appropriate interest rates across loans.

Finally, we seek to investigate the fairness of the LendingClub platform by performing a test for strong EMH. Our test is motivated by the landmark study of Keown (1981) [3], which found evidence for insider trading by examining the timing of market reactions to firm acquisition announcements. We apply the Power of the Pruned Exact Linear Time (PELT) test for change-point detection to examine whether there was a regime-shift in returns caused by the resignation of LendingClub's CEO. We find evidence for the occurrence of a change-point one month prior to the resignation of the CEO, but explore how this may be attributable to the IPO 17 months prior as well as the preceeding controversy that surrounded the resignation.

In studying the weak, semi-strong, and strong forms of EMH, we can draw conclusions about the stability and fairness of LendingClub's platform. In terms of market stability, the significant differences across grade basket returns exposes the LendingClub platform to possible market corrections which are detrimental for both investors and lendees. Similarly, the platform's inability to quickly adjust interest rates to macroeconomic trends creates unreasonable demands for lendees who are affected by these trends. In relation to fairness, we find evidence of the platform's reaction to the CEO resignation before the announcement of the event, suggesting insider activity.

3 Exploratory Data Analysis

To examine this data set under the lens of the efficient market hypothesis, we must understand what a rational investor would consider when analyzing the returns from a basket of loans in their portfolio. Using the 2014 to 2018 accepted loans data set, further augmented with data from 2007 onwards [4], we examine several important variables: term length, loan grade, loan amount, issue date, and loan status. Figure 1 shows descriptive statistics of the 2 million observations of accepted LendingClub loans.

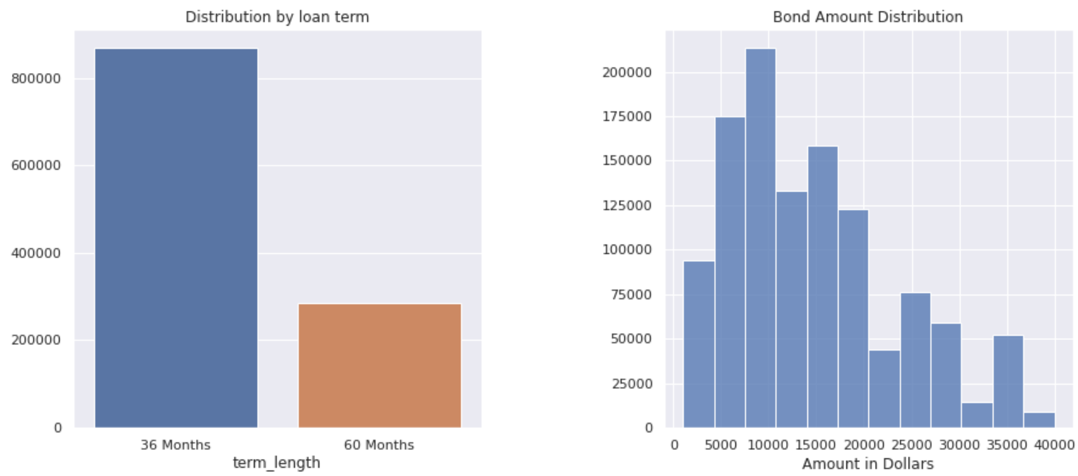

Figure 1: Count of loan term and amount on the LendingClub platform

LendingClub investors may only choose between loans with a term of 36 months or 60 months, with the shorter term occupying 76.5 % of all loans. As we discuss in section 4, the large prevalence of medium-term loans makes it less difficult to compare returns across loan maturities. LendingClub loans follow a right-skewed distribution with a mean loan amount of $14,482 and median of $12,000.

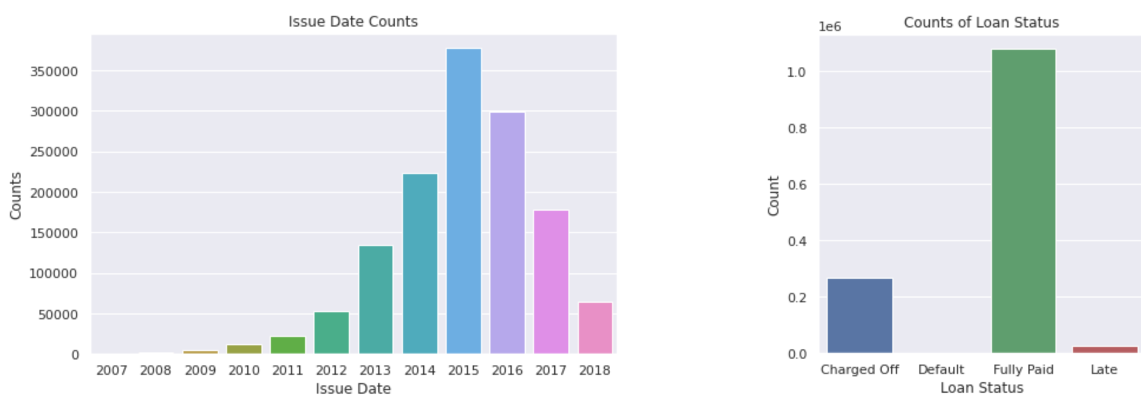

Figure 2: Count of issue date and loan status on the LendingClub platform

Figure 2 reveals further information about the distribution of our data set according to year and loan status. A majority of loans are under the Fully Paid category with the next most frequent being Charged Off. Since the Charged Off category refers to loans that have been written off as losses, we decide to count Charged Off in the default category. Ordinarily the default category includes just 40 observations. We see that from 2007 to 2015, there was a growth in observations in accepted loans. Since the beginning years (2007 to 2010) are minuscule compared to other years, we decided to ignore them in later time series analysis.

Loan grades are another major characteristic that determines the interest rate and profitability of individual loans. Accepted loans on LendingClub are given a grade from A to G representing their risk assessment from LendingClub's proprietary models. We first want to examine the distribution grades across the 2014 to 2018 data set and quantify its relationship with interest rates.

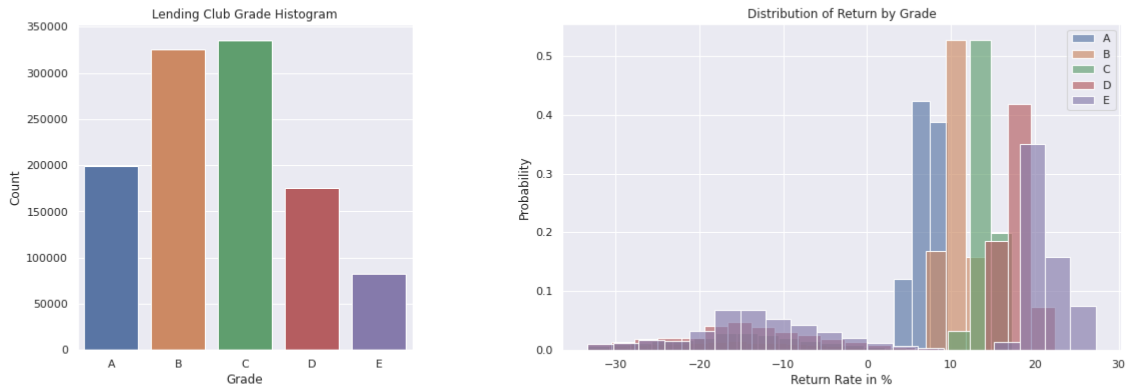

Figure 3: Analysis of Grade Classes Among LendingClub Loans

C grade loans make up the majority of LendingClub's offering with B, A, D, and E grade loans following. On average, the mean interest rate rises by 3.5 percent for each grade downgrade. Moreover, the variance of loan's interest rate also increases by 0.38 percent for each grade downgrade, showing greater uncertainty as risk increases. Interestingly, all loans have a base interest of at least 5 percent. This is due to LendingClub's proprietary models that begin with a base rate of 5.05% and add "risk-adjusted interest" determined by a loan's individual characteristics. Moreover, larger loan amounts and longer term maturity both result in smaller interest rates over time. As shown in the right graph on Figure 3, there is much greater variance in returns as grades decrease, resulting in greater volatility.

For investors, default rates are paramount for their risk assessment models. We wanted to understand the relationship between default rates and factors like time, and grade. First, we examine grade. We define default as when the loan_status column indicated "Charged Off" and not-default when it indicated "Fully Paid". We eliminate all loans with a status of "Current" due to lack of information surrounding their outcome.

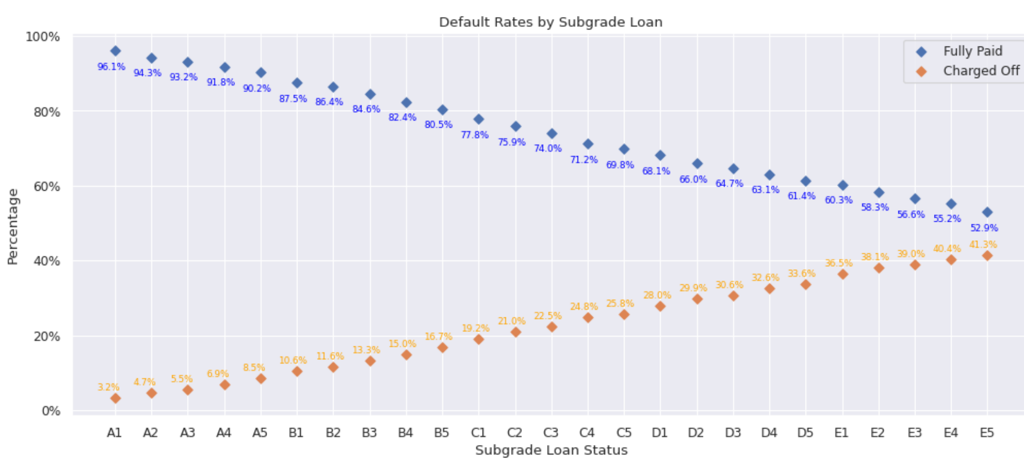

Figure 4: Default Rates over Sub-grade Loan Classes

As shown in Figure 4, there is a distinct linear relationship between sub-grade loan class and default rates. With sub-grade A1, default rates start at 3.2 percent and slowly climb to 41.3 percent at E5. An important consequence however is that interest rates as shown in Figure 3 are not linearly dependent on default rate. This is due to pressures from LendingClub's business model and promise to give out affordable loans. We next wanted to see how exactly default rates have shifted across time. Figure 5 uses an expanding mean to find this relationship.

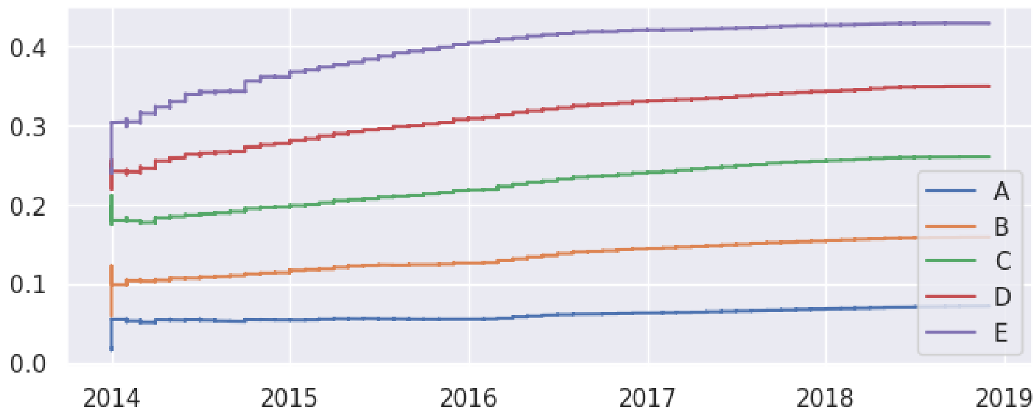

Figure 5: Expanding Mean of Default Rates over Time

Across the board, default rates have risen due to either increased business inflow, pressure to accept more loans, or leadership changes such as the 2016 CEO resignation event. This raises an important question for EMH, namely whether returns or pricing of these loans change over time. If default rates continue rising, then the market should adequately reflect these trends over time.

Overall, the implications of this exploratory data analysis surrounding default rates are such: a strong correlation between grade and default and a trend of rising default rates over time. Subsequently, factors like default rates, loan terms, amount, grade, and interest rates are important inputs to value and determine the returns of loans. Our analysis provides a detailed overlook of the loan characteristics of LendingClub.

4 Data Processing and Feature Engineering

4.1 Expected Return and Standard Deviation

With the important determinants of loan valuation in mind, we start by creating the return-related features, most importantly historical return and standard deviation. To do so, we calculate the end value of all loans in the dataset. For a particular loan i, this end value V_i equals the principal P_i plus the simple annual interest r_i earned for a maturity of m_i years. Additionally, the principal and interest are conditional on the loan not defaulting before maturity ($\mathbb{1}[\text{Default}_i = \text{False}] = D_i$). In the case that the loan does default, often there is an amount of principal + interest (RE_i) that has been or will be paid when a loan defaults, adding an extra conditional component to the valuation equation. A loan valuation incorporating the possibility of default can thus be expressed as follows:

$$V_i = P_i(1 + r_i \times m_i) \times D_i + D_i^C \times RE_i \tag{1}$$

$$\text{where}$$

$$V_i = \text{Loan Value}$$

$$P_i = \text{Loan Principal}$$

$$r_i = \text{Simple APR}$$

$$m_i = \text{Term length}$$

$$D_i = \mathbb{1}[\text{Default}_i = \text{False}]$$

$$RE_i = \text{Recovered Amount}$$

This value, as expressed in Equation 1, represents the main valuation of interest in this study. An alternative approach is to consider a discounted Net Present Value (NPV) for each interest-bearing loan. However, this would introduce several complications, since calculating NPV(RE_i) would require a precise date of recovery on both the interest and principal of the loan. Not discounting cash flows can introduce a bias depending on the maturity of the instruments and the risk-free discount rate. However, since the LendingClub loans are close in maturity, being either 3 or 5-years, this problem is alleviated to a large extent. The risk-free rate over the main period of interest (2014-2018) averages 1.42% for 3 year treasury bonds and 1.83% for 5 year treasury bonds. This small difference in the discount rates alleviates the valuation differential that may arise when comparing across the 2 maturities.

Taking V_i as the valuation, the annualised return on the loan is as follows:

$$R_i = \frac{V_i - P_i}{P_i} \div m_i$$

Thus for each loan grade category G comprised of n_G loans, we calculate a group mean and standard deviation:

$$\bar{R}_G = \mathbb{E}[R_G] = \frac{1}{n_G} \sum_{i=1}^{n_G} R_i$$

$$\sigma_G = \sqrt{\frac{\sum_{i=1}^{n_G} \left(R_i - \bar{R}_G\right)^2}{n_G - 1}}$$

We focus on $\bar{R}_G$ as the main cross-sectional variable of interest for our hypothesis testing in section 5. However, in our time series tests in section 6, we also consider a sequential monthly difference return for time-series stationarity.

Since our analysis focuses on a narrow set of factors influencing loan valuation, missing values were not a major issue. We discarded <10% of the rows. We also removed the data for E, F, and G grades, as it is very sparse. The more numerous D grade loans have similar characteristics, so we will only use these.

5 Weak-form efficiency: asset pricing in a cross-sectional setting

5.1 Theoretical Background

A central principle of the EMH states that it is impossible to identify assets that have superior returns with a comparable or superior level of risk compared to other assets [5]. This is a natural consequence of supply and demand in an efficient market. Take for instance in a loan market with two assets, loan grade 1 and loan grade 2. If loan grade 1 offers a lower return than loan grade 2, a risk-neutral investors would quickly start to demand loan grade 2, driving its interest rate down until it offered a comparable level of return to loan 1. This market adjustment is illustrated graphically in Figure 6.

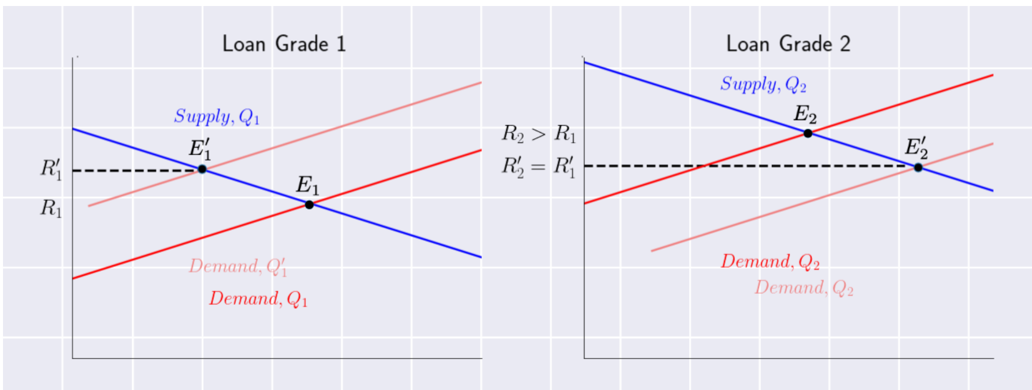

Figure 6: E_1 and E_2 show an inefficient loan market equilibrium where $R_2 > R_1$ for competing assets 1 and 2. E_1' and E_2' show how such an equilibrium fails to persist in the context of an efficient market. Eventually, demand for loan 2 will increase and demand for loan 1 will fall, pushing the returns to parity.

5.2 Testing Setup

As a first test of EMH in the LendingClub data, we examined whether certain loan grades offer higher expected returns than others. The null hypothesis under EMH is that each group mean difference is 0. If this difference were not 0 in an efficient market, a risk-neutral investor would reallocate capital to undervalued loans, pushing it back to 0.

We examine whether there are differences in the mean return across pairs of loan grade or subgrade categories.

The t-statistic for between-group differences is defined as follows:

$$t = \frac{\bar{R}_1 - \bar{R}_2}{\sigma_{\bar{R}_1,\bar{R}_2}}, \text{ where}$$

$$\sigma_{\bar{R}_1,\bar{R}_2} = \sqrt{\frac{\sigma_p^2}{n_1} + \frac{\sigma_p^2}{n_2}} = \text{Pooled standard error}$$

$$\sigma_p^2 = \frac{(n_1 - 1)\,\sigma_1^2 + (n_2 - 1)\,\sigma_2^2}{n_1 + n_2 - 2} = \text{Pooled variance}$$

$$n_1 = \text{Sample size of group 1}$$

$$n_1 = \text{Sample size of group 2}$$

This gives the following two-sided t-test for group mean differences:

$$H_0 : \bar{R}_1 - \bar{R}_2 = 0, \quad H_A : \bar{R}_1 - \bar{R}_2 \neq 0$$

5.3 Assumptions

5.3.1 Investor Risk-Neutrality

By only comparing the expected value of competing loans, our analysis assumes that investors are risk-neutral, meaning that they are indifferent to the volatility of pay-offs. However, while investors with limited capital may be risk averse and biased towards lower-variance payoffs, it is important to note that over 80 percent of investors on the LendingClub platform come from an institutional or managed account background [9]. Given their large amount of capital compared to for instance retail investors, it is more appropriate to assume risk-neutrality. Furthermore, relaxing the assumption of risk-neutrality is difficult in practice. Namely, common risk metrics like Sharpe ratios assume that returns are normally distributed. However, loan returns follow a bi-modal distribution with one mode arising in a default state and another for settled loans.

5.3.2 Independent and Identically Distributed (i.i.d) returns

Since this test is cross-sectional, comparing sample means across grade categories, it implicitly assumes that all returns for a particular group are sampled from the same distribution. That is, for a group G, we assume $R_G \sim P_G(\bar{R}_G, \sigma_G^2)$ for all samples. This applies for both the reference group and the comparison group. However, in section 6 and beyond, we start to relax this assumption by performing tests on the monthly difference of the return in a sequential time series.

5.3.3 Homogoneity of Variance

The homogoneity of variance assumption applies to all independent sample t-tests, stating that all comparison groups must have the same variance.[ref stats sols]. This assumption may be violated in our case, since lower grade groups have higher standard deviations. However, independent sample t-tests utilize t-statistics, which are robust to violations of the assumption as long as group sizes are close to equal. It was also for this reason that we excluded loan grades E, F and G, which have higher variance as well as being less numerous. Differences in variance between categories can also be caused by confounding variables, some of which which we control for section 6.

5.4 Results

We run the group means difference tests on all combinations of grades A-D. For each group combination, we report mean differences, confidence bounds and p-values in Table 1. In every case, we find the reference group (that with a better grade) to have a lower expected return than the control group. These results are significant at all levels for reference group A, suggesting that A grade loans offer a lower EV than all other grade categories. This pattern appears to follow when examining lower grade reference groups, although significance is mixed. This suggests that investors with risk-neutral preferences are acting in an irrationally risk-averse fashion. In the LendingClub market, investors prefer to invest in higher grade reference groups, despite those groups having a lower expected return. This suggests a 'factor premium' for A grade loans, which is a common feature of market inefficiency [9].

Group			**Confidence Interval**	
Reference (1)	Control (2)	$\bar{R}_1 - \bar{R}_2$	Lower	Upper
A	B	-1.1281***	-1.528	-0.728
	C	-1.5858***	-2.105	-1.067
	D	-1.9071***	-2.537	-1.277
B	C	-0.4577*	-1.052	0.137
	D	-0.7789**	-1.472	-0.085
C	D	-0.3213	-1.09	0.447

Table 1: Group means difference tests on mean return across grade categories.
Significance codes: 1% Level: ***, 5% Level: **, 10% Level: *

5.5 Analysis

Under the conditions listed in section 5.3, we find that the LendingClub market is inefficient by the weak EMH. However, before making any claim on market efficiency, we must evaluate the validity of our assumptions, specifically 5.3.2 and 5.3.3. We aim to test whether our data is independent and identically distributed and whether variance across grades is homogeneous by determining if the returns for each loan grade are identically distributed.

5.5.1 Kolmogorov Smirnov Test

To evaluate our assumptions, we look to the Kolmogorov Smirnov test (KS Test), a nonparametric exact test of equality. The two-sided KS test determines whether the samples originate from the same probability distribution with no assumptions about their distributions. The test evaluates the maximum distance between two distributions against the Kolmogorov distribution; however, since the true distributions are unknown, the KS test estimates them via empirical distribution functions. Empirical distribution functions are cumulative distribution functions that are constructed as step functions from the data. Mathematically, these functions are defined as:

$$F_n(x) = \frac{\#elements \leq x}{n}$$

For two empirical distributions functions, $F_{1,n}$ and $F_{2,m}$ with n and m observations, respectively, the KS test uses the following null and alternative hypothesis:

$$H_0 : F_{1,n} = F_{2,m}, \quad H_A : F_{1,n} \neq F_{2,m}$$

The test statistic is calculated via:

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)|$$

and for large samples, it rejects with significance level α when

$$D_{n,m} \geq c(\alpha) \cdot \sqrt{\frac{n+m}{nm}} \quad \text{where} \quad c(\alpha) = \sqrt{-ln(\frac{\alpha}{2}) \cdot \frac{1}{2}}$$

5.5.2 Results

We run the KS Test on grades A-D, excluding E, F and G as before. For each test, we find a test statistics, $D_{n,m}$, and its associated p-value as detailed below. To reduce type I error in our multiple hypothesis testing schema, we adjusted our resulting p-values via the Bonferroni method. This p-value correction method simply adjusts the significance level α to $\frac{\alpha}{n}$ for n observations. After doing so, we found that each grade pair rejected the null-hypothesis and thus, was drawn from a unique distribution. With this in mind, we cannot validate our conclusions from the mean deviation test and cannot disprove weak market efficiency.

	$D_{n,m}$	p-value
(A, B)	0.766	0.000
(A, C)	0.757	0.000
(A, D)	0.677	0.000
(A, E)	0.600	0.000
(B, C)	0.636	0.000
(B, D)	0.675	0.000
(B, E)	0.599	0.000
(C, D)	0.561	0.000
(C, E)	0.585	0.000
(D, E)	0.359	0.000

Table 2: KS test statistics with Bonferroni-adjusted p-values

Figure 7: Distribution of annual returns by grade

6 Semi-strong efficiency: return predictability with public information

As shown in in section 5.5 we cannot disprove the Weak Efficient Market Hypothesis due to modelling assumptions 5.3.2 and 5.3.3. For a more realistic analysis, we now relax these assumptions and test the Semi-Strong Efficient Market Hypothesis instead.

The Semi-Strong Efficient Market Hypothesis states that asset returns reflect all public information, so such observed information cannot be used to forecast future returns. To test this for peer-to-peer loan annual returns, we will use a time series model and check if predictors exist for any of the loan grades.

6.1 Theoretical Background

An equivalent way of saying that no public information can help forecast future asset returns is that the latter are mean independent of all current information. Therefore, a way to phrase the Semi-Strong EMH mathematically is:

$$\mathbb{E}[Y_{t+1}|\Omega_t] = \mathbb{E}[Y_{t+1}],$$

where Y_{t+1} are asset returns at time $t+1$, with t being present time, and Ω_t is all publicly available information at present time.

If we view this in the framework of a linear regression,

$$Y_t = \alpha + \vec{\beta}\Omega_{t-1} + u_t,$$

then the condition above implies that $\vec{\beta} = \vec{0}$. Therefore, to disprove the hypothesis, it is enough to find a subset of Ω_{t-1} for which the regression coefficients are statistically significant. We will test this for time series data of LendingClub loan annual returns in the next section.

6.2 Testing Setup

Consider an Autoregressive Distributed Lag (ADL) model of the annual return $Y_{t,G}$ from a basket of loans of grade G, issued at month t. If the term of the loans is Δt, an investor only has returns information from settled loans issued more than Δt months ago. Thus, a realistic regression equation is:

$$Y_{t,G} = \alpha_0 + \sum_{i=1}^{p} \beta_i Y_{t-\Delta t-i,G} + \sum_{j=1}^{q} \delta_{1j} X_{1,t-j} + ... + \sum_{j=1}^{q} \delta_{kj} X_{k,t-j} + u_{t,G},$$

where $X_{1,t}, ..., X_{k,t}$ are the time series of the subset of public information we use for prediction, and $u_{t,G}$ is the error term.

Then, under the EMH we must have $\beta_i = \delta_{1j} = ... = \delta_{kj} = 0$ for all i and j, as explained in the last section. This is what we will test. For our model to be rigorous, the four ADL assumptions must hold:

- Exogeneity: $\mathbb{E}[u_t|Y_{t-\Delta t-1,G}, Y_{t-\Delta t-2,G}, ..., X_{1,t-1}, ..., X_{k,t-1}, ...] = 0$
- Stationarity and weak dependence for $(Y_{t,G}, X_{1,t}, ..., X_{k,t})$
- No large outliers: $0 < \mathbb{E}[Y_{t,G}^4] < \infty$, $0 < \mathbb{E}[X_{1,t}^4] < \infty$, ..., $0 < \mathbb{E}[X_{k,t}^4] < \infty$
- No perfect multicollinearity

In the following section, we will use time series data of annual returns by grade basket along with time series data of macroeconomic trends to construct an ADL model that respects these assumptions. This model will allow us to correctly test the semi-strong Efficient Market Hypothesis.

6.3 Building the Model

To construct the time series for a grade basket, we take the mean over all individual annual returns from that grade, at every month:

$$Y_{t,G} = \frac{1}{n_{t,G}} \sum_{i=1}^{n_{t,G}} R_{t,i}$$

where $n_{t,G}$ is the total number of loans issued at month t in grade G, and $R_{t,i}$ is the annual return for an individual loan issued at month t in that grade.

Data from the first 24 months is sparse, causing outliers in $Y_{t,G}$, so we remove it. There are also too few data points for the 60-month loans, so we expect high standard errors for the regression parameters. Therefore, we restrict to 36-month loans. To satisfy the stationarity assumption, we work with the monthly differences in annual returns instead of the annual returns themselves:

$$Y_{t,G} \rightarrow Y_{t,G} - Y_{t-1,G}$$

This removes the trend in the time series, as shown in the graphs below.

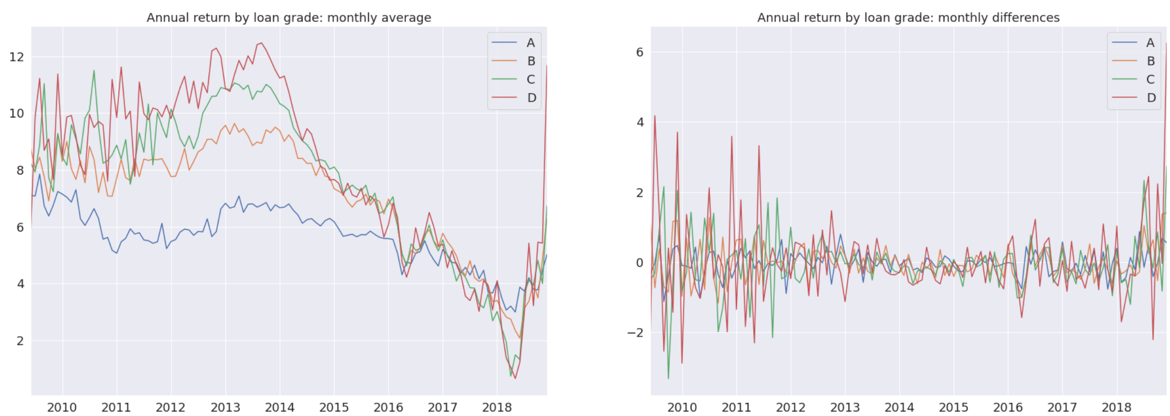

Figure 8: Time series before (left) and after (right) taking monthly differences

Processing the data by taking log differences or percent changes can also remove trends in time series. We used the Augmented Dickey Fuller (ADF) test to see which processing method achieves the best stationarity. Monthly differences is the most effective, as we can see from the table below.

Grade	DF stat (differences)	DF stat (pct changes)	DF stat (log differences)	lags
A	-5.0123***	-4.4916***	-4.5685***	4
B	-3.7305***	-2.1774	-2.5662	4
C	-4.2331***	-4.0616***	-4.8197***	4
D	-2.8893**	-2.8337*	-4.2432***	4

Table 3: Augmented Dickey Fuller test for unit roots
Significance codes: 1%: ***, 5%: **, 10%: *

Now, it is time to construct the time series $X_{1,t}, ..., X_{k,t}$ of past public information that we will use for forecasting. We used the following three macroeconomic trends, which we suspect are closely related to loan returns:

- 3-Year US Treasury Yield, $DGS3$. It is reasonable to expect that this risk-free rate sets a lower bound on the interest rate for peer-to-peer loans. The A grade basket is especially affected, as it has the lowest interest rate among the four baskets.

- Cboe Volatility Index, $VIXCLS$. This index quantifies the market's expectations for near-term price changes of the Standard and Poor's 500 Index. It forecasts economic volatility which may affect people's ability to pay their debts.
- US Unemployment Rate, $UNRATE$. This is an indicator of the financial health of low-income families. We expect it to be correlated with the C and D grade baskets.

Similar to before, we wish to remove the trends in these time series for stationarity. Furthermore, we want them to have a similar scale in order to achieve better regression results. We do this by taking the log differences:

$$DGS3_t \rightarrow ln(DGS3_t) - ln(DGS3_{t-1})$$
$$VIXCLS_t \rightarrow ln(VIXCLS_t) - ln(VIXCLS_{t-1})$$
$$UNRATE_t \rightarrow ln(UNRATE_t) - ln(UNRATE_{t-1})$$

As we can see from Figure 9, this accomplishes the desired stationarity.

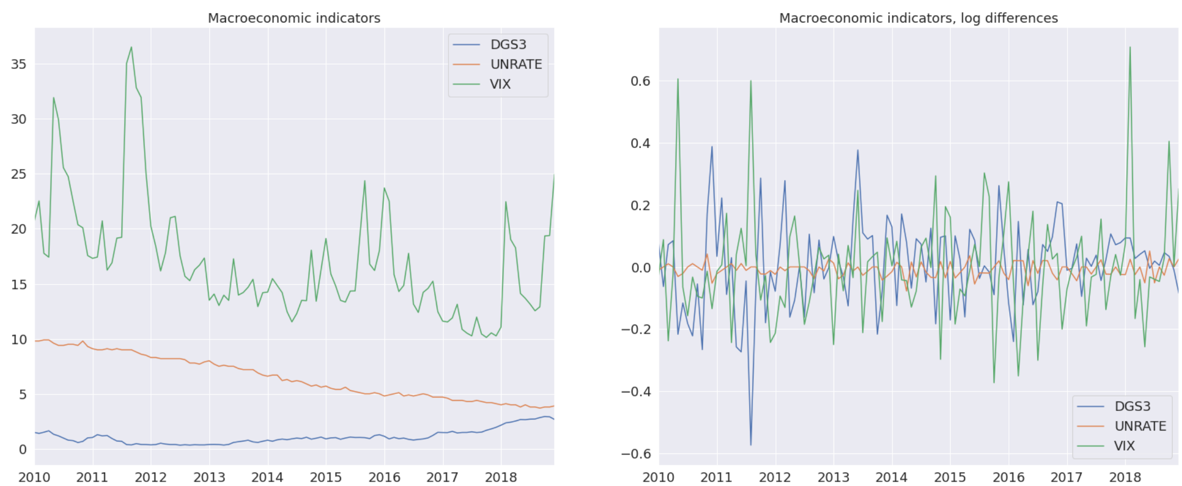

Figure 9: Macro indicators before (left) and after (right) taking log differences

We see that log differences and percent changes achieve a similar level of stationarity by running the ADF test. Monthly differences are not considered, as they would not scale the time series equally. Results are in the table below.

Regressor	DF stat (pct changes)	DF stat (log differences)	lags
DGS3	-4.7777***	-4.3034***	8
VIX	-4.2862***	-4.4853***	8
UNRATE	-3.5653***	-3.5583***	8

Augmented Dickey Fuller test for unit roots
Significance codes: 1%: ***, 5%: **, 10%: *

We choose log differences over percent changes because they better linearize the data, which helps validate the exogeneity assumption.

The last step in creating the regression model is deciding how many lag lengths to use, that is, determining the optimal pair (p, q). To do this, we will select the (p, q) pair that minimizes the Akaike Information Criterion (AIC) for each loan grade. This is a reasonable option, as the AIC is approximately equal to the logarithm of the Root Mean Squared Forecast Error (RMSFE). We also have the constraint $p <= 4$ and $q <= 8$. These are the maximum lags for which the time series can be considered stationary, as seen in the ADF tables above. We listed the pairs (p, q) that minimize AIC below.

Grade	p	q	AIC
A	2	5	51.33
B	4	4	84.07
C	4	1	170.5
D	4	1	258.3

Table 4: Optimal lag lengths p and q

With the regressors and lag lengths we chose, the no multicollinearity assumption is also satisfied, as the heatmap below shows.

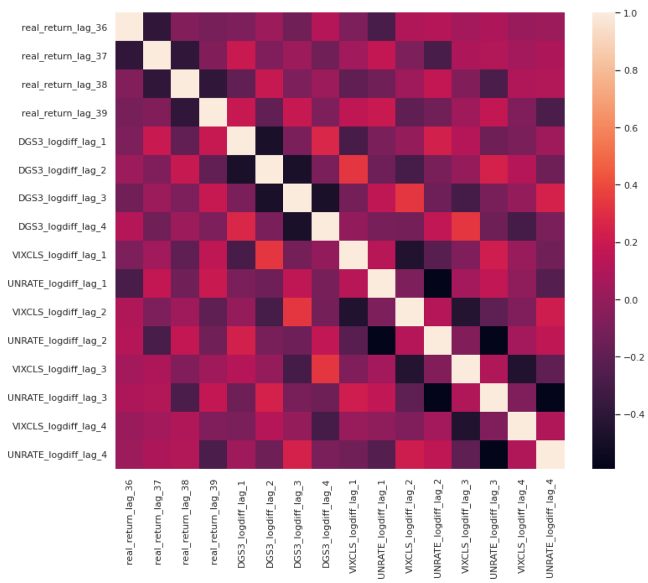

Figure 10: Correlation heatmap of regression variables

6.4 Results

We firstly fit several AR(2) models of lag return, which are reported in Table 5. We also run the ADL models separately for each grade, with the parameters we had in Table 4. For the A basket, this includes 2 lags of the difference in annual returns and 5 lags for the macroeconomic factors; for the B basket, 4 lags for each regressor; and for the C and D baskets, 4 lags for the difference in returns and 1 lag for the macroeconomic factors.

Across the 4 loan grades, we find that there are several lagged variables that correlate with return. For grade C, we see that previous lagged returns of 36 positively correlated with future return at a 5 percent significant level. This could be due to reactions about profitability opportunity from Grade C performance as they reached the end of their term.

On the individual grade-level models for macroeconomic variables, we see that lagged unemployment rate was negatively correlated with returns. We believe these results are intuitive in inefficient markets, as interest rates fail to account for default rises and declining interest rates during worsened economic condition. Treasury rates in grade A and D were also negatively correlated. These observations indicate that LendingClub's interest rates did not rise fast enough to match the rising risk free rate, resulting in less actualized returns. These factors lend credence to the alternative hypothesis that the LendingClub market is not efficient.

Grade		ϕ	Lower	Upper
A	**Intercept**	-0.0046	-0.080	0.070
	Lag 36	0.0490	-0.204	0.302
	Lag 37	0.1909	-0.051	0.433
B	**Intercept**	-0.0404	-0.132	0.051
	Lag 36	-0.0887	-0.245	0.067
	Lag 37	0.1385*	-0.013	0.290
C	**Intercept**	-0.0308	-0.184	0.122
	Lag 36	0.1213**	0.006	0.237
	Lag 37	-0.0159	-0.154	0.122
D	**Intercept**	0.0079	-0.282	0.298
	Lag 36	-0.0901	-0.274	0.094
	Lag 37	-0.0878	-0.266	0.091

Table 5: Autoregressive models fit on 4 grade classes of loans.
Significance codes: 1% Level: ***, 5% Level: **, 10% Level: *.

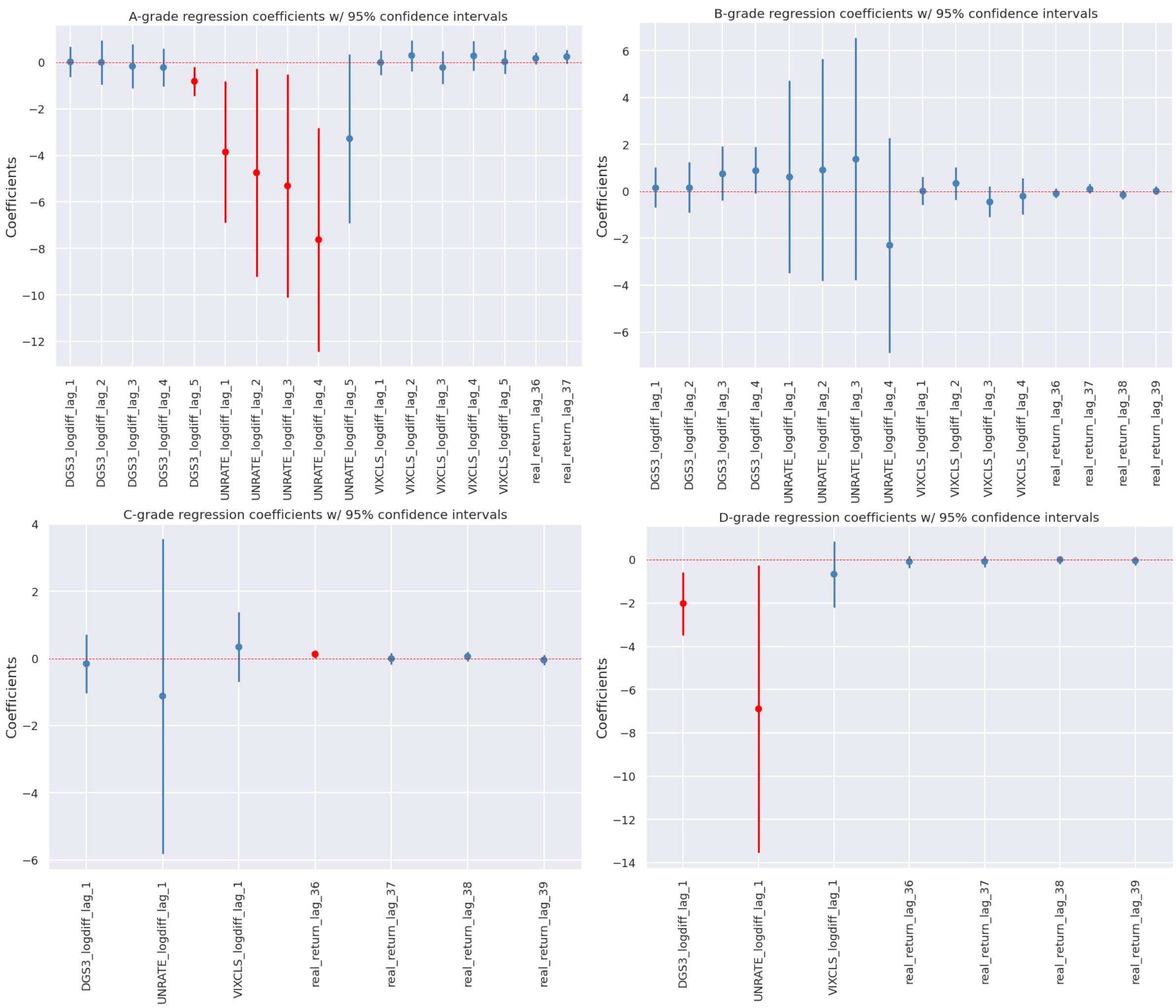

Figure 12: Autoregressive distributed lag (ADL) models fit on 4 grade classes of loans. Exogenous variables include the three-year treasury yield (DGS3), unemployment rate (UNRATE) and VIX Index (VIXCLS). Red lines indicate significance.

7 Strong Form Efficiency

The analysis from section 5 and section 6 suffices to show that the LendingClub market is neither weakly nor semi-strong efficient. Nevertheless, to investigate the fairness and ethics of the LendingClub market, we perform a final test for strong efficiency.

7.1 Theoretical Background

Where the semi-strong form of the EMH states that market prices are reflective of all public information, the strong form EMH states that markets are reflective of all public *and* private information. If the strong EMH holds, then returns will be influenced by private insider information, before that information has ever been made publicly available.

Strong market efficiency was famously uncovered in a landmark study by Keown (1981) [3], which found that news about acquisitions was reflected in market prices before ever being publicly announced. Keown's conclusion was that trader's of large public companies possessed insider information, which represented a clear violation of Rule 10b-5 of the Security Exchange Act of 1934 and evidence of strong-form EMH. Traders and institutional investors were able to profit on insider information, where retail investors relied on public sources.

7.2 Testing Setup

Renaud Laplanche, CEO and co-founder of LendingClub, famously resigned as CEO on 9 May 2016. In an analog to Keown's study, we test whether his resignation was reflected in loan returns before having been announced publicly. To do so, we use the Power of the Pruned Exact Linear Time (PELT) test for multiple change-point detection in time series. The PELT algorithm aims to partition a time series into homogenous blocks. Specifically, PELT tries to minimize a $\mathcal{C}$ cost function of within-block similarity, while penalising on the number of change points. For each loan grade, we minimize:

$$\arg\min_{\tau} \sum_{i=1}^{m+1} \left[\mathcal{C}\left(Y_{G,(\tau_{i-1}+1):\tau_i}\right)\right] + \beta m$$

where

$$\mathcal{C}\left(Y_{G,(\tau_{i-1}+1):\tau_i}\right) = -\max_{\theta} \sum_{i=\tau_{i-1}+1}^{\tau_i} \log f\left(Y_{G,i} \mid \theta\right)$$

$$\tau = \text{Change point locations}$$

$$\theta = \text{Maximum Likelihood Estimate}$$

$$m = \text{Number of change-points}$$

$$\beta = \text{Penalisation Parameter}$$

We set β = 2.5, which restricts the number of change points to at most 2 and is in line with the values used by [7]&[8] and we run the test for groups A-D testing the following null hypothesis:

$$H_0 : \text{No change point 3 months before CEO resignation on 9 May 2016}$$

$$H_A : \text{Change point occurs before and within 3 months of 9 May 2016}$$

7.3 Assumptions

7.3.1 No public information about the resignation was available before the date of resignation

Our change point study implicitly assumes that there was no information publicly available on the resignation of the CEO prior to the 9th May 2016. However, from plotting the Google search history for LendingClub during 2016 in Figure 13, it becomes clear that public information may have been known about the reason for the resignation as early as 1 May 2016. News report at the time also place initial controversy at this date.

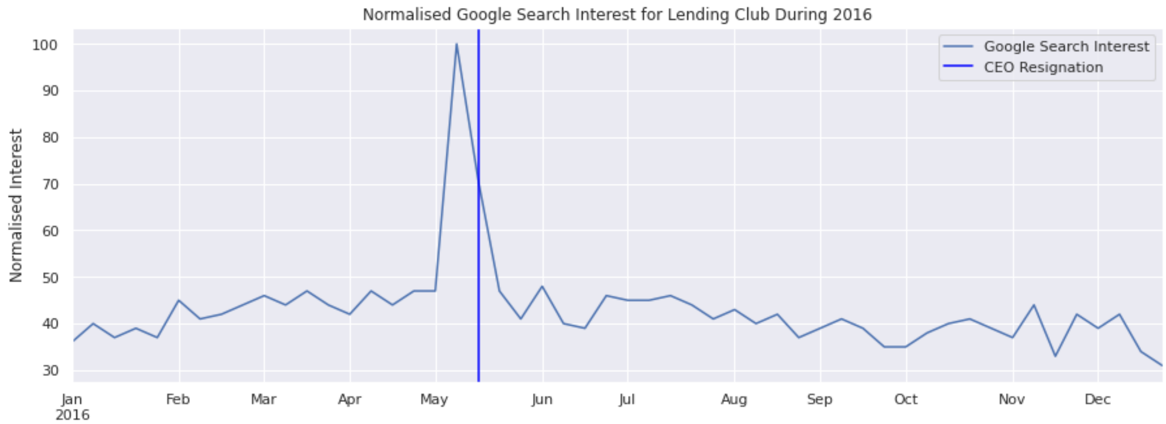

Figure 13: Google search interest for the LendingClub platform during 2016

It is for this reason that we present a modified null and alternative hypothesis:

$$H_0' : \text{No change point occurs 3 months before 1 May 2016}$$
$$H_A' : \text{Change point occurs before and within 3 months of 1 May 2016}$$

7.4 Results

The change points chosen by PELT at the 5% level of significance are reported for A-D grade returns in Figure 14. The figure marks suggestive evidence of an early reaction for A and C grade loans, with each of their latter change points being marked at April 2016. We caveat such results with the fact that the LendingClub IPO filing occurred approximately 16 months prior to the change point that PELT detected. Notwithstanding the potential bias introduced by the IPO, we find evidence to reject our null hypothesis H_0' in favour of the alternative hypothesis. The conclusion is that private information about the controversy and resignation may have been known by investors on the LendingClub platform.

Figure 14: Change-points chosen by the Pruned Exact Linear Time (PELT) algorithm for A-D grade returns on the LendingClub marketplace

8 Conclusion

Through analyzing the Efficient Market Hypothesis, we are able to show that LendingClub markets are inefficient under the three interpretations. The weak interpretation mentions that current prices only reflect past movement and cannot be predicted. Under the weak lens, we use a group means difference test alongside the Kolmogorov Smirnov Test to show that poorer loan grades offer systematically greater expected returns, which rational investors have not fully taken advantage of.

The semi-strong interpretation argues that prices reflect all public information in the market. Under the semi-strong lens, we build an ADL model with VIX, treasury yield, and the unemployment rate with statistically significant regression co-efficients. The strong interpretation builds on the semi-strong by also mentioning that private information as well is reflected in current prices. Under the strong-lens, we conclude using the PELT test to analyze the 2016 CEO resignation.

In conclusion, the significant differences across grade basket returns indicate detrimental market instability for both investors and lendees. The market's inertia to macroeconomic trends creates unhealthy fiscal demands for lendees. The market's preemptive response to the CEO's resignation suggests insider activity.

Given this result, there are various future research directions one can go. First, because LendingClub markets are inefficient, investors can capitalize on this inefficiency through methods like statistical arbitrage, data analysis, and prediction models. Future research can go into the specifics of such strategies. Secondly, our analysis assumes risk neutrality for all investors, as we could not find any risk-adjusted return metrics that could work under a bi-modal distribution. Future analysis could therefore find a risk-measure that works under this unique situation. Overall, efficiency is an important factor in producing healthy markets, as it enables stability from unpredictable volatility and fosters fairness by providing equal predictive power. By showing how LendingClub markets are inefficient, we can provide a starting point on how to improve the peer-to-peer loan model.

9 References

[1] Lending Club First Quarter 2016 Results. (2016, May 6). Retrieved July 24, 2022, from https://s24.q4cdn.com/758918714/files/doc_presentations/1Q16-Presentation.pdf.

[2] Fama, E. F. (1970). Efficient Capital Markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383. https://doi.org/10.2307/2325486

[3] Keown, Arthur., amp; PINKERTON, J. O. H. N. M. (1981). Merger announcements and insider trading activity: An empirical investigation. The Journal of Finance, 36(4), 855–869. https://doi.org/10.1111/j.1540-6261.1981.tb04888.x

[4] George, N. (2022). All Lending Club loan data. https://www.kaggle.com/datasets/wordsforthewise/lending-club

[5] Malkiel, B. C. (2003). The efficient market hypothesis and its critics. CFA Digest, 33(4), 40–41. https://doi.org/10.2469/dig.v33.n4.1367

[6] FRED Unemployment, VIX, amp; Treasury Yield Data - Federal Reserve Economic Data. (2022). https://fred.stlouisfed.org/

[7] Killick, R., Fearnhead, P., amp; Eckley, I. A. (2012). Optimal detection o7 changepoints with a linear computational cost. Journal of the American Statistical Association, 107(500), 1590–1598. https://doi.org/10.1080/01621459.2012.737745

[8] Dorcas Wambui, G. (2015). The power of the pruned exact linear time(pelt) test in multiple Changepoint detection. American Journal of Theoretical and Applied Statistics, 4(6), 581. https://doi.org/10.11648/j.ajtas.20150406.30

[9] Lending Club Fourth Quarter 2016 Results. (2016, May 6). Retrieved July 24, 2022, from https://s24.q4cdn.com/758918714/files/doc_presentations/4Q16-Presentation.pdf.

[10] Suh, D. (2009). Stock returns, risk factor loadings, and model predictions: A test of the CAPM and the Fama -French 3 -factor model . Graduate Theses, Dissertations, and Problem Reports.

[11] Jackson, B. (2004). An Algorithm for Optimal Partitioning of Data on an Interval. ArXiv.