

Portfolio return optimization under an endogenous investment set: evidence from P2P lending

Journal of
Derivatives and
Quantitative
Studies: 선물연구

Jaesung James Park

Korea Small Business Institute, Seoul, South Korea

Jonghyun Moon

State University of New York at Stony Brook, Stony Brook, New York, USA

HyeIn Yu

Graduate School of Information, Yonsei University, Seoul, South Korea

Hyomin Lee

School of Civil, Environmental and Architectural Engineering, Korea University, Seoul, South Korea

Hyebin Jeong

Department of Statistics, Kyungpook National University, Daegu, South Korea

Minseo Choi

Department of Mathematical Sciences, Seoul National University, Seoul, South Korea, and

Hyeseong Hwang

Department of Quantitative Risk Management, Yonsei University, Seoul, South Korea

Received 26 February 2026
Revised 18 March 2026
28 March 2026
Accepted 30 March 2026

Abstract

We develop a Sharpe ratio-maximizing decision framework for environments in which binary approval rules endogenously determine the investable set. In conventional portfolio theory, investors optimize weights within an exogenously given asset universe; in credit screening, however, the approval threshold τ simultaneously governs both the number and the quality of admitted assets. This endogeneity renders diversification a function of τ , causing idiosyncratic risk to re-emerge at the portfolio level. Because each loan generates only a single realized outcome, its idiosyncratic variance cannot be directly estimated. We compute expected returns using decile-specific conditional returns of predicted default probabilities and construct a risk proxy from the cross-sectional variance within expected-return intervals. Using LendingClub data from 2007 to 2020 with 50 repeated cross-validation runs, we obtain an average Sharpe ratio of 0.0919 and an average approval rate of 26.9% under the optimal threshold τ^* . The Sharpe ratio exhibits an inverted U-shaped pattern with respect to τ , confirming a trade-off in which stricter screening improves the quality of the approved loan pool while concentrating it on fewer loans. Whereas indiscriminate approval of all loans yields a negative Sharpe ratio, the proposed model achieves a positive one. A portfolio that underperforms Treasury securities when evaluated solely on expected returns can outperform them once risk information is incorporated. The findings demonstrate that when the investment set is endogenously determined by the screening rule, explicit incorporation of asset-level risk information – rather than expected return alone – is essential for investment decisions.

Keywords Idiosyncratic risk, Sharpe ratio, P2P lending, Credit screening, Endogenous investment set

Paper type Research article

© Jaesung James Park, Jonghyun Moon, HyeIn Yu, Hyomin Lee, Hyebin Jeong, Minseo Choi and Hyeseong Hwang. Published in *Journal of Derivatives and Quantitative Studies: 선물연구*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/>



Journal of Derivatives and Quantitative
Studies: 선물연구
Emerald Publishing Limited
e-ISSN: 2713-6647
p-ISSN: 1229-988X
DOI 10.1108/JDQS-02-2026-0013

1. Introduction

We develop a Sharpe ratio–maximizing decision framework in which binary approval rules endogenously determine the investment opportunity set. The empirical setting is a P2P (peer-to-peer) lending portfolio from LendingClub. In conventional portfolio theory, investors optimize portfolio weights within an exogenously given asset universe. In contrast, in a credit selection environment, the approval threshold τ determines the set of assets admitted to the portfolio. This endogeneity renders diversification a function of τ , creating conditions under which idiosyncratic risk—typically assumed away in standard theory—re-emerges at the portfolio level.

Borrower characteristics relevant to lending decisions—such as credit grade, annual income, debt-to-income ratio, and loan purpose—are ultimately summarized into a single scalar, the expected return, through default probability estimation. Even when only loan origination and realized outcome information are available, individual loan-level risk remains unobservable. We address this limitation by inferring risk from cross-sectional dispersion in realized returns and incorporating it into the lending rule. Specifically, we rank loans along a one-dimensional scale of expected returns, partition them into multiple intervals, and use the cross-sectional variance within each interval as a risk proxy for individual loans. This enables the formulation of a decision rule that simultaneously accounts for expected return and risk.

Since [Markowitz \(1952\)](#), the central problem of portfolio optimization has been the optimal management of the trade-off between expected return and risk. The CAPM implies that idiosyncratic risk is eliminated through diversification, so that only sensitivity to market risk (beta) is priced. Although foundational to modern finance, this conclusion relies on the assumption that investors can assign continuous weights to all assets.

The credit screening environment considered here does not satisfy this assumption. Lending decisions are discrete—either approval or rejection—and partial approval is infeasible. In equity markets, investors can continuously adjust portfolio weights after purchase; in lending, by contrast, the full loan amount is typically either approved or rejected. More fundamentally, the investment opportunity set itself is determined as the outcome of the screening rule. When the approval threshold τ is set stringently, only loans with high excess return per unit of risk are selected, improving individual loan quality; however, the portfolio becomes concentrated in a small number of similar loans, so that the outcome of any single loan exerts a larger influence on overall portfolio performance. Conversely, relaxing τ admits loans with diverse characteristics, diluting the influence of any individual loan, but may lower expected returns or include loans with higher risk.

The simultaneous determination of individual loan quality and the portfolio weight of risky assets as functions of τ resembles portfolio optimization under cardinality constraints. However, it differs in that the number of included assets is not imposed exogenously but instead arises endogenously from the screening rule. Portfolio optimization with cardinality constraints typically imposes an exogenous upper bound on the number of included assets and searches for the efficient frontier subject to that constraint; this problem is known to be NP-hard ([Bertsimas and Shioda, 2009](#); [Xu et al., 2023](#)). In the present setting, the cardinality constraint is not externally imposed but emerges from the screening threshold τ itself. The threshold simultaneously determines the portfolio's expected return level and the weight of risky assets admitted. The objective of this study is therefore to identify the value of τ that maximizes the Sharpe ratio under this structure.

Founded in 2007, LendingClub grew to become the world's largest P2P lending platform. Following the resignation of CEO Renaud Laplanche in 2016 and subsequent internal control failures, the firm encountered persistent difficulties in attracting investors [1]. In 2020, LendingClub acquired Radius Bank, shut down its P2P lending platform, and transitioned to a traditional banking model [2]. Although LendingClub operated under a marketplace model in which it did not directly bear loan risk, one factor contributing to its failure was that P2P investors were unable to adequately evaluate the risk–return structure of individual loans ([Morse, 2015](#)).

The experience of LendingClub suggests that optimal investment strategies cannot be established when investors rely solely on portfolio expected returns without properly assessing portfolio risk. For example, LendingClub's grade system reflects the ordinal ranking of default probabilities but fails to capture the substantial variation in risk implied by returns within the same grade. If investors screen loans exclusively on the basis of expected return, loans with identical expected returns but heterogeneous risk profiles may be indiscriminately included in the portfolio, thereby reducing risk-adjusted performance. The decision-making framework proposed in this study—utilizing cross-sectional variance within expected-return intervals—offers an alternative approach to this problem.

The remainder of this paper is organized as follows. [Section 2](#) reviews the literature on P2P loan default prediction and portfolio optimization under cardinality constraints. [Section 3](#) describes the LendingClub data and variables. [Section 4](#) presents the interval partitioning of the bimodal return distribution, the construction of the risk proxy, and the decision rule for Sharpe ratio maximization. [Section 5](#) reports the empirical results. [Section 6](#) concludes.

2. Literature review

Research on default prediction in P2P lending has progressed through the application of machine learning techniques aimed at improving traditional credit scoring models, with most studies focusing on enhancing predictive accuracy. [Serrano-Cinca et al. \(2015\)](#) analyze determinants of default using a logistic regression model applied to LendingClub data and demonstrate that the platform-assigned grade is the most powerful predictor of default; incorporating additional borrower information, such as debt levels, further improves predictive performance. [Kim \(2022\)](#) applies semi-supervised learning to LendingClub data and shows that performance comparable to fully supervised learning can be achieved using only a limited amount of labeled data. [Negi et al. \(2025\)](#) compare machine learning-based default prediction models—including XGBoost, random forests, and artificial neural networks—using Bondora platform data, and demonstrate that predictive outputs can be integrated with return estimation to optimize investment performance. However, these studies primarily focus on optimizing classification metrics such as AUC and F1-score, and efforts to link predictive outcomes directly to portfolio return optimization from the investor's perspective remain relatively limited.

In the context of portfolio optimization, [Byanjankar et al. \(2021\)](#) formulate P2P lending investment as a portfolio optimization problem and propose an expectation-based framework that selects a set of loans minimizing risk subject to a target return constraint. [Nayaka et al. \(2024\)](#) develop an integrated framework combining credit evaluation, return assessment, and portfolio optimization to simultaneously achieve investor-level risk management and return maximization.

The endogenous determination of the investment opportunity set induced by the approval rule, as emphasized in this study, is closely related to portfolio optimization under cardinality constraints. Portfolio optimization with cardinality constraints is NP-hard [3], and exact solutions typically rely on mixed-integer quadratic programming or metaheuristic approaches ([Bertsimas and Shioda, 2009](#)). [Xu et al. \(2023\)](#) propose an efficient global optimization algorithm based on Lagrangian relaxation for mean-variance optimization under cardinality constraints. In these studies, however, the cardinality constraint is imposed exogenously—often to reduce transaction costs or facilitate portfolio management—and does not address situations in which the number of investable assets is endogenously determined by a decision rule.

While the existing literature has advanced either default prediction accuracy or portfolio optimization techniques, it has not sufficiently examined how approval rules shape the investment opportunity set and how this process affects portfolio performance. To address this gap, the present study develops a Sharpe ratio-based investment decision framework that

maximizes risk-adjusted performance in a setting where the approval rule endogenously determines the size of the investment set.

3. Data

This study uses loan-level data publicly released by LendingClub, covering the period from June 2007 to September 2020. The raw dataset consists of approximately 1,755,295 loans, with 141 variables recorded for each loan. For each origination, loan terms, borrower characteristics, and credit history are observed once at the time of issuance. Because origination dates differ across loans, the dataset contains temporal information; however, it represents a pooled cross-section rather than a panel dataset tracking the same borrowers over time. To account for time-varying macroeconomic conditions and seasonality, the issuance year is included as an explanatory variable. The issuance month is transformed using trigonometric encoding— $\sin(2\pi \cdot \text{month}/12)$ and $\cos(2\pi \cdot \text{month}/12)$ —to preserve the cyclical structure in which December and January are contiguous. These transformed variables are used as inputs in the default probability model.

Model estimation for deriving the optimal investment strategy relies exclusively on variables observable at the time of loan application. Post-origination variables—such as repayment history or delinquency records—are excluded because they contain future information unavailable at the decision point and would otherwise induce data leakage, resulting in an upward bias in predictive performance.

Default is defined using the `loan_status` variable: loans classified as “Charged Off” or “Default” are treated as defaults, and “Fully Paid” loans are classified as non-defaults. Loans in intermediate states—such as “Current,” “In Grace Period,” or “Late”—are excluded because their final outcomes are not yet realized. A total of 639,139 observations are removed under this criterion.

Borrower characteristics used to predict ex ante investment performance are grouped into four categories, credit-related variables (e.g. credit grade, revolving utilization, delinquency history), financial variables (e.g. annual income, debt-to-income ratio, loan amount), employment-related variables (e.g. employment length, home ownership status), and loan-specific variables (e.g. loan purpose, term, interest rate).

For monetary variables—such as `avg_cur_bal`, `revol_bal`, and `tot_cur_bal`—a logarithmic transformation is applied to mitigate extreme right skewness. For variables characterized by structural missingness—such as `mths_since_last_delinq` and `mths_since_last_major_derog`, which are recorded only for borrowers who have experienced the relevant event—missing indicators are introduced to prevent information loss [4]. The average loan amount in the sample is USD 15,351.

The outcome variable used to evaluate investment performance is the realized return r_i for each loan, computed by annualizing the holding period return (HPR) based on the contractual term [5].

$$HPR_i = \frac{\text{Total Amount Received}_i - \text{Loan Amount}_i}{\text{Loan Amount}_i}, \quad r_i = (1 + HPR_i)^{12/T_i} - 1$$

Here, T_i denotes the contractual maturity of loan i (36 or 60 months). In the case of full repayment, the return is determined by the contractual internal rate of return (IRR). In the event of default, returns are generally distributed in the negative domain, depending on the timing of default and the amount recovered.

In the investment simulation, we retrospectively reconstruct investment decisions using realized LendingClub loan data to evaluate performance conditional on approving or rejecting each loan. If a loan is rejected, the corresponding capital is assumed to be invested in US Treasury securities with a maturity matching the loan term (36 or 60 months) at the loan’s issuance date (`issue_d`). US Treasury yields are obtained from FRED (Federal Reserve

Economic Data), using the 3-year (DGS3) and 5-year (DGS5) maturity series. Although all loans in the dataset were in fact originated, portfolio performance is computed ex post under the counterfactual assumption that selected loans were not funded and that the capital was instead allocated to Treasury securities.

4. Methodology

4.1 Bimodal distribution of returns and expected returns

The distribution of loan returns exhibits a bimodal pattern. Among 898,522 fully repaid loans, the mean return is 4.4% and the median is 4.1%, concentrated around the contractual interest rate (mean 12.6%), with a relatively low standard deviation of 2.4% points. In contrast, 217,634 defaulted loans have a mean return of -18.2% and a median of -16.4%, concentrated in the loss region, with a substantially larger standard deviation of 14.9% points. Defaults account for 19.5% of the total sample. Figure 1 illustrates the distribution of realized returns by default status.

Because this bimodal structure arises from the discrete states of default and non-default, it can be naturally represented by a two-component mixture model.

$$P(r_i | x_i) = (1 - p_i) \cdot g_0(r_i) + p_i \cdot g_1(r_i)$$

Here, $p_i = f(x_i; \theta)$ denotes the probability of default, g_0 is the return distribution conditional on full repayment, and g_1 is the return distribution conditional on default. Under this mixture structure, the expected return is given by $\hat{E}[r_i] = (1 - \hat{p}_i) \cdot E[g_0] + \hat{p}_i \cdot E[g_1]$. Since IRR_i is deterministically observed from the contractual terms and \bar{r}_{default} is computed as the sample mean of defaulted loans in the training data, a straightforward approach is to substitute the non-default component with the contractual IRR_i and the default component with the overall average realized return of defaulted loans, \bar{r}_{default} , obtained from the training sample.

$$\hat{E}[r_i] = (1 - \hat{p}_i) \cdot IRR_i + \hat{p}_i \cdot \bar{r}_{\text{default}}$$

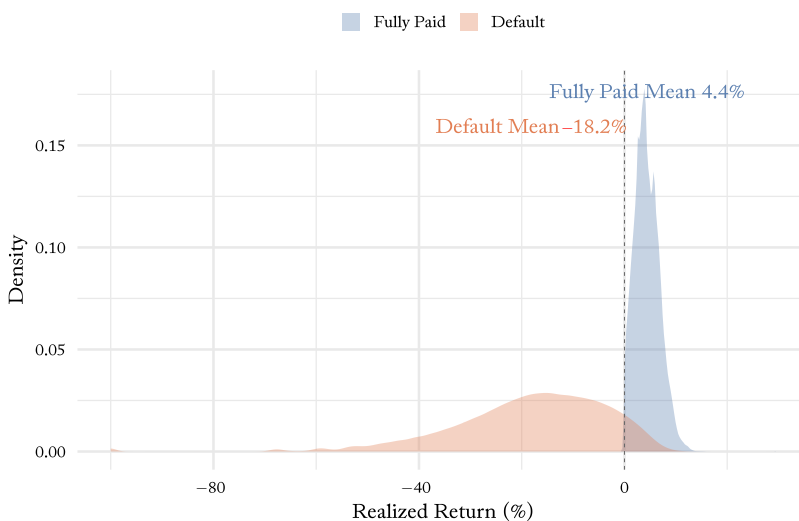


Figure 1. Distribution of realized returns by default status

However, this constant substitution approach has two limitations. First, loans with higher default probabilities tend to exhibit lower recovery rates upon default, yet a single \bar{r}_{default} fails to capture such heterogeneity. Second, IRR_i reflects the contractual return conditional on full repayment to maturity; therefore, realized returns may diverge from IRR_i in cases of prepayment.

To address these limitations, this study employs conditional returns by deciles of predicted default probability. Specifically, the predicted probabilities \hat{p}_i in the training data are partitioned into ten deciles $d = 1, \dots, 10$. Within each decile, we compute separately the average realized return of non-defaulted loans, $\bar{r}_{\text{normal},d}$, and the average realized return of defaulted loans, $\bar{r}_{\text{default},d}$. The expected return is then given by

$$\hat{E}[r_i] = (1 - \hat{p}_i) \cdot \bar{r}_{\text{normal},d(i)} + \hat{p}_i \cdot \bar{r}_{\text{default},d(i)}$$

Loans in the validation and evaluation samples are assigned to deciles according to the boundaries determined from the training data, thereby preserving consistency in interval classification. However, because the decile-specific conditional returns rely entirely on the empirical distribution of the training sample, the possibility of overfitting cannot be ruled out [6]. Ultimately, within this structure, the key determinant of expected return estimation is the prediction of the default probability \hat{p}_i . The multidimensional information contained in the feature vector x_i is compressed into a single scalar—expected return—through the mapping $x_i \rightarrow \hat{p}_i \rightarrow \hat{E}[r_i]$.

4.2 Construction of a risk proxy via interval partitioning

Expected return alone is insufficient for optimal lending decisions. Consider two loans with identical expected returns of 5%. Loan A has a 95% probability of full repayment and a 5% probability of a small loss, representing a relatively low-risk structure. Loan B, by contrast, is a high-interest loan with a contractual IRR of 30%, a 60% probability of full repayment, and a 40% probability of a large loss. Although both loans share the same expected return, their risk profiles differ substantially. Loan A exhibits low return volatility, with realizations concentrated near its expectation, whereas Loan B is highly dispersed between high returns and large losses. A rational investor would prefer Loan A when expected returns are equal. Incorporating this preference into a decision rule requires information on risk—specifically, the variance of returns—in addition to expected return.

However, each loan is observed with only a single realized return r_i , rendering its individual variance unobservable. To address this limitation, we construct intervals (bins) along the expected return dimension. Loans with similar expected returns are grouped within the same interval, allowing the cross-sectional variance of realized returns within each bin to be computed. This variance is then used as a risk proxy for individual loans belonging to that interval.

Specifically, the estimated expected returns $\hat{E}[r_i]$ in the training data are partitioned into J intervals B_1, \dots, B_J . For each interval B_j , we compute

$$\mu_j = \frac{1}{n_j} \sum_{i \in B_j} r_i, \quad \sigma_j^2 = \frac{1}{n_j - 1} \sum_{i \in B_j} (r_i - \mu_j)^2$$

Where n_j denotes the number of loans in interval B_j . These interval-specific statistics measure the extent to which loans with identical expected returns realize heterogeneous outcomes in practice. These statistics capture the intrinsic uncertainty arising from the fact that loan outcomes are realized only as full repayment or default.

4.3 Interval assignment for new loans and the structure of portfolio returns

Given a new borrower's feature vector x_{new} , the default probability model outputs $\hat{p}_{\text{new}} = f(x_{\text{new}}; \theta)$, and the expected return $\hat{E}[r_{\text{new}}]$ is computed using the decile-specific conditional returns. This value is then mapped onto the interval boundaries $\{b_0, b_1, \dots, b_J\}$ determined from the training data. If $\hat{E}[r_{\text{new}}] \in [b_{j-1}, b_j)$, the loan is assigned to interval B_j , and the corresponding σ_j is applied as the risk proxy.

Let $w_j = n_j/N$ denote the proportion of loans assigned to interval B_j . Raising τ to exclude low-return intervals increases the portfolio mean return. At the same time, because defaulted loans are concentrated in low-return intervals, which therefore exhibit large σ_j , excluding them also reduces the return variance among the remaining loans. That is, τ simultaneously shifts both the expected-return level and the risk profile of the portfolio.

4.4 Decision rule and sharpe ratio optimization

The approval or rejection decision for an individual loan is determined by whether its risk-adjusted excess return exceeds a screening threshold τ .

$$\frac{\hat{E}[r_i] - r_{\text{treasury},i}}{\hat{\sigma}(r_i)} > \tau$$

Here, $r_{\text{treasury},i}$ denotes the Treasury yield corresponding to the issuance date and maturity of loan i , and $\hat{\sigma}(r_i)$ is the standard deviation of returns for the expected-return interval to which loan i belongs (i.e. $\hat{\sigma}(r_i) = \sigma_{j(i)}$). The left-hand side is a Sharpe ratio-like measure at the individual-loan level: even when expected returns are identical, high-risk loans are screened out, and only loans with sufficiently high excess return per unit of risk are approved.

At the portfolio level, the Sharpe ratio is defined as

$$SR(\tau) = \frac{\bar{R}_p(\tau) - \bar{r}_{\text{treasury}}(\tau)}{\sigma_p(\tau)}$$

Here, $\bar{R}_p(\tau)$ is the weighted average return of the portfolio constructed under threshold τ . Approved loans contribute their realized returns r_i , whereas rejected loans are assigned the corresponding Treasury yield $r_{\text{treasury},i}$. The term $\bar{r}_{\text{treasury}}(\tau)$ denotes the weighted average of maturity-matched Treasury yields for the loans constituting the portfolio; because the portfolio composition varies with τ , it is itself a function of τ . The term $\sigma_p(\tau)$ is the weighted cross-sectional standard deviation of individual loan returns within the portfolio constructed under τ . Because each loan generates only a single realized outcome, time-series volatility cannot be computed at the individual level; the cross-sectional variance of loans comprising the same portfolio is therefore used as a proxy. As τ changes, the composition of admitted loans changes, making σ_p a function of τ . Because rejected capital is allocated to Treasury securities, the Treasury yield represents the opportunity cost of loan funding. Under this definition, if all loans are rejected and all capital is invested in Treasuries, then $\bar{R}_p = \bar{r}_{\text{treasury}}$, implying $SR = 0$. A positive Sharpe ratio is obtained only when approved loans generate excess returns above this opportunity cost.

Since $SR(\tau)$ is a function of τ , the optimal threshold τ^* is selected in the validation sample via grid search.

$$\tau^* = \underset{\tau}{\operatorname{argmax}} SR(\tau)$$

4.5 Model estimation procedure

Executing repeated K -fold cross-validation on the full dataset is computationally intensive. We therefore draw a 10% stratified random sample that preserves the default rate and conduct the analysis on this subsample.

The procedure is as follows.

Step 1 (Default Probability Model Estimation). We estimate a model predicting the default probability $p_i = f(x_i; \theta)$ from the feature vector x_i . The loss function is binary cross-entropy.

$$L(\theta) = -\frac{1}{n} \sum_{i=1}^n [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)]$$

We implement XGBoost. The hyperparameters (`max_depth`, `learning_rate`, `subsample`, `colsample_bytree`, `min_child_weight`) are selected in the first repetition via Bayesian optimization [7] to maximize AUC, and then fixed for subsequent repetitions. Although re-optimizing hyperparameters in every repetition would be ideal, doing so would require an additional $K \times 25$ training runs (25 search trials per repetition). To control computational cost, we reuse the first-repetition hyperparameters across all repetitions. Because these hyperparameters govern structural model complexity, we do not expect the optimal configuration to vary substantially across subsamples drawn from the same underlying data.

Step 2 (Estimation of Decile-Specific Conditional Returns and Computation of Expected Returns). The predicted default probabilities in the training sample are partitioned into ten deciles. Within each decile d , we compute the average realized return of non-defaulted loans, $\bar{r}_{\text{normal},d}$, and the average realized return of defaulted loans, $\bar{r}_{\text{default},d}$, and fix the decile boundaries accordingly. Using these quantities, we compute the expected return for each loan in the training sample: $\hat{E}[r_i] = (1 - \hat{p}_i) \cdot \bar{r}_{\text{normal},d(i)} + \hat{p}_i \cdot \bar{r}_{\text{default},d(i)}$.

Step 3 (Interval Partitioning and Construction of Risk Proxies). We partition the expected-return axis into J intervals and compute, for each interval, the mean realized return μ_j and the standard deviation σ_j .

Step 4 (Threshold Optimization). In the validation sample, we conduct a grid search over τ and select τ^* that maximizes the portfolio Sharpe ratio.

Step 5 (Final Evaluation). In the evaluation sample, we apply τ^* to compute the final portfolio Sharpe ratio.

Steps 2–5 are repeated K times. The data are split into training, validation, and evaluation sets in a 6:2:2 ratio. In each repetition, the data are randomly partitioned using a different seed, and the entire procedure—from model estimation to final evaluation—is conducted independently. We set $K = 50$. Random partitioning is adopted because the objective of this study is not to forecast future defaults in chronological order but to verify whether a screening rule constructed solely from information available at the time of loan origination generates consistent risk-adjusted excess returns across repeated sample splits.

5. Empirical results

5.1 Summary of repeated cross-validation results

Table 1 summarizes evaluation-sample performance across 50 repetitions. The default prediction model achieves an average AUC of 0.7196 (standard deviation 0.0038), indicating stable discriminatory power across repetitions. Under the optimal screening threshold τ^* , the portfolio attains an average Sharpe ratio of 0.0919, with a corresponding average approval rate of 26.9%.

5.2 Interval-specific statistics of expected returns

Table 2 reports interval-specific statistics computed from the training sample in the first split. With $J = 10$ intervals, the mean realized return μ_j increases with the interval index, while the

Table 1. Summary of repeated cross-validation results ($K = 50$)

Metric	Mean	Std. Dev	Min	Max
AUC	0.7196	0.0038	0.7119	0.7269
SR(τ^*)	0.0919	0.0096	0.0766	0.1203
τ^*	0.0680	0.0414	0.0000	0.1500
Approval rate	0.2691	0.0694	0.1009	0.4070
Portfolio return (%)	1.7264	0.0604	1.5754	1.8626
Portfolio volatility (%)	3.4381	0.6038	1.9202	4.6494

Note(s): * The approval rate and portfolio statistics are computed in the evaluation sample under the optimal screening threshold τ^* for each repetition

Table 2. Interval-specific statistics of expected returns (decile-based approach, first split)

Interval	E[r] lower bound	E[r] upper bound	μ_j	σ_j	n_j
1	-13.91	-3.52	-9.20	16.22	6,697
2	-3.52	-1.67	-3.52	14.95	6,697
3	-1.67	-0.47	-1.12	13.70	6,697
4	-0.47	0.39	0.21	11.96	6,697
5	0.39	0.99	1.16	10.86	6,697
6	0.99	1.40	2.03	9.08	6,696
7	1.40	1.63	2.24	7.66	6,697
8	1.63	1.80	2.54	6.23	6,697
9	1.80	1.99	2.60	5.57	6,697
10	1.99	2.42	2.60	4.12	6,697

Note(s): * μ_j and σ_j denote the mean and standard deviation of realized returns for loans within each interval

standard deviation σ_j is larger in low-return intervals and declines in high-return intervals. Low-return intervals contain loans with high default probabilities; dispersion between default and full-repayment outcomes amplifies within-interval variance. High-return intervals consist predominantly of fully repaid loans, with realized returns clustering around the contractual IRR.

5.3 Portfolio performance across screening thresholds

Figure 2 reports validation-sample performance (averaged across 50 repetitions) as τ varies from -1.0 to 2.0 in increments of 0.05 . The Sharpe ratio exhibits an inverted U-shaped pattern, peaking around $\tau \approx 0.05$. The individual-loan risk-adjusted excess return $(\widehat{E}[r_i] - r_{\text{treasury},i})/\sigma_j$ is concentrated within a narrow range around zero. When τ lies below the lower bound of this distribution (in the negative region), virtually all loans satisfy the approval criterion and the approval rate approaches 100%; however, after accounting for default losses, portfolio returns fall below Treasury yields, producing a negative Sharpe ratio. As τ increases toward zero, loans with negative risk-adjusted excess returns—i.e. inferior to Treasuries on a risk-adjusted basis—are excluded first, and the Sharpe ratio improves sharply. The optimal threshold τ^* averages 0.068 , close to zero, reflecting that risk-adjusted excess returns in P2P lending are only marginal relative to Treasuries, so modest screening eliminates most low-quality loans.

Beyond τ^* , the rapid decline in approvals results in a portfolio composed of a small number of loans, so that the default status of any individual loan exerts a disproportionately large effect on portfolio returns, and the Sharpe ratio declines. This inverted U-shaped pattern reflects the fundamental trade-off in an endogenous investment set: stricter screening improves the quality

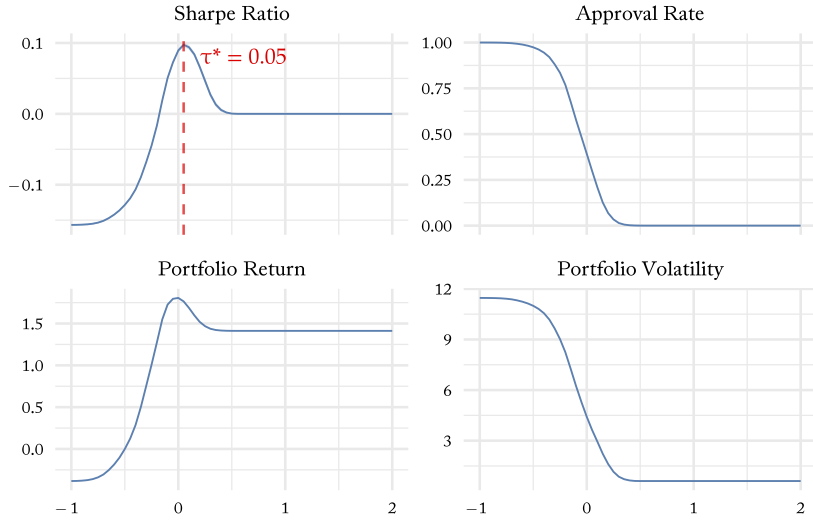


Figure 2. Portfolio performance metrics by screening threshold (50-iteration mean; shaded areas represent 95% confidence intervals)

of the approved loan pool but concentrates the portfolio on fewer loans, amplifying the influence of individual risk. As τ rises, the approval rate declines in an S-shaped pattern over the $\tau \approx -0.5$ to 0.5 range, the portfolio return peaks near $\tau \approx 0$ before declining gradually, and portfolio volatility decreases monotonically.

5.4 Distribution of the optimal screening threshold

Figure 3 illustrates the distribution of τ^* selected across the 50 repetitions. The optimal threshold ranges between 0.00 and 0.15, with a mean of 0.068 (standard deviation 0.041),

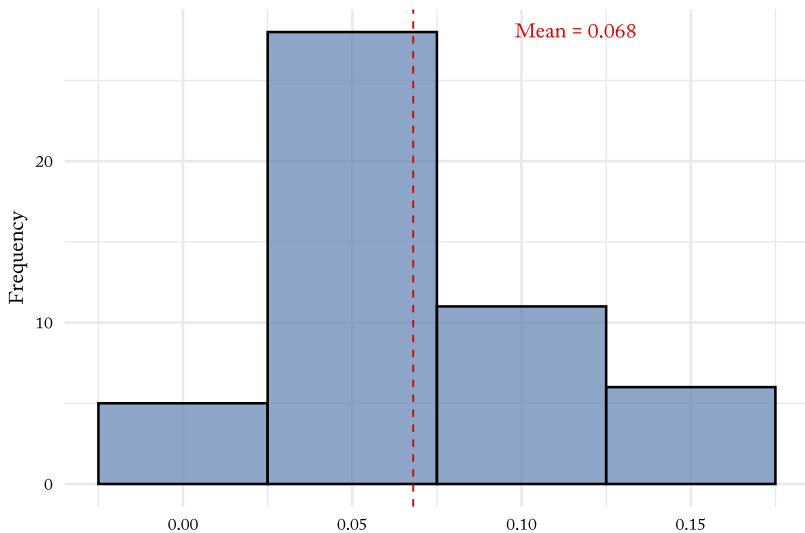


Figure 3. Distribution of the optimal screening threshold τ^* ($K = 50$)

concentrated around 0.05. This indicates that the optimal screening threshold is relatively stable across data partitions, suggesting strong practical applicability as an operational decision rule.

5.5 Sensitivity analysis with respect to the number of intervals J

Because the number of intervals J is a methodological choice determined ex ante by the researcher, the credibility of the analysis would be weakened if the results were highly sensitive to this parameter. To assess robustness, we conduct five repetitions for each of $J = 5, 10, 15, 20$. The results are presented in Figure 4 and Table 3. Across the four specifications, the average Sharpe ratio ranges from 0.0827 to 0.0836. The standard deviation intervals overlap substantially, and no statistically meaningful differences are observed. The results remain stable across alternative choices of the number of intervals.

5.6 Relationship between default prediction performance and portfolio performance

Figure 5 plots the relationship between AUC and the Sharpe ratio across 50 repetitions; the correlation coefficient is 0.228, indicating a positive but weak association. While improvements in default discrimination contribute in the expected direction, substantial dispersion in the Sharpe ratio remains even at similar levels of AUC. AUC measures discrimination between default and non-default outcomes, whereas the Sharpe ratio evaluates risk-adjusted excess return. Portfolio performance may remain constrained if expected returns do not sufficiently exceed Treasury yields or if dispersion of realized returns within an expected-return interval remains large.

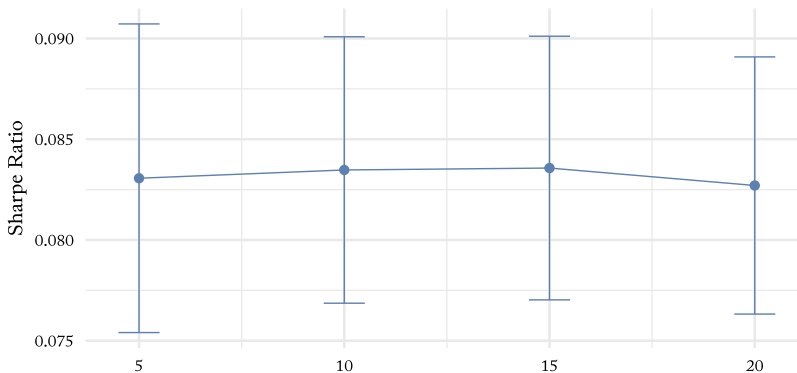


Figure 4. Sharpe ratio by number of intervals (J) (mean \pm standard deviation)

Table 3. Portfolio performance by number of intervals (J)

J	Mean SR	Std. Dev. of SR	Approval rate	Mean τ^*
5	0.0831	0.0077	0.2526	0.0700
10	0.0835	0.0066	0.2530	0.0700
15	0.0836	0.0065	0.2538	0.0700
20	0.0827	0.0064	0.2528	0.0700

Note(s): * Reported values are the mean and standard deviation computed over five repetitions for each J

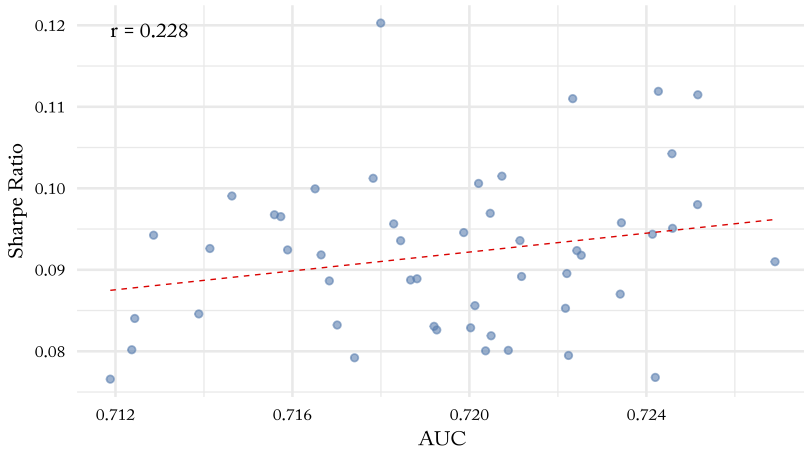


Figure 5. Relationship between AUC and sharpe ratio ($K = 50$)

5.7 Benchmark comparison

Table 4 compares the decile-based model with three benchmark strategies. The approve-all strategy yields a Sharpe ratio of -0.1536 , indicating that indiscriminate P2P investment underperforms Treasuries on a risk-adjusted basis. Although the average IRR of LendingClub loans is 13.2%, exceeding the average Treasury yield of 1.4% over the same period by 11.8% points, realized returns after accounting for default losses fall below Treasury yields. The reject-all strategy yields a Sharpe ratio of 0 by definition.

The decile-based model achieves a Sharpe ratio of 0.0919, an improvement of approximately 0.25 relative to the approve-all strategy (-0.1536). When scaled by the approve-all portfolio volatility (11.3%), this corresponds to an annualized risk-adjusted excess return of approximately 2.8% points, reflecting the economic value of credit screening based on default prediction. The Oracle strategy assumes perfect ex ante prediction of default and serves as an unattainable upper bound; the proposed model’s Sharpe ratio lies between the reject-all benchmark and the Oracle benchmark.

6. Conclusion

We develop a Sharpe ratio–maximizing decision framework for credit screening environments in which the approval threshold determines the investment opportunity set. By using the cross-sectional variance within expected-return intervals as a risk proxy, the framework incorporates

Table 4. Benchmark strategy comparison

Strategy	Sharpe ratio	Approval rate	Portfolio return (%)	Portfolio volatility (%)
Oracle (perfect default prediction)	0.9658	0.80	3.7543	2.4255
Approve-all	-0.1536	1.00	-0.3225	11.2890
Reject-all	0.0000	0.00	1.4117	0.6057
Decile model (50-run mean)	0.0919	0.27	1.7264	3.4381

Note(s): * Statistics for the decile model are averaged over 50 repetitions, The Oracle benchmark assumes perfect ex ante prediction of default status

risk information beyond expected returns. Using LendingClub data from 2007–2020 and conducting 50 repeated cross-validations, we obtain an average Sharpe ratio of 0.0919 and an average approval rate of 26.9% under the optimal screening threshold τ^* . The default prediction model maintains a stable AUC of 0.7196. The inverted U-shaped pattern of the Sharpe ratio with respect to τ confirms that the trade-off between the quality of the approved loan pool and portfolio diversification is an inherent structural feature of an endogenous investment set. Sensitivity analysis over $J = 5, 10, 15, 20$ intervals reveals no statistically significant differences in the Sharpe ratio, supporting robustness to the choice of interval count.

The empirical implications are as follows. First, although the average contractual IRR of LendingClub loans is 13.2%, exceeding Treasury yields by 11.8% points, the approve-all strategy yields a negative Sharpe ratio (-0.1536). The nominal credit spread is insufficient to compensate for realized default losses.

Second, credit screening improves the Sharpe ratio by approximately 0.25, turning excess performance from negative to positive; however, achieving substantially larger positive excess performance while expanding the approval rate beyond 26.9% likely requires additional information that more precisely captures borrowers' repayment capacity.

Third, the ordering approve-all ($SR < 0$) < reject-all ($SR = 0$) < proposed model ($SR > 0$) demonstrates that a portfolio that underperforms Treasuries when evaluated using expected-return information alone can outperform Treasuries once risk information is incorporated. Idiosyncratic risk should therefore not be treated as negligible merely because it is, in principle, diversifiable; because the degree of diversification is variable in an endogenous investment set, idiosyncratic risk must be explicitly incorporated into decision-making.

The findings demonstrate the importance of incorporating asset-level risk information into the decision process when screening rules endogenously determine the investment opportunity set. As τ is set more stringently, the portfolio becomes concentrated on fewer loans and thus more sensitive to individual loan outcomes; explicit incorporation of risk information, rather than expected return alone, is therefore essential. Although developed in the context of P2P lending, the framework can be extended to domains in which screening rules determine the investment set, including venture capital investment, insurance underwriting, and project financing.

From a practical perspective, the results suggest that P2P lending platforms can improve investor decision quality by providing information not only on expected returns but also on return volatility. LendingClub's grade system reflected only the ordinal ranking of default probabilities and failed to capture heterogeneity in return volatility within grades, which may be interpreted as one contributor to the limitations of the P2P model.

This study has several limitations. First, individual loan risk is proxied by cross-sectional variance within expected-return intervals, implicitly assuming risk homogeneity within each interval; in practice, risk may remain heterogeneous, and such heterogeneity is likely to increase as intervals widen. Second, because the LendingClub dataset contains only approved loans, information on rejected applications is unobserved, raising the possibility of sample selection bias. Third, results may depend on the choice of interval boundaries, and there is limited theoretical guidance for selecting the optimal number of intervals J . Fourth, default correlations across loans are not explicitly modeled. Although the issuance year and trigonometrically encoded issuance month are included as inputs to the default probability model, partially controlling for time-varying macroeconomic conditions and seasonality, this does not substitute for the correlation structure in which defaults cluster simultaneously during economic downturns. In the presence of systematic risk, cross-sectional variance may understate risk relative to normal periods, potentially shifting both the absolute level of the Sharpe ratio and the location of τ^* . However, the core logic of this study—that τ governs the scope of the portfolio, generating a trade-off between the quality of the approved loan pool and portfolio concentration—is judged to hold independently of the influence of systematic risk. Setting τ^* differentially across business-cycle phases or employing conditional variances that

incorporate default correlations as risk proxies are follow-up tasks that can further develop the present framework. Fifth, although this study verifies the stability of the screening rule through repeated random partitioning, out-of-time validation using a temporal split was not performed. In practice, models trained on historical data are applied to future loans; assessing the practical applicability of the proposed framework therefore requires follow-up validation based on a temporal split. These limitations remain important topics for future research.

Acknowledgments

The authors are grateful to Professor Keunkwan Ryu (Seoul National University) for insights into the economic meaning of the Sharpe ratio in P2P lending portfolio return optimization.

Notes

1. “LendingClub experienced problems in early 2016, with difficulties in attracting investors, a scandal over some of the firm’s loans and concerns by the board over CEO Renaud Laplanche’s disclosures leading to a large drop in its share price and Laplanche’s resignation.” Wikipedia, LendingClub.
2. “LendingClub acquired Radius Bank and announced that it would be shutting down its peer-to-peer lending platform.” Frankel (2021).
3. An NP-hard problem refers to a class of problems for which no algorithm is known that guarantees an optimal solution in polynomial time. Portfolio optimization under cardinality constraints falls into this category because the number of possible asset combinations increases combinatorially with the number of available assets.
4. After preprocessing, a total of 100 variables are included in the model estimation. Detailed descriptions of the variables used in the analysis are available upon request.
5. Two approaches may be used to annualize the holding period return: one based on the actual investment period and the other based on the contractual term. Annualization based on the actual investment period can substantially amplify short-term losses in cases of early default by extrapolating them to a one-year horizon. For example, consider a loan with a contractual maturity of 36 months that defaults after 6 months, with only 30% of principal recovered. The holding period return of -70% , when annualized over the actual period (6 months), yields $((1 - 0.70)^{12/6} - 1) \times 100 = -91.0\%$. By contrast, annualization based on the contractual term (36 months) yields $((1 - 0.70)^{12/36} - 1) \times 100 = -33.1\%$, which is approximately 2.8 times smaller in absolute value. Thus, annualizing returns using the actual investment period extrapolates short-term default losses to an annual scale, thereby overstating the average return of defaulted loans, \bar{r}_{default} . When such inflated values enter the computation of expected returns, most loans appear to underperform Treasury yields. Moreover, since *IRR* is reported as an annualized rate based on the contractual term, it is appropriate to apply the same time convention when computing realized returns in the event of default.
6. We implement both the constant and decile-based approaches under an identical data partition and compare the mean absolute error (MAE) between expected and realized returns in the validation sample. The decile-based approach yields an MAE no greater than that of the constant specification, suggesting that the likelihood of overfitting is limited. Further details are available from the author upon request.
7. We use Optuna (Akiba *et al.*, 2019), a hyperparameter optimization framework, with the Tree-structured Parzen Estimator (TPE) as the internal algorithm.

References

- Akiba, T., Sano, S., Yanase, T., Ohta, T. and Koyama, M. (2019), “Optuna: a next-generation hyperparameter optimization framework”, *The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2623-2631.
- Bertsimas, D. and Shioda, R. (2009), “Algorithm for cardinality-constrained quadratic optimization”, *Computational Optimization and Applications*, Vol. 43 No. 1, pp. 1-22, doi: [10.1007/s10589-007-9126-9](https://doi.org/10.1007/s10589-007-9126-9).

- Byanjankar, A., Mezei, J. and Heikkilä, M. (2021), "Data-driven optimization of peer-to-peer lending portfolios based on the expected value framework", *Intelligent Systems in Accounting, Finance and Management*, Vol. 28 No. 3, pp. e1490-129, doi: [10.1002/isaf.1490](https://doi.org/10.1002/isaf.1490).
- Frankel, M. (2021), "LendingClub is ending its P2P lending platform - now what?", *The Motley Fool*, available at: <https://www.fool.com/money/personal-loans/articles/lendingclub-ending-its-p2p-platform-now-what/> (accessed 28 April 2026).
- Kim, H. (2022), "Semi-supervised learning to predict default risk for P2P lending", *Journal of Digital Convergence*, Vol. 20 No. 4, pp. 185-192, doi: [10.14400/JDC.2022.20.4.185](https://doi.org/10.14400/JDC.2022.20.4.185).
- Markowitz, H. (1952), "Portfolio selection", *The Journal of Finance*, Vol. 7 No. 1, pp. 77-91, doi: [10.1111/j.1540-6261.1952.tb01525.x](https://doi.org/10.1111/j.1540-6261.1952.tb01525.x).
- Morse, A. (2015), "Peer-to-peer crowdfunding: information and the potential for disruption in consumer lending", *Annual Review of Financial Economics*, Vol. 7 No. 1, pp. 463-482, doi: [10.1146/annurev-financial-111914-041939](https://doi.org/10.1146/annurev-financial-111914-041939).
- Nayaka, P., Bhowmik, B. and Hegde, A. (2024), "Advancements in credit scoring, profit scoring, and portfolio optimization for P2P lending", *Proceedings of the 2024 International Conference on Communication, Control, and Intelligent Systems (CCIS)*, IEEE, doi: [10.1109/CCIS63231.2024.10932068](https://doi.org/10.1109/CCIS63231.2024.10932068).
- Negi, A., Nayaka, P., Pandey, A., Bhowmik, B. and Gautam, H. (2025), "Enhancing investment decisions in P2P lending using machine learning", *Proceedings of the 2025 Control Instrumentation System Conference (CISCON)*, IEEE.
- Serrano-Cinca, C., Gutiérrez-Nieto, B. and López-Palacios, L. (2015), "Determinants of default in P2P lending", *PLOS One*, Vol. 10 No. 10, e0139427, doi: [10.1371/journal.pone.0139427](https://doi.org/10.1371/journal.pone.0139427).
- Xu, W., Tang, J., Yiu, K.F.C. and Wen Peng, J. (2023), "An efficient global optimal method for cardinality constrained portfolio optimization", *INFORMS Journal on Computing*, Vol. 36 No. 2, pp. 690-704, doi: [10.1287/ijoc.2022.0344](https://doi.org/10.1287/ijoc.2022.0344).

Corresponding author

Jaesung James Park can be contacted at: duundich@gmail.com