risks

MDPI

Communication

Credit Scoring for Peer-to-Peer Lending

Daniel Felix Ahelegbey * and Paolo Giudici

Department of Economics and Management Sciences, University of Pavia, 27100 Pavia, Italy; paolo.giudici@unipv.it

* Correspondence: danielfelix.ahelegbey@unipv.it

Abstract: This paper shows how to improve the measurement of credit scoring by means of factor clustering. The improved measurement applies, in particular, to small and medium enterprises (SMEs) involved in P2P lending. The approach explores the concept of familiarity which relies on the notion that the more familiar/similar things are, the closer they are in terms of functionality or hidden characteristics (latent factors that drive the observed data). The approach uses singular value decomposition to extract the factors underlying the observed financial performance ratios of SMEs. We then cluster the factors using the standard k-mean algorithm. This enables us to segment the heterogeneous population into clusters with more homogeneous characteristics. The result shows that clusters with relatively fewer number of SMEs produce a more parsimonious and interpretable credit scoring model with better default predictive performance.

Keywords: clustering; credit scoring; factor models; FinTech; P2P lending; segmentation

1. Introduction

When it comes to the measurement of credit scoring, the traditional concept of one model fits all may work well only for firms or individuals with capacity, credit access, cash, and/or collateral. Such models usually do not work for people with no financial history or collateral even if they have payback capabilities. Continuing with the traditional credit scoring system will not help a significant proportion of SMEs. It is therefore vital and crucial to develop alternative credit scoring models tailored in a way to enable SMEs without traditional financial history but with payback capabilities based on alternative means to have a credit score that enables them to gain access to credit.

Recent advancements gradually transforming the traditional economic and financial system is the emergence of digital-based systems. Such systems present a paradigm shift from traditional intra-organizational systems to customer-oriented technological (digital) systems. Financial technological ("FinTech") companies are gradually gaining ground in major developed economies across the world. The emergence of business-to-customer (B2C), customer-to-customer (C2C), provider-oriented business-to-business (B2B), and peer-to-peer (P2P) platforms are typical examples of FinTech systems. Thus, Fintech offers solutions that differ from traditional institutions regarding the providers and the interaction types as well as regarding the banking and insurance processes they support (Haddad and Hornuf 2019; Puschmann 2017). These platforms aim at facilitating credit services by connecting individual lenders with individual borrowers without the interference of traditional banks as intermediaries. Such platforms serve as a digital financial market and have significantly improved the customer experience in terms of cost saving and speed of the services to both individual borrowers and lenders, as well as small business owners.

Despite the various advantages of current fintech systems, the existing digital platform systems inherit some of the challenges of traditional credit risk management. Credit scoring is purely based on the data available on a borrower that signals their financial worth and ability to pay back loans. In addition, they are characterized by the asymmetry of information and by a strong interconnectedness among their users (see, e.g., Chen et al.

Citation: Ahelegbey, Daniel Felix, and Paolo Giudici. 2023. Credit Scoring for Peer-to-Peer Lending. *Risks* 11: 123. https://doi.org/10.3390/risks11070123

Academic Editor: Mogens Steffensen

Received: 12 June 2023
Revised: 1 July 2023
Accepted: 4 July 2023
Published: 7 July 2023

2022; Giudici and Polinesi 2021; Giudici et al. 2020, 2022) that makes distinguishing healthy and risky credit applicants difficult, thus affecting credit issuers. There is, therefore, a need to explore methods that can help improve the credit scoring of individuals or companies that engage in P2P credit services (see El Annas et al. 2023; Jiang et al. 2018; Lyócsa et al. 2022; Ma et al. 2021; Xia et al. 2020).

This paper investigates how a factor clustering-based approach to segment a population in order to improve the statistical-based credit score for small and medium enterprises (SMEs) involved in P2P lending. The methodology employed in this paper extends the similarity of latent factors recently adopted by Ahelegbey et al. (2019a, 2019b). The approach explores the concept of familiarity as a signal for functional relationships among financial institutions. The familiarity concept relies on the notion that the more familiar things are, the closer they are in terms of functionality or hidden characteristics (latent factors that drive the observed data). By this reasoning, we postulate that the more familiar the latent factors of SME performance ratios, the closer they are in terms of either holding securities with similar features or pursuing similar financial strategies or models. Such features make SMEs become more identical and increase the probability of exposure to common risk factors.

We contribute to the literature on the application of factor models in finance (see, e.g., Dungey and Gajurel 2015; Dungey et al. 2005; Forbes and Rigobon 2002; Fox and Dunson 2015; Lopes and Carvalho 2007; Nakajima and West 2013). In this paper, we cluster the factors that drive the observed financial data, enabling us to segment the population and estimate a logistic model based on the sample segmentation.

Our empirical application contributes to the literature on credit scoring in P2P lending systems (see Ahelegbey et al. 2019a, 2019b; Andreeva et al. 2007; Barrios et al. 2014; Emekter et al. 2015; Giudici et al. 2020; Serrano-Cinca and Gutiérrez-Nieto 2016, for related works). We provide an application of our approach to a well-known dataset consisting of over 15,000 SMEs involved in P2P credit services across Europe. Our results show that segmentation of the heterogenous market based on latent characteristics presents a more efficient scheme to credit risk measurement that achieves higher performance than the conventional approach.

The paper is organized as follows. Section 2 presents the econometric methodology, Section 3 discusses the application to a credit scoring database provided by a European rating agency for P2P platforms, and Section 4 presents concluding remarks.

2. Econometric Methodology

2.1. Segmented Logistic Model

Let Y_i be a binary variable that represents the observed loan default status of firm-i, where $Y_i = 1$ if the firm has defaulted on its loan, and $Y_i = 0$ if it has not defaulted. Denote with X_i a p-dimensional vector of observed financial features of firm-i that predicts its creditworthiness. The conventional approach to credit scoring is to model the probability of the default given the observed firm financial features via a one-model-fits-all logistic regression. In this application, we assume the firms can be classified into groups according to some similarities in the latent characteristics of their observed features. Suppose there exist k non-overlapping groups of firms. We model the probability of the default of firms-i in group l as a logistic model via the log-odds function given by

$$\log\left(\frac{\pi_i^{(l)}}{1-\pi_i^{(l)}}\right) = \beta_0^{(l)} + \sum_{j=1}^{p} X_{ij}^{(l)} \beta_j^{(l)} \tag{1}$$

where $\pi_i^{(l)} = P(Y_i^{(l)} = 1 | X_i^{(l)})$ is the probability of default of firm-i in group l, $\beta_0^{(l)}$ is a constant term of the group, $\beta_j^{(l)}, j = 1, \ldots, p$, is a logistic regression coefficient that shows the change in the logit of the probability associated with a unit change in the j-th predictor holding all other predictors constant.

2.2. Factor Model with SVD

Let X be the observed data matrix of n institutions, each with p number of features measuring financial performance ratios. We denote with X_i, the i-th institution which corresponds to the i-th row of X. We proceed under the assumption that the observed data matrix X can be approximated via singular value decomposition (SVD) given by (Hoff 2007)

$$X = UDV' \tag{2}$$

where U is of dimensions $n \times r$, V is of dimensions $p \times r$, $r < p$, with the columns of U and V denoting the left and right singular vectors of X, respectively, and D is an r-dimensional diagonal matrix of the square roots of the non-zero eigenvalues of $X'X$ and XX'.

2.3. Clustering Latent Coordinates

To classify the n firms into k non-overlapping groups, we consider a clustering scheme such that "similar" firms belong to the same group and "different" firms go into different groups. In this application, we use the latent coordinates of the firms in U as points in a plane (or some higher-dimensional space).

Given that U is a matrix of coordinates of n points in an r-dimensional space, the i-th row represents the coordinates of the i-th firm, while the j-th column of U represents the coordinates of the institutions on the j-th axis. For simplicity, we plot the first three dimensions of U as the default dimension of the latent positions. This correlates with most applications involving multi-dimensional scaling and provides a convenient framework to visualize the position of agents/firms in a 3-D space.

Typical clustering methods discussed in the literature range from centroid-based methods (k-means) to density-based, distribution-based, and hierarchical methods. In this application, we follow the centroid-based method of k-means clustering due to its simplicity, popularity, and successful application in market segmentation (see James et al. 2013). The choice of the number of clusters can be considered as a model selection problem since each cluster corresponds to a different statistical model (see Bai and Ng 2002; D'Angelo et al. 2023; Fraley and Raftery 2002; Handcock et al. 2007; Hoff et al. 2002).

3. Empirical Application

3.1. Data

We evaluate the effectiveness of our proposed model by considering a well-known dataset for P2P studies (Ahelegbey et al. 2019a, 2019b; Giudici et al. 2020). The dataset consists of 15,045 SMEs engaged in P2P lending on digital platforms across Southern Europe, with each institution containing 24 financial features in ratios constructed from official financial records in 2015. The data were obtained from the European External Credit Assessment Institution (ECAI).

Table 1 presents a description and summary of the financial ratios based on default status. In all, the data consist of 1632 (10.85%) defaulted institutions and 13,413 (89.15%) non-defaulted companies. Due to differences in the scale of the values of the sampled financial variables, we standardize each series to a zero mean and unit variance.

Table 1. Description and summary statistics of the financial ratios based on default status.

Var	Formula (Description)	Active (Mean)	Defaulted (Mean)
V1	(Total Assets − Shareholders Funds)/Shareholders Funds	8.87	9.08
V2	(Longterm debt + Loans)/Shareholders Funds	1.25	1.32
V3	Total Assets/Total Liabilities	1.51	1.07
V4	Current Assets/Current Liabilities	1.6	1.06
V5	(Current Assets − Current assets: stocks)/Current Liabilities	1.24	0.79
V6	(Shareholders Funds + Non current liabilities)/Fixed Assets	8.07	5.99
V7	EBIT/Interest paid	26.39	−2.75
V8	(Profit (loss) before tax + Interest paid)/Total Assets	0.05	−0.13
V9	P/L after tax/Shareholders Funds	0.02	−0.73
V10	Operating Revenues/Total Assets	1.38	1.27
V11	Sales/Total Assets	1.34	1.25
V12	Interest Paid/(Profit before taxes + Interest Paid)	0.21	0.08
V13	EBITDA/Interest Paid	40.91	5.71
V14	EBITDA/Operating Revenues	0.08	−0.12
V15	EBITDA/Sales	0.09	−0.12
V16	Constraint EBIT	0.13	0.56
V17	Constraint PL before tax	0.16	0.61
V18	Constraint Financial PL	0.93	0.98
V19	Constraint P/L for period	0.19	0.64
V20	Trade Payables/Operating Revenues	100.3	139.30
V21	Trade Receivables/Operating Revenues	67.59	147.12
V22	Inventories/Operating Revenues	90.99	134.93
V23	Total Revenue	3557	2083
V24	Industry Classification on NACE code	4566	4624
	Total number of institutions (%)	13,413 (89.15%)	1632 (10.85%)

Figure 1 presents a 3-D scatterplot of the SMEs' latent positions based on singular value decomposition of the observed features. The coordinates of defaulted SMEs are in red circles and the non-default SMEs are in green triangles.

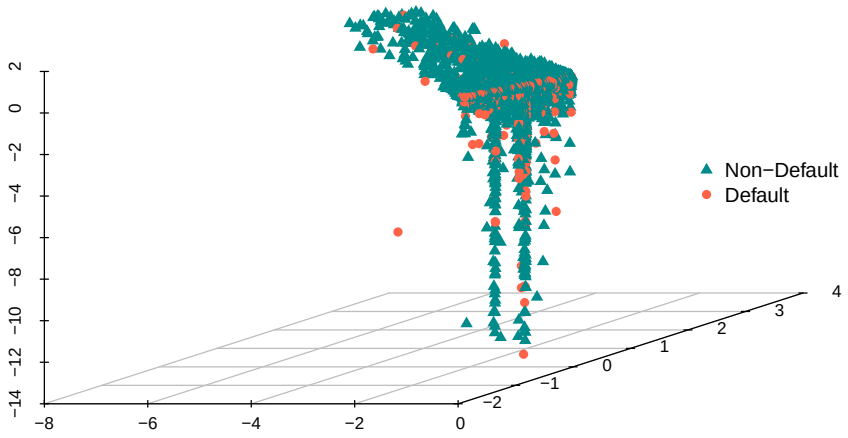

Figure 1. A 3-D scatterplot of the firm's latent positions based on singular value decomposition of observed features. Coordinates of defaulted SMEs are in red circles and non-default SMEs are in green triangles.

Table 2 shows the statistics of the number and percentage of the defaulted status of SMEs based on k-means clustering of the latent coordinates. From the 3-D plot in Figure 1 two clusters are evident. Thus, we choose $k = 2$. The table shows that 14,866 (98.81%) of

the SMEs are classified in Cluster 1, while the rest 179 (1.19%) are in Cluster 2. Of those in Cluster 1, 10.80% have defaulted and 89.20% have not. Of those in Cluster 2, 14.53% are defaulted SMEs, while 85.47% are not.

Table 2. Statistics of the defaulted status of SMEs according to k-mean clustering of the latent coordinates.

Status	Cluster 1		Cluster 2	
Default	1606	10.80%	26	14.53%
Non.Default	13,260	89.20%	153	85.47%
Total	14,866	98.81%	179	1.19%

3.2. Credit Score Modeling

Table 3 reports the estimated coefficients of the logistic regression for the full sample and the clustered samples. We remark that the results of the table are derived via a thorough activity of model selection, aimed at obtaining the best fit statistical model using stepwise logistic regression. The estimation of the models is carried out on the training sample which we set to be 70% of the sample. Given that 98.81% of the full sample is classified into Cluster 1, it is therefore not surprising that the credit score for the Full Sample and Cluster 1 have the same key drivers. However, we observe that for Cluster 2, the determinants of the credit score are somewhat different. There are however, some common drivers for Cluster 1 and 2, such as V3 (Total Assets/Total Liabilities), V4 (Current Assets/Current Liabilities), V14 (EBITDA/Operating Revenues), and V21 (Trade Receivables/Operating Revenues). Despite these common terms, the result shows that the majority of the key drivers of credit score for those in Cluster 1 are not significant determinants for those in Cluster 2.

Table 3. Stepwise logistic regression coefficients.

	Full Sample	Cluster (1)	Cluster (2)
V1	0.0022	0.0023	
V2			−0.1981 **
V3	−0.5779 ***	−0.5409 ***	−4.7532 **
V4	−0.2761 ***	−0.4121 ***	−0.5076 *
V5		0.1882	
V6	0.0023 *		
V7	0.0041 ***	0.0043 ***	
V8	−2.2462 ***	−2.2547 ***	
V9			−1.3393 **
V10	−0.3603 **	−0.2687 *	
V11	0.4119 ***	0.3458 **	2.8614
V12	0.1621 **	0.1610 **	
V13	−0.0023 **	−0.0024 **	
V14	−0.6840 ***	−0.7531 ***	5.0525 ***
V15			−3.9867 ***
V16	0.7136 ***	0.6889 ***	
V18	0.3775 *	0.4068 **	
V19	0.7888 ***	0.8021 ***	
V20		0.0007 **	
V21	0.0021 ***	0.0021 ***	0.0019 **
V22	0.0005 **	0.0005 *	0.0025
V23	−0.00002 ***	−0.00003 ***	
Constant	−2.2426 ***	−2.3916 ***	3.0007
Observations	12,035	11,892	143
Log Likelihood	−3167.4570	−3118.5110	−28.6924
Akaike Inf. Crit.	6370.9140	6275.0230	77.3849

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

3.3. Performance of Default Predicting Accuracy

We evaluate the default prediction accuracy of the estimated models on the testing sample and compare the performance in terms of the standard area under the curve (AUC) derived from the receiver operator characteristic (ROC) curve. The AUC depicts the true positive rate (TPR) against the false positive rate (FPR) depending on some threshold. TPR is the number of correct positive predictions divided by the total number of positives. FPR is the ratio of false positive predictions to overall negatives.

Table 4 shows the results of the area under the ROC curve of the full sample and clustered sample models. The result shows that the Full Sample model and Cluster 1 achieved an 82.34% prediction rate, while Cluster 2 reported a rate of 96.77%. However, the combined performance of the clustered sample model attains 82.62%. Thus, the clustered sample shows a slightly higher gain in predictive performance compared to the full sample approach.

Table 4. Comparing area under the ROC curve of the full sample and clustered sample models.

	Full Sample	**Cluster (1)**	**Cluster (2)**
AUC	0.8234	0.8234	0.9677
	Full Sample	Combine Cluster 1 and 2	
AUC	0.8234	0.8262	
	Statistic	*p*-value	Significance
DeLong test	−1.688	0.046	**

Note: ** $p < 0.05$.

Table 4 also reports the DeLong test of the pairwise comparison of the AUC of the full sample and that of the clustered sample models. The DeLong test is a statistical test for determining whether the AUCs of two models are significantly different (see DeLong et al. 1988). We conduct a one-sided DeLong test under the following hypothesis:

$$H_0: \text{AUC (Full sample)} \geq \text{AUC (Clustered sample)}$$
$$H_1: \text{AUC (Full sample)} < \text{AUC (Clustered sample)}$$

The results of the test as shown by the table indicates that the difference between the ROC of the clustered sample and the full sample leads to a DeLong test statistics equal to −1.688, with a corresponding *p*-value equal to 0.046, which indicates that the null hypotheses can be rejected and that the clustered sample achieves a performance significantly superior with respect to the full sample.

3.4. Implication of the Study

Market segmentation is the subdivision of markets into discrete groups that share similar characteristics. Such segmentation can be a useful scheme for credit risk measurement especially in systems like P2P, that are characterized by agents with heterogeneous characteristics. Market segmentation is simply a classification problem that involves assigning individuals in a population into discrete non-overlapping groups.

Our proposal could be useful to credit risk managers and investors involved in P2P lending. Because credit risk in P2P systems is not born by the platform but, rather by the investors, the ability of the latter to distinguish healthy and risky clients plays a crucial role in credit risk mitigation. Our study provides evidence that by clustering SMEs into two non-overlapping groups based on latent characteristics, investors, as well as risk managers, can identify and distinguish between common and dissimilar financial variables that drives the probability of default of a loan issued to a client.

4. Conclusions

This paper contributes to the strand of empirical studies to improve credit scoring for SMEs engaged in peer-to-peer platforms. We present a factor clustering-based approach to segment a heterogeneous population into groups with more homogeneous characteristics. The approach uses singular value decomposition to extract the factors underlying the observed financial performance ratios of SMEs. These factors are then classified into clusters via a k-mean algorithm. We then model the credit score for each sub-population via logistic regression.

The empirical application of our approach is demonstrated by applying our proposed methodology to evaluate the probability of default of 15,045 SMEs engaged in P2P lending across Europe. Our factor clustering approach to credit score modeling is shown to produce an efficient framework to analyze the latent positions of SMEs engaged in a P2P platform and provides a way to segment a heterogeneous market/customers into clusters with more homogeneous characteristics. The result shows that clusters with relatively fewer SMEs produce a more parsimonious and interpretable credit scoring model with better default predictive performance.

A limitation of our study could be related to our choice of clustering method and the number of clusters since such decisions strongly impact the results obtained. Thus, a more rigorous and robust clustering method can be advanced to improve the results of our study. Another possible limitation of this study is the application of latent-based clustering of SMEs instead of segmentation based on observed financial information. Future research can be advanced in these directions to improve the contribution of our study.

Author Contributions: Conceptualization, D.F.A. and P.G.; methodology, D.F.A. and P.G.; software, D.F.A.; validation, D.F.A.; formal analysis, D.F.A. and P.G.; investigation, D.F.A. and P.G.; resources, P.G.; data curation, D.F.A.; writing—original draft preparation, D.F.A.; writing—review and editing, P.G.; visualization, D.F.A.; supervision, P.G.; project administration, P.G.; funding acquisition, P.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Italian MUR PON, contract number n. 22-G-14923-1.

Data Availability Statement: Data is avalilable upon request to the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

Ahelegbey, Daniel Felix, Paolo Giudici, and Branka Hadji-Misheva. 2019a. Factorial Network Models To Improve P2P Credit Risk Management. *Frontiers in Artificial Intelligence* 2: 8. [CrossRef]

Ahelegbey, Daniel Felix, Paolo Giudici, and Branka Hadji-Misheva. 2019b. Latent Factor Models For Credit Scoring in P2P Systems. *Physica A: Statistical Mechanics and Its Applications* 522: 112–21. [CrossRef]

Andreeva, Galina, Jake Ansell, and Jonathan Crook. 2007. Modelling Profitability Using Survival Combination Scores. *European Journal of Operational Research* 183: 1537–49. [CrossRef]

Bai, Jushan, and Serena Ng. 2002. Determining the Number of Factors in Approximate Factor Models. *Econometrica* 70: 191–221. [CrossRef]

Barrios, Luis Javier Sánchez, Galina Andreeva, and Jake Ansell. 2014. Monetary and Relative Scorecards to Assess Profits in Consumer Revolving Credit. *Journal of the Operational Research Society* 65: 443–53. [CrossRef]

Chen, Xiao, Zhaohui Chong, Paolo Giudici, and Bihong Huang. 2022. Network Centrality Effects in Peer to Peer Lending. *Physica A: Statistical Mechanics and Its Applications* 600: 127546. [CrossRef]

D'Angelo, Silvia, Marco Alfò, and Michael Fop. 2023. Model-based Clustering for Multidimensional Social Networks. *Journal of the Royal Statistical Society Series A: Statistics in Society* 186: 481–507. [CrossRef]

DeLong, Elizabeth R., David M. DeLong, and Daniel L. Clarke-Pearson. 1988. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* 44: 837–45. [CrossRef]

Dungey, Mardi, and Dinesh Gajurel. 2015. Contagion and Banking Crisis–International Evidence for 2007–2009. *Journal of Banking and Finance* 60: 271–83. [CrossRef]

Dungey, Mardi, Renée Fry, Brenda González-Hermosillo, and Vance L. Martin. 2005. Empirical Modelling of Contagion: A Review of Methodologies. *Quantitative Finance* 5: 9–24. [CrossRef]

El Annas, Monir, Badreddine Benyacoub, and Mohamed Ouzineb. 2023. Semi-supervised Adapted HMMs for P2P Credit Scoring Systems with Reject Inference. *Computational Statistics* 38: 149–69. [CrossRef]

Emekter, Riza, Yanbin Tu, Benjamas Jirasakuldech, and Min Lu. 2015. Evaluating Credit Risk and Loan Performance in Online Peer-to-Peer (P2P) Lending. *Applied Economics* 47: 54–70. [CrossRef]
Forbes, Kristin J., and Roberto Rigobon. 2002. No Contagion, Only Interdependence: Measuring Stock Market Comovements. *The Journal of Finance* 57: 2223–61. [CrossRef]
Fox, Emily B., and David B. Dunson. 2015. Bayesian Nonparametric Covariance Regression. *The Journal of Machine Learning Research* 16: 2501–42.
Fraley, Chris, and Adrian E. Raftery. 2002. Model-based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association* 97: 611–31. [CrossRef]
Giudici, Paolo, and Gloria Polinesi. 2021. Crypto price discovery through correlation networks. *Annals of Operations Research* 1: 443–57. [CrossRef]
Giudici, Paolo, Branka Hadji-Misheva, and Alessandro Spelta. 2020. Network based credit risk models. *Quality Engineering* 2: 199–211. [CrossRef]
Giudici, Paolo, Thomas Leach, and Paolo Pagnottoni. 2022. Libra or Librae? Basket based stable coins. *Finance Research Letters* 44: 102504.
Haddad, Christian, and Lars Hornuf. 2019. The Emergence of the Global Fintech Market: Economic and Technological Determinants. *Small Business Economics* 53: 81–105. [CrossRef]
Handcock, Mark S., Adrian E. Raftery, and Jeremy M. Tantrum. 2007. Model-Based Clustering for Social Networks. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 170: 301–54. [CrossRef]
Hoff, Peter D. 2007. Model Averaging and Dimension Selection for the Singular Value Decomposition. *Journal of the American Statistical Association* 102: 674–85. [CrossRef]
Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock. 2002. Latent Space Approaches to Social Network Analysis. *Journal of the American Statistical Association* 97: 1090–98. [CrossRef]
James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning*. Berlin/Heidelberg: Springer, vol. 112.
Jiang, Cuiqing, Zhao Wang, Ruiya Wang, and Yong Ding. 2018. Loan Default Prediction by Combining Soft Information Extracted From Descriptive Text in Online Peer-to-Peer Lending. *Annals of Operations Research* 266: 511–29. [CrossRef]
Lopes, Hedibert Freitas, and Carlos Marinho Carvalho. 2007. Factor Stochastic Volatility with Time-Varying Loadings and Markov Switching Regimes. *Journal of Statistical Planning and Inference* 137: 3082–91. [CrossRef]
Lyócsa, Štefan, Petra Vašaničová, Branka Hadji Misheva, and Marko Dávid Vateha. 2022. Default or Profit Scoring Credit Systems? Evidence from European and US Peer-to-Peer Lending Markets. *Financial Innovation* 8: 32. [CrossRef]
Ma, Zhengwei, Wenjia Hou, and Dan Zhang. 2021. A credit Risk Assessment Model of Borrowers in P2P Lending Based on BP Neural Network. *PLoS ONE* 16: e0255216. [CrossRef]
Nakajima, Jouchi, and Mike West. 2013. Bayesian Analysis of Latent Threshold Dynamic Models. *Journal of Business and Economic Statistics* 31: 151–64. [CrossRef]
Puschmann, Thomas. 2017. Fintech. *Business & Information Systems Engineering* 59: 69–76.
Serrano-Cinca, Carlos, and Begoña Gutiérrez-Nieto. 2016. The Use of Profit Scoring as an Alternative to Credit Scoring Systems in Peer-to-Peer Lending. *Decision Support Systems* 89: 113–22. [CrossRef]
Xia, Yufei, Lingyun He, Yinguo Li, Nana Liu, and Yanlin Ding. 2020. Predicting Loan Default in Peer-to-Peer Lending Using Narrative Data. *Journal of Forecasting* 39: 260–80. [CrossRef]