Online Appendices for
Crowd, Lending, Machine, and Bias

A1. Summary Statistics Tables

Table A.1 Summary Statistics of the Full Dataset

	Full Sample	Funded Loans	Funded Loans with Matched Label	p-value
Sample size	247443	20668	19,529	-
Loan characteristics				
Listing amount	7879.14	6540.63	6623.36	0.15
Listing term	36	36	36	-
Monthly payment	290.97	234.00	236.86	0.16
Borrower Characteristics				
Stated monthly income	4181.11	4637.69	4664.11	0.51
Income verifiable	0.89	0.94	0.94	0.93
Debt to Income ratio	0.4980	0.3224	0.3250	0.78
Months employed	65.60	68.95	68.97	0.98
Employment status				
Full-time	0.79	0.86	0.86	0.76
Self-employed	0.09	0.08	0.08	0.79
Part-time	0.04	0.04	0.04	0.91
Retired	0.03	0.02	0.02	1.00
Not employed	0.02	0.01	0.01	0.88
Employed	0.03	0.00	0.00	-
Other	0.002	0.00	0.00	-
Has prior Prosper loans	0.05	0.11	0.11	0.62
Is homeowner	0.36	0.47	0.48	0.28
Is Prosper lender	0.11	0.25	0.25	0.49
Number of Public Records (last 10 years)	0.58	0.36	0.35	0.72
Number of Public Records (last 12 months)	0.06	0.03	0.03	0.67
Credit Characteristics				
Credit Grade				
AA	0.04	0.13	0.13	0.37
A	0.05	0.13	0.13	0.21
B	0.07	0.17	0.17	0.25

C	0.12	0.21	0.21	0.86
D	0.17	0.18	0.18	0.31
E	0.17	0.08	0.08	0.35
HR	0.38	0.08	0.08	0.11
ScoreX				
< 600	0.23	0.17	0.16	0.06
600-619	0.05	0.09	0.09	0.55
620-639	0.05	0.09	0.09	0.43
640-649	0.03	0.07	0.07	0.82
650-664	0.03	0.08	0.08	0.95
665-689	0.04	0.11	0.11	0.61
690-701	0.02	0.05	0.05	0.71
702-723	0.02	0.08	0.08	0.36
724-747	0.02	0.08	0.09	0.30
748-777	0.02	0.08	0.08	0.42
778+	0.02	0.08	0.08	0.54
Missing	0.46	0.00	0.00	-
Current credit lines	8.28	9.58	9.62	0.56
Open credit lines	7.29	8.22	8.25	0.62
Bank utilization	0.6297	0.5507	0.5462	0.23
Total open revolving accounts	5.96	6.36	6.38	0.68
Installment balance	28766.14	26158.08	26131.91	0.94
Real estate balance	97792.11	124228.62	126214.06	0.40
Revolving balance	17890.87	19061.02	19214.45	0.71
Total inquiries	11.51	9.04	9.04	1.00
Inquiries in last 6 months	3.95	2.60	2.59	0.92
Total trade items	25.88	24.60	24.64	0.82
Satisfactory accounts	19.18	20.26	20.35	0.52
Now delinquent derogatory	2.43	0.96	0.95	0.48
Was delinquent derogatory	4.27	3.38	3.34	0.38
Delinquencies over 30 days	9.45	6.67	6.60	0.52
Delinquencies over 60 days	4.37	2.75	2.72	0.49
Delinquencies over 90 days	8.74	4.93	4.86	0.51
Amount delinquent	3454.80	1100.30	1080.28	0.75
Length of credit history	4809.67	4907.40	4919.86	0.64
Outcomes				
Default		0.32	0.32	0.36
Principal Paid		4898.07	4936.77	0.44
Interest Paid		1341.93	1352.17	0.49

Table A.2 Summary Statistics for the Gender Groups

	All Listings			Funded Loans		
	Female	Male	p-value	Female	Male	p-value
Sample Size	15313	18187	-	1888	2909	-
Loan Characteristics						
Listing amount	6050.87	7734.51	<.01	5282.03	6981.21	<.01
Listing term	36	36	-	36	36	-
Monthly Payment	233.67	291.28	<.01	192.28	246.94	<.01
Borrower Characteristics						
Stated monthly income	3101.24	5174.16	<.01	3306.27	5197.60	<.01
Income verifiable	0.94	0.90	<.01	0.97	0.97	0.51
Debt to Income ratio	0.39	0.31	<.01	0.34	0.26	0.06
Months Employed	63.75	77.12	<.01	65.14	78.13	<.01
Employment status						
Full-time	0.90	0.85	<.01	0.93	0.93	0.94
Self-employed	0.02	0.09	<.01	0.01	0.05	<.01
Part-time	0.05	0.01	<.01	0.06	0.02	<.01
Retired	0.00	0.00	<.01	0.00	0.00	0.77
Not Employed	0.00	0.00	<.01	0.00	-	-
Employed	0.03	0.04	<.01	0.00	0.00	-
Other	0.00	0.00	0.54	0.00	0.00	-
Has prior prosper loans	0.09	0.11	<.01	0.10	0.12	0.01
Is homeowner	0.33	0.43	<.01	0.40	0.54	<.01
Is Prosper Lender	0.10	0.19	<.01	0.15	0.32	<.01
No. of Public Records (last 10 years)	0.57	0.51	<.01	0.43	0.33	<.01
No. of Public Records (last 12 months)	0.06	0.05	<.01	0.04	0.03	<.01
Credit Characteristics						
Credit Grade						
AA	0.03	0.06	<.01	0.08	0.16	<.01
A	0.04	0.07	<.01	0.09	0.14	<.01
B	0.07	0.10	<.01	0.14	0.18	<.01
C	0.13	0.16	<.01	0.22	0.21	0.58
D	0.19	0.19	0.31	0.23	0.16	<.01
E	0.18	0.15	<.01	0.11	0.08	<.01
HR	0.36	0.27	<.01	0.13	0.07	<.01
ScoreX						
< 600	0.54	0.42	<.01	0.23	0.15	<.01
600-619	0.10	0.09	<.01	0.12	0.07	<.01
620-639	0.09	0.10	0.13	0.11	0.09	0.02
640-649	0.05	0.05	0.54	0.08	0.06	0.05

650-664	0.05	0.06	<.01	0.09	0.09	0.83
665-689	0.05	0.08	<.01	0.10	0.12	0.11
690-701	0.02	0.03	<.01	0.05	0.05	0.31
702-723	0.03	0.05	<.01	0.07	0.08	0.01
724-747	0.03	0.04	<.01	0.06	0.09	<.01
748-777	0.02	0.04	<.01	0.06	0.09	<.01
778+	0.02	0.04	<.01	0.04	0.10	<.01
Current credit lines	8.51	8.77	<.01	9.28	9.74	<.01
Open credit lines	7.47	7.71	<.01	7.82	8.34	<.01
Bank utilization	0.65	0.60	<.01	0.58	0.53	<.01
Total open revolving accounts	5.73	5.53	<.01	6.20	6.11	0.50
Installment balance	25551.62	28169.07	<.01	21708.11	24723.79	<.01
Real estate balance	62000.01	99735.11	<.01	80476.52	133178.85	<.01
Revolving balance	11770.23	16468.26	<.01	12104.01	17943.58	<.01
Total inquiries	11.70	11.86	0.18	9.15	9.13	0.96
Inquiries in last 6 months	3.48	3.50	0.75	2.70	2.60	0.34
Total trade items	26.48	24.92	<.01	24.53	24.54	0.97
Satisfactory accounts	18.29	18.66	0.01	19.08	20.59	<.01
New delinquent derogatory	3.18	2.28	<.01	1.34	0.78	<.01
Was delinquent derogatory	5.00	3.99	<.01	4.11	3.18	<.01
Delinquencies over 30 days	11.37	8.76	<.01	8.46	6.01	<.01
Delinquencies over 60 days	5.55	3.95	<.01	3.60	2.27	<.01
Delinquencies over 90 days	11.10	7.51	<.01	6.09	3.98	<.01
Amount delinquent	2861.22	2467.65	<.01	1194.10	921.99	0.07
Length of credit history	4920.65	4598.94	<.01	4962.55	4734.01	<.01
Outcomes						
Default				0.36	0.27	<.01
Principal Paid				3879.56	5499.38	<.01
Interest Paid				1214.09	1355.24	<.01

Table A.3 Summary Statistics for the Race Groups

	All Listings			Funded Loans		
	Black	Non-Black	p-value	Black	Non-Black	p-value
Sample size	47866	8327	-	5831	1885	-
Loan Characteristics						
Listing amount	7085.72	7235.03	0.04	6483.98	6387.55	0.53
Listing term	36	36	-	36	36	-
Monthly Payment	270.11	277.74	<.01	231.80	230.89	0.87
Borrower Characteristics						
Stated monthly income	4470.06	3547.65	<.01	4615.89	4365.13	0.01
Income verifiable	0.87	0.86	<.01	0.95	0.94	0.08
Debt to Income ratio	0.34	0.38	<.01	0.30	0.31	0.44
Months Employed	67.63	73.53	<.01	70.47	74.06	0.09
Employment status						
Full-time	0.80	0.77	<.01	0.87	0.84	<.01
Self-employed	0.09	0.11	<.01	0.07	0.08	0.02
Part-time	0.03	0.04	0.10	0.03	0.04	0.01
Retired	0.03	0.04	<.01	0.02	0.02	0.27
Not Employed	0.02	0.02	0.26	0.01	0.01	0.33
Employed	0.03	0.03	0.51	0.00	0.00	-
Other	0.00	0.00	0.06	0.00	0.00	-
Has prior prosper loans	0.10	0.09	0.27	0.12	0.12	0.61
Is homeowner	0.41	0.44	<.01	0.52	0.51	0.41
Is Prosper Lender	0.14	0.13	0.04	0.24	0.23	0.16
Number of Public Records (last 10 years)	0.52	0.63	<.01	0.33	0.47	<.01
Number of Public Records (last 12 months)	0.05	0.07	<.01	0.03	0.04	<.01
Credit Characteristics						
Credit Grade						
AA	0.05	0.04	0.23	0.13	0.11	<.01
A	0.06	0.06	0.84	0.14	0.12	0.05
B	0.09	0.09	0.05	0.18	0.18	0.78
C	0.16	0.16	0.31	0.23	0.20	0.02
D	0.19	0.18	0.24	0.17	0.19	0.10
E	0.15	0.17	<.01	0.08	0.10	<.01
HR	0.30	0.30	0.89	0.08	0.10	<.01
ScoreX						
< 600	0.45	0.47	<.01	0.16	0.20	<.01
600-619	0.09	0.09	0.88	0.08	0.09	0.17
620-639	0.10	0.09	0.07	0.09	0.09	0.40
640-649	0.06	0.05	0.02	0.07	0.07	0.72
650-664	0.06	0.07	0.55	0.09	0.07	0.02

665-689	0.07	0.07	0.53	0.12	0.12	0.99
690-701	0.03	0.02	0.02	0.05	0.05	0.91
702-723	0.04	0.04	0.54	0.08	0.08	0.41
724-747	0.04	0.04	0.17	0.09	0.07	0.01
748-777	0.03	0.03	0.31	0.08	0.07	0.10
778+	0.03	0.03	0.53	0.08	0.07	0.17
Current credit lines	8.75	8.35	<.01	9.79	9.19	<.01
Open credit lines	7.73	7.37	<.01	8.40	7.88	<.01
Bank utilization	0.62	0.62	0.58	0.55	0.57	0.08
Total open revolving accounts	5.77	5.48	<.01	6.40	5.98	<.01
Installment balance	29660.53	25661.92	<.01	26912.95	24108.01	<.01
Real estate balance	89033.60	86536.73	0.26	116340.17	116623.84	0.96
Revolving balance	16145.69	15798.75	0.41	17613.95	18805.47	0.27
Total inquiries	10.81	10.40	<.01	8.36	9.44	<.01
Inquiries in last 6 months	3.09	2.98	0.01	2.28	2.68	<.01
Total trade items	26.38	25.98	0.01	25.27	25.13	0.71
Satisfactory accounts	19.31	18.55	<.01	20.93	20.05	0.01
New delinquent derogatory	2.75	2.69	0.21	0.97	1.21	<.01
Was delinquent derogatory	4.32	4.74	<.01	3.37	3.86	<.01
Delinquencies over 30 days	9.83	10.74	<.01	6.93	7.93	<.01
Delinquencies over 60 days	4.57	4.89	<.01	2.81	3.26	<.01
Delinquencies over 90 days	9.28	9.66	0.04	4.99	5.79	0.01
Amount delinquent	3123.89	3018.72	0.40	1124.25	1512.51	0.03
Length of credit history	4938.47	5041.86	<.01	4954.46	5114.79	0.02
Outcomes						
Default				0.33	0.34	0.35
Principal Paid				4790.35	4700.74	0.50
Interest Paid				1338.45	1412.68	0.07

A2. XGBoost Model

We use an XGBoost model (Chen and Guestrin, 2016) to fit a prediction function $\phi(X)$ that outputs the predicted probability of default $P(Y = 1|X)$. XGBoost is a scalable tree based boosting system that achieves state-of-the-art results in many ML challenges. An XGBoost model is a tree ensemble model that consists of multiple regression trees (also known as CARTs). Unlike normal decision trees (classification trees) that output class labels or use the portion of positive classes as the class probability, a regression tree performs a regression in each leaf node. In our case, the logistic regression is used and each tree outputs the probability of default. The final prediction of an XGBoost model is the sum of the predictions from each regression tree. Formally, the prediction for instance i with feature vector X_i is as follows:

$$\widehat{Y_\iota} = \phi(X_i) = \sum_{k=1}^{K} f_k(X_i), \quad f_k \in \mathrm{F}$$

where K is the number of regression trees, f_k is the mapping function of the k-th regression tree, and F is the space of regression trees.

To learn a model, we minimize an objective function that consists of the training loss and regularization term as follows:

$$L(\phi) = \sum_i l(\widehat{Y_\iota}, Y_i) + \sum_k \Omega\,(f_k)$$
$$l\left(\widehat{Y_\iota}, Y_i\right) = Y_i \ln\left(1 + e^{-\widehat{Y_\iota}}\right) + (1 - Y_i)\ln\left(1 + e^{\widehat{Y_\iota}}\right)$$
$$\Omega(f) = \gamma T + \frac{1}{2}\lambda\|w\|_1$$

where T is the number of leaves in a tree, and w is the vector of the leaf weights. γ and λ are hyperparameters. Intuitively, we aim to balance the prediction accuracy and the model's simplicity since we minimize the sum of the training errors and the model's complexity. The above objective function cannot be optimized using traditional optimization techniques, and the model is additively and greedily trained by learning a tree and adding it to the model in one iteration. We do this greedy search in each iteration until we finish training all K trees.

Tree based models are vulnerable to overfitting problems. In practice, we use several techniques to alleviate it. First, when fitting regression trees, we can specify a maximum depth so that a tree stops growing once it reaches the depth. Second, we can subsample instances and/or features and therefore create slightly different datasets for each tree. Third, after we learn a tree f_t, we usually shrink it when adding it to the model.

There are several hyperparameters in an XGBoost model: the number of trees K, the regularization terms γ and λ, the learning rate ϵ, the maximum tree depth, the instance subsample percentage and the feature subsample percentage. We tune these parameters using five-fold cross-validation on the training set. Because our hyperparameter space is relatively large and fitting the XGBoost model is computationally expensive, we choose the Bayesian Optimization instead of the more commonly used grid search or random search to search for the optimal hyperparameters. Bayesian Optimization balances the exploration-exploitation trade-off, which helps avoid getting stuck in the local minimum. In general, it is more efficient than the grid search or random search and enables us to find a near optimal set of hyperparameters with fewer trials.

Bayesian Optimization is a common method that is used for tuning hyperparameter in ML models. Hyperparameter searching is a maximization process in which we aim to find a set of parameters V that maximize our objective $h(V)$ (e.g., the cross-validated model accuracy score). The challenge is that $h(V)$ is a black box function since we do not know its explicit expression or its derivatives. The evaluation is restricted to getting a response value with a certain input. Grid search and random search enumerate large sets of possible input values and independently evaluate the responses in each iteration while Bayesian Optimization learns the shape of the objective $h(V)$ from the responses over the iterations. It starts from a Gaussian Process prior that is characterized by a mean function μ and a covariance function k as follows:

$$\hat{h}_0(V) \sim GP(\mu(v), k(v, v'))$$

In each iteration, the algorithm samples a data point ($V_i, h(V_i)$), adds it to the observed data set $D_{1:t}$, and produces an updated posterior function using Bayes theorem as follows:

$$P(D_{1:t}) \propto P\left(\hat{h}_{t-1}\right) P(\hat{h}_{t-1})$$

The algorithm then samples a new input V that corresponds to a potentially high value of h(V) based on its current posterior function. This process is repeated until the preset number of inputs have been sampled. Then, the algorithm tries to find the optimal point of the final posterior function as the final optimal values.

A3. Generating ROC Curves

The AUC-ROC is a common evaluation metric that is good for measuring predicted score performance when Y is unbalanced, as in our case where there are more completed loans ($Y = 0$) than default loans ($Y = 1$). To calculate the accuracy, we need a binary predicted label, and it is commonly derived by setting threshold for the predicted score as follows:

$$\hat{Y}(t) = 1 \; iff \; s \geq t$$

where s is the predicted score (in our case, $\phi(X|R^* = 1)$ or $p(X, Z \,|R^* = 1)$), and t is the discrimination threshold. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) as t varies. When $t > max\ (s)$, no loan is classified as being in default, and thus both TPR and FPR are 0. When $t \leq (s)$, all loans are classified as being in default, and thus both TPR and FPR are 1. When the threshold is in between, the TPR and FPR take different values and result in an ROC curve. The AUC-ROC is simply the area under this curve. It also has an intuitive interpretation as the probability that a randomly selected positive instance is ranked higher than a randomly selected negative instance using their predicted score s. The ROC of a random guess (no predictive power) is a diagonal line that goes through (0, 0) and (1, 1), and it has a corresponding AUC of 0.5. The ROC of a perfect prediction would go through the perfect classification point (0, 1), and its AUC would be 1. Most predictive models are between the two extreme cases, and therefore have an AUC that ranges from 0.5 and 1. Figure 1 plots the ROC curves of the machine predictions $\phi(X|R^* = 1)$ and the crowd predictions $p(X, Z \,|R^* = 1)$ on the holdout test set of 7812 loans. We can see that the curve of the machine predictions strictly dominates the curve of the crowd predictions with a higher AUC of 0.7406 compared to 0.6783. This suggests that, overall, our ML model has more predictive power than the crowd.

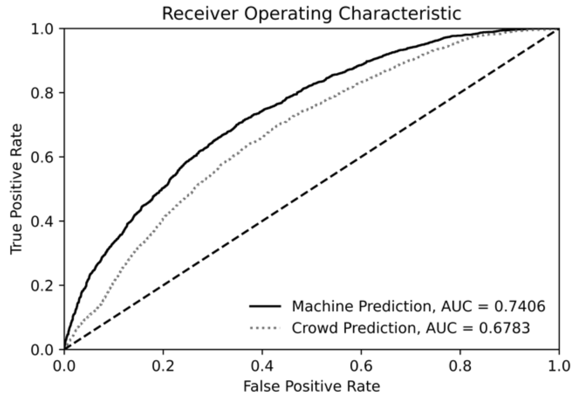

Figure A.1 ROC Curves and AUC

A4. Fixed Risk-Free Rate Model

In the main analysis of comparison (1), we use risk premium as a proxy for the crowd-predicted loan default risk and assume a time-varying risk-free rate when calculating the risk premium. The rationale is that the final interest rate has two sources of variation: risk-free rate and default risk premium. Since we are interested in the crowd's assessment of the loan default risk, it is important to remove the variation in the risk-free rate from the final inters rate. However, even though the risk-free rate fluctuated daily, the variation was small, and investors may consider a fixed risk-free rate when making investment decision during a relatively short period. In this case, the risk-free rate is absorbed into the constant term, and the interest rate formula can be re-written as:

$$r(X, Z|R^* = 1) = g\big(m(X, Z|R^* = 1)\big) + c$$

This means that the only varying part of interest rate is default risk premium, and interest rate itself preserves the ordinal information in the crowd predicted default risk. Therefore, under the constant risk-free rate assumption, we can use interest rate as a proxy for the crowd prediction, and the results based on this alternative proxy are shown in Figure A.2. When assuming a fixed risk-free rate, the crowd prediction accuracy slightly increases from 0.6840 to 0.6918, but the change is minimal and our results for the main analysis (that the machine produces more accurate predictions and leads to better investment decisions) still hold.

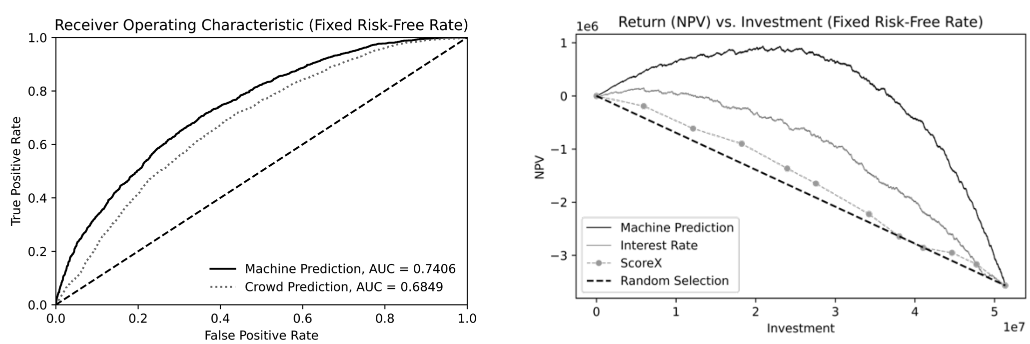

Figure A.2 ROC Curves and Return (NPV) - Investment Plot with Fixed Risk-Free Rate

A5. Multiple Time Periods

In the time period of our main analysis (March 2007 to October 2008), online P2P lending platforms were considered new and becoming more and more popular, and during this period, they may have attracted borrowers and lenders who were less rational and just wanted to "give it a try". While the machine has the only goal of accurately identifying the default risks and making profitable investments, some of the crowd investors might be exploring the platform and trying things out. To examine the potential shift in lenders' and borrowers' behavior during the growth period and mitigate the concern that our results are driven by the difference in the crowd investors' objective(s) and the machine's objective, we divide our sample into 4 time periods with equal length (148 days). The period-specific summary statistics are reported in Table A.4.

Table A.4 Summary Statistics of the Loans by Time Period

Period	Start Date	End Date	# of Loans	Default Rate	Avg. Interest Rate
1	2007-03-01	2007-07-26	4871	0.3631	0.1720
2	2007-07-27	2007-12-22	4142	0.3421	0.1714
3	2007-12-22	2008-05-18	5336	0.3024	0.1744
4	2008-05-18	2008-10-12	5180	0.2786	0.1910

We perform the same analysis for the loans in each period as in the main analysis: split the loans into a training set and a test set, train a XGBoost model using the training data, use the model to predict default risk for loans in the test set, use risk premium as the proxy for the risk prediction of the crowd, and compare the accuracy of the predictions made by the machine and by the crowd and the decisions based on the machine and the crowd predictions. Figure A.3 and Figure A.4 show the ROC curves and the Return (NPV)-Investment plots in the four periods for comparison (1), respectively. We do not perform comparison (2) on each of the periods because the number of crowd in each period is too small.

The number of loans remained roughly the same in each period, while the default rate steadily decreased over the time. The average interest rate was stable for the first three periods and had a small jump in the last period, which might be caused by the financial crisis which started to outburst at that time. This suggests that either the crowd were getting better at identifying risky loans and becoming selective over time, or the quality of the pool of potential loans was improving. Meanwhile, the decreasing default rate means fewer

Figure A.3 ROC Curves in Different Time Periods

positive labels (default loans), and therefore it is increasingly difficult for the machine to learn the patterns of defaulting behavior.

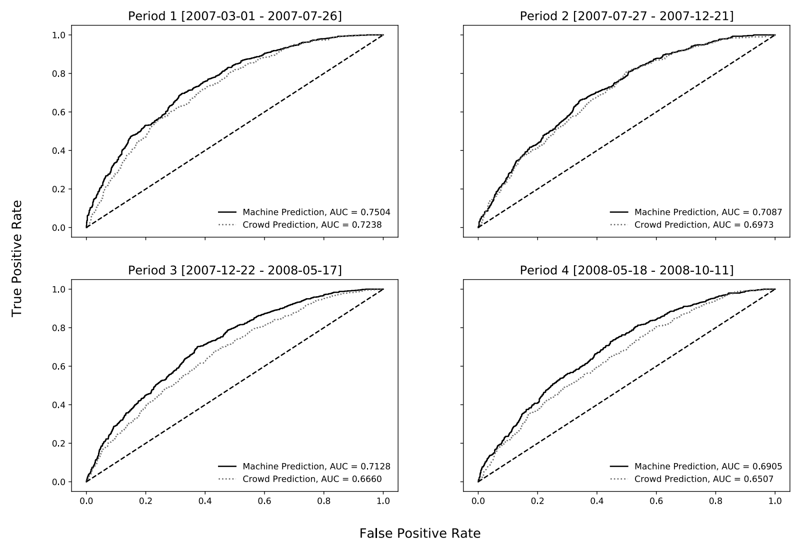

Figure A.3 ROC Curves in Different Time Periods

Figure A.3 and Figure A.4 show consistent results. The NPV of the loans that have been invested (in the test set) increases over time (compare the value on the vertical axis for right most point in each of the 4 plots), suggesting that the investors were improving their decisions. However, it becomes harder and harder to further identify risky loans from the selected set, and the prediction accuracy of machine prediction and crowd prediction both decline over time. It is still clear that, in all the periods, the machine makes more accurate predictions and leads to substantially higher returns. Moreover, we can see from the plots that as the crowd improved their investment decisions, the level of additional welfare gain that machine provides also increases. This may suggest that while the crowd learned to pick up certain signals to make better predictions over time, the machine may be able to capture patterns that the crowd had ignored, and therefore could generate more benefits, especially when it is difficult to identify defaulters from non-defaulters. Although Prosper might attract investors with reduced rationality or different objectives during its growth period, the overall quality of investment decisions made by the crowd noticeably improved. On top of it, the machine learning model still makes better predictions and provides an increasing level of improvement over time.

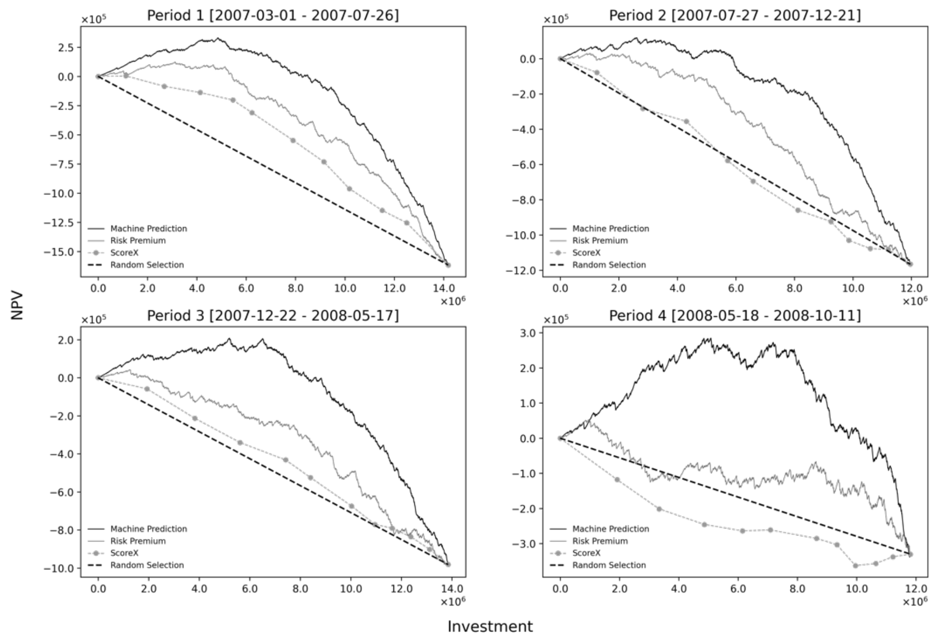

Figure A.4 Return (NPV)-Investment Plots in Different Time Periods

A6. Sample Portfolios by Investment Amount

Table A.5 Portfolios Comparison, By Investment Amount

	Using Machine Prediction		Using Risk Premium	
Investments (in millions)	NPV	IRR	NPV	IRR
1.00	75,614.92	0.0768	20,688.07	0.0416
2.50	207,815.02	0.0817	88,477.92	0.0514
5.00	366,299.27	0.0757	196,443.86	0.0540
10.00	653,039.22	0.0710	70,290.54	0.0328
15.00	791,477.63	0.0633	-185,288.88	0.0195
20.00	881,261.76	0.0577	-418,054.24	0.0134
25.00	875,970.71	0.0518	-675,057.11	0.0090
30.00	625,032.66	0.0423	-1,091,873.18	0.0022
35.00	170,775.40	0.0314	-1,575,215.47	-0.0042
40.00	-423,224.07	0.0205	-2,091,709.89	-0.0096
45.00	-1,433,494.00	0.0052	-2,752,343.99	-0.0164
50.00	-3,050,957.47	-0.0165	-3,409,180.78	-0.0219
51.34	-3,567,424.43	-0.0230	-3,567,424.4	-0.0230

A7. A Rolling-Window Analysis

In the main analysis, we randomly split all the loans in the market portfolio into training and test sets and use the training set to train the machine learning model. One concern is that the machine sees the outcomes of "future loans" when making predictions for the loans in the test set, while the crowd had only access to the past information, thus the machine vs crowd comparison is not fair. To address this concern, we build a dynamic rolling-window machine learning model. Instead of randomly splitting the loans into the training and test sets, we predict default risk for each loan in the test set using the loans that started within 180 days (6 months) prior to the target loan's creation date as the training data. In this way, we ensure that the machine only has access to the past information – loans that were funded in the past 180 days – when making predictions for each loan. Such a rolling-window model also has practical value, as it is dynamically adapted when new data becomes available and therefore remains relevant for new predictions.

Since this model requires sufficient past loans as training data, we only make predictions for loans that were funded after August 2007. Figures A.5 and A.6 show the results of Comparison (1) and Comparison (2) for this rolling-window model. It suggests that our main results continue to hold under this model.

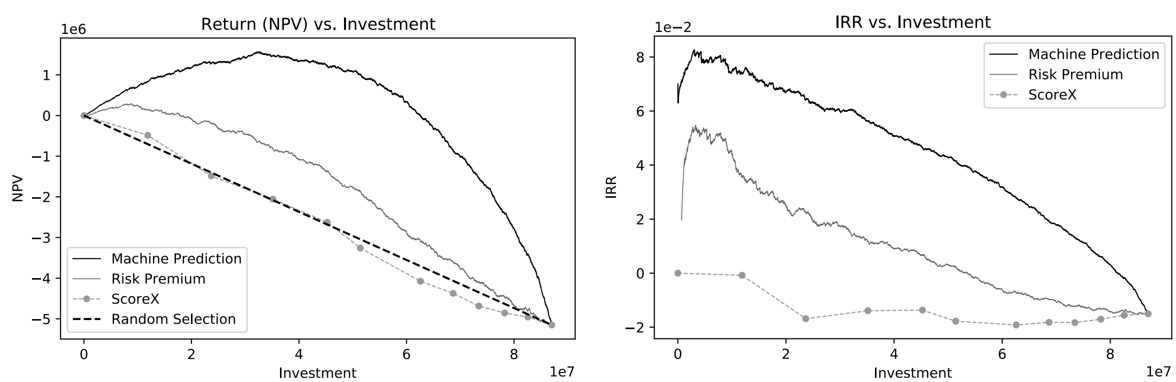

Figure A.5 Return-Investment Plot for Different Portfolios

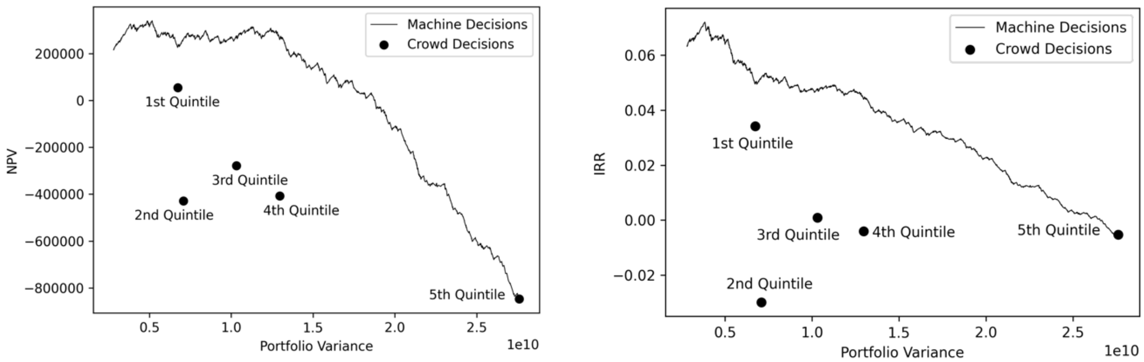

Figure A.6 Return-Variance Plot for Different Portfolios

A8. Loans Binned Based on Risk Premium

In the main analysis of Comparison (2), we divide all the funded loans into 100 quantile bins based on the machine predicted risk, and use the percentage of defaulted loans in a bin as an estimate of the true default probability for the loans in that bin. In this robustness check, we define the 100 quantile bins based on the risk premium, and redo comparison (2). Figures A.7 and A.8 show the results for the random assignment check and those for the comparison. They correspond to Figures 3 and 4 in the body of the paper. As we can see from the figures, the random assignment condition still holds, and the comparison results are similar to those presented in the body of the paper.

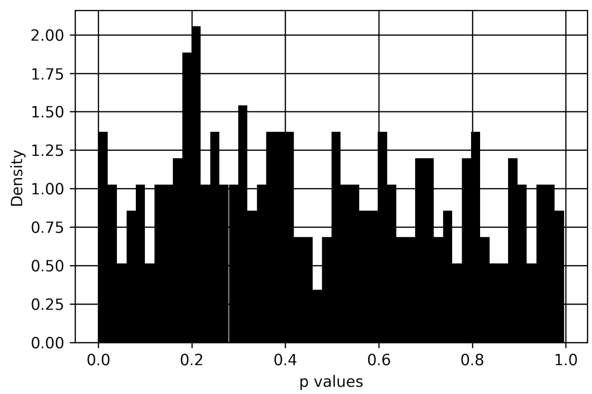

Figure A.7 Distribution of the p-values in the F-tests

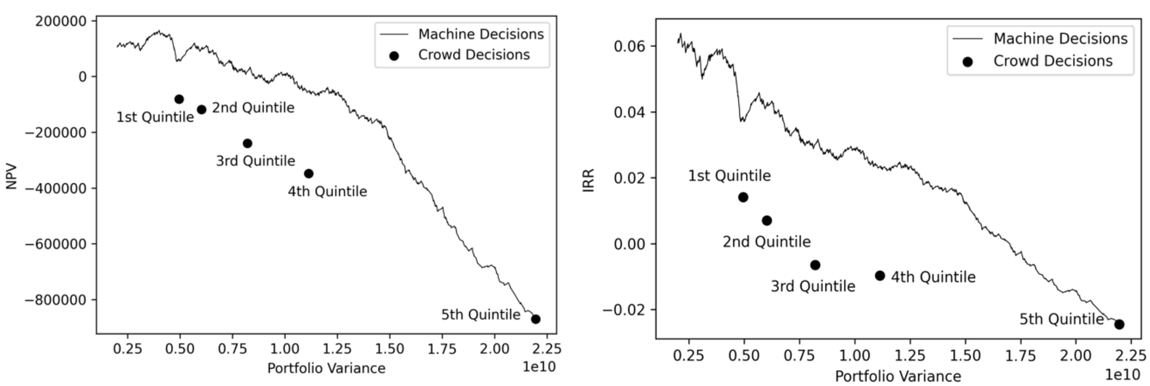

Figure A.8 Return-Variance Plot for Different Portfolios

A9. Debiasing Method

Continuous Variable

When X^j is a continuous variable, we first estimate the probability density function (PDF) of X^j from the data x_i^j using the kernel density estimation for each demographic group as follows:

$$\hat{f}_h(x|A=a) = \frac{1}{n_a h}\sum_{i=1}^{n} K(\frac{x-x_i^j}{h}), i \in A_a, \qquad a \in \{0,1\}$$

where n_a is the number of observations in each group, K is the kernel function that we set to be a standard normal density function, and h is a hyperparameter called the kernel bandwidth that is determined using Scott's Rule as $h = n_a^{-\frac{1}{5}}$. Let $\hat{F}_h(x^j)$ be the corresponding cumulative distribution function (CDF) as follows:

$$\hat{F}_h(x|A=a) = \int_{-\infty}^{x^j} \hat{f}_h(t|A=a)dt = P\left(x^j \le x \middle| A=a\right), a \in \{0,1\}$$

Then, we define a new random variable $\tilde{X}^j$ as follows:

$$\tilde{X}^j(A) = \hat{F}_h\left(X^j|A\right)$$

Each data point $\tilde{x}_i^j$, as a realization of $\tilde{X}^j$, is simply calculated as follows:

$$\tilde{x}_i^j = \hat{F}_h\left(x_i^j|A_i\right)$$

Now we prove that $\tilde{X}^j \perp A$:

$$P\left(\tilde{X}^j \le x|A=0\right) = P\left(\hat{F}_h\left(X^j|A=0\right) \le x\right)$$

$$= P\left(X^j \le \hat{F}_h^{-1}(x|A=0)|A=0\right)$$

$$= x, \forall x \in [0,1]$$

Similarly:

$$P\left(\tilde{X}^j \le x|A=1\right) = x, \forall x \in [0,1]$$

Thus:

$$P(\tilde{x}^j \leq x|A = 0) = P(\tilde{x}^j \leq x|A = 1), \forall x \in [0,1]$$

I.e., $\tilde{x}^j \perp A$.

Categorical (Nominal) Variable

When X^j is a categorical variable that takes k values, the conditional distribution is characterized by the following:

$$P(X^j = t \middle| A = a) = p_{at}, t = 1, \dots, k, \qquad a \in \{0, 1\}$$

We define a new discrete random variable $\tilde{X}^j$ that takes $2k$ values:

$$P(\tilde{X}^j = s \middle| A = 0) = P(\tilde{X}^j = s \middle| A = 1) = \alpha_s, s = 1, \dots, 2k$$

The mapping function $X^j = \sigma(\tilde{X}^j, A)$ is defined as follows:

$$\begin{cases} \sigma(\tilde{X}^j = t, A = 0) = t, t = 1, \dots, k \\ \sigma(\tilde{X}^j = t + k, A = 0) = t, t = 1, \dots, k \\ \sigma(\tilde{X}^j = 2t - 1, A = 1) = t, t = 1, \dots, k \\ \sigma(\tilde{X}^j = 2t, A = 1) = t, t = 1, \dots, k \end{cases}$$

We then solve for $\alpha_s (s = 1, \dots, 2k)$ that satisfies the following conditions:

$$\begin{cases} \alpha_t + \alpha_{t+k} = p_{0t}, t = 1, \dots, k \\ \alpha_{2t-1} + \alpha_{2t} = p_{1t}, t = 1, \dots, k \\ \alpha_s \geq 0, s = 1, \dots, 2k \end{cases}$$

The system of equations is guaranteed to have solution when $k = 2$, but not for more categorical values. When the system is not solvable, we can either choose to solve the equivalent least squares problem[1] or group some of the categories to decrease the k.

$\tilde{X}^j$ is by design independent of A, and each data point $\tilde{x}_i^j$, as a realization of $\tilde{X}^j$, is calculated as follows:

$$\tilde{x}_i^j = \sigma^{-1}(x_i^j, A)$$

[1] In that case, $\tilde{x}^j$ may be weakly dependent of A. The bias can be reduced but not removed.

Ordered Categorical (Ordinal) Variable

When X^j is an ordered categorical variable, we do not want to treat it as a nominal variable since it would lose the ordinal information during the process. Instead, we first convert the data into a continuous variable by randomly sampling value from a corresponding range as follows:

$$\hat{x}_i^j \sim unif(l_t, h_t), if\ x_i^j = t$$

where l_t and h_t are respectively the low and high values corresponding to the level t. It can be decided based on the nature of X^j. For example, if X^j is discretized bins, l_t and h_t can be the natural bin ranges. If X^j only takes a few sequential values, we can set the following:

$$l_t = \frac{2t-1}{2}, h_t = \frac{2t+1}{2}$$

Afterwards, we use the debiasing method for continuous variables that was described before to create $\tilde{X}^j$ from $\hat{X}^j$.

Mixed Variable

Some variables are continuous by nature but may have a point mass at a certain value due to missing data or censoring. For example, outstanding balance of prior prosper loans should be considered as a continuous variable, but because most the borrowers in our dataset do not have prior prosper loans, this field is 0 for them. We call such a variable a 'mixed variable' since it is a mix of continuous variables and binary variables. None of the methods described above can effectively create a new variable that is independent of A from a mixed variable, and we need a special way to handle it.

Let X^j be a mixed variable that has a point mass at a and a continuous distribution elsewhere. We define a binary variable u as follows:

$$u = 1 \text{ iff } X^j = a$$

Let $\tilde{u}$ be a debiased variable that we created from u, which is constructed to be a nominal variable. Therefore, $\tilde{u} \perp A$. Let $\widetilde{w}$ be a debiased variable that we created from $X^j | u = 0$, which is constructed to be a continuous variable. It follows that $\widetilde{w} \perp A | u = 0$.

Now, we define variable $\tilde{X}^j$ as follows:

$$\tilde{X}^j(\tilde{u}, A) = \begin{cases} \widetilde{w}(\tilde{u}, A), & if \sigma(\tilde{u}, A) = 0 \\ \widetilde{w}(\tilde{u}, 1 - A), & o.w. \end{cases}$$

By construction, $\tilde{X}^j \perp A \mid \tilde{u}$, and thus $P\left(A \middle| \tilde{X}^j, \tilde{u}\right) = P(A | \tilde{u})$. Since $\tilde{u} \perp A$, $P(A | \tilde{u}) = P(A)$. Therefore, we have $P\left(A \middle| \tilde{X}^j, \tilde{u}\right) = P(A)$, i.e., $\mathrm{A} \perp \tilde{\mathrm{X}}^j, \tilde{u}$, which implies that $\tilde{X}^j \perp A$.

A10. Robust Check Results for Fairness Check

In the main analysis, we assign loans that were requested by borrowers who were in high/low female concentrated occupations (female percentage greater than 75%/less than 25%) into the female/male group; we assign loans that were requested by borrowers who were in high/low black concentrated locations (black percentage greater than 75%/less than 25%) into the black/non-black group. The following table shows the results when the cutoff is set to 85% and 15%.

Table A.7 Fairness Checks

Panel A: Fairness Check for the Original XGBoost Model						
	Gender Groups			Race Groups		
Number of Observations	**1918**			**3087**		
AUC	0.7043			0.7348		
	Female	Male	Mean Difference	Black	Non-Black	Mean Difference
Prob. Of Being Funded	0.6054	0.6919	-0.0865**	0.7019	0.5898	0.1121***
True Positive Rate	0.7021	0.7746	-0.0725*	0.8041	0.6747	0.1294***
False Positive Rate	0.4340	0.4936	-0.0596	0.4873	0.4194	0.0680
Average Score of Positive Class	0.6934	0.7245	-0.0311**	0.7508	0.7078	0.0430***
Average Score of Negative Class	0.5458	0.5964	-0.0506*	0.5959	0.5514	0.0445*
Panel B: Fairness Check for the De-Biased Model						
	Gender Groups			Race Groups		
Number of Observations	**1918**			**3087**		
AUC	0.6927			0.7310		
	Female	Male	Mean Difference	Black	Non-Black	Mean Difference
Prob. Of Being Funded	0.6553	0.6641	-0.0088	0.6748	0.7051	-0.0303
True Positive Rate	0.7589	0.7406	0.0183	0.7781	0.7992	-0.0211
False Positive Rate	0.4717	0.4807	-0.0090	0.4581	0.5161	-0.0580
Average Score of Positive Class	0.7134	0.7133	0.0001	0.7384	0.7465	-0.0081
Average Score of Negative Class	0.5720	0.5922	-0.0203	0.5818	0.5974	-0.0156

Notes: (1) We do not directly observe gender and race in the data. We assign loans that were requested by borrowers who were in high/low female concentrated occupations (female percentage greater than 85%/less than 15%) into the female/male group; we assign loans that were requested by borrowers who were in high/low black concentrated locations (black percentage greater than 85%/less than 15%) into the black/non-black group.
(2).p<0.1; *p<0.05; **p<0.01; ***p<00.01