What's in a Face?

An Experiment on Facial Information and Loan Approval Decision

Internet Appendix

Appendix 1. Experimental Instruction and Related Documents

In this part, we list English-translated documents related to the main experiment.

- **Recruitment poster**

[Content] Click to make decisions on the computer.

[Time] About 90 minutes

[Payment] Through Alipay or Wechat.

[Sessions] Three periods can be selected from. A: 16:30-18:00; B: 18:30-20:00; C: 20:00-21:30

[Experimental location] Room 305, Building 1, Guanghua School of Management

[Requirements]

1. Peking University students who can use their Peking University gateway to log onto the Internet.

2. Students who can arrive on time, complete all tasks carefully, and do not use mobile phones during the experiment.

3. This research needs to extract participants' face information at the start of the experiment. The research team guarantees anonymization of the face information. Students who are concerned about this, please consider carefully before you sign up.

[Ways to register] Send a text message with "REGISTRATION + name + department + program + gender + age + all sessions that you are available (such as AC)" to 131XXXX1797. We will send you a confirmation message if your registration is successful. We will also remind you one day before the experiment via text messages.

- **Consent form**[1]

Hello! Welcome to our experiment. This study requires digitized variables of your facial information, so the experimenter will take a picture of you. We promise to use your photos as digital variables only in this research project, and we will never disclose your photo information in any way.

---

[1] Subjects need to sign this form on their arrival at our lab.

Financial Development Research Center of Peking University

Experimenter (signature): ____

Date: ____

- **[Material 1] Experimental instructions**[2]

In this experiment, you will see a series of loan applications, and decide whether to lend to each borrower according to the application information. The information is real historical data from one loan institution, but the personal identity information is hidden.

You have two tasks in total. Task 1 is to score the "creditworthiness" of the borrower's face photo. Task 2 is to decide to "accept" or "reject" the application.

The basic payment is 20 RMB. In addition, you can earn extra money in tasks 1 and 2. We will evaluate the quality of your answers according to the borrowers' repayment behavior. For those who were accepted based on real historical data, the repayment behavior is their real outcome. For those who were rejected based on real historical data, the repayment behavior is predicted via machine learning. You cannot distinguish these two situations during the experiment.

Introduction of Tasks

Task 1: Score the Face Photos

Your decision: You need to score 30 face photos which are randomly selected by the computer program. The score range is from 0 to 10, where a higher score represents higher creditworthiness. Here, the operational definition of "creditworthiness" is when the borrower repays within 90 days after the due day (including repaying in advance, repaying on the due day and delaying no more than 90 days) in each repayment period (if this is a multi-period loan).

Your payment: The computer program will calculate the correlation parameter ρ between your score and the borrowers' repayment behavior, which lies within a range [-1, 1]. You will gain $50 * \rho$ RMB. That is to say, the more positively related your score and repayment behavior are, the higher payment will be in this task.

Task 2: Lending Decision

[2] Subjects need to read this material before logging into the experimental website. They can keep this material in their hands throughout the experiment.

Description of information: Before you make the lending decision, you will see some information about the borrower, including four types (A/B/C/D). For each application, you will definitely see information A, and randomly see one of B/C/D or none of B/C/D. In other words, each application has four possible information sets: A, A+B, A+C, or A+D. Meanwhile, information A and information B/C/D appear in random order. The detailed content of information A/B/C/D is as follows:

A. Background information

   Meaning: The information that borrowers filled out in the mobile APP or the loan institution inquired from a linked database in the applying stage. There are more than 20 items, including personal information (demographic information, economic and background information, and social information), and loan application (loan amount, loan duration, loan interest). For a detailed explanation of each variable, please read the second material, *Application Information Details*.

B. Face photo of the borrower

   Meaning: The photo taken of borrowers in the applying stage.

   If this information appears, you will be asked to score the "creditworthiness" as task 1. The score in this situation will be pooled with task 1 and determine the payment in the scoring decision.

C. Credit score of the borrower given by algorithm

   Meaning: Our research team used about 75,000 observations out of the sample of this experiment and implemented machine learning based on the operational definition of "creditworthiness". We then applied the algorithm to the sample in this experiment to get the predicting credit score. If this information appears, you will see a credit score between 0-10, predicted by machine learning, but will not see the face photo.

D. Credit score of the borrower given by yourself in task 1

   If this information appears, you will see your score in task 1, but will not see the face photo.

Your decision: After viewing the information, you need to decide to "accept" or "reject" this application. You need to approve 120 applications in task 2, with 10 rounds as a group so there are 12 groups in total. You can choose to take a rest after finishing any group, with no time limit or frequency limit. You can leave your seat during the rest, but please keep quiet.

Your payment: The accuracy of your lending decisions will affect your payment. The same as the definition of "creditworthiness", here we define repayment as when the borrower repays within 90 days after the due day (including repaying in advance, repaying on the due day and delaying no more than

90 days) in each repayment period (if this is a multi-period loan). If your decision coincides with the repayment behavior, you will earn 10 RMB in this round. If not, you will earn 0. Namely, payment in task 2 has four possible outcomes:

	The borrower repays	The borrower defaults
You accept	10 RMB	0 RMB
You reject	0 RMB	10 RMB

After the experiment, the computer program will select 30 rounds randomly and pay you according to accuracy in these selected rounds. Each decision you make may influence the payment in task 2 potentially, so please decide carefully!

Lastly, please note that the webpage can only go forward and cannot go back. So you cannot return to look at the previous page. Please make sure you have read the information completely before you move onto the next page. If you click on "back" or "refresh", the experiment will be ended and we can only pay you the basic payment.

If you have any questions about the process of this experiment, please raise your hand. If you have no questions, please read the second material, *Application Information Details*. You can view these two materials at any time during this experiment.

- **[Material 2] Details of borrower information (Type A information)**[3]

The following figure is an example of the application information page.

[3] Subjects need to read this material before logging into the experimental website. They can keep this material in their hands throughout the experiment.

Borrower's information

Borrower's ID: 1607470

Loan information

Loan amount	1000	**Loan days**	30	**Interest rate**	**3%**
Payment	1030 RMB * 1 period				

Personal information

<table>
<tr><td>Gender</td><td>Male</td><td>Age</td><td>24</td><td>Location</td><td>Hubei</td></tr>
<tr><td>Zhima score</td><td>572</td><td>Monthly average phone bill</td><td>147.42</td><td>Occupation</td><td>White-collar</td></tr>
<tr><td rowspan="4">Phone number Provided</td><td>Family</td><td>Yes</td><td rowspan="4" colspan="2">Numbers of application on other platforms in recent month</td><td rowspan="4">13</td></tr>
<tr><td>Workplace</td><td>No</td></tr>
<tr><td>Friend</td><td>Yes</td></tr>
<tr><td>Other</td><td>No</td></tr>
</table>

Next

The following table introduces the meaning of the variables involved, please read them one by one. If you have any questions, please raise your hand to ask.

Variable name	**Variable meaning**
Loan information	
Loan amount	Range from 500 to 5,000 yuan
Loan term	Range from 2 to 150 days
Interest rate	Range from 6% to 36%. The borrowing rate represents the cost to the borrower of borrowing. Total repayment ≈ loan amount × (1 + interest rate) Repayment per period ≈ Total repayment ÷ Number of periods (The "≈" is taken because the actual formula of the loan company is equal to the principal and interest. It is similar to the above formulas.)
Borrower information	
Gender	Male / Female
Age	Calculated from the year of birth on the borrower's ID
Location	Filled in by the borrower. Very few borrowers are missing this item.

Zhima score	Based on five dimensions of information such as identity, performance, history, relationships, and behavior. Range from 350 to 950. Most people's scores are between 600 ~ 650.
Monthly phone call charges	Monthly average phone call charges in the last 12 months, obtained by the loan company from other sources. Very few borrowers are missing this item.
Occupational category	Borrowers choose one of the following categories: worker / teacher / white-collar worker / student / entrepreneur / self-employed / company employee / corporate legal person / online shop owner / unemployed / other. Very few borrowers are missing this item.
Number of applications on other platforms (within 1 month)	Number of times that the borrower has applied for loans on other platforms within 1 month on the application date, obtained by the loan company from other sources. ※ Please note that this variable only indicates the number of applications, not representing approval, or loan amount the person needs to repay. The average is 20 times.
Contact phone number provided for family / workplace / friend / other	Borrowers need to fill in several contacts and indicate which one they belong to: family, workplace, friend, or other. These four variables indicate whether the borrower provided contacts labeled “family” / “workplace” / “friend” / “other”.

- **Survey questions after experiment**

First, please answer the following questions related to this experiment. (Note: All questions are required.)

1. In the process of decision making, what is your decision on the weight of facial information and background information? (slider 1~100)
2. Please rank the importance of three forms of facial information (photos, your own score, AI score) in decision making. (ranking 1~3)
3. Which do you think is more accurate, your own score or the AI score? (multiple choice)

a. My own score; b. AI score; c. Almost the same

4. In which case do you think your decision is the most accurate? (multiple choice)

   a. Background information only;

   b. Background information with facial photos;

   c. Background information with my own score;

   d. Background information with AI score;

   e. Almost the same.

Next, please answer the following questions related to yourself. (Note: If it is not convenient for you to answer the following questions, please select "other" for multiple-choice questions and type "1" for fill-in-the-blank questions.)

1. Gender (multiple choice) a. Male; b. Female; c. Other
2. Age: ____. (fill in the blank, 1~99)
3. You study at ___(department / school), ____(university). The subject/field that you think is most suitable and interesting for you is ____. (fill in the blanks)
4. You are studying in the ___ program. (multiple choice)

   a. Undergraduate; b. Master; c. PhD; d. Other
5. You are from ____ (province), ____ (city). (fill in the blanks)
6. The education level of your father is: (multiple choice)

   a. Junior high school and below; b. Vocational high school; c. College; d. University; e. Master and above; f. Other
7. The education level of your mother is: (multiple choice)

   a. Junior high school and below; b. Vocational high school; c. College; d. University; e. Master and above; f. Other
8. Do you have a credit card? (multiple choice)

   a. Yes; b. No; c. Other
9. How many times have you borrowed on the Internet (including consumption loans and cash loans)? (multiple choice)

   a. Never; b. Once; c. 2-5 times; d. More than 5 times; e. Other
10. Has your family had any loan experience (including home loan, car loan, and education loan)? (multiple choice)

a. Yes; b. No; c. Other

- **Instruction of Online Supplementary Experiment**

Thank you for participating in our experiment!

The task of this experiment is to evaluate some face photos. There are four sets of questions, each with 100 photos. There are no standard answers to all questions. Please choose the most satisfying answer carefully. After this task is completed, please fill in an anonymous background questionnaire.

The experimental payment is divided into two parts:

(1) The basic payment for completing all the tasks is 20 RMB.

(2) There may be more than one question for each photo in the first two groups. To ensure fairness, we will award an additional 5-10 RMB according to the number of questions.

The experiment takes about 30 to 40 minutes. The experimenter will review whether your answer is valid within 24 hours, and then transfer the money to Alipay for you. If you have not completed all tasks, you cannot receive the payment. Therefore, please take about 40 minutes in a reliable network environment to ensure that the experiment is completed at one time.

Now, please use your mobile phone number (the mobile phone number used to receive Alipay transfers) as the username, and provide your gender, and then log onto this website.

Appendix 2. Formulas of Geographical Facial Features

This part introduces how to generate geographical facial features from photos. In the first step, we use a pre-made algorithm to convert facial photos into 81 pairs of coordinates of key points.[4] Figure A.1 shows how these 81 points portray the outline of a face. Next, we calculated a series of angles and distances according to *Handbook of Anthropometry* (Preedy 2012).

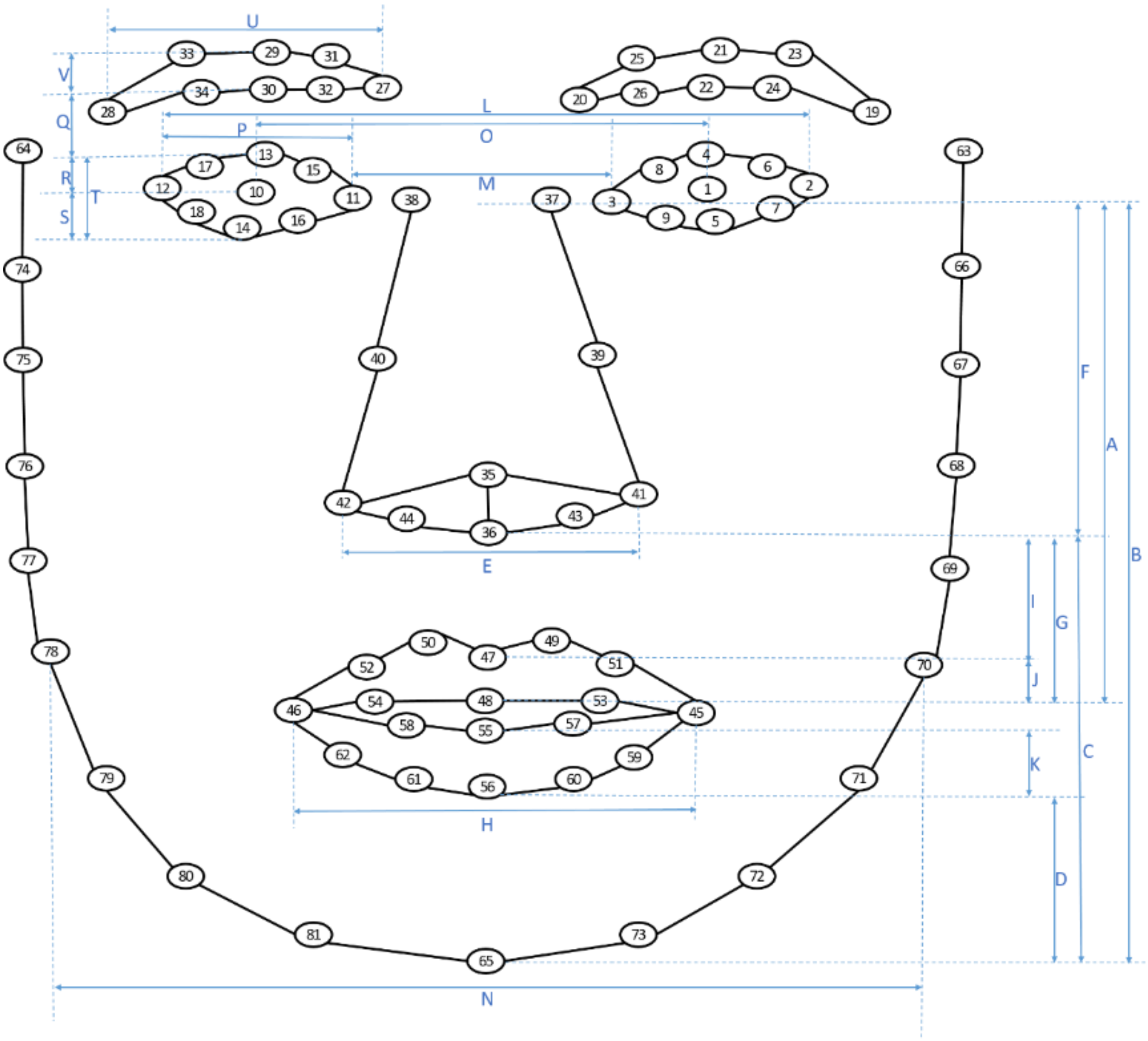


Figure A.1

81 key points of human face

---

[4] We use Face Landmarks SDK on https://www.faceplusplus.com.cn/sdk/face-landmarks.

Table A.1

Formulas of geographical facial features

Var	Formula (for symmetric variables, we only define the right side as examples)
Cheek Sharp	Largest cosine value of angle at point 70
Jaw Sharp	Cosine value of angle point 70-point 65-point 78
Eyes Upward Sloping	Cosine value of angle between facial symmetry axis and the line passing through point 2 and point 3
fWHR (Face Width to Height Ratio)	N/A in Figure A.1 i.e., the distance between point 78 and point 70, divided by the distance between midpoint of point 48 and midpoint of point 37 and point 38
Eyes Dist	M/L in Figure A.1 i.e., the distance between point 3 and point 11, divided by the distance between point 2 and point 12
Eyebrow Dist	Q/T in Figure A.1 i.e., the distance between point 4 and point 22, divided by the distance between point 4 and point 5
Eyeball Hor Loc	2*O/(L + M) in Figure A.1 i.e., twice the distance between point 1 and point 10, divided by the sum of the distance between point 3 and point 11 and the distance between point 2 and point 12
Eyeball Ver Loc	R/T in Figure A.1 i.e., the distance between point 1 and point 4, divided by the distance between point 4 and point 5
Brow Height	V/U in Figure A.1 i.e., the distance between point 21 and point 22, divided by the distance between point 19 and point 20
Brow Curv	Upper polyline distance/U in Figure A.1 i.e., the sum of polyline distance between point 19, point 23, point 21, point25 and point 20, divided by the distance between point 19 and point 20

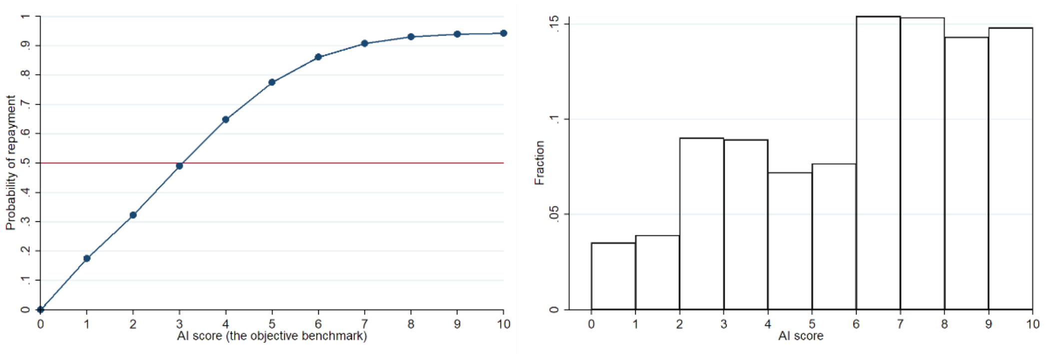
Jaw Height	D/C in Figure A.1 i.e., the distance between point 56 and point 65, divided by the distance between point 36 and point 56
Nose Mouth Dist	I/C in Figure A.1 i.e., the distance between point 36 and point 47, divided by the distance between point 36 and point 65
Mouth Width	H/N in Figure A.1 i.e., the distance between point 45 and point 46, divided by the distance between point 70 and point 78
Nose Height	F/A in Figure A.1 i.e., the distance between point 36 and midpoint of point 37 and point 38, divided by the distance between point 48 and midpoint of point 37 and point 38
Nose Width	E/N in Figure A.1 i.e., the distance between point 36 and point 47, divided by the distance between point 36 and point 65
Asymmetry	Asymmetry = (Asym1 + Asym2 + Asym3 + Asym4 + Asym5 + Asym6 + Asym7) / 7, where: Asym1 = distance between point 65 and point 78 / distance between point 65 and point 70; Asym2 = distance between point 12 and point 78 / distance between point 2 and point 70; Asym3 = distance between point 12 and point 46 / distance between point 2 and point 45; Asym4 = distance between point 11 and central axis / distance between point 3 and central axis; Asym5 = distance between point 12 and central axis / distance between point 3 and central axis; Asym6 = distance between point 42 and central axis / distance between point 41 and central axis; Asym7 = distance between point 46 and central axis / distance between point 45 and central axis

Appendix 3. Whether the facial signal is strong enough to affect the approval decision

A natural binary choice of whether to approve the loan application in the experimental design may lack the power to detect the change in prior and posterior beliefs, especially when both are on the same side of 50%. We have investigated this issue in depth and in the following four aspects.

(i) Is the signal of facial photo strong enough to alter the repayment behavior?

We would like to first understand whether the signal of facial photo is strong enough to alter the binary approval decision in our data. We run a probit regression of actual repayment behavior on AI score (the objective evaluation of the facial photo) and other background information. Then we calculate how the predicted probability of repayment changes with different levels of AI score. Appendix Figure A.2 below shows that when the AI score increases from 0 (the lowest) to 10 (the highest), the predicted repayment probability increases from about 0% to 95%, with 50% crossed around the AI score level of 3. The average of AI score is 6.1, and the distribution in (b) suggests that a non-trivial amount is distributed around level 3. These patterns together indicate that the signal of facial photo is strong enough to induce change in the binary repayment behavior in the data.



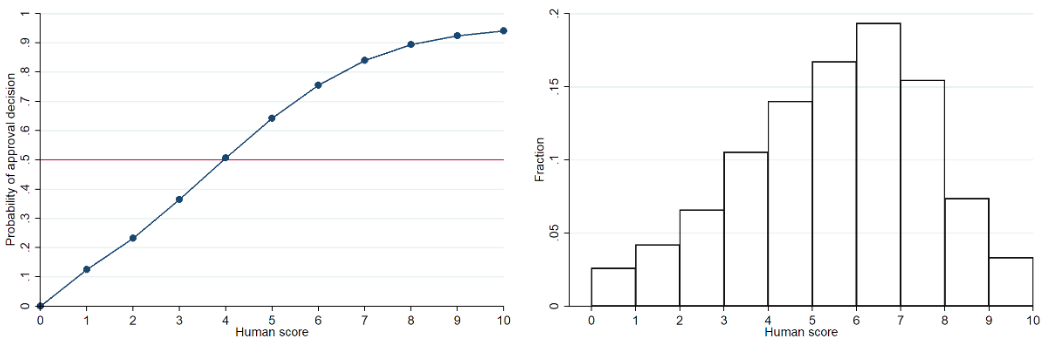
(a) Pr(repayment) and AI score　　　　(b) Distribution of AI score

Figure A.2

Objective predictability of facial signal: repayment behavior and AI score

(ii) Is the signal of facial photo strong enough to alter the binary approval decision in subjects' choices?

Despite the fact that the signal of facial photo is strong enough in the data, an equally important question is whether the same conclusion holds in subjects' choices. Similarly, we run a probit regression of subjects' approval decision on human score (the subjective evaluation of the facial photo) and other background information. Then we calculate how the predicted probability of repayment changes with different levels of human score. Appendix Figure A.3 below shows that when the human score increases from 0 (the lowest) to 10 (the highest), the predicted repayment probability also increases from about 0% to 95%, with 50% crossed around the human score level of 4. The average of human score is 5.39, and the distribution below suggests that a non-trivial amount is distributed around level 4. These patterns together indicate that the signal of facial photo is also strong enough to induce change in the binary approval decision in subjects' choices.



(a) Pr(approval) and human score (b) Distribution of human score

Figure A.3

Subjective predictability of facial signal: approval decision and human score

<u>(iii) Does the subject-level approval rate concentrate on all accepting or approving?</u>

We also draw the distribution of subject-level approval rate in Appendix Figure A.4. The average approval rate is 65.3%, and we can see that almost no subjects chose always approving or rejecting the loan applications. This pattern suggests that whatever priors subjects may have, the updated posterior goes across the 50% cutoff regularly. According to Appendix Figure A.4, the signal of facial photo is strong enough to contribute to this cutoff point crossing.

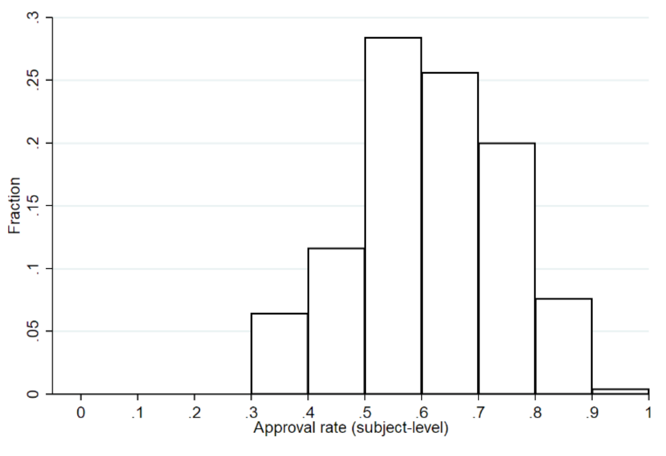


Figure A.4

Distribution of approval rate (subject-level)

(iv) Is there a significant effect of facial photo on the binary approval decision?

One final piece of evidence on this issue is from the estimation about how subjects weigh human score in their approval decisions. If the additional information of facial photo is not strong enough, or subjects' priors are sufficiently far from the 50% cutoff point, then subjects' loan approval decisions should not depend on the human score. However, Table III shows that the coefficient of human score on approval decision is significantly positive, suggesting that subjects' final decisions do react to the signals of human scores.

To sum up, both the objective relationship in the data and the decisions by the subjects suggest that the extreme case that subjects bunch at all approving or all rejecting, does not happen in our data. The additional signal of facial photo is strong enough for the subjects to alter their priors in sufficient amount of cases that even the binary approval decision can effectively detect the probability updating from facial photo. However, because we lack direct data on prior and posterior beliefs, we cannot evaluate the effect continuously, the estimated effect from the binary choice could be an under-estimation compared to the true effect. But given that we already find strong enough effects of facial photo in changing the approval decisions, the main conclusion is not substantially changed by this concern.

Appendix 4. The Training Process of Machine Learning Algorithm

This part describes the training processes of all the machine learning algorithm in detail.

1. Benchmark for the human score: use facial information to predict the repayment behavior

To set up an objective prediction of real repayment behavior based on facial information, we use another dataset of 75,000 observations (containing the facial photo and the real repayment behavior) from the same company to train a machine learning algorithm.

Basically, the training process is composed of two steps. In the first step, to reduce the dimensionality of the photo pixel information, we use a pre-made algorithm (Face Landmarks SDK on https://www.faceplusplus.com.cn/sdk/face-landmarks) to convert facial photos into 81 pairs of coordinates of key points, which portray the outline of a face. In this way, we keep the information of facial structure but eliminate other confounding factors such as clarity, brightness, and saturation that may affect the prediction.

In the second step, we develop a multi-layer algorithm to predict repayment likelihood using 81 key points as attributes and repayment behavior as labels. Since the raw data we obtain is an imbalanced sample with approximately 77% repaid borrowers and 23% default borrowers, we use the over-sampling method to construct a balanced sample to train the algorithm, which is a common practice to guarantee the validity of the algorithm in machine learning (e.g., Chawla et al. 2002). Specifically, we resampled the default borrowers with replacement and added white noises to the repeatedly sampled observations. Therefore, our algorithm does not “know” about the actual default rate when being trained.

We use keras, tensorflow, and numpy packages in the algorithm. Our neuro network has three layers: the first layer is convolutional neural network (CNN); the second layer is full convolutional network (FNN); the third layer is a max pooling layer. The parameters are as follows: 80% of observations are used as a training set and 20% are used as a testing set; our optimization method is the stochastic gradient descent with momentum; learning rate is equal to 0.01; momentum is equal to 0.9; batch size is 128; epoch number is 10; time cost of training each epoch is 45min.

In our experimental data without over-sampling, this algorithm gives an average score of 6.08 (sd = 2.62, min = 0, max = 10), and achieves an out-of-sample accuracy rate of 80.5%. To evaluate the performance of our algorithm with more indices, we divide our sample into “AI approved” and “AI rejected” following a critic value of 0.5. Then we compare this AI decision with real outcome to obtain the accuracy rate. The precision rate (the ratio of correctly predicted repaid observations to the total predicted repaid observations) is 92.96%, the recall rate (the ratio of correctly predicted repaid observations to the total repaid observations) is 80.75%, and the F1-score (the harmonic mean of precision rate and recall rate) is 0.864. All of the indices suggest that our algorithm perform quite well.

2. Benchmark for the Baseline Treatment: use background information to predict the repayment behavior

To set up an objective benchmark of the baseline treatment, we train another machine learning algorithm, which uses background information to predict the repayment behavior.

We train this algorithm in a random forest model. We use 9929 observations for training, with the repayment behavior as the output variables, and 54 background variables (including dummies) as input variables. The *mtry* (number of variables randomly sampled as candidates at each split) is 5. The *ntree* (number of trees to grow) is 500. In another sample with 9995 observations, this algorithm gives an out-of-sample accuracy of 78.6% (the cutoff value is 0.5).

3. Benchmark for the Score/Photo Treatment: use background information and AI score to predict the repayment behavior

To set up an objective benchmark of the Score/Photo treatment, we train another machine learning algorithm, which uses background information and AI score to predict the repayment behavior.

We train this algorithm in another random forest model. We use 9929 observations for training, with the repayment behavior as the output variables, and 54 background variables (including dummies) and the AI score as input variables. The *mtry* (number of variables randomly sampled as candidates at each split) is 8. The *ntree* (number of trees to grow) is 500. In another sample with 9995 observations, this algorithm gives an out-of-sample accuracy of 87.2% (the cutoff value is 0.5).

Appendix 5. Other tables and figures

Table A.2

T-test of differences in borrowers' background information across treatment groups

Treatment	B(Baseline)	S(Score)	P(Photo)	B-S	B-P
Stats	Mean (sd)	Mean (sd)	Mean (sd)	Diff (se)	Diff (se)
Loan amount (RMB in 1,000)	1.372	1.344	1.383	0.027*	-0.011
	(0.666)	(0.640)	(0.674)	(0.013)	(0.013)
Loan duration (in days)	55.462	53.967	56.058	1.495	-0.596
	(44.554)	(43.443)	(44.527)	(0.883)	(0.894)
Interest-principal proportion	0.227	0.226	0.226	0.001	0.001
	(0.070)	(0.071)	(0.071)	(0.001)	(0.001)
Age	27.488	27.559	27.556	-0.070	-0.068
	(4.995)	(5.172)	(5.062)	(0.102)	(0.101)
Male (dummy)	0.773	0.781	0.774	-0.008	-0.002
	(0.419)	(0.414)	(0.418)	(0.008)	(0.008)
Family (dummy)	0.998	0.998	0.997	0.001	0.002
	(0.040)	(0.047)	(0.057)	(0.001)	(0.001)
Firm (dummy)	0.392	0.374	0.389	0.019	0.004
	(0.488)	(0.484)	(0.488)	(0.010)	(0.010)
Friend (dummy)	0.959	0.969	0.963	-0.010**	-0.004
	(0.199)	(0.173)	(0.188)	(0.004)	(0.004)
Other (dummy)	0.032	0.028	0.026	0.004	0.006
	(0.176)	(0.166)	(0.159)	(0.003)	(0.003)
Number of applications in 1 month	25.487	25.707	25.252	-0.220	0.235
	(14.843)	(14.683)	(14.446)	(0.296)	(0.294)
Average monthly phone bill	183.173	181.900	184.984	1.273	-1.811
	(107.684)	(107.371)	(112.954)	(2.159)	(2.215)
Zhima credit score	604.704	605.075	605.361	-0.370	-0.656
	(56.422)	(55.387)	(57.951)	(1.122)	(1.148)
N	4,961	4,962	4,968	9,923	9,929

Note: The first three columns summarize background information in Baseline, Score Treatment and Photo Treatment, with mean in the cell and standard deviation in the bracket. The last three columns report t statistics between Baseline/Score Treatment and Baseline/Photo Treatment, with the difference in mean in cells and standard error in brackets. *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed t-tests, respectively.

Table A.3

T-test of differences in borrowers' background information across information sequences

Treatment	Score Treatment & Photo Treatment		
Background information sequence	F(First)	S(Second)	F-S
Stats	Mean (sd)	Mean (sd)	Diff (se)
Loan amount (RMB in 1,000)	1.352	1.375	0.022
	(0.652)	(0.663)	(0.013)
Loan duration (in days)	54.360	55.680	1.319
	(43.754)	(44.242)	(0.883)
Interest-principal proportion	0.227	0.225	-0.002
	(0.070)	(0.072)	(0.001)
Age	27.558	27.557	-0.001
	(5.060)	(5.175)	(0.103)
Male (dummy)	0.776	0.779	0.002
	(0.417)	(0.415)	(0.008)
Family (dummy)	0.997	0.998	0.001
	(0.056)	(0.047)	(0.001)
Firm (dummy)	0.376	0.386	0.010
	(0.484)	(0.487)	(0.010)
Friend (dummy)	0.966	0.966	-0.000
	(0.180)	(0.182)	(0.004)
Other (dummy)	0.026	0.028	0.002
	(0.160)	(0.165)	(0.003)
Number of applications in 1 month	25.404	25.556	0.151
	(14.723)	(14.405)	(0.292)
Average monthly phone bill	187.033	179.780	-7.253^{**}
	(114.629)	(105.385)	(2.211)
Zhima credit score	604.448	606.003	1.556
	(56.897)	(56.455)	(1.138)
N	5,015	4,915	9,930

Note: Background information is tested by whether it appears first or second in Score Treatment and Photo Treatment. The first two columns summarize background information in which background information appears first or second, with mean in cells and standard deviation in brackets. The last column reports t statistics between the first two columns, with difference in mean in cells and standard error in brackets. *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed t-tests, respectively.

Table A.4

Subjects' background information

Variable	N	Mean	SD	Min	Max
Gender_u	250	0.35	0.48	0	1
Age_u	246	21.13	2.48	17	32
Edu_u	247	1.32	0.60	1	3
Edu_f	234	2.93	1.22	1	5
Edu_m	236	2.83	1.21	1	5
Card_u	244	0.48	0.50	0	1
Edebt_u	241	0.47	0.98	0	3
Fdebt_u	233	0.69	0.46	0	1

Note: This table summarized the statistics of subjects in this experiment. Gender_u is coded 1 for male and 0 for female. Age_u is the subjects' self-reported ages. Edu_u is coded 1 for undergraduate, 2 for master's, and 3 for PhD. Edu_f and Edu_m are father's and mother's education, which is coded 1 for junior high school or below, 2 for vocational high school, 3 for college, 4 for university, and 5 for master's or above. Card_u is a dummy (having a credit card). Edebt_u is experiences of borrowing on the Internet, which is coded 0 for never, 1 for once, 2 for two to five times, and 3 for more than five times. Fdebt_u is a dummy (family's mortgage experience).

Table A.5

Session, time and number of subjects

Session	Date and time	Number of subjects
1	2019/5/12 13:00-14:30	23
2	2019/5/12 14:30-16:00	23
3	2019/5/19 09:00-10:30	27
4	2019/5/19 10:30-12:00	25
5	2019/5/19 13:00-14:30	27
6	2019/5/19 14:30-16:00	26
7	2019/5/28 16:30-18:00	34
8	2019/5/28 18:30-20:00	31
9	2019/5/28 20:00-21:30	34

Note: The first column is the session number. The second column is the date and time of each session. The third column is the number of subjects in each session. There are 9 sessions and 250 subjects in total.

Table A.6

Questions in supplementary online experiment

Key words	Questions
Panel A: Appearance	
Dressing formal	Please evaluate the dress code: A. Daily; B. Home
Makeup	Please rate the makeup of the person in the photo, 0 for completely plain, and 10 for completely heavy.
Glasses	Is the person in the photo is wearing glasses? A. Yes; B. No
Hair color	Which of the following colors is closest to the person's hair in the photo? A. Black; B. Gray; C. Brown; D. Other
Hair length	Please rate the length of hair of this person in this photo, 0 for the shortest and 10 for the longest.
Hair curvature	Please rate the curvature of hair of this person in this photo, 0 for the straightest and 10 for the most curved.
Body weight	Please rate the bodyweight of the person in the photo, 0 for the thinnest and 10 for the heaviest.
Beauty	Please rate the beauty of the person in the photo, 0 for the least beautiful and 10 for the most beautiful.
Panel B: Expression	
Mouth open	Please rate the mouth opening of the person in the photo, 0 for closed up and 10 for wide open.
Smile	Please rate the smile of the person in the photo, 0 for no smile and 10 for big smile.
Eyes open	Please rate the eye condition of the person in the photo, 0 for closed, 5 for naturally-opened, and 10 for wide open.
Look to left/right	Please rate the horizontal direction of the gaze of the person in the photo, 0 for the most left and 10 for the most right.
Look up/down	Please rate the vertical direction of the gaze of the person in the photo, 0 for the most downward and 10 for the most upward.

Table A.7

Summary statistics of evaluations in supplementary online experiment

Variable	N	Mean	SD	Min	Max
Panel A: Appearance					
Dressing Formal	5,001	0.65	0.48	0.00	1.00
Makeup	4,625	1.05	1.28	0.00	8.50
Glasses (dummy)	4,925	0.05	0.23	0.00	1.00
Haircurve	4,625	2.22	1.47	0.00	9.50
Hairlength	4,625	3.00	1.81	0.00	9.50
Haircolor (black=1, others=0)	4,625	0.90	0.30	0.00	1.00
Bodyweight	5,005	5.23	1.34	0.00	9.67
Beauty	5,033	3.62	1.32	0.25	9.00
Panel B: Expression					
Mouthopen	5,033	3.18	1.72	0.00	8.67
Smile	5,033	1.49	1.23	0.00	9.00
Eyes Open	4,925	4.76	1.12	1.33	10.00
Look to Left/Right	4,925	1.06	1.29	0.00	10.00
Look Downward	4,925	4.53	1.06	0.50	9.75

Note: Explanations of variables are in Appendix Table A.6.

Table A.8

Biases in human score in task 1

	(1)	(2)	(3)	(4)	(5)	(6)
DV	Outcome (= Repay if Subject = 0; = Human Score if Subject = 1)					
	Repay as dummy			Repay as continuous variable		
IV	Geo.	Res.	Sup.	Geo.	Res.	Sup.
Jaw Sharp	-0.209			-0.213*		
	(0.129)			(0.127)		
Cheek Sharp	-0.186			-0.297		
	(1.571)			(1.539)		
Eyes Upward Sloping	0.131			0.078		
	(0.192)			(0.187)		
Eyes Dist	-0.116			-0.075		
	(0.398)			(0.395)		
Eyebrow Dist	0.014			0.013		
	(0.011)			(0.011)		
Eyeballs Hor Loc	-0.954**			-1.004**		
	(0.412)			(0.407)		
Eyeballs Ver Loc	-0.168			-0.143		
	(0.200)			(0.196)		
Brow Height	-1.090			-1.081		
	(0.671)			(0.669)		
Brow Curve	0.737			0.721		
	(0.472)			(0.468)		
Jaw Height	0.252*			0.266*		
	(0.150)			(0.148)		
Nose Mouth Dist	-0.417			-0.448		
	(0.277)			(0.275)		
Mouth Width	-0.648***			-0.643***		
	(0.195)			(0.196)		
Nose Height	-0.005			-0.035		
	(0.289)			(0.288)		
Nose Width	0.121			0.080		
	(0.309)			(0.310)		
fWHR	-0.217**			-0.221***		
	(0.084)			(0.083)		
Asymmetry	0.072			0.076		
	(0.075)			(0.073)		

Subj. # Jaw Sharp	-0.037		-0.033	
	(0.143)		(0.141)	
Subj. # Cheek Sharp	-0.764		-0.653	
	(1.771)		(1.744)	
Subj. # Eyes Upward Sloping	-0.165		-0.113	
	(0.215)		(0.211)	
Subj. # Eyes Dist	0.155		0.114	
	(0.419)		(0.415)	
Subj. # Eyebrow Dist	-0.039***		-0.038***	
	(0.012)		(0.012)	
Subj. # Eyeballs Hor Loc	1.360***		1.410***	
	(0.441)		(0.435)	
Subj. # Eyeballs Ver Loc	-0.179		-0.204	
	(0.221)		(0.217)	
Subj. # Brow Height	1.202		1.193	
	(0.740)		(0.738)	
Subj. # Brow Curve	-0.609		-0.594	
	(0.523)		(0.518)	
Subj. # Jaw Height	-0.078		-0.092	
	(0.159)		(0.156)	
Subj. # Nose Mouth Dist	0.101		0.133	
	(0.308)		(0.307)	
Subj. # Mouth Width	0.619***		0.614***	
	(0.205)		(0.208)	
Subj. # Nose Height	-0.324		-0.294	
	(0.318)		(0.316)	
Subj. # Nose Width	-0.540*		-0.499	
	(0.317)		(0.316)	
Subj. # fWHR	0.059		0.063	
	(0.090)		(0.090)	
Subj. # Asymmetry	-0.055		-0.059	
	(0.082)		(0.080)	
Cosine Similarity		-0.116		-0.099
		(0.215)		(0.210)
Same Gender		0.004		0.001
		(0.015)		(0.015)
Subj. # Cosine Similarity		0.781***		0.764***
		(0.221)		(0.217)

Subj. # Same Gender	0.010		0.013	
	(0.016)		(0.016)	
Mouth Open		-0.002		-0.002
		(0.007)		(0.006)
Smile		-0.008		-0.008
		(0.009)		(0.009)
Eyes Open		0.004		0.005
		(0.010)		(0.010)
Look to Left/Right		0.005		0.003
		(0.008)		(0.008)
Look Downward		-0.001		-0.000
		(0.010)		(0.010)
Hair Curve		0.010		0.010
		(0.008)		(0.008)
Hair Length		0.004		0.003
		(0.010)		(0.010)
Makeup		0.013		0.011
		(0.009)		(0.009)
Body Weight		0.249***		0.258***
		(0.076)		(0.075)
Beauty		-0.020		-0.009
		(0.084)		(0.083)
Dressing Formal		0.020		0.017
		(0.023)		(0.022)
Hair Color		-0.038		-0.041
		(0.037)		(0.037)
Glasses		0.131***		0.136***
		(0.032)		(0.032)
Subj. # Mouth Open		-0.001		-0.001
		(0.007)		(0.007)
Subj. # Smile		0.012		0.012
		(0.009)		(0.009)
Subj. # Eyes Open		0.011		0.010
		(0.011)		(0.011)
Subj. # Look to Left/Right		-0.016*		-0.013
		(0.009)		(0.009)
Subj. # Look Downward		-0.007		-0.008
		(0.011)		(0.011)

Subj. # Hair Curve			-0.016			-0.016*
			(0.009)			(0.009)
Subj. # Hair Length			-0.010			-0.009
			(0.011)			(0.010)
Subj. # Makeup			-0.006			-0.004
			(0.010)			(0.010)
Subj. # Body Weight			-0.233***			-0.242***
			(0.082)			(0.081)
Subj. # Beauty			0.224**			0.213**
			(0.092)			(0.091)
Subj. # Dressing Formal			0.092***			0.095***
			(0.024)			(0.024)
Subj. # Hair Color			0.048			0.052
			(0.042)			(0.042)
Subj. # Glasses			-0.070*			-0.076**
			(0.039)			(0.039)
Male Borrower	-0.045**	-0.065***	0.047	-0.050***	-0.070***	0.042
	(0.018)	(0.015)	(0.043)	(0.018)	(0.015)	(0.042)
Subj. # Male Borrower	-0.056***	-0.028*	-0.137***	-0.050***	-0.022	-0.132***
	(0.019)	(0.016)	(0.047)	(0.019)	(0.016)	(0.046)
Subj.	-1.972	-1.319***	-0.508***	-1.934	-1.298***	-0.492***
	(1.842)	(0.219)	(0.121)	(1.827)	(0.215)	(0.121)
ID (250 dummies)	Yes	Yes	Yes	Yes	Yes	Yes
ROUND (110 dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9788	9788	4270	9788	9788	4270
Adjusted R^2	0.192	0.185	0.218	0.182	0.174	0.207

Note: This table uses OLS regression to study the biases in subjects' judgment of facial photos. The sample is restricted to the human score in task 1. The dependent variable is equal to Repay if Subject = 0, and equal to Photo Score if Subject = 1. In Column (1), (2), and (3), Repay is a dummy variable. In Column (4), (5) and (6), Repay is a continuous variable, as described in Appendix Figure A.5. Independent variables are respectively geometric features, facial resemblance and supplementary features, as described in Appendix Table A.1 and Appendix Table A.6. This table controls for individual fixed effect (ID), and round fixed effect (ROUND). Standard errors are clustered by individual subjects. *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed t-tests, respectively.

Table A.9

Fatigue effect in decision quality

	(1)	(2)	(3)
DV	Correct	Wrongly Reject	Wrongly Accept
Score Treatment	-0.042	0.019	0.047
	(0.044)	(0.048)	(0.058)
Photo Treatment	0.052	-0.034	-0.041
	(0.047)	(0.049)	(0.056)
Second Half	0.055	-0.055	-0.018
	(0.048)	(0.049)	(0.059)
Score * Second Half	0.049	0.024	-0.134
	(0.065)	(0.069)	(0.085)
Photo * Second Half	-0.067	0.026	0.078
	(0.065)	(0.066)	(0.077)
ID (250 dummies)	Yes	Yes	Yes
Text First (1 dummy)	Yes	Yes	Yes
Constant	Yes	Yes	Yes
Base group mean	0.596	0.281	0.123
P value of Score = Photo	0.032	0.243	0.124
Pseudo R2	0.027	0.050	0.044
N	9838	9838	9800

Note: This table uses probit regression to study the fatigue effect in decision quality and reports original estimations. The sample is restricted to rounds without feedback (1~80) in task 2. Three dependent variables are Correct (correctly approve or correctly reject = 1; wrongly approve or wrongly reject = 0), Wrongly Reject (wrongly reject = 1; otherwise = 0), and Wrongly Accept (wrongly accept = 1; otherwise = 0). Second Half is a dummy of whether this round is in the second half rounds (41~80). The baseline is the base group. This table controls for subjects' fixed effects (ID) and whether background information appears first (Text First). Standard errors are clustered by individual subjects. *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Table A.10

Learning effect in photo score and perceived effect of facial information

	(1)	(2)	(3)	(4)
Treatment	Photo		Score	Photo
DV	Repay	Repay	Approve	Approve
Photo Score * Feedback * Round Num	-0.006	0.004	-0.014	-0.012
	(0.015)	(0.017)	(0.019)	(0.031)
Photo Score * Feedback	0.877	-0.102	0.162	2.733
	(1.488)	(1.626)	(1.748)	(3.099)
Photo Score * Round Num	-0.002	-0.003	0.012*	0.003
	(0.005)	(0.006)	(0.006)	(0.007)
Feedback * Round Num	0.003	-0.003	0.022**	0.017
	(0.009)	(0.009)	(0.011)	(0.016)
Photo Score	0.713***	0.369	2.488***	4.522***
	(0.232)	(0.256)	(0.320)	(0.391)
Feedback	-0.484	0.182	-0.916	-2.024
	(0.875)	(0.921)	(1.019)	(1.568)
Round Num	0.001	0.002	-0.006	0.002
	(0.003)	(0.003)	(0.004)	(0.004)
BACKGROUND	No	Yes	Yes	Yes
ID (250 dummies)	Yes	Yes	Yes	Yes
Text First (1 dummy)	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Pseudo R^2	0.055	0.155	0.388	0.447
N	4963	4963	4786	4963

Note: This table uses probit regression to investigate learning effects in the photo score and perceived effects of facial information, and report original estimations. The sample is restricted to all rounds (round 1~120) in Score Treatment or Photo Treatment in task 2. The dependent variables are Repay (yes = 1; no = 0) and Approve (yes = 1; no = 0). Round Num is the round number. Feedback is a dummy of whether the round gives feedback. BACKGROUND is a set of variables of background information. This table controls for subject fixed effect (ID), round fixed effect (ROUND) and whether background information appears first (Text First). Standard errors are clustered by individual subjects. *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Table A.11

Autocorrelation in human score and approval decisions

	(1)	(2)
DV	Task 1: Human Score	Task 2: Approval Decision
Human Score [t-1]	0.045**	
	(0.018)	
Approval [t-1]		-0.044***
		(0.010)
ID (250 dummies)	Yes	Yes
Round in Task 1 (30 dummies)	Yes	No
Round in Task 2 (80 dummies)	No	Yes
Text First (1 dummy)	No	Yes
Treatment (3 dummies)	No	Yes
Constant	Yes	Yes
Adjusted/Pseudo R2	0.311	0.068
N	4523	9588

Note: This table shows how subjects' decision are autocorrelated with the last round in two tasks. Column (1) analyzes the autocorrelation in Task 1 using an OLS regression, controlling for subject fixed effect (ID), round fixed effect (Round in Task 1). Column (2) analyzes the autocorrelation using a probit regression and reports the coefficients of marginal effects at means and the significance levels at original estimates, controlling for subject fixed effect (ID), round fixed effect (Round in Task 2), whether background information appears first (Text First), and treatment fixed effect (Treatment). Standard errors are clustered at the subject level in two regressions. *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Table A.12

Heterogeneity in subjects' major

	(1)	(2)	(3)	(4)
DV	Approve	Correct	Wrongly Reject	Wrongly Accept
Econ * Round Num * Feedback	0.0037	-0.0024	-0.0018	0.0060
	(0.0040)	(0.0040)	(0.0041)	(0.0047)
Round Num * Feedback	0.0053***	0.0011	-0.0047**	0.0036
	(0.0020)	(0.0021)	(0.0023)	(0.0023)
Econ * Round Num	-0.0007	0.0002	0.0004	-0.0012
	(0.0012)	(0.0015)	(0.0015)	(0.0016)
Round Num	0.0014**	0.0011*	-0.0011	-0.0004
	(0.0006)	(0.0007)	(0.0007)	(0.0009)
Econ * Feedback	-0.3920	0.2520	0.1943	-0.6331
	(0.3908)	(0.3939)	(0.4061)	(0.4627)
Feedback (dummy)	-0.1901	0.0304	0.1690	-0.1756
	(0.1928)	(0.1969)	(0.2064)	(0.2250)
Treatment (2 dummies)	Yes	Yes	Yes	Yes
ID (250 dummies)	Yes	Yes	Yes	Yes
Text First (1 dummy)	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Observations	14643	14643	14643	14643
Pseudo R^2	0.062	0.025	0.053	0.029

Note: This table uses probit regression to study the learning effect in decision quality and reports original estimations. The sample is restricted to all rounds (1~120) in task 2. Four dependent variables are Approve (approve = 1; reject = 0), Correct (correctly approve or correctly reject = 1; wrongly approve or wrongly reject = 0), Wrongly Reject (wrongly reject = 1; otherwise = 0), and Wrongly Accept (wrongly accept = 1; otherwise = 0). Econ is a dummy of whether the student subject is major in economics, finance, or psychology. In our sample, 22.8% subjects major in economics, finance, or psychology. Feedback is a dummy of whether this round gives feedback. The baseline is the base group. Round Num is the round number. This table controls for subjects' fixed effects (ID) and whether background information appears first (Text First). Standard errors are clustered by individual subjects. *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Table A.13

Heterogeneity among subjects from the top ten provinces in GDP

	(1)	(2)	(3)	(4)
DV	Approve	Correct	Wrongly Reject	Wrongly Accept
High GDP * Round Num * Feedback	-0.0013	0.0042	-0.0021	-0.0028
	(0.0035)	(0.0035)	(0.0037)	(0.0040)
High GDP * Round Num	-0.0009	-0.0005	0.0013	-0.0014
	(0.0011)	(0.0012)	(0.0012)	(0.0015)
High GDP * Feedback	-0.0013	0.0042	-0.0021	-0.0028
	(0.0035)	(0.0035)	(0.0037)	(0.0040)
Round Num * Feedback	0.0069***	-0.0014	-0.0040	0.0060**
	(0.0023)	(0.0024)	(0.0025)	(0.0027)
Round Num	0.0015**	0.0014*	-0.0016**	-0.0002
	(0.0007)	(0.0007)	(0.0007)	(0.0009)
Feedback (dummy)	-0.3528	0.3108	0.0969	-0.5125**
	(0.2254)	(0.2273)	(0.2384)	(0.2581)
Treatment (2 dummies)	Yes	Yes	Yes	Yes
ID (250 dummies)	Yes	Yes	Yes	Yes
Text First (1 dummy)	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Observations	15182	15182	15182	15182
Pseudo R^2	0.062	0.026	0.054	0.029

Note: This table uses probit regression to study the learning effect in decision quality and reports original estimations. The sample is restricted to all rounds (1~120) in task 2. Four dependent variables are Approve (approve = 1; reject = 0), Correct (correctly approve or correctly reject = 1; wrongly approve or wrongly reject = 0), Wrongly Reject (wrongly reject = 1; otherwise = 0), and Wrongly Accept (wrongly accept = 1; otherwise = 0). High GDP is a dummy of whether the subject is from the top ten provinces in GDP. In our sample, 44.4% subjects are from the top 10 provinces in GDP. Feedback is a dummy of whether this round gives feedback. The baseline is the base group. Round Num is the round number. This table controls for treatment fixed effect (Treatment), subjects' fixed effects (ID), and whether background information appears first (Text First). Standard errors are clustered by individual subjects. *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

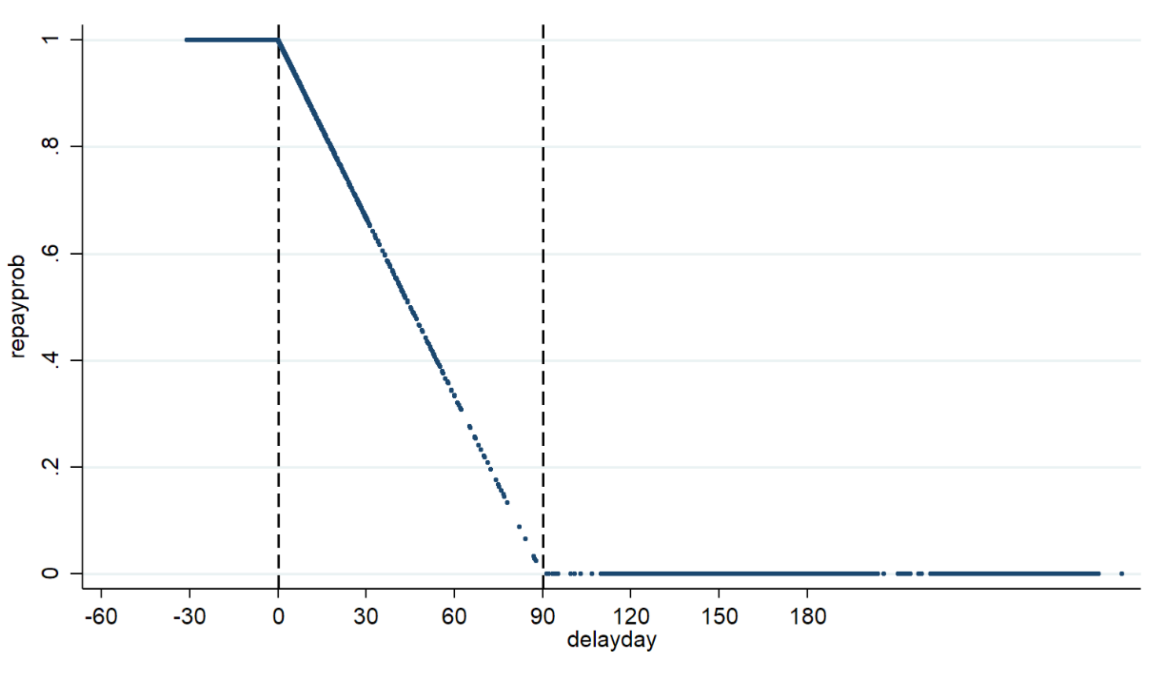
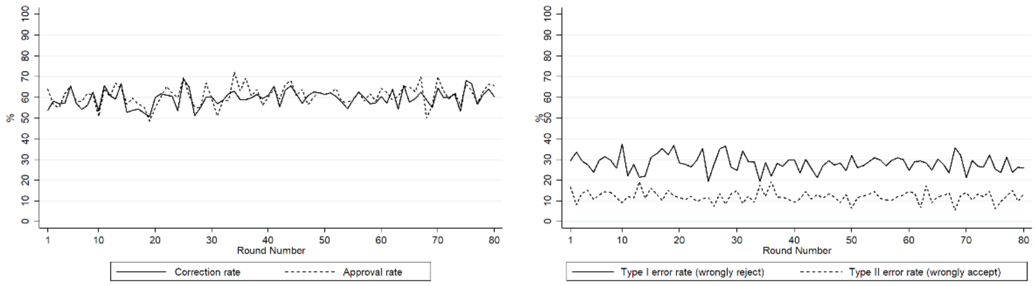


Figure A.5

Definition of repayment probability

Note: This figure visualizes how we define the repayment probability. Borrowers with nonpositive delayed days (i.e., repay in advance) are defined as having a repayment probability of 1. Borrowers with delayed days larger than 90 (i.e., default in this paper) are defined as having a repayment probability of 0. Borrowers with 0~90 delayed days are defined as having a repayment probability of $\frac{90-\text{delay days}}{90}$.



(a) Accuracy rate and approval rate **(b)Type I and type II error rates**

Figure A.6

Fatigue effect and learning effect in decision quality

Reference

Preedy, Victor R. 2012. *Handbook of Anthropometry: Physical Measures of Human Form in Health and Disease*. Springer Science & Business Media.