B. Informed consent and instructions [ONLINE APPENDIX]

Informed consent

Welcome,
You are invited to participate in a study investigating the processes that influence people's decision making. If you agree to participate, we will ask you to complete an interactive computerized task. You will be also asked to answer a short questionnaire at the end of the study. The estimated duration of the entire experiment is 120 minutes.

Before agreeing to participate in this study, it is important that you read and understand the following explanations, so you can make an informed decision about taking part in this study.

Purpose: This study is designed to investigate the processes involved in decision making.

Confidentiality: Data collected will remain strictly confidential. All data will be used for research purposes and to write a scientific paper about the nature of decision processes. Only researchers who are associated with the study will see your responses.

Your responses will not be associated with your name; instead, your name will be converted to a code number when the researchers store the data. No names or identifying information will be used in any publication or presentation.

Potential Risks and Discomforts: There are no anticipated risks associated with participation in this study.

Anticipated Benefits: The benefits associated with participating in this study are:

(1) you receive a participation payment of 12.000 COP,

(2) a payment which is based on your task performance.

(3) the satisfaction to contribute to the scientific understanding of how people make decisions. Upon the completion of the study, you will be given a thorough explanation of the study. You can also opt to receive a manuscript of any manuscript based on the research (or summaries of our results) upon completion.

Participation and Withdrawal: Your participation in this research is entirely voluntary. If you choose not to participate, it will not affect your relationship with any of the researchers involved or their institutes. If you decide to participate, you are free to withdraw your consent and discontinue your involvement at any time without penalty.

Questions: The experimenter will answer any questions about the research either now or during the course of the experiment. You can signal that you have a question by raising your hand and the experimenter will come to you promptly.

If you have other questions or concerns, you can address them to any of the following:
Nils Köbis (n.c.kobis@gmail.com)

Instructions: The instructions below inform you about the general procedure of the study you are about to take part in. It is conducted by the University of Amsterdam and University of Zürich.

Consent: I have had the opportunity to discuss this study and my questions have been answered to my satisfaction. I consent to take part in the study with the understanding that I may withdraw at any time. I am aware that an explanation about the rationale and predictions underlying this experiment will be presented upon completion of the study. I freely consent to take part in this study.

________________________________ Signature, date

General instructions

Thank you for agreeing to participate. During the study, we require your complete, undistracted attention. Please read the following instructions carefully. If you have questions at any point or do not understand the instructions, please raise your hand and one of the assistants will come and help you.

The study has two parts. Both parts have 15 rounds. After the information about the payment of the study you receive the instructions for the first part. Instructions for the second part will be distributed when the first part is over.

In both parts, there are several rounds of decision-making. Your decisions and those of other participants will determine your earnings. You will receive 7 Euros as a participation fee for this study and in addition you will be paid for six extra rounds of decision making. The computer will randomly draw three rounds of the first part and three rounds of the second part. The results of these rounds will be paid out privately to you and the others in cash at the end of today's session.

All the payoffs in the study will be expressed in points. At the end of the study your earnings will be converted in Euro at the conversion rate of

100 point = 3.000 COP.

All interactions among you and other participants will take place through computers. You are **not allowed to speak** to the other participants. If you do not follow that rule you can be excluded from the study. You will not know which specific participant made which decision and the other participants will not know the decisions you made. Your decisions and the decisions of all other participants are completely private.

In the following, the procedure for the first part of the study is described in detail.

Instructions for Part One

Roles

The computer will randomly choose 10 participants and make them into one group. Within such a group, the computer will randomly assign different roles to the participants.

The computer will randomly assign the role of **teacher** to TWO participants and the role of **student** to the remaining EIGHT participants. Other participants do not know your role and you do not know the roles of the other participants. Importantly, each participant will keep the role assigned to them by the computer throughout the entire study.

Basic Structure

In each round, teachers are paid a salary of **40 points** to teach a class and they can decide to ask students to pay a motivation fee of **10 points**. Students on the other hand have the opportunity to either join one of the two classes and obtain a diploma worth **250 points** or decide not to go to school and obtain **115 points**. If students decide to join one of the classes, they have two different ways to obtain the diploma depending on the decision of the teacher. If the teacher in their class is not asking for a motivation fee, they can only get the diploma exerting effort and paying the corresponding **effort cost.** The effort cost can change from student to student and from one round to the other.

If instead the teacher in their class is asking for a motivation fee, they can decide to obtain the diploma either by exerting effort or paying the motivation fee to the teacher. Paying the motivation fee, however, reduces the value of the diploma for all the students in the class. The sequences of decisions and the calculation of the payoffs are described in more detail below.

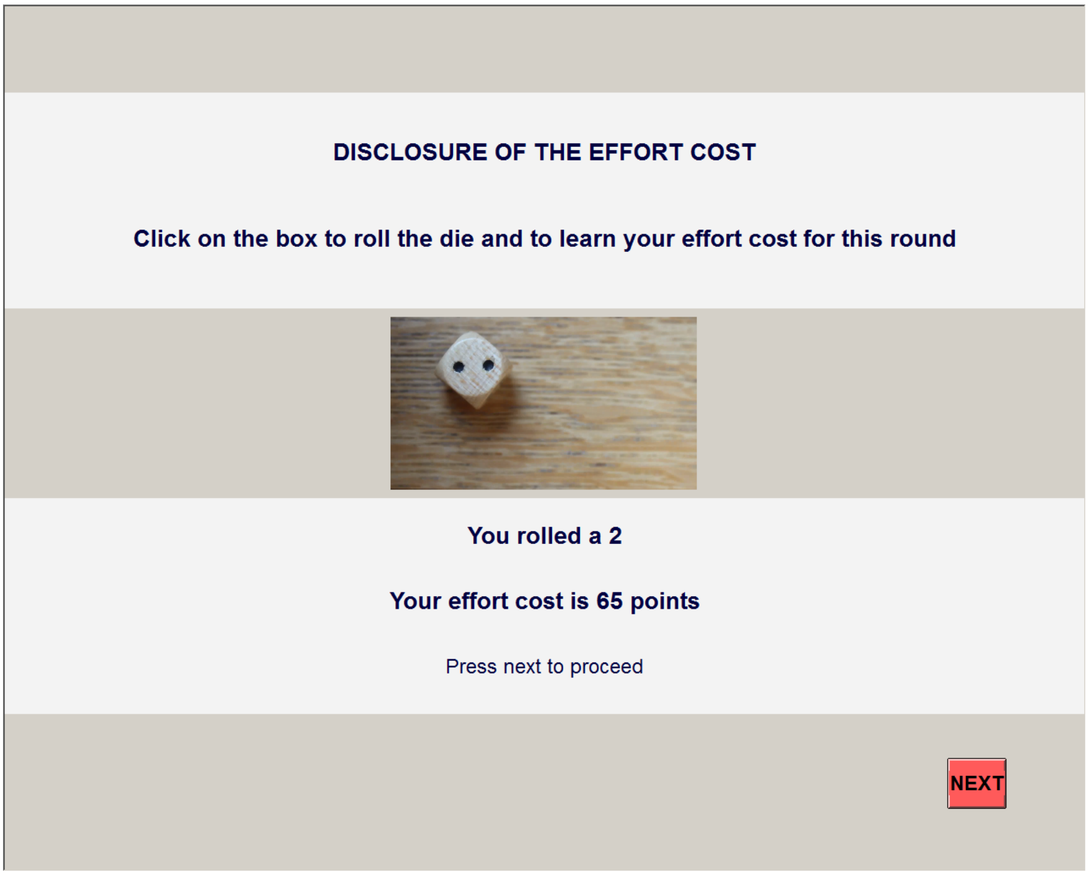
Sequence of Events

At the beginning of each round students are privately informed about their effort cost.

Effort cost

Students can obtain the degree by exerting effort. The cost of effort differs for each student, and per round. There are three levels of effort cost: **10, 65 and 160 points**. At the start of each round, the computer will determine the effort cost of a student with an independent roll of a die. A roll of a 1 leads to an effort cost of 10 points; a roll of 2, 3,4, 5 leads to an effort cost of 65 points; and a roll of 6 leads to an effort cost of 160 points. Each student is informed of her or his own effort cost, but not of the effort costs of the other students. The teachers are also not informed of the effort costs of the students.

The effort cost determines how much a student must pay to obtain the degree without motivation fee.

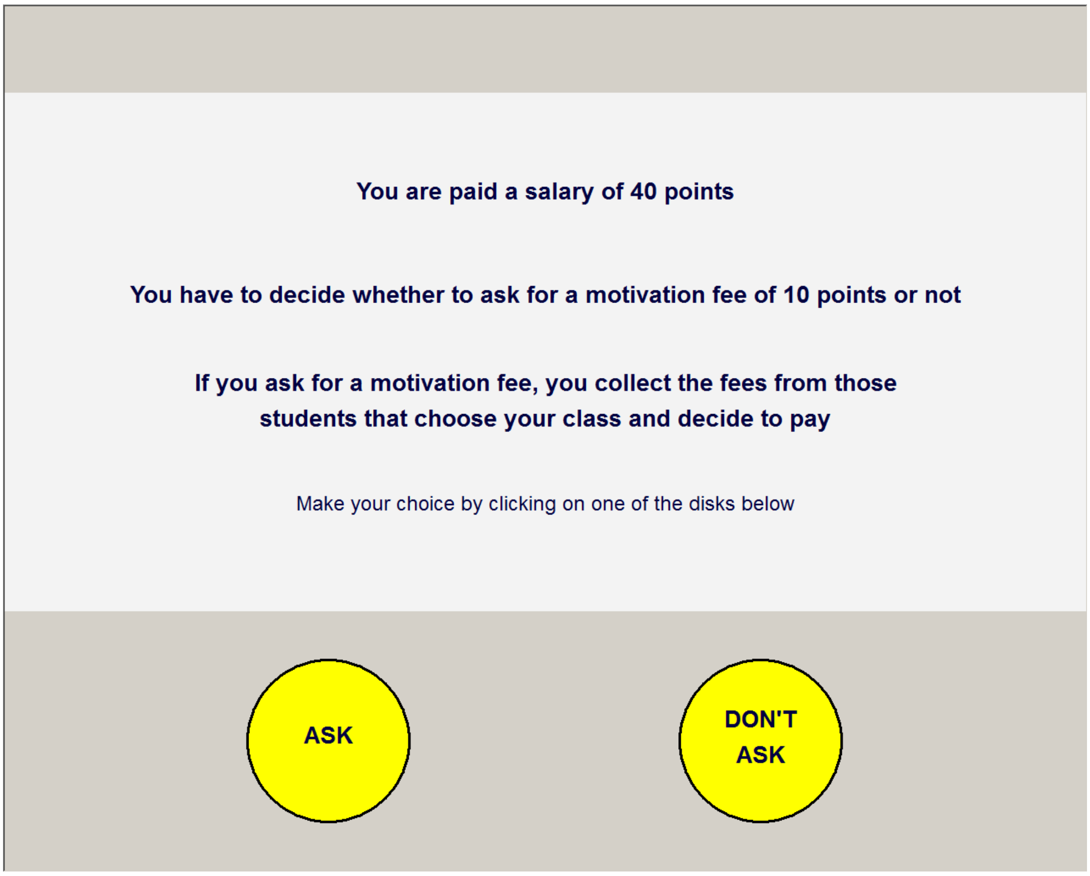


Decisions

Teachers' decision

After students are informed about their effort cost, each teacher decides at the same time whether to

1. ASK for a motivation fee OR
2. NOT ASK for a motivation fee



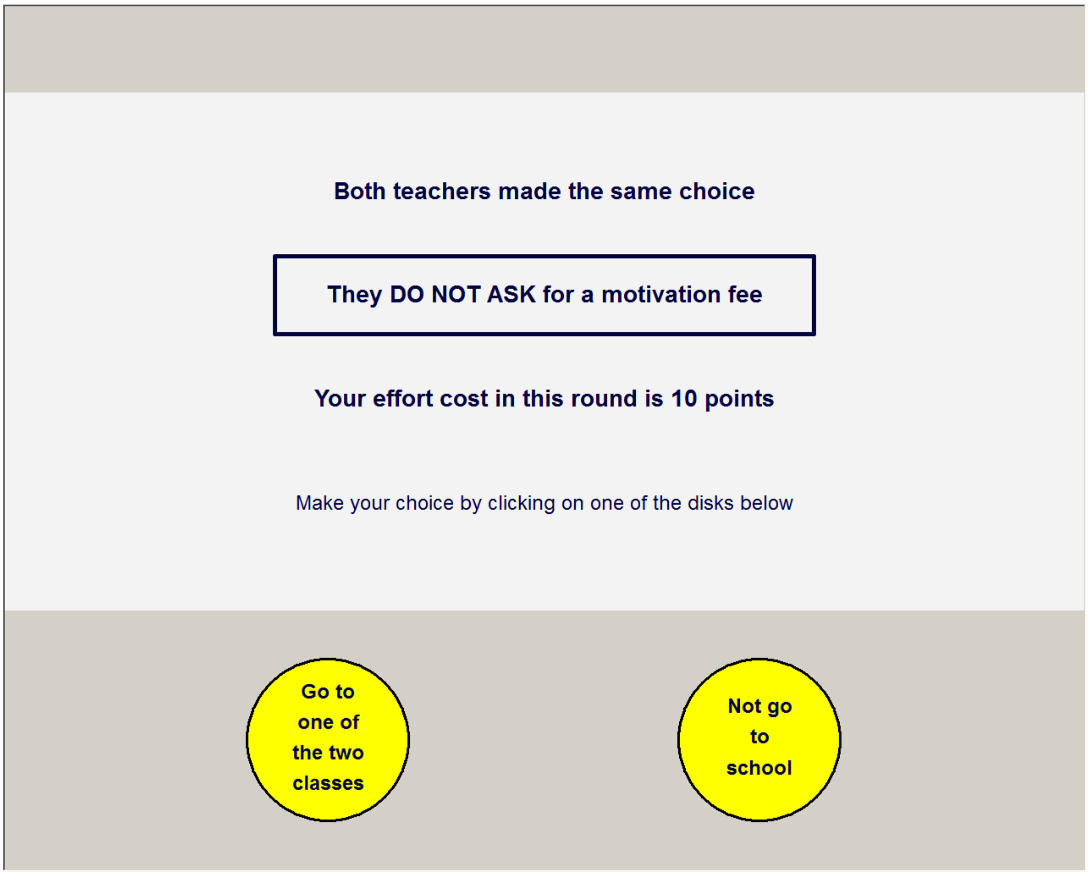
Students' decision

After both teachers decided whether to ask or not to ask for a motivation fee, students are informed about the teachers' decisions and are asked to make their own decisions. Depending on the decisions of the teachers, there are three possible scenarios.

1. Both teachers DO NOT ASK for a motivation fee

In this case students can choose either to

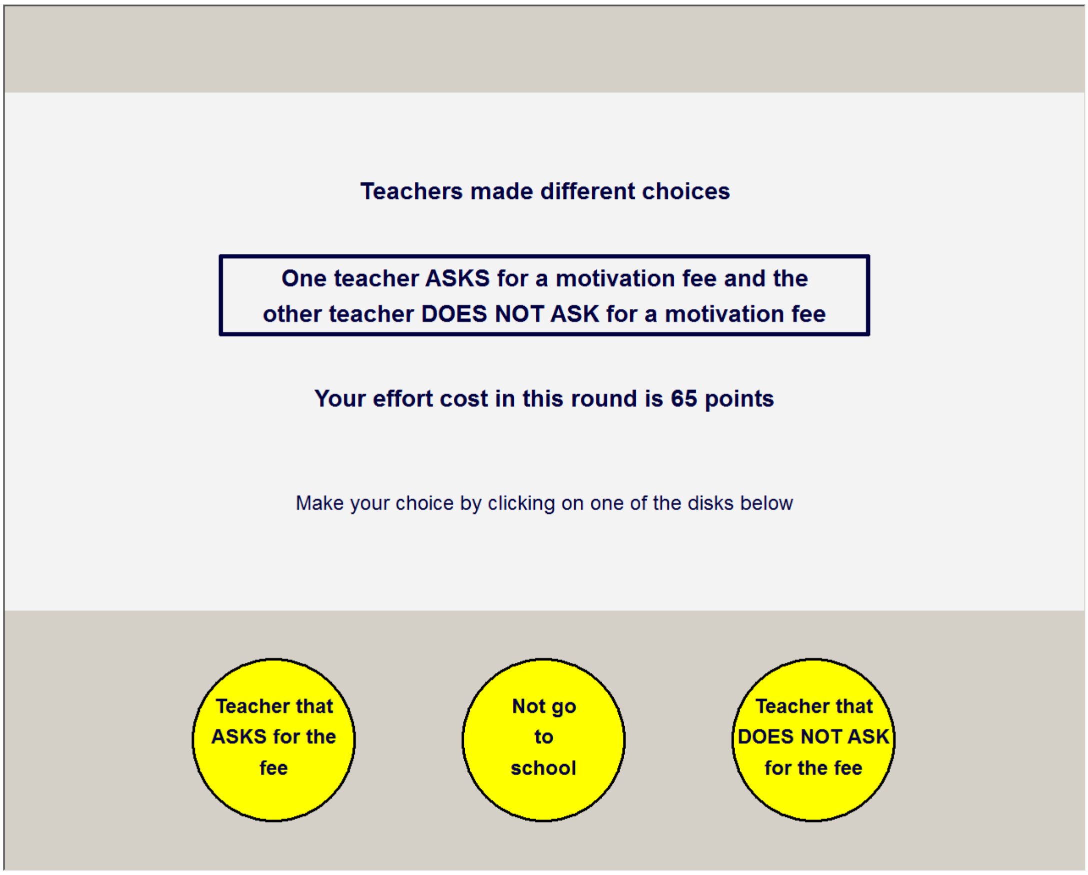
1. not go to school OR
2. be assigned to one of the two teachers and obtain the diploma by exerting effort. The computer will randomly assign half of the students choosing this option to one teacher and the other half to the other teacher.

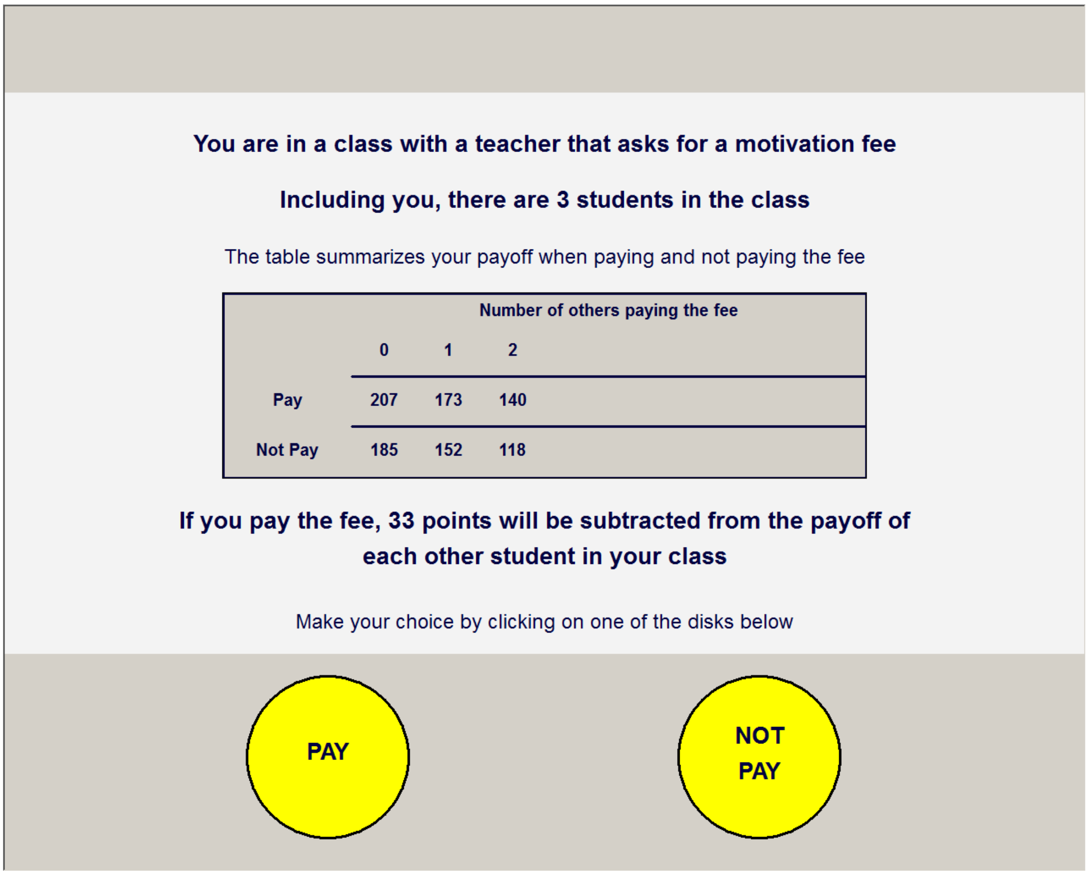


2. One teacher ASKS for a motivation fee and the other teacher DOES NOT ASK for a motivation fee

In this case, the students can choose either to:

1. not go to school OR
2. go to the teacher that DOES NOT ASK for a motivation fee and obtain the diploma by exerting effort. OR
3. go to the teacher that ASKS for a motivation fee and make a second choice whether
   (a) to PAY the motivation fee and obtain the diploma without exerting effort
   (b) NOT TO PAY the motivation fee and obtain the diploma by exerting effort


Teachers made different choices
One teacher ASKS for a motivation fee and the other teacher DOES NOT ASK for a motivation fee
Your effort cost in this round is 65 points
Make your choice by clicking on one of the disks below
Teacher that ASKS for the fee
Not go to school
Teacher that DOES NOT ASK for the fee



3. Both teachers ASK for a motivation fee. In this case students can choose to either:

1. not go to school OR
2. be assigned to one of the two teachers and make a second choice whether
   (a) to PAY the motivation fee and obtain the diploma without exerting effort
   (b) NOT TO PAY the motivation fee and obtain the diploma by exerting effort

   The computer will randomly assign half of the students choosing this option to one teacher and the other half to the other teacher.

Payoffs

Teachers' payoffs:

Teachers receive a fixed-wage for each round of **40 points.** If a teacher asks for a motivation fee, the teacher receives in addition **10 points** for each student who joins the teacher's class and pays the motivation fee.

Students' Payoffs

Not going to school leads to a payoff of **115 points**.

If students go to school the student receives a diploma worth **250 points** but there is a cost to getting the diploma.

If students are in a class where the teacher does not ask for a motivation fee, each student pays his/her individual effort cost to obtain the diploma.

> **Example:**
> In a class in which a teacher DOES NOT ASK for a motivation fee, a student with an effort cost of **10 points**, has to pay that effort cost to obtain the degree. The payoff for that student in that round therefore is **240 = 250 (value of the degree) - 10 (effort cost)**

If students are in a class where the teacher ASKS for a motivation fee, each student decides whether to pay the motivation fee or not.

The students who do not pay the motivation fee pay the effort cost and the students who pay the motivation fee do not pay the effort cost.

Independent of whether a student pays or does not pay the motivation fee, the value of the diploma is reduced according to the proportion of students in the class who pay the motivation fee. That means, if **all** students in a class pay the motivation fee the value of the diploma for each student in this class is reduced by **100 points**.

If **half** of the students pay the motivation fee the value of the diploma is reduced by $\frac{1}{2} * \mathbf{100 = 50}$ **points**

> **Example:**
> If there are **four students** in the class and **three of them pay** the motivation fee, the value of the diploma is reduced for all students by $\frac{3}{4} * \mathbf{100 = 75}$ **points**
> Therefore, if the student that does not pay the motivation fee has an effort cost of **65 points**, he obtains a payoff of **110 = 250 (value of the degree) – 65 (effort cost) – 75**.
> The students that pay the motivation fee, instead, have a payoff of **165 = 250 (value of the degree) – 10 (motivation fee) – 75** independently of their effort cost.

At the end of each round, students and teachers are informed of the results of the round before the next round is started.

Instructions for Part 1 are over. We will now ask you to answer some questions on your computer screen to ensure that you understand the instructions completely.

Summary of the decisions

Part 1 consists of 15 rounds. At the beginning of each round:

1. **Students** are **informed** about their **effort cost**
2. **Teachers** independently **decide** whether to **ask for a motivation fee** or **not**
3. **Students** are **informed** about the **decision of both teachers**. That means that they know that either: I) **both** teachers **ask for a motivation fee**

   II) **both** teachers **don't ask for a motivation fee**

   in these cases, **students decide** whether to

   (a) **not go to school at all**, OR
   (b) **be randomly assigned** to one of the classes

   III) **one teacher asks for a motivation fee** and the **one teacher doesn't ask for a motivation fee**

   in this case, **students decide** whether to either

   (a) **not go to school at all**, OR
   (b) **go** to the class in where the **teacher asks for a motivation** fee, OR
   (c) **go** to the class in where the teacher **does not ask for a motivation fee**.
4. **Students** who decide to **go to one of the two classes** will **receive information** about **how many other students are with them in the class**.
5. **If students** are in **class** with a **teacher** who **asks for a motivation fee**, **students decide** whether to **pay the motivation fee**.

Instructions for Part Two (FW)

Basic Structure

In the second part of the study, the salary for the teachers is **240 points**. This is the only change in the structure of the study compared to Part one. Below is a short summary of the payoffs for Part Two.

Teachers' payoffs

In part 2, teachers receive a fixed-wage for each round of **240 points**.

If a teacher asks for a motivation fee, the teacher receives in addition **10 points** for each student who pays the motivation fee.

Students' Payoffs

In part 2, the students' payoffs remain the same.

That means, that not going to school leads to a payoff of **115 points**.

Going to school and receiving a diploma is worth **250 points** but there is a cost to getting the diploma.

If students are in a class where the teacher does not ask for a motivation fee, each student pays his/her individual effort cost to obtain the diploma.

If students are in a class where the teacher ASKS for a motivation fee, each student decides whether to pay the motivation fee or not.

The students who do not pay the motivation fee, pay the effort cost and the students who pay the motivation fee, do not pay the effort cost.

Independent of whether a student pays or does not pay the motivation fee, the value of the diploma is reduced according to the proportion of students who pay the motivation fee. That means, if **all** students pay the motivation fee the value of the diploma for student is reduced by **100 points**.

If **half** of the students pay the motivation fee the value of the diploma is reduced by $\frac{1}{2} * \mathbf{100 = 50}$ **points**.

At the end of each round, students and teachers are informed of the results of the round before the next round is started.

Instructions for Part 2 are over. We will now ask you to answer some questions on your computer screen to ensure that you understand the instructions completely.

Instructions for Part Two (PR)

Basic Structure

In the second part of the study, the salary of the teachers depends on the number of students in their class. That means that, on top of the fixed-wage of **40 points**, a teacher receives an amount of **50 points** for each student in the class.

> **Example:**
> If 4 students are in the teacher's class, the teacher receives $\mathbf{40 + 4 * 50 = 240}$ **points** as a salary for this round.

This is the only change in the structure of the study compared to Part one. Below is a short summary of the payoffs for Part Two

Teachers' payoffs

In each round of Part 2, on top of the fixed-wage of **40 points**, teachers receive an amount of **50 points** for each student in the class.

If a teacher asks for a motivation fee, the teacher receives in addition **10 points** for each student who pays the motivation fee.

Students' Payoffs

In Part 2, the students' payoffs remain the same.

That means, that not going to school leads to a payoff of **115 points**.

Going to school and receiving a diploma is worth **250 points** but there is a cost to getting the diploma.

If students are in a class where the teacher does not ask for a motivation fee, each student pays his/her individual effort cost to obtain the diploma.

If students are in a class where the teacher ASKS for a motivation fee, each student decides whether to pay the motivation fee or not.

The students who do not pay the motivation fee, pay the effort cost and the students who pay the motivation fee, do not pay the effort cost.

Independent of whether a student pays or does not pay the motivation fee, the value of the diploma is reduced according to the proportion of students who pay the motivation fee. That means, if **all** students pay the motivation fee the value of the diploma for student is reduced by **100 points**.

If **half** of the students pay the motivation fee the value of the diploma is reduced by $\frac{1}{2} * \mathbf{100 = 50}$ **points**.

Instructions for Part 2 are over. We will now ask you to answer some questions on your computer screen to ensure that you understand the instructions completely.

C. Inequity aversion with the experimental parameters [ONLINE APPENDIX]

C.1 Inequity averse teachers with the experimental parameters

In this section, we derive predictions for the constellation of parameters used in the experiment, assuming that teachers are inequity averse. To do that, we assume that the teachers have the following utility function Fehr & Schmidt (1999)

$$U(x_i, x_{-i}) = x_i - \beta \sum_{j \neq i} \max(x_i - x_j, 0) - \alpha \sum_{j \neq i} \max(x_j - x_i, 0)$$

where: x_i is the payoff of the agent, x_{-i} are the payoffs of the other agents, α is the parameter measuring aversion to disadvantageous inequality, and β is the parameter measuring aversion to advantageous inequality.

C.1.1 Inequity averse teachers under fixed-wage (*F=40* and *F=240*)

We first derive the utility function of the teachers for three possible cases.

(i) Both teachers solicit bribes. When both T_1 and T_2 solicit bribes, all the students go to school and each teacher has 4 students in his/her class. The good students do not pay the bribe and the bad and intermediate students pay the bribe.

Let n_{Gi} be the number of good students in the class of teacher T_i. The payoff for all the students in this class is $\pi_s(n_{Gi}) = 250 - 10 - \frac{100 \cdot (4 - n_{Gi})}{4}$. The payoff for T_i is $\pi_i(n_{Gi}) = F + 10 \cdot (4 - n_{Gi})$, where F can be either 40 or 240, depending on the experimental part. Therefore, conditional on n_{G1} and n_{G2}, the utility of a teacher, when both teacher solicit bribes, takes the following form

$$\begin{aligned} U_{(1,1)}(n_{G1}, n_{G2}, \alpha, \beta) = \pi_1(n_{G1}) + & \\ & - \beta \max(\pi_1(n_{G1}) - \pi_2(n_{G2}), 0) + \\ & - \beta \left[4 \max(\pi_1(n_{G1}) - \pi_s(n_{G1}), 0) + 4 \max(\pi_1(n_{G1}) - \pi_s(n_{G2}), 0)\right] + \\ & - \alpha \max(\pi_2(n_{G2}) - \pi_1(n_{G1}), 0) + \\ & - \alpha \left[4 \max(\pi_s(n_{G1}) - \pi_1(n_{G1}), 0) + 4 \max(\pi_s(n_{G2}) - \pi_1(n_{G1}), 0)\right] \end{aligned}$$

From this, we can obtain the unconditional expected utility by summing over the possible realizations of the students' ability levels. This takes the following form

$$EU_{(1,1)}(\alpha, \beta) = \sum_{n_{G1}=0}^{4} \sum_{G_2=0}^{4} \binom{4}{n_{G1}} \left(\frac{1}{6}\right)^{n_{G1}} \left(\frac{5}{6}\right)^{4-n_{G1}} \binom{4}{n_{G2}} \left(\frac{1}{6}\right)^{n_{G2}} \left(\frac{5}{6}\right)^{4-n_{G2}} U_{(1,1)}(n_{G1}, n_{G2}, \alpha, \beta)$$

that, after simplification, becomes

$$EU_{(1,1)}(\alpha, \beta) = \frac{220}{3} - \frac{799755}{209952}\beta - \frac{140767775}{209952}\alpha$$

for $F = 40$ and

$$EU_{(1,1)}(\alpha, \beta) = \frac{820}{3} - \frac{196754975}{209952}\beta - \frac{799775}{209952}\alpha$$

for $F = 240$.

(ii) One teacher solicits bribes and the other teacher does not solicit bribes. When one teacher solicits bribes and the other does not, the good and intermediate students go to the teacher who does not solicit bribes and obtain a payoff $\pi_G = 250 - 10$, and $\pi_M = 250 - 65$, respectively. The bad students go to the teacher who solicits bribes and pay the bribe, obtaining a payoff of $\pi_B = 250 - 10 - 100$.

Let n_G be the number of good students and n_B be the number of bad students. The teacher soliciting bribes (T_1) obtains a payoff of $\pi_1(n_B) = F + n_B \cdot 10$. The teacher not soliciting bribes (T_2), instead, obtains a payoff of $\pi_2 = F$. Conditional on n_B and n_G, the utility of T_1 takes the following form

$$\begin{aligned} U_{(1,0)}(n_G, n_B, \alpha, \beta) = \pi_1(n_B) + & \\ - \beta \max(\pi_1(n_B) - \pi_2, 0) + & \\ - \beta \left[n_G \max(\pi_1(n_B) - \pi_G, 0) + (8 - n_G - n_B) \max(\pi_1(n_B) - \pi_M, 0) + n_B \max(\pi_1(n_B) - \pi_B, 0) \right] + & \\ - \alpha \max(\pi_2 - \pi_1(n_B), 0) + & \\ - \alpha \left[n_G \max(\pi_G - \pi_1(n_B), 0) + (8 - n_G - n_B) \max(\pi_M - \pi_1(n_B), 0) + n_B \max(\pi_B - \pi_1(n_B), 0) \right] & \end{aligned}$$

and the unconditional expected utility is

$$EU_{(1,0)}(\alpha, \beta) = \sum_{n_G=0}^{8} \sum_{n_B=0}^{8-n_G} \frac{8!}{n_G! n_B! (8 - n_G - n_B)!} \left(\frac{1}{6}\right)^{n_G + n_B} \left(\frac{4}{6}\right)^{8 - n_G - n_B} U_{(1,0)}(n_G, n_B, \alpha, \beta)$$

After simplifications, this becomes

$$EU_{(1,0)}(\alpha, \beta) = \frac{160}{3} - \frac{40}{3}\beta - \frac{3200}{3}\alpha$$

for $F = 40$ and

$$EU_{(1,0)}(\alpha, \beta) = \frac{760}{3} - \frac{1640}{3}\beta$$

for $F = 240$.

Similarly, one can compute the conditional and unconditional utilities given n_G and n_B for the teacher not soliciting bribes. In this case, the unconditional expected utility, i.e., $EU_{(0,1)}(\alpha, \beta)$, becomes

$$EU_{(0,1)}(\alpha, \beta) = 40 - \frac{3560}{3}\alpha$$

for $F = 40$ and

$$EU_{(0,1)}(\alpha, \beta) = 240 - \frac{1280}{3}\beta - \frac{40}{3}\alpha$$

for $F = 240$.

(iii) Both teachers do not solicit bribes. When both T_1 and T_2 do not solicit bribes, the good and intermediate students go to school and respectively obtain a payoff of $\pi_G = 250 - 10$ and $\pi_M = 250 - 65$, independent of the class they go to. The bad students, instead, do not go to school and obtain $\pi_B = 115$. Teachers obtain a payoff of $\pi_i = F$, independent of the number of students in the class.

Let n_G be the number of good students and n_B be the number of bad students. The utility of a teacher, conditional on n_B and n_G, is given by

$$\begin{aligned} U_{(0,0)}(n_G, n_B, \alpha, \beta) = \pi_1 & \\ & - \alpha \left[n_G \max(\pi_1 - \pi_G, 0) + (8 - n_G - n_B) \max(\pi_1 - \pi_M, 0) + n_B \max(\pi_1 - \pi_B, 0) \right] \\ & \quad - \beta \left[n_G \max(\pi_G - \pi_1, 0) + (8 - n_G - n_B) \max(\pi_M - \pi_1, 0) + n_B \max(\pi_B - \pi_1, 0) \right] \end{aligned}$$

and the unconditional expected utility is given by

$$EU_{(0,0)}(\alpha, \beta) = \sum_{n_G=0}^{8} \sum_{n_B=0}^{8-n_G} \frac{8!}{n_G! n_B! (8 - n_G - n_B)!} \left(\frac{1}{6}\right)^{n_G+n_B} \left(\frac{4}{6}\right)^{8-n_G-n_B} U_{(0,0)}(n_G, n_B, \alpha, \beta)$$

After simplifications, the unconditional expected utility becomes

$$EU_{(0,0)}(\alpha, \beta) = 40 - 1140\alpha$$

for $F = 40$ and

$$EU_{(0,0)}(\alpha, \beta) = 240 - 460\beta$$

for $F = 240$.

Payoff matrix for the teachers under fixed-wage.

After having obtained the utility of the teachers for all cases, we can write the payoff matrix for both Low fixed-wage and High fixed-wage.

Low fixed-wage. When $F = 40$, we have the following payoff matrix for the teachers

		T_2: $B_2 = 1$	T_2: $B_2 = 0$
T_1	$B_1 = 1$	$\frac{220}{3} - \frac{799755}{209952}\beta - \frac{140767775}{209952}\alpha$	$\frac{160}{3} - \frac{40}{3}\beta - \frac{3200}{3}\alpha$
	$B_1 = 0$	$40 - \frac{3560}{3}\alpha$	$40 - 1140\alpha$

Note that, conditional on T_2 soliciting bribes, T_1 prefers $B_1 = 1$ over $B_1 = 0$ when

$$\beta \leq \frac{1399680}{159955} + \frac{21675053}{159955}\alpha$$

and, conditional on T_2 not soliciting bribes, T_1 prefers $B_1 = 1$ over $B_1 = 0$ when

$$\beta \leq 1 + \frac{11}{2}\alpha$$

These inequalities identify regions with different equilibria of the game. These are reported in panel (a) of Figure 3 in the main paper.

High fixed-wage. When $F = 240$, we have the following payoff matrix for the teachers

		T_2	
		$B_2 = 1$	$B_2 = 0$
T_1	$B_1 = 1$	$\frac{820}{3} - \frac{196754975}{209952}\beta - \frac{799775}{209952}\alpha$	$\frac{760}{3} - \frac{1640}{3}\beta$
	$B_1 = 0$	$240 - \frac{1280}{3}\beta - \frac{40}{3}\alpha$	$240 - 460\beta$

Note that, conditional on T_2 soliciting bribes, T_1 prefers $B_1 = 1$ over $B_1 = 0$ when

$$\beta \leq \frac{1399680}{21435091} + \frac{399917}{21435091}\alpha$$

and, conditional on T_2 not soliciting bribes, T_1 prefers $B_1 = 1$ over $B_1 = 0$ when

$$\beta \leq \frac{2}{13}$$

These inequalities identify regions with different equilibria of the game. These are reported in panel (b) of Figure 3 in the main paper.

C.1.2 Inequity averse teachers under piece-rate

We first derive the utility functions for the teachers for all three possible cases. Note that, the payoffs for the students remain the same as in the fixed-wage cases. The only payoffs that can change are the ones for the teachers. These payoffs now depend on the number of students in the class.

(i) Both teachers solicit bribes. When both teachers solicit bribes, there are no differences compared to the fixed-wage case. Let n_{Gi} be the number of good students in the class of teacher T_i. The payoff for T_i is $\pi_i(n_{Gi}) = 40 + 4 \cdot 50 + 10 \cdot (4 - n_{Gi})$, which is the same as the payoff teachers obtain in the fixed-wage scenario with $F = 240$. Therefore, the expected utility for teachers who both solicit bribes is the same as in the fixed-wage with $F = 240$.

(ii) One teacher solicits bribes and the other teacher does not solicit bribes. Compared to the fixed-wage case, the payoffs for the two teachers change. Payoffs now depend on the number of bad students. Let n_B be the number of bad students. The teacher who solicits bribes (T_1) obtains a payoff $\pi_1(n_B) = 40 + n_B \cdot 50 + n_B \cdot 10$ and the teacher who does not solicit bribes (T_2) obtains a payoff $\pi_2(n_B) = 40 + (8 - n_B) \cdot 50$. The utility of T_1 conditional on n_B is

$$\begin{aligned}
U_{(1,0)}(n_G, n_B, \alpha, \beta) = \pi_1(n_B) + \\
- \beta \max(\pi_1(n_B) - \pi_2(n_B), 0) + \\
- \beta \left[n_G \max(\pi_1(n_B) - \pi_G, 0) + (8 - n_G - n_B) \max(\pi_1(n_B) - \pi_M, 0) + n_B \max(\pi_1(n_B) - \pi_B, 0)\right] + \\
- \alpha \max(\pi_2(n_B) - \pi_1(n_B), 0) + \\
- \alpha \left[n_G \max(\pi_G - \pi_1(n_B), 0) + (8 - n_G - n_B) \max(\pi_M - \pi_1(n_B), 0) + n_B \max(\pi_B - \pi_1(n_B), 0)\right]
\end{aligned}$$

In this case, the unconditional expected utility is

$$EU_{(1,0)}(\alpha,\beta) = \sum_{n_G=0}^{8}\sum_{n_B=0}^{8-n_G} \frac{8!}{n_G! n_B! (8-n_G-n_B)!} \left(\frac{1}{6}\right)^{n_G+n_B} \left(\frac{4}{6}\right)^{8-n_G-n_B} U_{(1,0)}(n_G, n_B, \alpha, \beta)$$

that, after simplification, becomes

$$EU_{(1,0)}(\alpha,\beta) = 120 - \frac{2151595}{26244}\beta - \frac{22796875}{26244}\alpha$$

Similarly, one can define the conditional utility for the teacher who does not solicit bribes and compute the unconditional expected utility. After simplifications, the equation for the expected utility $EU_{(0,1)}(\alpha,\beta)$ takes the following form

$$EU_{(0,1)}(\alpha,\beta) = \frac{1120}{3} - \frac{183564625}{104976}\beta - \frac{206545}{104976}\alpha$$

(iii) Both teachers do not solicit bribes. This case differs slightly from the previous cases. Let n_G and n_B be the number of good and bad students. Note that, when both teachers do not solicit bribes, they share the good and intermediate students. Therefore, when a teacher has $\left\lceil \frac{8-n_B}{2} \right\rceil$ students in the class the other teacher has $\left\lfloor \frac{8-n_B}{2} \right\rfloor$ students in the class. The payoff of the former is $\pi_+(n_B) = 40 + \left\lceil \frac{8-n_B}{2} \right\rceil \cdot 50$ and the payoff of the latter is $\pi_-(n_B) = 40 + \left\lfloor \frac{8-n_B}{2} \right\rfloor \cdot 50$. Moreover, a teacher obtains $\pi_+(n_B)$ or $\pi_-(n_B)$ with equal probability. Therefore, the utility of a teacher conditional on n_B and n_G is either

$$\begin{aligned} U^+_{(0,0)}(n_G, n_B, \alpha, \beta) = {} & \pi_+(n_B) + \\ & - \beta \max(\pi_+(n_B) - \pi_-(n_B), 0) + \\ & - \beta \left[n_G \max(\pi_+(n_B) - \pi_G, 0) + (8 - n_G - n_B) \max(\pi_+(n_B) - \pi_M, 0) + n_B \max(\pi_+(n_B) - \pi_B, 0) \right] + \\ & - \alpha \max(\pi_-(n_B) - \pi_+(n_B), 0) + \\ & - \alpha \left[n_G \max(\pi_G - \pi_+(n_B), 0) + (8 - n_G - n_B) \max(\pi_M - \pi_+(n_B), 0) + n_B \max(\pi_B - \pi_+(n_B), 0) \right] \end{aligned}$$

with probability $\frac{1}{2}$ or

$$\begin{aligned} U^-_{(0,0)}(n_G, n_B, \alpha, \beta) = {} & \pi_-(n_B) + \\ & - \beta \max(\pi_-(n_B) - \pi_+(n_B), 0) + \\ & - \beta \left[n_G \max(\pi_-(n_B) - \pi_G, 0) + (8 - n_G - n_B) \max(\pi_-(n_B) - \pi_M, 0) + n_B \max(\pi_-(n_B) - \pi_B, 0) \right] + \\ & - \alpha \max(\pi_+(n_B) - \pi_-(n_B), 0) + \\ & - \alpha \left[n_G \max(\pi_G - \pi_-(n_B), 0) + (8 - n_G - n_B) \max(\pi_M - \pi_-(n_B), 0) + n_B \max(\pi_B - \pi_-(n_B), 0) \right] \end{aligned}$$

with probability $\frac{1}{2}$. Therefore, the unconditional expected utility is

$$\begin{aligned} EU_{(0,0)}(\alpha,\beta) = {} & \sum_{n_G=0}^{8}\sum_{n_B=0}^{8-n_G} \frac{8!}{n_G! n_B! (8-n_G-n_B)!} \left(\frac{1}{6}\right)^{n_G+n_B} \left(\frac{4}{6}\right)^{8-n_G-n_B} \cdot \\ & \cdot \frac{1}{2} \left[U^+_{(0,0)}(n_G, n_B, \alpha, \beta) + U^-_{(0,0)}(n_G, n_B, \alpha, \beta) \right] \end{aligned}$$

After simplifications, it becomes

$$EU_{(0,0)}(\alpha,\beta) = \frac{620}{3} - \frac{36176125}{139968}\beta - \frac{9115645}{139968}\alpha$$

Payoff matrix for the teachers under piece-rate.

After having obtained the utility for the teachers in all cases, we can write the payoff matrix for the piece-rate regime. When $F = 40$ and piece-rate of $s = 50$, we have the following payoff matrix for the teachers:

		T_2: $B = 1$	T_2: $B = 0$
T_1	$B = 1$	$\frac{820}{3} - \frac{196754975}{209952}\beta - \frac{799775}{209952}\alpha$	$120 - \frac{2151595}{26244}\beta - \frac{22796875}{26244}\alpha$
	$B = 0$	$\frac{1120}{3} - \frac{183564625}{104976}\beta - \frac{206545}{104976}\alpha$	$\frac{620}{3} - \frac{36176125}{139968}\beta - \frac{9115645}{139968}\alpha$

Note that, conditional on T_2 soliciting bribes, T_1 prefers $B_1 = 1$ over $B_1 = 0$ when

$$\beta \geq \frac{466560}{3786095} + \frac{8593}{3786095}\alpha$$

and, conditional on T_2 not soliciting bribes, T_1 prefers $B_1 = 1$ over $B_1 = 0$ when

$$\beta \geq \frac{7278336}{14820571} + \frac{67480613}{14820571}\alpha$$

These inequalities identify regions with different equilibria of the game. These are reported in panel (c) of Figure 3 in the main paper.

C.2 Inequity averse students with the experimental parameters

Here we show that, for a constellation of parameters estimated in the literature (Goeree & Holt 2000, Blanco et al. 2011, Beranek et al. 2015), students' inequity aversion does not generate incentives to deviate from the selfish equilibrium strategy.

To check this, we verified that each type of student, i,e, bad, intermediate, and good, has no profitable deviation from the selfish strategy profile when assuming inequity aversion. The procedure is as follows: (i) we fix the type of one student and assume that the teachers and the other students follow the equilibrium strategy; (ii) we calculate the expected utility of this student for all his/her possible strategies;[12] (iii) we compare the expected utility of the equilibrium strategy and the utility of the alternative strategies.

Results of these calculations are reported in Table C.3. The table shows the results for the calculations for a range of estimated inequity aversion parameters, and for the following three cases: (i) Fixed Wage when teachers do not solicit bribes; (ii) Piece rate when both teachers do not solicit bribes; and (iii) Piece rate when both teachers solicit bribes. Comparing the expected utility of the equilibrium strategy to the expected utility of the best alternative strategy, we show that, in all these cases, students do not have an incentives to deviate.[13]

[12] For each strategy, we obtained the expected utility by calculating the utility over all the 128 possible combination of other students' types, multiplying each of these for the probability to observe that distribution, and taking the sum.

[13] Note that the equilibrium is not a strict equilibrium in case (iii)

Table C.3: Students' incentive to deviate from the equilibrium with Inequity Averse teachers. Estimated expected utility for the three cases.

Paper	BEN (2011) $\alpha = 0.910$; $\beta = 0.380$		GH prop (2000) $\alpha = 0.860$; $\beta = 0.660$		GH resp (2000) $\alpha = 0.860$; $\beta = 0.120$		BCG Nott. (2015) $\alpha = 0.754$; $\beta = 0.484$		BCG Izmir (2015) $\alpha = 1.227$; $\beta = 0.589$		BCG M-Turk (2015) $\alpha = 1.218$; $\beta = 0.410$	
(i) Fixed Wage when both teachers do not solicit bribes												
Student type	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.
Bad	-542.475	-772.225	-506.350	-724.850	-506.350	-724.850	-429.765	-624.415	-771.507	-1072.582	-765.005	-1064.055
Intermediate	-4.525	-542.475	-18.683	-506.350	25.417	-506.350	14.152	-429.765	-76.804	-771.507	-60.618	-765.005
Good	87.050	-542.475	-25.650	-506.350	191.700	-506.350	45.190	-429.765	2.928	-771.507	74.975	-765.005
(ii) Piece Rate when both teachers do not solicit bribes												
Student type	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.
Bad	-444.533	-719.142	-413.940	-674.684	-413.671	-674.684	-348.698	-580.432	-639.486	-1001.008	-633.865	-993.005
Intermediate	41.837	-444.533	23.564	-413.940	70.477	-413.671	51.685	-348.698	-14.691	-639.486	1.950	-633.865
Good	64.883	-444.533	-64.150	-413.940	184.700	-413.671	16.957	-348.698	-31.431	-639.486	51.058	-633.865
(c) Piece Rate when both teachers solicit bribes												
Student type	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.	EU eq	EU best alt.
Bad	-119.191	-119.191	-112.448	-112.448	-97.759	-97.759	-77.217	-77.217	-215.920	-215.920	-208.466	-208.466
Intermediate	-119.191	-119.191	-112.448	-112.448	-97.759	-97.759	-77.217	-77.217	-215.920	-215.920	-208.466	-208.466
Good	-32.885	-32.885	-49.170	-49.170	0.805	0.805	-12.472	-12.472	-113.264	-113.264	-94.965	-94.965

D. Additional analysis [ONLINE APPENDIX]

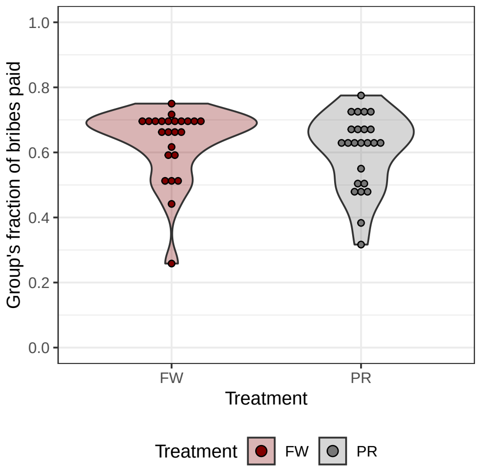
D.1 Overall bribing

Figure D.1 reports information regarding fraction of bribes paid. Note that bribes are not paid either because the students did not have the opportunity to pay or because the student had the opportunity and decided not to pay the bribe. Results are similar to the ones obtained looking at the fraction of teachers soliciting bribes, which are reported in the paper. The fraction of bribes paid in part 1 does not differ across the two treatments. A Wilcoxon rank sum test does not reject the null hypothesis that the probability of a random observation from one treatment exceeding a random observation from the other is one half. In part 2, the fraction of bribes paid is significantly lower in PR compared to FW (Wilcoxon rank sum test $p < 0.001$). As for the comparison of part 1 and part 2, we find a significant difference in overall bribing in PR (Wilcoxon signed rank test $p < 0.001$) but not in FW (Wilcoxon signed rank test $p = 0.170$).

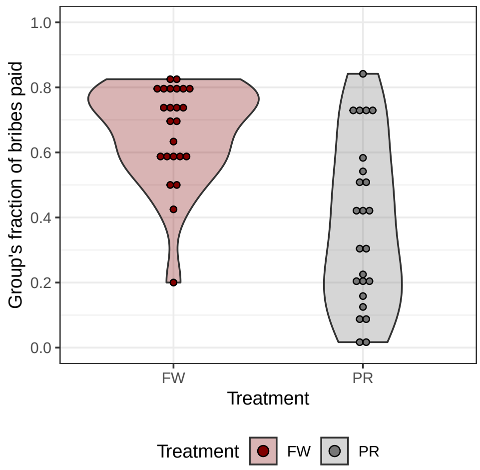
Panel (c) of Figure D.1 reports the theoretical fraction of bribes that should have been paid given the choice of the teachers and the ability of the students (crosses), along with the observed fraction of bribes paid (dots). With fixed-wage the actual frequency of bribes paid is slightly lower than the theoretically predicted fraction.

Figure D.1: Group's fraction of bribes paid in part 1 (a) and in part 2 (b) and fraction of students paying bribes over periods (c)

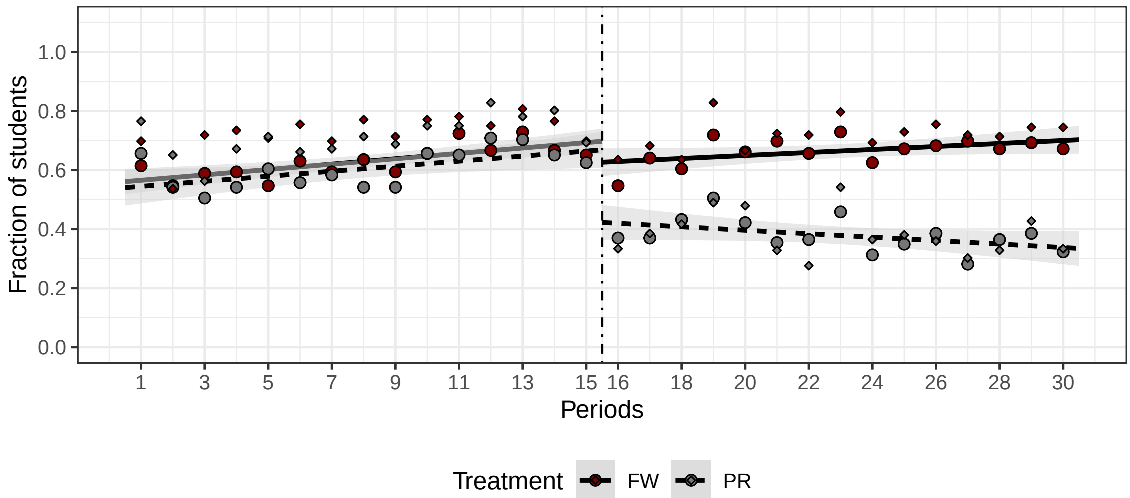
(a) Part 1



(b) Part 2



(c) Fraction of bribes paid over periods (diamonds are the predicted fractions when students follow the equilibrium strategy after observing the decisions of the teachers)

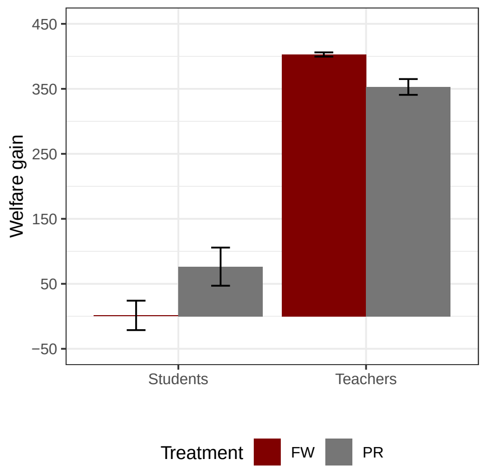


D.2 Social welfare

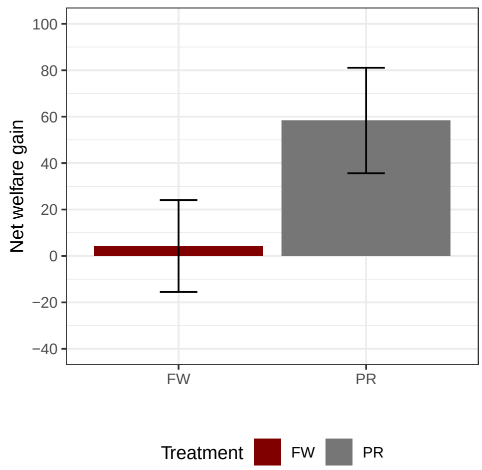
We analyze who benefits the most from the two interventions by comparing the increase in total payoffs for teachers and for students when moving from part 1 to part 2. As is apparent from Figure D.2 panel (a), which depicts the increase in the group total welfare for teachers and students by treatment, the introduction of a high fixed-wage particularly benefits teachers. They receive about 400 extra points per period (Wilcoxon signed rank test $p < 0.001$), while the students do not benefit at all (Wilcoxon signed rank test $p = 0.900$). The introduction of the piece-rate bonus, on the other hand, benefits both teachers and students who enjoy an extra payoff of about 350 and 80 points per period, respectively (Wilcoxon signed rank test $p < 0.001$ in both cases).

Figure D.2: Welfare difference between part 2 and part 1 for students and teachers by treatment (panel (a)) and total efficiency gain between part 1 and part 2, i.e., increase in group total payoff net of the public expenditure, by treatment (panel (b)). (Mean and 95% C.I.)

(a) Welfare gain by role



(b) Net welfare gain



Although students overall prefer the introduction of the piece-rate bonus over fixed-wage increase (Wilcoxon signed rank test $p < 0.001$ in both cases), differences in the preference across student types exists. That is, while the piece-rate scheme benefits students with low and medium effort costs, the minority of students with high effort costs prefer to obtain their diploma by paying a bribe (and not exerting effort). Teachers, on the other hand, would prefer the introduction of a high fixed-wage over a piece-rate bonus (Wilcoxon rank sum test $p < 0.001$). The reason lies in the disincentive to solicit bribes in the piece-rate regime which, in turn, produces an incentive to choose the outside option for the students with high effort costs. Hence, the total wage paid to the teachers becomes smaller. This, however, implies that the piece-rate bonus policy is on average less expensive.

Finally, to estimate the overall welfare benefits of the two interventions we account for the different public expenditure levels. Figure D.2 panel (b) illustrates that the fixed-wage increase fails to improve net welfare when moving from part 1 to part 2 (Wilcoxon signed rank test $p = 0.705$). The introduction of the piece-rate scheme, on the other hand, produces a significant welfare gain (Wilcoxon signed rank test $p < 0.001$). Taken together, the new results suggest that introducing a piece-rate scheme in corrupt education systems can help to reduce the occurrence of bribe transactions and improve overall social welfare, in particular for students (with low and medium effort costs).

D.3 Dynamic play of teachers

Here we look at the dynamics of teachers' behavior in the second part of the experiment. This analysis permits gaining deeper insights into how the shift from a corrupt system to one with less corruption occurs. In particular, we study how teachers react to their own choices, as well as, to the other teachers' choices. Table D.1 reports the fraction of times a teacher solicited bribes conditional on his/her choice, and the other teacher's choice in the previous period. In the *FW* treatment, teachers tend to solicit bribes independent of the previous period's choices. This is in line with selfish preferences and not with inequity aversion. In the *PR* treatment, instead, teachers tend to solicit bribes if the other teacher did so in the previous period (more strongly if both did so) and tend to abstain from soliciting bribes if the other teacher also abstained. This is not in line with the predictions based on selfish preferences. Instead, in combination with the pattern observed in treatment *FW*, it suggests that the dynamics are in line with a tit-for-tat type of behavior, which is compatible with incomplete conditional cooperation (Fischbacher & Gächter 2010).

Table D.1: Teachers' fraction of "solicit" choices conditional on previous period decisions (part 2 data); + means that behavior is in agreement with a model; - that it is in disagreement

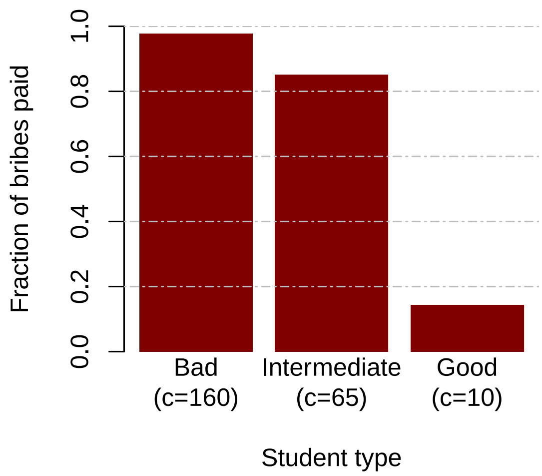
	Both do not solicit	Solicit and the other does not	Do not solicit and the other does	Both solicit
Piece-Rate	26.3%	33.9%	54.1%	83.7%
Inequity Aversion	+	+	+/−	+
Selfish Preferences	+	+	+/−	−
Tit-For-Tat	+	+	+/−	+
Fixed-Wage	75.0%	71.2%	88.1%	93.4%
Inequity Aversion	−	−	−	−
Selfish Preferences	+	+	+	+
Tit-For-Tat	+	+	+	+

D.4 Students individual behavior

D.4.1 Choice to pay the bribe or not

Figure D.3 shows the fraction of bribes paid by effort cost. These frequencies are close to the predicted frequencies of 1, 1, and 0 for the bad, the intermediate, and the good students, respectively.

Figure D.3: Fraction of bribes paid by effort level



Deviations from the equilibrium predictions are very rare when they are extremely costly, i.e., not paying the bribe when the effort cost is high, and they are more common when the cost of deviating is smaller, i.e., in case of medium and low effort cost.

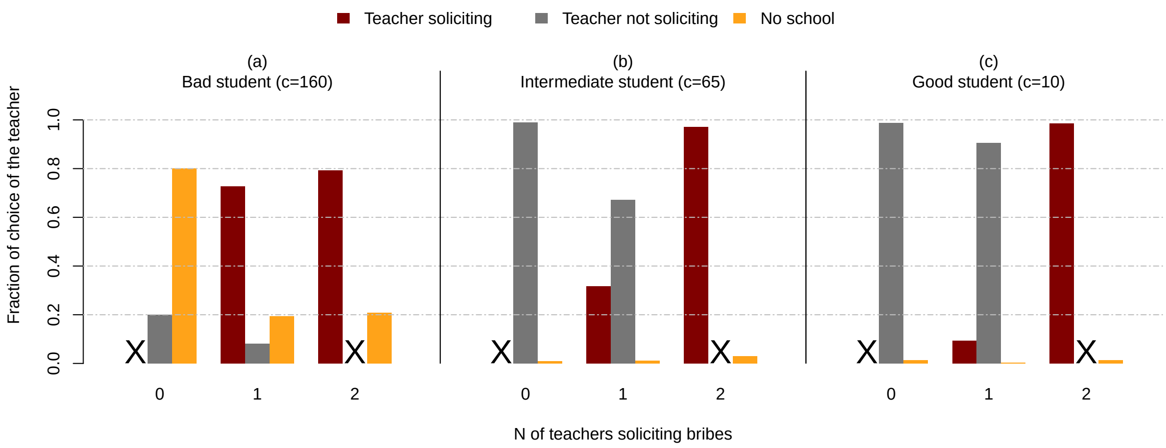
D.4.2 Choice to go to school

Figure D.4 reports the students' choice of the teacher by effort cost and by number of teachers who solicit bribes. When the number of teachers soliciting bribes is 0, the choice is between the "Teacher not soliciting" and "No school". When the number of teachers soliciting bribes is 2, the choice is between the "Teacher soliciting" and "No school". When the number of teachers soliciting bribes is 1, the choice is among all three available options.

Recall that the equilibrium predictions are as follows:

- When the number of teacher soliciting bribes is 0, bad students should choose "No school" and the other students should choose "Teacher not soliciting".
- When the number of teacher soliciting bribes is 1, bad students should choose "Teacher soliciting" and the other students should choose "Teacher not soliciting".
- When the number of teacher soliciting bribes is 2, all students' types should choose "Teacher soliciting".

Figure D.4: Students' choice of the teacher by student's type cost and options available



As for good students, panel (c) in Figure D.4 illustrates that the observed fractions of choices are largely in line with the equilibrium predictions. The highest deviation rate occurs when one teacher solicits bribes and the other teacher does not solicit bribes. In this case, 9.5% of the choices deviate from the ones predicted by the equilibrium.

For the intermediate students, choices are in line with predictions when both teachers make the same choice (0 and 2). When one teacher solicits and the other does not solicit bribes, about one third of the choices are for "Teacher soliciting" (31.6%). This is not compatible with the equilibrium predictions, but it can be profitable if other students do not pay the bribe when they should. Recall that "Teacher not soliciting" guarantees a payoff of 185. "Teacher soliciting", instead, can provide a higher amount only when paying the bribe and some other students do not pay the bribe. Empirically, choosing a "Teacher soliciting" and behaving optimally (pay when more than 2 students are in the class and not paying when alone in the class) gives a payoff $\geq$ 185 only 46 out of 425 times (10.82%).

As for the bad students, when "Teacher soliciting" is available, 20% of the choices are for "No school". This may be due to a preference for not paying the bribe. Indeed, these students are willing to give away about 25 points to avoid paying the bribe to the teacher (In case of a fully corrupt class of 4 students, choosing "Teacher soliciting" and paying gives 140 points and "No school"gives 115). Moreover, about 20% of the choices are for "Teacher not soliciting" when the "No school" option is available (payoff of 90 instead of 115).

D.5 Moral vs equilibrium play

Previous results show that students' behavior is close to the equilibrium predictions most of the time. Here, we look at whether students decisions can be explained by moral concerns. To do so, we compare two possible styles of play for the students: equilibrium play and moral play. Moral play is defined by choosing the best strategy from the set that excludes paying a bribe. Table D.2 summarizes the predicted choices for all possible situations for the two styles of play.

Table D.2: Summary of the choices for the different styles of play (N.S.= no school; B.F.S. = bribe free school (teacher not soliciting); B. S. = bribe school (teacher soliciting))

N. teachers soliciting bribes	Effort Cost	Equilibrium	Moral	Mistake
0	High	N.S.	N.S.	B.F.S.
1	High	B.S. + pay	N.S.	B.F.S. B.S. + not pay
2	High	B.S. + pay	N.S.	B.S. + not pay
0	Medium	B.F.S.	B.F.S.	N.S.
1	Medium	B.F.S.	B.F.S.	N.S. B.S. + pay B.S. + not pay
2	Medium	B.S. + pay	B.S. + not pay N.S.	—
0	Low	B.F.S.	B.F.S.	N.S.
1	Low	B.F.S.	B.F.S.	N.S. B.S. + pay B.S. + not pay
2	Low	B.S. + not pay	B.S. + not pay N.S.	B.S. + pay

Figure D.5 panel (a) shows the fraction of periods where choices are consistent with equilibrium play and moral play for each individual. Note that the sum of the fractions can be greater than 1. This is due to the fact that some choice patterns are compatible with both equilibrium play and moral play. In general, equilibrium play seems to better predict behavior—most of the individuals are in the bottom right corner—. The subjects with a fraction of equilibrium choices strictly higher than the fraction of moral choices are 348 out of 384.

Figure D.5: Students' individual fraction of choices in line with equilibrium play (x) and with moral play (y). Points represents participants. Panel (a) uses all choices and panel (b) only the choices where the predictions differ.

(a) All choices (b) Moral ≠ equilibrium

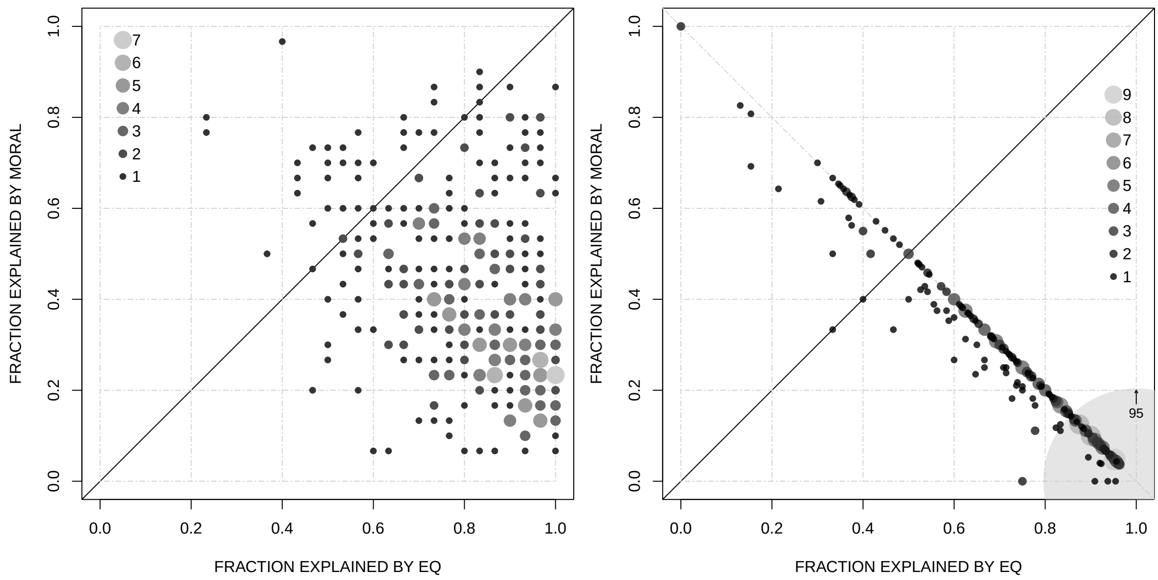


Figure D.5b panel (b) takes into account the fact that the predictions of the two styles may overlap. Therefore, the Figure considers only the data of those situations in which equilibrium play and moral play provide different predictions—i.e., when effort cost is low and there is at least one teacher soliciting bribes and when effort cost is medium and both teachers solicit bribes. In contrast to the panel (a), the sum of the fraction of choices explained by the two styles cannot be greater than one. Figure D.5b panel (b) illustrates that equilibrium play seems to better predict behavior (most of the individuals are in the bottom right corner). In this case, the subjects with a fraction of equilibrium choices strictly higher than the fraction of moral choices are 349 out of 384.

As a second step, we look at differences in the fraction of the two styles of play across treatments. Figure D.6 and Figure D.7 replicate the analyses in Figure D.5b for the PR and for the FW treatment separately. Comparing the number of subjects with a fraction of equilibrium choices strictly higher than the fraction of moral choices in the two treatments does not reveal significant differences in the style of play under piece-rate versus under fixed-wage (170 subjects out of 192 in PR and 179 out of 192 in FW; $\chi^2(1) = 2.546, p = 0.111$). Analyzing the number of subjects that consistently play according equilibrium predictions leads to the conclusion (45 subjects out of 192 in PR and 50 out of 192 in FW; $\chi^2(1) = 0.350, p = 0.554$).

Figure D.6: Students' individual fraction of choices in line with equilibrium play (x) and with moral play (y) in the PR treatment. Points represents participants. Panel (a) uses all choices and panel (b) only the choices where the predictions differ.

(a) All choices (b) Moral ≠ equilibrium

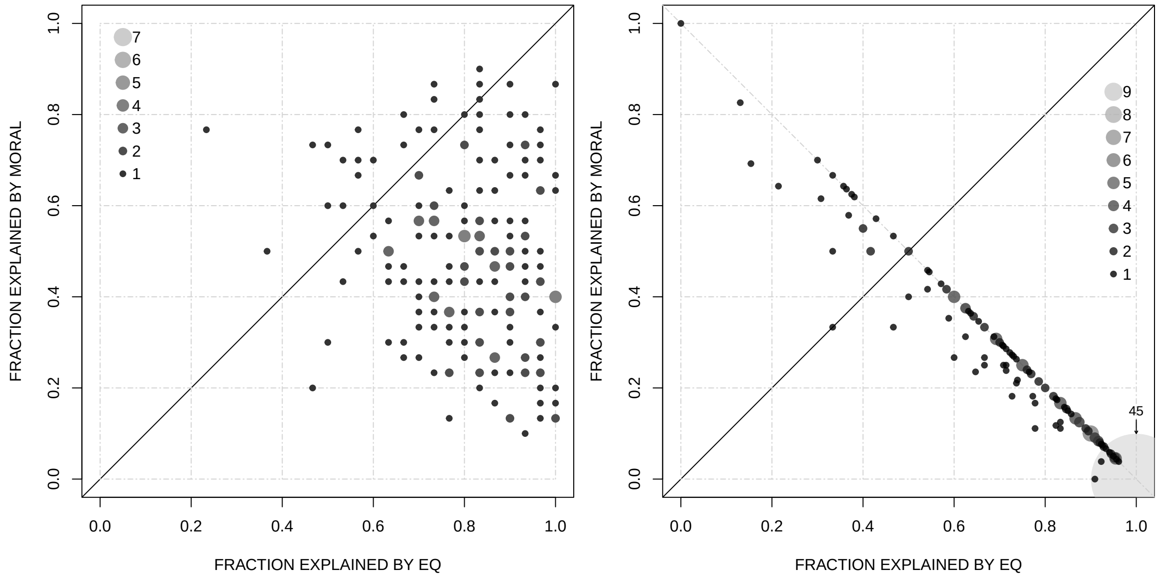


Figure D.7: Students' individual fraction of choices in line with equilibrium play (x) and with moral play (y) in the FW treatment. Points represents participants. Panel (a) uses all choices and panel (b) only the choices where the predictions differ.

(a) All choices (b) Moral ≠ equilibrium