I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13474

Selection and Incentives under Time Pressure: The Importance of Framing

Maria De Paola
Francesca Gioia
Valeria Pupo

JULY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13474

Selection and Incentives under Time Pressure: The Importance of Framing

Maria De Paola
University of Calabria and IZA

Francesca Gioia
University of Milan

Valeria Pupo
University of Calabria

JULY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Selection and Incentives under Time Pressure: The Importance of Framing*

In this paper we investigate whether the framing of the incentives used to foster participation into contexts characterized by high degrees of time pressure affects individuals' self-selection. At this aim we run a lab-in-the-field experiment structured in two parts. The first part investigates individual characteristics that affect performance under time pressure, while the second is devoted to analyze how the decision to work under time pressure is affected by the reward/punishment framing of incentives. We find that individuals characterized by a high degree of risk aversion perform worse under time pressure. Nonetheless, when facing a penalty incentive scheme these individuals are more likely to choose to work with strict term limits, suggesting that penalty contracts might generate adverse selection problems.

JEL Classification: C9, C91, D01, D91, J33

Keywords: time pressure, bonus, penalty, incentive schemes, framing, selection, lab-in-the-field experiment

Corresponding author:
Maria De Paola
Department of Economics, Statistics and Finance "Giovanni Anania"
University of Calabria
Via Ponte Bucci
87036 Arcavacata di Rende (CS)
Italy
E-mail: m.depaola@unical.it

* The authors would like to thank Vincenzo Scoppa, Marco Alberto De Benedetto and Janna Smirnova and participants to the Southern Europe Experimentalists Team (SEET) Workshop, Ibiza, 2019, for useful comments and suggestions. We are grateful to the Dean of the Faculty of Economics at the University of Calabria for allowing us to run the experiment. We would like to thank A. Appratto, P. Brogno, L. De Angelis, I. Sciommarello.

1. Introduction

Standard economic theory predicts no difference between positively framed schemes where bonuses are paid after targets are met and – economically equivalent – negatively framed contracts that change the reference point by prepaying bonuses and clawing them back if targets are unmet. However, a large number of recent papers show that people work harder under loss than gain contracts, suggesting that employers can improve effort simply by framing contracts as penalties (Hannan et al., 2005; Brooks et al., 2012; Fryer et al., 2012; Hossain and List, 2012; Armantier and Boly, 2015; Hong et al. 2015; Levitt et al. 2016; Imas et al., 2017; Pierce et al., 2020; Van der Stede, et al. 2018). A possible explanation is that the change in the reference point incentivizes loss-averse individuals to increase effort in order to avoid losses.[1]

Most of the existing studies focus exclusively on the incentive effects of framing. However, as pointed out by de Quidt (2018), it is also important to understand whether individuals are more or less likely to accept reward framed contracts than loss framed contracts. The theoretical prediction is that individuals will dislike penalty contracts and demand higher rewards to accept them. Nonetheless, the existing empirical evidence shows ambiguous results (Armantier and Boly, 2015; Imas et al. 2017; de Quidt, 2018).

When framing the contracts, the target to meet is often defined in terms of performance. However, this is not always the case. For example, for services and products where reliability is a key competitive factor, the target is often also defined in terms of timing or deadlines. The peculiarity of such contracts, aimed at incentivizing delivery on time, is that framing not only changes the reference point in terms of base pay, but also in terms of deadlines: a strict deadline (joined with a high base pay) is chosen with loss framed schemes, while a less stringent one (for a low base pay) is chosen in case of gain framed schemes. In these situations, the deadline becomes part of the target to be reached and according to the type of incentive scheme used it can be optimal to set different deadlines (El-Tannir, 2019). Individuals may be more or less likely to enter into such types of contracts depending on the bonus/penalty framing of the incentive scheme.

In this paper we contribute to the literature analyzing how framing affects selection by focusing on a contest characterized by time pressure and by investigating how individual self-select into time pressure environments according to the reward/punishment framing of incentives. Our aim is to understand whether individuals who self-select into the contract are those best suited to work under time pressure conditions. We answer our research question by running a lab-in-the-field experiment involving a sample of university students sitting a computerized test being randomly assigned to two different treatments: a *Bonus* treatment setting a less stringent deadline and allowing individuals to shift to the strict deadline and obtain a bonus, and a *Penalty* treatment defining as reference point a strict deadline that individuals can change by paying

[1] Armantier and Boly (2015) show that changing the reference point can have ambiguous effects on effort due to an inverse U-shaped relationship between incentives framing and effort provision. Moreover, Pierce et al. (2020) show that the additional effort obtained through loss-framed incentives can be directed toward actions that harm overall performance.

a penalty. Students were asked to choose whether to work under time pressure or not. Except for the different reference point that has to change by design, the incentive schemes were payoff equivalent.

We firstly tested the impact of time pressure on students' performance and find that working under strict time limits has a negative effect on performance, which is especially detrimental for risk averse students. This result is in line with findings by Dohmen et al. (2011), who, investigating the relationship between cognitive and non-cognitive abilities, show that more risk averse individuals perform worse in a test in which performance is related to speed and accuracy in applying a given correspondence rule under time pressure (see also Dohmen et al., 2018).

As regards our main research question, asking how framing incentives in terms of bonus or penalty affects selection into a time-pressured environment, we find that being assigned to the *Penalty* treatment significantly increases the probability of choosing to work under time pressure. This is consistent with results found by de Quidt (2018) who shows that the penalty framing increases job offer acceptance rate relative to the bonus framing and by Imas et al. (2017) providing evidence of a higher willingness to pay to participate in a real-effort task under the loss treatment than under the gain one. Interestingly, we also find that the acceptance rate is heterogeneous according to individual risk aversion: while the *Penalty* treatment for students in the first decile of the risk aversion distribution does not produce any effect on the probability of deciding to work under time pressure, for those in the last two deciles it generates a considerable increase. Then, the penalty incentive scheme, having strict deadlines as reference point, might generate adverse selection problems as it induces also individuals who are unsuited to work under time pressure to self-select into this environment.[2] Our results showing that framing effects interact with risk aversion are consistent with results found by a number of recent studies, highlighting that for more risk averse individuals the relationship between frame and choice is stronger (Huangfu, 2014; Biddle and Marasinghe, 2019; Tabesh et al., 2019).[3]

This evidence can be explained considering that contracts framed as penalty set the reaching of the goal as the default and give a penalty if one moves away from it (i.e. does not achieve the goal). In our experiment the goal was also set in terms of amount of time employed to perform the task. In addition, since by design the deadline is part of the reference point, in the penalty treatment time pressure is the default, inducing individuals who are more affected by framing to stick with it. It has also to be considered that the use of penalty contracts may bring individuals to perceive the base pay as uncertain (Brink and Rankin, 2013), thus more risk averse individuals may choose to work under time pressure to make the base pay certain and avoid the penalty due to using more time.

In the last part of the paper, we focus on the effect that bonus and penalty incentive schemes produce on performance. In particular, we investigate whether individuals who self-selected into the time pressure environment under the *Penalty* treatment obtain better results compared to individuals who made the same

[2] Imas et al. (2017) find that loss aversion is a relevant factor in explaining willingness to accept under the loss treatment, while de Quidt (2018) finds that acceptance behavior is not affected by loss aversion.

[3]The idea is that risk propensity moderates the relationship between frame and choice. For individuals with high risk propensity, the strong dispositional tendency for risk-taking dominates the situational uncertainties (e.g. different frames), and therefore, they are less influenced by situational mandates imposed by positive or negative frames.

choice under the *Bonus* treatment. Results show that even if individuals freely choose to work with stringent time constraints, their performance is negatively affected by time pressure. On the other hand, the *Penalty* treatment does not produce any statistically significant differentiated effect. Nonetheless, when we consider risk preferences, we find that more risk averse individuals perform worse into time pressure when they have decided to work with stringent time limits under the *Penalty* treatment (compared to the performance they get when the same choice has been made under the *Bonus* treatment).

Taken together, the results we find in the different steps of our experiment show that loss-framed incentives to use limited time can generate adverse selection problems. Indeed, the penalty incentive scheme, setting as default a strict deadline, induces individuals with a high degree of risk aversion to work under time pressure, even if they tend to perform worse under binding time limits.

We contribute to the existing literature in different ways. First, we add to the few works that investigate how framing affects selection. To the best of our knowledge, ours is the only study that investigates this issue into a time-pressured environment. While a number of papers have considered how bonus and penalty contracts affect performance under time pressure, no evidence is available on the effects they produce on selection. Hannan et al. (2005) found that even though individuals prefer bonus contracts, the effect of loss aversion leads to a higher effort choice under penalty contracts. Hossain and List (2012) observed significantly higher team productivity under payments framed in penalty terms in a high-tech manufacturing company in China. Fryer et al. (2012) and Levitt et al. (2016) examined the influence of framed contracts respectively for teachers and students showing higher effort under loss than gain contracts. Similar results are found by Imas et al. (2017) in a series of incentivized laboratory experiments. They also show that individuals are more willing to accept penalty contracts than gain contracts. A preference for the first type of contract is found also by de Quidt (2018) who runs six field experiments to study workers' preferences over bonus and penalty contracts and in four of them finds a higher take-up of penalty contracts (which disappears in the two experiments that manipulate salience). Nevertheless, having a higher number of individuals deciding to join a given type of contract does not imply that these individuals are well suited to perform under that contract. By comparing individual characteristics associated with good performance under time pressure with the characteristics of individuals choosing to work under binding deadlines we are able to offer some evidence on this issue. Finally, we innovate the existing evidence by collecting data through a lab-in-the-field experiment. That is, we combine the advantages of laboratory experiments (in terms of having a randomized setting where making *ceteris paribus* changes in the exogenous variables, enhancing control and observing variables that cannot be directly observed in the field) to the advantages of field experiments (in terms of observing subjects in a task belonging to their real life thus reducing perceived artificiality of the task and increasing the external validity of the results).

The paper is organized as follows. In Section 2 we describe the experiment. In Section 3 we present the data and conduct balance checks. In Section 4 we analyze the relationship between time pressure and performance. In section 5 we examine how framing affects selection into a time-pressured environment. In section 6 we analyze the impact of bonus and penalty incentive schemes on individual performance under time pressure. Section 7 concludes.

2. Experimental Design

We collect our data using a lab-in-the-field experiment. Like a lab experiment, our subjects perform a computerized task in a lab where each subject has his own computer and cannot look at other participants' screen nor interact with them. Instructions are read aloud at the beginning of the experiment and before each relevant part.[4] Like a field experiment, subjects are evaluated on an activity that belongs to their daily lives.

Our experiment involved 185 students enrolled in the academic year 2016-2017 at the course of Principles of Economics offered by the Degree Course in Law (5 years) at the University of Calabria.[5] The course was held during the second semester (from February to June), it was compulsory and worth 9 credits, corresponding to 63 hours of teaching and to a nominal 250 hours of study. All students (213) attended the lectures in the same room, at the same time and with the same instructor and teaching material.

At the beginning of the course, we informed students that they could choose between two different exam schemes: the "standard exam" to sit at the end of the course with questions and exercises covering the whole course program and an "alternative exam" scheme.[6]

The alternative exam scheme, related to the experiment and involving time pressure, was composed of two tests, each covering half of the program, to be taken right after the first half of the course (intermediate test) and at the end of it (final test), respectively.

The intermediate test – which we take into consideration in our experiment – was computerized using z-Tree (Fischbacher, 2007) and was structured into two parts. The first part was aimed at testing the impact of time pressure on students' performance in order to understand which are the individual characteristics associated with good performance under time pressure and at evaluating whether the time limits imposed to treated and control students were adequate to create the conditions associated to "time pressure" and "absence of time pressure". It was composed of 12 questions (correct answers, according to the degree of difficulty of the questions, were worth 1 or 2 points, while wrong and missing answers were worth zero points), which were divided into two blocks - equivalent in terms of difficulty - of 6 questions. For one block students had a time span of 6 minutes to answer, while for the other they had 12 minutes. The order of the blocks was randomized across students. Then, similarly to Kocher et al. (2019), time limits are imposed at "block level" and not on each question. The time available for completing each block of

[4] Students were not allowed to ask in private clarifying questions.

[5] The University of Calabria is a middle-sized public university located in the South of Italy. It has currently about 29,000 students enrolled in different Degree Courses and at different levels of the Italian University system. Since the 2001 reform, the Italian University system is organized into three main levels: First Level Degrees (3 years of legal duration), Second Level Degrees (2 further years) and Ph.D. Degrees. In order to gain a First Level Degree, students have to acquire a total of 180 credits. Students who have acquired a First Level Degree can undertake a Second Level Degree (acquiring 120 more credits). After having accomplished their Second Level Degree, students can apply to enroll for a Ph.D. However, in some degrees, such as Law and Medicine, the First and the Second Level Degrees are coupled together with a Degree lasting 5 and 6 years respectively.

[6] The instructions given to the students are reported in Appendix A.

questions was displayed on the computer screen of the student before the beginning of the block. The maximum grade that students could obtain in this part of the exam was 16 points.

The second part of the intermediate test was instead aimed at investigating the impact of framing on individual decisions to work under time pressure. More precisely, it was designed to analyze if framing the choice of time pressure as giving the possibility to gain bonus points instead of avoiding losing points (penalty) affects selection into time pressure. In this part, students were required to answer 12 questions. They could choose whether to answer in 24 minutes or 12 minutes: when answering in 24 minutes, correct answers were worth 1 or 2 points depending on the difficulty of the question and wrong and missing answers were worth zero points; when completing the task within 12 minutes, correct answers were worth 1.25 and 2.5 points, respectively and there was still no penalty for wrong and missing answers.

The choice was proposed under two economically equivalent and randomly assigned incentive schemes: *Bonus* treatment and *Penalty* treatment. Students in the *Bonus* treatment were described the task as requiring to answer 12 questions in 24 minutes and were given the opportunity to choose to reduce the time to 12 minutes to have a bonus of 0.25 and 0.5 for the questions worth 1 and 2, respectively, when using 24 minutes. In the *Penalty* treatment, students were described the task as requiring to answer 12 questions worth 1.25 or 2.5 points, depending on the difficulty, in 12 minutes and they were given the opportunity to decide to increase the time available to complete the test to 24 minutes at the cost of losing either 0.25 or 0.5 for the questions worth 1.25 or 2.5 points, respectively. Thus, in both treatments, thanks to the bonus obtained/penalty avoided, when working under time pressure students could gain at this part of the exam a maximum of 20 points instead of the 16 points they could get when working without time pressure.

After reading the instructions, students in the *Bonus* treatment were asked “Please, choose if you want to answer using 24 minutes or using 12 minutes and increasing by 0.5 and 0.25 the mark of the questions worth 2 and 1 points, respectively”. Students in the *Penalty* treatment were posed the same question with the following framing “ Please, choose if you want to answer using 12 minutes or using 24 minutes and decreasing by 0.5 and 0.25 the mark of the questions worth 2.5 and 1.25 points, respectively”. Below the question, they saw a button with written “12” and a button with written “24” and they had to press one of the two buttons in order to make the choice. Thus, even if by design the two treatments having a different framing involve a different default, the question used to study selection was posed in such a way to have students belonging to different treatments making the same choice between the two options (although framed differently) and not deciding whether to move away from the option set as the default. Students had a minute of time to make their decision.

As we wanted to investigate whether framing effects were short lived or persist also once individuals had the time to better realize their choices, after 8 minutes from their first decision we gave to students the opportunity to make again the same choice. Therefore, they could take the same decision or change it. More precisely, keeping the framing as described above, we asked whether they wanted to accomplish the exam in other 4 minutes or have 16 additional minutes of time. We also reminded students about the points gained (lost) in each case.

The time limits imposed at the first and second part of the test were decided based on our previous experience in order to create the conditions of time pressure and absence of time pressure, respectively. In fact, previous experience with students' behavior at a pilot test has shown that the average amount of time necessary to complete a block of 6 questions was about 9 minutes, while the average amount of time students' took to complete the second part of the test was 18 minutes. Thus, a time span of 12 and 24 minutes to answer the first and second set of questions respectively can be considered as well beyond the time even slow students typically need to accomplish the task (meaning that this does not introduce any time pressure). Instead, a time span of 6 minutes for a block of 6 questions and 12 minutes for the 12 questions of the second part does create time pressure. This judgment is supported by results presented in Section 3 showing that students' performance is negatively affected by time pressure.[7]

The grade obtained in the intermediate test was given by the sum of the grades obtained in both parts of the test. Thus, it could reach a maximum of 32 for students deciding to work without binding time limits in the second part of the test and a maximum of 36 points for students deciding to work under time pressure.

The final test – to be taken at the end of the course to complete the exam in the alternative scheme – consisted of questions and exercises covering all the remaining teaching material and allowing students to score a maximum of 32 points on the basis of their absolute performance (with a time of 90 minutes).[8] The final grade obtained by those who participated in the experiment was given by the average of the grades obtained in the intermediate and final tests.[9] Students scoring less than 18 on average did not pass the exam and had to retake the standard exam.

As required by the University, for ethical reasons, students were free to join the experiment and free to leave it at any point (after registering for the experiment or having taken the intermediate test) and sit the standard exam. In the Italian University system, usually students are allowed to retake an exam 5 times a year. However, students typically cannot retake during the year an intermediate exam and, if they fail it, they have to take the standard exam (without the possibility to split the class workload into two parts).

Students were given one week to choose whether to join the experiment or to take the standard exam. To join the experiment students had to fill out an on-line form in which they were also required to complete a short survey on their expectations on the exam, their family background and their loss aversion, risk and time preferences. The aim was that of collecting information on a number of individual characteristics that might drive both selection into time pressure and affect performance under strict time constraints. Students were assured that their answers would not be read before the tests were graded.

Once the list of participants had been obtained (211 students), they were stratified by their *High School Grade* (divided into quartiles) and gender and randomly assigned to two different treatments: *Bonus* and *Penalty*.

[7] We are aware that any exam induces a stressful environment by itself. However, this feeling of stress is present both for students performing under time pressure and for those with no strict time limits and then it does not hamper the internal validity of our experimental design.

[8] The final test, similarly to the standard exam, was conducted following the standard rules without any increase in time pressure. For this reason, we do not analyze the related outcomes.

[9] Pass grades range from 18 to "30 cum laude," which was obtained by students obtaining an average grade higher than 30.

A total of 185 students (about 88% of the 211 students who decided to join the experiment) showed up at the test, 95 students belonging to the *Bonus* treatment and 90 to the *Penalty* treatment.

3. Descriptive Statistics and Balance Checks

The design of the experiment produced three subsamples of students: those who enrolled in the course (213), those who joined the experiment (211) and those who actually showed up at the intermediate test (185).

In the first two columns of Table 1 we provide descriptive statistics separately for the first and last subsamples. We do not consider separately students who joined the experiment by registering for the intermediate exam because only two students enrolled in the course decided to not sign up for the alternative exam scheme. About 65% of students attending the course are female and this is very close to the percentage of women who actually took the test (66%). The average age of students is 20 in both subsamples.

Students enrolled in the course obtained an average *High School Grade* of 88.5 which rises to an average of 89 for students taking the test. About 66% of students enrolled in the course had studied in a *Lyceum*, while this percentage increases to 68% among those who actually took the test.[10]

We have also collected an additional measure of individual ability by including in the on-line survey a question asking the grade they expected to obtain at the exam. Students enrolled in the course report an average expected grade of 25.6, while students showing up at the exam have an average expected grade of 25.7. In addition, as the small literature investigating individual performance under time pressure has highlighted the relevance of risk aversion we have also collected data on this individual attitude. More precisely, we asked the following question: "We would like to ask you a hypothetical question that you should answer as if the situation was a real one. There is the lottery A which allows you to obtain, with the same probability, either €100 or €0 (that is, if you toss a coin, you will earn zero if it comes up heads and zero if it is tails). Please choose what you prefer in the following 9 decisions".[11] Students are then asked to choose between the lottery and the certain amount 10, 20, 30, 40, 50, 60, 70, 80 and 90. We built the variable *Risk Aversion* taking values from 0 (for participants who prefer the lottery to a certain amount of €90) to 9 (for people who prefer to receive €10 instead of playing the lottery). Some students gave inconsistent answers to the list of questions eliciting risk attitudes by switching more than once between the lottery and the certain amount. We exclude these students from our analysis. *Risk Aversion* is on average 5.6 in both subsamples (198 consistent students enrolled in the course and 172 showed up at the exam).[12]

[10] In Italy, after lower secondary school pupils choose between a 'more academically oriented track' (Lyceum), or a more labor market-oriented track (Technical or Vocational). Students coming from more educated families typically choose a Lyceum, while those from poorer socio-economic backgrounds tend to enroll at vocational schools.

[11] Given that our experiment relies on a real life situation it was difficult to use incentivized elicitations procedures. However, as shown by a number of papers, survey data are quite good predictors of actual risk taking behavior in lottery experiments (Dohmen et al., 2011; Lönnqvist et al., 2015; Chuang and Schechter, 2015).

[12] We have also asked a question to elicit students' loss aversion as in Essl and Jaussi (2017). Students had to make six choices of whether or not to play a lottery. In each lottery, the winning price was set at €60 and the losing price varied from €20 to €70. There was a 50/50 chance of receiving either €60 or the losing price. *Loss Aversion* takes

Since our alternative exam scheme allowed students to use the extra marks gained at the intermediate test (they could gain 4 extra marks as the maximum mark was set at 36 while the maximum grade in the Italian system is 30 cum laude, which is obtained by students getting 31 and 32) for the final test held at the end of the course, we have also collected information on students' time preferences. We asked the following hypothetical question: "Imagine that you have won €100.000 in a lottery. The full amount of money you have won will be available to you in one year, but you can have your winnings immediately if you give up a part. What would be the largest amount of money you would be prepared to give up in order to have your money immediately?" Respondents could select an amount of €0, €2.000, €4.000, €6.000, €8.000, €10.000, €12.000, €14.000, €16.000, €18.000, €20.000, € 22.000. We created the variable *Discount Rate* taking values from 0 to 0.22. On average the discount rate is 4.7% for students enrolled in the course and 4.6% for students taking the test, respectively.

In the second part of the intermediate test, about 79% of students decided to work under time pressure when firstly asked to choose (*TPFirstChoice*) and about 62% of them decided to work under time pressure when given the possibility to take again the same decision (*TPSecondChoice*). About 19% of the sample changed her\his mind when asked to choose for the second time and from the time pressure setting switched to the no time pressure one.

Students obtained on average a grade of 3.7 in the first part of the intermediate test when working under time pressure and 4.6 when working without time pressure (both grades range from 0 to 8). In the second part of the intermediate test, instead, students obtained a grade of about 10 (net of any bonus and ranging from 0 to 16). The points lost because students did not attempt the answer are 1.5, 0.16 and 0.97 in the first part with time pressure, without time pressure and in the second part, respectively. The points lost because students gave a wrong answer are 2.8, 3.3 and 7.5 in the first part with time pressure, without time pressure and in the second part, respectively. On average students obtained a bonus/avoided a penalty of 1.2 points (0.95 was the bonus obtained and 1.5 the penalty avoided).

values from 0 (for students who accept the lottery with a losing price of €70) to 6 (for students who refuse all lotteries) and is on average 4. However, 23 students reported inconsistent answers. Therefore, we decided to not consider this variable in our main analysis to increase the sample size and we report the results of our estimates including the indicator of *Loss Aversion* among controls in Appendix B. This measure is positively correlated to risk aversion (+0.311, *p*-value 0.000). As in de Quidt (2018), we find that loss aversion does not affect neither the choice to work under time pressure nor performance under strict deadlines. To handle the strict correlation between our measures of loss and risk aversion, we have also developed an across-individual ordinal measure of loss aversion, where loss aversion is the difference between a subject's percentile of safe choices in the risk task and their percentile of safe choices in the loss task. In addition, we have regressed loss task safe choices on risk task safe choices and used the predicted residual as a measure of loss aversion. In both cases results are similar to those found using *Loss Aversion* and reported in Appendix B.

Table 1. Descriptive Statistics. Mean and SD

	Enrolled at the course (1)	Taking the Exam (2)	Bonus (3)	Penalty (4)	Difference (*p*-value) (5)	Sample (6)	Sample: Diff (*p*-value) (7)
Penalty		0.513				0.512	
		(0.501)				(0.501)	
Female	0.653	0.665	0.674	0.656	0.018	0.680	0.003
	(0.477)	(0.473)	(0.471)	(0.478)	(0.795)	(0.468)	(0.964)
Age	20.086	20.007	20.030	19.982	0.048	20.018	0.063
	(1.195)	(1.064)	(1.146)	(0.975)	(0.760)	(1.086)	(0.703)
High School Grade	88.465	89.011	88.821	89.211	-0.390	89.122	-0.180
	(9.796)	(9.407)	(9.088)	(9.779)	(0.779)	(9.447)	(0.901)
Lyceum	0.657	0.681	0.695	0.667	0.028	0.692	0.026
	(0.476)	(0.467)	(0.463)	(0.474)	(0.684)	(0.463)	(0.714)
Expected Grade	25.592	25.724	25.505	25.956	-0.450	25.738	-0.465
	(2.926)	(2.924)	(3.420)	(2.283)	(0.296)	(2.923)	(0.299)
Risk Aversion	5.586	5.616	5.420	5.821	-0.401	5.616	-0.401
	(2.363)	(2.318)	(2.343)	(2.288)	(0.258)	(2.318)	(0.258)
Discount Rate	0.047	0.046	0.043	0.050	-0.007	0.045	-0.004
	(0.056)	(0.055)	(0.051)	(0.059)	(0.383)	(0.053)	(0.667)
TPFirstChoice		0.789	0.716	0.867	-0.151	0.779	-0.176
		(0.409)	(0.453)	(0.342)	(0.012)	(0.416)	(0.005)
TPSecondChoice		0.616	0.516	0.722	-0.206	0.610	-0.203
		(0.488)	(0.502)	(0.450)	(0.004)	(0.489)	(0.006)
TPSwitch		0.189	0.232	0.144	0.087	0.186	0.061
		(0.393)	(0.424)	(0.354)	(0.132)	(0.390)	(0.306)
Grade part I TP		3.703	3.695	3.711	-0.016	3.738	-0.046
		(1.779)	(1.631)	(1.933)	(0.950)	(1.792)	(0.867)
Grade part I not TP		4.578	4.411	4.756	-0.345	4.581	-0.399
		(1.887)	(1.854)	(1.916)	(0.215)	(1.904)	(0.170)
Grade part II		9.995	9.674	10.333	-0.660	10.048	-0.517
		(3.636)	(3.768)	(3.480)	(0.218)	(3.599)	(0.348)
Missings part I TP		1.514	1.484	1.544	-0.060	1.488	-0.046
		(2.003)	(2.128)	(1.873)	(0.839)	(1.963)	(0.878)
Missings part I notTP		0.157	0.200	0.111	0.089	0.169	0.097
		(0.693)	(0.846)	(0.484)	(0.385)	(0.718)	(0.378)
Missings part II		0.972	0.771	1.183	-0.412	1.023	-0.478
		(2.190)	(1.944)	(2.416)	(0.201)	(2.255)	(0.165)
Mistakes part I TP		2.784	2.821	2.744	0.077	2.773	0.092
		(1.961)	(1.968)	(1.963)	(0.791)	(1.956)	(0.759)
Mistakes part I notTP		3.265	3.389	3.133	0.256	3.25	0.302
		(1.891)	(1.853)	(1.933)	(0.359)	(1.908)	(0.300)
Mistakes part II		7.499	7.618	7.372	0.246	7.371	0.183
		(3.874)	(4.079)	(3.663)	(0.667)	(3.778)	(0.751)
Bonus/Penalty points		1.216	0.947	1.500	-0.553	1.217	-0.513
		(1.163)	(1.128)	(1.137)	(0.001)	(1.169)	(0.004)
Observations	213	185	95	90		172	

Notes: In columns (1) to (4) and (6) we report standard deviations in parentheses. In column (5) we report in parentheses *p*-values for the test of equality of means. In column (7) we report differences between *Bonus* and *Penalty* treatments in the sample and in parentheses *p*-values for the test of equality of means. In column (1) *Risk Aversion* refers to a sample of 198 students who made consistent choices; in column (2) to 172 students and in columns (3) and (4) to 88 and 84 students, respectively.

Treatment groups are evenly balanced in the subsample of students joining the experiment by design. In columns (3) and (4) we report descriptive statistics separately by treatment groups in the subsample of students showing up at the exam and in column (5) the differences of means and *p*-values of tests of equality of variables' means between treatments. Data regarding predetermined characteristics show that we are

unable to reject the hypothesis that the randomization was successful in creating comparable treatment groups as regards observable characteristics also in the subsample of students showing up at the exam. As regards experimental variables, it emerges that students in the *Penalty* treatment choose to work under time pressure significantly more than students in the *Bonus* treatment and that the bonus points gained are significantly smaller than the penalty points avoided.[13]

In columns (6) and (7) we report descriptive statistics and test of the differences between the *Bonus* and the *Penalty* treatment (as in columns 2 and 5) for the sample of students used in the main analysis, that is the subsample of students taking the exam that has consistent risk preferences. Also in this subsample we cannot reject the hypothesis that the randomization was successful in creating comparable treatment groups as regards observable characteristics.

4. Time pressure and students' performance

In this section, we analyze students' performance in the first part of the exam in which they were asked to answer to two sets of six questions, one under time pressure and the other with an adequate amount of time. If our design really induces time pressure, our hypothesis is to observe a better performance in the set of questions for which we allowed more time.

In order to investigate the effectiveness of our design, we reshape our dataset in such a way to obtain two observations for each student, one when working under time pressure and the other without time pressure. Thus, we create the dummy variable *TP* taking the value of 1 for the observations belonging to the time pressure environment and 0 otherwise, and a variable *Grade part I* which represents students' performance under time pressure when *TP* is equal to 1 and without time pressure when *TP* is equal to 0.

Table 2 reports individual fixed effects estimates for the impact of time pressure on students' academic performance. In all our regressions, standard errors are corrected for heteroscedasticity and clustered at the student's level. The first specification controls only for the impact of time pressure. We find that students' performance is significantly lower when they work under time pressure: on average the grade obtained is about 0.84 points lower compared to that obtained without time pressure. This reassures us that we have really induced time pressure.

In column (2) to investigate which are the individual characteristics that affect performance under time pressure, we have included among controls the interaction term *TP*Female*, the interaction terms between *TP* and our three measures of individual ability (*TP*High School Grade*, *TP*Expected Grade* and *TP*Lyceum* respectively) and the interaction terms between *TP* and our measures of risk (*TP*Risk Aversion*) and time preferences (*TP*Discount Rate*).[14] We find that students with a higher degree of risk

[13]Estimation results from a Linear Probability Model (not reported) on the full sample of students enrolled in the courses show that none of the above described predetermined characteristics significantly affects the probability of sitting the exam.

[14] Time preferences might be correlated to risk aversion and our results may be driven by the fact that more patient individuals are more risk averse (De Paola, 2013).

aversion tend to perform worse under time pressure[15]: the interaction term is negative and statistically significant at the 5 percent level. Individuals characterized by a risk aversion of 1 point higher than the mean (that is 5 in our sample) suffer a penalty of about 1.24 points.[16] However, none of the other interaction terms is statistically significant implying that the effect of time pressure is not related neither to individual ability[17] nor to time preferences (similar results are found by De Paola and Gioia, 2016).

As shown in column (3), the worsening in performance under time pressure as risk aversion increases holds true also when we exclude from our set of regressors all the interaction terms that were not statistically significant in the previous specification. Our result is in line with findings by Dohmen et al. (2011), who, investigating the relationship between cognitive and non-cognitive abilities, show that more risk averse individuals perform worse in a test under time pressure (see also Dohmen et al., 2018)

In columns (4) to (6) we investigate whether the worse individual performance under time pressure is driven by the fact that students do not attempt all the questions or by mistakes. In column (4) we consider as dependent variable the points lost due to students not attempting to answer the questions (*Missings part I*) (including among regressors our full set of controls). We find that time pressure significantly increases the number of points lost because of missing answers. However, we do not find statistically significant differences according to the degree of risk aversion: the interaction term *TP*Risk Aversion* is positive but far from being statistically significant.

In column (5) we consider as outcome variable the number of points lost due to writing incorrect answers (*Mistakes part I*). We find that the interaction term *TP*Risk Aversion* is positive implying that risk averse students tend to give more incorrect answers under time pressure. However, the effect is imprecisely estimated. In column (6) to investigate whether students instead of making mistakes tend to skip the questions when they do not know the answer or have doubts about it, we include among control variables the points lost due to skipped questions. We find that controlling for missing points, for students with a risk aversion equal to the mean the impact of time pressure on incorrect answers is positive but not statistically significant. On the other hand, for students with a degree of risk aversion higher than the mean, working under time pressure produces a statistically significant increase in the number of points lost due to mistakes. The interaction term between *TP*Risk Aversion* attracts a coefficient of about 0.19 statistically significant at the 5 percent level, implying that an increase of one point in risk aversion produces an increase in the points lost due to mistakes of about 0.08 SD.[18]

[15] We have checked whether our design selects out the risk seeking students and we have found no evidence of this. Mann-Whitney two-sample statistic shows that students enrolled at the course but not taking the exam and those taking the exam have the same risk aversion distribution.

[16] Qualitatively the same results are found when we include among controls the interaction variable *TP*Loss_Aversion*. While this interaction term is negative but not statistically significant at standard levels, the interaction term *TP*Risk_Aversion* remains negative and statistically significant at the 10 percent level. See Appendix B, Table B1 column (1).

[17] Ability matters but its effect is relevant independently of time pressure conditions (the interaction term is not statistically significant).

[18] The order of the time pressure/non-time pressure environments was varied in the first part of the intermediate test by randomly assigning some students to be exposed to time pressure in the first set of questions and other to experience time pressure after having already worked in a no time pressure environment. Our results are robust when we also control for order effects. Having experienced time pressure in the first set of questions does not affect the grade.

Table 2. Time Pressure and Students' Performance in Part I

	Grade part I (1)	*Grade part I* (2)	*Grade part I* (3)	*Missings Part I* (4)	*Mistakes Part I* (5)	*Mistakes Part I* (6)
TP	-0.8430***	-1.0364*	-0.7218***	1.0490**	-0.0126	0.7586
	(0.2414)	(0.5566)	(0.2472)	(0.4905)	(0.6341)	(0.5586)
TP*Risk Aversion		-0.2023**	-0.1966**	0.0584	0.1439	0.1868**
		(0.0945)	(0.0899)	(0.1339)	(0.1473)	(0.0946)
TP*Female		0.0905		0.0210	-0.1114	-0.0960
		(0.5145)		(0.4900)	(0.5862)	(0.4902)
TP*High S. Grade		0.0073		-0.0071	-0.0002	-0.0054
		(0.0274)		(0.0272)	(0.0331)	(0.0265)
TP*Expected grade		-0.0337		0.0693	-0.0356	0.0154
		(0.0793)		(0.0826)	(0.1142)	(0.0820)
TP*Lyceum		0.3707		0.3188	-0.6895	-0.4552
		(0.5252)		(0.5016)	(0.6443)	(0.5145)
TP*Discount Rate		-2.6795		1.4447	1.2348	2.2968
		(4.9921)		(4.1976)	(5.6479)	(4.8533)
Missing answers						-0.7351***
						(0.0898)
Constant	4.5814***	4.5814***	4.5814***	0.1686	3.2500***	3.3739***
	(0.1207)	(0.1196)	(0.1185)	(0.1174)	(0.1441)	(0.1128)
Observations	344	344	344	344	344	344
Adjusted R-sq	0.303	0.316	0.328	0.096	0.057	0.391

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

In Table 3 to check the robustness of our results to a different functional form of risk aversion, we consider among regressors the dummy variable *Averse*, which has the value of 1 for respondents with a degree of risk aversion higher than 5, and the interaction term between *Averse* and *TP*. We replicate the same specifications reported in Table 2. As shown in the first two columns, the negative impact of time pressure on performance is mainly driven by risk averse students: in both specifications the coefficient attracted by the dummy variable *TP* is negative but not statistically significant, while the interaction term *TP*Averse* is always negative and statistically significant, showing a worsening in the performance of these students of about 1 point. In column (3), where we analyze the impact of time pressure on points lost due to skipped questions, we do not find any statistically significant difference according to risk preferences. Instead, in columns (4) and (5) consistently with results reported in Table 2, we find that when working under time pressure students who are more risk averse tend to lose more points due to incorrect answers, even when we control for the number of skipped questions.

All the results described above hold qualitatively the same when instead of controlling for individual fixed effects we simply estimate an OLS model.

Table 3. Time Pressure and Students' Performance in Part I using as control the dummy *Averse*

	Grade part I (1)	*Grade part I* (2)	*Missings Part I* (3)	*Mistakes Part I* (4)	*Mistakes Part I* (5)
TP	-0.4003	-0.1277	1.2340*	-0.8337	0.0485
	(0.6695)	(0.4669)	(0.7062)	(0.8094)	(0.6567)
TP*Averse	-0.9903*	-0.9843*	-0.1871	1.1774*	1.0437*
	(0.5554)	(0.5421)	(0.6043)	(0.6903)	(0.5323)
TP*Female	0.0133		0.0339	-0.0472	-0.0230
	(0.5099)		(0.4847)	(0.5680)	(0.4811)
TP*High S. Grade	0.0052		-0.0024	-0.0028	-0.0045
	(0.0274)		(0.0269)	(0.0319)	(0.0261)
TP*Expected grade	-0.0370		0.0636	-0.0266	0.0189
	(0.0769)		(0.0827)	(0.1114)	(0.0796)
TP*Lyceum	0.3872		0.2872	-0.6745	-0.4691
	(0.5242)		(0.4984)	(0.6346)	(0.5117)
TP*Discount Rate	-2.3809		1.3207	1.0602	2.0043
	(5.1177)		(4.0980)	(5.6349)	(4.9585)
Missing answers					-0.7149***
					(0.0880)
Constant	4.5814***	4.5814***	0.1686	3.2500***	3.3705***
	(0.1198)	(0.1187)	(0.1175)	(0.1424)	(0.1123)
Observations	344	344	344	344	344
Adjusted R-squared	0.322	0.314	0.110	0.092	0.406

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

5. Bonus and Penalty framing and selection into time pressure environments

This section is devoted to analyze how the framing of the incentive to use limited time affects self-selection into time pressure. Research indicates that penalty framing, despite less prevalent than bonus framing, induces agents to exert more effort (Hannan et al., 2005; Brooks et al., 2012; Hossain and List, 2012; Imas et al., 2017). If individuals anticipate this effect, we would expect that under the penalty framing individuals are more likely to choose to work under time pressure (Imas et al. 2017). We will also investigate whether individuals endowed with particular characteristics are more sensitive to framing effects. In particular, we will analyze if individuals with different risk preferences react differently, this to inquire about adverse selection effects given the evidence discussed in the previous section showing that time pressure is especially detrimental for more risk averse students.

Table 4 reports OLS estimates for the impact of our treatment on the decision to work under time pressure. The dependent variable is the dummy variable *TPFirstChoice* taking the value of one for students who have decided to work under time pressure when firstly proposed to do so and zero otherwise. In this setting we have one observation for each student and in all our regressions standard errors are corrected for heteroscedasticity. The first specification controls only for the impact of our treatment using the dummy variable *Penalty*. Being assigned to the *Penalty* treatment significantly increases the probability of choosing to work under time pressure by about 18 percentage points. The effect we find is quite a large one: while individuals under the *Bonus* treatment have an average probability of choosing the time pressure setting of

about 70%, for individuals in the *Penalty* treatment this probability rises to about 87%. The treatment effect remains significant and similar in magnitude also when we control for individual predetermined characteristics, individual preferences and for performance in the first part of the exam (column 2). This result is consistent with results found by de Quidt (2018) and Imas et al. (2017).

In column (3), to investigate whether individuals with different risk attitudes respond differently to framing effects, we include among regressors the interaction term *Penalty*Risk Aversion* (demeaned). We find that an increase of one point in the degree of risk aversion increases the probability of choosing to work under time pressure by about 5 percentage points for students exposed to the *Penalty* treatment (the effect is statistically significant with a *p*-value=0.053). This implies that while the *Penalty* treatment for students in the first decile of the risk aversion distribution (a degree of risk aversion lower than two) does not produce any effect on the probability of deciding to work under time pressure, for those in the last two deciles (a degree of risk aversion higher than seven) it generates an increase of about 30 percentage points. These results hold true also when including the full set of interaction terms. As shown in column (4) the interaction term *Penalty*RiskAversion* is still positive but less precisely estimated (*p*-value=0.138). As regards the other interaction terms, we do not find statistically significant effects.

To better investigate the differentiated effect of the *Penalty* treatment according to individual risk attitudes, in columns (5) and (6) we run separate regressions for individuals with a degree of risk aversion below and above the mean value, respectively. We find that, while for individuals with a low level of risk aversion being assigned to the *Penalty* treatment does not affect the probability of choosing to work under time pressure, when individuals with a high degree of risk aversion are framed the possibility to work under time pressure as the chance to avoid losing points instead of gaining extra points, they are more likely to choose to work under time pressure. The effect for risk averse individuals is large (+29 percentage points) and highly statistically significant. In column (7) we test whether this difference is statistically significant running our estimates on the whole sample and including among regressors the dummy variable *Averse* and the interaction term *Penalty*Averse*. The interaction term is positive and statistically significant with *p*-value=0.018.

Similar results are obtained also when we estimate a Probit model (results not reported).

As far as control variables are concerned, it is interesting to notice that it is very difficult to identify specific characteristics that are correlated to a higher probability of deciding to perform under time pressure. Most of our observable individual characteristics do not show any statistically significant correlation.[19]

All in all, we find that framing affects selection into time pressure and that penalty framing is more likely to induce risk averse individuals to choose to work with stringent time limits. This evidence, coupled with our finding of the previous section, suggests that the penalty incentive scheme, changing the reference

[19] We only find that students with better *High School Grade* tend to prefer the environment without time pressure. More precisely, when we consider separately individuals with low and high risk aversion, we find that individuals characterized by high risk aversion who have obtained a good performance in terms of *High School Grade* have a lower probability of choosing to work under time pressure, while no effect emerges for individuals with lower levels of risk aversion.

point towards a strict deadline joined with a high base pay, negatively affects self-selection because it induces individuals to choose to work under time pressure even when they are not cut out for it.[20]

Table 4. Selection into time pressure environments. Dependent variable TPFirstChoice

	Whole	Whole	Whole	Whole	Low Risk Aversion	High Risk Aversion	Whole
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Penalty	0.1759***	0.1807***	0.1493**	0.1165	0.0799	0.2892***	-0.0638
	(0.0618)	(0.0656)	(0.0650)	(0.2541)	(0.0887)	(0.1076)	(0.2731)
Risk Aversion		-0.0075	-0.0325	-0.0302			
		(0.0135)	(0.0208)	(0.0215)			
Penalty*RiskAversion			0.0528*	0.0413			
			(0.0271)	(0.0277)			
Averse							-0.2338**
							(0.0946)
Penalty*Averse							0.3263**
							(0.1363)
Female		0.0007	0.0155	0.0303	0.2629**	-0.2458***	0.0118
		(0.0678)	(0.0683)	(0.1131)	(0.0994)	(0.0855)	(0.1116)
High School Grade		-0.0062	-0.0080**	-0.0141**	-0.0004	-0.0107*	-0.0143**
		(0.0039)	(0.0040)	(0.0059)	(0.0062)	(0.0054)	(0.0060)
Expected Grade		0.0109	0.0124	0.0175	-0.0123	0.0216*	0.0175
		(0.0124)	(0.0126)	(0.0155)	(0.0206)	(0.0126)	(0.0150)
Lyceum		0.0517	0.0641	0.1269	0.2055	0.0093	0.1202
		(0.0754)	(0.0762)	(0.1210)	(0.1284)	(0.0972)	(0.1195)
Discount Rate		-0.6399	-0.5581	-1.2276	-0.5897	0.2854	-1.1474
		(0.6296)	(0.6206)	(1.1204)	(0.8581)	(1.0376)	(1.1111)
Grade Part I TP		-0.0129	-0.0077	-0.0277	-0.0400	-0.0132	-0.0242
		(0.0204)	(0.0205)	(0.0352)	(0.0315)	(0.0307)	(0.0345)
Grade Part I NTP		-0.0095	-0.0066	-0.0052	-0.0024	-0.0032	0.0016
		(0.0184)	(0.0181)	(0.0308)	(0.0251)	(0.0269)	(0.0299)
Penalty*Female				-0.0185			0.0023
				(0.1433)			(0.1423)
Penalty*High School G				0.0120			0.0118
				(0.0082)			(0.0077)
Penalty*Exp. Grade				-0.0187			-0.0173
				(0.0298)			(0.0286)
Penalty*Lyceum				-0.1132			-0.1078
				(0.1505)			(0.1460)
Penalty*Discount Rate				0.9462			0.8758
				(1.3180)			(1.3053)
Penalty*Grade P. I TP				0.0321			0.0294
				(0.0439)			(0.0427)
Penalty*Grade P. I NTP				0.0025			-0.0065
				(0.0382)			(0.0375)
Constant	0.6932***	0.9360***	1.0327***	1.4967***	0.9594**	1.3131***	1.6657***
	(0.0494)	(0.3311)	(0.3413)	(0.4828)	(0.4598)	(0.4162)	(0.4856)
Observations	172	172	172	172	84	88	172
Ad. R-squared	0.039	0.036	0.050	0.035	0.088	0.122	0.056

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. In columns 2 to 7 we also control for dummies for town of residence. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

[20] Again our results do not change if we add among controls our measure *Loss Aversion* and the interaction term *Penalty*Loss Aversion*. Both these variables are not statistically significant. See Appendix B, Table B2 columns (1) and (2).

We suggest two possible explanations of this result. First, the use of penalty contracts of the type "an employee will receive a base salary of €15 and a penalty of €5 if the target goal (i.e. accomplishing the task within the more binding deadline) is not achieved" may make the base pay seem like it is not guaranteed (i.e. uncertain) (Brink and Rankin, 2013). Thus, more risk averse individuals may choose time pressure to make the base pay certain and avoid the penalty due to using more time. The second explanation relies on the fact that contracts framed as penalty set the reaching of the goal as the default and give a penalty if one moves away from it (i.e. does not achieve the goal). A default bias in individual behavior would explain both a higher level of effort under loss contracts (as shown by previous literature) and a higher likelihood, especially for risk averse individuals, to choose time pressure and reach the goal set as default (as shown by our results). This explanation seems less related to our framework because, as it is possible to see from the instructions reported in Appendix A, students have always to choose between the option "time pressure" and the option "no time pressure" instead the default bias is typically observed when individuals have to switch to the other option to avoid sticking with the option set as default. The results showing that more risk averse individuals are likely to be affected by framing is also consistent with the idea that risk propensity moderates the relationship between frame and choice: a higher risk propensity might dominate situational uncertainties deriving from different frames.

Since in our design after 8 minutes students are given the possibility to change their choice by choosing again whether or not to work under time pressure, in Table 5 we consider as dependent variable the choice made the second time students have to take the decision to work or not under time pressure (*TPSecondChoice*). While when posed with this question the first time about 78% of students decided to work under time pressure, now this percentage reduces to 61% implying that some of the individuals who had initially chosen the time pressure framework have decided to abandon it and perform under more relaxed time limits. 32 subjects who had previously decided to work under time pressure have changed their choice preferring to switch to the no pressure environment. Only 3 individuals have instead followed the opposite pattern switching from no time pressure to time pressure.

As shown in column (1) of Table 5, again we find a positive effect of being assigned to the *Penalty* treatment on the probability of choosing to work under time pressure. In column (2), to analyze how student's first choice affects her\his later decision, we add among control variables *TPFirstChoice* and the interaction term *TPFirstChoice*Penalty*. We find that individuals tend to keep their first decision (*TPFirstChoice* is positive and highly statistically significant). This is especially true for individuals who were exposed to the penalty incentive scheme (*TPFirstChoice*Penalty* attracts a large and positive coefficient, statistically significant at the 5 percent level – *p*-value=0.013).

In columns (3) and (4) we investigate the effect of risk aversion. In column (3), where we do not control for the choice previously made by individuals, we find that the interaction term *Penalty*Risk Aversion* is negative but not statistically significant. However, as shown in column (4), when we also include among regressors *TPFirstChoice, TPFirstChoice*Penalty, TPFirstChoice*Risk Aversion* and *TPFirstChoice*Penalty*Risk Aversion,* we find that while more risk averse individuals exposed to the

bonus incentive scheme are likely to confirm the choice to work under time pressure, those exposed to the penalty treatments behave in the opposite way. In fact, the interaction term *TPFirstChoice*Penalty*Risk Aversion* is negative and statistically significant at the 1 percent level, implying that more risk averse students who had firstly chosen to work under time pressure under the penalty incentive scheme are less likely to confirm this choice in a second moment. We report for completeness also a specification (column 5) where we also add the interaction terms between *Penalty* and our control variables as in column (4) of Table 4. Results remain qualitatively the same.

Table 5. Selection into time pressure environments. Dependent variable TPSecondChoice, TPSwitch

	TP Second choice (1)	*TP Second choice* (2)	*TP Second choice* (3)	*TP Second choice* (4)	*TP Second choice* (5)	*TP Switch* (6)
Penalty	0.2261***	-0.0901	0.2409***	-0.1233	0.0923	-0.2874
	(0.0746)	(0.0643)	(0.0759)	(0.0757)	(0.2163)	(0.2173)
TPFirstChoice		0.5567***		0.2379	0.2443	
		(0.0814)		(0.2052)	(0.2081)	
TPFirstChoice*Penalty		0.2467**		0.8810***	0.8687***	
		(0.0987)		(0.2669)	(0.3014)	
Penalty*Risk Aversion			-0.0249	0.0348	0.0326	0.0782**
			(0.0324)	(0.0302)	(0.0403)	(0.0301)
TPFirstChoice*RiskAversion				0.0603**	0.0583*	
				(0.0304)	(0.0324)	
TPFirstChoice*Penalty*RiskAv				-0.1137***	-0.1095**	
				(0.0420)	(0.0520)	
CONTROLS	YES	YES	YES	YES	YES	YES
PENALTY*CONTROLS	NO	NO	NO	NO	YES	YES
Observations	172	172	172	172	172	172
Adjusted R-squared	0.078	0.370	0.075	0.386	0.370	-0.006

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. Controls (not reported) as in Table 4 column 2. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

Results going in the same direction are found in column (6) where we investigate the behavior of students who change their mind and from the time pressure setting switch to the no time pressure one. We consider as outcome the dummy variable *TPSwitch* taking the value of one for those individuals who have firstly decided to work under time pressure and then have changed their mind shifting to the no time pressure environment. We find that the effect of having been exposed to the penalty incentive scheme on the probability of switching is heterogeneous according to risk aversion. The interaction term *Penalty*RiskAversion* is positive and highly statistically significant implying that an increase in risk aversion reduces the negative impact of *Penalty* on the probability of switching.[21]

All in all, our results show that individuals characterized by a high degree of risk aversion are initially more likely to select the time pressure environment but then they change their mind when it is given them the opportunity to do so.

[21] In Appendix B, Table B2 columns (3) and (4) we report the same estimates as in columns (4) and (6) including among controls *Loss Aversion*. The coefficient of *Loss Aversion* is not statistically significant and results on students' risk aversion are similar.

6. The effect of time pressure on performance with self-selected individuals

In this section, we analyze the effect produced by time pressure under the two alternative deadline-dependent incentive schemes on students' performance. Since students were free to choose whether to work under time pressure, we are only able to investigate whether individuals who have self-selected into the time pressure environment under the *Penalty* treatment obtain better results compared to individuals who have made the same choice under the *Bonus* treatment. Then, as sorting was allowed, the effects we find are the mixed result of incentives and selection.

We first focus on the results obtained by students during the first 8 minutes of the second test (before they were allowed to change their mind and select a time pressure environment different from the one firstly selected).[22] In table 6, column (1), we regress student's performance on *TPFirstChoice, Penalty* and the interaction term *TPFirstChoice*Penalty*. We find that students who have decided to work under time pressure under the *Bonus* scheme gain 1.15 points more compared to students working without time pressure. This initial advantage is smaller for individuals who have selected the time pressure environment under the *Penalty* scheme (0.49, *p*-value=0.232), although the difference is not statistically significant. In columns (2) and (3), we run separate regressions for individuals with low and high risk aversion, respectively. Results show that individuals with high risk aversion have an initial advantage when working under time pressure when this choice has been made under the *Bonus* scheme. On the other hand, no advantage is observed for these individuals when they have chosen the time pressure environment under the *Penalty* treatment (0.28, *p*-value=0.581). No difference between the two treatments is found for individuals characterized by low risk aversion.

In columns (4) to (6) we focus on students' final performance at the second test (the points gained during the whole duration of the test, that is 12 or 24 minutes).[23] In order to have a clean effect of time pressure we have excluded from our sample those who have changed their choice switching from an environment to another because these students have used part of their time working under a different time limit condition. In column (4) we control for the dummy variable *TPSecondChoice*, the dummy variable *Penalty* and the interaction term *TPSecondChoice*Penalty*. We find that students who have decided to perform under time pressure obtain a significantly worse performance compared to students who have decided to work without time pressure and that the *Penalty* treatment does not produce any statistically significant effect. These results, similarly to what found by Recalde et al. (2018), show that even if individuals freely choose to work with stringent time constraints, their performance is negatively affected by time pressure. In addition, as already shown by Kocher et al. (2019), individuals' worse performance under time pressure is at least partially related to their inability to efficiently allocate time across the

[22] To easily compare the magnitude of the effects with the estimates in Section 3 we have divided the number of points obtained in the first 8 minutes by two (as the maximum number of points they can get is now 16 instead of 8).
[23] Again to allow an easier comparison with estimates in Section 3 we have divided the total number of points by two.

different items. As a consequence, while initially performance is not harmed, in later questions, when less time is available, results obtained deteriorate harming the overall performance.

In columns (5) and (6), where we report results from separate regressions for individuals with low and high risk aversion, we find that the effect of time pressure is similar for both groups of students (we lose precision due to the reduction of sample size). This is probably due to the fact that the highly risk averse students who were poorly performing under time pressure had abandoned that environment.[24]

Table 6. The effect of time pressure on Grade part II (without bonus/penalty points)

	First 8 minutes			Whole part		
	Whole	Low Risk Aversion	High Risk Aversion	Whole	Low Risk Aversion	High Risk Aversion
	(1)	(2)	(3)	(4)	(5)	(6)
Penalty	0.3217	-0.3922	0.7096	0.3672	-0.9321	0.8624
	(0.4910)	(0.9185)	(0.5801)	(0.4301)	(0.7626)	(0.5581)
TPFirstChoice	1.1513***	0.7330	1.3750***			
	(0.3455)	(0.6068)	(0.3969)			
TPFirstChoice* Penalty	-0.6626	-0.0765	-1.0963*			
	(0.5345)	(0.9680)	(0.6484)			
TPSecondChoice				-0.6552*	-0.7305	-0.6708
				(0.3586)	(0.5847)	(0.5847)
TPSecondChoice* Penalty				0.3364	0.4763	-0.5998
				(0.5217)	(0.8870)	(0.6939)
CONTROLS	YES	YES	YES	YES	YES	YES
Observations	172	84	88	140	65	75
Adjusted R-squared	0.213	0.036	0.317	0.300	0.339	0.320

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. Controls (not reported) as in Table 4 column 2. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

We have also investigated the incentive effects of our treatments under time pressure by using as outcome variables the other measures of performance we have available: the points lost due to skipped questions and the points lost due to incorrect answers. Results are reported in tables C1 and C2 in Appendix C of the paper. Estimates on mistakes and missing answers show that the heterogeneous impact of the incentive during the first 8 minutes for high risk averse individuals is mainly due to points lost due to mistakes.[25]

All in all, our estimates show that, when we allow individuals to self-select into time pressure environments, framing does not produce any statistically significant effect on performance. This result does not imply that the two schemes produce the same impact as our findings are driven both by selection and incentives. It could be, for instance, that the penalty incentive scheme is more effective in inducing

[24] As regards control variables we find a positive impact of our indicator of ability, *High School Grade*, and of the performance in the first part of the exam under time pressure.

[25] We have also investigated the effect of time pressure under the two incentive schemes considering as outcome variable the number of points obtained comprehensive of bonus gained and penalty avoided (results are reported in Table C3 in Appendix C of the paper). We find similar results for the first 8 minutes while we do not find any statistically significant difference in performance between students working with and without time pressure when considering the whole duration of the test. This implies that students working under time pressure compensate their worse performance in terms of correct answers thanks to the additional points they get working under this condition. The performance of students working under the *Penalty* treatment is not different compared to those working under the *Bonus* treatment.

individuals to provide a high level of effort, but we cannot see this when looking at individual performance because of adverse selection effects (the penalty scheme persuades also individuals who are not cut out for working under time pressure to do so).

7. Concluding Remarks

In this paper we have investigated how the bonus versus penalty framing of the incentive used to induce individuals to work under time pressure affects self-selection into such environments.

Working under strict time constraints is of key importance in many economic environments, for example in all those services and manufacturing firms where delivery reliability is a key competitive factor. Performance in these situations is likely to be affected by the stress arising from the need to cope with limited time. In fact, a number of papers have shown that time pressure is detrimental for decision quality and performance (e.g., Sutter et al., 2003; Kocher and Sutter, 2006; De Paola and Gioia, 2016; El Haji et al., 2019; Buckert et al., 2017; Kirchler et al., 2017). While experimental studies that investigate this issue randomly allocate subjects into different time pressure conditions, in real life situations individuals self-select into contexts characterized by different degrees of time pressure on the basis of their ability to cope with this condition (Kocher et al. 2019). In order to induce individuals to choose to work under stressful conditions, it is necessary to reward them with schemes that can positively affect performance. Effort may be incentivized in several ways, for example using schemes that offer rewards for good performance if and after targets are met (bonus contracts) or schemes that prepay bonuses and punish bad performance if targets are unmet (penalty contracts).

While the literature investigating the incentive effects of framing is quite large, the study of how framing affects selection has been overlooked up to date. We add to the few paper analyzing this aspect by considering how framing affects selection in a specific context represented by all the situations in which individuals face binding time constraints. Firstly, we study the effect of (exogenously imposed) time pressure on performance and its determinants; then, we investigate the role of the framing of the incentive scheme in determining selection into time pressure environments.

As regards the first research question, in line with previous literature, we show that working under strict time limits has a negative effect on performance, particularly for risk averse individuals.

When looking at the impact of penalty and bonus framing on individuals' decision to work under time pressure, consistently with results found by de Quidt (2018) and Imas et al (2017), we find that self-selection is sensitive to framing: being assigned to the *Penalty* treatment significantly increases the probability of choosing to work under time pressure. Also, in addition to the existing evidence, we show that this effect is heterogeneous according to individual risk attitudes and it appears to be stronger for risk-averse individual. When we investigate whether individuals become aware of their difficulty to cope with strict time limits, we find that students exposed to the *Penalty* treatment at the top of the risk aversion distribution are more likely to shift to the no time pressure condition when they are given this opportunity. These results show that proposing a penalty incentive to enter into time pressure contexts, thus setting strict

deadlines as reference point, might generate adverse selection problems: risk averse individuals, who are less suited to work under pressure, are more likely to select the time restriction in the penalty treatment, suggesting that risk propensity moderates the relationship between frame and choice. This problem is relevant considering that this this type of contract is quite common in situations where delivery on time is important for overall performance.

Our evidence is supported by the results that we find when investigating how the reward/punishment framing of incentives affects performance under time pressure. Even if the *Penalty* treatment does not produce any statistically significant effect on overall performance regardless of risk preferences, we find that risk averse students who were assigned to the *Penalty* treatment and have decided to work under time pressure obtain a worse performance compared with the no time pressure environment and the *Bonus* treatment. This is likely to derive from adverse selection.

Our analysis improves the understanding of the interplay between framing of incentives and time pressure especially in all those situations in which individuals have the opportunity to decide whether or not to work under strict time limits. It highlights how risk averse individuals perform worse under binding time limits and nonetheless are more likely to choose time pressure when incentivized with a penalty scheme.

Additional research, aimed at providing further evidence on both incentive and selection effects of different types of contracts and on the heterogeneous response that individuals with different characteristics might show, would help to better understand why some contracts are more widely used than others.

References

Armantier, O., Boly, A. (2015). Framing of incentives and effort provision. *International Economic Review*, 56(3), pp. 917-938.

Biddle, N., Marasinghe, D. (2019). Risky Data: The Combined Effect of Framing, Trust and Risk Preferences on the Intended Participation in the Consumer Data Right. Tax and Transfer Policy Institute-Working paper, 9

Brink, A. G., Rankin, F. W. (2013). The effects of risk preference and loss aversion on individual behavior under bonus, penalty, and combined contract frames. *Behavioral Research in Accounting*, 25(2), pp. 145-170.

Brooks, R.R.W., Stremitzer, A., Tontrup, S. (2012). Framing contracts: why loss framing increases effort, *Journal of Institutional of Theoretical Economics*, 168 (1), pp. 62–82.

Buckert, M., Oechssler, J., Schwieren, C. (2017). Imitation under stress, *Journal of Economic Behavior and Organization*, 139, pp. 252–266.

Chuang Y. Schechter L. (2015), Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results, *Journal of Development Economics*, 117, pp.151-170.

de Quidt, J., Fallucchi, F., Kölle, F., Nosenzo, D., Quercia, S. (2017). Bonus versus penalty: How robust are the effects of contract framing?. *Journal of the Economic Science Association*, 3(2), pp. 174-182.

de Quidt, J. (2018). Your loss is my gain: a recruitment experiment with framed incentives. *Journal of the European Economic Association*, 16(2), pp. 522-559.

De Paola M. (2013). The Determinants of Risk Aversion: The Role of Intergenerational Transmission, *German Economic Review*, 14(2), pp. 214-234.

De Paola, M., Gioia, F. (2016). Who performs better under time pressure? Results from a field experiment. *Journal of Economic Psychology*, 53, pp. 37–53.

Dohmen, T., Falk A., Huffman D., Sunde U., Schupp J., Wagner G. G. (2011). Individual risk attitudes: measurement, determinants and behavioral consequences, *Journal of the European Economic Association*, 9, pp. 522-550.

Dohmen T., Falk A., Huffman D., Sunde U. (2018). On the Relationship Between Cognitive Ability and Risk Preference. *Journal of Economic Perspectives*, 32(2), pp. 115–134.

Essl, A., Jaussi S. (2017). Choking under time pressure: The influence of deadline-dependent bonus and penalty incentive schemes on performance, *Journal of Economic Behavior & Organization*, 133, pp. 127–137.

El Haji, A., Krawczyk, M., Sylwestrzak, M., Zawojska, E. (2019). Time pressure and risk taking in auctions: A field experiment. *Journal of behavioral and experimental economics*, 78, 68-79.

El-Tannir, A. (2019). Optimal project deadlines for mean-variance incentive contract designs. *Computers & Industrial Engineering*, 137, 106018.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), pp. 171–178.

Fryer, R., Levitt, S., List, J., Sadoff, S. (2012). Enhancing the efficacy of teacher incentives through loss aversion: a field experiment. National Bureau of Economic Research, No. 18237.

Hong, F., Hossain, T List, J. A. (2015). Framing manipulations in contests: a natural field experiment. *Journal of Economic Behavior & Organization*, 118, 372–382.

Hannan R.L., Hoffman V.B., Moser D.V. (2005). *Bonus versus Penalty: Does Contract Frame Affect Employee Effort?* In: Rapoport A., Zwick R. (eds) Experimental Business Research, pp. 151–169. Springer, Boston, MA

Hossain, T., List, J.A. (2012). The behavioralist visits the factory: increasing productivity using simple framing manipulations. *Management Science*, 58 (12), pp. 2151–2167.

Huangfu, G. (2014). Individual differences in risk-taking tendency and framing effect. *Social Behavior and Personality: an international journal*, 42(2), 279-284.Imas, A., Sadoff, S., Samek, A. (2017). Do people anticipate loss aversion? *Management Science*, 63 (5), pp. 1271–1284.

Levitt S., List J., Neckermann S. and Sadoff S. (2016), The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance, *American Economic Journal: Economic Policy*, 8(4), pp.183-219.

Lönnqvist J., Verkasalo M., Walkowitz, G., Wichardt P. (2015), "Measuring individual risk attitudes in the lab: Task or ask? An empirical comparison", *Journal of Economic Behavior & Organization*, 119, pp. 254-266.

Kirchler, M., Andersson, D., Bonn, C., Johannesson, M., Sørensen, E. Ø., Stefan, M. (2017). The effect of fast and slow decisions on risk taking, *Journal of Risk and Uncertainty*, 54, pp. 37–59.

Kocher, M. G., Sutter, M. (2006). Time is money: Time pressure, incentives, and the quality of decision-making, *Journal of Economic Behavior and Organization*, 61, pp. 375–392.

Kocher, M.G., Schindler, D., Trautmann, S.T. et al. Martin G. Xu Y. (2019). Risk, time pressure, and selection effects, *Experimental Economics*, 22(1), pp. 216-246.

Leuven E., Oosterbeek H., Sonnemans J., van der Klaauw B. (2011). Incentives versus Sorting in Tournaments: Evidence from a Field Experiment, *Journal of Labour Economics*, 29(3), pp. 637-658.

Pierce, L., Rees-Jones, A., Blank, C. (2020). The Negative Consequences of Loss-Framed Performance Incentives (No. w26619). National Bureau of Economic Research.

Sutter, M., Kocher, M. G., Strauß, S. (2003). Bargaining under time pressure in an experimental ultimatum game. *Economics Letters*, 81, pp. 341–347.

Tabesh, P., Tabesh, P., Moghaddam, K. (2019). Individual and contextual influences on framing effect: Evidence from the Middle East. *Journal of General Management*, 45(1), 30-39.

Van der Stede, W., Wu, A., Wu, S. (2018). An empirical analysis of employee responses to bonuses and penalties.https://coller.tau.ac.il/sites/collerenglish.tau.ac.il/files/media_server/Recanati/management/conferences/accounting/2018/pdf/wim-van-der-stede.pdf

Appendix A. Instructions to students

The test is divided into two parts. You will have to answer 12 questions in the first part and 12 in the second one. At the beginning of each part we will give you detailed instructions on the score of the questions and the time available.
Now the first part of the exam starts. This part is composed of 12 questions. Some questions are worth 2 points, while others 1 point. You will find the score next to each question. Wrong answers are worth zero points. The maximum score that you can obtain in this part is 16.
The questions are divided into two blocks of 6 questions. For one block you will have a time span of 6 minutes to answer, while for the other block you will have 12 minutes. The time available to complete each block of questions will be displayed on your computer screen before the beginning of each block of questions. When the time available will be over, the screen with the questions will disappear and you will see a new screen that will indicate the time available for the new block of questions.
Now you have to answer these two blocks of questions and then we will move to the second part of the exam.
After the end of the first part

Bonus treatment
Now the second part of the exam starts.
In this part, you will have to answer 12 questions in 24 minutes. Some questions are worth 2 points, while others 1 point. You will find the score next to each question. Wrong answers are worth zero points.
You are given the opportunity to decide to reduce the time available to complete the test to 12 minutes. If you decide to reduce the time available, you will gain a maximum of 4 points. This means that the questions where you read 1 point, will be worth 1.25 points and the questions where you read 2 points will then be worth 2.5 points. If you choose to answer in 12 minutes, the maximum score that you can obtain in this part will be 20 instead of 16. This means that the maximum score that you can obtain at the test is 36 (16 at the first part and 20 at the second part) if you decide to use 12 minutes and 32 (16 at the first part and 16 at the second part) if you decide to use 24 minutes.
Please make your choice by answering the question that you see on your computer screen. You have a minute of time to make a decision.
On the screen: Please, choose if you want to answer using 24 minutes or using 12 minutes and increasing by 0.5 and 0.25 the mark of the questions worth 2 and 1 points, respectively.

After 8 minutes
8 minutes have passed. We give to you the opportunity to make again the same choice. That is, regardless of the previous choice, you can choose whether to have 16 additional minutes to accomplish the exam or to use only another 4 minutes. As explained before, if you choose 4 minutes you will gain a maximum of 4 points (the maximum score at this part will be 20 instead of 16). That is you will gain 0.25 for each question worth 1 point (that will be evaluated 1.25) and 0.5 for each question worth 2 points (that will be evaluated 2.5).
Please make your choice by answering the question that you see on your computer screen. You have a minute of time to make a decision.
On the screen: Please, choose if you want to accomplish the exam using 16 minutes or using 4 minutes and increasing by 0.5 and 0.25 the mark of the questions worth 2 and 1 points, respectively.

Penalty treatment

Now the second part of the exam starts.

In this part, you will have to answer 12 questions in 12 minutes. Some questions are worth 2.5 points, while others 1.25 points. You will find the score next to each question. Wrong answers are worth zero points.

You are given the opportunity to decide to increase the time available to complete the test to 24 minutes. If you decide to increase the time available, you will lose a maximum of 4 points. This means that the questions where you read 1.25 points, will be worth 1 point and the questions where you read 2.5 points will then be worth 2 points. If you choose to answer in 24 minutes, the maximum score that you can obtain in this part will be 16 instead of 20. This means that the maximum score that you can obtain at the test is 36 (16 at the first part and 20 at the second part) if you decide to use 12 minutes and 32 (16 at the first part and 16 at the second part) if you decide to use 24 minutes.

Please make your choice by answering the question that you see on your computer screen. You have a minute of time to make a decision.

On the screen: Please, choose if you want to answer using 12 minutes or using 24 minutes and decreasing by 0.5 and 0.25 the mark of the questions worth 2.5 and 1.25 points, respectively.

After 8 minutes

8 minutes have passed. We give to you the opportunity to make again the same choice. That is, regardless of the previous choice, you can choose whether to have 16 additional minutes to accomplish the exam or to use only another 4 minutes. As explained before, if you choose 16 minutes you will lose a maximum of 4 points (the maximum score at this part will be 16 instead of 20). That is you will lose 0.25 for each question worth 1.25 point (that will be evaluated 1) and 0.5 for each question worth 2.5 points (that will be evaluated 2).

Please make your choice by answering the question that you see on your computer screen. You have a minute of time to make a decision.

On the screen: Please, choose if you want to accomplish the exam using 4 minutes or using 16 minutes and decreasing by 0.5 and 0.25 the mark of the questions worth 2.5 and 1.25 points, respectively.

Appendix B. Loss Aversion

In this Appendix we report our main estimates considering among controls also an indicator of students' loss aversion.

In order to elicit loss aversion, we have asked the same question used in Essl and Jaussi (2017). Students had to make six choices of whether or not to play a lottery. In each lottery, the winning price was fixed at €60 and the losing price varied from €20 to €70. There was a 50/50 chance of receiving either €60 or the losing price. We have created the variable *Loss Aversion* taking values from 0 (for students who accept the lottery with a losing price of €70) to 6 (for students who refuse all lotteries). 23 students reported inconsistent answers therefore the sample reduces to 162 students. *Loss Aversion* is on average 4.

In table B1 we study the relationship between time pressure and students' performance in part I considering among regressors also the variable *Loss Aversion* interacted with time pressure. We find that students' loss aversion is not a significant determinant of their performance in time pressure environments and the only variable that attracts a significant coefficient is still *Risk Aversion*. Since the two measures are correlated (+0.311, p-value 0.000) in column (2) we run the same estimate excluding *Risk Aversion* from the set of controls. The coefficient of *Loss Aversion* is still negative but not statistically significant.

Table B1. Time Pressure and Students' Performance in Part I controlling for *Loss Aversion*

	Grade part I	*Grade part I*
	(1)	(2)
TP	-1.0739*	-1.1890**
	(0.5679)	(0.5585)
TP*Risk Aversion	-0.1637*	
	(0.0949)	
TP*Female	0.1911	0.1770
	(0.5425)	(0.5405)
TP*High S. Grade	0.0048	-0.0027
	(0.0297)	(0.0298)
TP*Expected grade	-0.0351	-0.0277
	(0.0830)	(0.0835)
TP*Lyceum	0.3205	0.3630
	(0.5290)	(0.5352)
TP*Discount Rate	-2.0572	-1.6612
	(4.8890)	(4.9827)
TP*Loss Aversion	-0.1025	-0.1746
	(0.1678)	(0.1693)
Constant	4.6049***	4.6049***
	(0.1228)	(0.1240)
Observations	324	324
Adjusted R-sq	0.330	0.317

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

In table B2 we add our variable *Loss Aversion* in the specifications studying the determinants of the choice to work under time pressure. In the first two columns we look at the first time that students make this choice and find that *Loss Aversion* does not significantly affect their decision. In column (3) we look at the second time that students are asked to choose whether or not to work under time pressure and find again no effect of loss aversion. The same result holds when we consider *TPSwitch* in column (4).

Table B2. Selection into time pressure environments controlling for *Loss Aversion*

	TP First Choice (1)	TP First Choice (2)	TP Second Choice (3)	TP Switch (4)
Penalty	0.1491**	0.0887	-0.1374*	-0.3140
	(0.0708)	(0.2606)	(0.0806)	(0.2241)
RiskAversion	-0.0335	-0.0346	-0.0085	-0.0475*
	(0.0224)	(0.0244)	(0.0193)	(0.0268)
Penalty*RiskAversion	0.0572**	0.0530*	0.0372	0.0793**
	(0.0288)	(0.0302)	(0.0329)	(0.0343)
Loss Aversion	-0.0069	0.0076	-0.0132	0.0203
	(0.0227)	(0.0386)	(0.0233)	(0.0392)
Penalty*Loss Aversion		-0.0226		-0.0277
		(0.0466)		(0.0507)
TPFirstChoice			0.2544	
			(0.2158)	
TPFirstChoice*Penalty			0.8639***	
			(0.2847)	
TPFirstChoice*RiskAversion			0.0543*	
			(0.0321)	
TPFirstChoice*Penalty*Risk Aversion			-0.1092**	
			(0.0452)	
CONTROLS	YES	YES	YES	YES
PENALTY*CONTROLS	NO	YES	NO	YES
Observations	162	162	162	162
Adjusted R^2	0.045	0.027	0.382	-0.033

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively. In columns (1) and (2) we report the same estimates as in Table 4 columns (3) and (4). In columns (3) and (4) we report the same specifications as in Table 5 columns (4) and (6).

We have also developed two additional measures of loss aversion in order to handle the strict correlation emerging in our data between the measures of loss and risk aversion. The first is an across-individual ordinal measure of loss aversion, where loss aversion is the difference between a subject's percentile of safe choices in the risk task and their percentile of safe choices in the loss task. The second measure is obtained by regressing loss task safe choices on risk task safe choices and using the predicted residual as a measure of loss aversion. In both cases, results are similar to those reported in this Appendix.

Appendix C. Effect of time pressure on other measures of performance

In this appendix we investigate the incentive effects of the bonus and penalty schemes under time pressure by using as outcome variables three measures of performance: the points lost due to skipped questions, the points lost due to incorrect answers and students' payoff (that is their performance plus the bonus points gained/penalty points avoided). Results are reported in tables C1 to C3.

We find that the heterogeneous impact of the bonus/penalty incentive on the points gained during the first 8 minutes for high risk averse individuals, highlighted in Table 6, is due to points lost because of mistakes, instead the negative impact of time pressure on the final performance is due to points lost because of missing answers.

When we consider as outcome variable the number of points got comprehensive of bonus gained and penalty avoided (Table C3), we find results similar to those emerging in Table 6 for the first 8 minutes while we do not find any statistically significant difference in performance between students working with and without time pressure when considering the whole duration of the test. This implies that students working under time pressure compensate their worse performance in terms of correct answers they gave thanks to the additional points they get working under this condition. The performance of students in terms of payoff is not influenced by the treatment status.

Table C1. The effect of time pressure on points lost due to missing answers

	First 8 minutes			Whole part		
	Whole	Low Risk Aversion	High Risk Aversion	Whole	Low Risk Aversion	High Risk Aversion
	(1)	(2)	(3)	(4)	(5)	(6)
Penalty	0.1041	0.3494	-0.1058	0.2096	0.4396	-0.0771
	(0.2089)	(0.3268)	(0.2764)	(0.2190)	(0.3447)	(0.3290)
TPFirstChoice	0.1237	-0.1527	0.1411			
	(0.1326)	(0.1436)	(0.2683)			
TPFirstChoice* Penalty	0.1926	-0.0584	0.4900			
	(0.2651)	(0.3839)	(0.4270)			
TPSecondChoice				0.3741**	0.0230	0.3912
				(0.1643)	(0.1989)	(0.3039)
TPSecondChoice* Penalty				0.0811	-0.2198	0.5182
				(0.2968)	(0.4255)	(0.4919)
CONTROLS	YES	YES	YES	YES	YES	YES
Observations	172	84	88	140	65	75
Adjusted R-squared	0.060	0.161	0.013	0.063	0.135	0.040

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. Controls (not reported) as in Table 4 column 2. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

Table C2. The effect of time pressure on points lost due to incorrect answers

	First 8 minutes			Whole part		
	Whole	Low Risk Aversion	High Risk Aversion	Whole	Low Risk Aversion	High Risk Aversion
	(1)	(2)	(3)	(4)	(5)	(6)
Penalty	-0.3588	0.1837	-0.6723	-0.4578	0.7040	-0.8280
	(0.4517)	(0.8041)	(0.5445)	(0.3931)	(0.7326)	(0.5269)
TPFirstChoice	-1.1954***	-0.6419	-1.4212***			
	(0.3399)	(0.6165)	(0.3865)			
TPFirstChoice* Penalty	0.5940	0.1114	0.9358			
	(0.5033)	(0.8583)	(0.6207)			
TPSecondChoice				0.4933	0.7186	0.4962
				(0.3547)	(0.5976)	(0.4480)
TPSecondChoice* Penalty				0.3014	-0.3622	0.3686
				(0.4877)	(0.8254)	(0.6706)
CONTROLS	YES	YES	YES	YES	YES	YES
Observations	172	84	88	140	65	75
Adjusted R-squared	0.286	0.130	0.399	0.287	0.340	0.309

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. Controls (not reported) as in Table 4 column 2 including also the number of points lost due to missing answers. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.

Table C3. The effect of time pressure on student's payoff

	First 8 minutes			Whole part		
	Whole	Low Risk Aversion	High Risk Aversion	Whole	Low Risk Aversion	High Risk Aversion
	(1)	(2)	(3)	(4)	(5)	(6)
Penalty	0.3408	-0.4194	0.6902	0.4456	-0.8078	0.8791
	(0.5352)	(1.0132)	(0.6538)	(0.4772)	(0.8900)	(0.6570)
TPFirstChoice	2.0798***	1.6304**	2.3064***			
	(0.3894)	(0.6736)	(0.4613)			
TPFirstChoice* Penalty	-0.7606	-0.1553	-1.1797			
	(0.5981)	(1.0917)	(0.7498)			
TPSecondChoice				0.4681	0.5052	0.3839
				(0.4180)	(0.6954)	(0.5791)
TPSecondChoice* Penalty				-0.4069	0.2698	-0.5557
				(0.6026)	(1.0494)	(0.8357)
CONTROLS	YES	YES	YES	YES	YES	YES
Observations	172	84	88	140	65	75
Adjusted R-squared	0.282	0.116	0.375	0.230	0.265	0.239

Notes: Standard errors (corrected for heteroskedasticity) are reported in parentheses. Controls (not reported) as in Table 4 column 2. The symbols ***, **, * indicate that the coefficients are statistically significant at the 1, 5 and 10 percent level, respectively.