

IZA DP No. 8708

**Who Performs Better under Time Pressure?
Results from a Field Experiment**

Maria De Paola
Francesca Gioia

December 2014

Who Performs Better under Time Pressure? Results from a Field Experiment

Maria De Paola

*University of Calabria
and IZA*

Francesca Gioia

University of Edinburgh

Discussion Paper No. 8708
December 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Who Performs Better under Time Pressure? Results from a Field Experiment^{*}

We investigate whether and how time pressure affects performance. We conducted a field experiment in which students from an Italian University are proposed to choose between two exam schemes: a standard scheme without time pressure and an alternative scheme consisting of two written intermediate tests, one of which to be taken under time pressure. Both exam schemes consist of a verbal and a numerical part, each carrying half of the total marks. Students deciding to sustain the alternative exam are randomly assigned to a “time pressure” and a “no time pressure” group. Students performing under time pressure at the first test perform in absence of time pressure at the second test and vice versa. We find that being exposed to time pressure exerts a negative and statistically significant impact on students’ performance both at the verbal and at the numerical task. The effect is driven by a strong negative impact on females’ performance, while there is no statistically significant effect on males. Gender differences in handling time pressure are relevant only when dealing with the verbal task. Using data on students’ expectations, we also find that the effect produced by time pressure on performance was correctly perceived by students. Female students expect a lower grade when working under time pressure, while males do not.

JEL Classification: C93, D03, I23, J71

Keywords: time pressure, time constraints, gender differences, student performance

Corresponding author:

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

^{*} We are grateful to the Dean of the Faculty of Economics at the University of Calabria for allowing us to carry out the experiment, to Sara Laurita and Antonella Caparelli for providing us the data on students’ characteristics. We would like to thank Daniele Checchi, Marco Leonardi, Giuseppe Rose, Vincenzo Scoppa and seminar participants at the University of Calabria for useful comments and suggestions. Financial support from Economic and Social Research Council is gratefully acknowledged.

1. Introduction

Economic decisions are often taken under time pressure. Individuals taking part in negotiations and having to decide quickly whether to accept or refuse an offer experience time pressure; so do people buying products which are available only in limited quantity. Trading decisions in financial markets are also subject to severe time pressure. Similar pressure is faced in many occasions by managers deciding about business strategies or by people bidding in auctions. Researchers themselves face time pressure in their activity since someone else could come up with the same idea.

Performance in such jobs is likely to be affected by the stress arising from the need to cope with limited time. In our paper we look at a real life situation, that is students sitting their final exam, to investigate how and to what extent being exposed to time pressure affects individual performance and whether there is heterogeneity in the ability to handle time pressure.

The psychological literature has long investigated the effects produced by time pressure on individual decisions showing that time pressure is detrimental for decision quality (Diederich, 1997; Busemeyer and Diederich, 2002; Diederich and Busemeyer, 2003). Such a negative effect of time pressure is related to the worsening of the reasoning process and to individuals' tendency to ignore important information and rely on heuristics (Kruglanski and Freund, 1983; Gigerenzer et al., 1999; Rieskamp and Hoffrage, 2008).

In spite of the importance that time pressure has in many economic decisions, economists have typically ignored this issue. Only recently, a small number of papers has devoted attention to the impact of time pressure on individual decisions. Kocher and Sutter (2006) run a laboratory experiment using the beauty-contest game to investigate the existence of a trade-off between the quality of decision-making and time pressure and the effect exerted by time-dependent incentive schemes on such a trade-off. They show that, in absence of a time-dependent incentive scheme, the depth of reasoning decreases under time pressure even in interactive contexts. Similarly, Sutter et al. (2003) examine the effects of time pressure on bargaining behavior in an ultimatum game, showing that time pressure has high efficiency costs by leading to significantly higher rejection rates of offers, despite the effect vanishes with repetition. Kocher et al. (2013) and Bollard et al. (2007) find that time pressure changes individual attitudes toward risk. In addition, time pressure can change individual behavior by rising physiological stress, which in turn increases risk taking (Starcke et al., 2008; Putman et al., 2010; Buckert et al., 2014) and inhibits strategic thinking (Leder et al., 2013).

Even less is known on gender differences in response to time pressure. Some physiological studies (see Voyer, 2010 for a review) show that in some cognitive tests (such as mental rotation tasks) gender differences in favor of men are significantly larger when the task is administered with time constraints compared to when such constraints are absent. Shurchkovy (2012) shows that among factors that make women less effective than men in certain competitive environments (Gneezy, Niederle, and Rustichini, 2003; Niederle and Vesterlund, 2007) a crucial role is played by the ability

to handle time pressure. In a laboratory experiment, she finds that gender inequality in performance is due to men and women reacting differently to time pressure: women perform significantly better than men in competitive verbal tasks without time constraints.

In this paper, we provide new evidence both of a negative impact of time pressure on individual performance and of gender differences in handling time pressure.

Understanding whether males and females react differently to time pressure is relevant in trying to explain why women, even if as educated as men, continue to be heavily under-represented in many professions involving risky and high-pressure activities, such as executives, financial traders, entrepreneurs etc. An increasing literature documents how gender differences in preferences have a role in explaining gender differences in economic and social outcomes (Croson and Gneezy, 2009; Bertrand, 2010; Niederle and Vesterlund, 2011). Adding to this literature, in our work we focus on gender differences in the ability to face time pressure: these differences, as those in attitudes towards competition, risk aversion, time and social preferences, already widely investigated by the economic literature, might help at explaining job sorting and labor market outcomes.

However, compared to many of the existing works on gender differences in preferences (and all the few papers dealing with time pressure) that rely on laboratory experiments, we run a field experiment allowing us to observe individuals in a real life environment, in which they have strong incentives to perform well.

Our experiment has involved 220 students enrolled in an economic undergraduate course at the University of Calabria in the academic year 2013-2014. At the beginning of the course, students were given the opportunity to choose whether to sit their final exam according to the standard scheme (consisting of a single test on the whole program to be taken at the end of the course) or to an alternative scheme consisting of two intermediate tests (to sit right after the first half of the course and at the end of it, respectively) one of which to be taken under time pressure. More precisely, for each intermediate test, there was a treatment group having to take the exam under “time pressure”, that is in half an hour, and a control “no time pressure” group with a time limit of one hour. Students assigned to the “time pressure” group at the first test were assigned to the “no time pressure” group at the second test and viceversa. Time limits were decided based on our previous experience in order to create the conditions of time pressure and absence of time pressure, respectively. Answers given by students participating in the experiment to a question included in an ex-post survey, asking their opinion about the time limits faced at each test, confirmed our evaluation.

Students deciding to sit the exam under the alternative scheme could switch back to the standard exam at any point. Both the standard exam and the intermediate tests had the same structure and were composed by the same number of questions, some requiring verbal answers and some requiring numerical and graphical solutions.

Students who joined the experiment (by deciding to sit the exam according to the alternative examination scheme) were randomly assigned (on the basis of a stratified assignment procedure) to

take the first test under time pressure (110 students) or in absence of time pressure (110 students). The assigned time limit was announced to students immediately before they started to work on their test. Even if 17% of students joining the experiment did not show up at the first test, (we ended up with 92 and 90 students in the treatment and control group, respectively) we still have comparable groups since the choice to take the test was not related to the treatment status (students knew that they had to work under time pressure in one of the two intermediate tests but they did not know whether they would had to face time pressure at the first test).

38 students out of the 182 taking part in the first test did not show up at the second test. Contrary to what happened for the choice to sit the first test, the decision to take also the second intermediate test was strongly influenced by the treatment status. Female students and students residing in the same area in which the University is located were less likely to take also the second test when they had performed at the first test under time pressure (even if they would have been assigned to the “no time pressure” group in the second test). Instead, older students were more likely than younger students to sit the second test when assigned to the time pressure group during the first test. The subsample of 144 students sitting the second test is therefore no longer randomized and does not produce causal inference.

Using data on students’ performance at the first test, we find that students working under time pressure obtain an average grade of about 3 points lower than the grade achieved by students working in absence of time pressure. Moreover, when splitting the grade at the first test in its two components, we find that time pressure exerts a statistically significant negative effect both on the verbal and on the numeric part of the test.

When looking at gender differences in handling time pressure, we find that time pressure does not affect males in a statistically significant way, but it negatively affects females. A female student having to complete the exam under time pressure obtains a grade at the test which is about 5 points lower than a female student with a non-binding completion time (the effect is statistically significant at the 1 percent level). When splitting the overall performance in its two components, we find that in numerical tasks, which are typically thought as masculine, there are no significant gender differences in performance. Instead, differences between males and females are large in magnitude and statistically significant when looking at the verbal component of the test. These results confirm findings by previous works showing that women and men do not react in the same way to time pressure and that time pressure is particularly detrimental for females’ performance in verbal tasks (Shurchkovy, 2012).

Using data on students’ expectations after having completed the first test we also find that the negative effect produced by time pressure on performance was correctly perceived by students. Female students expect a grade of about 3 points lower when working under time pressure. Instead, male students do not expect lower grades when assigned to the “time pressure” group.

The paper is organized as follows. In Section 2 the design of the experiment is explained. Section 3 provides descriptive statistics and randomization checks. In Section 4 we analyze the relationship between time pressure and performance and investigate gender differences in reaction to time pressure. Section 5 looks at students' expectation about their performance. Section 6 concludes.

2. Experimental Design

To investigate the effects of time pressure on individual performance, we run a field experiment involving students enrolled in the academic year 2013-2014 at the Personnel Economics course, offered by the First Level Degree Course in Business and Administration at the University of Calabria.¹ The course was taught to students during the first half of the second semester (teaching period from March to April). All students attended the lectures in the same room, at the same time and with the same instructor and teaching material.

The instructor and the course's teaching material and schedule were the same as in previous years. The only difference concerned the final exam. In fact, our experimental protocol consists in having students perform their final exam under a scheme introducing time pressure in a randomized setting.

At the beginning of the course, we told students that they could choose between two different exam schemes: the "standard exam" and an "alternative exam" scheme.

The alternative exam consists of two intermediate tests, to be taken right after the first half of the course and at the end of it, respectively. The structure of the two intermediate tests is equivalent to that one of the standard exam that usually students take at the end of the course: students have to answer a total of 12 questions (worth a maximum of 32 points), half of them requiring verbal answers not based on computation or graphical analysis (worth a maximum of 16 points) and half requiring numerical and graphical solutions (worth a maximum of 16 points). However, while the standard exam is about the whole course program, each of the two tests composing the alternative exam scheme is about only half of the program (the first test has questions regarding only the first half of the program, while the second test has questions regarding only the second part of the program). Moreover, even if both the standard exam and the two tests of the alternative exam are evaluated with scores ranging from 0 to 32, for students undertaking the two tests the final grade at the exam is given by the average

¹ The University of Calabria is a middle-sized public university located in the South of Italy. It has currently about 35,000 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 around three main levels: First Level Degrees (3 years of legal duration), Second Level Degrees (2 years more) 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 enroll in a Ph.D. degree.

of the marks obtained at the two tests. Only students reaching a final grade of at least 18 (the minimum passing line in Italy) pass the exam.

The standard exam gives students 120 minutes to complete their examination form. Previous experience with students' behavior at the standard exam of Personnel Economics has shown that the average amount of time necessary to complete the 12 questions composing the exam is about 80-90 minutes. Moreover, previous experience with exams of courses requiring students to sit both a midterm exam (on half of the program) and a final exam (on the whole program) has shown that the majority of students complete the midterm exam in less than half of the time they need for the final exam.² Thus, a time span of 60 minutes to answer a test 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 30 minutes does create time pressure, inducing some feeling of stress, but at the same time it is sufficient to correctly answer to all the questions.³ Based on this consideration, to investigate the impact of being exposed to time pressure on students' performance, we randomly create two groups: a treatment group called "time pressure" group that is required to accomplish the test in 30 minutes and a "no time pressure" group that has 60 minutes to complete the test and serves as our control group. Students who are assigned to the treatment group when taking the first test, switch to the control group for the second test and viceversa.

As required by the university administration, students joining the experiment were free to leave it at any point (after having registered for the alternative exam, after having taken the first test or after having taken also the second test) and to sit the standard exam, scheduled for the week after the second intermediate test. Students were informed of that and were given one week of time to choose whether to sit the standard exam or the two tests. They had to manifest their decision by signing up online for the respective list. Almost all of the 223 students enrolled at the course signed up to join the alternative exam scheme (only 3 students did not sign up).

On the basis of the available administrative information on students' characteristics, we proceeded to the stratification of students deciding to sit the alternative exam according to gender, high school grade (which has been divided into quartiles), typology of high school attended and a dummy variable denoting whether the student was from the same area in which the University is located. As a result of this procedure, 32 groups were formed and within each of such groups, students were randomly assigned to the treatment and the control group. We ended up with a "time pressure" group and a "no time pressure" group of 110 students each. The two groups of students took each test on the same day and at the same time but in two different rooms. For the first intermediate test, students were informed about the room in which they had to sit it two days before the examination day

² At the Faculty of Economics at the University of Calabria, usually students have 120 minutes for final exams and 90 minutes for midterm exams. Some courses give students 150 minutes for final exams and 105 minutes for the midterm exam.

³ As we will explain later, our judgment is supported by the answers that students gave to a survey question asked at the end of each test.

but the information about the assigned time to complete the test was only announced immediately before students started to work on their test.

Even if almost all students signed to join the alternative exam, 17% of them did not show up at the first intermediate test (shifting to the standard scheme). Thus, 182 students took part to the first test: 92 in the “time pressure” group and 90 in the “no time pressure” group.

Tests were marked and students were informed about the mark obtained (professors marking the test had no information on whether students were assigned to the treatment or the control group).

A week after the end of teaching classes students undertook the second test. When deciding whether to undertake the second test and accomplish the alternative exam or to shift to the standard exam, students had more information compared to when they decided for the first test: student then knew in advance whether they would be required to work under time pressure; moreover, they had obtained a feedback both in terms of the grade scored at the first test and in terms of self-evaluation of the time needed to complete the test. 38 (21%) of the 182 students who sat the first test decided to not show up at the second test. Instead, 144 of them decided to undertake the second test (76 in the “time pressure” group and 68 in the “no time pressure” group).

At the end of each test, students were asked to complete a short survey asking questions about their expectations as regards their performance and their opinion about the time given to complete the test. We reassured students that we were not going to look at their answers before tests were graded announcing that we would have put their answers in a closed envelope signed by two of them and that the envelope would have been opened at the presence of students after the whole process of examination was over.⁴

We use the survey question on their opinion about the time given to complete the test to understand whether the time limits imposed to the treated and control students were adequate to create the conditions associated to “time pressure” and “no time pressure”, respectively.

The question asked was as follows: the amount of time you had to accomplish the test was: 1) barely sufficient; 2) sufficient; 3) plenty. As we can see from Table 1, both in the first and in the second intermediate test the majority of students perceived the amount of time as either barely sufficient or sufficient; only 15% in the first test and 8% in the second one found it plenty. When comparing treatment and control groups, we find that 89% of students assigned to the “time pressure” group at the first test answered that the time they had to accomplish the test was barely sufficient (this percentage decreases at 76% for the second test), 9% stated that it was sufficient and about 1.5% that it was plenty. On the other hand, about 3% of students assigned to the “no time pressure” group at the first test answered that the time was barely sufficient, while 69% and 28% answered that the time was sufficient or plenty. At the second test these percentages were about 84% and 14%, respectively.

⁴ Filling in the survey was not mandatory. Students could decide whether to not answer or to answer only to some questions.

These answers together with the experience we had during previous years make us confident that, even if all students were working under time constraints, only for students assigned to the time pressure group the constraint was binding and therefore time pressure was relevant in shaping individual behavior.

Table 1. Perceived time pressure

	First Test			Second Test		
	Overall	Time Pressure Group	No Time Pressure Group	Overall	Time Pressure Group	No Time Pressure Group
	Mean (Standard Deviation)					
<i>Examination Time Span:</i>						
<i>Barely sufficient</i>	0.441 (0.498)	0.892 (0.312)	0.028 (0.167)	0.4 (0.492)	0.761 (0.430)	0.016 (0.126)
<i>Sufficient</i>	0.404 (0.493)	0.092 (0.292)	0.690 (0.466)	0.523 (0.501)	0.224 (0.420)	0.841 (0.368)
<i>Plenty</i>	0.154 (0.363)	0.015 (0.124)	0.282 (0.453)	0.077 (0.268)	0.015 (0.122)	0.143 (0.353)
<i>Observations</i>	136	65	71	130	67	63

3. Experimental Data

3.1. Descriptive statistics

The design of the experiment, allowing students to change their mind at any point and sit the standard exam instead of the alternative exam, produces three subsamples of students: those who signed up to sit the alternative exam (thus joining the experiment), those who actually showed up at the first intermediate test and those who showed up also at the second intermediate test.⁵

In Table 2, we provide descriptive statistics separately for the three subsamples of students. About 56% of students joining the experiment are females. Females also represent 56% of students who decided to take the first test and are 58% of students who also took the second test. Students showing up at the first test have an average high school grade of 88.33, which is slightly higher than the average high school grade of students initially deciding to sit the alternative exam (87.42). The average high school grade for students undertaking the second test is about 89. 49% of students joining the experiment and 50% of students taking the first test have studied in a Lyceum. This percentage increases to 54% for students taking the second test. Students in all samples are on average 22 years old. About 47% of both students deciding to join the experiment and students actually

⁵ Only 3 students enrolled at the course decided to sit the standard exam instead of the alternative exam therefore it is not relevant to distinguish between those two samples.

showing up at the first test are from the same area in which the University is located, while this percentage decreases to 53% for students undertaking the second test.

Treatment and control groups are evenly balanced in the subsample of students joining the alternative exam scheme and remain almost so also in the subsample of students taking the first intermediate test. Instead, in the subsample of students taking the second intermediate test there is a slightly higher percentage of students performing under time pressure (53%).

Table 2. Descriptive Statistics

Variables	Enrolled at the Course and joining the experiment	Showing up at the first test	Showing up at the second test
	Mean (Standard Deviation)		
<i>Female</i>	0.564 (0.497)	0.560 (0.498)	0.576 (0.496)
<i>High School Grade</i>	87.418 (8.790)	88.330 (8.603)	88.833 (8.275)
<i>Lyceum</i>	0.486 (0.501)	0.500 (0.501)	0.535 (0.501)
<i>Age</i>	21.905 (0.295)	21.912 (0.284)	21.910 (0.288)
<i>Same area</i>	0.473 (0.500)	0.473 (0.501)	0.431 (0.497)
<i>Time pressure</i>	0.500 (0.501)	0.505 (0.501)	0.528 (0.501)
Observations	220	182	144

Notes: High School Grade ranges from 60 to 100.

In Table 3 we show descriptive statistics of our outcome variables, i.e. students' actual and expected performance, for the overall sample and for the treatment and the control group (separately for each test). The average grade obtained by students at the first test is 20.38. Students working under time pressure at the first test obtain an average grade of 18.76, while the average grade obtained by students working in absence of time pressure is 22. At the second test students perform relatively better: the average grade scored by students is 24.36; 23 and 25.8 for students working with and without time pressure, respectively.

When splitting the overall grade in its two components, we see that in both tests students in the time pressure group have on average the same grade in the numerical and in the verbal part of the exam, while students in the no time pressure group on average perform better in the numerical part of the exam.

We built our indicators of expected performance using students' answers to the following two questions of the short survey submitted at the end of each test: "What grade do you expect to get at this test?"; "What grade do you expect that the other students will get on average at this test?". For both questions respondents could select a grade ranging from 0 to 32.

Students in the treatment group expect to obtain a lower grade (21.191 and 23.432 for the first and second test, respectively) compared to those in the control group (23.105 and 24.754 for the first and second test, respectively). The treatment status does not affect the expectation regarding the average performance of the other students taking the test. In the first test, on average students expect to perform worse than their peers and this effect is especially high for treated students. In the second test, on average treated students expect to perform slightly worse than their peers while control students display a tiny positive relative expectation.

Table 3. Descriptive Statistics of Student Performance

	First Test			Second Test		
	Overall	Time Pressure Group	No Time Pressure Group	Overall	Time Pressure Group	No Time Pressure Group
	Mean (Standard Deviation)					
<i>Grade</i>	20.379 (6.856)	18.755 (6.677)	22.039 (6.671)	24.358 (5.928)	23.066 (6.223)	25.801 (5.258)
<i>Grade Verbal Task</i>	10.126 (4.425)	9.462 (4.465)	10.806 (4.304)	12.035 (3.953)	11.592 (3.984)	12.529 (3.888)
<i>Grade Numerical Task</i>	10.253 (4.041)	9.293 (3.939)	11.233 (3.926)	12.302 (3.293)	11.434 (3.631)	13.272 (2.566)
<i>Observations</i>	182	92	90	144	76	68
<i>Expected Grade</i>	22.021 (4.487)	21.191 (4.500)	23.105 (4.266)	24.078 (3.857)	23.432 (4.067)	24.754 (3.531)
<i>Expected Average Grade</i>	23.771 (1.597)	23.696 (1.557)	23.868 (1.655)	24.258 (1.706)	24.080 (1.914)	24.444 (1.449)
<i>Relative Expected Grade</i>	-1.750 (4.192)	-2.505 (4.462)	-0.763 (3.613)	-0.181 (3.782)	-0.648 (4.111)	0.310 (3.367)
<i>Observations</i>	143	81	62	129	66	63

Notes: *Grade at First Test* ranges from 2 to 32. However, grades for passed exams range from 18 to “30 cum laude” (set equal to 31 and 32). Both *Grade Numerical Task* and *Grade Verbal Task* range from 0 to 16. *Relative Expected Grade* is computed as the difference between *Expected Grade* and *Expected Average Grade*.

3.2. Comparable treatment and control groups

To investigate the effect that time pressure produces on individual performance we need two comparable groups: one facing time pressure and the other working in absence of time pressure. At this aim enrolled students who decided to sit the alternative exam have been randomly assigned to a treatment and a control group. However, our experimental setup allows students to leave the experiment at each stage thus introducing a self-selection element that could potentially invalidate our random design: after signing up for the alternative exam, students can decide whether to show up at the first test (and then realize which group they have been assigned to) and whether to take also the second test (for which they know in advance whether they are going to perform under time pressure or not).

In this section we analyze students' decision to participate in the two tests at the aim of checking whether the randomization is successful in both subsamples. Estimation results from a probit model are reported in Table 4. The reported coefficients represent the average marginal effect of each regressor on the outcome variable.

Columns (1) and (2) of Table 4 report the determinants of the decision to undertake the first test. As explained in the experimental design section, when deciding whether to sit the first test, students only know that they have to perform under time pressure one of the two tests of the alternative exam scheme, without knowing exactly which one. Therefore, by showing up at the first test students are somehow expressing their intention to work under time pressure: even if there is a chance that the first test will be taken in absence of time stress, in order to accomplish the exam they have to take also the second test, which, if time pressure was absent at the first test, will be performed under time pressure.

In the first specification, we only control for students' individual characteristics (gender and age), cognitive ability (measured by the typology of high school attended and the high school final grade) and province of residence. We find a positive and statistically significant effect for *High School Grade*: students with a higher level of cognitive ability, measured by a higher final grade at high school, are more likely to actually participate in the alternative exam by sitting the first test. *Female*, *Lyceum*, *Age* and the dummy variable *Same Area* do not affect the decision to sit the first test in a statistically significant way. In column (2) we add among regressors the dummy variable for the treatment status, *Time Pressure First Test*, and the interaction terms between this dummy and all other control variables. As expected, we find that the decision to take the first test was not affected by the assigned treatment group at this test and that treated and control students do not differ in terms of observable characteristics.

Table 4. Determinants of students' participation decisions

Variables	Showing up at the First Test (1)	Showing up at the First Test (2)	Showing up at the Second Test (3)	Showing up at the Second Test (4)
<i>Female</i>	-0.0616 (0.0503)	-0.0999 (0.0772)	-0.0489 (0.0418)	0.0293 (0.0602)
<i>High School Grade</i>	0.0116*** (0.0029)	0.0115*** (0.0041)	0.0002 (0.0024)	-0.0005 (0.0035)
<i>Lyceum</i>	0.0730 (0.0493)	0.0734 (0.0684)	-0.0080 (0.0443)	-0.0009 (0.0660)
<i>Age</i>	0.1080 (0.0803)	0.1621 (0.1201)	0.0046 (0.0955)	-0.1833* (0.1098)
<i>Same Area</i>	-0.0351 (0.0490)	0.0010 (0.0669)	-0.1444*** (0.0425)	-0.0613 (0.0626)
<i>Grade First Test</i>			0.0332*** (0.0021)	0.0319*** (0.0065)
<i>Time Pressure First Test</i>		2.0159 (3.7432)		-5.9155* (3.1478)
<i>Time Pressure First Test*Female</i>		0.0663 (0.1036)		-0.1395* (0.0807)
<i>Time Pressure First Test*High School Grade</i>		-0.0001 (0.0063)		0.0007 (0.0048)
<i>Time Pressure First Test*Lyceum</i>		0.0121 (0.1000)		0.0375 (0.0857)
<i>Time Pressure First Test*Age</i>		-0.0915 (0.1663)		0.2713* (0.1474)
<i>Time Pressure First Test*Same Area</i>		-0.0747 (0.0978)		-0.1558* (0.0869)
<i>Time Pressure First Test*Grade First Test</i>				0.0037 (0.0098)
<i>Log pseudolikelihood</i>	-92.777	-92.160	-45.587	-41.781
<i>Pseudo R-Squared</i>	0.0836	0.0897	0.5111	0.5519
<i>Observations</i>	220	220	182	182

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 (3) and (4) we investigate the determinants of the decision to attend the second test in the subsample of (182) students who took the first test. When deciding whether to sit the second test, students already know under which treatment condition the second test has to be performed (time pressure or no time pressure) and, more importantly, they have received feedback on their performance at the first test: they have learned from their previous experience how they have performed in the assigned time and how they could perform in the time that will be assigned to them for the second test. Then, it could be that only students who consider themselves more able to handle time pressure decide to sit the second test, while the others shift to the standard exam. In addition, these students know the score they have obtained at the first test: as to pass the exam it is necessary a minimum average score at the two tests of 18, it could be that students with low scores at the first test decide to shift to the standard exam without even trying to sit the second one.

As shown in column (3), in which we analyze the probability to take part in the second test in relation to gender, *High School Grade*, *Lyceum*, *Age*, the dummy variable *Some Area* and the grade obtained at the first test, the only statistically significant (at the 1 percent level) variables are the dummy *Some Area* and the grade achieved in the first test. An increase of one point in the grade at the

first test increases the probability of attending the second test by 3.3 percentage points, while coming from the same area reduces the probability of attending the second test by 14 percentage points (these students face lower costs and as a consequence their incentives to timely accomplish their degree program might be lower).

In column (4) we check the random assignment of this subsample by investigating whether the decision to show up at the second test is related to having performed under pressure at the first test (which implies that the student is going to perform without time pressure at the second test). At this aim we include among regressors the dummy *Time Pressure First Test* and the interaction term between *Time Pressure First Test* and the other control variables. We find that the grade obtained at the first test is still a strong determinant of the decision to sit the second test, independently from the treatment status in the first test. However, having performed at the first test under time pressure produces heterogeneous effects according to gender, age and area of residence. Females who have performed at the first test under time pressure are less likely to accomplish the alternative exam by taking part in the second test (even if they will now be assigned to the “no time pressure” group). As we will see below, this is due to the fact that females’ performance is undermined by time pressure. Once female students become aware of the negative effects produced by time pressure on their performance, they prefer to leave the experiment and shift to the standard exam (allowing them to perform without time pressure).

Similarly, students residing in the same area in which the University is located who have performed at the first test under time pressure are less likely to accomplish the alternative exam by taking part in the second test. Again, this might be due to the lower costs these students sustain to pay their university studies which reduce the incentive to timely graduation. The opposite motivation might explain why older students are more likely than younger students to sit the exam when assigned to the time pressure group in the first test.

All in all our estimates show that randomization is no longer successful for students taking part in the second test. In fact, for this subsample we find significant differences according to treatment status for student’s gender, area of residence and age. As a consequence, to analyze the effect that time pressure produces on individual performance we only focus on students’ performance at the first test. Students’ performance at the second test is analyzed in the Appendix of the paper.

4. Time Pressure and Students' Performance

4.1. Results

To investigate the effect that time pressure produces on student's performance we estimate the following linear regression model:

$$[1] \quad \text{Grade At First Test}_i = \beta_0 + \beta_1 \text{Time Pressure}_i + \beta_2 X_i + \varepsilon_i$$

where *Grade At First Test* is the dependent variable of the model, that is the grade that the student scores at the first test of the alternative exam; *Time Pressure* is the dummy for the treatment group; X denotes the vector of student's predetermined characteristics and cognitive ability; ε is an error term.

Table 5 reports OLS estimates for the impact of time pressure on students' academic performance. In the first two columns, students' academic performance is measured using the overall grade obtained at the first test. The first specification controls only for the impact of time pressure. It emerges a negative and statistically significant relationship between time pressure and the grade obtained at the test: treated students who had to complete the test in half an hour obtain an overall grade that is 3.284 points lower than the grade obtained by control students who had a completion time of one hour.

The second specification adds among controls students' personal characteristics and cognitive abilities. The impact of time pressure is still negative and statistically significant at the 1 percent level and it is almost of the same magnitude as in the specification without controls.

As far as control variables are concerned, the impact of the dummy *Female* is positive and statistically significant at the 5 percent level; similarly, students' cognitive abilities, represented by high school grade and having attended a lyceum, are positively and statistically significantly correlated with the grade obtained. *Age* and the dummy variable *Same Area* do not produce any statistically significant effect on performance.

Unfortunately, we do not have information on family background (parents' education and type of occupation) for the whole sample of students taking the first test. We have this information only for 146 students. Using this subsample we find that the negative effect of time pressure remains negative and statistically significant (at the 1 percent level) also when controlling for students' family background (results, not reported, are available upon request).

In columns (3) to (6) we estimate the same specifications of columns (1) and (2) by splitting the overall grade in its two components: verbal and numerical, respectively. When the time to perform a task is limited, individuals have to choose how to manage it in order to obtain the best possible outcome. Since the test has posed students with both numerical and verbal questions, those facing time pressure might have decided to devote more time to the task in which they are more (less) capable.

Table 5. Time Pressure and Performance: OLS estimates

Variables	Grade at first test (1)	Grade at first test (2)	Verbal Task (3)	Verbal Task (4)	Numerical Task (5)	Numerical Task (6)
<i>Time Pressure</i>	-3.2835*** (0.9895)	-3.2489*** (0.9304)	-1.3436** (0.6500)	-1.2677** (0.6137)	-1.9399*** (0.5831)	-1.9812*** (0.5724)
<i>Female</i>		1.9259** (0.9501)		1.4212** (0.6631)		0.5047 (0.6196)
<i>High School Grade</i>		0.2236*** (0.0531)		0.1279*** (0.0372)		0.0957*** (0.0354)
<i>Lyceum</i>		3.6820*** (0.9610)		2.2541*** (0.6401)		1.4279** (0.5736)
<i>Age</i>		-1.2462 (1.6692)		-0.8392 (1.2065)		-0.4070 (0.8037)
<i>Same area</i>		-1.4062 (0.9461)		0.0639 (0.6231)		-1.4701** (0.5783)
R-squared	0.0576	0.2142	0.0232	0.1674	0.0579	0.1451
Observations	182	182	182	182	182	182

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.

We find that being exposed to time pressure exerts a negative and statistically significant impact on students' performance both at the verbal and at the numerical task. As shown in column (3), students who had to complete their test under time pressure have obtained a grade at the verbal task that is 1.34 points lower compared to students who had a completion time of one hour. Instead, student's performance at the numerical task reduces by 1.9 points under time pressure (column 5). The size of the effects remains substantially unchanged when controlling for ability and individual characteristics (see columns 4 and 6).

4.2. Heterogeneous Effects according to Gender

After having assessed the existence of a negative impact of time pressure on students' performance, we inquire whether the reactions of students to time pressure differ according to gender and whether this heterogeneity depends on the typology of task considered. At this aim, we estimate the same specifications reported in Table 5 and include among regressors the interaction term between the treatment status *Time Pressure* and the dummy *Female*. Results are reported in Table 6.

In column (1) we report OLS estimates obtained using as outcome variable the overall grade gained by students and including among regressors only the dummies *Time Pressure*, *Female* and the interaction term *Time Pressure*Female*. We find that the negative effect of time pressure on student's performance is mainly due to females' reactions. The effect of time pressure on male performance is negative but not statistically significant; instead we find a negative and highly statistically significant impact on female performance. A female student having to complete the test under time pressure obtains an overall grade which is about 5 points lower than a female student with a non-binding completion time (the effect is statistically significant at the 1 percent level). The difference in the effect of time pressure due to gender is of about 4.2 points, statistically significant at the 5 percent

level. In column (2), we add among controls students' characteristics and cognitive ability. In this specification as well, the impact of working under time pressure is heterogeneous according to gender with effects of about the same magnitude of those found in column (1). As far as control variables are concerned, we find again that higher ability students (i.e. those with a higher high school grade and coming from a lyceum) perform better at the first intermediate test while *Age* and *Same Area* do not exert any significant effect on students' performance.

Table 6. Heterogeneous Effect of Time Pressure according to Gender: OLS estimates

Variables	Grade at first test (1)	Grade at first test (2)	Verbal Task (3)	Verbal Task (4)	Numerical Task (5)	Numerical Task (6)
<i>Time Pressure</i>	-0.8559 (1.5953)	-1.0920 (1.4607)	0.2544 (1.0886)	0.1648 (0.9970)	-1.1103 (0.9449)	-1.2567 (0.9325)
<i>Female</i>	4.4170*** (1.4143)	3.9550*** (1.3426)	2.9884*** (0.9202)	2.7688*** (0.9482)	1.4286* (0.8550)	1.1862 (0.8735)
<i>Time Pressure*Female</i>	-4.1880** (1.9971)	-3.8691** (1.8631)	-2.7517** (1.3225)	-2.5697** (1.2601)	-1.4363 (1.1944)	-1.2995 (1.1612)
<i>High School Grade</i>		0.2166*** (0.0524)		0.1232*** (0.0367)		0.0934*** (0.0357)
<i>Lyceum</i>		3.6151*** (0.9444)		2.2097*** (0.6356)		1.4054** (0.5668)
<i>Age</i>		-1.5664 (1.6222)		-1.0518 (1.1649)		-0.5146 (0.8073)
<i>Same area</i>		-1.4067 (0.9374)		0.0636 (0.6171)		-1.4703** (0.5778)
R-squared	0.1081	0.2337	0.0788	0.1880	0.0731	0.1514
Observations	182	182	182	182	182	182
Time pressure Effect on females	-5.0438***	-4.9611***	-2.4973***	-2.4049***	-2.5465***	-2.5562***
p_value	0.0000	0.0000	0.0009	0.0015	0.0005	0.0003

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 the remaining columns we investigate whether such heterogeneity depends on the typology of task students had to complete. We replicate the specifications reported in columns (1) and (2) considering as dependent variable the grade obtained in the Verbal Task and in the Numerical Task, respectively.

We find (columns 3 and 4) that while males' performance in the verbal test is not harmed by time pressure, females' performance in this task reduces by about 2.5 points under time pressure (the effect is significant at the 1 percent level). The gender difference in the impact of time pressure is statistically significant at the 5 percent level. Instead, when looking at the numerical task, which is typically considered as a masculine one (see Gneezy, Niederle, and Rustichini, 2003; Niederle and Vesterlund, 2007; Shurchkovy, 2012), we find that males and females react similarly to time pressure. As shown in columns (5) and (6), time pressure produces a negative effect on the performance in the numerical task for both males and females, but the effect is not statistically significant at conventional levels neither for males nor for females.

All in all our results suggest that the feeling of stress induced by the need to cope with the limited time produces a strong negative effect on the performance of female students while leaving

almost unchanged males' performance. Furthermore, if one considers the verbal and the numerical tasks as a typically feminine and masculine task, respectively (as is the case in the literature), our estimates suggest that when feeling the pressure of limited time, students dedicate more time to the task in which they are "less capable". In fact, when splitting the overall performance in its two components, it emerges that in numerical tasks there are no significant gender differences in performance; instead, differences between males and females are large in magnitude and statistically significant when looking at the verbal component of the test.

5. Time Pressure and Students' Expectations about Performance

In this section we investigate whether students, once performed the test, have perceived the effect that time pressure has produced on their performance. At this aim we use students' answers to questions concerning their expectations, included in the survey that we proposed to them at the end of each test, to build our two outcome variables: student's *Expected Grade* and student's *Relative Expected Grade* (the difference between this expectation and the average grade the student expects to be obtained by his/her peers).⁶

We estimate an OLS model on the subsample of 143 students answering the two questions about performance related expectations. In all specifications we include the full set of our controls. Results are reported in Table 7.

The first two specifications have as outcome variable student's *Expected Grade*. As shown by estimates reported in column (1), on average students working under time pressure expect to obtain a grade of about 1.7 points lower than the grade expected by students working in absence of time pressure. In column (2), to investigate gender differences in grade expectations in relation to time pressure, we include among controls the interaction term *Time Pressure*Female*. We find that male students working under time pressure do not expect to obtain a lower grade. Instead, working under time pressure produces a negative and statistically significant effect on the grade expectations of female students: on average, female students facing time pressure at the first test expect to obtain a grade which is about 3 points lower compared to that expected by a female student working in absence of time pressure (the effect is significant at the 1 percent level and the gender difference at the 5 percent level).

⁶ To assess whether students have provided honest answers when predicting their grades, we have regressed the Expected Grade on the effective Grade they obtained. We find a positive high coefficient (0.93) on Grade, with a t-stat of 9.2 and a R-squared equal to 0.39.

Table 7. Grade expectations and Time Pressure: OLS estimates

Variables	Expected Grade (1)	Expected Grade (2)	Relative Expected Grade (3)	Relative Expected Grade (4)
<i>Time Pressure</i>	-1.6897** (0.7241)	0.1532 (1.0977)	-1.5345** (0.6567)	-0.0365 (0.9375)
<i>Female</i>	-0.0663 (0.7388)	1.7586 (1.1032)	-0.8503 (0.6754)	0.6331 (0.9569)
<i>Time Pressure*Female</i>		-3.1684** (1.4416)		-2.5756* (1.3151)
<i>High School Grade</i>	0.0822* (0.0417)	0.0765* (0.0410)	0.0688* (0.0374)	0.0642* (0.0371)
<i>Lyceum</i>	2.2690*** (0.7354)	2.2615*** (0.7242)	2.3101*** (0.6698)	2.3040*** (0.6638)
<i>Age</i>	0.2901 (1.0437)	0.0526 (1.0901)	-0.2621 (0.9762)	-0.4552 (0.9659)
<i>Same area</i>	-0.1692 (0.7102)	-0.3184 (0.7095)	-0.2450 (0.6663)	-0.3662 (0.6725)
R-squared	0.1183	0.1477	0.1394	0.1616
Observations	143	143	143	143
Time pressure Effect on Females		-3.015***		-2.612***
p_value		(0.002)		(0.004)

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.

Very similar results emerge when looking at the relative expected grade. Students performing under time pressure expect to obtain a grade about 1.5 points lower than the average grade expected for their peers (column 3). This negative relative expectation is particularly marked for treated female students: on average, a female student who performed the test in the time pressure group expects to obtain a relative grade which is about 2.6 points lower compared to that expected by a female student who performed the test in the no time pressure group.

Our evidence suggests that females realize that time pressure harmed them immediately after having performed at the first test and this reflects in their expectations, both in absolute and in relative terms. As shown in the previous section, female students' expectations of a low grade obtained at the test when performing under time pressure were confirmed by the grades students effectively obtained at the test.

6. Concluding Remarks

In today's society, characterized by strong competition, rapid changes in technology and innovation, the faster things can be done, the better it is. Therefore, many decisions and activities are made under time constraints. This may generate feelings of stress, which in turn can affect the performance in the task.

This paper adds to a small economic literature investigating the effects of time pressure on individual performance. Using data collected from a randomized field experiment in which we compare a group of undergraduate economics students taking the final test under time pressure to a control group having a not binding test completion time, we show that having to perform the test under

time pressure reduces the grade that students achieve, with an effect that is similar for the verbal and the numerical component of the test. However, females handle time pressure worse compared to males. Indeed, whilst time pressure does not affect males in a statistically significant way, females suffer from a strong worsening of their performance when exposed to time pressure. Gender differences in performance are large in magnitude and statistically significant especially in the verbal component of the test. We also find that students correctly perceive the effect produced by time pressure on their performance: female students working under time pressure expect a lower grade, while male students do not expect lower grades when assigned to the “time pressure” group.

These results confirm findings by previous works showing that women and men do not react in the same way to time pressure and that time pressure is particularly detrimental for females’ performance in verbal tasks (Shurchkovy, 2012). This evidence suggests females’ difficulty to manage time pressure as a possible cause of the lower share of women in high-qualified jobs requiring exposure to binding time constraints. However, the few works on this topic do not allow to reach conclusive results and further research is required to understand how men and women react to time pressure and how this reaction is related to other environment’s characteristics. It would be also interesting to study whether the repeated exposition to time pressure helps at reducing its negative effect on performance and to investigate whether there exists an “optimal” level of time pressure that can actually improve individuals’ performance.

References

- Bollard, A., Liu, R., Nursimulu, A.D., Rangel, A., Bossaerts, P. 2007. Neurophysiological evidence on perception of reward and risk: Implications for trading under time pressure. Working paper, CalTech, Pasadena, CA
- Buckert, M., Schwieren, C., Kudielka, B., and Fiebach, C. J. 2014. Acute stress affects risk taking but not ambiguity aversion. *Front. Neurosci.* 8:82
- Gneezy, U., Niederle, M., Rustichini, A., 2003. Performance in competitive environments: gender differences. *Quarterly Journal of Economics.* 118, 1049–1074
- Kocher, M. G., & Sutter, M. 2006. Time is money–Time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization*, 61(3), 375-392
- Kocher, M. G., Pahlke, J. & Trautmann, S. T. 2013. Tempus Fugit: Time Pressure in Risky Decisions, *Management Science*, 59(10), 2380-2391
- Leder, J., Häusser, J. A., & Mojzisch, A. 2013. Stress and strategic decision-making in the beauty contest game. *Psychoneuroendocrinology*, 38(9), 1503-1511
- Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much?. *Quarterly Journal of Economics*, 122 (3), 1067–1101

Putman, P., Antypa, N., Crysovergi, P., and van der Does, W. A. 2010. Exogenous cortisol acutely influences motivated decision making in healthy young men. *Psychopharmacology* 208, 257–263

Shurchkov, O. 2012, Under Pressure: Gender Differences In Output Quality And Quantity Under Competition And Time Constraints. *Journal of the European Economic Association*, 10: 1189–1213

Starcke, K., Wolf, O. T., Markowitsch, H. J., and Brand, M. 2008. Anticipatory stress influences decision making under explicit risk conditions. *Behav. Neurosci.* 122, 1352–1360

Sutter, M., Kocher, M. & Strauß, S. 2003. Bargaining under time pressure in an experimental ultimatum game. *Economics Letters*, 81(3), 341-347

Voyer, D. 2011. Time limits and gender differences on paper-and-pencil tests of mental rotation: a meta-analysis. *Psychonomic bulletin & review*, 18(2), 267-277

APPENDIX. Time pressure and students' performance: second test

In this appendix we investigate the impact of time pressure on students' performance in the second test. Table A1 shows OLS estimation results for the specification including all our control variables plus the grade achieved in the first test.

In the first three columns we investigate whether having to complete the second test under time pressure has affected students' performance at the test, both overall and in the two different tasks and whether the effect is heterogeneous according to gender. Our sample reduces to 144 students who attended both the first and the second test. We find that being exposed to time pressure reduces students' grade at the second test both overall and in the two components of the test. However, this effect is common to both male and female students.

In columns (4) and (5) we investigate the effect of time pressure on students' absolute and relative expectations using the subsample of 129 students who took both tests for which we have such information. We find that students working in the time pressure group expect a lower absolute and relative grade compared to control students' expectations. However, as for actual performance, expectations also do not differ between treated males and females.

However, as we show in the paper, students deciding to sit also the second intermediate test are no longer randomized over the assigned group for the test. Moreover, they received a feedback after the first test, both in terms of the grade scored and in terms of self-evaluation of the time needed with respect to the complexity of the exam.

Table A1. Effect of Time Pressure: Second Intermediate Test. OLS Estimates

Variables	Grade at second test (1)	Verbal Task (2)	Numerical Task (3)	Expected Grade (4)	Relative Expected Grade (5)
<i>Time Pressure</i>	-5.3435*** (1.1679)	-2.1395** (0.8900)	-3.2755*** (0.8319)	-2.4508** (1.1075)	-1.8077* (1.0684)
<i>Female</i>	1.5676* (0.9019)	1.2621* (0.6876)	0.3054 (0.6246)	0.1380 (0.9001)	-0.3498 (0.8455)
<i>Time Pressure*Female</i>	0.8640 (1.4895)	-0.1959 (1.1025)	1.1109 (1.0143)	0.8818 (1.3128)	0.3993 (1.2701)
<i>High School Grade</i>	0.0805* (0.0480)	0.0615* (0.0371)	0.0205 (0.0277)	0.0408 (0.0396)	0.0108 (0.0429)
<i>Lyceum</i>	-1.0848 (0.7202)	-0.9628* (0.5351)	-0.0705 (0.4817)	1.1509* (0.6612)	0.9504 (0.6642)
<i>Grade First Test</i>	0.6526*** (0.0712)	0.3777*** (0.0531)	0.2760*** (0.0439)	0.2307*** (0.0661)	0.2681*** (0.0657)
<i>Age</i>	-0.4355 (1.2542)	-0.0245 (0.9869)	-0.4277 (0.8639)	0.5800 (1.5443)	0.6527 (1.4235)
<i>Same area</i>	1.7100** (0.6775)	1.6346*** (0.4989)	0.0494 (0.4860)	-0.1227 (0.5932)	-0.1210 (0.6175)
R-squared	0.5327	0.4171	0.3435	0.2140	0.2019
Observations	144	144	144	129	129
Time pressure Effect on females	-4.4794	-2.3354	-2.1646	-1.5690	-1.4084
p_value	0.0000	0.0003	0.0002	0.048	0.064

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.