Lack of Control: An experiment

Prissé, Benjamin and Jorrat, Diego

Universidad Loyola Andalucía

25 September 2021

Lack of Control: An experiment *

Benjamin Prissé[†], Diego Jorrat[‡]

September 26, 2021

Abstract

We ran an experiment to study whether lack of control has an effect on experimental results. Subjects who were recruited following standard procedures completed the experiment online or in the laboratory. The experimental design is otherwise identical between conditions. Results suggest that there are no differences between conditions, except for a larger percentage of laboratory subjects donating nothing in the Dictator Game.

Keywords: Time Preferences, CTB, Experiments.

JEL-codes: C91, C93, D15.

---

*Thanks to Pablo Brañas-Garza and Praveen Kujal for sharing the data. Karen Khachatryan conducted the experiments and Alex Bueno programmed the software. Financial support from Middlesex Business School and Ministerio de Ciencia y Tecnología (PGC2018-093506-B-100) and Excelencia Junta de Andalucía (PY18-FR-0007) is gratefully acknowledged. Authors declare they have no competing interests.

[†] *LoyolaBehLab*, Universidad Loyola Andalucía. E-mail: benjamin.prisse@gmail.com. Address: Calle Escritor Castilla Aguayo, 4, 14004 Córdoba, Spain.

[‡] *LoyolaBehLab*, Universidad Loyola Andalucía, Spain.

1 INTRODUCTION

An increasing number of papers run experiments online, using platforms like Amazon Mechanical Turk (AMT). It gives access to a large and diverse panel of subjects, rapidly and at cheap cost. It also simplifies the logistical burden of organizing experiments.
Experimental economists are concerned about the validity of these experiments because they differ from standard experiments in three dimensions: stakes of payoffs, heterogeneous samples and lack of control. Resolving these three issues would validate the use of online experiments.
The first issue refers to whether the smaller payments of online experiments influence results, but Mason and Watts (2009) and Marge et al (2010) showed that AMT wages influenced the quantity but not the quality of work, in accordance with the meta-analysis of Camerer and Hogarth (1999).
The second issue refers to subject characteristics affecting results in economic paradigms, but online subjects display the same cognitive biases and logical fallacies (Goodman et al, 2013, Horton et al, 2011 and Paolacci et al, 2010) or the same behavior in economic games (Amir et al, 2012, Arechar et al, 2018, Brañas-Garza et al, 2018 and Jorrat, 2021) than traditional participants.
This paper focuses on the remaining lack of control issue: we do not know what subjects are doing when answering the task, meaning whether they remain concentrated on the task or are helped by a partner when making choices. We ran an experiment to answer this question by only varying this aspect of the experimental design, similarly to Hergueux and Jacquemet (2015) or Arechar et al (2018).
We made subjects perform several standard economic tasks related to the measurement of economic preferences, creating a complex environment likely to elicit any potential differences. Results are identical between conditions, except for laboratory subjects donating to charities more frequently in the Dictator Game. We conclude that online environments are as valid as laboratory environments.

2 EXPERIMENTAL DESIGN

The key feature of our design is the use of two experimental treatments only differing on the location where subjects complete the experiment:

- Laboratory: subjects are placed at the standard laboratory environment.
- Online: subjects are placed at home on their personal computer.

Comparing the two conditions allow us to evaluate the potential effect of not controlling subjects.

2.1 Experimental Tasks

Because the potential differences are numerous, we created a rich environment likely to elicit any potential differences by making subjects successively reply to several standard economic tasks:

- The Convex Time Budget (CTB) of Andreoni and Sprenger (2012).
- The Multiple Price Lists (MPL) of Andreoni et al (2015) in a modified version.
- The risk-aversion task (HL) of Holt and Laury (2002).
- The Dictator Game (DG) with subjects donating their total earnings.
- The Cognitive Reflection Task (CRT) of Frederick (2005)
- A Numeracy (Num) task related to percentages.

We recruited Middlesex University students in two experimental sessions: one in November 2014 and the other in March 2015. We sent a total of 22544 invitations to the six schools[1] of the university and registered 520 students in total. These subjects were randomly assigned to Lab/Online conditions but only 257 subjects completed the experiment (113 Lab, 144 Online). The attrition rate was 50.58% (52.92% Lab, 48.57% Online) with a t-test not rejecting the equality between conditions (p=0.217). This sample size allows us to find an effect size of 0.3 SD or higher with a 90% significance level. Any effect below this threshold could be considered a low effect according to Cohen´s effects size.

2.2 Questionnaire

Subjects also answered to a questionnaire measuring control variables. First, we verify if subjects are alone when answering the Online experiment. Regarding the location of reply, 75% reply at home, 18.75% at university, 5% somewhere else and 1.25% indicated nothing. It suggests that a significant proportion of subjects could be in presence of others, but when asked about their social environment 98.60% reported being "Alone" or "Mostly Alone", suggesting that Online subjects are as isolated as Lab subjects.

Second, it allows us to verify that samples are comparable. We run difference mean test between groups to check the balance in different controls variables. Table 1 provides results of these tests, with Romano-Wolf adjusted $p-value$ for multiple testing in the last column. It suggests that the two groups are similar in their characteristics (p ≥0.238), allowing us to estimate the causal effect of online setting on results. We conclude that the two groups are comparable and that the experiment measures the causal effect of experimental conditions.

[1] Art and Design, Business, Health and Education, Law, Media and Performing Art, Science and Technology

Table 1: Balance check

	n	$mean_L$	$mean_O$	$L-O$	$p-value$	$adj.p-value$*
Exercise	255	0.57	0.42	-0.15	0.016	0.238
Medical check up	254	0.71	0.62	-0.10	0.109	0.711
Smoke	256	0.18	0.21	0.03	0.513	0.987
Weight	213	77.27	73.85	-3.42	0.444	0.983
Height	206	166.57	163.20	-3.37	0.215	0.855
Work	256	0.44	0.45	0.01	0.936	0.987
English native	255	0.44	0.58	0.14	0.032	0.374
Age at admission	251	21.42	21.15	-0.27	0.650	0.987
Female	251	0.56	0.58	0.02	0.752	0.987
Numeracy score	257	4.40	4.60	0.21	0.180	0.840
CRT score	224	0.57	0.81	0.24	0.060	0.535
Trust - experimenter	252	0.87	0.89	0.02	0.581	0.987
Trust - donation	243	5.14	4.90	-0.24	0.318	0.937
All nighters	255	9.73	19.71	9.98	0.028	0.355
Want credit card	252	17.43	9.79	7.64	0.075	0.600

* Refers to Romano-Wolf step-down adjusted p-values.

3 RESULTS

The following section displays results for time, consistency and the answers subjects gave in each task. We find almost no differences between Lab and Online conditions.

3.1 Convex Time Budget

Time

Columns 1 and 2 in Table 2 provide the regressions results on *time response*[2]. Column 1 shows that subjects take on average 18.65 and 17.97 minutes in the Lab and Online setting respectively. The *online* variable is not significant ($p = 0.578$). Column 2 shows the same result when controlling for age, gender, numeracy and CRT score ($p = 0.217$). We also perform a Kolmogorov-Smirnov (KS) test confirming that the distribution of time response is similar in both settings ($p = 0.426$). Further analyses are presented in Section B.1.1 of the Appendix. They suggest the same no result when separating subjects by cognitive levels or whether they are below or above the median time.

Result 1a: Online environment has no impact on CTB response time.

[2]We exclude one outlier (68.70 min) in the Online condition from the time analysis.

Table 2: OLS estimations of the impact of online setting on time response, consistency and number of future allocations in the CTB task.

	(1) Time (in minutes)	(2) Time (in minutes)	(3) Consistency	(4) Consistency	(5) # future allocations	(6) # future allocations
online	-0.676	-1.568	-0.025	-0.021	1.417	1.589
	(1.213)	(1.267)	(0.062)	(0.067)	(1.906)	(2.152)
	[0.578]	[0.217]	[0.684]	[0.756]	[0.457]	[0.460]
Constant	18.651***	23.570***	0.407***	-0.061	2.124***	1.674*
	(0.861)	(4.425)	(0.046)	(0.218)	(0.166)	(0.931)
	[0.000]	[0.000]	[0.000]	[0.781]	[0.000]	[0.072]
Observations	256	218	257	219	257	219
R-squared	0.001	0.074	0.001	0.049	0.002	0.005
Controls	No	Yes	No	Yes	No	Yes

Note: Robust standard errors in parentheses and p values in brackets. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Consistency

Defining consistency in CTB is complex because the 45 budget choices create 920 opportunities to make an inconsistent choice, thus do not allow to use the usual complete consistency. Because a single trial can rapidly increase the number of inconsistent choices, we give some margin of errors to subjects by considering them consistent if at least 80% potential inconsistencies are avoided. Columns 3 and 4 of Table 2 show regression results for consistency. Column 3 shows that we have 40.70% consistent subjects in the Lab and 38.19% Online with the *online* variable again not significant ($p = 0.684$), and Column 4 shows that this result holds when adding controls ($p = 0.756$). Further analyses are presented on Section B.1.2 and results suggest that there are no heterogeneous effects across different cognitive levels, between early (first half of trials) and late (second half) parts of the CTB, or trials with similar monetary amounts.

Result 1b: Online environment has no impact on CTB consistency and type of inconsistencies.

Allocations of Tokens

Another potential difference between conditions is how subjects allocate tokens in CTB. We start by looking at the number of future allocations[3]. Column 5 in Table 2 suggests that future allocations are used on average 8.36 times in the Lab and 9.76 times Online, with the *online* variable not being significant ($p = 0.453$). Columns 6 shows that this result holds when adding controls ($p = 0.418$) and a Kolmogorov-Smirnov test confirms that the distribution of allocations to the future are similar between conditions ($p = 0.494$) Additionally, Section B.1.4

[3]Columns 5 and 6 are estimated with a negative binomial regression model.

suggests that there are no differences in the use of future allocations across different cognitive levels, except for middle cognitive abilities subjects allocating more to the future in the Online condition ($p = 0.048$). We also find that there are no differences in the use of present and interior allocations between conditions, or how subjects increase their future allocations decision-by-decision.

Result 1c: Online environment has no impact on CTB allocations of tokens.

3.2 Multiple Price Lists

Time

Subjects answer the MPL in 2.06 minutes on average in the Lab and 2.43 minutes Online. Column 1 of Table 3 shows that the *online* coefficient is significant at 5% ($p = 0.021$) but Column 2 shows that this result disappears when adding controls ($p = 0.102$). However, a Kolmogorov-Smirnov test rejects the null hypothesis of equality of distribution functions ($p = 0.029$). This difference is explained by the standard deviation of response time being 37% higher in the Lab, confirmed by a variance ratio test ($p = 0.001$). However, a Kruskal–Wallis rank test ($p = 0.333$) does not reject similarity of populations. Further analyses in Section B.2.1 show that we do not find differences across different cognitive levels.

Result 2a: Online environment has no impact on MPL response time.

Table 3: OLS estimations of the impact of online setting on time response, consistency and number of future allocations in the MPL task.

	(1) Time (in minutes)	(2) Time (in minutes)	(3) Consistency	(4) Consistency	(5) # future allocations	(6) # future allocations
online	0.370**	0.275	0.048	0.039	-0.010	-0.514
	(0.160)	(0.167)	(0.038)	(0.039)	(0.591)	(0.640)
	[0.021]	[0.102]	[0.215]	[0.317]	[0.987]	[0.422]
Constant	2.062***	1.545**	0.876***	0.681***	11.885***	15.233***
	(0.122)	(0.609)	(0.031)	(0.146)	(0.456)	(2.471)
	[0.000]	[0.012]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	256	218	257	219	257	219
R-squared	0.021	0.021	0.006	0.048	0.000	0.070
Controls	No	Yes	No	Yes	No	Yes

Note: Robust standard errors in parentheses and p values in brackets. Asterisks denote significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Consistency

Subjects are consistent in MPL if they do not make multiple switching. We have 87.61% of consistent subjects in the Lab and 92.36% Online. Column 3 shows that this difference is not significant ($p = 0.215$) and Column 4 that this result is robust to adding controls ($p = 0.317$). As before, Section B.2.3 shows that there are no differences on consistency across different cognitive levels. Results of Column 4 are robust to a probit model estimation.[4].

Result 2b: Online environment has no impact on MPL consistency.

Allocations to Future

Subjects on average make 11.89 allocations to the future in the Lab and 11.79 Online. Column 5 of Table 3 shows that the *online* coefficient is not significant ($p = 0.987$) and Column 6 that this result holds when adding controls ($p = 0.422$). Additional analysis in Section B.2.3 shows that there are no differences on allocations to the future across different cognitive levels and MPL task between conditions.

Result 2c: Online environment has no impact on MPL allocations to the future.

3.3 Holt-Laury

Time response

Columns 1 of Table 4 shows that subjects take 5.38 minutes in the Lab and 5.69 minutes in the Online setting with the *online* coefficient not being significant ($p = 0.467$)) and Column 2 shows that this result is robust to adding controls ($p = 0.527$)[5]. However, a Kolmogorov-Smirnov test rejects the equality of distribution functions ($p = 0.061$) and a variance ratio test rejects equality of standard deviations between conditions ($p = 0.000$) with Online being more than the double of the Lab[6]. Further analysis in Section B.3.1 suggests that Online subjects differently apprehend the HL task: fastest subjects are faster and slowest subjects are slower than their Lab counterparts.

Result 3a: Online environment has no impact on HL response time, but increases the variance.

Consistency

Subjects are consistent in HL if they do not make multiple switching. Column 3 of Table 4 shows that consistency is 54.01% in the Lab and 57.71% Online, but the *online* coefficient is not significant ($p = 0.560$). Column 4 shows that this result is robust to adding controls ($p = 0.615$) and Section B.3.2 shows that there are no significant differences between Lab and Online conditions across

[4]The coefficient (p-value) of the online variable is 0.030 ($p = 0.413$).

[5]We excluded three outliers in the Online condition from the time analysis because they took 192.50, 202.30 and 342.08 minutes to answer the task.

[6]Online: 4.519, Lab: 2.043.

Table 4: OLS estimations of the impact of online setting on time response, consistency and number of future allocations in the HL task.

	(1) time (in minutes)	(2) time (in minutes)	(3) consistency	(4) consistency	(5) # safe choices	(6) # safe choices
online	0.311	0.247	0.037	0.032	0.954*	0.737
	(0.426)	(0.389)	(0.063)	(0.063)	(0.524)	(0.582)
	[0.467]	[0.527]	[0.560]	[0.615]	[0.070]	[0.206]
Constant	5.380***	3.371**	0.540***	0.001	11.053***	14.240***
	(0.192)	(1.454)	(0.047)	(0.221)	(0.388)	(1.947)
	[0.000]	[0.021]	[0.000]	[0.995]	[0.000]	[0.000]
Observations	254	218	257	219	257	219
R-squared	0.002	0.041	0.001	0.141	0.013	0.042
Controls	No	Yes	No	Yes	No	Yes

Note: Robust standard errors in parentheses and p values in brackets. Asterisks denotes significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

the different tasks and cognitive levels.

Result 3b: Online environment has no impact on HL consistency.

Number of safe choices

Column 5 of Table 4 shows that the average number of safe choices is 11.05 in the Lab and 12.00 Online. This difference is marginally significant ($p = 0.070$) but Column 6 shows that this result disappears after adding controls ($p = 0.206$) and a Kolmogorov-Smirnov test does not reject the equality of distributions ($p = 0.476$). Section B.3.3 shows that the number of safe choices is similar over cognitive levels and across conditions.

Result 3c: Online environment has no impact on HL number of safe choices.

3.4 Dictator Game

The endowment of subjects in the dictator game was their total earnings, and they were informed about it before answering the task. We refer to share as the percentage of earnings donated. We follow the analysis made by Brañas-Garza et al (2021). Table 5 shows regression results for the dictator game. Columns 1 shows that there is a negative and marginally significant effect of *online* setting on the share ($p = 0.070$). However, Column 2 shows that the amount of money donated is not affected by treatment ($p = 0.273$). Columns 3 to 7 explain the share reduction by a 14.5% increase in the percentage of *online* subjects donating nothing ($p = 0.033$), with a Kruskal–Wallis test rejecting the similarity of distributions between conditions ($p = 0.056$). We remark that Lab subjects who consider themselves in the presence of others donate to charity

20.97% more than Online subjects who consider themselves alone ($p = 0.029$), suggesting Hawthorne effect.

Result 4: Online environment decreases the percentage of subjects donating in the DG.

Table 5: OLS estimations of the impact of online setting on share, giving and type of altruistic behavior in the charity dictator game.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	share	*giving*	$share = 0$	$share = 50$	$share = 100$	$share < 50$	$share > 50$
online	-5.776*	-0.919	0.145**	-0.032	-0.015	0.057	-0.025
	(3.169)	(0.837)	(0.068)	(0.029)	(0.029)	(0.041)	(0.030)
	[0.070]	[0.273]	[0.033]	[0.275]	[0.603]	[0.165]	[0.416]
Constant	-12.999	-5.197	0.706***	-0.013	-0.176	1.193***	-0.179
	(13.487)	(3.723)	(0.227)	(0.088)	(0.141)	(0.154)	(0.139)
	[0.336]	[0.164]	[0.002]	[0.879]	[0.215]	[0.000]	[0.199]
Observations	219	219	219	219	219	219	219
R-squared	0.086	0.097	0.059	0.011	0.055	0.060	0.066
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses and p values in brackets. Asterisks denotes significance level: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

4 DISCUSSION

In this paper, we investigated whether lack of control over subjects influence experimental results. We designed a rich environment made of several standard economic tasks to elicit any potential differences between subjects, then have subjects answer them in the Laboratory or Online. Results suggest that there are no differences across conditions for CTB, HL and MPL tasks.
However, we find that a higher share of laboratory subjects donate in the DG. A potential explanation is that the laboratory environment influence subjects to modify their behavior when in presence of others, suggesting Hawthorne effect. In conclusion, results validate the use of online experiments by showing that lack of control does not influence the pattern of results, and suggest that future studies should investigate whether online environment elicit more truthful behavior from subjects by removing the Hawthorne effect.

REFERENCES

Amir, O., & Rand, D. G. (2012). Economic games on the internet: The effect of $1 stakes. *PloS one, 7*(2), e31461.

Andreoni, J., Kuhn, M. A., & Sprenger, C. (2015). Measuring time preferences: A comparison of experimental methods. *Journal of Economic Behavior & Organization, 116*, 451-464.

Andreoni, J., & Sprenger, C. (2012). Estimating time preferences from convex budgets. *The American Economic Review, 102*(7), 3333-56.

Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting interactive experiments online. *Experimental economics, 21*(1), 99-131.

Brañas-Garza, P., Capraro, V., & Rascón-Ramírez, E. (2018). Gender differences in altruism on Mechanical Turk: Expectations and actual behaviour. *Economics Letters, 170*, 19-23.

Brañas-Garza, P., Jorrat, D., Kovářík & López, MC. (2021). Hyper-altruistic behavior vanishes with high stakes. *PLoS ONE, 16*(8), e0255668.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty, 19*(1), 7-42.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives, 19*(4), 25-42.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*(3), 213-224.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review, 92*(5), 1644-1655.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental economics, 14*(3), 399-425.

Jorrat, D. (2021). Recruiting experimental subjects using WhatsApp. *Journal of Behavioral and Experimental Economics, 90*, 101644.

Marge, M., Banerjee, S., & Rudnicky, A. I. (2010, March). Using the Amazon Mechanical Turk for transcription of spoken language. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 5270-5273).

IEEE.

Mason, W., & Watts, D. J. (2009, June). Financial incentives and the" performance of crowds". *In Proceedings of the ACM SIGKDD workshop on human computation* (pp. 77-85).

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision making, 5*(5), 411-419.

Online appendix

Lack of Control: An experiment

Benjamin Prissé and Diego Jorrat

A Additional Information on the Recruitment of Subjects

In this section, we give additional details on the recruitment of subjects.

The average age of experimental subjects was 23.16 years old and 56.97% of them were female. The average payment was £19.48. Among the 22544 subjects who were invited to the experiment, we invited 1291 individuals to the first experimental session and 21253 individuals to the second experimental session. Among the 520 subjects (240 Lab, 280 Online) who registered to participate in the experiment, 194 subjects (73 Lab, 121 Online) registered to the first experimental session and 326 subjects (167 Lab, 159 Online) to the second experimental session. We have 257 subjects (113 Lab, 144 Online) who completed the experiment, 92 subjects (33 Lab, 59 Online) completed the experiment in the first experimental session and 165 subjects (80 Lab, 85 Online) in the second experimental session. We lost 10 subjects in the first session in what seemed to have been a technical issue. Additionally, all others subjects who attended the experiment completed it. The attrition rate was 52.58% in the first cohort (54.79% Lab, 51.24% Online) with a t-test not rejecting the equality of attrition rates between conditions (p=0.710), and the attrition rate was 49.39% in the second cohort (52.09% Lab, 46.54% Online) with a t-test not rejecting the equality of attrition rates between conditions (p=0.191). T-tests do not reject that attrition rates are equal between the two Lab experimental sessions (p=0.899) and the two Online experimental sessions (p=0.55) either.

B Additional Analysis

In this section, we present additional analysis complementing the results found in the main paper.

B.1 Convex Time Budget

B.1.1 Response time

We analyze whether the effect of online setting is heterogeneous or not with the cognitive level of subjects. We assigned each subject to a cognitive level according to their score in the CRT task: low (no correct answer), middle (one correct answer) and high cognitive level (two or three correct answers). Figure S1 presents results with gray lines specifying 95% CI, confirming that there are no difference across cognitive levels.

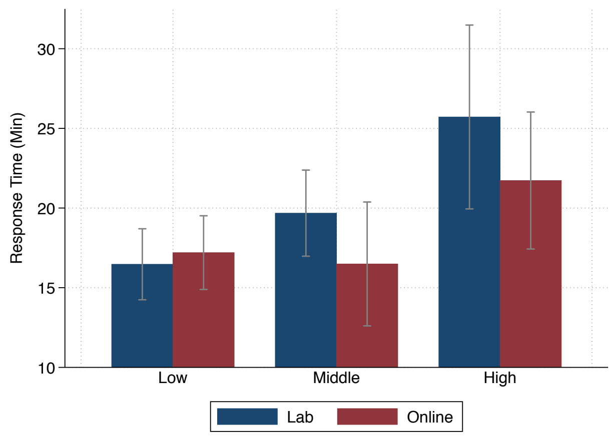

Figure S1: Time response in CTB by cognitive level

We can also suspect that some subjects lose time in the Online condition because they are distracted by the environment. These subjects would most likely be among the slowest Online subjects, therefore we separate subjects according to the median time in each condition and compare similar groups across conditions. If we focus on the 50% fastest subjects, the average time is 11.49 minutes in the Lab and 10.65 minutes Online with a t-test not rejecting the equality across conditions (p=0.243). If we focus on the 50% slowest subjects, the average time is 25.79 minutes in the Lab and 25.26 minutes Online with a t-test not rejecting the equality across conditions (p=0.711). It suggests that Online subjects are not less concentrated on the task.

B.1.2 Consistency

As before, Figure S2 shows that there are no differences in terms of consistency over different cognitive levels between Lab and Online conditions.
We can also investigate whether how subjects make inconsistent choices is different across conditions. First, we look at whether the number of inconsistencies in the early (first half of trials[7]) and late (second half of trials[8]) parts of the experiment differ across conditions. The proportion of inconsistencies in the early part of the experiment is 28.68% in the Lab and 27.13% Online with a t-test not rejecting equality across conditions (p=0.572). The proportion of inconsistencies in the late part of the experiment is 27.3% in the Lab and 27.6% Online with a t-test not rejecting equality across conditions (p=0.907). When comparing the percentage of inconsistencies between early and late parts of the experiment within conditions, t-tests do not reject equality for the Lab (p=0.621) and Online (p=0.846). We conclude that the likeliness of subjects making mistakes is

[7] until (d=7, k=70) and (a_d=0.18, a_{d+k}=0.20), rationality trial not taken into account.
[8] from (d=7, k=70) and (a_d=0.16, a_{d+k}=0.20)

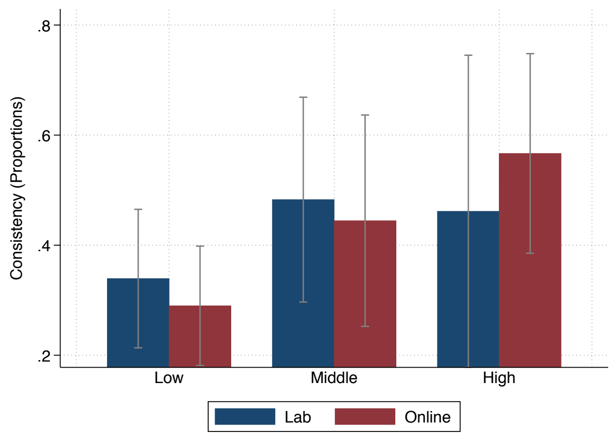

Figure S2: Consistency in CTB by cognitive level

not influenced by advancement in the task.

We also look at whether subjects make inconsistencies in similar trials across conditions. We classify trials in three types: Standard (0.XX in the present and 0.20 in the future), 20/25 (0.20 in the present and 0.25 in the future) and 20/20 (0.20 in the present and 0.20 in the future). The proportion of inconsistencies is 28.70% in the Lab and 28.24% Online for Standard trials with a t-test not rejecting equality across conditions (p=0.869), 28.77% in the Lab and 28.70% Online for 20/25 trials with a t-test not rejecting equality across conditions (p=0.976), and 20.76% in the Lab and 25.16% Online for 20/20 trials with a t-test not rejecting equality across conditions (p=0.257). We conclude that subjects are similarly inconsistent in different types of trials across conditions.

B.1.3 Number of future allocations

Figure S3 shows that there are no differences in terms of number of allocations to the future over different cognitive levels across conditions, except for middle cognitive abilities subjects who allocate 7.257 times more to the future in the Online condition (p=0.048). Further analysis does not reject that these subjects make a similar use of present (p=0.678) and interior (p=0.131) allocations, and a Kolmogorov-Smirnov test does not reject the equality of distribution of allocations to the future across conditions (p=0.383), suggesting that middle cognitive abilities subjects does not differ much across conditions.
Alternatively, we can study the use of present and interior allocations in CTB. Subjects use present allocations on average 2.70 times in the Lab and 3.20 times Online with a t-test not rejecting equality across conditions (p=0.544), and subjects use interior allocations 33.94 times in the Lab and 32.03 times

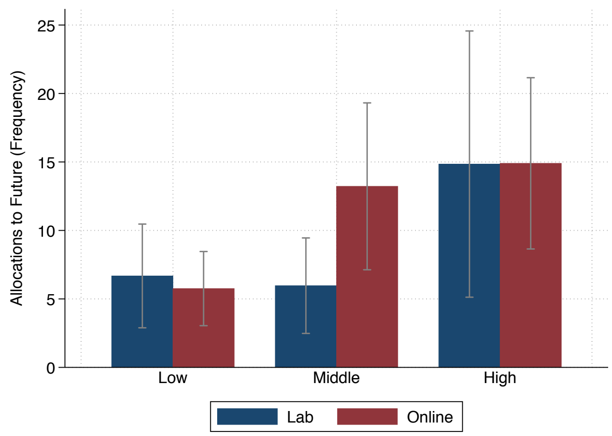

Figure S3: Number of Future Allocations in CTB by cognitive level

Online with a t-test not rejecting equality across conditions (p=0.363). We find similar results if we decompose by early payment date (d=0,7,35) or by parts of the experiment (early,late) and run t-tests across conditions, with the two significant t-tests having small effect size. We conclude that there are no differences in how subjects allocate across conditions.

B.1.4 Progression in Allocations to the Future

We can also study how subjects increase their allocations to the future when advancing to the next trial for trials with similar (d,k). More precisely, we measure progression in the future as the difference in allocations to the future between one trial and the previous one for trials with similar (d,k). The rationality test and 20/25 trials are not included. When comparing progressions to the future between the Lab and Online for identical trials with different early dates, we have three t-tests reaching significance: the first progression to the future for d=7, k=70 (p=0.032), d=35, k=70 (p=0.070) and d=7, k=98 (p=0.071) with Lab subjects systematically making larger progression to the future than Online subjects. No other t-tests reach significance.

B.2 Multiple Price List

B.2.1 Response time

Figure S4 shows that for middle and high cognitive levels there are no differences in MPL response time by treatment status, while for low cognitive levels there is a marginal difference (p=0.057) of 27 seconds that is not economically important.

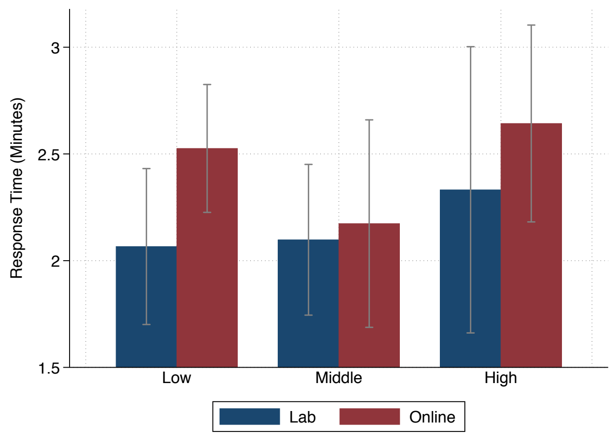

Figure S4: Time response by cognitive level in the MPL task.

B.2.2 Consistency

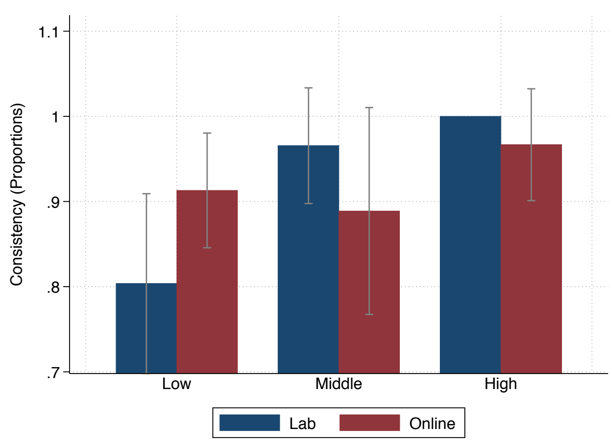

Figure S5: Consistency by cognitive level in the MPL task.

As before, Figure S5 shows that there is no difference in consistency between Lab and Online conditions across middle and high cognitive levels, while Lab subjects with low cognitive level are less consistent but the difference is marginally significant ($p = 0.087$).

If we compare the different MPL tasks, we also find no differences. Consistency of subjects is 92.03% in the Lab and 94.44% in Online setting during the first task and the difference is not significant (p=0.452). For the second MPL task, it is 94.69% in the Lab and 95.83% Online with the difference not being significant (p=0.672). And in the third MPL task, it is 90.27% in the Lab and 93.06% Online with the difference not being significant (p=0.428).

B.2.3 Number of future allocations

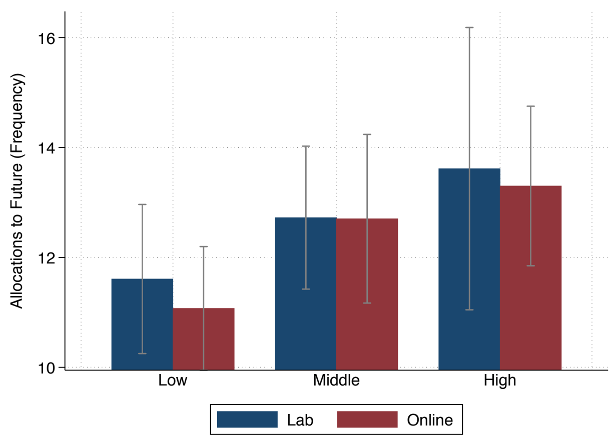

Figure S6: Number of Future Allocations in MPL by cognitive level

As before, Figure S6 shows that there are no differences in terms of number of allocations to the future in MPL across different cognitive levels between Lab and Online conditions.

Taking only consistent subjects, we observe identical switch to the future in each MPL task across conditions. In the first MPL task subjects on average switch to the future at the 3.36 choice in the Lab and the 3.47 choice Online with a t-test not rejecting equality across conditions (p=0.576). In the second MPL task subjects on average switch to the future at the 4.56 choice in the Lab and the 4.72 choice Online, with a t-test not rejecting equality across conditions (p=0.492). And in the third MPL task, subjects on average switch to the future at the 3.21 choice in the Lab and the 3.38 choice Online, with a t-test not rejecting equality across conditions (p=0.454). Overall, subjects switch to the future at the 3.71 choice in the Lab and at the 3.81 choice Online with a t-test not rejecting the equality across conditions (p=0.603). Results suggests that there are no differences in how subjects reply to the MPL task between the Lab and Online conditions.

In conclusion, we saw that subjects reply similarly to the MPL task in the Lab and Online conditions. They are similar in consistency and number of allocations to the future, but subjects with low cognitive level take more time in the Online condition. However, the effect size is reduced, so we can conclude that subjects are similar in the MPL task across conditions.

B.3 Holt Laury

B.3.1 Response Time

Figure S7 shows that there are no differences in terms of response time across different cognitive levels between Lab and Online conditions.

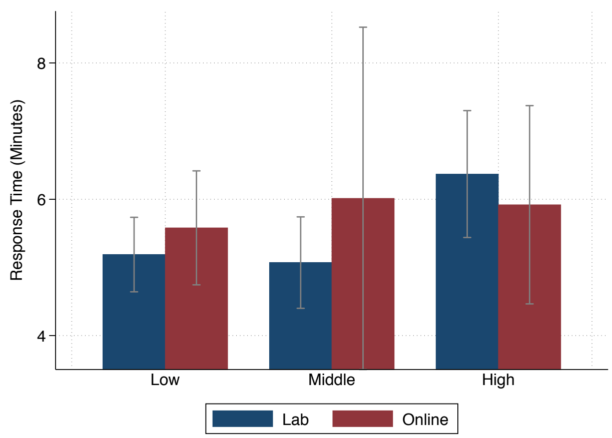

Figure S7: Response Time by cognitive level in the HL task

We can also separate subjects in each condition by the time they take to answer the task. Once again, we separate subjects by the median time in each condition (50% fastest, 50% slowest) and compare similar groups across conditions:

- Fastest subjects take 226.36sec in the Lab and 184.90sec Online, with a t-test rejecting equality of time between tasks at 1% (p<0.001)
- Slowest subjects take 418.82sec in the Lab and 504.81sec Online, with a t-test rejecting equality of time between tasks at 5% (p=0.030).

Therefore results suggest that for subjects in the Online condition, fastest subjects are faster and slowest subjects are slower than their Lab counterparts.

B.3.2 Consistency

Figure S8 shows that there are no differences in terms of consistency across different cognitive levels between Lab and Online conditions.
Additionally, we remark that we have a quite high number of subjects who are very inconsistent in at least one HL task. Subjects are very inconsistent in one HL task if they make two inconsistencies in this task. We have 35.40% in the

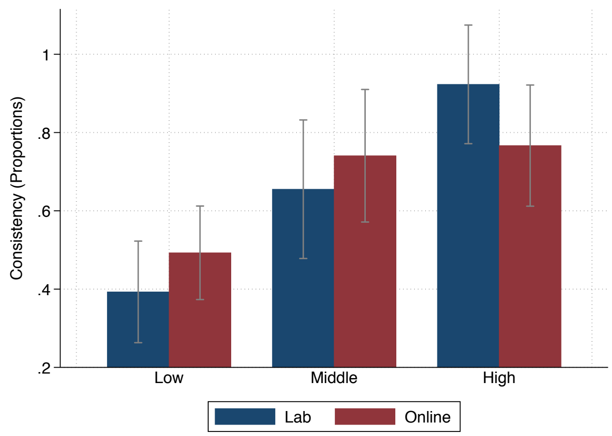

Figure S8: Consistency by cognitive level in the HL task

Lab and 33.33% Online very inconsistent subjects in the HL task, suggesting that our subjects have difficulties replying to the HL task.

B.3.3 Number of Safe Choices

Figure S9 shows that there are no differences in terms of number of safe choices across different cognitive levels across Lab and Online conditions.

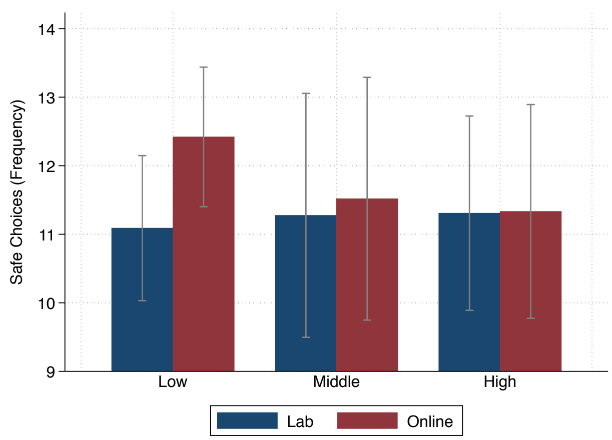

Figure S9: Number of Safe Choices by cognitive level in the HL task

Comparing the number of safe choices for subjects that are consistent in the task between conditions, we find a marginally significant (p=0.077) tendency for Online subjects to make 1.33 additional safe choices than Lab subjects.