Risky decision under laboratory deadline with experience and indirect self-selection*

Priyodorshi Banerjee† Tanmoy Das‡

August 2019

Abstract

Prior research has found that variation in the length of deadlines, or their presence, can affect risky choice, which is inconsistent with the rational benchmark. We study whether such results are affected if subject experience and some form of self-selection are allowed. We find that choices of experienced and self-selected subjects are immune to temporal aspects, making for potential consistency with the benchmark. Such immunity fails if experienced subjects are not allowed self-selection opportunities. Our overall evidence, on choice, task completion, and response time, indicates that experience as well as self-selection may be issues of importance, particularly in conjunction.

*This essay is based on an earlier article entitled "The impact of a deadline on decision under risk", derived from the third chapter of Tanmoy Das's doctoral thesis. Versions of that paper benefitted from useful comments by Sourav Bhattacharya, Sujoy Chakravarty, Sanmitra Ghosh, Parul Gulati, Matthias Klaes, Arijit Sen and P. Srikant. For the current version, we thank Martin Kocher, Sebastian Kube, Jengyi Meng, Shagata Mukherjee, Doron Sonsino, Joachim Weimann, and participants at the 7th Xiamen International Workshop on Experimental Economics, Xiamen, December 2018 and the 15th Western Economic Association International Conference, Tokyo, March 2019 for numerous helpful suggestions. The authors are solely responsible for all errors.

†*banpriyo@isical.ac.in*, Economic Research Unit, Indian Statistical Institute, Kolkata, India.

‡*tanmoy@xub.edu.in*, School of Economics, Xavier University Bhubaneshwar.

1 Introduction

Decision is often made under time pressure. A large experimental literature has emerged studying decision under time pressure, usually implemented using time constraints or deadlines. A broad finding, cutting across domains such as individual decision under risk (Ben-Zur and Bresnitz 1981), bargaining (Sutter et al 2003), and cooperation in social dilemmas (Rand et al 2012, Tinghög et al 2013), is that the extent of deadlines, or their presence, can affect choice.

The result is puzzling as economic theories, which abstract from time pressure, do not directly require information on time availability to predict choice, resulting in a deadline being a source of information formally irrelevant for decision. It has led to arguments that process models of choice, which require such information, may have advantages over static models, which do not (Goldstein and Weber 1996, Spiliopoulos and Ortmann 2018). The finding has policy implications as well. For decision under risk for instance, any policy aimed at influencing household portfolio risk would ideally have to be contingent on time pressure characteristics of target groups if household choices when allocating savings between safe and risky assets depended on time constraints faced.

A potential concern with existing findings is external validity. This paper presents a laboratory investigation of validity of results pertaining to risky decision in the pure gain domain under deadlines (the literature is reviewed in Section 4) with respect to two interrelated issues: self-selection into decision environments based on time pressure or deadlines, and experience.[1] Our primary concern is self-selection. Since individuals differ widely in terms of response time (RT; for decision problems), or normal completion time (NCT; for tasks or sets of problems) in the absence of deadlines (Benson and Beach 1996, Wegier and Spaniol 2015, Kocher et al 2019), they may have preferences over deadlines. This can prompt self-selection on the basis of deadlines in real settings where opportunities for such selection are present, leading to correla-

[1]Another concern is internal validity. Some subjects may fail to complete the task within the deadline. Internal validity is then a question if a) analysis of choice is based only on data harvested from subjects who complete, and b) such subjects differ systematically from those failing to complete. The issue has recently been scrutinized for decision under risk by Kocher et al (2019), who find grounds for concern.

tion between deadline and individual characteristics of active decision makers. For illustration, suppose an individual when facing a task would prefer the deadline to be at least some minimum. With self-selection, the group of individuals active at a particular deadline could then predominantly be those with a minimum preferred deadline (MPD) no more than the observed level. This is typically not accounted for in the literature, where deadlines are not set with reference to NCT or other individual characteristics, and randomly assigned to the subject, raising the possibility that behavior recorded at a particular deadline in the laboratory is driven by subjects with MPD more than the level set. Although discussion of this external validity concern (Benson and Beach 1996, Kocher et al 2019) has influenced the way deadlines are set (Ordóñez and Benson 1997, Nursimulu and Bossaerts 2014, Wegier and Spaniol 2015), the issue has not yet to our knowledge received systematic attention.

Our second concern in experience. Time constrained risky choices are commonly made in situations of familiarity or experience. For example, a fund manager has experience in the environment in which she has to select a portfolio under time pressure. If she thereby has some understanding of the underlying task structure, she may be more immune to the effects of time pressure than if she had no experience.[2] This can raise an independent external validity concern, as the typical approach in the literature does not account for experience, with subjects deciding under deadlines with no pre-exposure to the task.[3] In our case, an individual in the field may resolve issues of selection on the basis of experience with the task or related ones, or without such experience. The two processes are likely to have different properties. In this article, we study how opportunities for self-selection can affect the relationship between deadline and choice in settings where subjects are experienced.

Prior literature has mostly viewed the relationship between time pressure and choice through the lens of cognitive load: time pressure induces cognitive load, which impacts choice (e.g. Ben-

[2]Evidence suggests that experience vis-à-vis inexperience can activate different areas of the brain (Goel et al 2004) and prompt different reasoning strategies (Arocha et al 2005), even for formally identical tasks.

[3]It has been shown, though not for decision under risk, that experience can attenuate the impact of deadline on choice: see Rand et al 2012, Study 9.

Zur and Bresnitz 1981, Young et al 2012).[4] Our query thus stands as asking if time pressure continues to impose immoderate cognitive load when subjects are allowed experience and face deadlines which are individually optimal in some sense. For this purpose, our design approaches the issue of experience through pre-exposure, i.e., by recording choices of subjects in two separate episodes, as is standard in the literature (Kagel and Levin 1986, Hussam et al 2008), with no deadlines in the first episode. A direct approach to self-selection would require that subjects first choose a deadline from a menu, at some cost varying with deadline, and then choose for the risky task. How to determine the cost and the set of opportunities can be a difficulty with this approach, especially if subjects with preferences over deadlines also bear preferences over such trade-offs and opportunities, which can raise a further validity concern. Another problem could arise in comparing with existing results, generated by subjects who do not choose deadlines, and do not necessarily face ones individually optimal in any sense. Differences across the two groups would then not be uniquely attributable to whether they arise due to some subjects choosing personal deadlines, or due to those deadlines being personally optimal.

An alternative would be an indirect approach, which would remove menu choice from subjects, the source of problems with any direct approach, and use some rule instead to assign deadlines based on individual behavioral or other characteristics. One aspect of any indirect approach is that it would necessarily have to record relevant subject variables prior to assigning deadlines. More importantly, since an optimal deadline can presumably be revealed only through choice, deadlines produced using any such rule can at best be ‘pseudo’-optimal.

We nonetheless adopt an indirect approach, as it affords an unconfounded comparison. We develop a sorting procedure which can construct a deadline for an experienced subject equal to her estimated NCT. One group of experienced subjects are assigned such deadlines, which we call endogenous. Another group are randomly assigned these same deadlines, which we call exogenous. The effects of self-selection are examined by comparing across the two groups. The benchmark case, examined in the overwhelming majority of the literature, is where sub-

[4] See however e.g. Svenson and Maule (1993) for affective interpretations of the relationship between time pressure and choice.

jects are inexperienced and deadlines are exogenous. We examine the effects of experience by comparing benchmark behavior with that of experienced subjects facing exogenous deadlines. Our endogenous sorting procedure and how it compares to procedures used in the literature are detailed in Section 3.3.

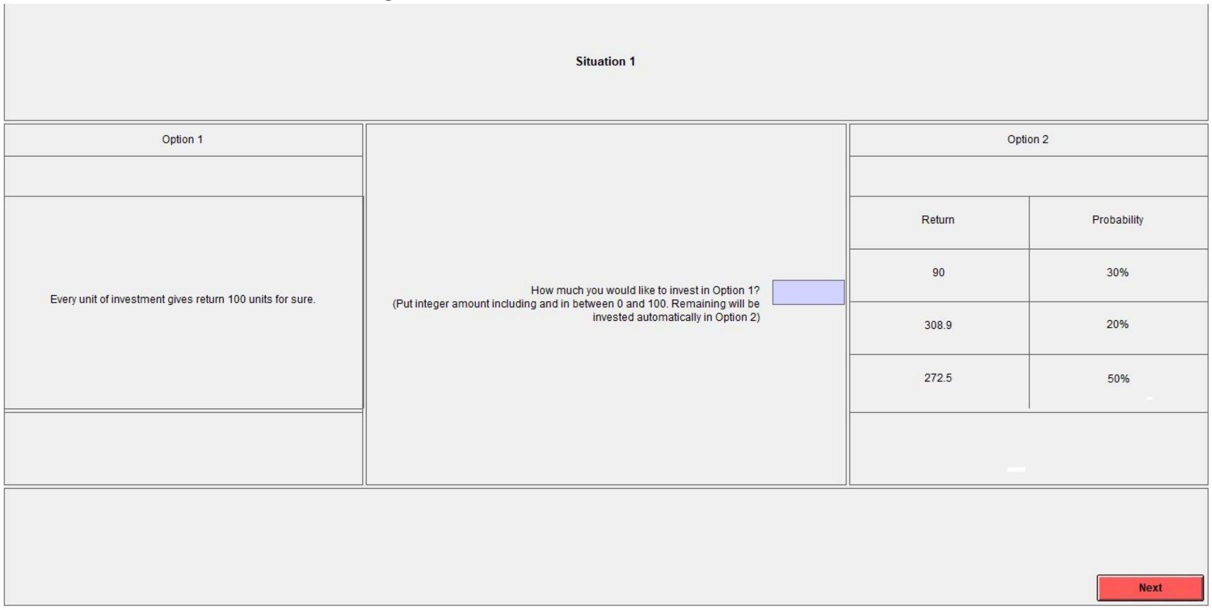
We find in the benchmark case that the presence of deadlines or variation in their length can affect decision. Specifically, imposing a deadline or reducing its length tend to make choice riskier. This is inconsistent with static models of choice and replicates the majority finding in the literature. For experience, we pursue the conjecture that it lessens the impact of exogenous deadlines. Our evidence supports this conjecture, though it shows that decisions of experienced subjects facing exogenous deadlines are also inconsistent with static models of choice. For self-selection, we pursue the conjecture that experienced subjects should be less affected by deadlines when these are endogenous rather than exogenous. Our evidence supports this conjecture, and suggests that time pressure may fail to impact an experienced individual's decision if her deadline is (pseudo-)optimal, indicating that decisions of experienced and self-selected individuals may be consistent with static models of choice. Overall, our findings, on choice, task completion, and RT, indicate that both concerns may be of importance, particularly in conjunction. The significance of these factors in influencing choice suggests that caution should be exercised when extending existing findings to understand behavior in the wild.

The paper is organized as follows. The next section describes experimental design and procedure. Section 3 contains preliminary analysis, while Sections 4 and 5 present the main results. Section 6 concludes. Instruction sheets and other experimental details are in the Appendices.

2 Design and procedure

We first describe risky decision tasks and problems. Any problem faced by any subject required the allocation of a budget of 100 across two assets, one safe and one risky. The safe or risk-free asset yielded 100 per unit of investment. The risky asset had 3 possible positive returns per

Figure 1: Decision screen (no deadline)



unit of investment, one less and two more than 100, with specified probabilities. Our study is thus in the pure gain domain.[5] The risky asset had a mean return of approximately 225 per unit of investment, though this information was not communicated to subjects. Figure 1 gives an example, and also shows the decision screen faced by any subject in the absence of a deadline.

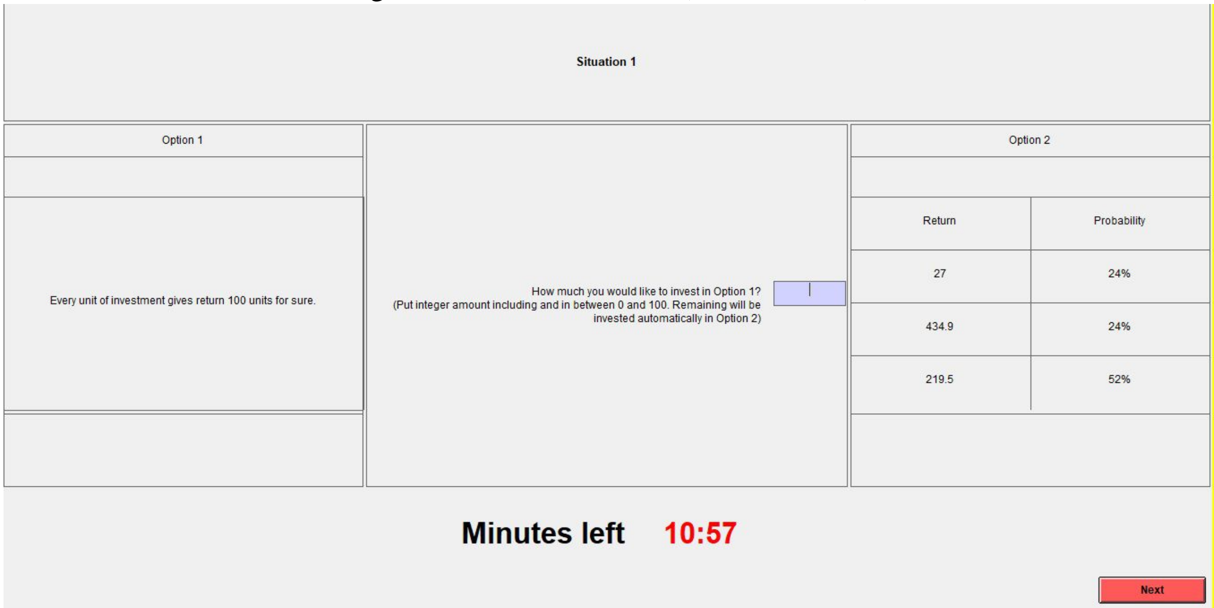
Risky assets in all problems faced by a given subject had a fixed variance per unit of investment. Half the subjects faced risky assets with variance approximately 8,000. These subjects were in the low-risk or **L** sub-conditions. The other half faced risky assets with variance approximately 20,000. These subjects were in the high-risk or **H** sub-conditions. This was the only difference between sub-conditions, given any condition (see Section 2.1). Variance information was not communicated to subjects.[6] An example of a low-risk problem is given in Figure 1 above. An example of a high-risk problem is given in Figure 2 below.

Twenty problems in a fixed sequence constituted one task. Two tasks were used in the **L**

[5]This is under the assumption that receiving less than 100 is not viewed as a loss, i.e., no reference point is formed at the return yielded by the risk-free asset.

[6]We introduced this variation as some prior papers had found that the relationship between deadline and choice could be dependent on the level of risk; see Section 4.

Figure 2: Decision screen (with deadline)



(resp. **H**) sub-conditions, 1L and 2L (resp. 1H and 2H), with no problem in common. Details of tasks and problems can be found in Appendix A. Subjects faced one task per episode.

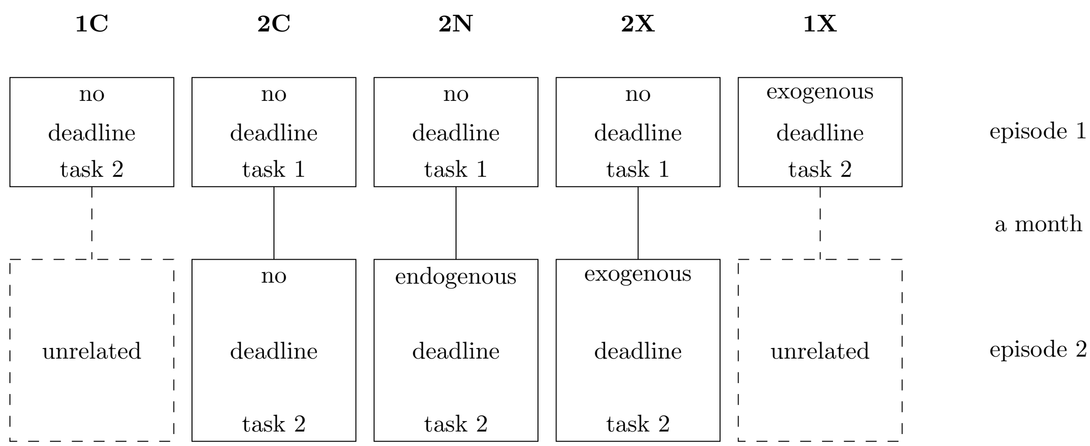
Any deadline set was at the level of the task/episode. Deadlines can be set at the level of the task, or at the level of the problem. Both types of implementation exist in the literature (e.g., Kirchler et al 2017 set deadlines for every problem, while Kocher et al 2019 set deadlines for tasks), and both have empirical relevance (Gabaix et al 2006, Kocher et al 2019). We impose deadlines at the level of the task as it eases analysis by suppressing sequence effects. Figure 2 shows the decision screen faced by any subject in the presence of a deadline.

2.1 Conditions

The experiment is between subjects and consists of five conditions, two controls and three treatments, each divided into two sub-conditions, **L** and **H**. Subjects made two appearances, with about a month separating the two episodes.[7] Subjects never faced deadlines in the control

[7]Inter-exposure intervals used in the literature on decision under risk vary. Studies of stochastic choice (Mosteller and Nogee 1951) use a short interval (within the same session), while test-retest reliability studies

Figure 3: Conditions



Each condition divided into two sub-conditions, **L** and **H**

conditions. The main control condition is **2C**. Subjects faced task 1 in episode 1 (task 1L in **2CL**, and task 1H in **2CH**), and task 2 in episode 2 (task 2L in **2CL**, and task 2H in **2CH**). The subsidiary control condition is **1C**. Episode 2 of this condition is unrelated to the current study. Subjects in **1C** faced task 2 in episode 1 (task 2L in **1CL**, and task 2H in **1CH**).

For the treatments, subjects faced task 1 in episode 1 of conditions **2N** and **2X** (task 1L or 1H depending on sub-condition), in the absence of any deadline, and task 2 in episode 2 (task 2L or 2H depending on sub-condition), in the presence of a deadline. Episode 2 of condition **1X** is unrelated to the current study. Subjects in **1X** faced task 2 in episode 1 (task 2L in **1XL**, and task 2H in **1XH**), in the presence of a deadline. Figure 3 lists all conditions. Deadlines actually implemented are listed in Section 3.3.

(Camerer 1989) use a gap of a few days (usually a week or less). Our choice is guided by the literature on intertemporal stability (Section 4.3), which uses an interval of at least one month (Stott 2006, Zeisberger et al 2012).

Even though tasks 1 and 2 are a priori similar, the use of different tasks creates potential confounds, such as in comparing across episodes for **2C**, **2N** and **2X**, and in comparing episode 1 of **1X** with episode 1 of **2X** (and of **2C** and **2N**). Such confounds are eliminated if tasks 1 and 2 are indeed similar. We use **1C** to check for similarity of tasks; details are in Section 3.1.

The reason why we use different tasks is as follows. Debriefing of and comments from a few subjects in some pilot experiments which used the same problems across episodes led us to understand that some subjects a) realized that some or all problems were the same across episodes, and b) tried to recall and give the same answer to a problem in episode 2 as the one they had given in episode 1, being under the impression that that was what the experimenter wanted. The simplest way to avoid such apparent experimenter demand effects, triggered by certain memory states, was to use different problems/tasks across episodes. We used task 2 in episode 1 of **1X** for ease of comparison with episode 2 of **2X**.

2.2 Procedure

Subjects were undergraduate students from a variety of disciplines studying in colleges and institutions in the Calcutta area, and had been recruited to participate "in a research study on financial decision making". They knew at the time of recruitment that two appearances would be required, separated by about a month. No other information about episodes was given at this time. Subjects were randomly assigned to conditions/sub-conditions, and no subject participated in more than one.

Data collection was on an individual basis. On arriving for any episode, subjects were given a pencil and a sheet of paper, and seated in front of a computer, after which they were administered instructions for that episode (see Appendix B for instructions). The instructions included an example, and were followed by two unpaid trial problems, none of which involved deadlines. The decision phase commenced at the end of the second trial problem. The decision screen in the absence of a deadline was common across trial and payoff-relevant problems (see Figure 1). A countdown was added if there was a deadline (see Figure 2). No information about episode 2 was given during or after the end of episode 1.[8]

All conditions required participation in two episodes. We only use data from subjects who completed both episodes. About 90% of subjects who completed episode 1 returned to complete episode 2.[9] We had 45 subjects per sub-condition, leading to 90 per condition, for a total of 450. We collected gender and family income information at the end of the second episode for each subject (the overall sample was about 40% male). No outcome information was conveyed for any payoff-relevant decision in any episode until the end of episode 2.

Subjects received INR 1 for every 1200 points (units of return) earned for problems completed across the two episodes.[10,11] Randomization was independent across subjects and problems. Salient payment rules for an episode were announced at its beginning. Decisions in reality usually lead to some payoff. Our choice to pay for all decisions was in part an emulation given our concern with external validity. It further made it easy to communicate the penalty associated with non-decision in the presence of a deadline. The choice was also guided by results of pilots which revealed insubstantial outcome difference compared to a payment scheme where one decision is randomly selected for payment, and a faint subject preference in favor of all decisions being paid. Subjects additionally received a show-up fee of INR 50 for the first episode, and INR 100 for the second episode. The show-up fee for the first episode was given at the end of the episode. All other earnings were given at the end of the second episode. Subjects received about INR 600 on average, including show up fees, across the two episodes.[12]

[8] Subjects in **2N** were thus not made aware that outcomes from episode 1 would determine deadline in episode 2. Note there was no deception, as no statements in the instruction were false or ambiguous. The effect of pre-announcing the sorting rule could also be of research interest, and is left for future work.

[9] Tests indicated episode 1 behavior was indistinguishable across subjects who eventually returned to participate in episode 2, and those who did not.

[10] No return accrued for any of the 20 problems left uncompleted in an episode with a deadline.

[11] The purchasing power parity exchange rate between the Indian Rupee and the US Dollar for 2010 was 16.84 rupees to a dollar according to the Penn World Tables (see Heston et al 2012).

[12] Nominally about USD 8.65 on 1 August, 2019.

3 Preliminaries

Subjects in episode 1 of **2C**, **2N** and **2X** faced the same task (given sub-condition) in the absence of any deadline. As such, we expect subjects to be behaviorally equivalent across **2CL**, **2NL**, and **2XL**, and also across **2CH**, **2NH**, and **2XH**. We first present results in this regard. These set the stage for the manipulation check with respect to task similarity (Section 3.1).

Subjects in episode 1 of these conditions took 20 decisions. This was true for subjects in condition **1C** as well. As a convention to be maintained throughout the paper, we take decision for a problem to be the allocation to the safe asset in it. For each subject we calculated her average allocation to the safe asset in episode 1. We also calculated her average RT per problem (in seconds) across the 20 problems. Table 1 below gives mean allocation to the safe asset and mean RT across subjects. Subjects in condition **1X** faced a deadline in episode 1, and some of them did not complete all problems within it. Averages for these subjects (for choice as well as RT) were computed only across decisions actually taken within the deadline.[13]

Table 1: Means - episode 1 (average across subjects)

	Per problem allocation to safe asset				
	1C	**2C**	**2N**	**2X**	**1X**
L	49.54	46.8	48.81	48.73	37.27
H	56.72	52.43	55.15	53.66	49.18
	Per problem response time (seconds)				
L	28.09	28.3	32.5	31.15	17.99
H	31.13	31.17	33.37	31.16	18.04

90 subjects in each condition (45 in each sub-condition)

Tables 2 and 3 below presents OLS regression results for comparisons of choice and RT respectively across relevant sub-condition pairs. The dependent variable is average allocation to the safe asset for each case in Table 2, and average RT for each case in Table 3. A condition dummy is the main independent variable and we expect insignificance in all cases. Gender and income were used as control variables. The upper and lower panels of the tables report results for comparisons across **2CL**, **2NL**, and **2XL**, and across **2CH**, **2NH**, and **2XH** respectively.

[13]See Table 15 in Section 5.1 for data on the number of decisions not taken.

Table 2: Comparing mean episode 1 allocations across **2C**, **2N** and **2X**

	Dep. var: average allocation to safe option		
	2CL1 vs **2NL1**	**2CL1** vs **2XL1**	**2NL1 vs 2XL1**
condition	1.28	1.461	-0.056
dummy	(3.374)	(3.861)	(3.877)
gender	5.247	12.158**	4.059
	(3.537)	(3.91)	(4.041)
income	0.594	0.154	-0.614
	(1.173)	(1.361)	(1.39)
constant	34.056**	26.389**	43.93***
	(10.49)	(8.379)	(10.685)
Adj. R^2	-0.003	0.074	-0.021
	2CH1 vs **2NH1**	**2CH1** vs **2XH1**	**2NH1 vs 2XH1**
condition	4.113	1.648	-1.851
dummy	(3.266)	(2.808)	(3.343)
gender	9.353**	6.056*	5.214
	(3.309)	(2.945)	(3.448)
income	0.469	0.081	0.104
	(1.319)	(1.031)	(1.377)
constant	27.035*	41.243***	48.694***
	(11.804)	(7.311)	(8.857)
Adj. R^2	0.061	0.017	-0.005

*No. of observations = 90. Standard errors are in parentheses. Condition dummy takes values 0 or 1. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. *, ** and *** respectively indicate significance in terms of two-tailed p-values at the 5%, 1% and 0.1% levels*

The tables expectedly show equivalence of choice as well as RT across conditions, given sub-condition (all p-values were more than 15%). The results thus confirm the a priori hypothesis of behavioral equivalence. We also conducted equivalent tobit regressions, comparison tests, classical and non-parametric, single- and multi-sample, and distributional comparison tests, classical and non-parametric, which overwhelmingly yielded the same outcome.[14]

Table 3: Comparing mean episode 1 response times across **2C**, **2N** and **2X**

	Dep. var: average response time per problem		
	2CL1 vs **2NL1**	**2CL1** vs **2XL1**	**2NL1 vs 2XL1**
condition	3.38	3.039	-2.541
dummy	(3.746)	(2.672)	(3.817)
gender	7.226	2.283	3.349
	(3.927)	(2.706)	(3.979)
income	-2.228	0.629	-2.796*
	(1.302)	(0.942)	(1.368)
constant	17.273	21.168***	38.44***
	(11.646)	(5.799)	(10.52)
Adj. R^2	0.056	-0.006	0.021
	2CH1 vs **2NH1**	**2CH1** vs **2XH1**	**2NH1 vs 2XH1**
condition	3.214	0.088	-2.866
dummy	(4.615)	(2.868)	(4.267)
gender	3.401	0.545	4.318
	(4.675)	(3.008)	(4.4)
income	1.779	0.305	2.77
	(1.864)	(1.053)	(1.757)
constant	13.062	29.186***	21.126
	(16.678)	(7.467)	(11.304)
Adj. R^2	-0.016	-0.034	0.002

*No. of observations = 90. Standard errors are in parentheses. Condition dummy takes values 0 or 1. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. * and *** respectively indicate significance in terms of two-tailed p-values at the 5% and 0.1% levels*

[14]Equivalent tobit regressions were conducted for all OLS results reported in the paper. We report only the OLS results, as tobit results were qualitatively identical in all cases.

3.1 Manipulation check I: are tasks 1 and 2 similar?

Given sub-condition, tasks 1 and 2 contain problems which are structurally similar. This leads to the ex ante expectation that the two tasks are similar, by which we mean they should generate equivalent behavior. We check for equivalence in this section. Results presented above suggest that behavior is equivalent for episode 1 across **2C**, **2N** and **2X**, given sub-condition. We therefore pool data from these conditions together and analyze whether choice and RT for these conditions for episode 1, generated by subjects facing task 1, are comparable to choice and RT for episode 1 of **1C**, generated by subjects facing task 2.

Table 4 below reports OLS results comparing outcomes in **1C** to those in **2C**, **2N** and **2X**, pooled together, for episode 1. The first and second columns give results from the comparisons of choices, and the third and fourth columns give results from the comparisons of RT. For the former (latter) case, the first (third) column gives results for the **L** sub-conditions, while the second (fourth) column gives results for the **H** sub-conditions. The dependent variable for the first two (last two) columns is average allocation to the safe asset (average RT). A condition dummy is the main independent variable for each column and we expect insignificance in all cases. Gender and income were used as control variables.

The table shows that choices and RT are the same across the conditions, given sub-condition (all p-values were more than 15%).[15] We conclude therefore that the tasks are equivalent. Results presented in this section so far hence create a basis for comparing a) across episodes for **2C**, **2N** and **2X**, b) across **2C**, **2N** and **2X** for episode 2, and c) episode 1 of **1X** with episode 1 of **2X** (and of **2C** and **2N**).

3.2 Manipulation check II: does pre-exposure yield experience?

Pre-exposure is a routine manipulation in experimental research designed to generate experience. This manipulation is rarely checked however. An exception is, e.g., Nosofsky and

[15] Similar results obtained from binary comparisons of **1C** separately with each of **2C**, **2N**, and **2X**.

Table 4: Comparing episode 1 means: **1C** versus **2C**, **2N** and **2X** pooled

	Dependent variable: per problem			
	average allocation		average response time	
	L	**H**	**L**	**H**
condition	-2.784	-3.409	2.161	0.999
dummy	(3.042)	(2.493)	(3.407)	(3.422)
gender	7.509**	5.423*	2.507	-0.465
	(2.666)	(2.203)	(2.985)	(3.024)
income	-0.206	-0.525	0.066	1.384
	(0.901)	(0.804)	(1.009)	(1.104)
constant	39.235***	50.17***	24.181***	27.198***
	(5.606)	(5.046)	(6.278)	(6.926)
Adj. R^2	0.028	0.029	-0.01	-0.007

*No. of observations = 180. Standard errors are in parentheses. Condition dummy is 0 for **1C**, and 1 otherwise. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. *, ** and *** respectively indicate significance in terms of two-tailed p-values at the 5%, 1% and 0.1% levels*

Palmeri (1997), who conclude on the basis of cross-exposure stabilization of choice and RT in the context of a perceptual classification task that an inter-exposure interval of a minimum of a few days is required for familiarization or experience.

It may be particularly important to check this manipulation in the current case because of our concern with external validity. Choice data is not appropriate for the purpose of constructing a test as there is no objective standard for gauging correctness of subjects' responses for decision under risk. The problem is compounded because of the absence of full agreement regarding the nature of underlying preferences in this domain. Hence an objectively experienced subject may or may not display choice stability, and observation of any particular pattern of inter-exposure choice variation may not necessarily indicate experience or its lack.

We appeal to RT data to construct a test. We assume that an experienced subject in the absence of any deadline should complete a task faster than an inexperienced one. We therefore take data from **2C**, with no deadlines in either episode, to study whether average RT falls from episode 1 to 2. Table 5 gives means of subjects' average RT in episodes 1 and 2.

The table above indicates that average RT fell from episode 1 to 2. We formally investigated

Table 5: Mean RT per problem in seconds - condition **2C**

	episode 1	episode 2
L	28.3	18.45
H	31.17	18.68

45 observations for each cell

this possibility through regression analysis. We took average RT per problem as the dependent variable.[16] There were two observations per subject, one per episode. We used subject specific dummies to control for subject fixed effects, and cluster correction at the level of subject to allow adjustment of the variance-covariance matrix to the lack of independence of observations.

The main independent variable is an episode dummy, taking value 0 for episode 1, and 1 for episode 2. The hypothesis is that this variable should have a negative coefficient, indicating a lower NCT for episode 2. Subjects' average allocation to the safe asset for episode 1 was used as a control variable, as were gender, and income. Table 6 below give the output from an OLS model (subject specific dummy variables are not reported for brevity). The first and second columns give results respectively for the **L** and **H** sub-conditions. The table shows that average RT is lower in episode 2 than episode 1 for **2C**. This confirms the null hypothesis and suggests that pre-exposure indeed generates experience in our sample. We note that the legitimacy of this approach rests on tasks 1 and 2 being equivalent, as demonstrated in Section 3.1.

3.3 Constructing and assigning deadlines

We first discuss endogenous deadlines, faced by subjects in **2N** for episode 2. A subject in this condition faced a deadline equal to her estimated NCT for the task in episode 2, rounded to the nearest minute.[17] This is her optimal deadline under two assumptions: one, that her optimal deadline is her MPD, and two, that her MPD is her NCT, i.e., the time she would have taken to complete the task had she faced it without a deadline. We use data from subjects in **2C** to

[16] See Section 4.3 for a comparison of choice across episodes. Analysis of how RT develops within an episode (not reported for brevity) indicates that RT tended to fall over the sequence of problems.

[17] Rounding was done as pilot studies indicated that subjects preferred to face a deadline expressed in whole minutes.

Table 6: Comparing response times across episodes: **2C**

	Dep var: avg RT	
	L	**H**
episode	-9.853***	-12.484***
dummy	(2.307)	(2.365)
avg allocation	0.276***	-0.328***
(episode 1)	($1.08e^{-15}$)	($9.19e^{-}16$)
gender	-0.078***	-2.697***
	($5.99e^{-}14$)	($3.04e^{-}14$)
income	-3.893***	-1.034***
	($1.27e^{-}14$)	($8.02e^{-}15$)
constant	43.836***	64.859***
	(3.46)	(3.548)
R^2	0.79	0.845
Obs (clusters)	90 (45)	

*Subject dummies not reported. Robust standard errors are in parentheses. Episode dummy is 0 for episode 1, and 1 otherwise. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. *** indicates significance in terms of two-tailed p-values at the 0.1% level* estimate NCT in episode 2 as a function of observed behavior from episode 1, and apply this estimated relationship for each subject in **2N**.

We estimate NCT for the task in episode 2, for **2C**, using a regression approach. Estimation is done for each sub-condition separately. The dependent variable in these equations is average RT per problem in seconds for episode 2. As independent variables, we used average RT per problem in seconds and average allocation in episode 1. No variable other than average RT in episode 1 was significant in any specification, and this variable was significant in every specification in which it was included. We used OLS, and adjusted R-square as the criterion for selection. This yielded in either case the specification with average RT in episode 1 as the only independent variable. These estimates are given below in Table 7.[18]

Table 7: Response time prediction equation - condition **2C**
Best fit estimates

Dep. var: average RT - episode 2		
	L	**H**
average RT (episode 1)	0.36*** (0.089)	0.431*** (0.072)
constant	8.273** (2.75)	5.25* (2.485)
Adj. R-sq	0.258	0.442
No. of obs	45	

*Standard errors are in parentheses. * , ** and *** respectively indicate significance in terms of two-tailed p-values at the 5%, 1% and 0.1% levels*

Any of these estimating equations yield average RT per problem for episode 2 in seconds. Division by 3 (multiplication by 20, the number of problems, and then division by 60) converts this figure to NCT for the task in minutes, possibly with a fractional part. Rounding off yields whole minutes, which is what we used to set deadlines. As an example, a subject in **2NL** whose

[18]Gender and income could not be used as independent variables, as these pieces of information were collected at the end of episode 2. Thus, although we had information on these variables for subjects in **2C** at the time the equations were estimated, we did not have the relevant information for subjects in **2N** at the time the estimated equations were applied to generate deadlines. As it happens, these variables never assumed significance, and the specifications reported in Table 7 remained the ones with the best fit.

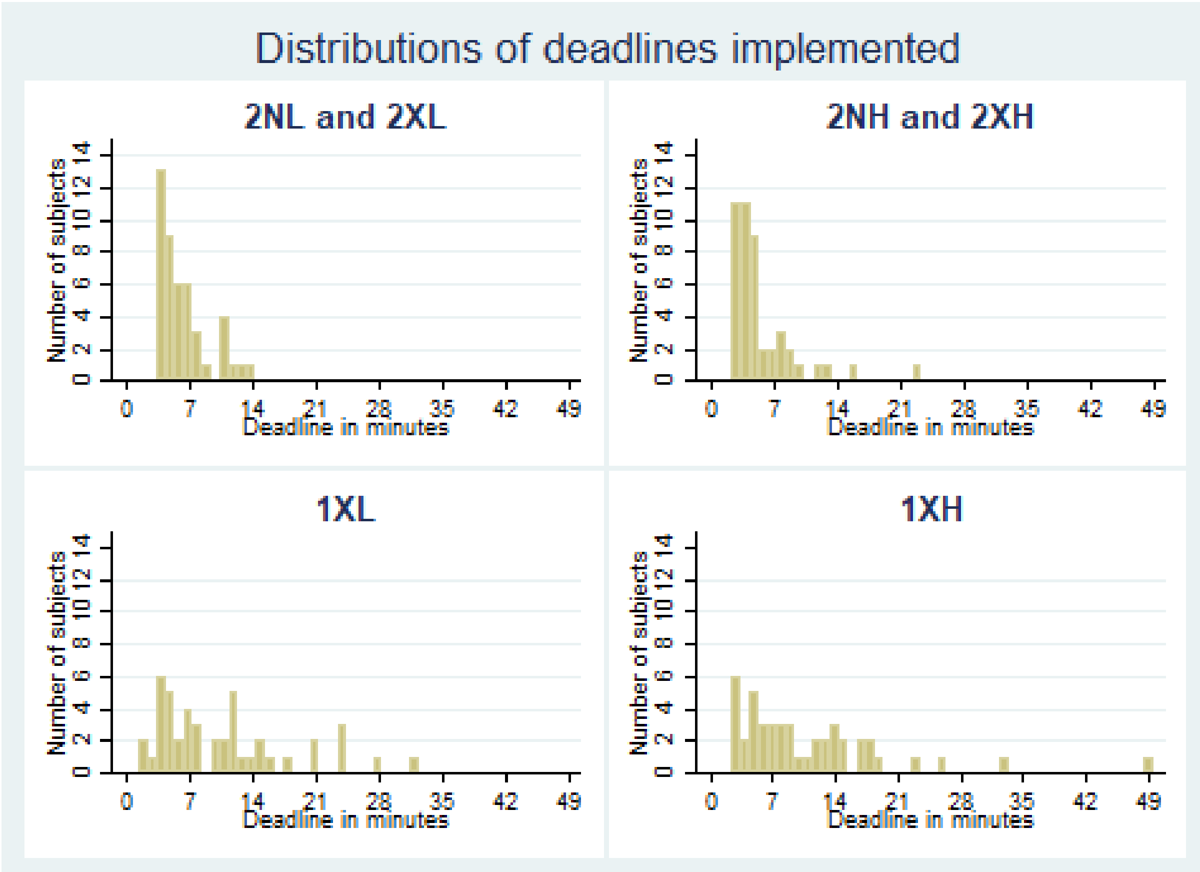


Figure 4: Deadline distributions

average RT in episode 1 is 31 seconds would face a deadline of 6 minutes in episode 2. Our procedure ensures that the average deadline faced by subjects in episode 2 of **2NL** (resp. **2NH**) is approximately equal to average NCT of subjects in episode 2 of **2CL** (resp. **2CH**).

We now turn to exogenous deadlines. We first discuss condition **2X**. Subjects in **2XL** (resp. **2XH**) faced the same set of deadlines as subjects in **2NL** (resp. **2NH**). The difference was that while the deadline for a subject in **2N** was linked to her episode 1 NCT through the sorting procedure discussed above, there was no such necessary linkage for any subject in **2X**, as the deadlines were randomly assigned to these subjects. Nevertheless, the average deadline faced by subjects in episode 2 of **2XL** (resp. **2XH**) continues to be approximately equal to average NCT of subjects in episode 2 of **2CL** (resp. **2CH**). Figure 4 below gives the distributions of deadlines implemented, with the top panel showing deadlines for conditions **2N** and **2X**.

We now discuss condition **1X**. We recall that subjects in this condition faced deadlines in episode 1. One option was to implement the same deadlines in **1XL** (resp **1XH**), as in **2NL** or **2XL** (resp. **2NH** or **2XH**). Since NCT falls with episode (Section 3.2), and deadlines in **2N** or **2X** are benchmarked to NCT for episode 2, this would have meant that average deadlines in **1X** would have been much less than average NCT for episode 1.

The other option was to maintain equality of average deadlines for **1X** with average NCT for episode 1. We chose this option. Since episode 1 NCTs from **2N** were used to generate episode 2 deadlines for **2N** and **2X**, we used episode 1 NCTs from **2N** to generate deadlines for **1X** as well. Specifically, we took the set of measured NCTs from **2NL** (resp. **2NH**), and set these as deadlines in **1XL** (resp. **1XH**), after rounding, with random assignment. The bottom panel of Figure 4 above shows the distributions of deadlines implemented in condition **1X**.

The traditional approach in the literature has been to set exogenous deadlines. Earlier papers (e.g., Ben-Zur and Bresnitz 1981) had deadlines determined by the experimenter, without any explicit foundation. Following the self-selection critique of Benson and Beach (1996), Ordóñez and Benson (1997) introduced an alternative procedure, whereby treatment groups face deadlines benchmarked to average NCT as measured behaviorally for a control group. Another alternative procedure has been introduced by Nursimulu and Bossaerts (2014), whereby treatment groups face deadlines benchmarked to average NCT as measured neuroeconomically for (external) control groups. Both these alternative procedures remain exogenous.

Our procedure is similar to that of Wegier and Spaniol (2015), who present the only prior example to our knowledge of non-exogenous deadlines. They had a practice session immediately prior to the main session. The task in the practice session was a truncation of the task in the main session. They measured RT for problems in the practice session at the level of the subject and set a subject's deadline for corresponding problems in the main session equal to 75% of her RT in the practice session (no justification is provided for this specific discount). The evidence from Nosofsky and Palmeri (1997) cited above suggests that subjects in Wegier and Spaniol (2015) are not experienced prior to the main session, which is an important difference with our work. Additionally, they do not approach the issue of self-selection, as they present no comparison with subjects facing exogenous deadlines which are equivalent on average.

4 Analysis of choice

We first study the benchmark case, with inexperienced subjects facing exogenous deadlines, on average equal to episode NCT. Prior findings in the pure gain domain are not uniform. Kocher et al (2013) found that subjects' choices under a tight deadline did not differ from their choices when they were effectively unconstrained. All other papers have found that the presence of a deadline, or variation in the length of one, can impact choice. Young et al (2012), Hu et al (2015) and Saqib and Chan (2015) report that deadlines induce more risky choices. Wegier and Spaniol (2015) report similarly, though their finding is weak. Kirchler et al (2017), on the other hand, found that a deadline leads to weakly less risk-taking relative to when no deadlines are present. Finally, Dror et al (1999) find that deadlines can induce less risky choices at low risk levels and more risky choices at high risk levels.[19]

Two pieces of analysis are involved. One, determining whether choices of subjects in episode 1 of **1X** vary with deadline faced. OLS regression results are presented below in Table 8. The first and second columns give results respectively for the **L** and **H** sub-conditions, while the third column gives aggregated results. The dependent variable is average allocation to the safe asset, with averaging over all problems completed within the deadline. We follow this averaging method for average allocation to the safe asset, as well as average RT, for all summary statistics and regression results presented in the rest of the paper.

The main independent variable is the deadline faced. The majority result in prior literature is that a shorter deadline causes choice to become riskier. We posit as the null hypothesis therefore a positive sign for this variable. Other variables for all three columns are the number of problems

[19]Lack of uniformity of results characterize studies, of which there are fewer, in mixed (e.g. Ben-Zur and Bresnitz 1981, Busemeyer 1985) and pure loss domains as well (e.g. Young et al 2012, Kocher et al 2013, Kirchler et al 2017).

completed within the deadline, gender, and income. The level of risk is also included for the aggregated case. The null hypothesis is confirmed for the **H** sub-condition and the aggregated data. A significant relationship exists for the **L** sub-condition at the 10% level.

Table 8: Choice and the extent of the deadline: **1X**

	Dep var: avg episode 1 alloc to safe asset		
	L	**H**	Aggregated
deadline	0.688†	0.477*	0.573**
(minutes)	(0.407)	(0.223)	(0.208)
risk level	-	-	11.141***
	-	-	(2.878)
no. of probs	-0.957	-0.177	-0.653
	(0.816)	(0.773)	(0.55)
gender	6.823	7.779	7.712*
	(5.6)	(3.908)	(3.265)
income	-2.983*	-1.945	-2.534**
	(1.341)	(1.102)	(0.859)
constant	47.228**	40.825**	39.725**
	(17.253)	(14.154)	(11.049)
Adj. R^2	0.256	0.241	0.359
No. of obs	45		90

*Standard errors are in parentheses. Risk dummy is 0 if **L**, 1 for **H**. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. †, *, ** and *** respectively indicate significance in terms of two-tailed p-values at the 10%, 5%, 1% and 0.1% levels*

The second piece of analysis involves determining whether choices in episode 1 of subjects in **1X** differ from corresponding choices of subjects in **1C**, **2C**, **2N** and **2X**, all pooled:[20] results presented earlier in Table 1 hint that subjects made riskier choices in **1X**. OLS regression results are presented below in Table 9. The first and second columns give results respectively for the **L** and **H** sub-conditions, while the third column gives aggregated results. The dependent variable is average allocation to the safe asset.

The main independent variable is a condition dummy, which takes value 1 for condition **1X**, and 0 otherwise. Prior majority results and the findings reported in Table 1 leads us to posit a

[20]The comparison between **1X** and **1C** is direct, as all these subjects faced task 2. Pooling across **1C**, **2C**, **2N** and **2X** is permitted by the results from Section 3.1. Results do not differ across the two cases, and we report below only from the pooled comparison.

negative sign for this variable as the null hypothesis. Other variables for all three columns are the number of problems completed within the deadline, gender, and income. The level of risk is also included for the aggregated case. The null hypothesis is confirmed in all cases.

Table 9: Comparing mean allocations across **1X** and **1C**, **2C**, **2N** and **2X**, pooled

	Dep var: avg episode 1 alloc to safe asset		
	L	**H**	Aggregated
cond dum	-11.52***	-6*	-8.883***
	(3.014)	(2.515)	(1.958)
risk level	-	-	7.196***
	-	-	(1.484)
no. of probs.	-0.455	-0.204	-0.378
	(0.745)	(0.68)	(0.505)
gender	8.215**	6.102**	7.282***
	(2.354)	(1.968)	(1.528)
income	-1.15	-0.95	-1.043*
	(0.754)	(0.672)	(0.506)
constant	48.046**	51.961***	47.166***
	(15.288)	(13.467)	(10.225)
Adj. R^2	0.107	0.059	0.131
No. of obs.	225		450

*Standard errors are in parentheses. Condition dummy is 1 if condition **1X**, 0 if any other condition. Risk dummy is 0 if **L**, 1 for **H**. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. *, ** and *** respectively indicate significance in terms of two-tailed p-values at the 5%, 1% and 0.1% levels*

We thus replicate the majority finding from earlier literature, and conclude for the benchmark case that deadlines, if they are present or shorter, make choice riskier.[21] Our evidence does not confirm findings of Dror et al (1999) that the relationship between deadline and choice can be inverted by a change in the level of risk (see also Busemeyer 1985). We shall hence only present regression results based on data pooled across the **L** and **H** sub-conditions in the rest of the analysis.[22] Overall, our findings militate against a static model as well as a sequential sampling model interpretation of the data in the benchmark case.

[21]Increased risk-seeking induced by deadlines does not necessarily indicate in our set up that decisions are "better" in any sense under time pressure. We approach the question of the quality of decision making from alternate perspectives in Section 5.

[22]Results are not necessarily identical across sub-conditions, but invertions were never detected for any case.

An assumption in much of the literature (for the benchmark case) is that the relationship between time pressure and choice is mediated by cognitive load. In Saqib and Chan (2015) and Kirchler et al (2017) for instance, time pressure is analyzed in a dual process framework, with cognitive load triggering a move from system 2 to system 1 thinking, the latter modelled prospect theoretically (see also Kahneman 2011). We now proceed toward our external validity concerns, to understand if time pressure continues to impose cognitive load for experienced subjects, especially when they face individually optimal deadlines. Much of the rest of our analysis concerns choices in episode 2. We first present summary statistics in this regard. Table 10 below shows average allocation to the safe asset in episode 2.

Table 10: Average allocation to safe asset - episode 2

	2C	**2N**	**2X**
L	45.51	44.53	39.59
H	51.9	53.31	50.44

90 observations in each condition (45 in each of **L** and **H**)

4.1 Choices of experienced subjects

We now analyze choices of experienced subjects, i.e., in episode 2 of conditions **2C**, **2X** and **2N**. The question is whether choice is affected by 'temporal aspects' (Spiliopoulos and Ortmann 2018), i.e., episode 2 deadline and episode 1 RT/NCT.[23] Subjects in **2C**, who do not confront a deadline, should face no cognitive load. We therefore expect their episode 2 choices to be consistent with static models such as the rational benchmark and hence unaffected by their episode 1 RTs. On the other hand, in a regression model where average allocation in episode 2 is explained by risk level, gender, income, and average allocation and average RT in episode 1, a significant coefficient estimate for average episode 1 RT would be supportive of process models which correlate choice and RT, e.g., Busemeyer and Townsend (1993). OLS output

[23]The latter because evidence from Section 3.3 indicates that NCT in the absence of a deadline is strongly correlated across episodes.

from the specification above is reported in column 1 of Table 11. The coefficient estimate is insignificant, as hypothesized (p-value = 0.373).

Table 11: Choices of experienced subjects

	Dep var: avg episode 2 allocation to safe asset				
	2C	**2C** vs **2X**	**2X**	**2C** vs **2N**	**2N**
cond dum	-	-5.006*	-	-0.398	-
	-	(2.271)	-	(2.214)	-
deadline	-	-	-0.288	-	-0.008
(minutes)	-	-	(0.522)	-	(1.916)
avg alloc	0.61***	0.631***	0.618***	0.619***	0.607***
(episode 1)	(0.104)	(0.071)	(0.099)	(0.071)	(0.103)
avg RT	-0.095	0.065	0.321*	-0.069	-0.068
(episode 1)	(0.107)	(0.087)	(0.147)	(0.056)	(0.264)
risk level	2.605	4.709*	7.593*	4.358	5.569
	(2.958)	(2.251)	(3.422)	(2.226)	(3.62)
no. of probs	-	-1.092	-0.554	5.332*	5.486
	-	(1.072)	(1.195)	(2.684)	(2.924)
gender	4.515	3.43	3.186	3.892	3.304
	(3.275)	(2.401)	(3.611)	(2.324)	(3.541)
income	1.223	0.725	-0.079	-0.159	-1.884
	(1.027)	(0.794)	(1.223)	(0.818)	(1.368)
constant	8.607	27	7.281	-93.776	-91.566
	(7.089)	(22.112)	(24.345)	(54.666)	(57.849)
Adj. R^2	0.373	0.384	0.392	0.362	0.344
No. of obs.	90	180	90	180	90

*Standard errors are in parentheses. Condition dummy is 0 if **2C**, 1 if **2X** or **2N**. Risk dummy is 0 if **L**, 1 for **H**. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. * and *** respectively indicate significance in terms of two-tailed p-values at the 5% and 0.1% levels*

The presence of a deadline in **2N** and **2X** can induce cognitive load. The validity of our procedure for indirect self-selection however implies that an endogenous deadline may fail to generate such load. OLS results in this regard for condition **2N** are presented in column 5 of Table 11. The specification is the same as in column 1, except for the presence of two additional independent variables, the length of the deadline faced, and the number of problems completed within it. We expect consistency with pseudo-optimal deadlines being approximately optimal on average, i.e., deadline and average episode 1 RT (which are highly correlated here) to produce insignificant coefficient estimates. The coefficient estimates are found to be insignificant, as hypothesized (p-values are 0.997 and 0.797 respectively).

Corresponding hypotheses for **2X** are unclear a priori. Experience may make a subject behave more in accordance with the rational benchmark, but exogeneity of the deadline may impose immoderate cognitive load. OLS results in this regard, using the same specification as in column 5, are presented in column 3. Deadline and average episode 1 RT are orthogonal here. The coefficient on deadline is insignificant (p-value = 0.583). This finding differs from the corresponding one in the benchmark case (compare with Table 8), providing relevance to our external validity concern with respect to experience. The coefficient on average episode 1 RT is significant, allowing for consistency with some process models of decision making. The finding in this case differs from that for **2N**, providing relevance to our external validity concern with respect to self-selection. The coefficient has a positive sign, indicating that experienced subjects with longer NCTs make safer choices given exogenous deadline.

Determination of whether the presence of deadlines affects decisions involves comparisons of episode 2 choices across **2C**, and **2X** or **2N**. OLS regression results in this regard are presented in the columns 2 and 4 of Table 11, for **2X** and **2N** respectively. In either case, the specification is the same as in column 1, except for the presence of two additional independent variables, the number of problems completed, and a condition dummy, the main independent variable, which takes value 0 for condition **2C**, and 1 otherwise.

Results in columns 1 and 5 lead us to expect an insignificant coeffient estimate on the dummy for **2N**. The hypothesis is confirmed by results in column 4 (p-value = 0.858). The coefficient on the dummy is significant and negative for **2X** (column 2). The presence of an exogenous deadline hence causes choices of experienced subjects to become riskier, a finding parallel to that for inexperienced subjects under exogenous deadlines in the benchmark case, though the effect appears weaker here (compare with Table 9). Differences in results across columns 2 and 4 indicate that the effect of the presence of deadlines on choice depends on whether these are endogenous or exogenous. Overall, our results prevent us from ruling out the presence of cognitive load amongst subjects in **2X**.

4.1.1 Optimality of endogenous deadlines I

A central construct in our approach is that an experienced subject's NCT is her optimal deadline, as both equal her MPD. Results for **2N** (column 5 of Table 11), where experienced subjects face such endogenous deadlines, showing that choice is independent of deadline and average episode 1 RT, are consistent with this construct being valid.

If a longer deadline involves no cost, then optimality of the MPD should imply optimality of a deadline longer than the minimum preferred one. Indifference across all deadlines at least the minimum preferred one in turn rests on a subject's behavior when she faces her MPD being equivalent to her behavior when she faces a longer deadline. The validity of our construct thus implies that behavior of subjects whose deadlines are respectively greater and lesser than their NCTs should differ.

We can use data from **2X** to assess this implication, as any subject there faces a deadline randomly assigned to her without any reference to her estimated NCT. We split the sample and designate those subjects with deadline more than NCT (superoptimal deadline) as facing weak time pressure (**W**), and the remaining subjects, with deadline less than NCT (suboptimal deadline), as those facing severe time pressure (**S**).[24] The first set of results in the comparison across the **W** and **S** groups, based on analysis of choice, is presented in this section. The remainder, based on analysis of task completion, is in Section 5.1.1. Some comparisons are in favor of the construct, though overall support for it is mixed.

Analysis of episode 2 choice in **2X** is presented in column 3 of Table 12 below. The specification is identical to that presented above in column 3 of Table 11, except for the presence of a time pressure dummy, taking value 0 for **W**, and 1 for **S**. Results pertaining to deadline and average episode 1 RT are identical in the current specification (p-value for deadline = 0.314) as in the previous one. The time pressure dummy is insignificant (p-value = 0.376); hence there is

[24] There were no subjects with estimated NCT equal to deadline, both calculated in seconds.

no evidence that choice on average differs across the **W** and **S** subjects.

Table 12: Time pressure and choice: condition **2X**

	Dep var: avg episode 2 allocation to safe asset				
	W	**S**	pooled **2X**	**2C** vs **W**	**2C** vs **S**
time pressure	-	-	-4.689	-	-
	-	-	(5.265)	-	-
deadline	-0.521	-1.737	-0.733	-	-
(minutes)	(0.972)	(1.704)	(0.723)	-	-
cond dum	-	-	-	-7.013*	-3.964
	-	-	-	(3.017)	(2.599)
avg alloc	0.692**	0.608***	0.608***	0.652***	0.625***
(episode 1)	(0.192)	(0.124)	(0.1)	(0.089)	(0.079)
avg RT	0.134	0.568*	0.412*	-0.045	0.029
(episode 1)	(0.346)	(0.241)	(0.179)	(0.102)	(0.093)
risk level	8.58	9.882*	8.38*	3.39	4.38
	(7.204)	(4.205)	(3.539)	(2.721)	(2.428)
no. of probs	-2.742	0.25	-0.532	-4.409	-0.855
	(5.146)	(1.259)	(1.197)	(3.819)	(1.105)
gender	-5.123	7.425	3.215	1.651	5.268
	(6.973)	(4.445)	(3.616)	(2.89)	(2.626)
income	-1.205	0.114	-0.211	0.999	0.896
	(3.02)	(1.388)	(1.233)	(0.986)	(0.827)
constant	71.163	-9.531	10.107	98.219	20.228
	(96.489)	(26.349)	(24.582)	(75.526)	(23.015)
Adj. R^2	0.22	0.46	0.391	0.36	0.395
No. of obs.	35	55	90	125	145

Standard errors are in parentheses. Time pressure dummy is 0 if ***W****, 1 for* ***S****. Condition dummy is 0 for* ***2C****, 1 otherwise. Risk dummy is 0 if* ***L****, 1 for* ***H****. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. *, ** and *** respectively indicate significance in terms of two-tailed p-values at the 5%, 1% and 0.1% levels*

Columns 1 and 2 present results from the analysis of choice in **2X**, respectively restricting attention to **W** and **S** subjects. The specifications are identical and equivalent to that in column 3 of Table 11. Choice is independent of deadline for both groups, as in the pooled sample (p-values: **W** = 0.596, **S** = 0.313). The groups differ in terms of the impact of NCT. There is no dependence for the **W** subjects (p-value = 0.702), whereas the coefficient on average episode 1 RT is positive and significant for the **S** subjects, as for the pooled sample. This is expected under the assumption that the **W** subjects face optimal deadlines, whereas the **S** subjects face suboptimal ones. This comparison thus provides evidence supporting the construct.

Next we compare choices of the **W** and **S** groups separately with those in **2C**. These results are presented in columns 4 and 5 respectively. The specifications are identical and equivalent to that in column 2 of Table 11. The results provide some evidence against our construct, since the condition dummy is negative and significant in the comparison involving **W** subjects, as in the pooled sample, but not in the one involving **S** subjects (two-sided p-value = 0.13). Together, the findings above show that behavior of **W** subjects is similar to that of subjects in **2N**, and dissimilar to that of **S** subjects, as far as sensitivity of choice to length of deadline and NCT are concerned, while behavior of **S** subjects is similar to that of subjects in **2C**, and dissimilar to that of **W** subjects, as far as choice on average is concerned.

4.2 Validity concerns

We now investigate our validity concerns, and compare across conditions with deadlines. We compare episode 1 choice in condition **1X** with episode 2 choice in condition **2X** to address the issue of experience. OLS regression results in this regard are presented in the first cloumn of Table 13 below. Examination of the effect of allowing opportunities for self-selection is conducted by proxy, by comparing choices under exogenous and endogenous deadlines of experienced subjects, i.e., by comparing episode 2 choices across **2X** and **2N**. OLS regression results in this regard are presented in the second cloumn of the table. We also study the combined effect of experience and self-selection by comparing episode 1 choice in **1X** with episode 2 choice in **2N**. OLS regression results in this regard are presented in the third column of the table. We note that the legitimacy of condition comparisons involving **1X** rests on stability of choice with respect to task 2 across episodes. Results presented above in Section 3.1 (Table 4), establishing equivalence of tasks 1 and 2, and below in Section 4.3 (the first column of Table 14), establishing stability of choice across episodes, give grounds for such as assumption.

The dependent variable is average allocation to the safe asset. The main independent variable is a condition dummy, which takes values 0 for **2X**, and 1 for **1X** for the first column, 0 for **2N**, and 1 for **2X** for the second column, and 0 for **2N**, and 1 for **1X** for the third column. Our null hypotheses based on findings reported earlier is that this variable is negative and significant, i.e., choice becomes safer as we move from **1X** to **2X** to **2N**. Other independent variables common across the three cases are the length of deadline faced, the level of risk, the number of problems completed within the deadline, gender, and income. Additional independent variables for the second column are average allocation in episode 1, and average RT in episode 1.

Table 13: Examination of external validity concerns

	Dep var: average allocation to safe asset		
	2X2 vs **1X1**	**2N2** vs **2X2**	**2N2** vs **1X1**
cond dum	-4.849†	-3.324	-7.127*
	(2.9)	(2.394)	(2.915)
deadline	0.55*	-0.581	0.39†
(minutes)	(0.22)	(0.483)	(0.222)
avg alloc	-	0.631***	-
(episode 1)	-	(0.071)	-
avg RT	-	0.085	-
(episode 1)	-	(0.085)	-
risk level	11.396***	6.263*	10.514***
	(2.511)	(2.424)	(2.451)
no. of probs	-1.045	0.081	-0.193
	(0.587)	(1.087)	(0.638)
gender	8.196**	2.916	6.341*
	(2.662)	(2.501)	(2.676)
income	-1.543	-0.628	-2.465**
	(0.816)	(0.899)	(0.834)
constant	18.625	9.436	42.61**
	(17.876)	(21.926)	(13.136)
Adj. R^2	0.19	0.361	0.194
No. of obs.	180		

*Standard errors are in parentheses. Condition dummy for column **i** vs **j** is 0 for **i**, 1 for **j**. Risk dummy is 0 if **L**, 1 for **H**. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. †, *, ** and *** respectively indicate significance in terms of two-tailed p-values at the 10%, 5%, 1% and 0.1% levels*

The sign of the coefficient on the condition dummy is negative in all cases, as conjectured. The coefficient is insignificant for the second column (p-value = 0.167), suggesting the issue of self-selection may be insubstantial, given experienced subjects. This is at odds with findings in Section 4.1 above, which suggested behavioral difference across **2X** and **2N**. Overall, we conclude that that concern surrounding self-selection is of some, though limited, importance. The coefficient for the first column is significant, though only at the 10% level. This suggests the issue of experience may be of muted importance, given exogeneity of deadlines. The coefficient is significant for the third column. Findings so far hence indicate that both external validity concerns have relevance, though weakly in isolation, while the combined effect of providing experience and allowing indirect self-selection is significant, and is to produce safer choice.

The effect of variation in the length of deadline faced on choice is noteworthy. The variable is insignificant for the second column (p-value = 0.231), suggesting that choices of experienced subjects are immune to the extent of the deadline. This coincides with results presented above in Section 4.1. The variable is significant for the first column, and also for the third, though at the 10% level for the latter. Such behavior appears to be driven in either case by the presence of inexperienced subjects.

4.3 Intertemporal stability

An important assumption in any theory of decision under risk that aims to make predictions or offer policy prescriptions is the stability of risk preference over time. Are choices for risky problems stable over time? Studies starting with McGlothlin (1956), have mostly found that aggregate choice is intertemporally stable. Experimental analyses (e.g., Wehrung et al 1984, Harrison et al 2005, Zeisberger et al 2012) compare risky choices of subjects across two episodes, separated by at least a month, in the absence of any deadline.[25] Our design, recording choices in two episodes for conditions **2C**, **2N** and **2X**, with an inter-exposure interval of approximately one month, hence permits an examination of intertemporal stability. The benchmark case, corresponding to the environment considered in prior literature, is given by **2C**, which does not confront subjects with deadlines.

[25] Stability of choice in this sense is thus equivalent to invariance of choice to experience.

We analyze, using data from condition **2C**, whether the aggregate intertemporal stability finding from earlier studies is replicated in our setting. OLS regression results in this regard are presented in the first column of Table 14. The dependent variable is average allocation to the safe asset. There were two observations per subject, one per episode. We used subject specific dummy variables to control for subject fixed effects (subject specific dummy variables are not reported for brevity), and cluster correction at the level of subject to allow adjustment of the variance-covariance matrix to the lack of independence of observations.

The main independent variable is a dummy, taking value 0 for episode 1, and 1 for episode 2. The null hypothesis, based on prior findings, is that this variable has an insignificant coefficient estimate. Other independent variables are average RT for episode 1, risk level, gender, and income. We find that the coefficient on the episode dummy is insignificant (p-value = 0.674). The null is thus confirmed, and the aggregate intertemporal stability finding replicated.

Our contribution to this research lies in asking whether aggregate intertemporal stability is robust to the presence of a deadline. Adopting the perspective of expected utility theory and interpreting a deadline as a source of choice-irrelevant information argues in favor of robustness. The question may be of importance as risky decision in real settings often have to be taken in the presence of relevant as well as apparently irrelevant information.

We use data from **2N** and **2X** to address this question. OLS regression results in this regard are presented respectively in the second and third columns of Table 14. The dependent variable and specifications are identical to that for the first column, except that two additional independent variables are introduced for both the second and the third columns. These are the length of the deadline faced, and the number of problems completed within the deadline. The specification for **2X**, in the third column, also includes the time pressure dummy (0 if **W**, 1 if **S**) as an independent variable.

The finding of stability from condition **2C** and results reported earlier in Section 4.1 lead us to posit stability for **2N**, where subjects faced endogenous deadlines, and instability for **2X**, where subjects faced exogenous deadlines. For the latter case, we expect choice to become

Table 14: Intertemporal stability: comparing choice across episodes

	Dep var: avg allocation to safe asset		
	2C	**2N**	**2X**
episode	-0.911	-3.065	-6.183*
	(2.162)	(2.498)	(2.536)
deadline	-	-4.358***	-1.299***
(minutes)	-	($6.38e^{-13}$)	($4.16e^{-14}$)
time pressure	-	-	-18.832***
	-	-	($3.83e^{-13}$)
avg RT	-0.264***	0.57***	0.529***
(ep 1)	($1.37e^{-14}$)	($1.01e^{-13}$)	($7.14e^{-15}$)
risk level	-1.426***	10.789***	22.925***
	($6e^{-13}$)	($1.19e^{-12}$)	($4.98e^{-13}$)
no. of probs.	-	1.363***	-1.154***
	-	($5.83e^{-13}$)	($4.51e^{-14}$)
gender	-5.812***	-2.672***	20.454***
	($6.07e^{-13}$)	($1.37e^{-12}$)	($4.32e^{-13}$)
income	-0.715***	18.205***	3.398***
	($1.26e^{-13}$)	($8.51e^{-13}$)	($8.7e^{-14}$)
constant	80.544***	-35.815***	24.358***
	(3.242)	(3.747)	(3.803)
R^2	0.805	0.785	0.803
No. of obs.	180		
No. of clusters	90		

*Subject dummies not reported. Robust standard errors are in parentheses. Episode dummy is 0 for episode 1, 1 for episode 2. Time pressure dummy is 0 if **W**, 1 for **S**. Risk dummy is 0 if **L**, 1 for **H**. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. * and *** respectively indicate significance in terms of two-tailed p-values at the 5% and 0.1% levels* riskier in the second episode. The null hypotheses for the episode dummy are thus insignificance for **2N**, and negativity and significance for **2X**. Results support the null hypotheses in either case (p-value for **2N** = 0.223). These findings reiterate the importance of our primary external validity concern for risky decision under deadline with respect to self-selection.

5 Analyses of task completion and response time

Given the payment scheme, a subject who took a decision for a problem within the deadline, if any, was guaranteed a positive payoff, while a problem left undecided at the expiration of the constraint, if any, yielded 0. Non-decision is hence inefficient in our environment, with inefficiency increasing in the number of undecided problems. Measurement of this inefficiency is tantamount to measurement of the shadow value of a deadline, if a deadline is viewed as a constraint on decision making. Results pertaining to this issue are presented in Section 5.1.

It is known that imposition of deadlines may cause acceleration in decision making (Ben-Zur and Bresnitz 1981), which may be a coping strategy under time pressure involving accelerated processing of information (Miller 1960). Some have argued on this basis that deadlines can impair decision, if there is a speed/accurary tradeoff under time pressure (Swensson and Thomas 1974). Others have argued that deadlines may force reallocation of cognitive resources, leading to simultaneous improvement in speed and accuracy (Beach and Mitchell 1978). Even if deadlines impair decision, accelerated choice may release time for other decision problems. The impact of deadlines on efficiency, from the angle of decision speed, is thus ex ante unclear. We present an analysis of the effect of deadlines on response time in Section 5.2.

5.1 Inefficiency

One can use the non-completion rate, i.e., the proportion of subjects not completing the task within the deadline, as a measure of inefficiency. Table 15 gives the number of subjects who did not complete the task, in each case out of 45. These are given in the first three columns. Classical proportions tests showed the non-completion rate was positive for every sub-condition.

Inefficiency can also be measured by the average number of problems completed within the deadline, out of 20. The middle three columns of Table 15 give the means for all subjects, and the last three columns of the table give the means only for subjects not completing the task. For the average number of problems completed by all subjects, a) two-tailed t-tests and Snedecor-Cochran tests indicated that it was less than 20 for **1XH**, **1XL** and **2XH** at the 5% level, b) one-tailed tests gave ambivalent answers for **2XL** and **2NH**, with t-tests yielding significant difference at the 5% level, and Snedecor-Cochran yielding such difference only at the 10% level, and c) it equaled 20 for **2NL** by all tests at all conventional levels of significance. Inefficiency thus appears to be highest in **1X**, followed by **2X**, and least in **2N**.

Table 15: Inefficiency

	Non-completion			Mean number of problems, subjects:					
	No. of subs			All			Non-completing		
	2N	**2X**	**1X**	**2N**	**2X**	**1X**	**2N**	**2X**	**1X**
L	1	4	11	19.91	19.67	18.8	16	16.25	15.09
H	4	7	12	19.84	19.42	18.67	18.25	16.29	15

We can analyze inefficiency within condition. We execute probit and OLS regressions for each of conditions **2N**, **2X** and **1X**. The probit dependent variable is a dummy taking value 1 if a subject completed the task within the deadline, and 0 if not. The OLS dependent variable is the number of problems completed within the deadline. The specifications were the same for the probit and OLS regressions for any condition. Independent variables common to all conditions were the length of deadline faced, the level of risk, gender and income. Average allocation to the safe asset and average per problem RT, both for episode 1, were additional independent variables for **2N** and **2X**. The specification for **2X** also included the time pressure dummy as an independent variable. The top and bottom panels of Table 16 below respectively show the probit and OLS output.[26] Given our finding above that inefficiency mostly exists in the current

[26]Logit regressions equivalent to probit ones were conducted for all reported probit regressions. These gave comparable results, and we only report the probit output.

environment, and viewing a deadline as a constraint on decision making, the null hypotheses are that the coefficient on the deadline variable should be positive and significant in all cases.

There is a strong link between the length of the deadline faced and inefficiency for **1X**, with longer deadlines reducing inefficiency. A similar, though weaker (with significance only at the 10% level), link is present for **2X** and **2N**. RT/NCT also appears to have some explanatory power for these two conditions, with significance in the probit regressions (p-values for OLS: **2X** = 0.142, **2N** = 0.138). It has a negative sign, implying subjects with longer NCT tend toward greater inefficiency. These results reinforce the hint that the impact of deadlines on inefficiency is highest in **1X**, confirming our validity concern with respect to experience.

We can also analyze inefficiency between conditions. We execute probit and OLS regressions, with the same dependent variables as in the within condition regressions above. The top and bottom panels of Table 17 below respectively show the probit and OLS output. A condition dummy is the main independent variable in every case, and takes values 0 for **2X**, and 1 for **1X** for the first column, 0 for **2N**, and 1 for **2X** for the second column, and 0 for **2N**, and 1 for **1X** for the third column. Our null hypotheses are that this variable is negative and significant, i.e., inefficiency is reduced as we move from **1X** to **2X** to **2N**. Other common independent variables are deadline, risk level, gender, and income. We retain the null hypotheses that the coefficient on the deadline variable is positive and significant in all cases. The specifications for the comparison between **2N** and **2X** also use average allocation and average RT, both from episode 1, as independent variables. We note that the legitimacy of condition comparisons involving **1X** rests on stability of task completion across episodes. Our design does not allow any independent assessment on this matter. Additionally, while choice is stable across episodes in the absence of any deadline (see the first column of Table 14), response time is not (Table 6). We therefore draw conclusions from results presented in the first and third columns with caution.

Deadline appears significant for all regressions, and tighter deadlines produce greater inefficiency. Combined with results from Table 16, we conclude there is support for the view that a deadline is acting as a binding constraint on decision making. Our null hypotheses with re-

Table 16: Analysis of inefficiency: within condition

	Probit dep. var: 1 if task done		
	2N	**2X**	**1X**
deadline	0.858†	0.31†	0.178***
(minutes)	(0.471)	(0.158)	(0.045)
time pressure	-	0.37	-
	-	(0.847)	-
avg alloc	0.015	-0.011	-
(episode 1)	(0.017)	(0.013)	-
avg RT	-0.087*	-0.057*	-
(episode 1)	(0.042)	(0.028)	-
risk level	-0.032	-0.269	-0.079
	(0.702)	(0.453)	(0.336)
gender	0.141	0.97*	-0.192
	(0.539)	(0.43)	(0.356)
income	-0.046	0.153	0.031
	(0.258)	(0.149)	(0.094)
constant	-1.046	-0.029	-0.583
	(2.145)	(1.25)	(0.697)
Pseudo R^2	0.132	0.25	0.266
	OLS dep. var: # of probs. done		
	2N	**2X**	**1X**
deadline	0.119†	0.111†	0.147***
(minutes)	(0.071)	(0.066)	(0.038)
time pressure	-	0.091	-
	-	(0.486)	-
avg alloc	0.004	-0.01	-
(episode 1)	(0.004)	(0.009)	-
avg RT	-0.015	-0.024	-
(episode 1)	(0.01)	(0.016)	-
risk level	-0.008	-0.166	-0.203
	(0.136)	(0.326)	(0.567)
gender	-0.011	0.752*	-0.214
	(0.133)	(0.323)	(0.643)
income	-0.007	0.076	0.373*
	(0.051)	(0.114)	(0.164)
constant	19.482***	18.699***	16.247***
	(0.378)	(0.939)	(1.278)
Adj. R^2	0.156	0.063	0.142

*No. of observations = 90. Standard errors are in parentheses. Risk dummy is 0 if **L**, 1 for **H**. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. †, * and *** respectively indicate significance in terms of two-tailed p-values at the 10%, 5% and 0.1% levels*

Table 17: Analysis of inefficiency: between conditions

	Probit dep. var: 1 if task done		
	2X2 vs **1X1**	**2N2** vs **2X2**	**2N2** vs **1X1**
cond dum	-1.051***	-0.236	-1.527***
	(0.265)	(0.332)	(0.304)
deadline	0.165***	0.32**	0.172***
(minutes)	(0.039)	(0.107)	(0.042)
avg alloc	-	-0.002	-
(episode 1)	-	(0.01)	-
avg RT	-	-0.04**	-
(episode 1)	-	(0.015)	-
risk level	-0.175	-0.201	-0.232
	(0.246)	(0.329)	(0.273)
gender	0.236	0.622*	0.006
	(0.251)	(0.31)	(0.278)
income	0.052	0.076	0.044
	(0.075)	(0.114)	(0.085)
constant	-6.451***	0.125	0.698
	(1.636)	(0.879)	(0.59)
Pseudo R^2	0.222	0.217	0.29
	OLS dep. var: # of probs. done		
	2X2 vs **1X1**	**2N2** vs **2X2**	**2N2** vs **1X1**
cond dum	-1.579***	-0.352*	-1.824***
	(0.355)	(0.166)	(0.317)
deadline	0.124***	0.104**	0.122***
(minutes)	(0.027)	(0.033)	(0.025)
avg alloc	-	-0.003	-
(episode 1)	-	(0.005)	-
avg RT	-	-0.015**	-
(episode 1)	-	(0.006)	-
risk level	-0.256	-0.036	-0.077
	(0.323)	(0.17)	(0.291)
gender	0.321	0.375*	-0.087
	(0.343)	(0.173)	(0.318)
income	0.253*	0.035	0.241*
	(0.103)	(0.063)	(0.097)
constant	8.163***	19.208***	18.543***
	(2.222)	(0.472)	(0.678)
Adj. R^2	0.141	0.067	0.18

*No. of observations = 180. Standard errors are in parentheses. Condition dummy for column **i** vs **j** is 0 for **i**, 1 for **j**. Risk dummy is 0 if **L**, 1 for **H**. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. *, ** and *** respectively indicate significance in terms of two-tailed p-values at the 5%, 1% and 0.1% levels* spect to the condition dummy also largely fail to get rejected (except for the probit comparison between **2X** and **2N**; two-sided p-value = 0.477), suggesting that provision of a) experience and b) opportunities to avail of self-selection, under deadline, can not only stabilize choice, as seen in Section 4, but also reduce inefficiency.

5.1.1 Optimality of endogenous deadlines II

We continue the inquiry inititated in Section 4.1.1 on the validity of the MPD construct. OLS regression results are presented below in Table 18 analyzing the number of problems completed within the deadline in **2X** for episode 2, disaggregated by the **W** and **S** groups. The dependent variable is the number of problems completed and the specifications are the same as for column 2 of Table 16, except that the time pressure dummy does not appear. The first and second columns respectively give results restricting attention to **W** and **S** subjects.[27]

We saw from column 2 of Table 16 above that there was a tendency for efficiency of subjects in **2X** to be dependent on NCT and the length of the deadline faced. This pattern in replicated in the **S** group, where efficiency is seen to be significantly dependent on both deadline and average episode 1 RT. No such pattern emerges for the **W** group (p-values for deadline and RT are respectively 0.217 and 0.498). These results are consistent with **S** subjects facing cognitive load due to severe time pressure engendered by suboptimal deadlines, and **W** subjects under weak time pressure and superoptimal deadlines facing insignificant cognitive load. This comparison is thus supportive of the MPD construct, suggesting, in conjunction with results from Section 4.1, that endogenous or pseudo-optimal deadlines can be considered as approximately optimal.

5.2 Acceleration

Much of the analysis in this section concerns RT in episode 2. Table 19 below shows average RT per completed problem in seconds in episode 2.

[27]Corresponding probit output is not presented as the model was not estimable for the sample of **W** subjects, due to collinearities between the inefficiency dummy (dependent variable) and the risk, gender and income indicators.

Table 18: Time pressure and inefficiency: **W** and **S** groups

	Dep. var: # of probs.	
	W	**S**
deadline	0.044	0.46*
(minutes)	(0.035)	(0.184)
avg alloc	0.009	-0.024
(episode 1)	(0.007)	(0.014)
avg RT	-0.009	-0.058*
(episode 1)	(0.013)	(0.026)
risk level	-0.091	-0.087
	(0.264)	(0.482)
gender	0.265	1.117*
	(0.251)	(0.483)
income	0.175	-0.06
	(0.106)	(0.159)
constant	18.372***	18.865***
	(0.708)	(1.309)
Adj. R^2	0.047	0.135
No. of obs.	35	55

*Standard errors are in parentheses. Risk dummy is 0 if **L**, 1 for **H**. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. * and *** respectively indicate significance in terms of two-tailed p-values at the 5% and 0.1% levels*

Table 19: Mean RT in secs - episode 2

	2C	**2N**	**2X**
L	18.45	10.74	13.71
H	18.68	11.3	11.9

90 observations in each condition, 45 in each sub-condition

We first determine if acceleration can be detected in our environment, by studying whether RT falls due to the imposition of deadlines. This involves comparing each deadline condition with its corresponding control. OLS regression output in this regard, with average RT as the dependent variable, is presented below in the first three columns of Table 20. For **1X**, in the first column, we compare with **1C**, **2C**, **2N** and **2X**, pooled (for episode 1), while for **2X** and **2N**, in the second and third columns respectively, we compare with **2C** (for episode 2). The main independent variable is a condition dummy which in each case takes value 0 for the control, and 1 for the treatment. The null hypothesis is of acceleration, i.e., a negative and significant coefficient on this dummy. Other independent variables, common to all three, are the level of risk, the number of problems completed within the deadline, gender and income. Average allocation to the safe asset and average RT, both for episode 1, also appear as independent variables for the comparisons with respect to **2X** and **2N**. We find confirmation of the null hypothesis in every case. The extant result is thus replicated in our sample, and we conclude that the presence of deadlines causes acceleration in decision making, irrespective of whether deadlines are endogenous or exogenous, and whether subjects are experienced or inexperienced.

We can also conduct comparisons across conditions involving deadlines. However, the legitimacy of any such comparison involving **1X** would have to rest on the stability of RT across episodes in the absence of any deadline, which is negated by the results of Section 3.2 (Table 6). We therefore only consider the comparison across **2N** and **2X**.

The presence of acceleration in our environment prompts the question of efficiency. Deadlines in our design are not endogenized through opportunity cost tradeoffs. We cannot therefore conduct an examination of efficiency enhancing impacts of a deadline arising due to decision time being released because of accelerated choice. The question remains as to whether acceleration is efficiency impairing in our case because of lower quality of decision making due to the speed/accuracy tradeoff, or whether it is efficiency enhancing due to higher quality of decision making because of refocused cognitive application. The comparison between **2N** and **2X** allows an approach in this regard.

Table 20: Acceleration

	Dep var: avg RT per problem			
	1X1	**2X2**	**2N2**	**2N2** vs **2X2**
	vs control			
cond dum	-14.082***	-6.331***	-8.846***	1.574*
	(2.262)	(1.063)	(0.924)	(0.647)
deadline	-	-	-	1.34***
(minutes)	-	-	-	(0.131)
avg alloc	-	0.028	0.046	0.037
(episode 1)	-	(0.033)	(0.03)	(0.019)
avg RT	-	0.292***	0.3***	0.079**
(episode 1)	-	(0.041)	(0.023)	(0.023)
risk level	1.23	-1.271	-0.522	-0.324
	(1.714)	(1.054)	(0.929)	(0.655)
no. of probs	-0.792	-0.1	-2.275*	-1.132***
	(0.583)	(0.502)	(1.12)	(0.294)
gender	1.968	-1.608	-0.292	-0.749
	(1.765)	(1.124)	(0.97)	(0.676)
income	0.568	-0.145	0.357	0.043
	(0.585)	(0.372)	(0.341)	(0.243)
constant	41.021***	14.236	70.125**	21.968***
	(11.812)	(10.353)	(22.807)	(5.926)
Adj. R^2	0.073	0.306	0.566	0.638
No. of obs.	450	180		

*Standard errors are in parentheses. For the first column, condition dummy is 1 if **1X**, and 0 if **1C**, **2C**, **2N**, or **2X**. For the second (third) column, condition dummy is 0 for **2C** and 1 for **2X** (**2N**). For the last column, condition dummy is 0 for **2N** and 1 for **2X**. Gender is 0 for male, 1 for female. Income takes values 1 through 6, higher incomes taking higher values. *, ** and *** respectively indicate significance in terms of two-tailed p-values at the 5%, 1% and 0.1% levels*

Results reported above have shown that choice is invariant across the conditions (column 2 of Table 13), while efficiency as measured by the number of problems completed is higher in **2N** (column 2 of the bottom panel of Table 17). These results suggest that experienced subjects find it easier to cope with time pressure under endogenous compared to exogenous deadlines, and hence lead to the expectation that efficiency of decision making in general should be higher in **2N**. Under the interpretation of acceleration as a coping strategy under time pressure, we should thus expect greater acceleration in **2N** if it induces better decisions, and lesser if there is a speed/accuracy trade-off.

The fourth column of Table 20 presents OLS regression results in this regard. The specification is exactly the same as those used for the second and third columns, except that the length of deadline faced appears as an additional independent variable. We retain the null hypotheses that the coefficient on this variable is positive and significant. The main independent variable is a condition dummy, which takes value 0 for **2N**, and 1 for **2X**. We find it is significant and positive, i.e., there is greater acceleration in **2N**. We interpret this finding as suggesting that acceleration under deadlines is associated with quality enhacement, rather than accuracy impairment. It buttresses the centrality of our main external validity concern with respect to self-selection. The coefficient on deadline is also positive and significant, as expected, reinforcing the idea of deadline as a decision constraint.

6 Summary and conclusion

Prior research has found that variation in the length of deadlines, or their presence, can affect risky choice. Such a finding is consistent with time pressure imposing cognitive load, and contradicts the predictions of static models. These findings have been obtained in environments where subjects are inexperienced and deadlines are exogenous. In this paper, we study whether such results are affected if subject experience and self-selection are allowed. We manipulate experience through pre-exposure, while the approach to self-selection is indirect and involves comparing effects of endogenous and exogenous deadlines.

We find, as in the existing literature, that choices of inexperienced subjects can be impacted by exogenous deadlines. At the opposite end, we find that choices of experienced subjects may be immune to deadlines if they are endogenous. This suggests potential applicability of static models of decision making. Choices of experienced subjects are found to be dependent on temporal aspect however if deadlines are exogenous, suggesting potential applicability of process models of decision making. We additionally find that choices may be intertemporally stable on average in our sample, with aggregate intertemporal stability robust to the insertion of endogenous deadlines, but not exogenous ones. Our overall evidence indicates that experience as well as self-selection can be factors influencing choice, rendering external validity concerns in their respect relevant.[28]

Similar conclusions are yielded by an analysis of efficiency in terms of task completion, where we find that while deadlines can constrain decision, and paralyze some decision makers, inefficiency of experienced subjects facing endogenous deadlines may be close to negligible. Analysis of RT suggests that deadlines induce decision acceleration, with some evidence of acceleration resulting from improved cognitive focus. Systematic investigations of the efficiency implications of deadlines in terms of forcing non-completion, and enabling release of decision time and superior deployment of cognitive resources are left for future research.

References

[1] Arocha, J., Wang, D., and Patel, V. (2005): Identifying reasoning strategies in medical decision making: a methodological guide, *Journal of Biomedical Informatics*, 15, 54 - 171.

[28]Our results satisfy internal validity, as we only report findings based on data from all subjects. Regression analyses (not reported for brevity) further show that results are largely unaffected if we exclude non-completing subjects from the sample.

[2] Beach, L. and Mitchell, T. (1978): A contingency model for the selection of decision strategies, *Academy of Management Review*, 3, 439 - 49.

[3] Benson, L., and Beach, L. (1996): The effects of time constraints on the pre-choice screening of decision options, *Organizational Behavior and Human Decision Processes*, 67, 222 - 28.

[4] Ben-Zur, H., and Bresnitz, S. (1981): The effects of time-pressure on risky choice behavior, *Acta Psychologica*, 47, 89 - 104.

[5] Busemeyer, J. (1985): Decision making under uncertainty: a comparison of simple scalability, fixed-sample, and sequential sampling models, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 11, 538 - 64.

[6] Busemeyer, J., and Townsend, J. (1993): Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment, *Psychological Review*, 100, 432 - 59.

[7] Camerer, C. (1989): An experimental test of several generalized utility theories, *Journal of Risk and Uncertainty*, 2, 61 - 104.

[8] Dror, I., Busemeyer, J., and Basola, B. (1999): Decision making under time pressure: An independent test of sequential sampling models, *Memory and Cognition*, 27, 713 - 25.

[9] Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006): Costly information acquisition: Experimental analysis of a boundedly rational model, *American Economic Review*, 96, 1043 - 68.

[10] Goel, V., Makale, M., and Grafman, J. (2004): The hippocampal system mediates logical reasoning about familiar spatial environments, *Journal of Cognitive Neuroscience*, 16, 654 - 64.

[11] Goldstein, W., and Weber, E. (1996): Content and discontent: Indications and implications of domain specificity in preferential decision making, in Busemeyer, J., Medin, D., and Hastie R. (eds.), *The Psychology of Learning and Motivation: vol. 32 - Decision Making from a Cognitive Perspective.* New York: Academic Press.

[12] Harrison, G., Johnson, E., McInnes, M., and Rutstrom, E. (2005): Temporal stability of estimates of risk aversion, *Applied Financial Economics Letters*, 1, 31 - 5.

[13] Heston, A., Summers, R., and Aten, B. (2012): Penn world table version 7.1, Center for International Comparisons of Production, Income and Prices, University of Pennsylvania.

[14] Hu, Y., Wang, D., Pang, K., Xu, G., and Guo, J. (2015): The effect of emotion and time pressure on risk decision-making, *Journal of Risk Research*, 18, 637 - 50.

[15] Hussam, R., Porter, D., and Smith, V. (2008): Thar she blows: Can bubbles be rekindled with experienced subjects? *American Economic Review*, 98, 924 - 37.

[16] Kagel, J., and Levin, D. (1986): The winner's curse and public information in common value auctions, *American Economic Review*, 76, 894 - 920.

[17] Kahneman, D. (2011): *Thinking, fast and slow.* New York: Farrar, Straus and Giroux.

[18] Kirchler, M., Andersson, D., Bonn, C., Johannesson, M., Sørensen, E., Stefan, M., Tinghög, G., and Västfjäll, D. (2017): The effect of fast and slow decisions on risk taking, *Journal of Risk and Uncertainty*, 54, 37 - 59.

[19] Kocher, M., Pahlke, J., and Trautmann, S. (2013): Tempus fugit: Time pressure in risky decisions, *Management Science*, 59, 2380 - 91.

[20] Kocher, M., Schindler, D., Trautmann, S., and Xu, Y. (2019): Risk, time pressure, and selection effects, *Experimental Economics*, 22, 216 - 46.

[21] McGlothlin, W. (1956): Stability of choices among uncertain alternatives, *American Journal of Psychology*, 69, 604 - 15.

[22] Miller, J. (1960): Information input overload and psychopathology, *American Journal of Psychiatry*, 116, 695 - 704.

[23] Mosteller, F. and Nogee, P. (1951): An experimental measurement of utility, *Journal of Political Economy*, 59, 371 - 404.

[24] Nosofsky, R. and Palmeri, T. (1997): An exemplar-based random walk model of speeded classification, *Psychological Review*, 104, 266 - 300.

[25] Nursimulu, A., and Bossaerts, P. (2014): Risk and reward preferences under time pressure, *Review of Finance*, 18, 999 - 1022.

[26] Ordóñez, L., and Benson, L. (1997): Decisions under time pressure: How time constraint affects risky decision making, *Organizational Behavior and Human Decision Processes*, 71, 121 - 40.

[27] Rand, D., Greene, J., and Nowak, M. (2012): Spontaneous giving and calculated greed, *Nature*, 489, 427 - 30.

[28] Spiliopoulos, L., and Ortmann, A. (2018): The BCD of response time analysis in experimental economics, *Experimental Economics*, 21, 383 - 433.

[29] Stott, H. (2006): Cumulative prospect theory's functional menagerie, *Journal of Risk and Uncertainty*, 32, 101 - 30.

[30] Sutter, M., Kocher, M., and Strauß, S. (2003): Bargaining under time pressure in an experimental ultimatum game, *Economics Letters*, 81, 341 - 47.

[31] Svenson, O. and Maule, A. (1993): *Time pressure and stress in human judgement and decision making.* New York: Plenum.

[32] Swensson, R., and Thomas, R. (1974): Fixed and optimal stopping models for two choice discrimination times, *Journal of Mathematical Psychology*, 11, 213 - 36.

[33] Tinghög, G., Andersson,, D., Bonn, C., Böttiger, H., Josephson, C., Lundgren, G., Västfjäll, D., Kirchler, M., and Johannesson, M. (2013): Intuition and cooperation reconsidered, *Nature*, 498, E1 - E2.

[34] Wegier, P. and Spaniol, J. (2015): The effect of time pressure on risky financial decisions from description and decisions from experience, *PLoS One*, 10, e0123740.

[35] Wehrung, D., MacCrimmon, K., and Brothers, K. (1984): Utility assessment: domains, stability and equivalence procedures, *INFOR*, 22, 98 - 115.

[36] Young, D., Goodie, A., Hall, D., and Wu, E. (2012): Decision making under time pressure, modeled in a prospect theory framework, *Organizational Behavior and Human Decision Processes*, 118, 179 - 88.

[37] Zeisberger, S., Vrecko, D., and Langer, T. (2012): Measuring the time stability of prospect theory preferences, *Theory and Decision*, 72, 359 - 86.

Appendix A: Risky assets

Any problem faced by any subject required the allocation of a budget of 100 across two assets, one safe and one risky. The safe asset yielded 100 per unit of investment. The 3 possible outcomes and corresponding probabilities for the risky assets varied across problems. The tables below list the possible outcomes and corresponding probabilities for all risky assets used in the experiment. For each asset, outcomes and probabilities are arranged separately and vertically within the cell in the same order as seen by subjects (see Figures 1 and 2). Percentage signs indicate probabilities, as in the experimental interface.

There were four tasks (1L, 1H, 2L, 2H), each with 20 problems in a fixed sequence. Risky assets from these problems, in implemented sequences, are given in Tables 22 for tasks 1L and 1H, and 23 for tasks 2L and 2H. In either episode, before her task, each subject saw an example problem followed by two trial problems, the latter two in a fixed sequence. These were low variance problems in the **L** sub-conditions, and high variance ones in the **H** sub-conditions. The same problems were used in all **L** (**H**) sub-conditions, irrespective of episode or condition. Risky assets from these problems are listed in Table 21.

Table 21: Assets in example and trial problems

	Examples		Trials 1		Trials 2	
	47	18%	50	20%	39	18%
L	231.7	45%	232.9	20%	251.1	57%
	303.5	37%	280.7	60%	299.4	25%
	34	23%	50	25%	39	18%
H	110.1	22%	121.7	25%	184.4	57%
	350.8	55%	364.1	50%	451.6	25%

Appendix B: Instructions

Instructions are given below. After the instruction screens, subjects were shown the example. They faced a screen which first said: The tables below give the possible returns and the corresponding chances for both options. How much of your 100 units will you invest in option 1, i.e., how many units of option 1 will you buy (whatever remains will be used to buy units of option 2)?

The example problem followed, and the screen ended with: Remember, your answer here will be the number of units of option 1 you are buying out of 100, the remainder being the number of units of option 2 you are buying.

Then they saw a screen which showed based on that problem what the possible total returns (points) were, and with what chances, using as an example an allocation of 70 to the safe asset. This was followed by a screen which said: You can now practice using 2 trial situations, before

Table 22: Risky assets from problems in task 1 (mean = 225)

High risk (variance 20,000 approx.)						Low risk (variance 8000 approx)					
	50	15%		49	21%		90	30%		46	10%
1	116.9	40%	11	119.9	31%	1	256.9	20%	11	194	59%
	379.4	45%		369.9	48%		293.3	50%		341.8	31%
	80	13%		29	22%		66	13%		39	18%
2	112.2	45%	12	177.3	41%	2	161.7	34%	12	232.3	25%
	390.8	42%		394.4	37%		304.5	53%		280.6	57%
	99	23%		66	35%		100	10%		73	21%
3	425.2	33%	13	129.6	15%	3	356.9	30%	13	244.8	67%
	140.7	44%		364.9	50%		179.9	60%		380	12%
	41	33%		33	19%		76	23%		53	16%
4	244.1	33%	14	134.3	35%	4	222.1	34%	14	188.3	31%
	385	34%		373.3	46%		307	43%		298.3	53%
	61	35%		81	30%		15	15%		22	12%
5	398.6	35%	15	122.8	26%	5	245	40%	15	195.7	40%
	213.8	30%		383.6	44%		277	45%		300.2	48%
	27	28%		21	11%		56	11%		47	18%
6	208.6	31%	16	122.5	45%	6	169.2	41%	16	224.7	37%
	372.7	41%		380.9	44%		311.3	48%		296.4	45%
	59	20%		51	33%		36	15%		36	15%
7	101.2	30%	17	190.2	23%	7	214.2	43%	17	213.1	42%
	365.7	50%		373.7	44%		303.5	42%		302.5	43%
	39	18%		77	27%		71	21%		37	11%
8	423.6	31%	18	135.2	32%	8	208.2	31%	18	133.2	21%
	169.9	51%		392.5	41%		303.1	48%		283.7	68%
	40	30%		31	31%		53	18%		41	10%
9	228.7	40%	19	256.8	41%	9	172.1	15%	19	205.2	67%
	405.1	30%		393.2	28%		283	67%		362.5	23%
	70	25%		90	23%		100	20%		32	11%
10	379.9	45%	20	381.3	45%	10	163.8	30%	20	182	41%
	121.8	30%		102.3	32%		311.6	50%		305.9	48%

Table 23: Risky assets from problems in task 2 (mean = 225):

High risk (variance 20,000 approx.)						Low risk (variance 8000 approx)					
	27	24%		50	15%		90	30%		46	10%
1	434.9	24%	11	132.3	45%	1	308.9	20%	11	147.9	31%
	219.5	52%		394.9	40%		272.5	50%		295.8	59%
	29	22%		80	13%		66	13%		39	18%
2	166.2	37%	12	102.6	42%	2	192.9	53%	12	251.1	57%
	383.3	41%		381.2	45%		335.7	34%		299.4	25%
	66	40%		10	23%		10	10%		82	28%
3	169.1	10%	13	211.6	44%	3	209	60%	13	276.3	56%
	363.4	50%		392.8	33%		328.7	30%		295.8	16%
	31	16%		41	33%		76	23%		53	16%
4	108.3	33%	14	246.3	34%	4	232	43%	14	217.1	53%
	361.4	51%		387.1	33%		316.9	34%		327.1	31%
	52	33%		61	39%		15	15%		22	12%
5	211.8	30%	15	260.9	30%	5	246.9	45%	15	205.2	48%
	390	37%		396.6	31%		279.1	40%		309.7	40%
	23	19%		27	28%		56	11%		86	29%
6	111.9	27%	16	231.2	41%	6	180.3	48%	16	271.5	42%
	352.6	54%		395.5	31%		322.5	41%		296.7	29%
	82	23%		59	26%		77	25%		77	25%
7	104.1	31%	17	144.3	31%	7	256.6	56%	17	221.9	19%
	378	46%		383.6	43%		326.7	19%		292.1	56%
	44	11%		38	17%		37	11%		71	21%
8	129.1	50%	18	106.3	32%	8	212.7	68%	18	228.7	48%
	399	39%		361.8	51%		363.2	21%		323.6	31%
	31	31%		36	26%		41	10%		88	23%
9	231.2	28%	19	146.6	24%	9	128.4	23%	19	212.3	40%
	367.5	41%		360.9	50%		285.7	67%		323.9	37%
	21	23%		60	22%		100	20%		32	11%
10	171.4	33%	20	124.8	33%	10	200.8	50%	20	191.8	48%
	371.8	44%		379.2	45%		348.6	30%		315.7	41%

facing the 20 actual situations. You will face the trial situations one after the other. The points you earn for the trials will not be counted.

They then faced the two trial problems over two screens, and then saw a screen giving them their points for the two trial problems, with the reminder that these will not be counted for final points. In episodes with a deadline, they then saw a screen which reminded them of their deadline, after which they could proceed to the decision problems. In episodes without a deadline, the intermediate deadline reminder screen was absent.

For episode 1 of all conditions, the program ended after the 20th problem was over. For episode 2 of ***2C**, **2N** and **2X**, subjects were shown their accrued points after the 20th problem was over, and then faced a questionnaire collecting gender, income, and other information. Subjects in **1C** and **1X** also saw their accrued points and faced a questionnaire at the end of their second episode.*

I. Instruction screen for **1CL**, **1CH**, **2CL**, **2CH**, **2NL**, **2NH**, **2XL**, **2XH**, all episode 1

You will face 20 situations one after the other. In each you have decided to invest 100 units by buying units of financial options. There are two options available for each situation. After finishing a situation, please press the NEXT button, and the next situation will appear. You will now see a sample situation.

II. Instruction screens for **2CL**, **2CH**, episode 2

1. Welcome Back!! As in the previous session, you will face 20 situations one after the other. The situations will be similar to those you faced before. You will also get points for each situation as earlier.

2. And now the situations. In each you have decided to invest 100 units by buying units of financial options. There are two options available for each situation. After finishing a situation, please press the NEXT button, and the next situation will appear. You will now see a sample situation.

III. Instruction screens for **2NL**, **2NH**, **2XL**, **2XH**, all episode 2

1. Welcome Back!! As in the previous session, you will face 20 situations one after the

other. The situations will be similar to those you faced before. You will also get points for each situation as earlier.

The difference is here is now a time limit: your time limit is 11 minutes. Your points will accrue, but only those situations you complete before the time limit will be counted for final points. You will not receive points for situations completed after the time limit is over.

If you cross the limit without completing all 20 situations, you will be informed that the time limit is over. Please continue in that case till the end for the program to finish.

2. And now the situations. In each you have decided to invest 100 units by buying units of financial options. There are two options available for each situation. After finishing a situation, please press the NEXT button, and the next situation will appear. You will now see a sample situation.

IV. Instruction screen for **1XL**, **1XH**, episode 1

You will face 20 situations one after the other. In each you have decided to invest 100 units by buying units of financial options. There are two options available for each situation. After finishing a situation, please press the NEXT button, and the next situation will appear.

You will get points for each situation. However there is a time limit: your time limit is 11 minutes. Your points will accrue, but only those situations you complete before the time limit will be counted for final points. You will not receive points for situations completed after the time limit is over.

If you cross the limit without completing all 20 situations, you will be informed that the time limit is over. Please continue in that case till the end for the program to finish.

You will now see a sample situation.