CEE-M
CENTER FOR ENVIRONMENTAL ECONOMICS - MONTPELLIER

Hard vs. soft commitments: Experimental evidence from a sample of French gamblers

Paul Bettega
Paolo Crosetto
Dimitri Dubois
&
Rustam Romaniuc

CEE-M Working Paper 2023-09

CEE-M

INRAE

Hard *vs.* soft commitments: Experimental evidence from a sample of French gamblers*

Paul Bettega† Paolo Crosetto‡ Dimitri Dubois§ Rustam Romaniuc¶

August 31, 2023

Abstract

People use commitment devices to formalize and facilitate their goals. Self-commitments are ubiquitous and may take different forms: soft, when the commitment can be broken at a low cost, or hard, when that cost is high. The effects of soft and hard commitments have usually been studied separately. We conduct an online experiment with 1527 individuals representative of a big gambling company's client population to study the comparative effects of hard and soft commitment devices in a risk taking game. Our results show that asking for a hard limit leads subjects to reduce their risk-taking even when the limit turns out to be non-binding, i.e., the commitment is ex-post soft. Hard commitments lead to slightly lower levels of risk taking.

Keywords: soft commitment, hard commitment, risk taking, self-control
JEL Codes: C93, D02, D91

*We would like to thank our partner on this experimental project, *la Française des Jeux (FDJ). FDJ* provided financial support and access to a database of their clients but had not interfered at the data collection stage, data analysis, or writing of the paper.

†no affiliation

‡INRAE, GAEL. Email: paolo.crosetto@inrae.fr

§CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France. Email: dimitri.dubois@umontpellier.fr

¶Montpellier Business School. Email: r.romaniuc@montpellier-bs.com

1 Introduction

Self control is an important non-cognitive skill that is associated with favorable economic and social outcomes (Laibson et al., 1998; Heckman et al., 2006; Alan and Ertac, 2015). However, abundant evidence suggests that people have a hard time controlling their instantaneous passions and often succumb to temptation (Thaler and Shefrin, 1981; Milkman et al., 2014; Schilbach, 2019; Milkman, 2021). Consider the two-pack a day cigarette smoker who went through many attempts to quit but was never successful in kicking the habit. In New Year's resolutions, one intends to eat more healthy foods in the future, exercise more regularly, and watch television less often, but many of these intentions fail because of self-control problems. One popular solution to self-control problems is to use a commitment device. For example, in "How to get ready for retirement: Save, save, save", Rankin (1993) suggests to "Use whatever means possible to remove a set amount of money from your bank account each month before you have a chance to spend it." There is ample empirical evidence showing that commitments that restrict one's action space – called hard commitments (Giné et al., 2010; Schwartz et al., 2014; Royer et al., 2015; Burke et al., 2018; Sadoff and Samek, 2019) – work well to reduce one's tendency to err in the direction of instantaneous gratification, as judged by the person's own standards. It has been shown, for instance, that some people use specific ordering strategies that enforce watching "high brow" movies (Read et al., 1999) or that some individuals accept to put their money in temporarily locked savings accounts in order to keep it away from themselves (for more examples, see Milkman (2021)). Hard commitments allow individuals to "bind" themselves as Ulysses did before setting out to the Syrens (Elster, 1984). However, people often avoid using hard commitments because of their lack of flexibility. As pointed out by Ashraf et al. (2006), when being offered to commit to a locked saving account, only 28% of customers accepted the offer.

An alternative to hard commitments that can be used to prevent oneself from succumbing to temptation while providing more flexibility are soft commitment devices, i.e., commitments that can be easily broken (see Bryan et al. (2010) for an extensive review of the literature on commitment devices). Deviating from soft commitments involves psychological costs such as shame (if the commitment was made public) or guilt (if it was made privately) or some degree of both shame and guilt. Examples of soft commitments backed by non-pecuniary costs include taking a fixed amount of money when going out with friends (one can always borrow on the spot, so the commitment is soft), brushing one's teeth earlier in the evening to avoid late night snacking (the cost of redoing it is low), renting a place in an open space to avoid taking a nap when working from home (couches may also be available in open spaces). Given the ubiquity of soft commitment devices, it is important to understand to what extent making a commitment soft as compared to hard changes people's behavior. However, to our knowledge, there is limited empirical evidence on the effects of soft commitments relative to hard and no commitments in a specific situation where temptation is present. Our aim in this paper is to fill this gap.

To study the comparative effects of soft relative to hard and no commitments, we ran an online artefactual field experiment on a representative sample of French gamblers. We partnered with *la Française des Jeux*, the operator of France's national lottery games, to get a sample of 1527 participants, representative of *la Française des Jeux*'s gamblers.[1] As gambling behaviors may be viewed as a sign of a lack of willpower, we believe that it is important to study the comparative effects of hard and soft commitment devices in a population that may be less prone to self-restrain especially when it comes to making decisions in a risky game.

The experiment consists of two conditions, *Baseline* and *Commitment*, built on top of a modified Balloon Analogue Risk Task (BART, Lejuez et al., 2002). The BART is a risk-elicitation game in which subjects pump air in a fictional balloon, and collect money proportional to the number of pumps, unless the balloon bursts, in which case they lose the accumulated payoff. In the *Commitment* condition, after five rounds of the BART subjects were given the opportunity to select an upper limit on the number of pumps for five further rounds. Subjects were informed that the limit would be binding with a 25% chance. After subjects made their choice regarding the self-imposed limit, they were informed whether the limit was binding or not. The fact that the limit was binding with a 25% chance allows us to capture the demand for a commitment device as well as to compare decisions under two different environments: (1) when commitment is hard given that the limit is binding and (2) when commitment is soft given that the subject is free to choose any number of

[1] This is one of the rare studies on commitment devices that uses such a large sample of participants that are representative of the population of gamblers of a major national operator of lottery games. While there are field studies on self control that employ a non-student population (Ashraf et al., 2006; Milkman et al., 2014), we are not aware of any study on the topic of commitment devices to increase self-control using a sample of subjects with similar characteristics to our participants.

pumps but knowing that they had expressed a preference to limit their behavior to a certain extent. Given that the limit has a positive probability of being applied, revealing one's true preference is a dominant strategy.

We find substantial demand for commitment, with more than a third of subjects setting themselves a limit, that in half of the cases can be said to be restrictive – i.e., lower than previous play. We further find that the risk-taking reduction is significant for both the *hard* and the *soft* commitment groups – i.e., risk taking decreases *also* for the group of subjects having asked for a limit but to whom the limit was not applied. However, soft commitments do not seem to be as effective as hard commitments.

Our paper relates to the literature that seeks to test whether and how commitment devices can make people succumb less to temptation. The literature on hard commitment devices has hitherto been considered quite independently from research that investigates the effects of soft commitments. For instance, Trope and Fishbach (2000), Ariely and Wertenbroch (2002), and Houser et al. (2018) compare hard commitments to a control without any commitment. Despite the importance of such comparisons, a better understanding of hard commitments requires a comparison between hard and soft devices. Indeed, hard commitments impose both a pecuniary and a non-pecuniary cost in case of breach of commitment. The pecuniary cost can go to infinity if the individual decides to remove altogether the tempting option from their choice set. At the same time, hard commitments come also with non-pecuniary costs in case of a breach such as shame or guilt (Kast et al., 2012). The decision to deviate from one's commitment may also signal to the individual a lack of willpower, which may represent another source of psychological discomfort (Bénabou and Tirole, 2004). On the other hand, soft commitments are backed solely by non-pecuniary costs. When present, pecuniary costs are mostly symbolic. Our experimental study allows us to compare the impact of hard commitments relative to soft and no commitments. We show that self-commitments have non pecuniary costs that explain why a self-imposed limit that can be crossed without incurring any consequence can still change people's behavior – albeit less so than a non-revertible hard limit. It is worth noting that we are not able to say whether the demand for soft commitments is different than the demand for hard commitments, which limits our capacity to draw more general policy implications from our experiment.[2] However, given previous research showing that people generally avoid using hard commitments because of their lack of flexibility, our results indicate that soft commitments do have an effect on behavior. Offering soft commitment devices results in behavioral change, even if the extent to which people change their conduct is less important than with hard commitments.

Another aspect of our study that should be stressed concerns the nature of commitment. As in Hoong (2021), in our experiment the commitment choice is not binary but continuous. Thus, our study captures the degree of commitment that is endogenously chosen by the subjects. However, Hoong (2021) studied commitment devices to limit phone, Facebook, and Instagram usage by adolescents while our focus is on risk taking. Furthermore, their study analyzes the effect of soft commitments while we compare the effects of hard and soft commitments. Our work is also related to the theoretical investigation of commitment by Gul and Pesendorfer (2001). In their model, there is a cost of avoiding the most tempting item in a choice set at the moment of decision-making. Individuals, therefore, benefit from removing these items in advance. Our experiment allows subjects that might be tempted by taking too much risk, as is likely the case in our gamblers sample, to eliminate tempting options from their choice set by choosing an upper limit on the number of pumps that they will be able to select. However, Gul and Pesendorfer (2001) are interested in modelling the demand for such commitments while we empirically analyze their behavioral effects after some participants to our study decide to eliminate tempting options from their choice set.

2 Experimental Design and Procedure

2.1 Experimental conditions

Our experiment consists of two experimental conditions: a *Baseline* and a *Commitment* condition.

In each condition, subjects played 10 rounds of a modified version of the BART. The screen showed a small simulated balloon. Each subject had to choose a number of pumps between 1 and 64, knowing that each pump would inflate the balloon and would yield a gain of €0.15. However, each pump could result in the explosion of the balloon. The probability that a balloon would explode is given by a draw without replacement from an urn containing 64 tokens, as in Lejuez et al. (2002). This

[2] On demand for commitment, see Carrera et al. (2021) and Ek and Samahita (2022).

gives an additional probability of a burst of 1/64 for each pump. If the subject chose 20 pumps out of the 64 possible, then the probability that the balloon would explode was 20/64. As in Lejuez et al. (2002), choosing a higher number of pumps (i) increased the amount to be lost because of an explosion and (ii) decreased the relative gain of any additional pump. In our experiment, the average break point was 32 pumps. That is, a risk-neutral subject would maximize their gains by choosing 32 pumps.

Note that in the original study by Lejuez et al. (2002) subjects had to click on a pump button to inflate the balloon and they had to click on it as many times as they wanted knowing that the balloon could explode at any moment. We modified the original BART study along two dimensions. First, because we are interested in behaviors under risk rather than ambiguity, we decided to inform subjects about the range of outcomes and that a priori each pump is equally likely to result in an explosion. Second, we asked subjects to choose the desired number of pumps before the balloon started to inflate. That is, a subject had to indicate a specific number of pumps and only then the balloon started to inflate until it reached the chosen number of pumps or exploded – whichever happened first. This way, we did not constrain the number of chosen pumps on balloons that exploded, which allowed us to avoid the truncation of the data usual for BART studies.

In all conditions, at the end of each round, subjects were informed about the outcome of the balloon task (whether it exploded or not before it reached the indicated number of pumps) and about their earnings in that particular round. At the end of the experiment, one round out of the ten was randomly chosen for payment and this was common knowledge from the outset of the experiment. The "pay one" approach can help to avoid wealth effects and hedging (Charness et al., 2016).

In the *Baseline* condition, subjects played 5 rounds of the BART followed by a 10 seconds pause where they saw a message informing them that the game would resume after a few seconds. After the pause, they had to play for 5 more rounds that were identical to the first 5 rounds. Subjects were informed at the beginning of the experiment that there was a total of 10 rounds. The 10 seconds pause was implemented to mimic the break that we implemented in the *Commitment* treatment, with the exception of the commitment mechanism introduced in the latter condition but absent in the *Baseline*.

In the *Commitment* condition, the first 5 rounds were identical to the *Baseline*. However, at the end of round 5, instead of the pause, subjects were offered the possibility to select an upper limit on the number of pumps that they could choose in all of the rounds that would follow – i.e., round 6 to 10. All subjects were informed that the limit would be binding with a 25% chance. For those who opted for no limit, they went on to round 6 of the BART as in the *Baseline*. For those who opted for a limit, after subjects made their choice, they were informed whether the chosen limit was binding or not. Then, subjects proceeded to round 6 of the BART. In case the limit was binding, subjects could choose a number of pumps between 1 and the chosen limit. If the limit was not binding, subjects could choose any number of pumps between 1 and 64, as in the *Baseline* condition.

Two design choices are worth discussing. First, the choice of having the first 5 rounds identical across the two conditions. This sequence was implemented for two reasons: (i) to allow subjects to get accustomed with the game and (ii) to capture subjects' "natural" risk preferences in the absence of any commitment device. This way, we can ensure that our subjects have overall similar risk profiles across the two conditions by looking at behaviors in rounds 1-5. Additionally, we make sure that those who are offered the possibility to choose a limit after round 5 have been exposed to the game and that they had gotten a feeling of their temptation level.

The second design choice concerns the stochastic aspect of the limit. The fact that the limit was binding with a 25% chance allows us to capture the demand for a commitment device as well as subjects' self-control when the limit is effectively implemented (thus, making it a hard commitment) compared to when it is non binding (making it a soft commitment). Since the limit has a positive probability of being applied, revealing one's true preference is a dominant strategy.

2.2 Participants

The recruitment process started in October and ended in December 2019. 803 subjects participated in the *Baseline* and 724 in the *Commitment* condition. Subjects were recruited by *Bilendi*, a private company, within the framework of a partnership that some of the authors of this study concluded with *la Française des Jeux* (FDJ), which is the operator of France's national lottery games.[3] Bilendi recruited subjects for this study from a pool of more than 1 million individuals who had a personal

[3]For more information about the two companies, see their respective websites: FDJ and Bilendi

account with FDJ. The study includes individuals who had declared that they played at least once one of FDJ's games during the 12 months prior to the study (in one of FDJ's physical sale points or online). Therefore, all our subjects have some appetite for gambling, which makes this study original compared to using a population that may be less prone to temptation.[4]

The sample in our experiment was 40.52% female and 59.48% male, and relatively more evenly distributed than traditional student samples when it comes to subjects' age: 7.50% of participants were between 18 and 24 years old, 20.67% between 25 and 34 years old, 37.43% between 35 and 49 years old, 24.89% between 50 and 64 years old and 9.51% were 65 years old or older. Considering the employment status of our subjects, the sample was 46.25% blue collar, 26.78% white collar, 14.05% retired, 9.45% inactive (not working or looking for a job) and 3.47% students.

2.3 Procedure

The experiment was developed using oTree (Chen et al., 2016). Each subject could participate only once, and strict anonymity was ensured.[5] At the end of the experiment, the LEEM sent a file with the identifiers and associated payments to *Bilendi*, which then paid the subjects their experimental earnings via PayPal.

The experiment lasted on average 20 minutes. At the very end of the experiment, subjects were informed about the round number that was randomly selected to count for payment. The average payoff was €1.60 (std 1.84). In addition to the earnings related to their decisions, each subject was paid a €5 participation fee.

3 Results

The data as well as all scripts to replicate all analyses described in this section can be found online in this github repository.

[4]For example, Milkman et al. (2014) studied temptation and commitment in a sample of gym users. It is quite possible that people who have a gym subscription have greater willpower than people who have no gym subscription. Other studies, for example Ariely and Wertenbroch (2002), Casari (2009), and Houser et al. (2018) relied on a population of students.

[5]To ensure this, the Laboratory for Experimental Economics in Montpellier (LEEM) created a dedicated "Room" on its oTree server with unique identifiers. *Bilendi* sent personalized invitation emails with a unique identifier. This ensured that on the LEEM server we only had the identifiers and decisions while Bilendi had the identifiers and identifying information of the subjects, but did not have access to the decisions of the subjects in the experiment.

803 subjects participated in the *Baseline* treatment, playing 10 rounds of our modified BART. Their data serves as a counterfactual. 724 subjects participated in the *Commitment* treatment, and, after round 5, were offered the opportunity to set an upper limit on the number of pumps.

We chose not to consider the observations for the first round, as behavior in that round is starkly different from all other rounds. Visual inspection of the data makes this abundantly clear (see Figure 2 in Appendix A), and tests confirm stark differences both in distribution (Kolmogorov-Smirnov test, p-value < 0.001) and mean (27.28 for round 1, 20.95 for all other rounds combined, t-test p-value < 0.001) in both the baseline and commitment conditions. Furthermore, it is well known in the experimental literature studying dynamic games that decisions in the first round are different than decisions made in the rest of the game because of learning effects (Andreoni, 1988; Crosetto and Filippin, 2013). Round 2 is also somewhat different from some of the other rounds, but the difference is small and does not warrant exclusion. We therefore restrict our analysis to rounds 2 to 10.

This section will be organized around two main axes: analysis of the demand for commitment by our subjects, and the effects of applying a soft or a hard commitment on risk-taking behavior. We can only identify demand for *hard* commitment, as subjects ex-ante knew that the commitment could be hard with 25% chance; on the other hand, as subjects were free to set the limit they wanted, we can look at the intensive margin of demand for commitment – how *harsh* the self-imposed limit was with respect to previous play. When looking at the impact of commitment we can distinguish *soft* and *hard* commitment, as ex-post subjects were randomly sorted into these two possibilities.

3.1 Demand for commitment

Theoretically, *if* subjects follow a wide class of decision models that do not include temptation, multiple selves or multiple conflicting utility functions, the possibility to set a limit cannot improve their situation. Indeed, if the constraint is applied it only makes the choice of a higher number of pumps (i.e., of riskier choices) impossible. But note that without the constraint, subjects have no incentives to choose a higher number of pumps than the one they intrinsically desire to choose. Theoretically, hence, we expect subjects not to self-impose a limit or to be indifferent between choosing a limit that they would never cross anyway (in the absence of the commitment device) and choosing no limit.

Notwithstanding, 35.1% of subjects set themselves a limit. This share is consistent with results observed by Houser et al. (2018) (28.6% chose a limit when it had no cost) and Toussaert (2018) (35.8%) but lower than what Alcott et al. (2022) found in the context of digital consumption (78% of their subjects chose to limit their use of most popular online platforms) and higher than the share of subjects who chose to self-restrict their game time in Acland and Chow (2018) (25%). Thus, in line with previous research, we find that there is a demand for a limit that *can* be binding. Note that when setting the limit, subjects do *not* know if the commitment will be soft or hard; they must assume it to be hard, though, as that possibility is there. Subjects setting themselves a limit were older than subjects choosing not to (46.2 *vs.* 43.5 years on average, t(492)=2.40, p = 0.017, d = 0.19), and were more likely to be males (64.8% *vs.* 56.3%, $\chi^2(1)$=4.43, p = 0.035); the two groups cannot be said to differ with respect to their profession ($\chi^2(4)$=3.26, p = 0.516) nor with respect to their self-reported gambling behavior (t(522)=1.11, p = 0.269, d = 0.08).

Among the subjects choosing to set a limit, 42.5% set a *restrictive* limit, i.e. a limit that was lower than the maximum number of pumps chosen in rounds 2-5, and 13% more set a *weakly* restrictive limit, i.e. a limit *equal* to the maximum of previously chosen pumps. 20.1% set a more stringent limit, i.e. a limit that was lower than the *mean* of their pumps in rounds 2-5, and a further 3.9% set a weakly binding stringent limit. These subjects are referred to in the literature as *sophisticated* types, aware of their self-control problems and opting for a commitment device (Milkman, 2021). Most subjects, instead, chose a limit that was not restrictive given their past behavior; we cannot know whether they intended to increase their pumps in further periods, and hence subjectively found the chosen limit restrictive given their intentions on future periods.

In our experiment the harshness of the limit is endogenous – i.e. subjects could place the limit anywhere, relative to their behavior in the preceding rounds. We compute an index $\mathcal{H}$ of limit harshness as

$$\mathcal{H} = 1 - \frac{\max_p}{L},$$

where $\max_p$ is the maximum number of pumps in periods 2-5 and L is the limit set at the end of period 5. The index ranges from 1 – indicating that the limit is set at 0 pumps, i.e,. is the harshest possible limit – to negative numbers – indicating that the limit is *not* restrictive. A harshness of 0 indicates that the limit was set exactly at the maximum level of pumps in the first part of the experiment; a limit of -1 indicates that the limit was set at *twice* the maximum number of pumps in previous play.

The distribution of harshness (cut to -1 for visual convenience) is given in Figure 3 in Appendix A. The overall average harshness is -0.75 (st.dev; 4.69), that is, on average subjects set a limit 75% *higher* than their previous maximum level of risk taking, but with large heterogeneity. Among the subjects who did set a restrictive limit, average harshness is 0.35 (st.dev. 0.25), that is, the limit is on average 35% lower than the maximum number of pumps in previous periods. Overall, the (limited number of) subjects that *did* set themselves a restrictive limit, chose to restrict themselves quite substantially. The harshness of the limits does not correlate with demographics or with self-reported gambling behavior.

3.2 Effects of soft and hard commitments

In this section we focus on the change in behavior shown by subjects when facing no, a soft, or a hard commitment. Only 1 subject in 4 had to really face the self-imposed limit, i.e. faced a hard commitment. For 3 out of 4 subjects who chose a limit, the commitment turned out to be soft. Subjects' behavior with respect to their self-imposed limit was markedly different between the subjects facing a soft and a hard commitment.

	Share of subjects who pump ... the self-imposed limit		
	...below...	...at...	...above...
Periods 2–5			
Soft commit	43.85%	13.90%	42.25%
Hard commit	46.27%	10.45%	43.28%
Periods 6–10			
Soft commit	50.27%	19.25%	30.48%
Hard commit	59.70%	40.30%	0

Table 1: Behavior before and after having set a self-imposed limit, by type of commitment faced

Table 1 shows the percentage shares of subjects whose maximum number of pumps in the first and second phase is below, at, or above the self-set limit, separately for the 1/4 of subjects whose commitment turned out to be hard and the 3/4 for which it turned out to be soft.

As expected, behavior of the soft and hard commitment groups with respect to the self-set limit is virtually indistinguishable for periods 2-5 ($\chi^2 = 0.531$, p = 0.77). As shown above, about 42% of subjects set a restrictive commitment – i.e. they set a limit *below* the maximum number of pumps chosen before setting the limit. About 10 to 14% set a limit exactly equal to the maximum amount of pumps; and about 43 to 46% set non-restrictive limits.

The effect of the limit on periods 6-10 is different across groups facing a hard or a soft commitment ($\chi^2 = 30.1$, $p < 0.001$). Subjects facing a hard commitment are forced to comply with the limit, so no subject exceeds the limit after having set it. The share of subjects complying exactly with the limit jumps from 10 to 40%; while the number of subjects restricting themselves further increases from 46 to 59%. These differences are significant ($\chi^2 = 41.9$, $p < 0.001$). Subjects facing a soft commitment *also* see a significant effect of the commitment ($\chi^2 = 5.99$, p = 0.05). This is indication that subjects face indirect, psychological costs of not complying with a self-set limit even when this limit simply does not apply. The number of subjects overshooting their self-imposed limit decreases from 42 to 30%, while the share of subjects exactly complying with the limit or restricting their pumps even more increases by 6 percentage points, each.

Our data also allows us to know *how much* did subjects restrict themselves – or overshot their self-imposed limit. To do this we propose a saturation indicator:

$$\mathcal{S} = \frac{P}{L},$$

where L is the limit and P the mean number of pumps in periods 6-10, after the setting of the limit. This indicator yields 1 if the subject chose a number of pumps exactly equal to the self-set limit; it is higher than 1 if the limit is exceeded, and lower than 1 if the subjects restricted their behavior more than the pre-set limit.

In order to make this indicator directly comparable between the soft-commitment subjects – to whom the limit was *not* applied – and the hard-commitment subjects – to whom it was – we also compute a *truncated saturation* indicator:

$$\mathcal{S}_t = \begin{cases} \dfrac{P}{L} & \text{if } P \leq L \\ \\ 1 & \text{if } P > L \end{cases}$$

This truncated indicator treats every subject exceeding the limit as being exactly on the limit, and is as such not influenced by the application of the limit. This allows us to more readily compare the behavior of subjects to whom the limit has not been applied with that of those to whom it has. Table 2 shows aggregate statistics for the saturation indicator and its truncated version before and after the imposition of the limit for both hard and soft commitment groups.

	Rounds 2 – 5	Rounds 6 – 10	Difference
Saturation ratio			
Soft commit	0.98 (1.58)	0.84 (0.88)	-0.14
Hard commit	0.83 (0.75)	0.65 (0.23)	-0.18
Truncated saturation ratio			
Soft commit	0.67 (0.29)	0.67 (0.28)	0.00
Hard commit	0.68 (0.25)	0.65 (0.23)	-0.03

Table 2: Mean (st.dev.) limit saturation by phase and type of commitment

Subjects reduced substantially their risk-taking with respect to the limit in the second part of the experiment – more so in the hard commitment group, but also among those who faced a soft commitment. The fact that the truncated indicator is substantially lower than its not-truncated counterpart indicates that subjects exceed the constraint if they could. The difference, though, mostly disappears when looking at the truncated saturation indicator – this seems not to be different across soft and hard commitment groups, nor before or after the limit is set – with the except of the hard commitment group slightly decreasing their saturation after setting the limit. This indicates that there is little difference between soft and hard commitment groups apart from the mechanical effect of the hard limit; in both cases subjects do limit themselves; in the hard commitment case, subjects are further mechanically forced to comply.

Table 2 hides the large heterogeneity of subjects when setting a limit, and when deciding how to behave with respect to this limit. Some subjects set a constraining limit, and then complied with it; others set a non-constraining limit, and then largely ignored it, not changing their behavior across phases. Further, some others set a limit *higher* than their previous behavior, and then *increased* their risk taking, saturating the constraint *more* after the limit was set than before. To get an idea of this heterogeneity, Figure 4 in Appendix A shows a raincloud plot of the saturation and truncated saturation ratios for both hard and soft commitment, before and after the limit was set.

Table 3 gives additional insights on these different strategies by subjects with respect to the self-imposed limit. It splits the population by the nature of the limit that they self-imposed (lower than the mean of previous pumps – biting – or not) and identifies if the number of pumps was reduced, stood put, or increased after setting the limit.

Nature of the limit	Change in behavior	Share of subjects	Mean change
Soft commitment			
Biting (26%)	Reduction	79.17%	-0.94
	No change	8.33%	–
	Increase	12.5%	0.43
Non-biting (74%)	Reduction	43.88%	-0.18
	No change	0.72%	–
	Increase	55.4%	0.22
Hard commitment			
Biting (19%)	Reduction	92.31%	-1.01
	No change	7.69%	–
	Increase	–	–
Non-biting (81%)	Reduction	50%	-0.15
	No change	1.85%	–
	Increase	48.15%	0.18

Table 3: Mean change with respect to limit after limit set

Table 3 shows that even in the soft-commitment case, a large share of subjects reduced their risk taking after having set a limit. Setting a biting limit seems to indicate a higher willingness to reduce risk taking with respect to setting a limit that was not biting. In this case, the reduction is rather high – the saturation ratio decreases by nearly 1. This might indicate that these are subjects who set a very stringent limit with respect to previous behavior, and stuck to it. In the case of hard commitment, subjects who set a biting limit have no room to increase their risk taking; but subjects setting a limit higher than their mean risk taking could still do it, and about half did so.

To summarize, we observe that our subjects do not saturate their constraint. Indeed, the saturation indicator is less than 1. Subjects exceed their limit when they can do so. We find that the saturation reduction is important when the limit is binding. Finally, we also observe a reduction for subjects for whom the constraint is not binding, thus suggesting that soft commitments can work to change risk-taking behavior.

3.3 Self-imposed limits and risk taking

We can further look at the effect of setting (or not) and then facing (or not) a commitment on risk taking. We characterize how much risk did subjects take before and after the setting of the limit, depending on the limit choice and whether the limit was applied. We focus on five groups of subjects: those in the *Baseline*, for whom nothing changed; those in the *Commitment* condition that *refused* a limit; those in the *Commitment* condition that set themselves a limit, but to whom the limit was not applied (soft commitment); those in the *Commitment* condition that set themselves a limit, and to whom the limit was applied (hard commitment); furthermore, to put the soft *vs* hard commitment more in perspective, we apply to the soft commitment group the limit they set themselves. This builds a counterfactual that shows us how much the soft commitment group deviated from their own self-set limits, and makes them comparable with the hard commitment group.

The descriptive statistics of this exercise are given in Table 4. We refrain from running traditional tests on these variables, though, as the behavior of a large part of subjects is very erratic, making it difficult to interpret mean pumps. Some subjects held a behavior that is consistent with the one expected if they had a clear preference and acted on it – i.e., chose a number of pumps and did not deviate from it. But the majority of subjects displayed a very erratic behavior, submitting very different choices across periods. This is especially so for the less risk averse or risk loving subjects. This erratic behavior is typical of Balloon-like tasks, and has been observed also in experiments with no or delayed feedback (for instance, see Crosetto and Filippin (2013)). The dynamic behavior of all subjects – pumps across periods – is visually displayed in Figure 5 in Appendix A, where subjects are classified according to their mean number of pumps – our proxy for their risk attitude –

and its standard deviation across periods – a proxy for erratic behavior.

	Rounds 2 – 5	Rounds 6 – 10	Difference
Baseline	21.24 (12.15)	21.09 (12.9)	-0.15 (9.46)
Limit refused	21.06 (12.04)	21.68 (12.81)	0.62 (9.09)
Soft commit	19.87 (12.21)	19.14 (11.68)	-0.74 (9.11)
Hard commit	20.16 (11.49)	18.81 (10.47)	-1.35 (8.4)
Soft commit counterfactual	19.87 (12.21)	17.55 (11.02)	-2.33 (9.15)

Table 4: Mean (st.dev.) of pumps before and after the setting of the limit, by subject group

Given this erratic behavior, we prefer to run a Markov-chain Monte Carlo (MCMC) analysis. This method allows us to obtain a distribution for the mean rather than a point estimate and thus to compare the parameters on their distribution rather than on a hypothetical normal distribution that is not borne out by the data.

We implement MCMC in python using the PyMC package (Salvatier et al., 2016). As MCMC is a Bayesian algorithm, we had to make assumptions on the distribution of the data and on the associated parameters. We chose as a prior a Poisson distribution, whose parameter is distributed according to an exponential law with a mean equal to the observed mean of the data. The Poisson distribution is justified in that the observed data are discrete and for the values of the studied parameters the number of theoretical observations outside the sample is very low.[6]

The distribution of the means of the difference between the first and second part, for each different subject group plus the counterfactual is presented in Figure 1; the mean of this mean distribution and its credible interval at 95% are in Table 5.

	Mean difference	95% credible interval
Baseline	-0.16	[-0.38, 0.05]
Limit refused	0.63	[0.34, 0.91]
Soft commit	-0.60	[-1.03, -0.18]
Hard commit	-1.04	[-1.75, -0.32]
Soft commit counterfactual	-2.16	[-2.59, -1.75]

Table 5: MCMC results: means and credible intervals of the mean difference of pumps across the two phases

The Table and Figure highlight several findings. First, in the Baseline, where nothing happens,

[6]Moreover, this hypothesis yields better results than the alternative of using a binomial distribution.

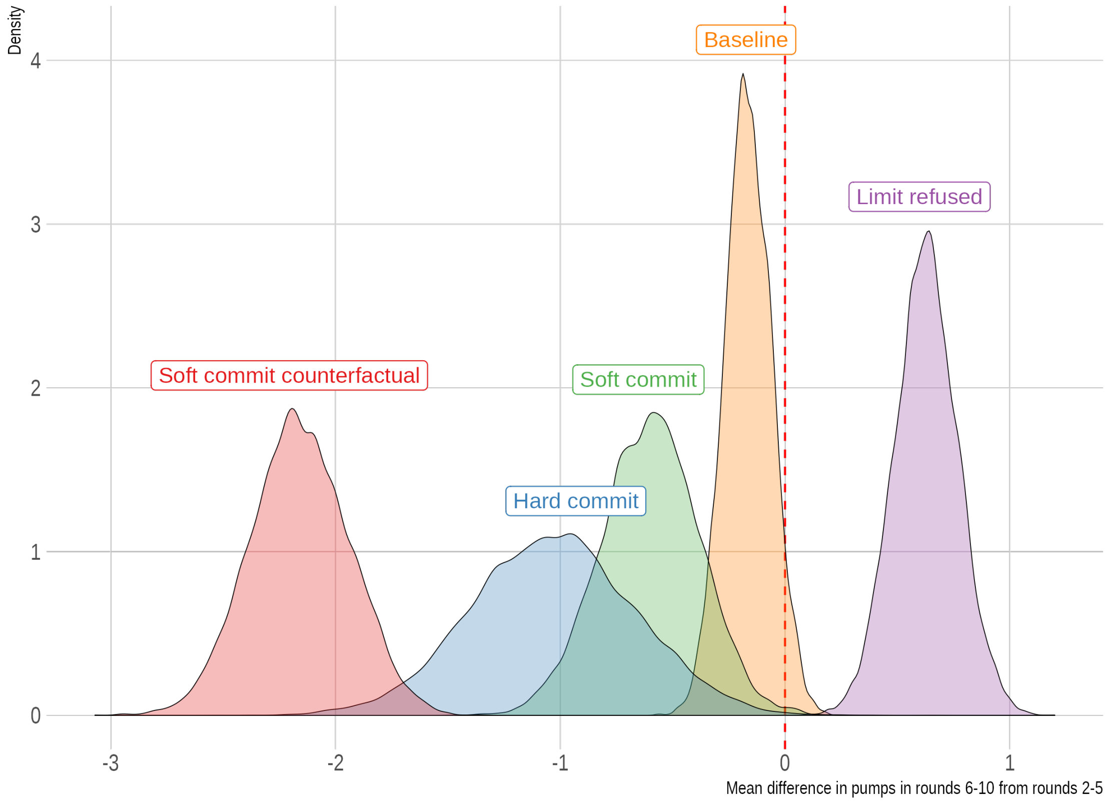

Figure 1: MCMC results: distribution of the mean difference of pumps across the two phases

there is no credible change from rounds 2 to 5 to rounds 6 to 10. In the Commitment treatment, a selection effect operates at first. Subjects not imposing themselves any limit tend to significantly increase their risk taking in the second part of the experiment. Subjects setting themselves a limit end up significantly decreasing their risk taking, with the hard commitment group being mechanically forced to adhere to the limit and hence decreasing more. Soft commitment subjects still decrease significantly their risk-taking, though.

The largest decrease in risk taking is recorded in the counterfactual group – created by applying the limit to the soft commitment group. This means that the soft commitment group had committed to a much harder limit than subjects ending up in the hard commitment group, and, had that limit been implemented, they would have seen a large decrease in risk taking. This puts the strength of reduction in risk taking by the soft commitment group in perspective. Soft commitments are, in the end, less effective than hard ones when we compare them to the actual bar subjects had set to themselves.

The analysis of the effects of offering subjects to choose a limit allows us to make the following conclusions. First of all, if the commitment treatment does not create a difference in terms of risky behavior compared to the baseline this is due to the fact that even if subjects who asked for a limit reduced their risk taking, subjects who refused to self-impose a limit increased it. Second, we show that asking for a limit leads subjects to reduce their risk-taking even when the limit turns out to be non binding. This shows that subjects are able to comply to some extent with the restrictions that they self-imposed even when deviations are free from consequences. However, soft-commitments do not seem to be as effective as hard commitments because the reduction in risk-taking is (slightly) greater when the limit requested by the subjects is binding. Finally, we see that the reduction in risk-taking is even greater when the limit is applied ex-post. This confirms what we have seen with the saturation of the constraint, i.e., that soft constraints do work, but up to a point. This observation is consistent with the idea of convex self-control costs – even if we cannot offer any conclusion on this conjecture with our data, and this hypothesis would require additional experimental inquiry.

4 Conclusion

The extant literature has investigated the effects of hard commitment devices that impose financial and non-financial costs in case of breach of commitment (e.g., Ashraf et al. (2006)) separately from soft mechanisms, such as expressing a preference to limit one's use of some activity without any form of external punishment or restriction. Soft commitments are most often solely backed by guilt and shame (e.g., Bhanot (2017)). As not everyone is comfortable with the idea of a commitment device that imposes significant penalties or restricts future freedoms, those who cannot stomach the thought of hard commitments may do better with a different form of commitment device. It is, therefore, important to understand to what extent soft commitments are a good substitute for hard ones using an identical environment where the two types of commitment devices can be compared.

We conducted an online experiment with 1527 individuals that are representative of a big gambling company's client population to study the impact of a hard commitment device relative to soft and no commitments, thus underlying the effect of non-pecuniary costs on one's capacity to resist a tempting activity. Compared to more standard experiments that examined temptation and self-control in the laboratory, our subject pool is more diverse in terms of age, occupation, and is composed of individuals who may have self-control problems to a greater extent than a standard population. Our study of commitment devices thus focuses on how to change the behavior of a population that may be most in need of solutions to resist tempting activities that they may want to avoid in the first place but have a hard time to do so. Furthermore, compared to some field experiments (e.g., Ashraf et al. (2006); Burger et al. (2011)), our study retains the advantages of laboratory studies in terms of control over the subjects' decision environment as we implemented an online balloon analogue risk task (Lejuez et al., 2002).

Our results can be summarized as follows. First, a minority of subjects in our experiment voluntarily opt for a commitment device (a limit), a result in line with previous experimental studies (Houser et al., 2018; Toussaert, 2018). Second, when comparing soft vs hard commitment groups, we observe a reduction in risk-taking behavior for both groups when compared to the baseline. That is, asking for a limit leads subjects to reduce their risk-taking even when the limit is not applied. The reduction is however slightly greater when the limit requested is binding.

Our results are thus in line with the previous literature showing that soft (Kast et al., 2012; Hoong, 2021) and hard (Laibson et al., 1998; Trope and Fishbach, 2000; Ashraf et al., 2006) commitments work to increase self-control. The contribution of this paper is to compare the two under identical conditions and show that soft commitments are an appealing instrument. Hard commitments only slightly overperform soft commitment devices in terms of reducing risk-taking behavior.But soft commitment devices are more attractive than hard ones and demand for more flexible commitments would presumably be higher than for binding mechanisms. One of the main limitations of our study is that we do not quantify the demand for each type of commitment device, which prevents us from drawing additional implications from our experiment.

5 Declarations

Data Availability. All data used in this experiment, as well as all scripts used to produce tables, figures and test results described in this paper are available at the paper's github repository

Competing Interests. PB and PC declare no competing interests. RR and DD received financial compensation for consultancy work from the organization that financially supported this research project (FDJ). However, FDJ may not gain or lose financially from the publication of this article and the authors (including RR and DD) received no financial compensation for this research.

Funding. This research project has been funded by *La Française des Jeux (FDJ).* However, FDJ may not gain or lose financially from the publication of this article. FDJ did not interfere with the design, data collection, data analysis or writing of this research article.

Contribution Statement. Overall idea and conception: PB (commitment treatment), DD and RR (risk and adaptation of the BART); Experimental software & data collection: DD; data analysis: PB, PC, DD; writing: PB, PC, DD, RR.

References

Acland, D., Chow, V., 2018. Self-control and demand for commitment in online game playing: Evidence from a field experiment. Journal of the Economic Science Association 4, 46–62.

Alan, S., Ertac, S., 2015. Patience, self-control and the demand for commitment: Evidence from a large-scale field experiment. Journal of Economic Behavior & Organization 115, 111–122.

Alcott, H., Gentzkow, M., Song, L., 2022. Digital addiction. American Economic Review 112, 2424–63.

Andreoni, J., 1988. Why free ride?: Strategies and learning in public goods experiments. Journal of Public Economics 37, 291–304.

Ariely, D., Wertenbroch, K., 2002. Procrastination, deadlines, and performance: Self-control by precommitment. Psychological Science 13, 219–224.

Ashraf, N., Karlan, D., Yin, W., 2006. Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. The Quarterly Journal of Economics 121, 635–672.

Bénabou, R., Tirole, J., 2004. Willpower and personal rules. Journal of Political Economy 112, 848–886.

Bhanot, S.P., 2017. Cheap promises: Evidence from loan repayment pledges in an online experiment. Journal of Economic Behavior & Organization 140, 246–266.

Bryan, G., Karlan, D., Nelson, S., 2010. Commitment devices. Annual Review of Economics 2, 671–698.

Burger, N., Charness, G., Lynham, J., 2011. Field and online experiments on self-control. Journal of Economic Behavior & Organization 77, 393–404.

Burke, J., Luoto, J.E., Perez-Arce, F., 2018. Soft versus hard commitments: A test on savings behaviors. Journal of Consumer Affairs 52, 733–745.

Carrera, M., Royer, H., stehr, M., Sydnor, J., Taubinsky, D., 2021. Who chooses commitment? evidence and welfare implications. Review of Economic Studies , forthcoming.

Casari, M., 2009. Pre-commitment and flexibility in a time decision experiment. Journal of Risk and Uncertainty 38, 117–141.

Charness, G., Gneezy, U., Halladay, B., 2016. Experimental methods: Pay one or pay all. Journal of Economic Behavior & Organization 131, 141 – 150. URL: http://www.sciencedirect.com/science/article/pii/S0167268116301779, doi:https://doi.org/10.1016/j.jebo.2016.08.010.

Chen, D.L., Schonger, M., Wickens, C., 2016. otree—an open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9, 88 – 97. URL: http://www.sciencedirect.com/science/article/pii/S2214635016000101, doi:https://doi.org/10.1016/j.jbef.2015.12.001.

Crosetto, P., Filippin, A., 2013. The "bomb" risk elicitation task. Journal of Risk and Uncertainty 47, 31–65.

Ek, C., Samahita, M., 2022. Pessimism and overcommitment: An online experiment with tempting youtube content. UCD Center for Economic Research Working Paper 22/03 .

Elster, J., 1984. Ulysses and the sirens, revised edition. Cambridge: Cambridge .

Giné, X., Karlan, D., Zinman, J., 2010. Put your money where your butt is: a commitment contract for smoking cessation. American Economic Journal: Applied Economics 2, 213–35.

Gul, F., Pesendorfer, W., 2001. Temptation and self-control. Econometrica 69, 1403–1435.

Heckman, J.J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor Economics 24, 411–482.

Hoong, R., 2021. Self control and smartphone use: An experimental study of soft commitment devices. European Economic Review 140, 103924.

Houser, D., Schunk, D., Winter, J., Xiao, E., 2018. Temptation and commitment in the laboratory. Games and Economic Behavior 107, 329–344.

Kast, F., Meier, S., Pomeranz, D., 2012. Under-savers anonymous: Evidence on self-help groups and peer-pressure as a savings commitment device. NBER Working Paper 18417.

Laibson, D.I., Repetto, A., Tobacman, J., Hall, R.E., Gale, W.G., Akerlof, G.A., 1998. Self-control and saving for retirement. Brookings Papers on Economic Activity 1998, 91–196.

Lejuez, C.W., Read, J.P., Kahler, C.W., Richards, J.B., Ramsey, S.E., Stuart, G.L., Strong, D.R., Brown, R.A., 2002. Evaluation of a behavioral measure of risk taking: the balloon analogue risk task (bart). Journal of Experimental Psychology: Applied 8, 75.

Milkman, K., 2021. How to change: The science of getting from where you are to where you want to be. Portfolio.

Milkman, K.L., Minson, J.A., Volpp, K.G., 2014. Holding the hunger games hostage at the gym: An evaluation of temptation bundling. Management Science 60, 283–299.

Rankin, D.M., 1993. How to get ready for retirement: Save, save save. New York Times 3, 1993.

Read, D., Loewenstein, G., Kalyanaraman, S., 1999. Mixing virtue and vice: combining the immediacy effect and the diversification heuristic. Journal of Behavioral Decision Making 12, 257–273.

Royer, H., Stehr, M., Sydnor, J., 2015. Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a fortune-500 company. American Economic Journal: Applied Economics 7, 51–84.

Sadoff, S., Samek, A., 2019. Can interventions affect commitment demand? a field experiment on food choice. Journal of Economic Behavior and Organization 158, 90–109.

Salvatier, J., Wiecki, T.V., Fonnesbeck, C., 2016. Probabilistic programming in python using pymc3.

PeerJ Computer Science 2, e55.
Schilbach, F., 2019. Alcohol and self-control: A field experiment in india. American Economic Review 109, 1290–1322.
Schwartz, J., Wyper, L., Maroba, J., Patel, D., Ariely, D., 2014. Healthier by precommitment. Psychological Science 25, 538–546.
Thaler, R.H., Shefrin, H.M., 1981. An economic theory of self-control. Journal of Political Economy 89, 392–406.
Toussaert, S., 2018. Eliciting temptation and self-control through menu choices: A lab experiment. Econometrica 86, 859–889.
Trope, Y., Fishbach, A., 2000. Counteractive self-control in overcoming temptation. Journal of Personality and Social Psychology 79, 493.

A Additional Figures and Tables

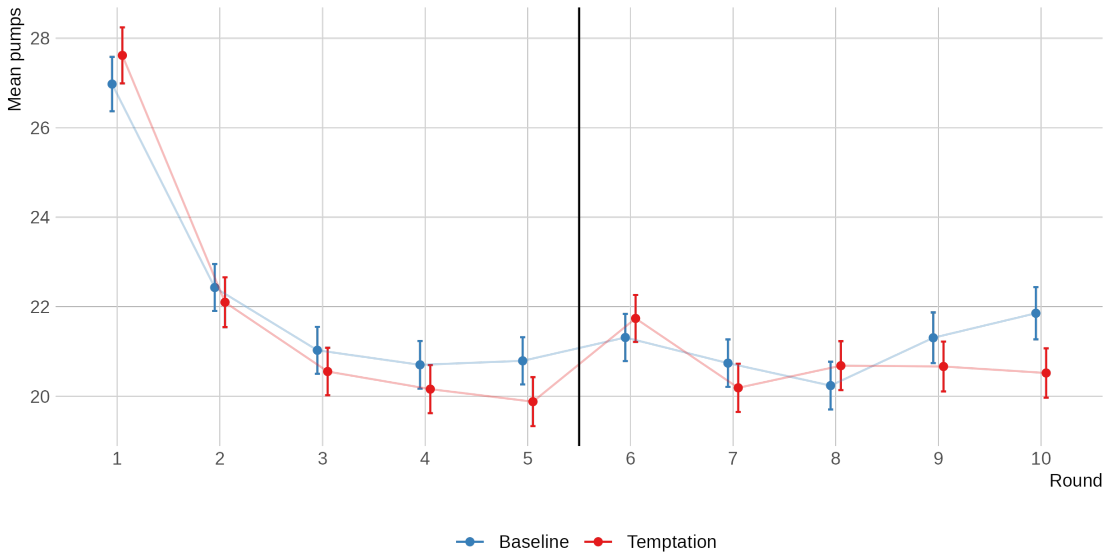

Figure 2: Mean pumps by round and treatment

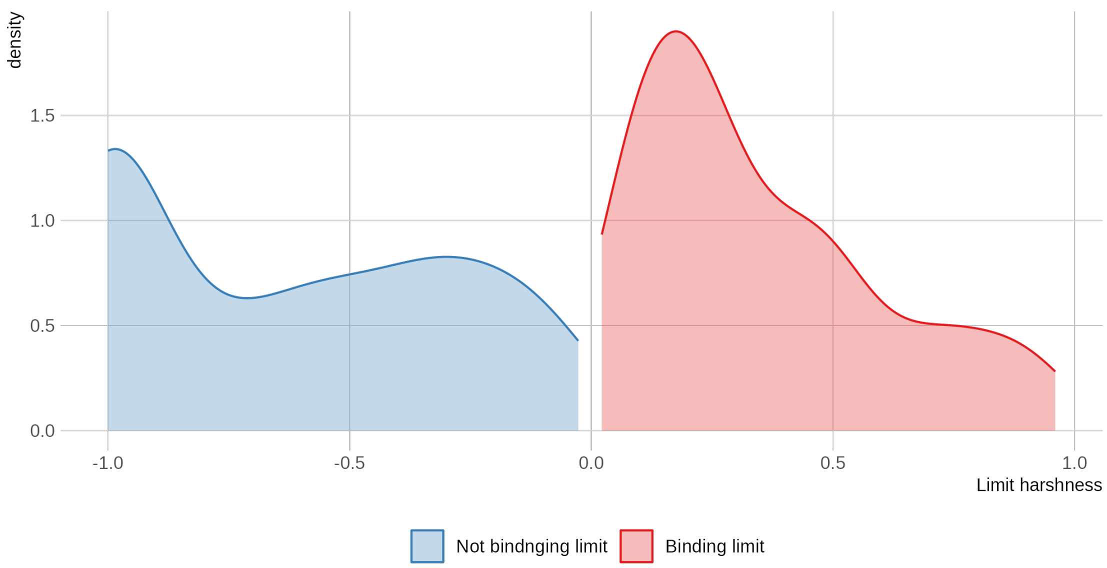

Figure 3: Harshness of the self-set limit

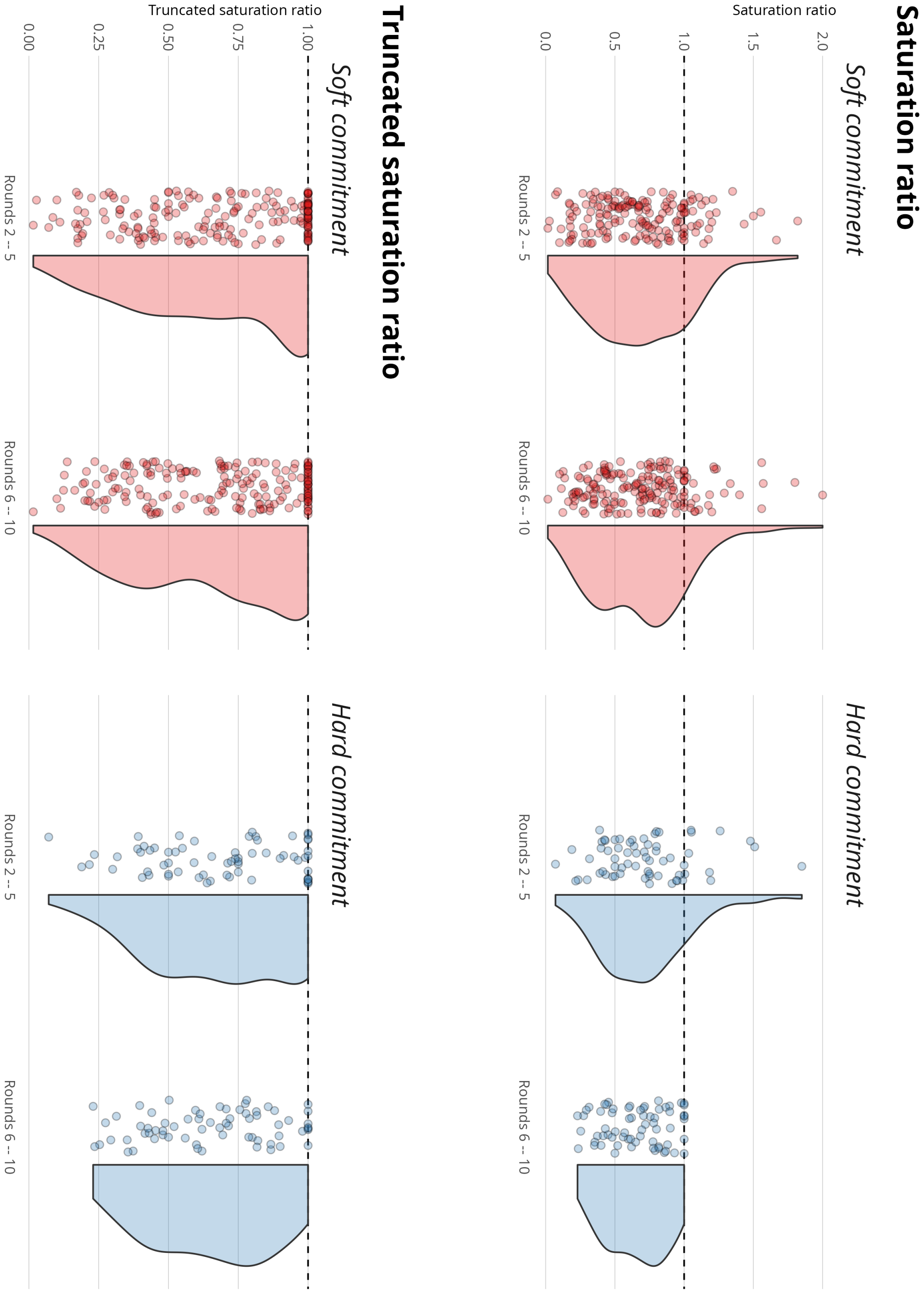

Figure 4: Saturation and Truncated saturation ratio – limited to a ratio of 2 for visualization's sake (8 excluded subjects)

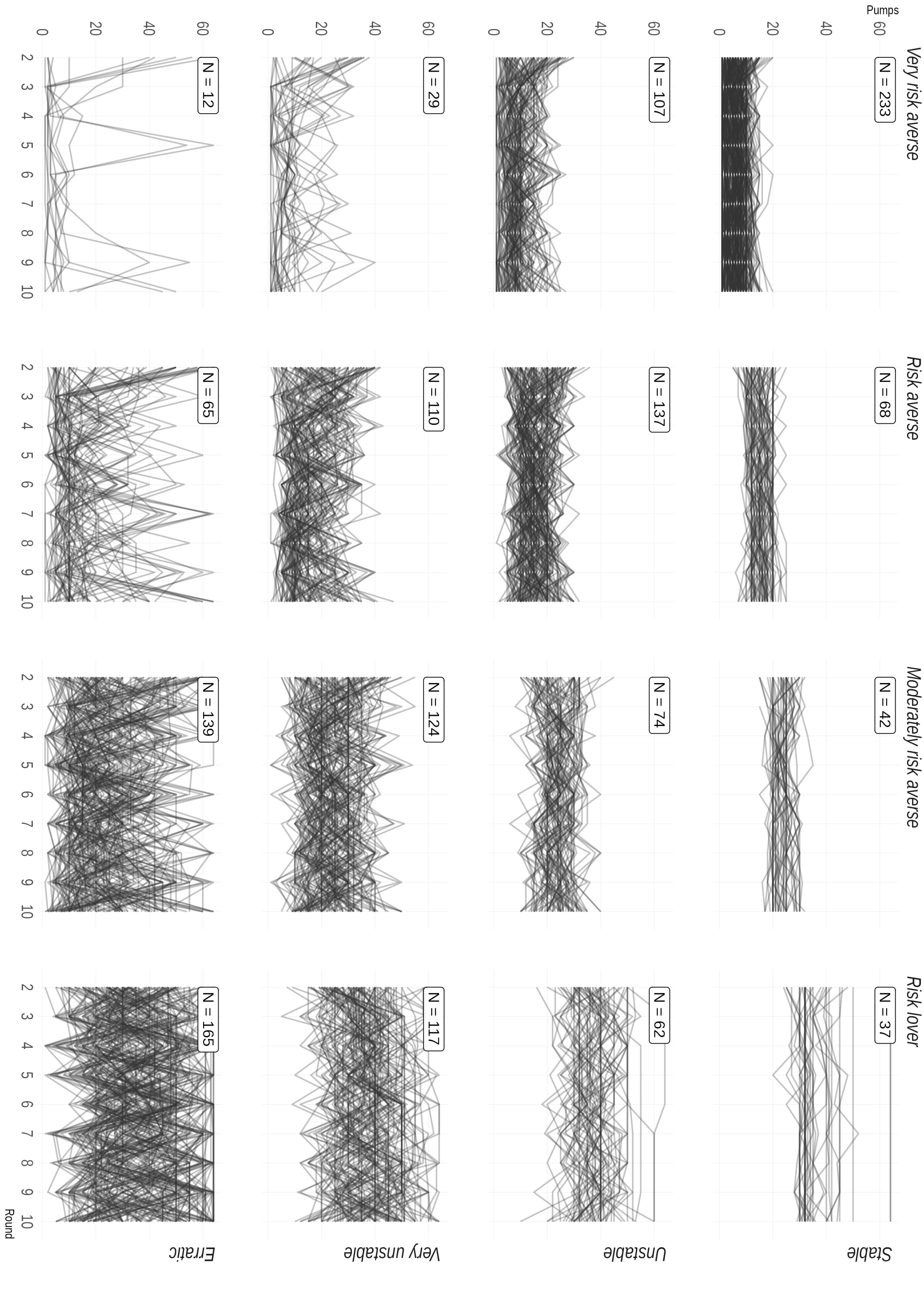

Figure 5: Number of pumps across periods, by risk attitude and stability of behavior

B Experimental instructions

(screenshots with texts translated from french)

Welcome

You are going to participate in a game that lasts about 15 minutes. At the end of the game, you will earn €5.00 and a bonus that will depend on your decisions. Your earnings will be sent to you by email as an Amazon gift certificate a few days after the end of the game.

The game has 10 round. In each round, you start with a deflated ball. Each time you inject air into the balloon, you win euros. The more the balloon is inflated, the more money you earn.

The balloon can withstand a maximum of 64 injections but may explode with each injection. The computer chooses at random the number of the injection at which the balloon explodes, i.e. a number between 1 and 64. The time at which the balloon explodes is therefore determined randomly.

You have to decide how many air injections you want to send into the balloon. Once your decision is made the balloon inflates automatically, injection by injection.

You earn €0.15 per injection if the balloon does not explode before the number of injections you choose. If the balloon explodes before, you earn €0.00.

At the end of the game, one of the 10 round will be randomly selected and your payoff for this round will be your payoff for the game.

Next

Figure 6: Screenshot of the Instructions screen

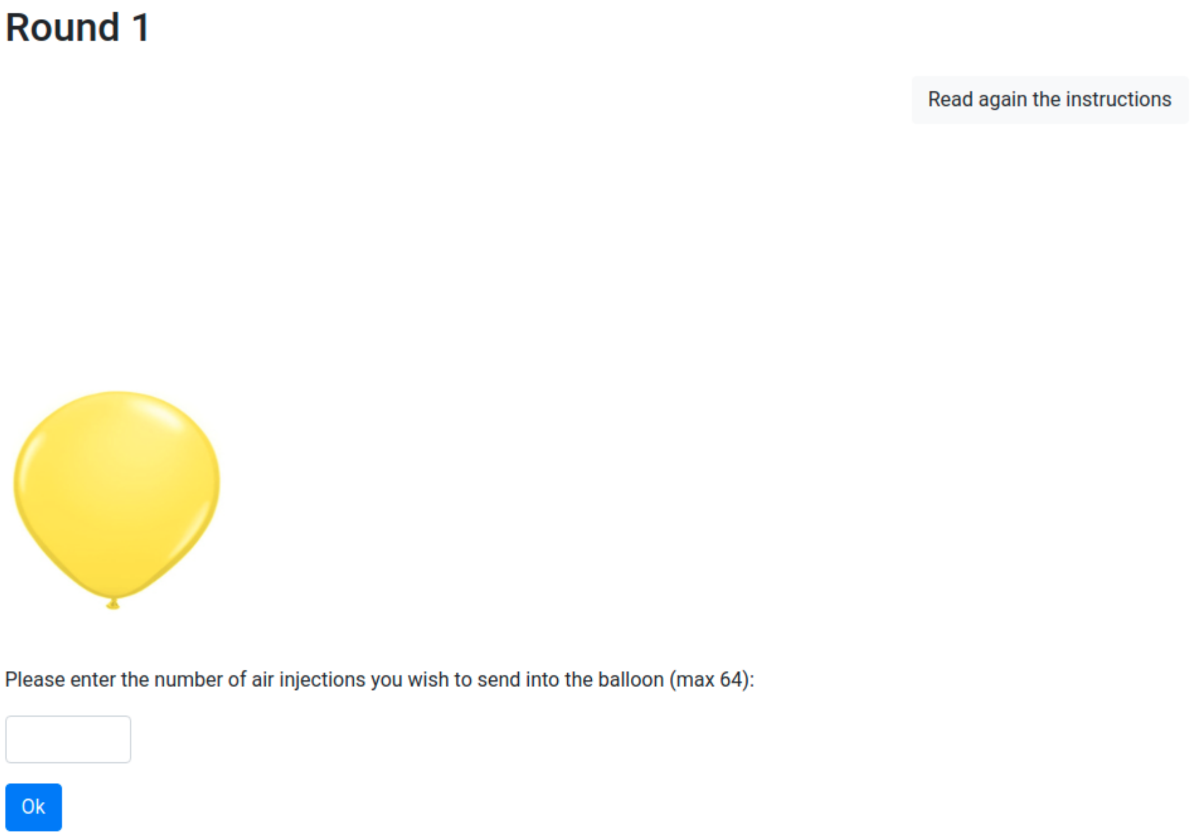

Figure 7: Screenshot of the decision screen

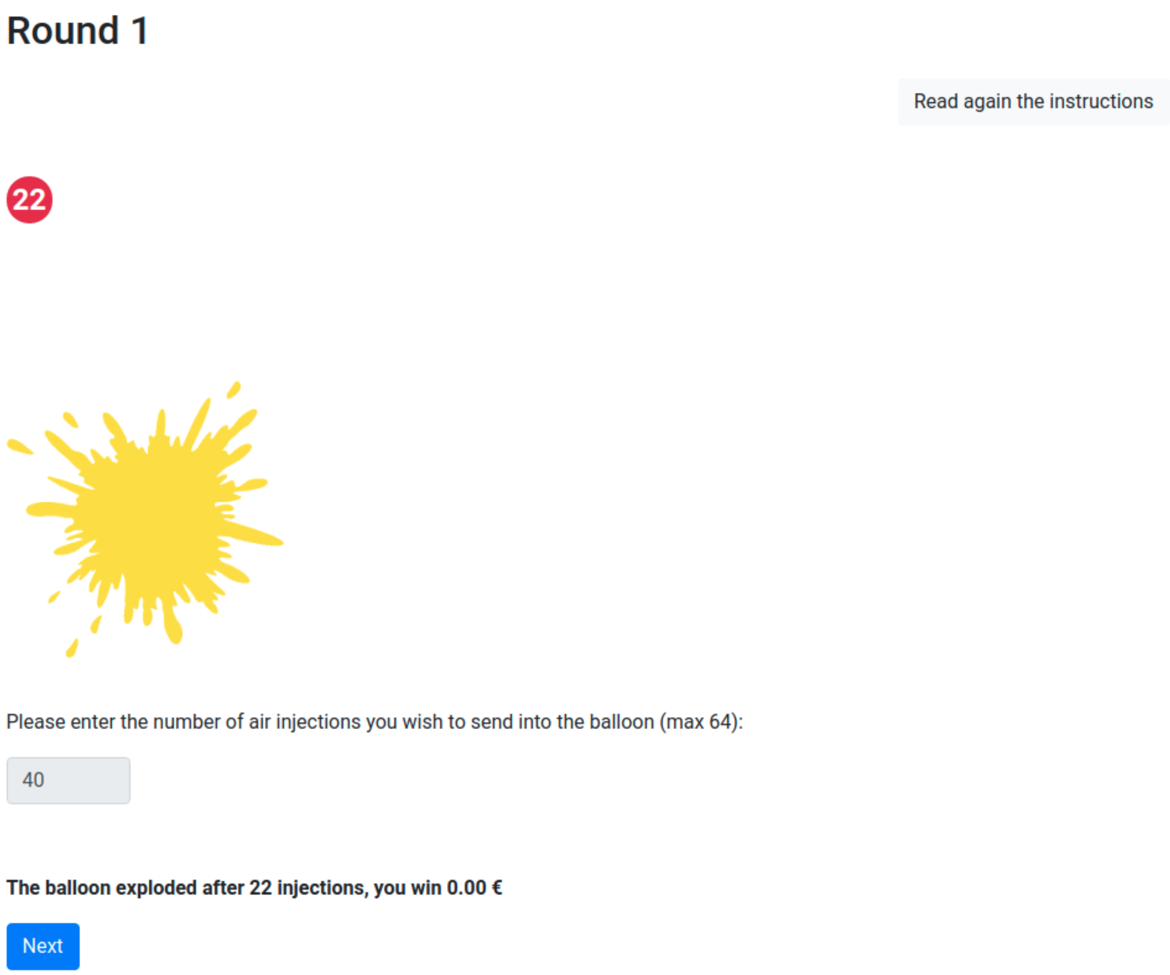

Figure 8: Screenshot of the decision screen when the balloon exploded

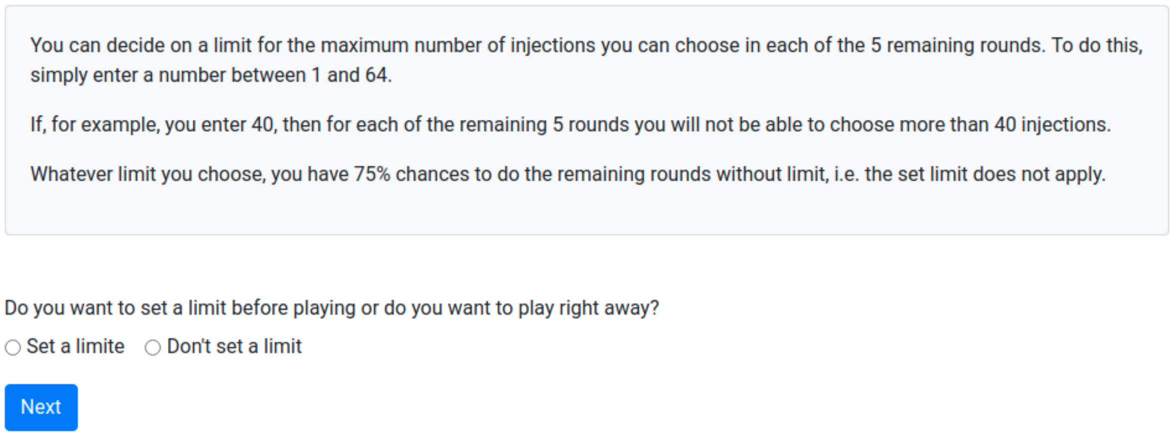

Figure 9: Screenshot of the temptation instruction screen

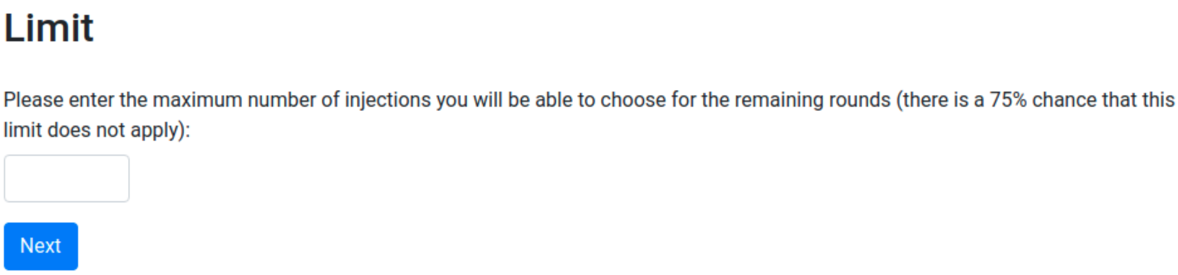

Figure 10: Screenshot of the temptation decision screen

C Questionnaire used before the BART

In the course of the last 12 months: (items presented in a random order) Response options: never, sometimes, often, very often

1. Have you made any bets that involved more money than what you could financially bear in case of a loss?
2. Have you felt that you had to make larger bets in order to feel the same excitement?
3. Have you played right after losing money (but on a different day) in order to get your money back?
4. Have you sold a personal item or borrowed money in order to gamble?
5. Have you ever felt that you may have a gambling problem?
6. Gambling has caused you health issues, including stress or anguish.
7. Has anyone criticized your gambling habits or said that you have gambling issues?

8. Your gambling habits have caused financial problems for you personally or for someone close to you.
9. Have you felt some guilt related to your gambling habits or to the consequences that your gambling habits produce.

CEE-M Working Papers[1] - 2023

WP 2023-01	Pauline Castaing & **Antoine Leblois** « Taking firms' margin targets seriously in a model of competition in supply functions »
WP 2023-02	Sylvain Chabé-Ferret, Philippe Le Coënt, Caroline Lefébvre, **Raphaële Préget**, François Salanié, **Julie Subervie** & **Sophie Thoyer** « When Nudges Backfire: Evidence from a Randomized Field Experiment to Boost Biological Pest Control »
WP 2023-03	**Adrien Coiffard, Raphaële Préget** & **Mabel Tidball** « Target versus budget reverse auctions: an online experiment using the strategy method »
WP 2023-04	**Simon Mathex, Lisette Hafkamp Ibanez** & **Raphaële Préget** « Distinguishing economic and moral compensation in the rebound effect: A theoretical and experimental approach »
WP 2023-05	**Simon Briole, Augustin Colette** & **Emmanuelle Lavaine** « The Heterogeneous Effects of Lockdown Policies on Air Pollution »
WP 2023-06	Valeria Fanghella, **Lisette Ibanez** & John Thøgersen « To request or not to request: charitable giving, social information, and spillover »
WP 2023-07	Francisco Cabo, Alain Jean-Marie & **Mabel Tidball** « Positional and conformist effects in public good provision. Strategic interaction and inertia »
WP 2023-08	**Adrien Coiffard, Raphaële Préget** & **Mabel Tidball** « Conservation auctions: an online double constraint reverse auction experiment »
WP 2023-09	Paul Bettega, Paolo Crosetto, **Dimitri Dubois** & Rustam Romaniuc « Hard vs. soft commitments: Experimental evidence from a sample of French gamblers »

[1] **CEE-M Working Papers / Contact** : laurent.garnier@inrae.fr
- RePEc https://ideas.repec.org/s/hal/wpceem.html
- HAL https://halshs.archives-ouvertes.fr/CEE-M-WP/