What Drives Demand for Loot Boxes? An Experimental Study

Simon Cordes, Markus Dertwinkel-Kalt, and Tobias Werner

Discussion Paper #2024/05

DFG Research Unit: “Consumer Preferences, Consumer Mistakes, and Firms’ Response” (FOR 5392)

What Drives Demand for Loot Boxes? An Experimental Study*

Simon Cordes Markus Dertwinkel-Kalt Tobias Werner

May 27, 2024

Abstract

The market for video games is booming, with in-game purchases accounting for a substantial share of developers' revenues. Policymakers and the general public alike are concerned that so-called "loot boxes" — lotteries that offer random rewards to be used in-game — induce consumers to overspend on video games. We provide experimental evidence suggesting that common design features of loot boxes (such as opaque odds and positively selected feedback) indeed induce overspending by inflating the belief of winning a prize. In combination, these features *double* the average willingness-to-pay for lotteries. Based on our findings, we argue for the need to regulate the design of loot boxes to protect consumers from overspending.

Keywords: Gaming, Video Games, Gambling, Loot Boxes, Microtransactions.

*We thank Mats Köster for his sustained support during the project. Furthermore, we thank Daniel Evans, Jonas Frey, Valentin Kecht, Hans-Theo Normann, Vasilisa Petrishcheva, Hannah Schildberg-Hörisch, Georg Schneider, and Florian Zimmermann for helpful comments. Cordes: University of Bonn, simon.cordes@uni-bonn.de. Dertwinkel-Kalt: University of Münster & Max Planck Institute for Research on Collective Goods, markus.dertwinkel-kalt@wiwi.uni-muenster.de. Werner: Center for Humans and Machines at the Max Planck Institute for Human Development & Düsseldorf Institute for Competition Economics, werner@mpib-berlin.mpg.de. We gratefully acknowledge financial support by the German Research Foundation (DFG project No. 462020252).

1 Introduction

The market for (online) video games has been booming in recent years, with in-game purchases accounting for a substantial share of developers' revenues. In 2020 alone, so-called "loot boxes" generated $15 billion of worldwide revenue, and projections suggest that 230 million people will spend money on loot boxes by 2025.[1] Loot boxes are digital lotteries in video games that — similar to gambling — offer *random* rewards to be used in-game. While loot boxes share many similarities with gambling (Drummond and Sauer, 2018), surprisingly little regulation is in place that would restrict their design and the way they are priced.[2] At the same time, policymakers and the general public alike are concerned that loot boxes induce consumers — in particular, those susceptible to gambling — to overspend on video games.[3]

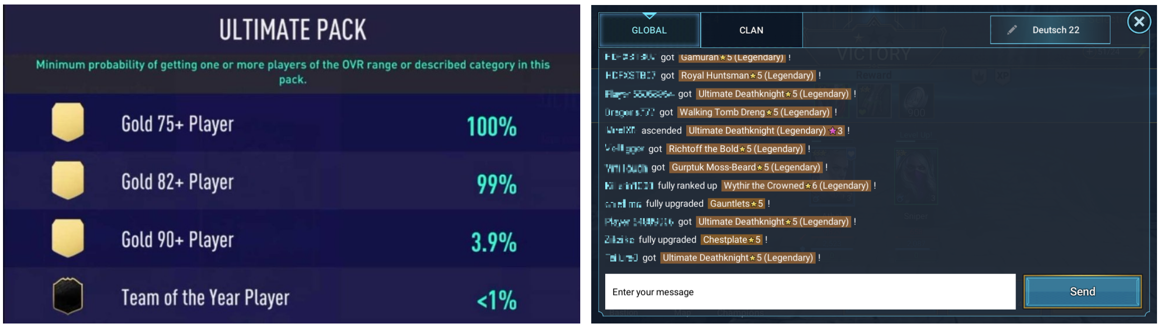


Figure 1: Typical design features, censored odds (left) and selected feedback (right), of loot boxes.

This concern is amplified by the fact that loot boxes are designed and marketed in ways that obfuscate the chances of winning different rewards. First, the odds are often censored. As a specific example, consider the football simulation *FIFA Ultimate Team*, where gamers build a team of players that vary in strength. Gamers can buy packs that offer lotteries over players. The odds, however, are provided, if at all, only for a coarse set of intervals, bunching together players of very different strengths (see the left panel of Figure 1). At the extreme, the worst player in an interval is around

[1]See, for instance, `https://www.juniperresearch.com/press/video-game-loot-boxes-to-generate-over-$20-billion` (accessed on September 16th, 2022).

[2]Recently, 20 consumer organizations from 18 European countries have suggested that loot boxes should be classified as gambling and therefore regulated (The Norwegian Consumer Council, 2022). Additionally, the Federal Trade Commission (FTC) is investigating loot boxes following concerns from U.S. legislators that they may be similar to gambling (Federal Trade Commission, 2020).

[3]See, for instance, `https://www.theguardian.com/society/2021/apr/02/video-game-loot-boxes-problem-gambling-betting-children` (accessed on September 16th, 2022).

1000 times less valuable than the best player. The Norwegian Consumer Council (2022) argues that gamers, therefore, overestimate the value of these lotteries. Moreover, the FTC asserts that false or inaccurate odds may violate the Federal Trade Commission Act, specifically Section 5, which prohibits firms' unfair or deceptive practices (Federal Trade Commission, 2020).

Second, gamers often receive highly selective feedback on the rewards other gamers have obtained. In the mobile game *Raid: Shadow Legends*,[4] for example, gamers receive a notification whenever another player wins a rare reward (see the right panel of Figure 1). This feature leads to a constant but selected stream of signals about rewards from loot boxes. As only rare rewards are reported, this provides them with a biased sample of the reward distribution.

Going further, game developers not only pay content providers (on *Youtube* or *Twitch*) to open loot boxes on their shows, but they allegedly also offer them better odds.[5] According to The Norwegian Consumer Council (2022, p.44), observing such a biased sample of the reward distribution "reinforces the players' belief that they might be similarly lucky." While one could easily imagine that both features contribute to overspending on loot boxes, there is a lack of systematic evidence supporting this claim.

In order to fill this gap, we experimentally investigate what drives the willingness-to-pay for loot boxes. In a between-subject design, we focus on the effects of censoring the odds and providing gamers with a selected sample of the reward distribution. We do so because these design features arguably do not provide any utility to gamers, while their sole purpose seems to be making consumers overspend on loot boxes. Indeed, we find evidence that both features substantially increase the willingness-to-pay for lotteries. Censoring the odds of a lottery increases a subject's willingness-to-pay compared to a baseline treatment. Also, simply providing subjects with a selective sample of the reward distribution increases their willingness-to-pay. Combining censored odds with a selected sample increases the willingness-to-pay by 100%. In a between-subjects design, we demonstrate that both features—censoring and sampling—increase the sub-

[4]In 2022, three years after its release, the game surpassed $1bn in lifetime revenue. For details, see `https://gamingonphone.com/news/raid-shadow-legends-surpasses-1-billion-in-lifetime-revenue/` (accessed on January 2nd, 2023).

[5]See, for instance, `https://gamerant.com/ftc-loot-boxes-better-odds-sponsored-streamers/` (accessed on December 22nd, 2022).

ject's willingness-to-pay by inflating her belief of winning a high reward. It is consistent with existing work that, for instance, shows that selected feedback affects economic behavior via a belief channel (e.g., Barron et al., 2019). Overall, our results suggest that the design of loot boxes — combining censored odds with selected feedback — contributes to overspending in the video game market and thus supports a case for regulating loot boxes.

We introduce our experimental design in Section 2. Subjects repeatedly state their willingness-to-pay (WTP) for different monetary lotteries with three potential prizes, one of which is zero. In a *Control* condition, we transparently describe the odds of the lotteries and do not provide additional information to the subjects. We assume that this *Control* condition identifies a subject's true WTP, and define overspending relative to this benchmark. We implement three treatments that capture the features of loot boxes discussed above. In *Censored*, subjects only learn the total probability of winning a non-zero prize, but not the exact probability of winning the highest prize. In *Sample*, we provide subjects with the full prize distribution and a selected sample thereof; that is, they observe the five highest outcomes in a sample of 400 draws. Finally, *Joint* combines both: subjects observe the censored prize distribution *and* a selected sample thereof. This last treatment resembles the current design of loot boxes most closely. Notably, our experimental design eliminates all features of loot boxes that may provide utility beyond winning a reward, such as a nice design or visual effects. Instead, we isolate the features of loot boxes that almost certainly do not affect a gamer's material utility and can thus be interpreted as inducing mistakes.

Section 3 presents our main results. Compared to the *Control* condition, the average WTP increases by roughly 45% in *Censored* and *Sample*, respectively. The average WTP doubles in *Joint* compared to the control condition. The subjects' beliefs show a similar pattern as the WTP across the different treatment conditions. Moreover, we identify the beliefs as the main channel of our treatment effects. Once we control for stated beliefs (i.e., the mediator), the average WTPs in *Censored*, *Sample*, and *Joint* do not differ significantly from that in *Control* anymore. It demonstrates that censored odds and selected feedback increase the subjects' WTPs by inflating their beliefs of winning a high prize.

In Section 4, we provide additional evidence on the underlying mechanism as well as for the relevance of our results. First, we study correlates of a survey measure on loot-box overspending. Consistent with the prior literature (e.g., Zendle and Cairns, 2018), survey measures of gambling behavior correlate with overspending on loot boxes. Controlling for these measures of gambling behavior, we still find a positive association between the average WTP for the lotteries in our experiment and survey measures of overspending on loot boxes. This speaks to the external validity of our findings and supports our view that our experimental measure of gambling behavior is informative for real-world overspending on loot boxes.

Second, we restrict the sample to decisions for which stated beliefs are "realistic" in that they are consistent with the provided information. Here, we find a precisely estimated zero difference in average beliefs between *Censored* and *Joint*. In either case, subjects tend to assign equal probabilities to the non-zero prizes. This is consistent with evidence on people naively applying a "50-50 heuristic" when being uncertain about the problem they face (e.g., Fischhoff and Bruine De Bruin, 1999).

Third, we ran a robustness experiment to address the concern that the lotteries are offered by the experimenter, not a firm, trying to maximize profits. The treatment *Info* replicates *Joint* but adds unbiased information on the reward distribution and explicitly tells the subject that the odds are not 50-50. While the additional information makes subjects less optimistic about winning the highest prize, it does not affect the average WTP. Our findings highlight the robustness of our main treatments and show that even in the presence of further unbiased information, the design features of loot boxes promote overspending.

We conclude in Section 5 by discussing tentative policy implications and challenges in their implementation. Our results highlight the complexities of regulating loot boxes. Current regulatory efforts, like those in Germany, focus on labeling video games that contain loot boxes. This may fall short as such a regulation does not address the core issue of loot box design, which we find to cause overspending. Additionally, our findings extend beyond digital loot boxes to other markets, such as trading cards and online gambling, where similar regulatory measures could help to reduce overspending.

Related Literature First, we contribute to the literature on gaming, specifically loot boxes. A series of papers have established a positive correlation between survey measures of gambling and overspending on loot boxes (Drummond and Sauer, 2018; Zendle and Cairns, 2018; Drummond et al., 2020). We strengthen this link by providing causal evidence on how key design features of loot boxes affect the WTP for lotteries. Chen et al. (2021) develop a model of optimal loot box pricing, assuming gamers maximize expected utility (EU). Gan (2022) shows that for buyers that have prospect theory preferences and are naive about them, selling a product via a loot box (that delivers the good with some constant probability in each period) represents the seller's uniquely optimal selling mechanism. Complementary to this work, we single out those features of loot boxes that are irrelevant for EU-maximizers but inflate demand by behavioral consumers.

The fact that game developers introduce "ambiguity" by censoring the odds of loot boxes connects our paper to the behavioral literature on choice under risk and ambiguity. Actually, firms should shy away from introducing payoff ambiguity as people generally have a strong distaste for ambiguity (see the large literature on ambiguity aversion building on Gilboa and Schmeidler, 1989; Schmeidler, 1989). If people are averse to ambiguity, ambiguity in payoffs should *lower* their willingness-to-pay. In our context, however, consumers appear to be ambiguity *seeking* as ambiguity *increases* their willingness-to-pay.

This observation aligns with recent literature that has challenged the universality of ambiguity aversion (e.g., Trautmann and Van De Kuilen, 2015; Kocher et al., 2018), finding that people tend to seek ambiguity when winning probabilities are small (see, e.g., Dimmock et al., 2016; Chandrasekher et al., 2022, and references therein). We enrich the existing literature by examining (i) a specific form of ambiguity that emerges from censoring the upper tail of a probability distribution, which firms can utilize to increase profits, and (ii) its application in the important domain of gaming, where firms can exploit it particularly effectively.

Notably, not only ambiguous or "vague" probabilities, but also vague outcomes have been investigated in the literature: subjects are vagueness-seeking if vagueness regards outcomes (Du and Budescu, 2005). In our loot box context, however, only vagueness in probabilities plays a role.

Further, our work relates to the literature studying real-world state lottery choice (e.g., Clotfelter and Cook, 1990; Rogers and Webley, 2001; Lockwood et al., 2199; Kachurka et al., 2021). In contrast to this strand of literature, we focus on an application to loot box design—a lottery design that uses censoring and selected sampling—methods typically not observed in real-world state lotteries. Further, the odds of winning in real-world lotteries are typically available—unlike in our setting.

Finally, we add to the literature on biased inferences from (non-)disclosed data. Empirical evidence from the lab and field suggests that individuals often draw wrong inferences from selectively disclosed data in strategic settings (Bolton et al., 2007; Koehler and Mercer, 2009; Brown et al., 2012; Benndorf et al., 2015; Deversi et al., 2021; Jin et al., 2021, 2022) and non-strategic ones (Esponda and Vespa, 2018; Barron et al., 2019; Enke, 2020; López-Pérez et al., 2022). While the former work relates to the Censored condition, the latter work resembles the Sample condition. We find that subjects naively bias censored probabilities towards a uniform distribution. By closely resembling common features of loot boxes, our design allows us to provide more nuanced insights into how to regulate their design.

2 Experimental design

2.1 Experimental setup

We develop an experimental design that allows us to test for the effects of two key features of loot boxes on the WTP for lotteries. First, the displayed odds of loot boxes are often censored. Second, gamers typically receive (positively) selected feedback on the reward distribution. We implement three treatments and one control condition to identify the effect of each feature in isolation — i.e., the effect of censored odds and selected feedback — as well as how both features interact with each other.

All subjects sequentially state their WTP for five lotteries.[6] Each lottery pays a non-

[6]Collecting several decisions within-subject has many advantages. First, it allows us to reduce individual-level decision noise that might be caused by factors orthogonal to the lottery design. Second, it increases statistical power, enabling us to measure treatment effects more precisely. Third, specific characteristics of one lottery might systematically distort participants' behavior. Using a large set of lotteries

zero prize with probability q% and zero otherwise. The non-zero prize is either 10 Coins or x Coins. Both probabilities and prizes vary across decisions: in each decision, we independently draw (without replacement) a probability $q \in \{10, 20, 30, 40, 50\}$ and a prize $x \in \{100, 120, 140, 160, 180\}$. Probability and prize pairs are drawn at the subject level, so that different subjects may observe different lotteries in a different order. The high prize of x Coins is always realized with probability 1%. For example, if $q = 10\%$ and $x = 100$, the lottery pays 0 Coins with probability 90%, 10 Coins with probability 9%, or 100 Coins with probability 1%. Before stating their WTP for a lottery, we ask subjects to state their belief on how often they would win this high prize in 100 draws (see Figure 2a for a screenshot and Section 2.2 for an interpretation).

In a between-subject design, we vary the amount of information that subjects receive on the lotteries across four conditions. We next describe these four conditions in more detail. Screenshots of the instructions and decision screens can be found in Appendix C.

Control In the control condition, subjects learn the full probability distribution of the different lotteries. More specifically, as illustrated in Figure 2a, they learn the exact probability of winning 10 Coins and x Coins, respectively. Subjects do not get any additional information on the lotteries.

Censored In the treatment *Censored*, subjects observe censored versions of the lotteries. For example in Figure 2b, they only learn the probability of receiving a non-zero prize, but not the exact probabilities of receiving 10 Coins or $x = 100$ Coins, respectively. This mimics the censoring strategies of video game designers such as *EA Sports* who do not provide gamers with the full probability distribution (see the left panel of Figure 1). Other than in *Control*, to assess the value of a lottery, subjects have to form a belief about the probability of receiving $x = 100$ Coins. Based on existing research (e.g., Fischhoff and Bruine De Bruin, 1999), we expect subjects to overestimate this probability, likely biasing it towards 50%.

allows us to rule out lottery-level effects. Reassuringly, we find no order effects in stated WTP as shown in Figure A.1.

Sample In the treatment *Sample*, subjects again learn the full probability distribution of each lottery, but on top, observe a sample from this distribution. As illustrated in Figure 2c, we present subjects the 5 highest outcomes in a sample of 400 actual draws from the lottery drawn at the subject level.[7] Notably, subjects receive transparent information on how the outcomes are chosen. This treatment is motivated by the common practice in video games of announcing rare prizes other players have obtained (see the right panel of Figure 1). Importantly, because subjects observe the full probability distribution, the sample does not contain any new information regarding the value of the lottery. Still, existing research (e.g., Barron et al., 2019) suggests that observing a series of high draws from a distribution may increase a subject's WTP.

Joint The treatment *Joint* combines both of the above: subjects observe a censored version of the lotteries together with the 5 highest outcomes in a sample of 400 draws. Unlike in *Sample*, the sample does contain information about the underlying probability distribution in this case. With censored odds, subjects arguably overestimate the low probability of winning x Coins initially. Then, if all subjects were Bayesian, the average belief upon observing the sample should decrease, moving closer to the truth. If, in contrast, subjects naively infer from a series of good draws that the lottery has to be even better than they initially thought, the average belief upon observing the sample should go up. Hence, compared to *Censored*, also the average WTP should increase.

Discussion of the Design Our design aims to contain the most essential and generalizable aspects of loot boxes while keeping the design sufficiently simple for subjects. We thus omit some aspects that are worth discussing. Firstly, firms that offer loot boxes are incentivized to choose the odds to maximize their profit. We deliberately omitted this feature in our setup. Recent experimental evidence by Deversi et al. (2021) and Jin et al. (2021) demonstrates that subjects often do not account for the strategic disclosure motives of other parties in sender-receiver games. Subjects lack skepticism regarding undisclosed information, even if these motives are transparently communicated and repeated interaction occurs. Therefore, it is plausible that video gamers do not consider the profit motive when evaluating the associated probabilities and purchasing loot boxes.

[7]For each subject, samples were drawn independently.

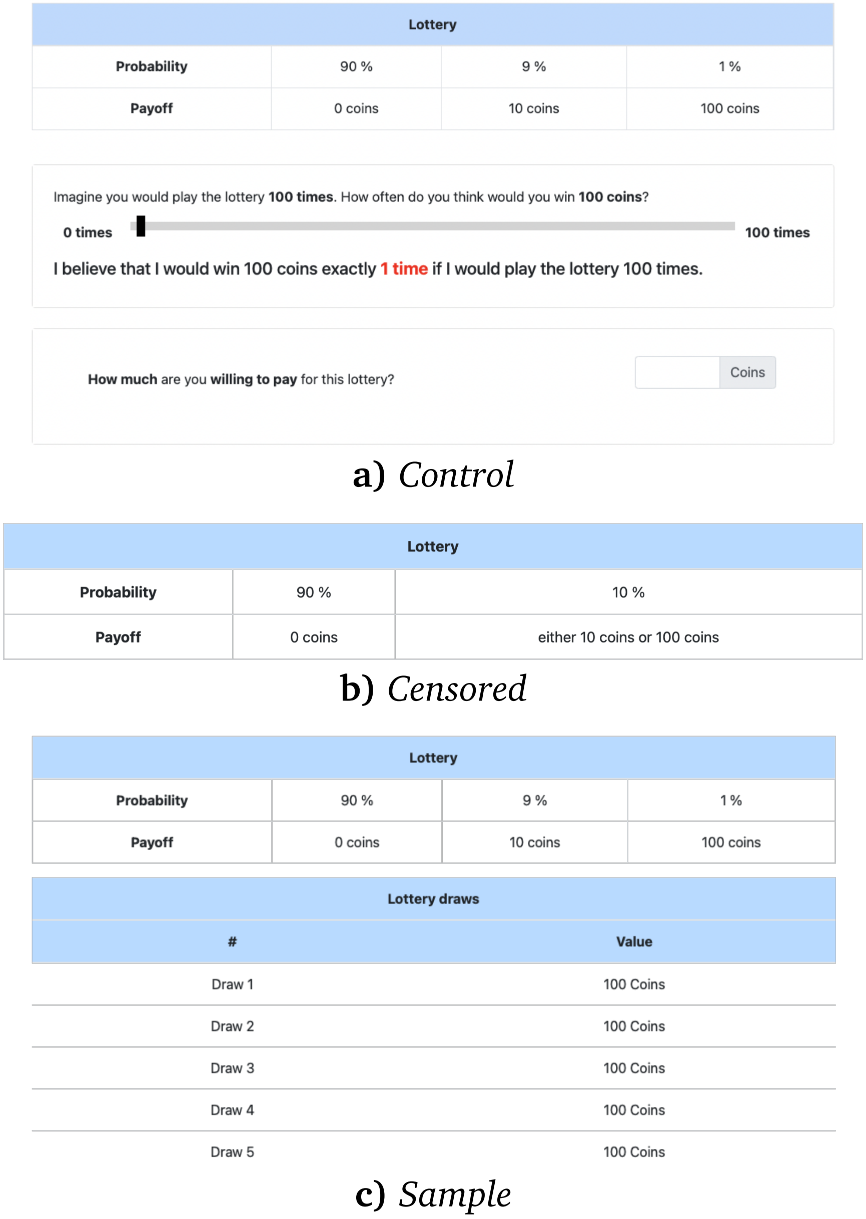


Figure 2: Screenshots of the different conditions in the experiment. (a) Control condition, (b) Censored condition, (c) Sample condition.

Additionally, video gaming is a leisure activity for most people, and the firm's incentives are likely not at the top of their minds. Secondly, unlike most video games, we chose to be explicit about the selection process in the *Sample* condition. It allows us to keep beliefs about the selection process constant. If anything, this should attenuate treatment effects, as participants should be less optimistic in the *Sample* condition compared to the real world. Consequently, we identify a lower bound treatment effect in the *Sample* condition. To further scrutinize the argument that those differences to real-world loot boxes do not invalidate our experimental design, we conduct an additional treatment discussed in Section 4.3 that provides further information to the subjects. It highlights the generalizability and robustness of our experimental design.

2.2 Conceptual framework

We sketch a simple model that motivates our experimental design as well as our analysis of the experimental data. Consider a lottery Z with a distribution G^* over prizes (such as player cards in *FIFA Ultimate Team*). The lottery is presented in a "format" f (such as our different treatments), which captures the description of the odds or the feedback provided to gamers. We assume that gamers form a subjective belief G_f over the prize distribution that depends on the format.

Willingness-to-Pay We assume that gamers aim to maximize their subjective expected utility. Denote as $u(z)$ the utility derived from prize z. A gamer's WTP for the loot box Z under format f is then given by

$$\mathbb{E}_{G_f}[u(Z)] = \mathbb{E}_{G^*}[u(Z)] + \underbrace{\mathbb{E}_{G_f}[u(Z)] - \mathbb{E}_{G^*}[u(Z)]}_{=:\phi(f)},$$

where $\phi(f)$ captures a bias that operates through the gamer's subjective belief.

To link the above to our experimental design, we impose the following central assumption.

Assumption 1. *The control condition eliminates any bias in the WTP.*

Under Assumption 1, our control condition identifies the average consumption value of a lottery. Moreover, a simple linear regression of the stated WTPs on treatment indicators identifies the average overspending on these lotteries due to censored odds and selected feedback.

Beliefs To test whether any bias in WTPs operates via beliefs, we ask subjects how often they believe to win the high prize of x Coins in 100 draws. Denote as b_i the belief of subject i for a lottery Z (under format f). We think of this belief as follows:

$$b_i = \mathbb{P}_{G_f}[Z = x] + \epsilon_i,$$

where the "noise" term ϵ_i includes implementation errors or general optimism.

Under the assumption that this noise is independent of the format, a simple linear regression of stated beliefs on treatment indicators identifies the bias in beliefs induced by censored odds and selected feedback on the reward distribution. Note that overly "optimistic" subjects can state beliefs that contradict the objective information they have.[8] We will also provide analyses separately for those subjects who state realistic beliefs.

2.3 Implementation and logistics

As is common practice for WTP elicitations, we use the BDM mechanism to incentivize subjects (Becker et al., 1964). We do not incentivize the belief elicitation, however. Recent work by Danz et al. (2022) suggests that standard incentivization techniques (such as the binarized scoring rule) systematically distort reported beliefs. Moreover, we view the belief question as an input to (or mediator of) a subject's stated WTP, which is our primary outcome of interest. To minimize anchoring effects, we use a slider without an initial value to elicit beliefs (see Figure A.9). To ensure that subjects engage with the lotteries, they could state beliefs (and afterwards WTPs) only after a 5 second delay.

The design was pre-registered in the AEA RCT registry as trial AEARCTR-0009501.[9] We collected data from 617 subjects located in the United Kingdom (UK) via *Prolific* in July 2022. The experiment consisted of 3 modules. First, we screened out inattentive participants via an attention check at the beginning of the experiment and after the instructions via comprehension questions. Second, all subjects who passed both tests stated their WTPs and beliefs for five lotteries. Third, we collected demographics and potential correlates of interest. Table A.6 suggest that randomization was successful. Screenshots of all parts of the experiment (including additional survey questions) can be found in Appendix B. Subjects earned a base fee of £1.50 for participation. In addition,

[8] For example, consider the lottery that pays 0 Coins with 90% probability, 10 Coins with 9% probability, and 100 Coins with 1% probability. Any belief larger than 10% would exceed the joint probability of winning a non-zero prize and contradict the available information.

[9] The pre-registration can be found at `https://doi.org/10.1257/rct.9501-3.0`.

1 out of 6 participants received a bonus payment depending on the WTP stated for one randomly selected lottery. Conditional on receiving a bonus, the average bonus paid was £5.35. The experiment took, on average, 11 minutes to complete.

3 Main results

Next, we turn to our main result. Figure 3 (left panel) displays the average WTP, separately for each treatment. First, we observe that all treatments increase the average WTP relative to the Control condition: the *Sample* condition by 43% ($p = 0.03$), the *Censored* condition by 45% ($p < 0.01$), and the *Joint* condition by 100% ($p < 0.01$). All treatment effects are highly statistically significant and meaningful in magnitude. Common design features of loot, therefore, induce significant overspending in our context. Table 1 Column (1) confirms this result in a regression format, and Column (2) shows robustness to adding lottery fixed effects and demographic controls. Interestingly, effect sizes in *Censored* and *Sample* are statistically indistinguishable from each other ($p = 0.78$), while the differences between *Censored* and *Joint* ($p < 0.01$) and *Sample* and *Joint* ($p = 0.01$) are statistically significant.

Result 1. *Both censoring and sampling induce overspending of almost 50%.*

Furthermore, consistent with our conceptual framework, the effect of censoring and sampling operates through the belief of winning the high prize. Figure 3 (right panel) displays the average stated belief, separately for each treatment. Relative to the *Control* condition, all treatments increase the average belief: the *Sample* condition by 3.1 p.p. ($p = 0.04$), the *Censored* condition by 12.3 p.p. ($p < 0.01$), and the *Joint* condition by 17.4 p.p. ($p < 0.01$). Table 1 Columns (4) and (5) confirm these findings in regression format. Again, Column (4) displays regression results without controls, while Column (5) adds lottery fixed effects and demographic controls. Examining the pairwise differences between the three interventions we find that differential effect sizes between *Sample* and *Censored* ($p < 0.01$), *Censored* and *Joint* ($p < 0.01$), and *Sample* and *Joint* ($p < 0.01$).

Our conceptual framework and results suggest that beliefs are the mediating channel that leads to overspending. To make this argument explicit, we control for beliefs when regressing the WTPs on the treatment indicators. Under the null of our conceptual framework, beliefs fully mediate the treatment effects on WTP. Thus, including the potential mediator in our regression allows us to test this hypothesis of full mediation. The results of this mediator analysis are in Column (3) of Table 1. Beliefs are statistically significant at any conventional level ($p < 0.01$) and are positively related to the WTPs. Furthermore, the effect of Censored, Sample, and Joint becomes insignificant. This cleanly demonstrates that censored odds and biased samples increase the willingness-to-pay via a beliefs channel.

Result 2. *Inflated beliefs drive participants' overspending.*

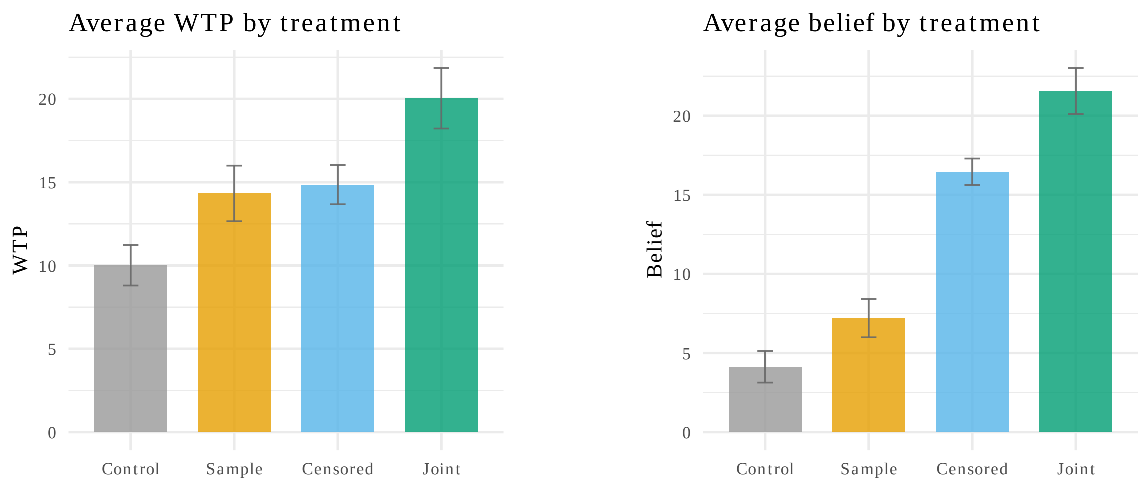


Figure 3: Average willingness-to-pay and beliefs by treatment. We include all subjects that finished the experiment. WTP is the willingness to pay for a lottery. Belief is a subjective estimate of the frequency of winning the lottery in 100 independent draws. Whiskers are the standard error of the mean.

One might speculate that overspending on loot boxes could be driven by self-control problems, as people with self-control problems tend to act impulsively and give in to immediate desires. The excitement and unpredictability of loot boxes can trigger this impulsive behavior, leading to excessive spending on loot boxes. Using the survey measure proposed in Tangney et al. (2004) — that tests for self-control problems at the hand of 13 established psychological questions — we find no (significant) relation between self-control problems and loot box overspending in our experiment (see Table A.2 in the Appendix for details).

Table 1: Regression results — main specification

	WTP			Belief	
	No controls	Controls	Controls + Belief	No controls	Controls
	(1)	(2)	(3)	(4)	(5)
Censored	4.84***	4.70***	-1.18	12.3***	12.4***
	(1.55)	(1.51)	(1.75)	(1.18)	(1.17)
Sample	4.31**	4.28**	2.88	3.08**	2.94*
	(2.00)	(1.99)	(1.79)	(1.55)	(1.56)
Joint	10.0***	10.2***	1.85	17.4***	17.6***
	(1.99)	(2.00)	(2.25)	(1.62)	(1.63)
Belief			0.476***		
			(0.068)		
Observations	3,085	3,085	3,085	3,085	3,085
Lottery FE		x	x		x

Notes: Results from ordinary least squares (OLS) regressions on treatment dummies. The outcome variable in columns (1), (2) and (3) is the willingness to pay and in (4) and (5) the beliefs. Columns (1) and (4) do not include control variables. Columns (2), (3) and (5) control for age, gender and monthly available budget, as well as lottery fixed effects. Standard errors clustered at the subject level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4 Additional results

4.1 Correlates of overspending on loot boxes

In order to get an idea about the external validity of our findings, we also examine whether there is a relationship between our experimental measures and real-world loot-box overspending. For this purpose, we asked subjects at the end of the experiment a series of questions on their daily usage of video games and knowledge of loot boxes. Specifically, we elicited (i) whether they know what loot boxes are and if the answer is positive, (ii) how much money they spend on loot boxes per month, and (iii) whether they have ever spent more money on loot boxes than they initially planned to. In our sample, 59% play video games, documenting that gaming is a pervasive phenomenon throughout society. Our subjects spend, on average, 1.2 hours playing video games daily. Moreover, 69% of our subjects know what loot boxes are. Conditional on knowing what loot boxes are, the average participant spends $15 on loot boxes per year. This average shrouds important heterogeneity: 75% of participants do not spend at all, while the 95th percentile spender invests \$100 or more per year. A substantial minority (11%) states to have ever spent more on loot boxes than initially planned.

Furthermore, we elicit the gambling behavior of the subjects using the nine survey questions from the Problem Gambling Severity Index (Ferris and Wynne, 2001). Consistent with the prior literature (Zendle and Cairns, 2018), we find a positive correlation ($\rho = 0.26$, $p < 0.01$) between loot-box usage and survey measures of gambling behavior.[10] Overall, our sample thus seems well-suited to study the demand for loot boxes.

Table 2: Regression results — predictors for real world overspending

	=1 if subject overspent on loot boxes			
	(1)	(2)	(3)	(4)
WTP	0.004***			0.003**
	(0.001)			(0.001)
Belief		0.002		0.0001
		(0.001)		(0.002)
Gambling Score			0.869***	0.823***
			(0.271)	(0.260)
Observations	425	425	425	425

Notes: Results from linear probability regression model of a dummy that is one if the subject state to have spent more than planned on loot boxes in the real world. WTP is the willingness to pay. Belief is a subjective estimate of the frequency of winning the lottery in 100 independent draws. Gambling score is the score from a self reported gambling questionnaire, scaled from 0 to 1. All variables are subject averages. We include subjects who have stated that to know what loot boxes are. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We are mainly interested in whether subjects who have overspent on loot boxes in the past differ in systematic ways from those who did not and whether our experimental measures pick up part of this variation. We asked subjects: “Have you ever spent more than you planned to on loot boxes?” We then regress a dummy variable that takes a value of one if the answer was yes and a value of zero otherwise on a subject’s average WTP and stated beliefs (see Table 2). We find a positive association between a subject’s WTP and overspending on loot boxes. An increase in the average WTP from the $5th$ percentile to the median (an increase of 10 coins) is associated with an increase in the probability to overspend of 3.28 percentage points. While admittedly small, the effect

[10]We define subjects as loot-box users if they have either (a) ever spent more than they planned to on loot boxes or (b) have positive yearly spending on loot boxes.

remains significant even when controlling for survey measures of gambling behavior. Hence, the average WTP for monetary lotteries picks up part of the variation in loot-box overspending that these survey measures cannot explain. On the other hand, stated beliefs are not correlated with the tendency to overspend on loot boxes.

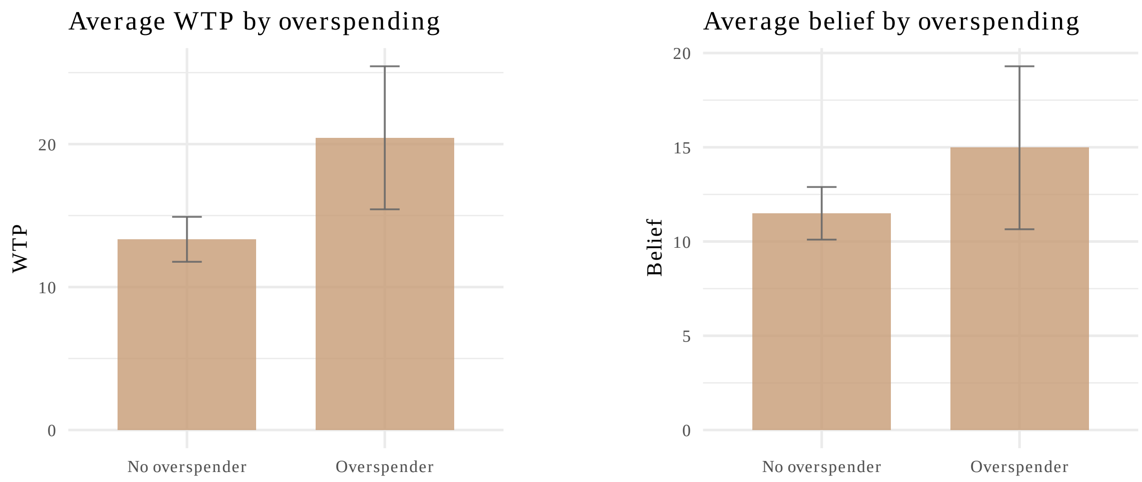


Figure 4: Average willingness-to-pay and beliefs for subjects that overspent on loot boxes in the past. WTP is the willingness to pay for a lottery. Belief is a subjective estimate of the frequency of winning the lottery in 100 independent draws. Whiskers are the standard error of the mean.

Figure 4 visualizes those results. The WTPs are higher for subjects who overspent on loot boxes in the past than those who have not: the average WTP for overspenders is 20.44, compared to a WTP for non-overspenders of 13.34 ($p < 0.01$). Beliefs show a similar pattern. The average stated belief for overspenders is 14.97, while the average belief for non-overspenders is 11.50 ($p = 0.12$). Note that the sample size of this analysis is comparably low ($N = 66$ for overspenders), such that the results should be interpreted as suggestive evidence. This evidence supports the argument that overspending in the experiment reflects overspending on actual loot boxes, highlighting the external validity to the study's findings.

4.2 Treatment effects under realistic beliefs

Next, we restrict our sample to decisions in which subjects stated "realistic" beliefs that could be interpreted as conditional probabilities consistent with the information they observed. Importantly, subjects do not directly report the probability of winning the highest lottery prize. However, they provide an estimated number of wins out of 100 plays, leading to potential overestimation from being overly optimistic or perceived luck. Interpreting this belief as a winning probability leads to beliefs that appear to be unrealistic. Consider, for example, a lottery that pays 10 Coins with 39% probability and 100 Coins with 1% probability. If subjects observe this distribution (in *Control* and *Sample*), the only realistic belief is exactly 1%. In *Control*, subjects stated the realistic belief in 85% of the decisions, while in *Sample*, they stated it 62% of the time. In treatments *Censored* and *Joint*, subjects only learn the probability with which the lottery pays a non-zero prize; here, 40%. Hence, any belief between 0% and 40% is realistic in this case. This leaves us with 97% of the decisions in *Censored* and 86% of the decisions in *Joint*.

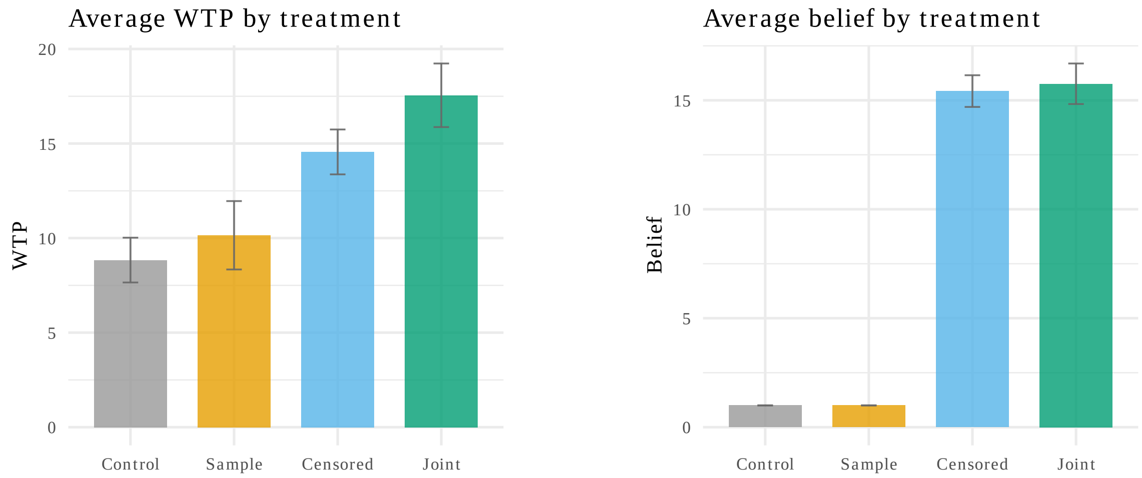


Figure 5: Average willingness-to-pay and beliefs by treatment in the main experiment. We include all subjects that finished the experiment. We exclude all decisions in which a subject stated a belief that is larger than the probability of winning a non-zero prize. WTP is the willingness to pay for a lottery. Belief is a subjective estimate of the frequency of winning the lottery in 100 independent draws. Whiskers are the standard error of the mean.

Figure 5 and Table A.3 show the results. By construction of the sample, stated beliefs are identical in *Control* and *Sample*. More interestingly, conditional on stating a realistic belief, there is also no significant difference in stated beliefs across *Censored* and *Joint*. At the same time, even realistic beliefs are massively inflated (compared to the truth) in these treatments. On average, subjects assign almost equal probabilities to the events of receiving 10 Coins and x Coins. To see this, consider the right panel of Figure 5 and note that the highest average realistic belief across all lotteries is 30%, with a mid-point of 15%.[11] Next, we calculate the difference between the realistic beliefs and the mid-point

[11]We consider lotteries that have a non-zero winning probability of $q \in \{10, 20, 30, 40, 50\}$.

of the respective lotteries for each subject. The average difference in *Censored* and *Joint* is only -0.44, which is not statistically significantly different from zero (see Table A.4). Furthermore, using a two-one-sided t-tests procedure for equivalence testing, we can reject the hypothesis that this difference falls outside of a five p.p. interval around zero ($p < 0.01$ for all tests).

This is consistent with existing evidence on people — when being uncertain — biasing probabilities towards a uniform distribution (e.g., Fischhoff and Bruine De Bruin, 1999; Dimmock et al., 2016). Across the four conditions, the average WTPs follow the same qualitative patterns as before. However, neither the difference in WTP between *Control* and *Sample* nor the difference in WTP between *Censored* and *Joint* is statistically significant in this smaller subsample.

4.3 Robustness experiment

One caveat of our design is that the lotteries are offered by the experimenter, not a firm trying to maximize profits. This might affect the inferences that subjects draw from observing censored probabilities or a selected sample, and it might result in higher WTPs compared to a market setting. We address this concern in a second experiment.[12]

In the treatment *Info*, we make it clear to subjects that the outcomes with censored probabilities are *not* equally likely, and we further provide them unbiased information on the probability distribution (on top of a selected sample). A total number of 414 subjects completed the experiment on *Prolific* in November 2022. The instructions, screenshots of the decision screens, and details on the implementation can be found in Appendix B.

Figure 6 summarizes our findings. The additional information significantly decreases the average (conditional) belief of winning the high prize compared to *Joint* ($p < 0.01$). While the WTP in *Info* is also slightly below the one in *Joint*, the effect is not significant at the 10% level ($p = 0.69$). Importantly, both the belief and the WTP are significantly higher in *Info* compared to *Control* (see Table A.5 in the Appendix). Overall, it highlights the robustness of our results and points to the significance of the design features of loot boxes on WTPs.

[12]The pre-registration is available at `https://doi.org/10.1257/rct.10506-1.0`.

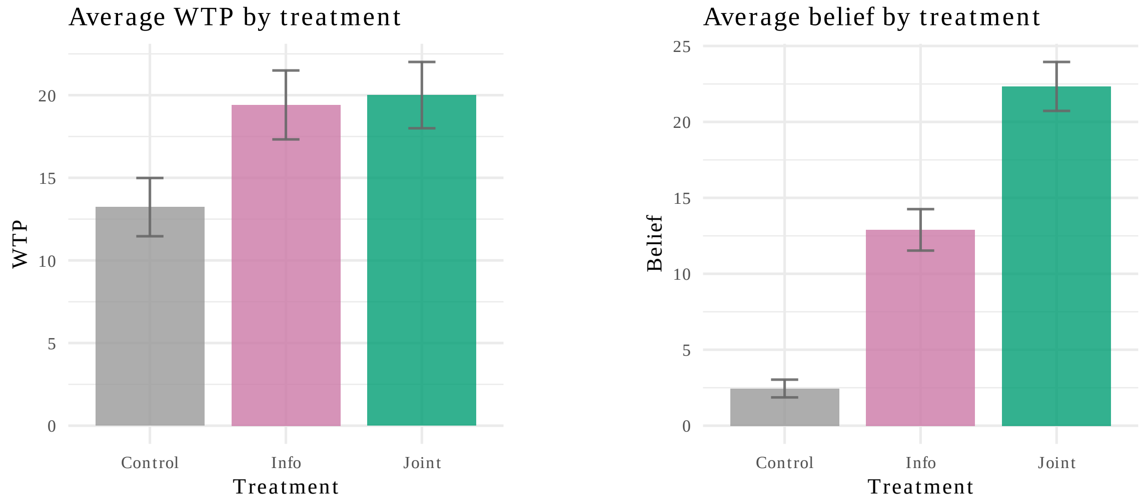


Figure 6: Average willingness-to-pay and beliefs by treatment in the Robustness Experiment. We include all subjects that finished the Robustness Experiment. WTP is the willingness to pay for a lottery. Belief is a subjective estimate of the frequency of winning the lottery in 100 independent draws. Whiskers are the standard error of the mean.

5 Conclusion

In a laboratory experiment, we document that two main characteristics of loot boxes, namely, censored odds and selective feedback, increase the demand and the willingness-to-pay for lotteries, arguably without providing additional utility for consumers.

We think that it is plausible to assume that we have estimated a lower bound on the distortion of loot-box demand. Real-world loot boxes are embedded into enjoyable video games and come along with contextual features that could enhance a gamer's utility. In such an environment, censored odds and selective feedback can be expected to increase gamers' demand likewise, if not even to a larger degree, which means that we estimate a lower bound on the demand-enhancing effect of loot box characteristics. For example, if gamers get distracted by fancy visual effects, they might become more prone to make statistical errors. A nice design of loot boxes might also result in a more favorable view of the game developer and the odds it offers. Our design, which abstracts from features of loot boxes that directly generate utility, then allows us to estimate a lower bound on the bias in loot-box demand.

Our results support a case for regulating loot-box design, but it is not apparent what regulation would be effective. Current plans for regulation in Germany include labels for games with loot boxes.[13] However, our results show that the design of loot boxes, rather than the random rewards they provide, encourages players to overspend. Hence, this regulation may not be effective in reducing overspending. While it should be easy to enforce a transparent display of odds, it is not clear that gamers will use this information when making their purchase decisions. Our robustness experiment, for instance, suggests that additional information may not affect WTPs. Moreover, even when learning the full probability distribution over *many* prizes, gamers might not act on it because it is simply too much information to be considered. Instead, regulators must find ways of communicating the odds of loot boxes in an easily understandable way.[14] Preventing gamers from being confronted with selected feedback on the reward distribution comes with the additional challenge that it is not only game developers who provide it to gamers. Even if game developers are not allowed to announce prizes others have won selectively, gamers may get similar (biased) feedback from talking to their peers or watching their videos on *Youtube* or *Twitch*.

Our insights are not restricted to loot boxes but carry over to other offline and online markets. For instance, trading cards (such as Panini or Pokémon cards) and many types of online gambling use and arguably exploit the same features as loot boxes. Given our findings, similar regulations to loot boxes should be imposed on those markets to protect consumers from severe overspending. Further, our findings can also be applied to other settings, e.g., state lotteries, where the odds typically are not saliently presented to participants. However, the media often reports when a participant won the lottery—oftentimes even with articles that feature the lucky winners. Our Sample feature can be interpreted as mirroring this behavior—only information about the highest payoffs is presented to consumers. An implication of our experiment is then that news coverage about lottery winners might have harmful spillover effects on lottery consumption, as new consumers will likely overspend.

[13]See, for instance, `https://usk.de/jugendschutzgesetz-aktualisiert-usk-bereitet-sich-auf-aenderungen-vor/` (accessed on September 19th, 2022).

[14]Alternatively, regulators could ban loot boxes altogether. As the case of Belgium shows, however, such a ban can only work if regulators also introduce proper enforcement mechanisms (Xiao, 2022).

References

Barron, Kai, Steffen Huck, and Philippe Jehiel (2019) "Everyday econometricians: Selection neglect and overoptimism when learning from others," *Working Paper* (SP II 2019-301).

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak (1964) "Measuring utility by a single-response sequential method," *Behavioral Science*, 9 (3), 226–232.

Benndorf, Volker, Dorothea Kübler, and Hans-Theo Normann (2015) "Privacy concerns, voluntary disclosure of information, and unraveling: An experiment," *European Economic Review*, 75, 43–59.

Bolton, Patrick, Xavier Freixas, and Joel Shapiro (2007) "Conflicts of interest, information provision, and competition in the financial services industry," *Journal of Financial Economics*, 85 (2), 297–330.

Brown, Alexander L., Colin F. Camerer, and Dan Lovallo (2012) "To Review or Not to Review? Limited Strategic Thinking at the Movie Box Office," *American Economic Journal: Microeconomics*, 4 (2), 1–26.

Chandrasekher, Madhav, Mira Frick, Ryota Iijima, and Yves Le Yaouanq (2022) "Dual-Self Representations of Ambiguity Preferences," *Econometrica*, 90 (3), 1029–1061.

Chen, Ningyuan, Adam N. Elmachtoub, Michael L. Hamilton, and Xiao Lei (2021) "Loot Box Pricing and Design," *Management Science*, 67 (8), 4809–4825.

Clotfelter, Charles T. and Philip J. Cook (1990) "On the economics of state lotteries," *Journal of Economic Perspectives*, 4 (4), 105–119.

Danz, David, Lise Vesterlund, and Alistair J Wilson (2022) "Belief elicitation and behavioral incentive compatibility," *American Economic Review*, 112 (9), 2851–2883.

Deversi, Marvin, Alessandro Ispano, and Peter Schwardmann (2021) "Spin doctors: An experiment on vague disclosure," *European Economic Review*, 139, 103872.

Dimmock, Stephen G, Roy Kouwenberg, and Peter P Wakker (2016) "Ambiguity attitudes in a large representative sample," *Management Science*, 62 (5), 1363–1380.

Drummond, Aaron and James D. Sauer (2018) "Video game loot boxes are psychologically akin to gambling," *Nature human behaviour*, 2 (8), 530–532.

Drummond, Aaron, James D. Sauer, Lauren C. Hall, David Zendle, and Malcolm R. Loudon (2020) "Why loot boxes could be regulated as gambling," *Nature human be-*

haviour, 4 (10), 986–988.

Du, Ning and David V Budescu (2005) "The effects of imprecise probabilities and outcomes in evaluating investment options," *Management Science*, 51 (12), 1791–1803.

Enke, Benjamin (2020) "What You See Is All There Is," *The Quarterly Journal of Economics*, 135 (3), 1363–1398.

Esponda, Ignacio and Emanuel Vespa (2018) "Endogenous sample selection: A laboratory study," *Quantitative Economics*, 9 (1), 183–216.

Federal Trade Commission (2020) "FTC Video Game Loot Box Workshop," *Federal Trade Commission*, `https://www.ftc.gov/system/files/documents/reports/staffperspective-paper-loot-box-workshop/loot_box_workshop_staff_perspective.pdf`, Available at: `https://www.ftc.gov/system/files/documents/reports/staffperspective-paper-loot-box-workshop/loot_box_workshop_staff_perspective.pdf`.

Ferris, Jacqueline Ann and Harold James Wynne (2001) *The Canadian problem gambling index*: Canadian Centre on substance abuse Ottawa, ON.

Fischhoff, Baruch and Wändi Bruine De Bruin (1999) "Fifty-Fifty=50%?" *Journal of Behavioral Decision Making*, 12 (2), 149–163.

Gan, Tan (2022) "Gacha Game: When Prospect Theory Meets Optimal Pricing," *arXiv preprint arXiv:2208.03602*.

Gilboa, Itzhak and David Schmeidler (1989) "Maxmin expected utility with non-unique prior," *Journal of mathematical economics*, 18 (2), 141–153.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin (2021) "Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure," *American Economic Journal: Microeconomics*, 13 (2), 141–173.

——— (2022) "Complex disclosure," *Management Science*, 68 (5), 3236–3261.

Kachurka, Raman, Michał Krawczyk, and Joanna Rachubik (2021) "State lottery in the lab: an experiment in external validity," *Experimental Economics*, 1–25.

Kocher, Martin G, Amrei Marie Lahno, and Stefan T Trautmann (2018) "Ambiguity aversion is not universal," *European Economic Review*, 101, 268–283.

Koehler, Jonathan J. and Molly Mercer (2009) "Selection Neglect in Mutual Fund Advertisements," *Management Science*, 55 (7), 1107–1121.

Lockwood, Benjamin, Hunt Allcott, Dmitry Taubinsky, and Afras Y. Sial, "What drives demand for state-run lotteries? evidence and welfare implications."

López-Pérez, Raúl, Ágnes Pintér, and Rocío Sánchez-Mangas (2022) "Some conditions (not) affecting selection neglect: Evidence from the lab," *Journal of Economic Behavior & Organization*, 195, 140–157.

Rogers, Paul and Paul Webley (2001) ""It could be us!": Cognitive and social psychological factors in UK national lottery play," *Applied Psychology*, 50 (1), 181–199.

Schmeidler, David (1989) "Subjective probability and expected utility without additivity," *Econometrica: Journal of the Econometric Society*, 571–587.

Tangney, JP, RF Baumeister, and AL Boone (2004) "High self-control predicts good adjustment, less pathology, better grades, and interpersonal success.," *Journal of Personality*, 72 (2), 271–324.

The Norwegian Consumer Council (2022) "Insert Coin: How the Gaming Industry Exploits Consumers Using Loot Boxes,"Technical report, `https://fil.forbrukerradet.no/wp-content/uploads/2022/05/2022-05-31-insert-coin-publish.p`.

Trautmann, Stefan T and Gijs Van De Kuilen (2015) "Ambiguity attitudes," *The Wiley Blackwell handbook of judgment and decision making*, 2, 89–116.

Xiao, Leon Y (2022) "Breaking Ban: Belgium's ineffective gambling law regulation of video game loot boxes," `https://osf.io/hnd7w/`.

Zendle, David and Paul Cairns (2018) "Video game loot boxes are linked to problem gambling: Results of a large-scale survey," *PloS one*, 13 (11), e0206767.

A Additional Tables and Figures

Table A.1: Regression results — Difference relative to *Joint*

	WTP			Belief	
	No controls	Controls	Controls + Belief	No controls	Controls
	(1)	(2)	(3)	(4)	(5)
Censored	-5.18***	-5.50***	-2.83	-5.31***	-5.11***
	(1.91)	(1.90)	(1.84)	(1.35)	(1.37)
Sample	-5.71**	-5.97***	1.47	-14.8***	-14.4***
	(2.29)	(2.24)	(2.31)	(1.68)	(1.71)
Belief			0.504***		
			(0.070)		
Observations	2,330	2,330	2,330	2,330	2,330
Lottery FE		x	x	x	

Notes: Results from ordinary least squares (OLS) regressions on treatment dummies. The outcome variable in columns (1), (2) and (3) is the willingness to pay and in (4) and (5) the beliefs. We restrict the sample to *Joint*, *Sample* and *Censored*. Columns (1) and (4) do not include control variables. Columns (2), (3) and (5) control for age, gender and monthly available budget, as well as lottery fixed effects. Standard errors clustered at the subject level in parentheses. * $p < 0.1$, $^{**}p < 0.05$, *** $p < 0.01$

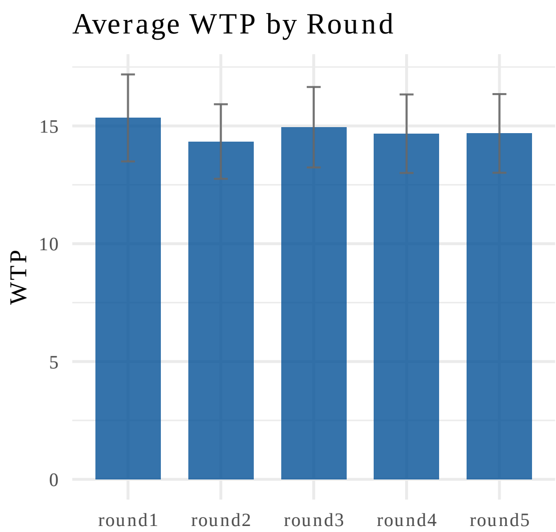


Figure A.1: Average willingness-to-pay, separately for each round. WTP is the willingness to pay for a lottery. Whiskers are the standard error of the mean.

Table A.2: Regression results — self-control

	WTP			Belief	
	No Controls	Controls	Controls + Beliefs	No Controls	Controls
	(1)	(2)	(3)	(4)	(5)
Censored	4.97***	4.84***	-1.03	12.3***	12.4***
	(1.56)	(1.53)	(1.76)	(1.19)	(1.18)
Sample	4.33**	4.30**	2.90	3.07**	2.94*
	(2.00)	(2.00)	(1.80)	(1.56)	(1.56)
Joint	10.1***	10.3***	1.96	17.4***	17.6***
	(2.00)	(2.02)	(2.26)	(1.63)	(1.64)
Self control	3.20	4.00	4.15	-0.719	-0.316
	(3.73)	(3.93)	(3.50)	(2.99)	(3.10)
Belief			0.476***		
			(0.067)		
Observations	3,085	3,085	3,085	3,085	3,085
Lottery FE		x	x		x

Notes: Results from ordinary least squares (OLS) regressions on the self control score by Tangney et al. (2004). We sum all survey items and standardize the score to be between 0 and 1. Larger values indicate a higher level of self-control. The outcome variable in Columns (1) - (3) is the willingness to pay. Columns (4) and (5) consider the belief as an outcome variable. Belief is a subjective estimate of the frequency of winning the lottery in 100 independent draws. We control for age, gender and monthly available budget, as well as lottery fixed effects. Standard errors clustered at the subject level in parentheses. * $p < 0.1$, $^{**}p < 0.05$, *** $p < 0.01$

Table A.3: Regression results — main specification — realistic beliefs

	WTP			Belief	
	No controls	Controls	Controls + Beliefs	No controls	Controls
	(1)	(2)	(3)	(4)	(5)
Censored	5.35***	5.34***	-1.75	13.7***	13.7***
	(1.43)	(1.38)	(1.68)	(0.466)	(0.462)
Sample	2.46	2.47	1.89	1.24***	1.12***
	(1.72)	(1.72)	(1.70)	(0.366)	(0.347)
Joint	8.34***	8.47***	1.23	14.1***	14.0***
	(1.84)	(1.85)	(2.15)	(0.606)	(0.592)
Belief			0.516***		
			(0.077)		
Observations	2,872	2,872	2,872	2,872	2,872
Lottery FE		x	x		x

Notes: Results from ordinary least squares (OLS) regressions on treatment dummies. The outcome variable in columns (1), (2) and (3) is the willingness to pay and in (4) and (5) the beliefs. Columns (1) and (4) do not include control variables. Columns (2), (3) and (5) control for age, gender and monthly available budget, as well as lottery fixed effects. Standard errors clustered at the subject level in parentheses. * $p < 0.1$, $^{**}p < 0.05$, *** $p < 0.01$

Table A.4: Regression results — Testing against 50:50 beliefs in *Censored* and *Joint*

	Belief difference	
	(1)	(2)
Constant	-0.436	
	(0.340)	
Censored		-0.364
		(0.416)
Joint		-0.521
		(0.555)
Observations	1,407	1,407

Notes: Results from ordinary least squares (OLS) regressions of the difference between the mid-point of the non-zero probability of winning the lottery and the belief on an intercept. We restrict the sample to the *Censored* and *Joint* treatment and to decisions with realistic beliefs. Standard errors clustered at the subject level in parentheses. * $p < 0.1$, $^{**}p < 0.05$, *** $p < 0.01$

Table A.5: Overspending: robustness experiment

	WTP		belief	
	No controls	Controls	No controls	Controls
	(1)	(2)	(3)	(4)
Info	6.19**	5.44**	10.4***	10.2***
	(2.49)	(2.46)	(1.37)	(1.33)
Joint	6.78***	6.27**	19.9***	19.7***
	(2.60)	(2.63)	(1.55)	(1.57)
Observations	2,070	2,070	2,070	2,070
Lottery FE		x		x

Notes: Results from ordinary least squares (OLS) regressions on treatment dummies. The outcome variable in columns (1) and (2) is the willingness to pay and in (3) and (4) the belief. The independent variables are indicators that equal 1 if the participant is in the respective condition and 0 else. Columns (1) and (3) do not include control variables. Columns (2) and (4) control for age, gender, monthly available budget and lottery fixed effects. Standard errors clustered at the subject level in parentheses. * $p < 0.1$, $^{**}p < 0.05$, *** $p < 0.01$

Table A.6: Summary statistics

	Treatments				
	Control	Sample	Censored	Joint	p(F)
	(1)	(2)	(3)	(4)	(5)
Female	0.56	0.57	0.55	0.49	0.43
	(0.50)	(0.50)	(0.50)	(0.50)	
Age (years)	40.11	38.81	39.05	40.34	0.69
	(13.93)	(13.19)	(14.63)	(13.53)	
College	0.74	0.68	0.75	0.69	0.37
	(0.44)	(0.47)	(0.43)	(0.47)	
Monthly budget	696.67	734.75	642.66	686.66	0.73
	(775.53)	(741.92)	(727.80)	(676.25)	
Play Video Games	0.62	0.54	0.61	0.60	0.51
	(0.49)	(0.50)	(0.49)	(0.49)	
Time Spent Playing / Week	1.11	1.15	1.33	1.31	0.72
	(1.60)	(1.82)	(2.38)	(2.31)	
Knows Loot Boxes	0.70	0.63	0.74	0.69	0.22
	(0.46)	(0.48)	(0.44)	(0.46)	
Gambling Index	0.03	0.04	0.04	0.05	0.09*
	(0.06)	(0.09)	(0.08)	(0.12)	
Self-Control Index	0.83	0.82	0.79	0.80	0.20
	(0.18)	(0.18)	(0.19)	(0.19)	
Observations	151	159	157	150	

Notes: Summary statistics. We include all subjects who completed the study.

Table A.7: Loot Box Statistics

	Treatments				
	Control (1)	Sample (2)	Censored (3)	Joint (4)	p(F)
Female	0.50	0.51	0.51	0.40	0.35
	(0.50)	(0.50)	(0.50)	(0.49)	
Age (years)	36.82	36.15	36.75	37.21	0.94
	(11.97)	(12.12)	(13.65)	(12.45)	
College	0.80	0.69	0.74	0.68	0.17
	(0.40)	(0.46)	(0.44)	(0.47)	
Monthly budget	767.30	736.35	626.14	677.68	0.48
	(853.01)	(802.62)	(605.51)	(635.04)	
Play Video Games	0.77	0.71	0.74	0.75	0.80
	(0.42)	(0.46)	(0.44)	(0.44)	
Time Spent Playing / Week	1.46	1.59	1.72	1.76	0.73
	(1.76)	(2.07)	(2.65)	(2.61)	
Knows Loot Boxes	1.00	1.00	1.00	1.00	1.00
	(0.00)	(0.00)	(0.00)	(0.00)	
Gambling Index	0.03	0.05	0.04	0.06	0.05*
	(0.06)	(0.11)	(0.08)	(0.14)	
Self-Control Index	0.82	0.80	0.78	0.78	0.43
	(0.17)	(0.18)	(0.20)	(0.19)	
Observations	105	100	116	104	

Notes: Summary statistics on loot boxes. We include all subjects who state to know what loot boxes are (69% of the total sample).

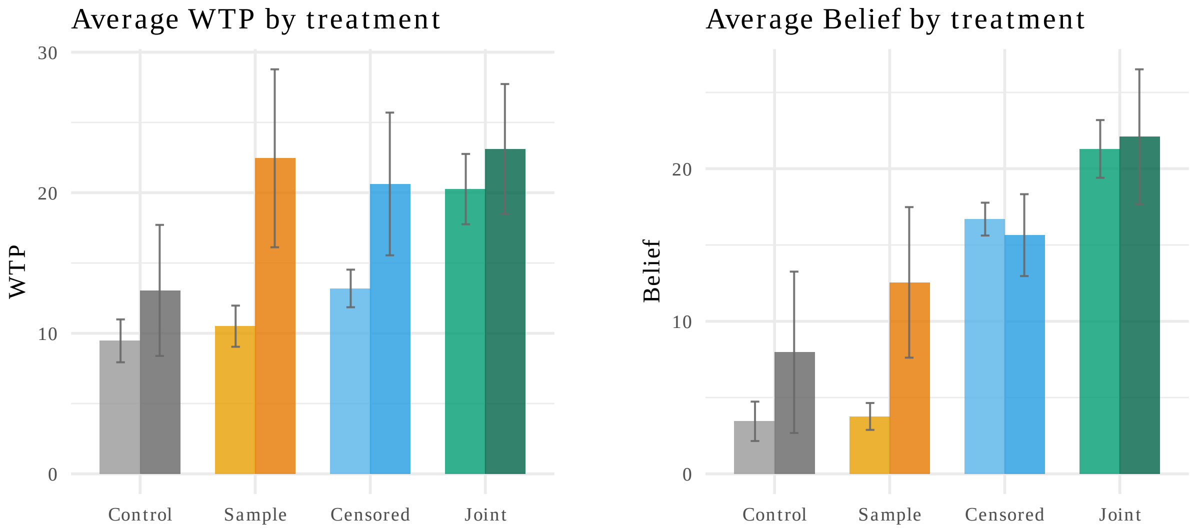


Figure A.2: Average willingness-to-pay and beliefs for subjects that overspent on loot boxes (dark) and those that did not (light) in the past for each treatment. WTP is the willingness to pay for a lottery. Belief is a subjective estimate of the frequency of winning the lottery in 100 independent draws. Whiskers are the standard error of the mean.

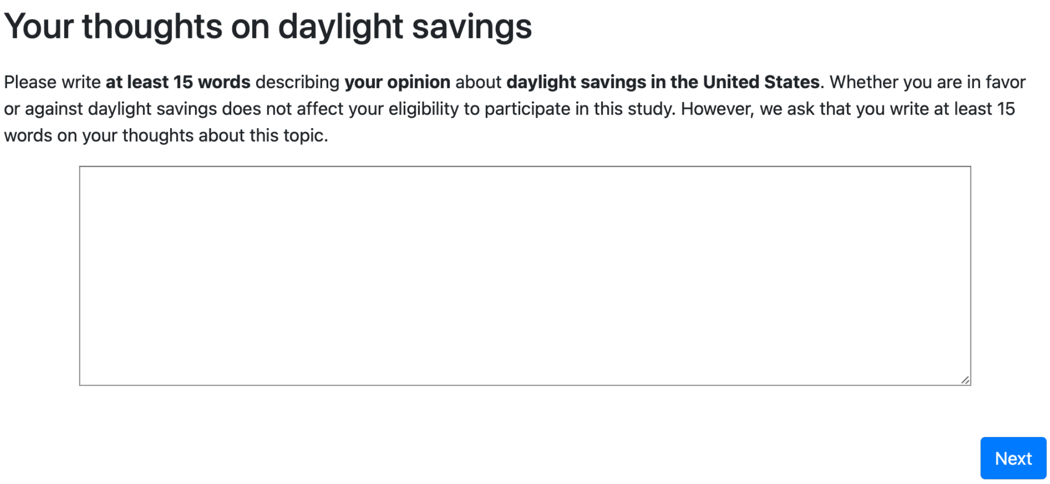
B Experimental Setup: Robustness Experiment

In this section, we present the design of *Robustness Experiment*, which is conducted in a between-subjects design. The experiment consists of three conditions: the *Control* and *Joint* conditions from the *Main Experiment*, as well as a new *Info* condition. The *Info* condition is an extension of the *Joint* condition and includes supplementary information concerning the prevalence of medium and high outcomes in a sample of 50 draws, as well as the information that the probabilities of these outcomes occurring are not equal. All other characteristics of the experiment are equivalent to those of the main Experiment, as described in Section 2.1.

We collected data from 414 subjects located in the UK via *Prolific* in November 2022. Screenshots of all parts of the experiment can be found in Appendix C. Subjects earned a base fee of £1.50 for participation. In addition, 1 out of 6 participants received a bonus payment depending on the WTP stated for one randomly selected lottery. Conditional on receiving a bonus, the average bonus paid was £5.08. The experiment took, on average, 11 minutes to complete.

C Experimental Details

Below we provide screenshots for all pages of the experiment. It includes the initial attention check, instructions, comprehension checks, and all survey questions. The screenshots are presented in the order in which participants progress through the experiment.

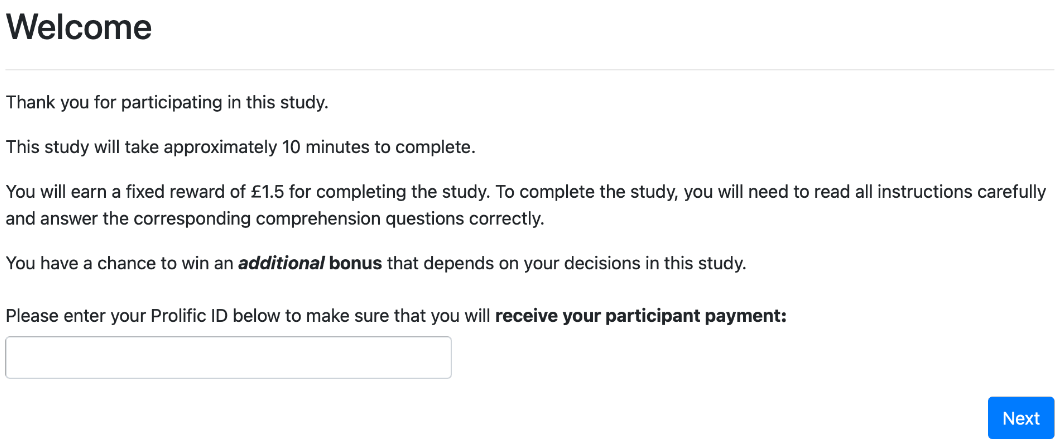


Your thoughts on daylight savings

Please write **at least 15 words** describing **your opinion** about **daylight savings in the United States**. Whether you are in favor or against daylight savings does not affect your eligibility to participate in this study. However, we ask that you write at least 15 words on your thoughts about this topic.

Next

Figure A.3: *Attention check at the beginning of the experiment.*



Welcome

Thank you for participating in this study.

This study will take approximately 10 minutes to complete.

You will earn a fixed reward of £1.5 for completing the study. To complete the study, you will need to read all instructions carefully and answer the corresponding comprehension questions correctly.

You have a chance to win an ***additional* bonus** that depends on your decisions in this study.

Please enter your Prolific ID below to make sure that you will **receive your participant payment:**

Next

Figure A.4: *Welcome page.*

Instructions

Please read these instructions carefully. There will be comprehension checks.

The Lotteries

In this study, we will ask you about your perception of, and preferences toward, different lotteries. Each lottery offers the chance to win different monetary prizes. The probabilities of these prizes may differ. You can find an example below:

Lottery			
Probability	40 %	30 %	30 %
Payoff	20 coins	40 coins	60 coins

Here, the lottery pays either 20 coins with 40% probability or 40 coins with 30% probability or 60 coins with 30% probability. Importantly, a lottery always only pays out one of the amounts; that is, it pays **either** 20 coins **or** 40 coins **or** 60 coins.

The Decisions and Bonus Payment

In total, you will see **5 lotteries**. For each lottery, we will ask you two questions. First, we will ask you **how often you think you would win the highest prize** of the lottery (i.e., 60 coins in the example above). After that, we will ask you for your **willingness to pay** for each lottery; that is, for the maximum amount of money you would pay to participate in this lottery.

Please indicate your **true** willingness to pay for each lottery. A payment mechanism will determine your potential bonus payment based on your stated willingness to pay. **The payment mechanism is designed in a way that gives you an incentive to indicate your true willingness to pay.** If you are interested in the details, you can find a complete description of the payment mechanism at the bottom of the page.

At the end of the experiment, the computer will randomly select one of the 5 lotteries for payment. The computer will then determine, based on your stated willingness to pay, your potential bonus payment. One out of every 6 subjects will be randomly selected to receive this additional bonus.

During the experiment, we use a fictitious currency called coins. For the bonus payment coins will be converted into British pound at a rate of 13 coins = 1 British pound.

Details on the payment mechanism | Next

Figure A.5: *Instructions on the experiment.*

Details on the payment mechanism | Next

The actual price of a lottery is set randomly. For this, a price will be chosen randomly. This randomly chosen price is between 0 and the maximum amount that the respective lottery can pay out. If the price is lower than or equal to your willingness to pay, you play the lottery. Your payoff will then be determined by a random draw from that lottery. Otherwise, the randomly selected price will be your payoff.

Example 1:

The lottery pays exactly 50 coins with a probability of 40% and nothing otherwise. You state that you are willing to pay a maximum of 20 coins to play the lottery. A random number between 0 and 50 is then drawn, which determines the price of 15 coins. Since this price is lower than your willingness to pay, you play the lottery. A random draw from the lottery determines your payoff.

Example 2:

The lottery pays exactly 0 coins with a probability of 20%, exactly 60 coins with 30% probability and exactly 70 coins with 50% probability. You state that you are willing to pay a maximum of 30 coins to play the lottery. A random number between 0 and 70 is then drawn, which determines the price of 35 coins. Since this price is higher than your maximum price, you do not play the lottery. You will receive the price of 35 coins as a payoff.

Figure A.6: *Additional info box with details on the payment mechanism.*

Comprehension questions

The questions below test your understanding of the instructions.

Important: If you fail to answer any one of these questions correctly, you will not be allowed to participate in the rest of the study, and you will not be able to earn a bonus payment.

1. Which payoffs are possible for the following lottery?

Lottery			
Probability	20 %	40 %	40 %
Payoff	0 coins	30 coins	80 coins

Please select one of the statements:

- It is possible that I get paid both 30 coins and 80 coins, i.e., I may receive a total amount of 110 coins from this lottery.
- I receive EITHER 30 coins OR 80 coins OR 0 coins from this lottery.
- I will receive at least some coins with certainty.

2. What is the probability to get a payoff of 30 coins for the following lottery?

Lottery			
Probability	50 %	40 %	10 %
Payoff	0 coins	30 coins	200 coins

Please select one of the statements:

- The probability to receive 30 coins is 60 %.
- The probability to receive 30 coins is 40 %.
- The probability to receive 30 coins is 0 %.

Next

Figure A.7: *Comprehension questions.*

You have correctly answered the comprehension questions. We will now start with the study.

The first decision begins on the next page.

Next

Figure A.8: *Start of the experiment.*

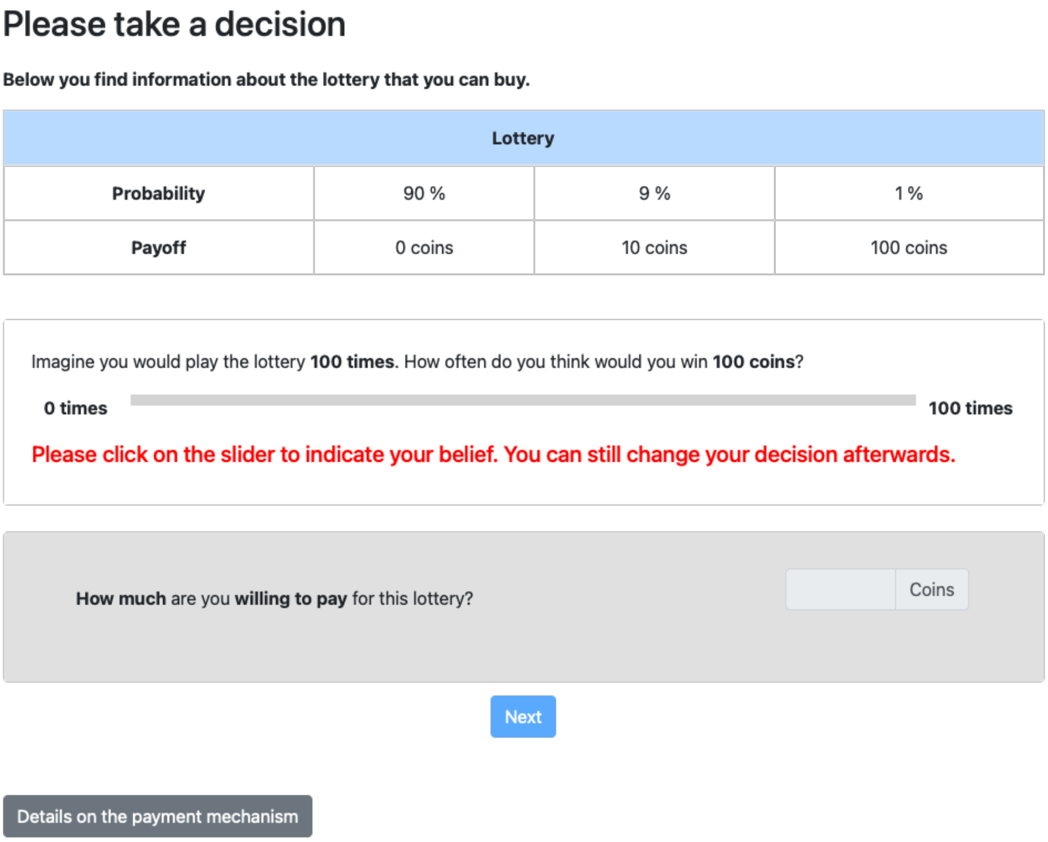
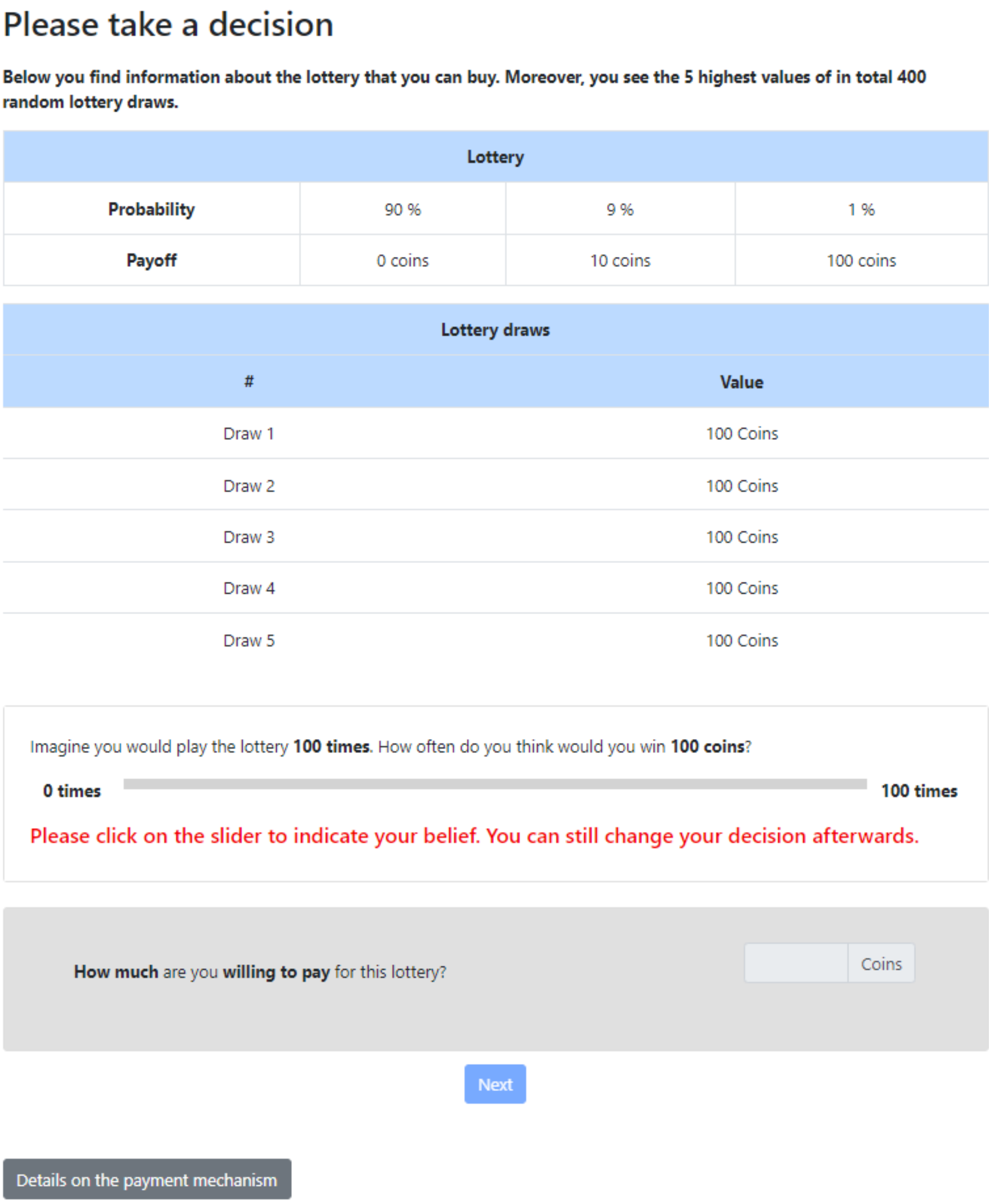


Figure A.9: *Belief & WTP elicitation in the "Control" treatment.*



Please take a decision

Below you find information about the lottery that you can buy. Moreover, you see the 5 highest values of in total 400 random lottery draws.

Lottery			
Probability	90 %	9 %	1 %
Payoff	0 coins	10 coins	100 coins

Lottery draws	
#	**Value**
Draw 1	100 Coins
Draw 2	100 Coins
Draw 3	100 Coins
Draw 4	100 Coins
Draw 5	100 Coins

Imagine you would play the lottery **100 times**. How often do you think would you win **100 coins**?

0 times **100 times**

Please click on the slider to indicate your belief. You can still change your decision afterwards.

How much are you **willing to pay** for this lottery? Coins

Next

Details on the payment mechanism

Figure A.10: *Belief & WTP elicitation in the "Sample" treatment.*

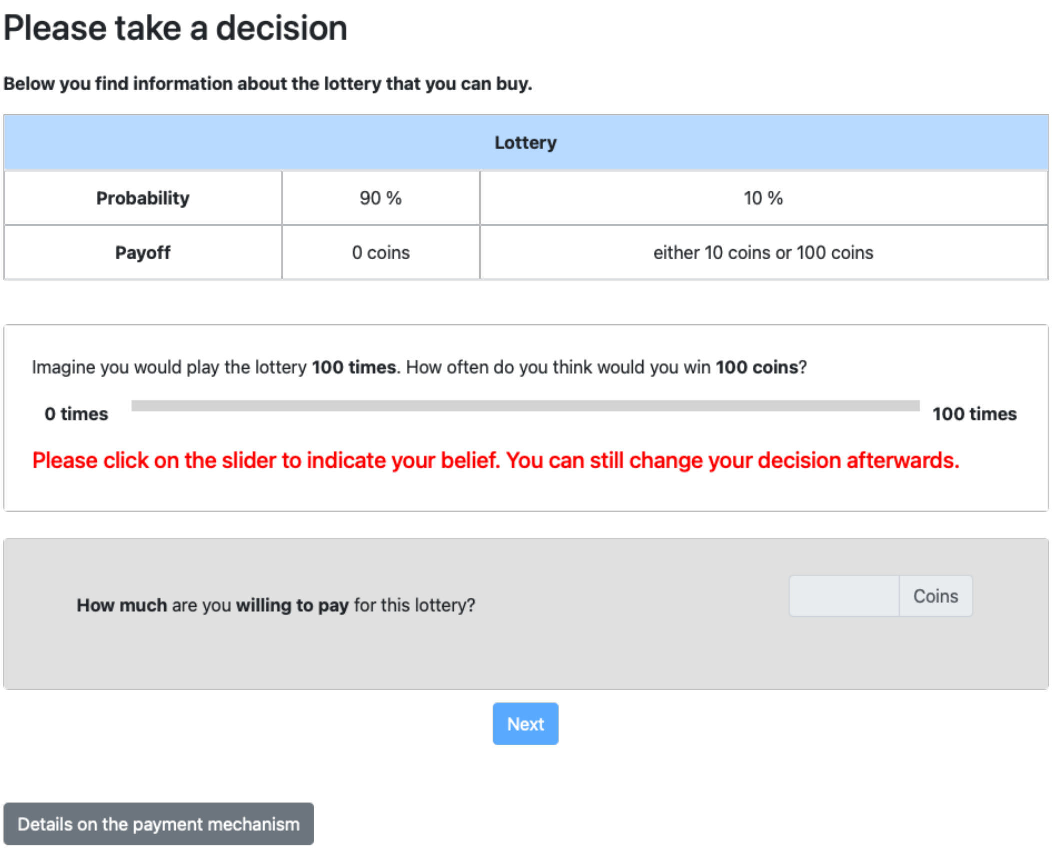


Figure A.11: *Belief & WTP elicitation in "Censored" treatment.*

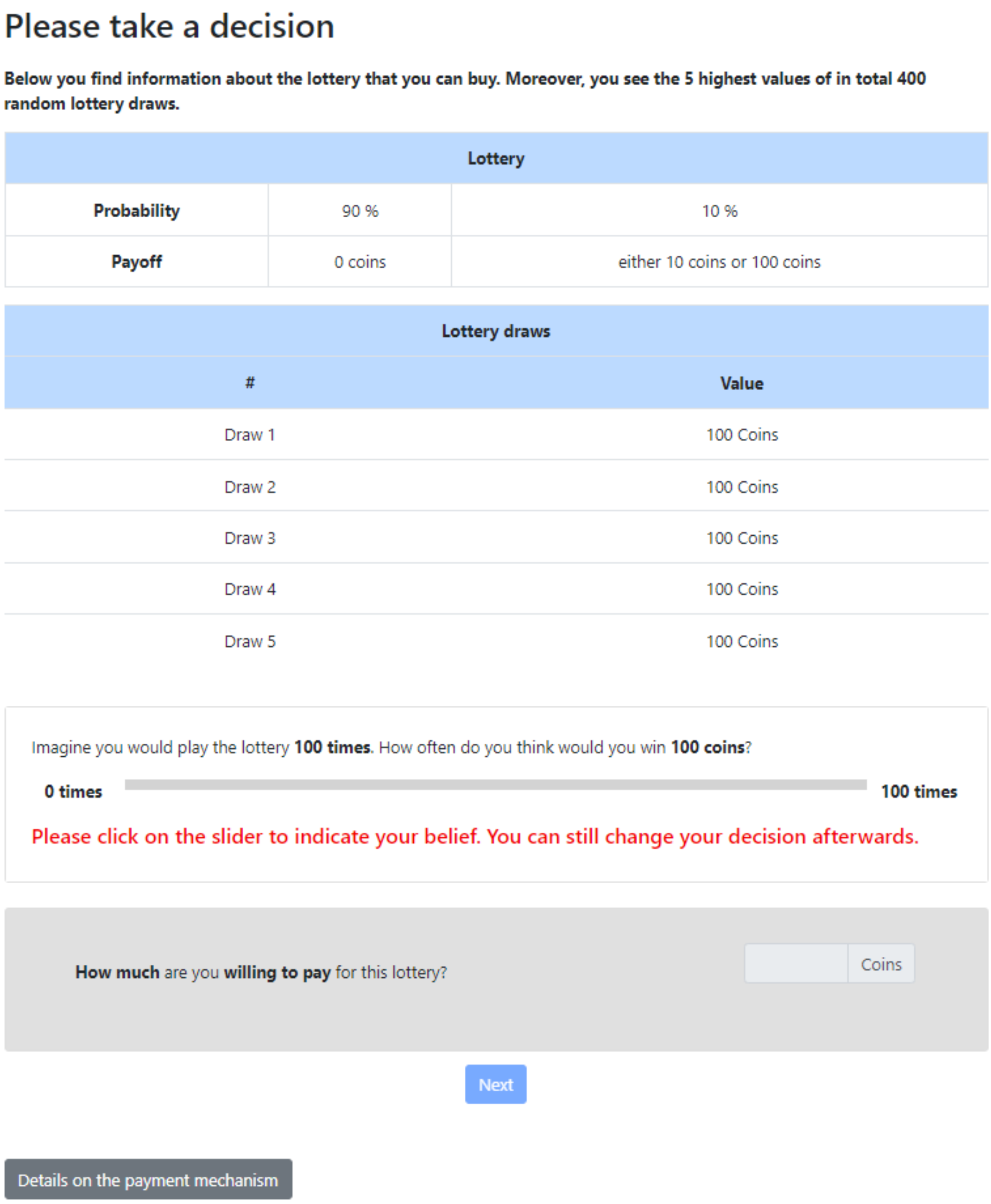


Figure A.12: *Belief & WTP elicitation in "Joint" treatment.*

Please take a decision

Below you find information about the lottery that you can buy. Moreover, you see the 5 highest values of in total 400 random lottery draws.

Lottery		
Probability	90 %	10 %
Payoff	0 coins	either 10 coins or 100 coins

Lottery draws	
#	**Value**
Draw 1	100 Coins
Draw 2	100 Coins
Draw 3	100 Coins
Draw 4	100 Coins
Draw 5	100 Coins

Note:

- 10 and 100 coins are **not equally likely.**
- Of the last 50 random lottery draws, **4 draws were equal to 10 coins** and **0 draws were equal to 100 coins.**

You will be able to enter your belief and your willingness to pay after 15 seconds.

Figure A.13: *Additional info in "Info" treatment.*

Please take a decision

Below you find information about the lottery that you can buy. Moreover, you see the 5 highest values of in total 400 random lottery draws.

Lottery		
Probability	90 %	10 %
Payoff	0 coins	either 10 coins or 100 coins

Lottery draws	
#	**Value**
Draw 1	100 Coins
Draw 2	100 Coins
Draw 3	100 Coins
Draw 4	100 Coins
Draw 5	100 Coins

Note:

- 10 and 100 coins are **not equally likely.**
- Of the last 50 random lottery draws, **4 draws were equal to 10 coins** and **0 draws were equal to 100 coins.**

Imagine you would play the lottery **100 times**. How often do you think would you win **100 coins**?

0 times **100 times**

Please click on the slider to indicate your belief. You can still change your decision afterwards.

How much are you **willing to pay** for this lottery? Coins

Next

Details on the payment mechanism

Figure A.14: *Belief & WTP elicitation in "Info" treatment.*

Please answer the following questions

Thank you for participating. Please answer the following questions. The survey will last approximately for another 10 minutes. Afterwards, you will receive your payment.

How old are you?

What is your gender?

What is your highest level of education:

What do you study? / What is your occupation?

How much money do you have available each month (after deducting fixed costs such as rent, insurance, etc., British pound)?

Next

Figure A.15: *General control questions.*

Please answer the following questions

Loot boxes in video games

In the following questions we will ask you about your experiences with loot boxes in video games. Loot boxes are digital items that contain a random reward. In many games it is possible to buy such loot boxes for real world money. Examples include Packs in Fifa 22, Hextech Chests in League of Legends or Eggs in Pokémon Go. Below you also see an exemplary picture of a loot box.

How many hours do you spend playing videogames per day on average?

Have you ever heard of loot boxes in videogames before?

- Yes
- No

Next

Figure A.16: *Questions about loot box experience.*

Please answer the following questions

How much did you spend on loot boxes during the last year on average? (in British pound):

Have you ever spend more than you planned to on loot boxes?

- Yes
- No

Next

Figure A.17: *Questions for loot box users.*

Please answer the following questions

Using the scale provided, please indicate how much each of the following statements reflects how you typically are.

	Not at all				Very much
I am good at resisting temptation.	1	2	3	4	5
I have a hard time breaking bad habits.	1	2	3	4	5
I am lazy.	1	2	3	4	5
I say inappropriate things.	1	2	3	4	5
I do certain things that are bad for me, if they are fun.	1	2	3	4	5
I refuse things that are bad for me.	1	2	3	4	5
I wish I had more self-discipline.	1	2	3	4	5

Figure A.18: *Self-control survey questions (Part 1).*

People would say that I have iron self-discipline.	○ 1	○ 2	○ 3	○ 4	○ 5
Pleasure and fun sometimes keep me from getting work done.	○ 1	○ 2	○ 3	○ 4	○ 5
I have trouble concentrating.	○ 1	○ 2	○ 3	○ 4	○ 5
I am able to work effectively toward long-term goals.	○ 1	○ 2	○ 3	○ 4	○ 5
Sometimes I can't stop myself from doing something, even if I know it is wrong.	○ 1	○ 2	○ 3	○ 4	○ 5
I often act without thinking through all the alternatives.	○ 1	○ 2	○ 3	○ 4	○ 5

Next

Figure A.19: *Self-control survey questions (Part 2).*

Please answer the following questions

In this part, we are interested in your gambling behavior.

Thinking about the last 12 months ...

Have you bet more than you could really afford to lose?	○ Never	○ Sometimes	○ Most of the time	○ Always
Have you needed to gamble with larger amounts of money to get the same feeling of excitement?	○ Never	○ Sometimes	○ Most of the time	○ Always
Have you gone back on another day to try to win back the money you lost?	○ Never	○ Sometimes	○ Most of the time	○ Always
Have you borrowed money or sold anything to gamble?	○ Never	○ Sometimes	○ Most of the time	○ Always
Have you felt that you might have a problem with gambling?	○ Never	○ Sometimes	○ Most of the time	○ Always

Figure A.20: *Gambling survey questions (Part 1).*

Have people criticised your betting or told you that you had a gambling problem, whether or not you thought it was true?	○ Never	○ Sometimes	○ Most of the time	○ Always
Have you felt guilty about the way you gamble or what happens when you gamble?	○ Never	○ Sometimes	○ Most of the time	○ Always
Has gambling caused you any health problems, including stress or anxiety?	○ Never	○ Sometimes	○ Most of the time	○ Always
Has your gambling caused any financial problems for you or your household?	○ Never	○ Sometimes	○ Most of the time	○ Always

Next

Figure A.21: *Gambling survey questions (Part 2).*

You have finished all tasks

You have completed this study in its entirety. **You will receive the fixed reward of <u>£1.5</u> credited to your Prolific account.**

You are eligable for the bonus payment. After the end of the entire study, one out of every 6 subjects will be randomly selected to receive an additional bonus. If you are selected for payment, your additional bonus will be determined as follows:

The computer will randomly select one lottery for payout, with equal probability, and determine a price for the lottery.

The actual price of a lottery is set randomly. For this, a price will be chosen randomly. This randomly chosen price is between 0 and the maximum amount that the respective lottery can pay out. If the price is lower than or equal to your willingness to pay, you play the lottery. Your payoff will then be determined by a random draw from that lottery. Otherwise, the randomly selected price will be your payoff.

If you won the bonus, we will inform you in the next days using your Prolific-ID.

Next

Figure A.22: *Last page in the experiment.*