Temptation: Immediacy and certainty

J. Lucas Reddinger[1][2]

8 December 2022

Job Market Paper

Abstract

Is an option especially tempting when it is both immediate and certain? To study the effect of risk on present bias, I conduct an online experiment in which workers allocate about thirty minutes of real-effort tasks between two weeks. I compare choices made two days before the first workday against choices made when work is imminent. In baseline treatments, one choice is randomly implemented; meanwhile, one treatment implements a particular allocation with certainty. By assuming that effort costs are not affected by the mechanism (and thus independent of risk preferences), my novel design permits estimation of present bias using a decision with a consequence both immediate and certain. I find the average intensity of present bias is far greater under certainty than under risk. I find no evidence that present bias is more pervasive among individuals, suggesting instead that present-biased individuals become more myopic.

***JEL* Codes** D03, D81, D91

Keywords intertemporal choice, time preferences, risk preferences, dynamic inconsistency, present bias, nonstationarity, immediacy effect, certainty effect, common difference effect, common ratio effect, non-expected utility, Allais paradox

[1]Department of Economics, University of California, Santa Barbara, 93106.

[2]Present affiliation: Menard Family Initiative, College of Business Administration, University of Wisconsin-La Crosse, 1725 State St., 339C Wimberly Hall, La Crosse, WI 54601; Phone (608) 785-6653; Email reddinger@ucsb.edu.

1. Introduction

While risk and time preferences are fundamental to the theory of decision-making, much remains unknown about the interplay of these two dimensions. Future prospects are inherently risky if any circumstance may arise that precludes consumption of the outcome. As a result of this simple tenet, outcomes may only be viewed as certain if obtained without delay. Further, this notion implies that a preference for certain outcomes results in a preference for immediate outcomes as well. When consumption is delayed, consumption becomes uncertain; conversely, risk diminishes the appeal of an immediate reward. I explore in this paper the relationship between certainty and immediacy with an experiment of dynamic decision-making over risky and delayed prospects.[3]

I focus on the effect of stated risk on estimates of the present-bias factor used in the quasi-hyperbolic discounting model, a prominent model that exhibits the immediacy effect. The model nests exponential discounting, but gives additional weight to the present moment relative to the future. This disproportionate weight on the present moment describes an immediacy effect. Accordingly, the decision-maker is present-biased, and she may change prior plans, exhibiting a preference reversal. Being one of the simplest dynamically inconsistent models, it is widely used for quantifying present bias. I experimentally investigate the robustness of this popular model to varying *types* and *degrees* of risk in decision-making.

To this end, I conduct an online experiment in which workers allocate a workload between two weeks. Each worker makes decisions in advance of the first workday, then makes additional decisions on the first workday with work being imminent. A present-biased worker will in advance choose some

[3]The term "uncertainty" often refers to Knightian uncertainty, which entails ambiguity regarding the probability distribution of consequences. My study involves no ambiguity, because subjects are informed of all probabilities and the selection mechanisms used. I use "uncertainty" and "risk" interchangeably, meaning more than one consequence is possible. This is in accord with *the certainty effect*, which implies no ambiguity or risk.

allocation between the two weeks, but then on the first workday prefer an allocation with less work for the present day. One allocation decision is selected to actually matter, and the worker must complete the tasks allocated to each week to earn a substantial payment. Workers may face risk regarding two dimensions—the day from which a decision is selected, as well as the relative productivity of the two weeks. The novel experimental design individually eliminates each channel of risk, measuring the effect of each on dynamic present-bias, as well as the interaction that approximates certainty and immediacy. I hypothesize that present bias is far more significant when the outcome is certain.

By measuring real effort supplied by workers in an online marketplace with substantial stakes, the results are undeniably meaningful for the labor market. I conduct my project on Amazon Mechanical Turk (AMT), an online spot labor market where "requesters" post small jobs ("HITs") for wages of their choosing, and freelance workers complete the jobs at will. My real-effort tasks asked workers to count the number of zeros in binary strings; other common jobs on the platform include tagging photos with keywords and transcribing sales receipts. In my "counting project," workers earn a premium wage for AMT—$9.50 for completing all three days of participation in the project, which consist of roughly 30 to 50 minutes of work in total.[4] These outcomes translate directly into a wide swath of labor market participation common in the "gig economy." Many people who drive for ride-hailing companies face nearly identical decisions as those posed in the present field experiment. Many such drivers must commit to providing a ride before learning the length of the ride, the destination, or the compensation to be earned. My results suggest that a driver who has an income target across multiple days may procrastinate less when facing uncertainty regarding compensation. With recent California legislation changing the terms for subcontractor versus employee status, ride-hail companies have modified the

[4]Subjects were first paid $1.50 for reading the experimental instructions and attempting a comprehension quiz; this component did not guarantee participation in the project.

information provided to drivers before committing to a specific ride, as well as consequences for backing out (Rana 2020). Such commitment and uncertainty regarding independent labor contracting may thus have a large unforeseen effect on labor market participation.

While results from this experiment translate directly to booming real-world labor markets, the theory being tested is fundamental to microeconomics and is thus widely applicable. Decisions regarding real-effort tasks present more than a field experiment in labor supply, but a general approach to induce precisely-timed consumption utility.[5] To address questions regarding dynamic inconsistency, precise timing of both decisions and the resultant consumption is crucial. When experimental subjects are paid with monetary stakes, the actual timing of resultant consumption is nebulous. Most individuals are unlikely to be so liquidity-constrained as to immediately consume \$9.50 of earnings; individuals are more likely to maintain a positive cash balance and incorporate the earnings into a smoother consumption plan. Although the counterfactual is unobservable—perhaps working an additional ten minutes of counting zeros directly displaces ten minutes of an equivalent task—real-effort tasks remain one of the most precisely-timed utility flows.

I find that the immediacy effect—measured with the quasi-hyperbolic present-bias parameter β—is significantly attenuated by the introduction of risk. In my experiment, decision-makers on average discount the future by a factor of $\hat{\beta} = 0.78$ relative to the present under certainty. In the treatment with the most risk, I find no statistically significant present bias, with $\hat{\beta} = 0.99$ when each decision has a 10% implementation probability. These results suggest that researchers should pay keen attention to prospects that are both risky and delayed, including risk introduced by randomization mechanisms. Finally, these results give grounds for more research on the relationship between various types of dy-

[5] Because workers give consent to participate in an academic research study, but their experience is otherwise commonplace for a worker on AMT, the experiment may be classified as a *framed field experiment* (Harrison and List 2004).

namic risk, resolution, and ambiguity.

1.1. Discounting, dynamic inconsistency, and present bias

Let's first discuss how an individual's plan may change over time. Suppose that it's Friday, and Charlie makes personal plans regarding when she will do chores and when she will relax on Saturday and Sunday. On Friday night, she makes Saturday plans to eat breakfast, do two hours of chores, then go to the beach. On Saturday morning, after finishing her breakfast, she has a change of heart. She reflects that she's simply in no mood for chores; she decides instead to go to the market for fun, leaving twice as many chores for Sunday. Charlie's revision of plans at the last moment is an example of dynamic inconsistency—choices that differ depending on when the choices are made.[6] This section presents exponential discounted utility (EDU), a dynamically consistent model, followed by quasi-hyperbolic discounted (QHD) utility, a model in which individuals may change plans regarding immediate consumption.

To model intertemporal decision-making, Samuelson (1937) introduced *exponential discounted utility*, which describes how an individual may value utility flows (of consumption goods, such as leisure) that occur across time. If utility flows $u(c_{t+\tau})$ result from consumption $c_{t+\tau}$ at time $t+\tau \in \mathbb{N}$, given a constant discount factor $\delta \in [0,1]$, the model gives an intertemporal value at time t of

$$U_t^{\text{EDU}} = \sum_{\tau=0} \delta^\tau u(c_{t+\tau}).$$

[6]Here we assume that Charlie's preferences are time-invariant (Halevy 2015).

We can compare value functions for a decision-maker at time $t = 1$ and $t = 2$ as

$$U_1^{\text{EDU}} = u(c_1) + \delta u(c_2) + \delta^2 u(c_3) \quad \text{and} \tag{1a}$$

$$U_2^{\text{EDU}} = u(c_2) + \delta u(c_3). \tag{1b}$$

On Day 1, the decision-maker allocates scarce utility between Day 2 and Day 3 to optimize her discounted utility, U_1. We may characterize her plan using the fact that she values Day 3 consumption utility at a factor δ of Day 2 utility, as in Equation (1a) and Figure 2a. On Day 2, she has the opportunity to change her plans—she could allocate leisure between Day 2 and Day 3 in the same proportion, or she could choose a different proportion. Upon optimizing her new discounted utility U_2, she chooses to allocate leisure in the same proportion. Once again, she values Day 3 consumption utility at a factor of δ that of Day 2 utility, as in Equation (1b) and Figure 2b. Because her substitution factor between Day 2 and Day 3 remains the same, her optimal plan as decided on Day 1 is the same as the optimal plan chosen on Day 2. Her choices are *dynamically consistent* between these two decision days. This model does not describe Charlie's preference reversal, wherein she decides to do fewer chores when imminent.

To capture a preference for immediate utility such as that of Charlie, Laibson (1997) introduces a present-bias parameter $\beta \in [0, 1]$ to discount all future utility flows contra present utility. This results in a preference for immediacy, also referred to as *present bias* or as an *immediacy effect*. This quasi-hyperbolic discounted utility model has an intertemporal value at time t of

$$U_t^{\text{QHD}} = u(c_t) + \beta \sum_{\tau=1} \delta^\tau u(c_{t+\tau}).$$

Figure 1: An exponential discounted utility maximizer is dynamically consistent

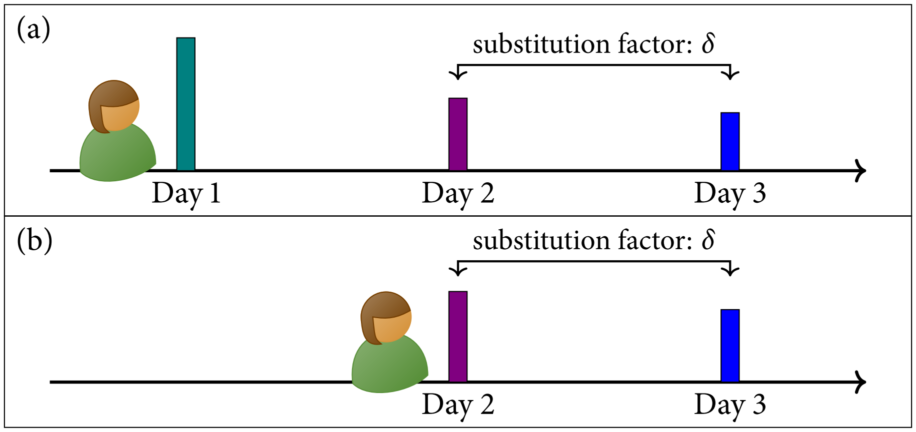
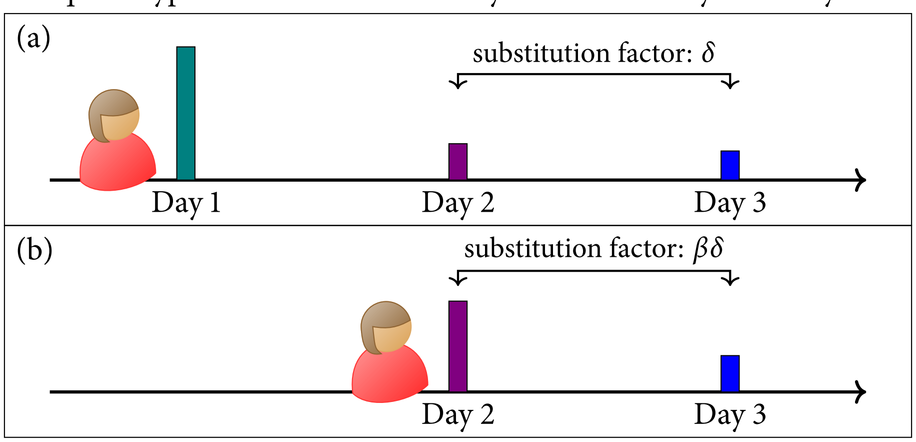


Figure 2: A quasi-hyperbolic discounted utility maximizer is dynamically inconsistent



Now the value functions at time $t = 1$ and $t = 2$ are respectively

$$U_1^{\text{QHD}} = u(c_1) + \beta\delta u(c_2) + \beta\delta^2 u(c_3) \quad \text{and} \tag{2a}$$

$$U_2^{\text{QHD}} = u(c_2) + \beta\delta u(c_3). \tag{2b}$$

On Day 1, Charlie is willing to substitute Day 3 consumption utility with factor δ as much Day 2 utility, as in Eq. (2a) and Fig. 3a. But on Day 2, her preferences for consumption between Day 2 and Day 3 are suddenly different, assuming $\beta < 1$. On Day 2, she values her immediate utility more highly, relative to Day 3 utility. Day 3 utility is now discounted by $\beta\delta$ relative to Day 2, whereas it was previously discounted by only δ, as seen in Eq. (2b) and Fig. 3b. So given that Charlie has $\beta < 1$, her consumption plan becomes more present-focused on Day 2. Because her plan has changed between Day 1 and Day 2, she has *dynamically inconsistent* preferences, a direct result of Charlie's *present bias*.

The present study will use this simple model of present bias to capture dynamic inconsistency. When $\beta = 1$, the decision-maker is not present-biased, and her consumption plans will not change over time. However, when $\beta < 1$, the decision-maker is present-biased, and her plans will change. To identify this present-bias factor β, we need to observe choices made in advance of their consequences, and then comparable decisions made when the consequences are immediate. Finally, to study the effect of risk on present bias, my experiment will vary the implementation likelihood of these decisions.

1.2. Diminishing sensitivity to delay or to risk

Allais (1953) introduced the notion that individuals exhibit diminishing sensitivity to risk. In a low-risk scenario, a marginal change in risk has a larger effect on behavior, while in a high-risk scenario, a marginal change in risk has a smaller effect. Recall from the previous section that present bias is an example of diminishing sensitivity to delay—an individual is more impatient regarding a delay that happens immediately than a delay that occurs in the distant future. In this section, I discuss diminishing sensitivity to risk and how it translates to diminishing sensitivity to time delay.

The *common ratio effect* is the prototypical example of diminishing sensitivity to risk, first presented by Allais (1953). Consider a menu $\mathcal{A}$ consisting of two simple prospects—gamble a yields 1 util with probability 0.9 (nothing otherwise), and a' yields 2 utils with probability 0.6 (nothing otherwise). Compare this to a menu $\mathcal{B}$ that consists of identical gambles, except that the probabilities are scaled by a common ratio of 2/3.

Menu $\mathcal{A}$:	$a = (1 \circ 0.9)$	or	$a' = (2 \circ 0.6)$
Menu $\mathcal{B}$:	$b = (1 \circ 0.6)$	or	$b' = (2 \circ 0.4)$

Because expected utility is linear in probabilities, the preference relation is maintained if the probabilities are multiplied by a common ratio. That is, under expected utility, a is preferred to a' if and only if b is preferred to b'. Allais noted that this relationship is often violated, usually with an individual preferring the safe alternative when risk is low, but then preferring the riskier option when both options involve more background risk.

The *certainty effect* is a special case of the common ratio effect, when one of the gambles obtains with certainty (probability of one). Now consider the following menus:

Menu $\mathcal{C}$:	$c = (3 \circ 1.0)$	or	$c' = (4 \circ 0.8)$
Menu $\mathcal{R}$:	$r = (3 \circ 0.5)$	or	$r' = (4 \circ 0.4)$

Here a common-ratio effect (with c preferred to c' and r' preferred to r) captures a preference for certainty that violates expected utility. Again notice that c obtains 25% more often than c', as does r relative to r'. An explanation is that increasing the likelihood by one-quarter is more valuable at 0.8 than at 0.4, simply because it guarantees the outcome in question. The 0.8 probability becomes 1.0, which obtains with certainty, while the 0.4 probability becomes 0.5, which still entails risk. Intuitively this behavior is consistent with a decision-maker who becomes less risk-averse in the presence of greater risk.

The common ratio effect also translates neatly into *time stationarity* (sometimes referred to as the *common difference effect*). Suppose that Ziggy will survive any given day with an independent probability of 0.8. This means that Ziggy will survive τ days with probability of 0.8^{τ}. We may thus translate probabilities into time delays and vice versa.[7] The following menus result:

Menu $\tilde{\mathcal{C}}$:	$C = 3$ now	or	$C' = 4$ in 1 day
Menu $\tilde{\mathcal{R}}$:	$R = 3$ in 3 days	or	$R' = 4$ in 4 days

Notice that the choices in both menus have the same delay of one day. Time stationarity of preferences implies that preferences over an equivalent delay is independent of when that delay occurs; this would

[7]With loose approximation, 0.5 is Ziggy's probability of survival after $\ln 0.5/\ln 0.8 \approx 3$ days, and 0.4 is his probability of surviving $\ln 0.4/\ln 0.8 \approx 4$ days.

mean that Ziggy has the same preferences over a one-day delay that begins today and a one-day delay that begins in three days. A preference for immediacy, the time-delay analog of the certainty effect, occurs when C is preferred to C', but R' is preferred to R.

While I often use terminology such as *hazard* and *survival*, bear in mind that any number of circumstances may arise over the course of a delay, not only death. Some circumstances may be foreseen, the risk of which may be carefully considered. Other circumstances may be unforeseen, and others still may carry a risk that is difficult to quantify.

Accordingly, assuming a constant, independent survival probability, the common-ratio effect in risk translates to non-stationarity in time. The certainty effect and the immediacy effect, respective special cases, are then theoretically equivalent. Chakraborty, Halevy, and Saito (2020) establish the equivalence of quasi-hyperbolic discounting (which distinguishes an immediacy effect) and rank-dependent utility with a certainty effect. This theoretical relationship has also been studied by Prelec and Loewenstein (1991), Baucells and Heukamp (2012), and Epper and Fehr-Duda (2018) with a variety of approaches. The implication of this equivalence is that the certainty effect and the immediacy effect are two sides of the same coin. That is, one can only capture certainty with immediate outcomes and vice versa. Because my study is among the first to use certain and immediate outcomes, I can test whether prevailing methodology has failed to fully capture these effects.

1.3. Evidence of risk moderating present bias

Multiple studies have shown that risk moderates present bias, although these studies have used only hypothetical or nearly-hypothetical monetary stakes. Keren and Roelofsma (1995) provided the first evidence of risk affecting present bias (the immediacy effect). Subjects make a single choice of either a smaller-sooner or larger-later hypothetical monetary reward that is obtained with some probability p. As shown in Table 1, treatment varies in two dimensions; payment occurs in the imminent future or the remote future, and this reward is obtained with probability p of 1.0, 0.9, or 0.5. At $p = 1$, the modal subject chooses the smaller-sooner reward for the imminent future and the larger-later reward for the remote future. Because the difference in time (4 weeks) and the difference in reward ($5) are the same for both choices, together these choices imply non-stationary (present-biased) time preferences. At $p = 0.5$, the modal subject chooses the larger-later reward at both time horizons, implying stationary time preferences. We conclude that present bias evaporates with the introduction of risk.

Table 1: Results from Keren and Roelofsma (1995), Table 1

	Probability p of monetary reward		
	1	0.9	0.5
Imminent future			
A. $50 now	82% (49)	54% (38)	39% (39)
B. $55 in 4 weeks	18% (11)	46% (32)	61% (61)
Remote future			
C. $50 in 26 weeks	37% (22)	25% (20)	33% (33)
D. $55 in 30 weeks	63% (38)	75% (59)	67% (67)
Difference (freq. of A minus C)	45%	29%	6%

Notes: Each subject in this 2×3 experimental design is assigned menu $\{A, B\}$ or menu $\{C, D\}$ and a probability of reward $p \in \{1, 0.9, 0.5\}$. The subject then makes a single pairwise choice from their assigned menu, given probability p of receiving the prize. The experiment was conducted using Dutch currency, here converted as Fl. 2.00 to $1.00.

Weber and Chapman (2005) replicate the results of Keren and Roelofsma (1995) using an iterated decision procedure to obtain indifference intervals for each subject, again with hypothetical monetary stakes. Baucells and Heukamp (2010) use real monetary incentives and find a similar relationship between risk and immediacy. Their results are the only evidence of this relationship using real stakes. Of the study's 221 MBA student subjects, three are selected for payment, and one of each of these students' 17 decisions is paid. This implies that each decision has fewer than 1 in 1,000 odds of being implemented, making the decisions nearly hypothetical. Regardless, these authors also conclude that present bias is moderated by risk.

In total, this body of evidence suggests that risk moderates the immediacy effect. However, these studies use either hypothetical stakes or monetary stakes with an extremely low probability of implementation. None of these studies have attempted to approximate real stakes that occur with certainty. Accordingly, a special relationship between the certain and immediate is plausible, and I aim to verify this hypothesis using decisions with consequences that are certainly implemented and immediately experienced. Specifically, I contribute estimates of the present-bias factor for various levels and types of risk, using an innovative methodology that approximates certainty (with a decision implemented with certainty) and immediacy (using imminent real-effort tasks).

1.4. Empirical estimates of discounting

Estimates of discount factors are extremely idiosyncratic. Frederick, Loewenstein, and O'Donoghue (2002) provide a survey of estimates of the exponential discount factor δ that span $[0, 1]$ with great spread. These estimates come from experimental and observational studies, using both real and hypothetical stakes, suggesting that measurement of time preferences depends on a variety of factors.

Attempts to jointly estimate β and δ of the quasi-hyperbolic discounting model have borne similar heterogeneity. Laibson, Repetto, and Tobacman (2007) use observational lifecycle consumption data to estimate $\hat{\beta} = 0.7031$ and $\hat{\delta} = 0.9580$. Andreoni and Sprenger (2012) minimize differential transaction costs and introduce the *convex time budget* (CTB) methodology in an attempt to separately identify risk and time preferences. Their most precise estimate of the daily discount factor is $\hat{\delta} = 0.99928$

(reported as $\hat{r} = 0.300$ with $\mathrm{SE}(\hat{r}) = 0.064$) with a corresponding present-bias factor of $\hat{\beta} = 1.004$ with $\mathrm{SE}(\hat{\beta}) = 0.002$. In sum, these results may suggest that evidence of present bias in consumption can be found in observational field data, but evidence of present bias in experimental monetary payments is evasive.

We should be skeptical of present bias in monetary rewards. An individual would need to be severely liquidity constrained to value \$5 today far more highly than \$5 tomorrow, holding transaction costs constant. Indeed, the intertemporal models of consumption discussed in Section 1.1 address precisely-timed utility flows. Any consumption resulting from a small cash payment is likely to be extremely diffuse in timing and flow.

Augenblick, Niederle, and Sprenger (2015) address this by studying dynamic inconsistency using real-effort tasks. The subjects' (cognitive or time) cost of completing these tasks provides the experimentalist the means to directly affect subjects' consumption utility with precise timing. They notably find no evidence of significant present bias in monetary rewards, but find significant present bias in real-effort tasks. Their primary experimental study motivates mine—subjects allocate a workload between two days using CTB, making allocations both when work is delayed and again when imminent. Using monetary payments, the authors estimate $\hat{\beta} = 0.974$ (SE = 0.009) and $\hat{\delta} = 0.988$ (SE = 0.003). Using two types of real-effort tasks, however, they estimate $\hat{\beta} = 0.888$ (SE = 0.033) and $\hat{\delta} = 0.999$ (SE = 0.025). These estimates constitute some of the most convincing experimental evidence of present bias, suggesting that individuals discount the future with a factor of 0.89.

Indeed, since the seminal study by Augenblick, Niederle, and Sprenger (2015), many papers have studied present bias in real effort. Imai, Rutter, and Camerer (2021) examine 220 estimates of present bias from 28 studies that use the convex time budget methodology. The authors find a high degree of heterogeneity in estimates of the present-bias factor β. The type of reward—monetary or real effort—best explains present bias. Consequently, we may conclude that immediate consumption utility, such as real-effort, must be used to capture a preference for immediacy.

1.5. Random incentive scheme

The experimentalist often faces trade-offs when choosing a mechanism, and some mechanisms can have unintended consequences. Many experiments in economics pay subjects for all decisions made (Charness, Gneezy, and Halladay 2016). While this approach preserves the scale of marginal incentives and risk, it can shift background wealth and reference points, confounding observed preferences. Consequently, the random incentive scheme (RIS) is often the preferred mechanism by paying each subject for one of their randomly-selected decisions. Azrieli, Chambers, and Healy (2018) argue that this mechanism is incentive-compatible by avoiding complementarity between outcomes.

In the context of my experiment, consider a subject who makes multiple real-effort allocations for a specific day. Further assume that the subject has convex costs for effort within a day. If multiple decisions are realized, then an outcome obtained from one decision directly affects the marginal cost in the other decision problems for that day. In contrast, the RIS would select only one decision for realization, properly isolating each decision from the other decisions with regard to background effort.

The RIS is not without potential shortcomings, however. Starmer and Sugden (1991) present some of the first experimental evidence of bias in RIS, demonstrating a failure of isolation. The authors do not pinpoint the underlying mechanism; they instead note many possible explanations, such as prospect theory and regret theory. Since then, evidence investigating the underlying mechanism has been mixed—Beattie and Loomes (1997) find a difference between RIS and single choice in one of their four experiments, and Cubitt, Starmer, and Sugden (1998) find no such difference.

Non-expected utility rationalizes isolation failure with non-linearity in probabilities, resulting in behavior consistent with the Allais paradox. Freeman, Halevy, and Kneeland (2019) find substantial evidence of the certainty effect when comparing pairwise choices to choice-list data, suggesting an interaction between the two. Cox, Sadiraj, and Schmidt (2015) do not find evidence of the Allais paradox, despite finding some failure to isolate. Freeman and Mayraz (2019) find some evidence of the Allais paradox, but the effect is independent of the outcome mechanism used.

Of general concern is the ability of subjects to isolate each individual decision made from the others, treating each as if it were a single, isolated choice. In this regard, isolation implies no framing effects between the various decisions. Further, a mechanism may be preferred if it is incentive-compatible with a single decision. Freeman and Mayraz (2019) conduct a careful experiment to study how various attributes of price lists can affect decisions made by individuals. The authors focus on decomposing a difference in decision-making between a price list and a single choice, by considering three conditions (see Table 2 for an example price list). With "R-list" treatment, all rows of pairwise choices are shown and one is selected uniformly at random, thus implementing RIS. Given "K-list" treatment, all rows of pairwise choices are shown, but one pairwise choice is selected with certainty (namely, Row 4). Finally, in a "single choice" treatment, only one pairwise choice is displayed (again, Row 4), and this choice is implemented with certainty. Conducted between-subjects, the authors fail to reject a null incentive effect (the difference between R- and K-list), but strongly reject a null framing effect (the difference between K-list and single choice). Accordingly, this evidence of isolation failure seems to be driven more by the presentation of choices than the incentive scheme. While the authors *do* find a certainty effect in additional Allais treatments, the effect is not differentially attenuated depending on the mechanism.

Table 2: A price list used by Freeman and Mayraz (2019)

Row	Option A	Option B
1.	\$1.00 with 100% chance	\$1.40 with 100% chance
2.	\$1.00 with 100% chance	\$1.40 with 95% chance
3.	\$1.00 with 100% chance	\$1.40 with 90% chance
4.	\$1.00 with 100% chance	\$1.40 with 85% chance
5.	\$1.00 with 100% chance	\$1.40 with 80% chance
6.	\$1.00 with 100% chance	\$1.40 with 75% chance
7.	\$1.00 with 100% chance	\$1.40 with 70% chance

Notes: "R-list" treatments use the random incentive scheme, with one row selected with uniform probability for implementation; "K-list" treatments offer identical choices, but one row (namely, Row 4) is implemented with certainty.

Brown and Healy (2018) add to this body of evidence by separating each pairwise choice onto a separate screen, comparing decision frequencies against price lists with all decisions juxtaposed on one screen. The sequence of separated decisions is drawn randomly for each subject in a separated treatment. Like Freeman and Mayraz (2019), treatments include an R-list, a K-list, and a single choice, but also a separated R-list and separated K-list. In contrast to Freeman and Mayraz, Brown and Healy reject a null incentive effect ($p = 0.041, N_1 = N_2 = 60$) between standard R- and K-lists, but the effect found is a *reverse* common ratio violation (with less frequent risk-taking behavior in the presence of more risk). The framing effect, captured between standard K-list and single choice treatments, is marginally significant ($p = 0.051, N_2 = 60, N_3 = 61$). However, with separated decisions, the difference between R-list and K-list treatments disappears ($p = 0.857, N_4 = 61, N_5 = 63$). Accordingly, use of separated decisions may be the most *empirically* incentive-compatible methodology for preference elicitation.

1.6. Convex time budgets

While pairwise choices are simple for subjects to understand and a mainstay of modern experimental economics, the resultant choice data are fairly limited. In most situations, if monotonicity holds, a single switch point is revealed within an interval. For example, if a subject chooses Option A in Rows 1–5 and Option B in Rows 6–7 from the price list depicted in Table 2, then we only know that indifference lies weakly between 75% and 80%. As previously discussed, this assumes that RIS is incentive-compatible, which is frequently empirically violated.

To elicit a preference for certainty, the subject must be offered an option with a certain outcome, and that decision cannot be randomized with the RIS payment mechanism. Accordingly, either a K-list or a single choice could be used, as both present one decision implemented with certainty, with differences only in framing. But a single decision cannot elicit an incentivized indifference interval.

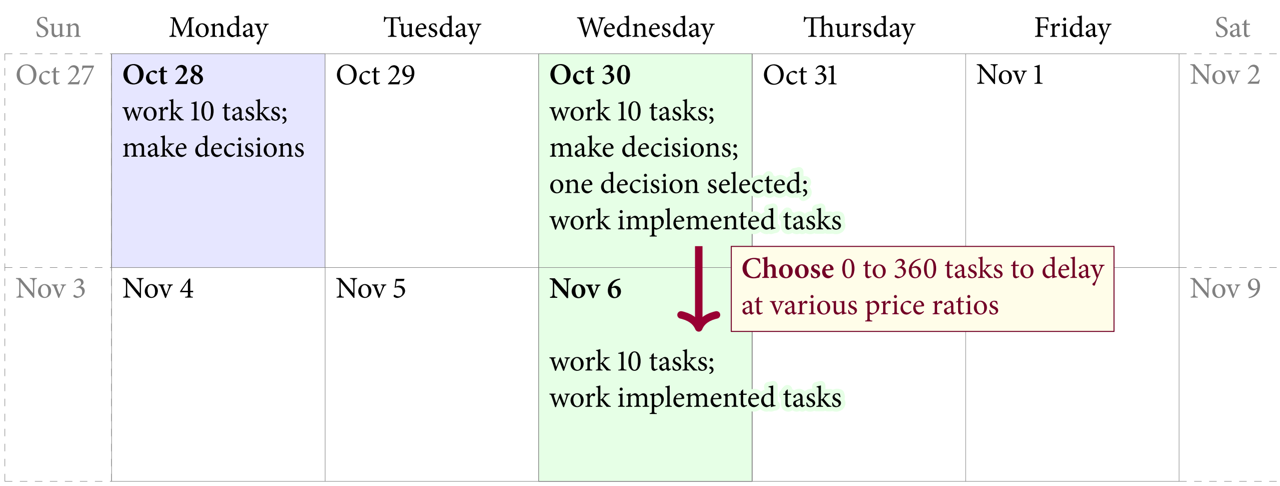
Accordingly, I turn to the *convex time budget* (CTB) approach pioneered by Andreoni and Sprenger (2012) and used to study dynamic inconsistency with real-effort tasks by Augenblick, Niederle, and Sprenger (2015). The CTB asks an individual to allocate a convex budget between two time periods. By making allocations at multiple price ratios, utility curvature can be identified while holding time preferences constant. Further, by making allocations with variation in time delay, time preferences can be identified. With monetary stakes, the identification of within-period utility curvature yields risk preferences; by using real-effort tasks, effort cost curvature is identified.

Essential to my experiment is the ability to obtain meaningful choice data from a subject using a decision that will be implemented with certainty. The CTB facilitates this with an interior allocation from the convex budget. That is, as long as a subject does not allocate all her work to the earlier workday or the later workday, then we learn about her preferences.

2. Experiment

I conduct an experiment in which individuals allocate real-effort between two weeks. Subjects participate on Monday (Day 0), Wednesday (Day 2), and the following Wednesday (Day 9). Subjects have a budget of 360 tasks to complete on Day 2, but any number of these tasks may be allocated to Day 9 at various price ratios using convex time budgets. To this end, subjects are asked on both Days 0 and 2 to allocate these tasks between Days 2 and 9. Figure 3 depicts the effort allocation problem in a calendar format.

Figure 3: Experimental timeline in a calendar format



At the beginning of the session on each of the three days, subjects are asked to complete ten *mandatory* real-effort tasks, in addition to the allocable tasks. These mandatory tasks give subjects recent experience with the tasks before making decisions on Days 0 and 2. The mandatory tasks also ensure no differential fixed costs between Days 2 and 9. For example, if a subject chooses to allocate all work to Day 2, she will still need to complete the 10 mandatory tasks on Day 9 in order to receive her substantial payment on Day 9. No decisions are made on Day 9; the subject simply completes the mandatory tasks and the allocated tasks. The experimental timeline is further detailed in Table 3.

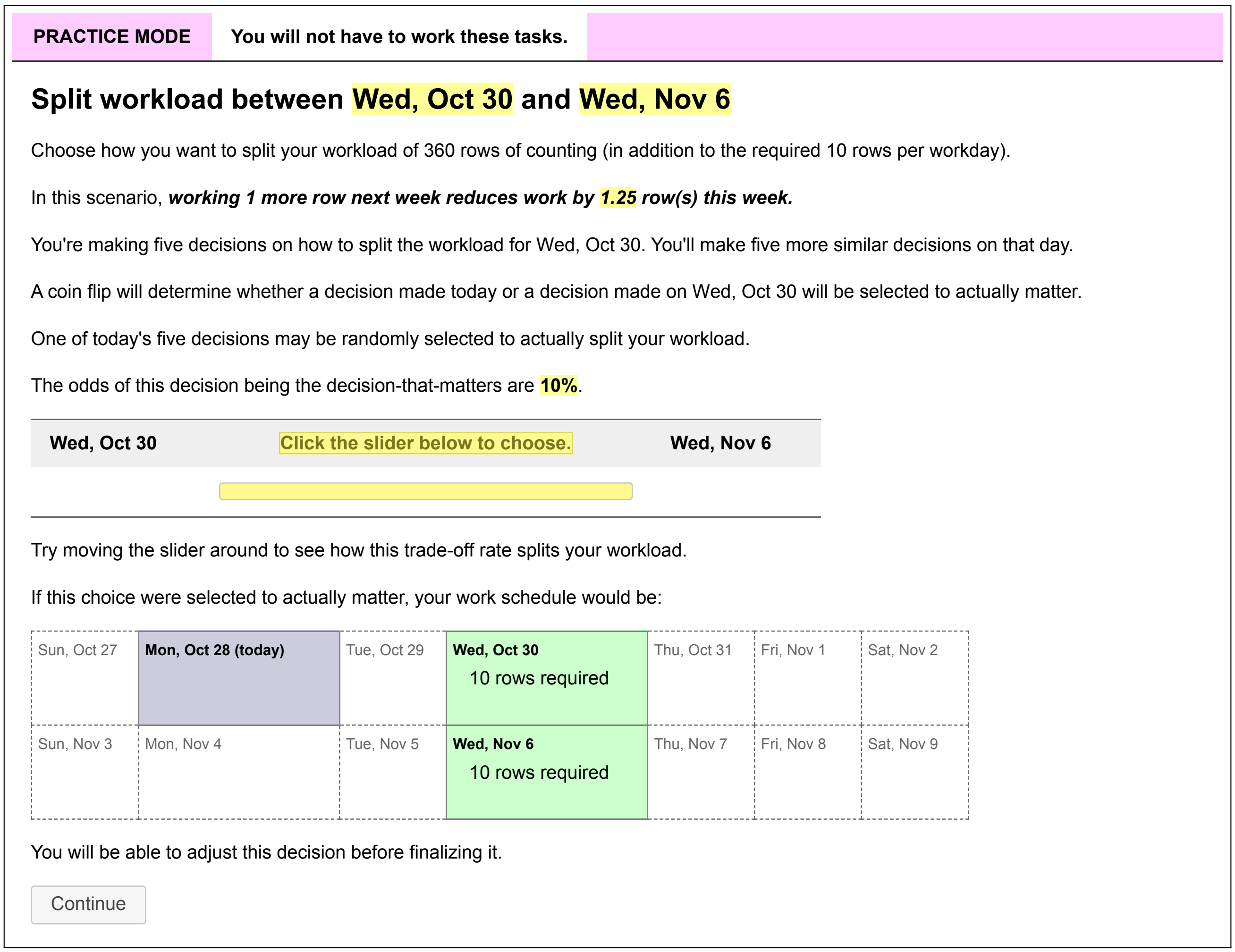
Intertemporal effort constraint A subject chooses how many of the 360 tasks for Day 2 that she wishes to delay to Day 9, at each of a variety of price ratios. Further, subjects make these choices on both Day 0 and Day 2. Accordingly, let $e_{i,d}^{t}$ denote effort chosen at price ratio R_i on decision-day d to be expended on workday t. For example, $e_{1,0}^{\text{Day 2}}$ is the effort chosen at price R_1 on Day 0 to be worked on Day 2. Then subjects face the constraint

$$e_{i,d}^{\text{Day 2}} + R_i e_{i,d}^{\text{Day 9}} = 360, \text{ for each } R_i \in \mathcal{R} := \langle 1.5, 1.25, 1, 0.75, 0.5 \rangle \text{ and } d \in D := \{0, 2\}.$$

These price ratios (which can also be interpreted as productivity ratios or gross interest rates) entail substantial income effects. For example, at $R_1 = 1.5$, a subject may choose to delay all 360 tasks from Day 2 to Day 9, thereby reducing the total work to 240 tasks on Day 9. On the other hand, if she chooses to delay all 360 tasks at $R_5 = 0.5$, she would need to complete 720 tasks on Day 9. I use these particular price ratios and a one-week delay to facilitate comparison with the results from Augenblick, Niederle, and Sprenger (2015).

Choice process Each subject thus makes ten decisions in total during the experiment. On Day 0, she chooses an allocation of tasks between Day 2 and Day 9 for each of the five price ratios. On Day 2, she again makes five such decisions. On each of these decision days, a subject makes a choice for each of the five price ratios, with each decision problem on a separate screen. Each subject is assigned to one of two possible sequential orderings of the presentation of these rates. Under price order $\mathcal{R}^A := \langle 1.25, 0.75, 1, 1.5, 0.5 \rangle$, the first screen elicits a choice for price $R_1 = 1.25$, the second screen for $R_2 = 0.75$, and so forth. Price order $\mathcal{R}^B$ is the converse of $\mathcal{R}^A$, with prices in the opposite order. After a subject makes choices at each of the five prices with each on a separate screen, she is then shown all five

Figure 4: Allocation interface



of her choices in a juxtaposed manner, sorted by R_i ascending in order. The subject may then make any final adjustments to her allocation choices before finalizing her decisions. I chose this process because it likely improves the quality of the choice data (Brown and Healy 2018; Freeman and Mayraz 2019). Figure 4 shows the separated allocation interface. After making choices at all five prices, the subject is shown their five choices juxtaposed and allowed to make any final adjustments (Appendix Figure 12).

Earnings A subject earns $1.50 for completing each daily session, which is paid within 24 hours. A subject earns a $5 bonus for completing all three sessions, paid within 24 hours of the last session. These payments, which total $9.50 for about 40 minutes of work, provide relatively large incentives for workers on Amazon Mechanical Turk (AMT). I withhold a large proportion of the total compensation until the the subject completes Day 9, thereby minimizing attrition. If a subject does not complete all of the decisions and tasks due on a participation day, she is not paid that day's earnings, nor may she participate further in the project.

Real-effort task Subjects are asked to count the number of zero digits ("0") in each of a sequence of binary strings and enter each count in an adjacent text field. If the response is correct, a green check ticks off the row; if the response is incorrect, a red cross appears with an error message. The subject must correct any mistakes before continuing. Figure 5 displays the real-effort task interface.

2.1. Treatments

Crucial is how one of these ten decisions is selected to be implemented. The baseline treatment randomizes uniformly across all ten decisions, so that the subject faces risk regarding which *price* will be implemented as well as risk regarding the *day* from which a decision will be chosen. In this baseline

Figure 5: Task interface

PRACTICE MODE **The correct answers are already filled in to save you time.**

Complete 10 required rows of counting

Please count the number of zeros (“0”) on each line and enter it in the box.

Each row will be marked correct or incorrect. You must correct errors before submission.

Row No.	String	Count (“0”)
1	1000110011100011	8
2	1000010100000001	12
3	1110001110000011	8
4	0100110010101111	7
5	0000100010101110	10
6	1101001011001010	8
7	0000111001010001	10
8	1011100110010010	8
9	1110110100011111	5
10	0110001100111001	8

Check responses and save

Table 3: Experimental timeline

Day 0 "Qualification HIT" Payment of $1.50 within 24 hours of completion

1. Instructions
2. Consent
3. Comprehension questionnaire
4. Demographic questionnaire

A subject is qualified for the next HIT if and only if all comprehension answers are correct.

Day 0 "Monday's HIT" Payment of $1.50 within 24 hours of completion

1. Instructions
2. *Practice:* Mandatory 10 tasks that would need to be completed
3. *Practice:* Effort allocation between Day 2 and Day 9, presented separately
4. *Practice:* Effort allocation between Day 2 and Day 9, presented juxtaposed
5. *Practice:* How today's decisions are used (resolution of decision-day risk)
6. *Practice:* How today's decisions are used (resolution of price risk)
7. *Practice:* View implemented tasks that would need to be completed
8. Complete the mandatory 10 tasks
9. Effort allocation between Day 2 and Day 9, presented separately
10. Effort allocation between Day 2 and Day 9, presented juxtaposed

A subject is qualified for the next HIT if and only if all parts of this HIT are completed.

Day 2 "This Wednesday's HIT" Payment of $1.50 within 24 hours of completion

1. Instructions
2. *Practice:* Mandatory 10 tasks that would need to be completed
3. *Practice:* Effort allocation between Day 2 and Day 9, presented separately
4. *Practice:* Effort allocation between Day 2 and Day 9, presented juxtaposed
5. *Practice:* How today's decisions are used (resolution of decision-day risk)
6. *Practice:* How today's decisions are used (resolution of price risk)
7. *Practice:* View implemented tasks that would need to be completed
8. Complete the mandatory 10 tasks
9. *Certain Day treatment only:* One day is selected for implementation
10. Effort allocation between Day 2 and Day 9, presented separately
11. Effort allocation between Day 2 and Day 9, presented juxtaposed
12. *Risky Day treatment only:* One day is selected for implementation
13. *Risky Price treatment only:* One price is selected for implementation
14. Complete the implemented tasks for today

A subject is qualified for the next HIT if and only if all parts of this HIT are completed.

Day 9 "Next Wednesday's HIT" Payment of $6.50 within 24 hours of completion

1. Instructions
2. Complete the mandatory 10 tasks
3. Complete the implemented tasks for today

Risky Price, Risky Day treatment, each decision has a 10% probability of being implemented, identical to the replication study of Augenblick, Niederle, and Sprenger (2015). Subjects are fully informed of this mechanism in advance, with the use of an intuitive interface that is described in the next section. In this baseline treatment with both a risky price and a risky day, both dimensions of risk are resolved on Day 2 after all decisions are made.

Certain Price (CP) treatment One dimension of my experiment eliminates risk regarding the price ratio to be implemented. In these treatments, subjects are informed that $R_2 = 1.25$ will certainly be implemented, making decisions for all prices $R_i \neq 1.25$ merely hypothetical. This treatment dimension thus presents a certain price to subjects.

Certain Day (CD) treatment The other dimension of risk concerns the day from which a decision is selected. On Day 0, prior to decision-making, all subjects are informed that either a decision from Day 0 or Day 2 will be selected with equal probability. However, subjects in a Certain Day treatment are informed that the day from which a decision is selected will be revealed *before* their Day 2 decisions. Accordingly, the day to be implemented is risky for *all* subjects on Day 0. Only subjects with Certain Day treatment learn which day's decisions matter before making Day 2 decisions.

Table 4: Probability of decision implementation

	Decision on	
Treatment	Day $d = 0$	Day $d = 2$
Risky Price, Risky Day	1/10	1/10
Risky Price, Certain Day	1/10	1/5
Certain Price, Risky Day	1/2	1/2
Certain Price, Certain Day	1/2	1

Notes: Probabilities of implementing effort allocation choice $e_{2,d}$ (chosen on decision-day d for $R_2 = 1.25$).

The resultant experiment is conducted between subjects, with a 2×2 factorial design between Certain Price and Certain Day treatments. Table 4 shows the stated probability of a decision being implemented by treatment condition. The interaction cell, with both Certain Price and Certain Day, is the primary interest of my study. Subjects in this treatment condition never face risk regarding the price—throughout the experiment, these subjects are aware that the price ratio $R_2 = 1.25$ will be selected with certainty. However, on Day 0, these subjects do not know whether their Day 0 decision or their Day 2 decision will be selected for implementation. On Day 2, prior to making a decision, they then learn which day's decisions matter. As a result, this Day 2 decision has all risk resolved. These decisions then ought to capture any special interaction between certainty and immediacy.

Each dimension of certainty necessarily converts some incentive-compatible decisions into hypothetical decisions. My analysis does not use hypothetical choice data, which indeed reduces the number of observations in these certainty treatment cells. Approximately half of the Certain Day subjects are informed that Day 0 has been selected, making their Day 2 decisions hypothetical. Because I exclude all hypothetical decisions from my analysis, I compensate by recruiting twice as many subjects for Certain Day treatment. I considered state-dependent resolution of risk, but I instead settled on the design as presented for its clarity and unadulterated integrity.[8]

[8] This design choice reduced by one-third my usable subject pool, compared to the following alternate design that I considered: If Day 0 is randomly selected, withhold this information until after Day 2 decisions are made; however, if Day 2 is selected, inform some subjects before Day 2 decisions are made. The subject would not have complete prior information about the precise timing of the resolution of risk, nor that this timing is state-dependent. In the spirit of complete prior information, informed consent, and avoidance of any potential deception, I rejected this alternate design at the cost of statistical power.

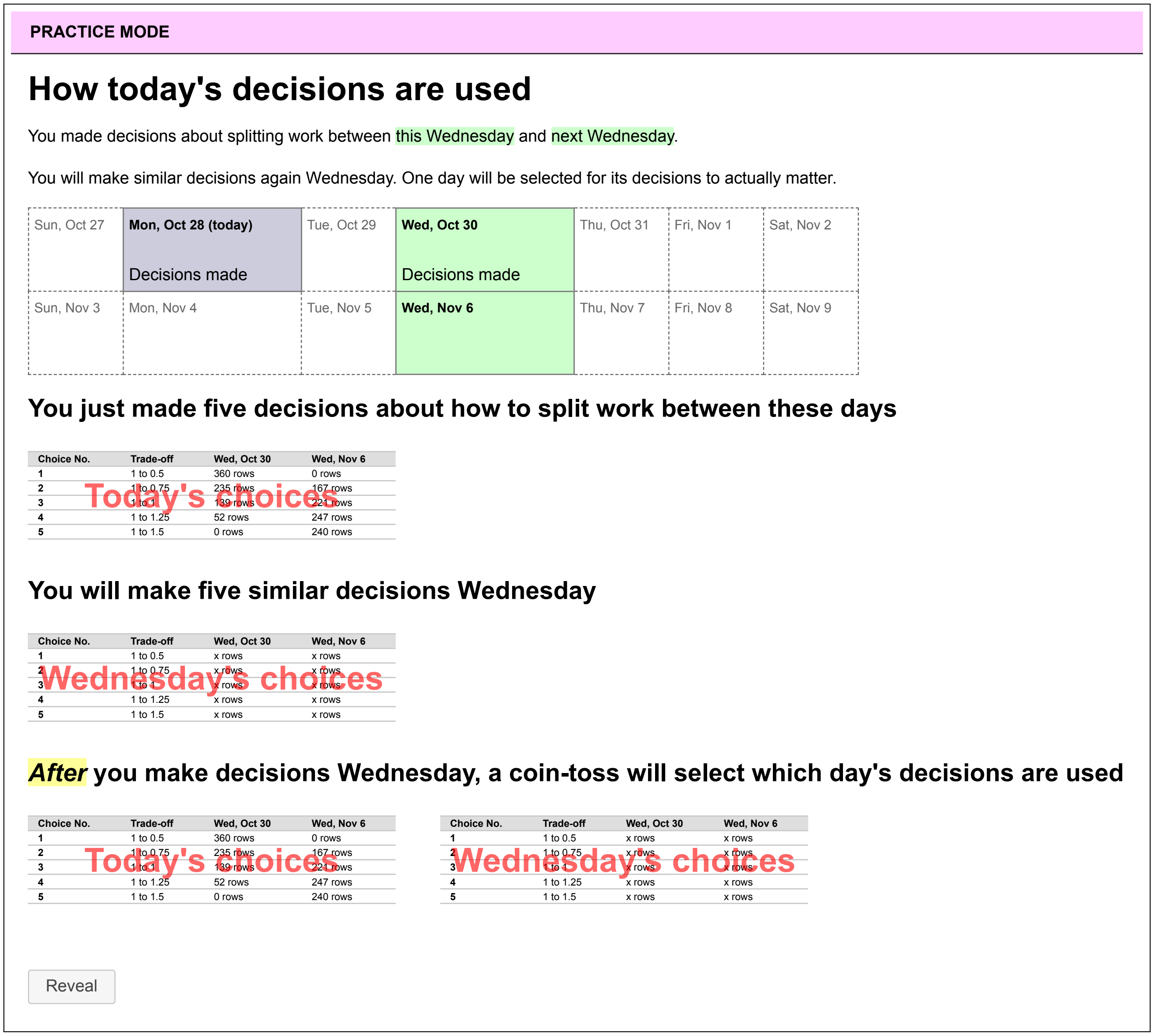
2.2. Interface

Of paramount importance is the subject's understanding of how her decisions will be used to actually affect the timing and amount of her work. Every subject is given complete prior information regarding her assigned mechanism with an *interactive practice round*. To ensure thorough understanding, the subject is guided through the *complete* practice round at the beginning of each participation day before the real tasks and decisions.

First, the subject is shown the ten mandatory tasks that are required to be completed at the beginning of every participation day. In the practice round, the text boxes are automatically populated with the correct answers, as shown in Figure 5. Next, the subject makes practice decisions for each of the five price ratios, with each decision on a separate screen, as shown in Figure 4. Next, the practice choices are shown juxtaposed as in Figure 12, allowing the subject to make any final changes. Having made practice decisions, the subject is now shown in a visually intuitive manner how these practice decisions would be used, as described below.

Day selection The subject is shown (in a list format, sorted by R_i ascending) the practice decisions she has just made. Her practice decisions from the alternate decision day are also shown on the same page. This interface is depicted in Figure 6. The page visualizes random selection between the two decision days, emulating a coin toss, by alternately highlighting the days in quick succession before one is finally selected. When the subject first clicks the button to try a coin toss, the alternate day is always selected. Then the subject is asked to try another practice coin toss, upon which the current decision day is selected. The practice round then proceeds with the present day's practice decisions selected as the practice decisions that matter.

Figure 6: Selection of risky day



PRACTICE MODE

How today's decisions are used

You made decisions about splitting work between this Wednesday and next Wednesday.

You will make similar decisions again Wednesday. One day will be selected for its decisions to actually matter.

Sun, Oct 27	**Mon, Oct 28 (today)** Decisions made	Tue, Oct 29	**Wed, Oct 30** Decisions made	Thu, Oct 31	Fri, Nov 1	Sat, Nov 2
Sun, Nov 3	Mon, Nov 4	Tue, Nov 5	**Wed, Nov 6**	Thu, Nov 7	Fri, Nov 8	Sat, Nov 9

You just made five decisions about how to split work between these days

Today's choices

Choice No.	Trade-off	Wed, Oct 30	Wed, Nov 6
1	1 to 0.5	360 rows	0 rows
2	1 to 0.75	235 rows	167 rows
3	1 to 1	139 rows	221 rows
4	1 to 1.25	52 rows	247 rows
5	1 to 1.5	0 rows	240 rows

You will make five similar decisions Wednesday

Wednesday's choices

Choice No.	Trade-off	Wed, Oct 30	Wed, Nov 6
1	1 to 0.5	x rows	x rows
2	1 to 0.75	x rows	x rows
3	1 to 1	x rows	x rows
4	1 to 1.25	x rows	x rows
5	1 to 1.5	x rows	x rows

After you make decisions Wednesday, a coin-toss will select which day's decisions are used

Today's choices

Choice No.	Trade-off	Wed, Oct 30	Wed, Nov 6
1	1 to 0.5	360 rows	0 rows
2	1 to 0.75	235 rows	167 rows
3	1 to 1	139 rows	221 rows
4	1 to 1.25	52 rows	247 rows
5	1 to 1.5	0 rows	240 rows

Wednesday's choices

Choice No.	Trade-off	Wed, Oct 30	Wed, Nov 6
1	1 to 0.5	x rows	x rows
2	1 to 0.75	x rows	x rows
3	1 to 1	x rows	x rows
4	1 to 1.25	x rows	x rows
5	1 to 1.5	x rows	x rows

Reveal

PRACTICE MODE | You will not have to work these tasks.

Implement workload for Wed, Oct 30 and Wed, Nov 6

For this practice round, the coin-toss selected today's decisions to actually matter.

Accordingly, here are the decisions you made today to split the workload of 360 rows of counting.

One rate and your corresponding choice is randomly selected to actually matter.

These work amounts are in addition to the 10 rows of counting required on each day.

Choice No.	Trade-off	Wed, Oct 30	Wed, Nov 6
1	1 to 0.5	360 rows	0 rows
2	1 to 0.75	235 rows	167 rows
3	1 to 1	139 rows	221 rows
4	1 to 1.25	52 rows	247 rows
5	1 to 1.5	0 rows	240 rows

Reveal

Figure 7: Selection of risky price

Price ratio selection The subject is now shown the five practice decisions she made on the present day, juxtaposed in a familiar way, as shown in Figure 7. If the subject has Certain Price treatment, then Row 4 is highlighted, and the subject is reminded that the choice for Row 4 with $R_2 = 1.25$ will certainly be implemented. However, if the subject has Risky Price treatment, no line is initially highlighted. The subject clicks a button labeled "Reveal," which initiates a roulette-wheel sequence, highlighting each row quickly in succession. After traversing the table twice, the highlight finally stops on the randomly selected decision line. In this manner the subject is shown how one of her five decisions from that day will be randomly selected.

After realizing which practice decision was ultimately selected, the subject is shown a practice task interface that requests the corresponding amount of work to be completed on the present day. Of course, the subject does not have to work these practice tasks. Finally, the subject clicks a link to exit the practice mode and begin an identical sequence with real decisions and real-effort tasks to complete.

3. Model

Assuming that the subject has power-function effort costs within a day, background effort of $\omega = 10$ tasks, and maximizes utility with quasi-hyperbolic discounting (QHD), then her optimization problem on decision-days $d = 0$ and $d = 2$ is to

$$\min_{e_{i,d}^{t}} \beta^{\mathbb{1}(d=0)}(e_{i,d}^{\text{Day 2}} + 10)^{\alpha} + \beta\delta^{7}(e_{i,d}^{\text{Day 9}} + 10)^{\alpha}, \text{ subject to } e_{i}^{\text{Day 2}} + R_i e_{i}^{\text{Day 9}} = 360, \tag{3}$$

for each price ratio R_i in $\{1.5, 1.25, 1, 0.75, 0.5\}$. This model uses δ as a daily pure exponential discount factor, while β discounts all future effort. Assuming the independence axiom (consistent with expected utility), the resultant intertemporal Euler equation is then

$$\left(\frac{e_{i,d}^{\text{Day 2}} + 10}{e_{i,d}^{\text{Day 9}} + 10}\right)^{\alpha-1} = \frac{\beta^{\mathbb{1}(d=2)}\delta^{7}}{R_i}. \tag{4}$$

We can use logarithms to linearize this equation as

$$\underbrace{\ln \frac{e_{i,d}^{\text{Day 2}} + 10}{e_{i,d}^{\text{Day 9}} + 10}}_{\text{log-effort-ratio}} = \underbrace{\frac{\ln\delta}{\alpha - 1}}_{\theta_{\text{delay}}} 7 + \underbrace{\frac{-1}{\alpha - 1}}_{\theta_{\text{lnrate}}} \ln R_i + \underbrace{\frac{\ln\beta}{\alpha - 1}}_{\theta_{\text{present}}} \mathbb{1}(d = 2). \tag{5}$$

This simple theoretical result implies that log-effort-ratio is the correctly specified choice variable to be considered in this problem, the left side of Equation (5). Further, the coefficients in this model are then differences in the intertemporal elasticity of substitution. For simplicity, let us define the regressand

$$E_{i,d} \coloneqq \ln \frac{e_{i,d}^{\text{Day 2}} + 10}{e_{i,d}^{\text{Day 9}} + 10}. \tag{6}$$

Adding an error term gives an estimable reduced-form model, with s indexing subjects, of

$$E_{i,d,s} = \theta_{\text{delay}} 7 + \theta_{\text{lnrate}} \ln R_i + \theta_{\text{present}} \mathbb{1}(\text{pr})_d + \varepsilon_{i,d,s}, \tag{7}$$

where $\mathbb{1}(\text{pr})$ indicates that the present day is a workday decision (when decision-day $d = 2$).

To fully account for any differences between treatments, I consider a model that allows a different β_T and δ_T for each of the four treatment conditions T. To this end, let $\mathbb{1}(\text{tr-cp})$ indicate treatment with Certain Price and $\mathbb{1}(\text{tr-cd})$ indicate treatment with Certain Day. We then arrive at the following estimable pooled reduced-form regression model:

$$\begin{aligned} E_{i,d,s} = {} & \theta_{\text{delay}} 7 + \theta_{\text{lnrate}} \ln R_i + \theta_{\text{present}} \mathbb{1}(\text{pr})_d + \theta_{\text{pr-cp}} \mathbb{1}(\text{pr})_d \mathbb{1}(\text{tr-cp})_s \\ & + \theta_{\text{pr-cd}} \mathbb{1}(\text{pr})_d \mathbb{1}(\text{tr-cd})_s + \theta_{\text{pr-cp,cd}} \mathbb{1}(\text{pr})_d \mathbb{1}(\text{tr-cp})_s \mathbb{1}(\text{tr-cd})_s \\ & + \theta_{\text{tr-cp}} \mathbb{1}(\text{tr-cp})_s + \theta_{\text{tr-cd}} \mathbb{1}(\text{tr-cd})_s + \theta_{\text{tr-cp,cd}} \mathbb{1}(\text{tr-cp})_s \mathbb{1}(\text{tr-cd})_s + \varepsilon_{i,d,s} \end{aligned} \tag{8}$$

Structural parameters are then recovered as follow.

$$\begin{aligned} \alpha &= 1 - \theta_{\text{lnrate}}^{-1} & & \\ \delta &= \exp \frac{\theta_{\text{delay}}}{-\theta_{\text{lnrate}}} & \delta_{\text{cp}} &= \exp \frac{\theta_{\text{delay}} + \theta_{\text{tr-cp}}}{-\theta_{\text{lnrate}}} \\ \delta_{\text{cd}} &= \exp \frac{\theta_{\text{delay}} + \theta_{\text{tr-cd}}}{-\theta_{\text{lnrate}}} & \delta_{\text{cp,cd}} &= \exp \frac{\theta_{\text{delay}} + \theta_{\text{tr-cp}} + \theta_{\text{tr-cd}} + \theta_{\text{tr-cp,cd}}}{-\theta_{\text{lnrate}}} \end{aligned}$$

$$\beta = \exp \frac{\theta_{\text{present}}}{-\theta_{\text{lnrate}}} \qquad \beta_{\text{cp}} = \exp \frac{\theta_{\text{present}} + \theta_{\text{pr-cp}}}{-\theta_{\text{lnrate}}}$$

$$\beta_{\text{cd}} = \exp \frac{\theta_{\text{present}} + \theta_{\text{pr-cd}}}{-\theta_{\text{lnrate}}} \qquad \beta_{\text{cp,cd}} = \exp \frac{\theta_{\text{present}} + \theta_{\text{pr-cp}} + \theta_{\text{pr-cd}} + \theta_{\text{pr-cp,cd}}}{-\theta_{\text{lnrate}}}$$

I will focus on results for these parameters, which are simple non-linear transformations of the reduced-form coefficients θ.

3.1. Identification of present bias

As represented by the factor β, present bias is identified from a two-day window. On Monday, Day 0, I assume the the decision-maker views both Wednesday, Day 2, and the next Wednesday, Day 9, as part of the future. Then on Day 2, I assume that she views that same day as part of the present moment, but she still views Day 9 as part of the future. This is how a two-day window identifies present bias in my experiment.

One could reasonably argue that Monday and Wednesday of the same week may both feel relatively present, while the following week may feel relatively distant. This would imply that present bias would be better identified from a week-long delay, as in Augenblick, Niederle, and Sprenger (2015). However, this is an empirical question, and Augenblick (2018) studies exactly how present bias varies with short delays. Using similar real-effort tasks, he finds that present bias quickly diminishes within three days, with two days capturing most present bias.[9] In my present study, the use of a two-day window will

[9]This experiment was originally designed with a three-day front-end delay and thus planned for participation days of Day 0 (Monday, October 28), *Day 3* (Thursday, October 31), and *Day 10* (Thursday, November 7). On Sunday, October 27th, I realized that Day 3 (the second decision-day) would be Halloween. Concerned about *any* potential confound, I moved the workdays to the Wednesdays, thereby reducing the front-end delay to two days. Because Augenblick (2018) suggests that short-run discounting levels off after about three days, I recommend at least three days for future studies. A week would be better, as used by Augenblick, Niederle, and Sprenger (2015), but the resultant three-week project may increase attrition on AMT. Also, when planning a labor field experiment, think carefully about weekday effects and holidays—including Halloween.

yield conservative estimates of β (biased upward). A week-long delay may better identify present bias, but it may also induce more attrition, which is a concern on AMT.

3.2. Identification of discounting

The daily discount factor δ is used to exponentially discount the future. In this setting, allocating work between Day 2 and Day 9 identifies this parameter, although this identification is more complicated than that of β.

Suppose that marginal cost of effort is constant within a day, so that the effort cost curvature parameter $\alpha = 1$. Then the ratio δ^7/R_i determines how the decision-maker allocates her workload between Day 2 and Day 9. If $\delta^7/R_i = 1$, she is indifferent to how the workload is split; otherwise she will allocate the entire workload to one day. For example, if she discounts the future ($\delta < 1$) but she can trade Day 2 and Day 9 work one-for-one (when $R_i = 1$), she will choose to do all of the work on Day 9.

Instead, assume that subjects do have increasing marginal cost of effort, with $\alpha > 1$. Then if $R_i = 1$ and $\delta = 1$, the subject would divide the workload evenly between Day 2 and Day 9. This is because the subject values smoothing effort between workdays, since additional effort becomes more costly within a day. Then as either the price ratio R_i or the discount factor δ changes, the subject chooses a different workload split between Day 2 and Day 9. Here, the subject weighs the benefit of smoothing effort against the discounted price ratio.

Because the ratio δ^7/R_i and effort cost convexity α jointly determine how a subject allocates her workload between days, the parameters α and δ are jointly identified. In Equations (7) and (8), parameters θ_{delay} and θ_{lnrate} determine α and δ. To this end, variation in the price ratio R_i obtains identification of these parameters.

3.3. Hypotheses

The primary hypothesis of this study is that there is an interaction between the immediacy effect and the certainty effect. This means that present bias at certainty should significantly differ from present bias with any amount of risk. Additional hypotheses are that each dimension of certainty will increase the severity of present bias. Treatment with Certain Price makes the decision at $R_2 = 1.25$ five times more likely to occur. Meanwhile, treatment with Certain Day makes the Day 2 decision twice as likely to occur. When treated with *both* Certain Price and Certain Day, the Day 2 choice for $R_2 = 1.25$ is implemented with certainty (probability of one).

Hypothesis 1 (Certainty and immediacy) Present bias is more intense under certainty (with both Certain Price and Certain Day treatment) than when the decision involves risk (the day is risky, the price is risky, or both). Then

(i) $\beta_{\text{cp,cd}} < \beta_{\text{cp}}$, present-bias factor given implementation probability of 1 versus 1/2,

(ii) $\beta_{\text{cp,cd}} < \beta_{\text{cd}}$, probability of 1 versus 1/5, and

(iii) $\beta_{\text{cp,cd}} < \beta$, probability of 1 versus 1/10.

Hypothesis 2 (Intermediate risk and immediacy) Away from certainty, present bias is more intense for decisions with higher probability of implementation. Then

(i) $\beta_{\text{cp}} < \beta$, present-bias factor given implementation probability of 1/2 versus 1/10,

(ii) $\beta_{\text{cd}} < \beta$, probability of 1/5 versus 1/10, and

(iii) $\beta_{\text{cp}} < \beta_{\text{cd}}$, probability of 1/2 versus 1/5.

These six hypotheses exhaustively compare present bias between treatments. Note that I do not hypothesize how the type of risk may matter. For example, controlling for the implementation probability, perhaps risk regarding the price ratio has a large effect, potentially driven by the income effect of these price ratios. Alternately, perhaps risk regarding the decision-day to be implemented is more important to the decision-maker, possibly driven by dynamic inconsistency. Research regarding types of risk and the underlying mechanisms is left to future work.

4. Results

I screened 389 workers on Amazon Mechanical Turk (AMT) on Monday, October 28th, 2019 using a "Human Intelligence Task" (HIT) to recruit workers for my experiment.[10] This "Qualification HIT" consisted of instructions, a comprehension quiz regarding the instructions, and a demographic survey.[11] Workers were promised $1.50 for a complete submission; every submission was compensated within 24 hours. Of these submissions, 220 answered all comprehension questions correctly and gave informed consent; these workers were subsequently enrolled in the experiment as subjects.

Subjects were asked to complete one HIT on that same day (October 28th, Day 0), one HIT on Wednesday (October 30th, Day 2), and one HIT on the following Wednesday (November 6th, Day 9). All three HITs were titled "multi-day counting project" and began with similar instructions that reminded subjects of the timing and payments involved. Subjects were paid $1.50 for successful completion of each of these HITs, as well as a $5 bonus for completing the entire project (meaning that they came back on Day 9 and completed all their tasks).

[10] I pre-registered this study with the American Economic Association as AEARCTR-0004651 (Reddinger 2019). The Human Subjects Committee at the University of California, Santa Barbara exempted my Protocol 56-19-0621.

[11] Only workers who had previously completed 1,000 HITs in total on AMT, had an approval rate of at least 98%, and were residents of the United States or Canada were able to view and accept the Qualification HIT.

4.1. Treatment effects on effort-share

Although log-effort-ratio is the correct choice variable for our structural estimation, let us also consider the simpler outcome of effort-share. This outcome will facilitate our interpretation of average treatment effects. Let us define effort-share as the proportion of total possible Day 2 effort that is chosen for Day 2; or $\varphi \coloneqq e_i^{\text{Day 2}}/360$. Because the Certain Price treatment only incentivizes $R_2 = 1.25$, let's only look at choice data at this price. Notice also that at this $R_2 = 1.25$ price ratio, effort is split evenly between the days if $e_i^{\text{Day 2}} = e_i^{\text{Day 9}} = 160$, which is when $\varphi_i = 160/360 = 0.44$. (Recall that $R_2 = 1.25$ results in a positive income effect as more tasks are delayed, permitting less work overall.)

Let us first consider the number of subjects who choose a lower share of effort for Day 2, φ, when deciding on Day 2 than was previously chosen on Day 0. That is, let's look at subjects who revise their Day 2 effort to a lesser amount than previously chosen. I assume that such dynamic inconsistency indicates present bias (as opposed to preferences that change over time). The proportion of subjects who make such present biased choices does not vary significantly by treatment (see Table 5). Therefore, these treatments that remove risk (with Certain Price, Certain Day, or both) do not seem to make an individual present biased if she was not already so.

Figure 8 depicts effort-share choices by treatment, with the fraction of choices made on Day 0 in blue and choices made on Day 2 in red. One striking observation is that many subjects revise their Day 2 effort share by choosing a very low amount, 0–10%. These choices at the lowest allocation of effort suggest that both dimensions of certainty treatment result in greater present bias. The proportion of present biased subjects in these treatments is not greater; rather, the *severity* of these subjects' present bias seems to be significantly greater with both dimensions of certainty treatment.

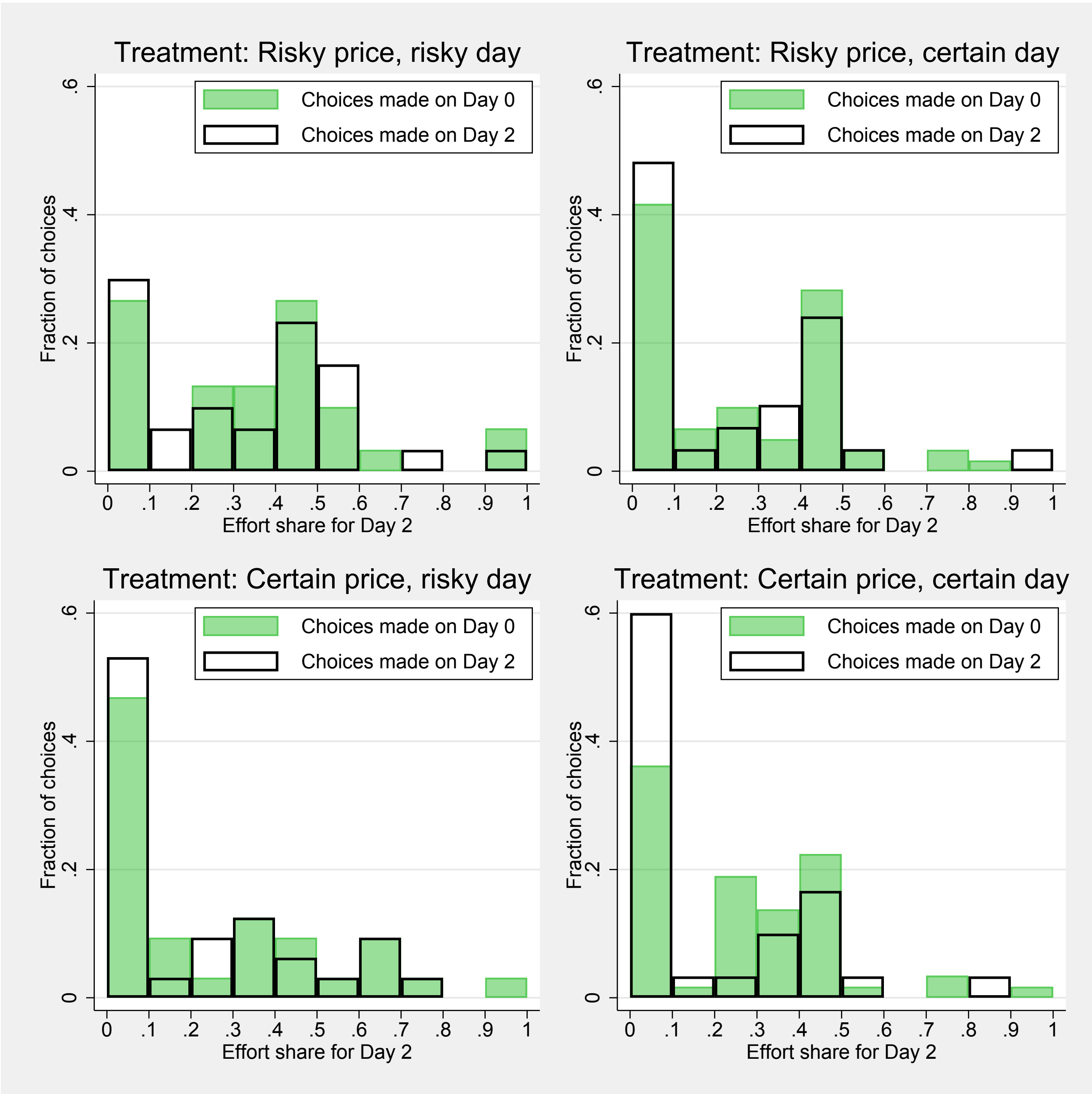
Next, let's look at treatment effects by day, to assess whether effort choices differ across treatments

Table 5: Present biased subjects by treatment

Treatment	Risky Day	Certain Day
Risky Price	12 of 30	6 of 29
Certain Price	9 of 32	9 of 30

Notes: Only choices made at the $R_2 = 1.25$ price ratio are used for comparability.

Figure 8: Histograms of effort-share chosen for Day 2 at $R_2 = 1.25$ for each treatment



within a given day. The first column of results in Table 6 shows us that on Day 0, subjects with a certain price choose 17.47% less of their 360 tasks for Day 2, compared to the baseline treatment with both types of risk. Then when deciding on Day 2, subjects treated with a certain price choose 21.54% less of their 360 tasks to complete on that same day, relative to the baseline treatment. I include choices made at all price ratios, given that the decisions are incentivized. Standard errors are clustered on subject, which yields about 30 clusters per 2×2 treatment cell, with 121 subject clusters in total. Note that results are similar for both the effort-share outcome and the log-effort-ratio outcome.

Table 6: Effort share by treatment regression results

	Effort-share choices $\varphi_{i,d}$			
	On Day $d = 0$		On Day $d = 2$	
Certain Price	−0.1747	$p = 0.030^{*}$	−0.2154	$p = 0.008^{**}$
	(0.0806)		(0.0811)	
Certain Day	−0.0436	$p = 0.236$	−0.0512	$p = 0.299$
	(0.0368)		(0.0493)	
Certain Price and Day	0.0669	$p = 0.461$	−0.0231	$p = 0.843$
	(0.0908)		(0.1166)	
$\ln R_i$	−0.6848	$p = 0.000^{***}$	−0.5907	$p = 0.000^{***}$
	(0.0738)		(0.0869)	
Constant	0.4887	$p = 0.000^{***}$	0.4451	$p = 0.000^{***}$
	(0.0339)		(0.0344)	
N (Decisions)	540		357	
G (Subjects)	180		121	
N_l (Left-censored)	92		68	
N_u (Uncensored)	391		257	
N_r (Right-censored)	57		32	

Notes: Excludes attrited subjects. Robust standard errors in parentheses, clustered on subject, from a two-limit Tobit model. $^{+}p<0.10$, $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.

We may then conclude that having a certain price makes subjects less patient. This is consistent with the Chakraborty, Halevy, and Saito (2020) explanation of the risk-delay relationship. One interpretation is that subjects experience delay as some risk of non-delivery. Further suppose that individuals have diminishing sensitivity to risk. Then the elimination of risk regarding the price induces subjects to be more risk-averse. Because delay is risky, subjects are less patient. Accordingly, subjects allocate more effort to the future.

Next we will use decisions from both Day 0 and Day 2 pooled together in one regression. This approach will give an increase in efficiency by combining all observations in one regression. This will also permit recovery of structural parameters.

4.2. Parametric results for the unrestricted model

Now let us return to our model of log-effort-ratio specified in Equation (8), which allows β_T and δ_T to vary by each treatment condition T. The estimated results for this unrestricted model are shown in Figure 9 and Table 7. For our econometric tests based on the hypotheses in the previous section, let us consider the null hypothesis that present-bias factors β_T are equal across treatments T. We reject that $\beta_{\text{cp,cd}} = \beta$ (at $p = 0.0177$). We also find marginal evidence against the null hypothesis $\beta_{\text{cp,cd}} = \beta_{\text{cd}}$ (with $p = 0.0527$), and we fail to reject $\beta_{\text{cp,cd}} = \beta_{\text{cp}}$. These latter tests seem to be underpowered, owing to the low number of observations in the certainty treatments. We also fail to find significant differences in present bias among intermediate levels of risk. We may in total conclude that the severity of present bias at certainty is sufficiently strong to be statistically different than the present bias with both types of risk. This evidence supports an interaction between certainty and immediacy. When one dimension of risk is already resolved, further eliminating the other dimension of risk seems to significantly affect present bias. These results are shown in Figure 9 and Table 7.

Figure 9: Estimates of β_T from the unrestricted model

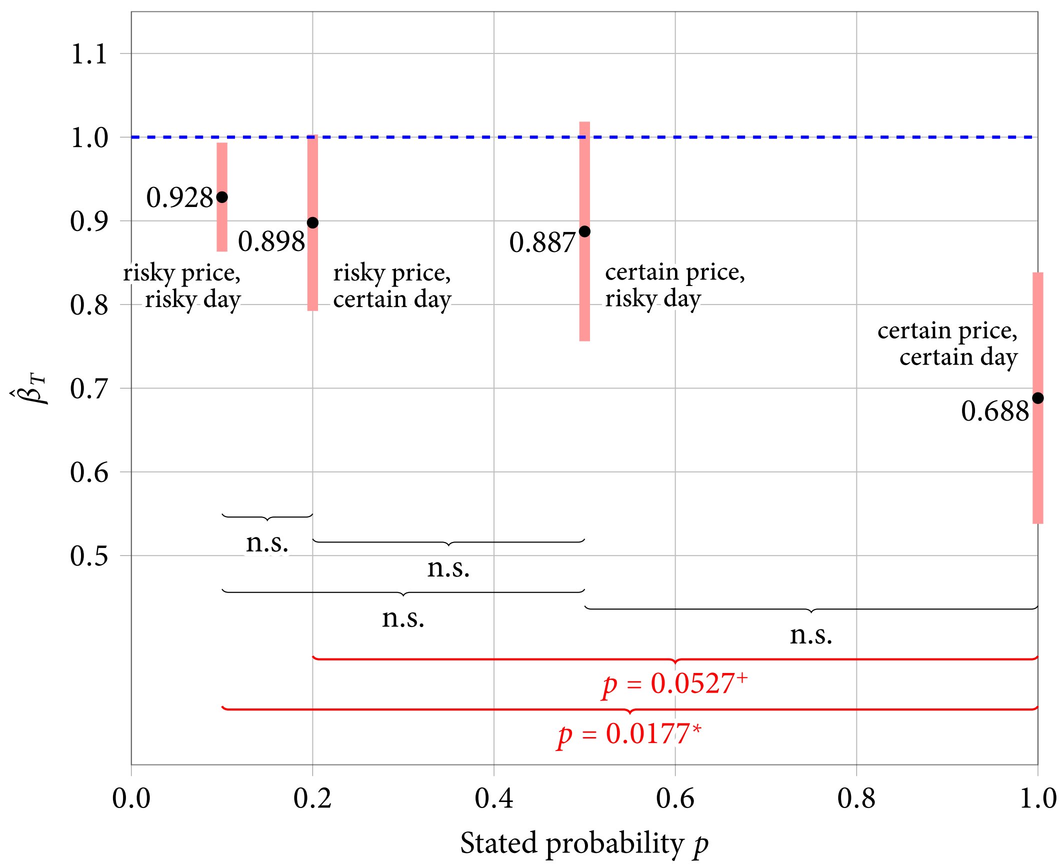


Table 7: Regression results from the unrestricted model

Parameter	Estimate	Hypothesis p-value that the parameter is equal to			
		1	β_{cd}	β_{cp}	$\beta_{cp,cd}$
β	0.9283	0.0294*	0.6027	0.6741	0.0177*
β_{cd}	0.8978	0.0558+		0.9213	0.0527+
β_{cp}	0.8873	0.2314			0.1357
$\beta_{cp,cd}$	0.6882	0.0020**			
		1	δ_{cd}	δ_{cp}	$\delta_{cp,cd}$
δ	0.9970	0.7603	0.4324	0.0175*	0.0226*
δ_{cd}	0.9415	0.3417		0.0321*	0.0366*
δ_{cp}	0.6822	0.0136*			0.6034
$\delta_{cp,cd}$	0.7461	0.0157*			
α	1.2824	0.0000***			

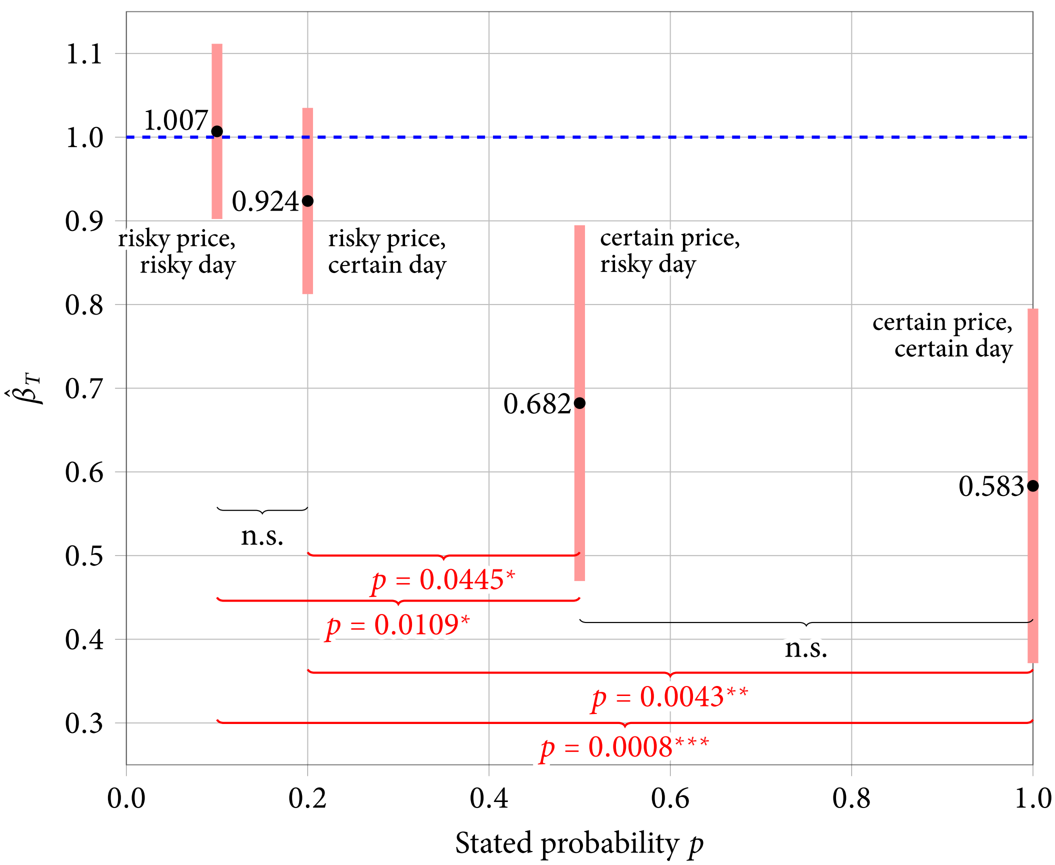
Notes: N = 897 observations from G = 180 subjects, with N_l = 160 left- and N_r = 89 right-censored observations. Excludes attrited subjects. Robust standard errors in parentheses, clustered on subject, from a two-limit Tobit model.
$^{+}p<0.10$, $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.

Table 8: Regression results from the restricted model

Parameter	Estimate	Hypothesis p-value that the parameter is equal to			
		1	β_{cd}	β_{cp}	$\beta_{cp,cd}$
β	1.0069	0.8965	0.3404	0.0109**	0.0008***
β_{cd}	0.9237	0.1766		0.0445*	0.0043**
β_{cp}	0.6822	0.0033**			0.4562
$\beta_{cp,cd}$	0.5833	0.0001***			
δ	0.9857	0.0006***			
α	1.2824	0.0000***			

Notes: N = 897 observations from G = 180 subjects, with N_l = 160 left- and N_r = 89 right-censored observations. Excludes attrited subjects. Robust standard errors in parentheses, clustered on subject, from a two-limit Tobit model.
$^{+}p<0.10$, $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.

Figure 10: Estimates of β_T from the restricted model



4.3. Parametric results for the restricted model

This section considers the model specified in Equation (8), with the additional restriction that δ is identical across treatments. Accordingly, a separate β_T is estimated for each treatment T, but only one δ is estimated for all four treatments. Because the Certain Price treatment had a different baseline time preference (as shown in Table 6 and discussed in Section 4.1), we may expect the β_{cp} estimate to be much different under this model, as it will partially compensate for the restriction on δ_{cp}. Given this restricted model, we obtain the regression results as in Figure 10 and Table 8.

With this restricted model, we again reject the null hypothesis that $\beta_{\text{cp,cd}} = \beta$ (now at $p = 0.0008$). This provides clear evidence that the elimination of all types of risk significantly increases present bias. Where we found only marginal evidence in the unrestricted model, we may now reject $\beta_{\text{cp,cd}} = \beta_{\text{cd}}$ (at $p = 0.0043$). As with the unrestricted model, we again fail to reject that $\beta_{\text{cp,cd}} = \beta_{\text{cp}}$; as before, this test seems underpowered. Nevertheless, we can confidently conclude that the elimination of all risk significantly increases average present bias, a result that is robust to both models.

This restricted model finds evidence that present bias differs for intermediate levels of risk. Specifically, we reject $\beta_{\text{cp}} = \beta_{\text{cd}}$ (at $p = 0.0445$) and also reject $\beta_{\text{cp}} = \beta$ (at $p = 0.0109$). These results come with the caveat that this model may be misspecified for the Certain Price, Risky Day treatment. Accordingly, we would best consider these results for intermediate levels of risk to be inconclusive.

4.4. Attrition

Let's consider the various stages at which a subject may quit the experiment. Recall that once a subject fails to complete a session of the experiment on the required day, she may not participate further. Of the 220 subjects who completed the intake questionnaire correctly, 208 enrolled in and completed the

Day 0 session. Accordingly, the remaining 12 were excluded from further participation. On Day 2, 192 subjects returned and completed that day's session, with again the remainder becoming excluded. On Day 9, 180 subjects completed the session, representing 87% retention of the subjects who completed the Day 0 session (see Table 13 for a complete tabulation).

We should be concerned that attrition might be driven by resolution of either dimension of risk. For example, on Day 2, if the subject learns that $R = 0.5$ has been randomly selected, she then must complete the most work possible (assuming she has monotonic preferences). We could imagine a subject deciding to quit the experiment after receiving such bad news. Alternately, which day's decisions are selected for implementation could also affect whether the subject does not continue.

Let's review the session timeline to consider the possibility of endogenous attrition. At the beginning of the Day 2 session, subjects in the Certain Day treatment receive resolution of day-uncertainty; these subjects are told whether their Day 0 choices or Day 2 choices will matter. Next, subjects in all treatments make allocation decisions. Then subjects with Risky Day treatment learn which day d is selected for its choices. Next, all Risky Price treatments learn which rate R_i is chosen. Finally, each subject s is asked to complete the amount of work they chose for their personally-drawn state of the world $\{i, d\}$. (This timeline is further described in Table 3.)

So let's look at whether subjects quit during the Day 2 session, after between the resolution of risk and before completing the real-effort tasks. *Two* subjects in this experiment quit during the Day 2 session—one subject after realizing $R = 1$ (the expected value), the other realizing $R = 1.25$ (better than expected). However, some subjects who completed Day 0 did not return on Day 2, and some subjects who completed Day 2 did not return for Day 9. This attrition could be selective—for example, a subject may draw $R = 0.5$ and choose to quit after the end of the Day 2 session. I predict attrition on either day $a = 2, 9$ by which state of the world (i, d) is drawn for subject s, using the OLS model

$$\mathbb{1}(\text{attrition-day})_a = \ln R_i + \mathbb{1}(\text{decision-day } d = 0)_s + \varepsilon_{i,d,s}.$$

Estimates show that neither the price R_i nor the decision-day d predicts attrition for Day 2 or Day 9.

All subjects who do not return on either Day 2 or Day 9 are excluded entirely from further analysis. I exclude these subjects because attrition may regardless be confounded with our variables of interest. For example, extreme present bias could induce some subject to not return on Day 2; if these subjects are also sophisticated regarding their time preferences, their Day 0 decisions may also be affected. A variety of such concerns exist, so I simply exclude any subjects who attrit.

4.5. Other considerations

Effort curvature Using the power functional form $c(e) = (e + 10)^{\alpha}$ for effort cost on a day with chosen effort e and background effort of 10 tasks, we reject the hypothesis that $\alpha \leq 1$ in all estimations. This ensures us that subjects are indeed minimizing effort costs (our second-order condition).

Survival We may easily interpret a single pooled estimate of δ as a continuation probability. Of the subjects who completed Day 2, 93.75% went on to complete Day 9. The weekly discount factor δ^7 can be interpreted as a weekly survival probability; if a subject is sophisticated about her likelihood of remaining in the project, this weekly discount factor would represent her belief that she would continue to participate from Day 2 to Day 9. I estimate this parameter as $\hat{\delta}^7 = 0.9202$ with a 95% confidence interval of $(0.8926, 0.9477)$. Because the retention rate falls into this confidence interval, I cannot reject the hypothesis that subjects are (in aggregate) sophisticated regarding their likelihood of continuing in the project between Days 2 and 9.

Some treatments include ten observations per subject, and others, two. Section A.1 shows that my primary result holds even when using only data for $R_i = 1.25$, so that each treatment has exactly two observations per subject. Section A.2 demonstrates that my results hold for a wide range of background effort levels as well.

4.6. Within-subject estimation

This within-subject regression model permits subject heterogeneity in β_s and δ_s for each subject s, while effort-cost curvature α is estimated from all observations pooled together. Because many subjects only face one or two incentivized decisions, estimation is highly collinear, and thus the parameter estimates are statistically imprecise. The distribution of $\hat{\beta}_s$ across subjects (as depicted in Figure 11) suggests that most subjects in all treatments have β_s near unity. However, subjects in the certainty treatments are more likely to have lower values of β_s. Notably, a substantial proportion of subjects in the Certain Price, Certain Day treatment have $\hat{\beta}_s < 0.5$. Many of these subjects chose a non-trivial amount of work for Day 2 when asked on Day 0, but then revised this amount significantly downward on Day 2. This behavior is observed far more often in the Certain Price, Certain Day treatment than the other treatments.

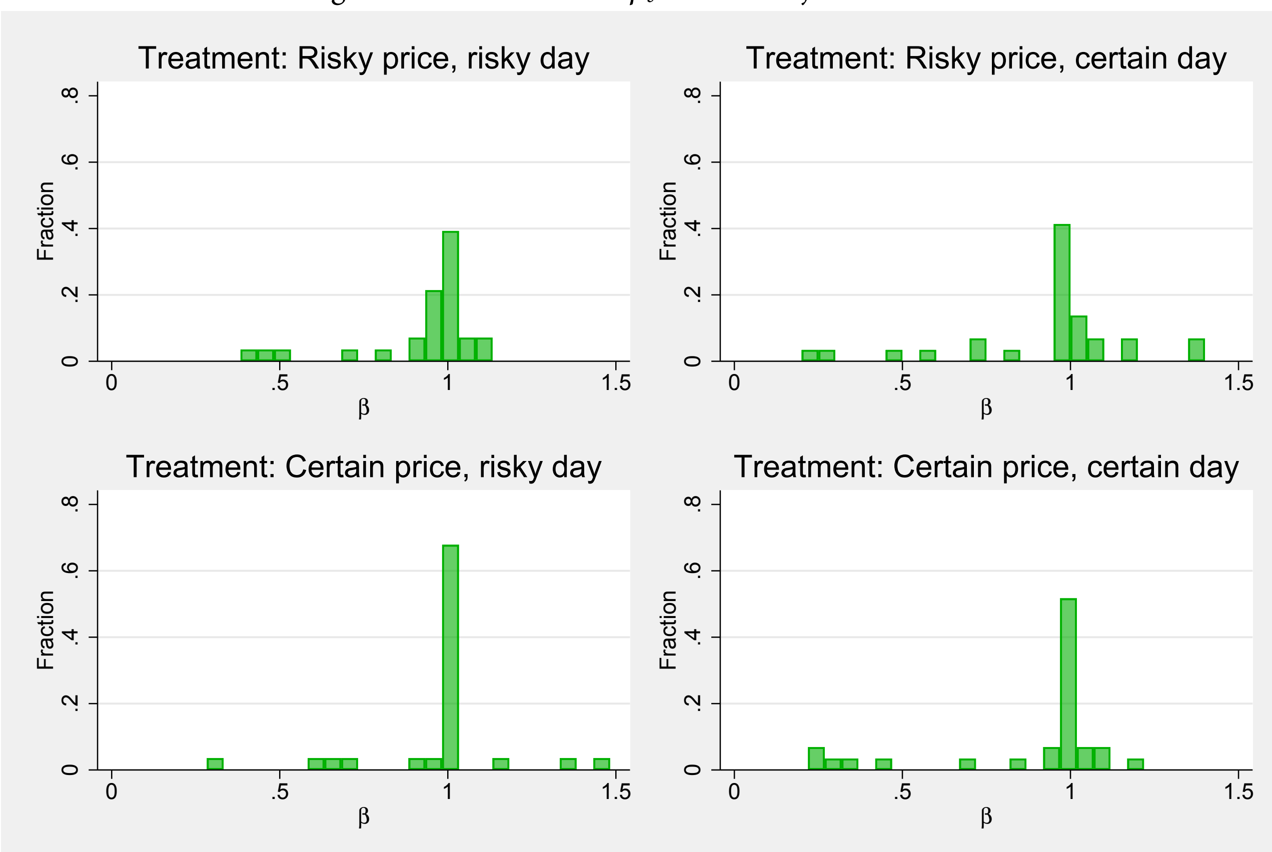
Table 9: Subjects with $\hat{\beta}_s < 1$ by treatment

Treatment	Risky Day	Certain Day
Risky Price	15 of 29	12 of 29
Certain Price	7 of 30	9 of 29

Notes: From a structural fixed-effects regression which estimates $\hat{\beta}_s$ and $\hat{\delta}_s$ for each subject (using only incentivized choice data).

We may also use these within-subject estimates to classify subjects as present-biased if $\hat{\beta}_s < 1$; the proportions of present-biased subjects by treatment are tabulated in Table 9. We previously looked at the extent of present bias by comparing only the choices made at $R_2 = 1.25$, as were presented in Table 5. The average present-bias factor is now estimated for each subject using her choices made at each of the five prices (if incentivized). With this approach, we now find a greater extent of present bias in the Risky Price treatments. This provides further evidence that the certainty treatments do not give rise to additional subjects behaving in a present-biased manner; instead, the subjects who are present-biased are more severely present-biased.

Figure 11: Distribution of β_s within-subject estimates



5. Conclusion

This study of dynamic inconsistency in real-effort task provision finds that risk diminishes the intensity of present bias. This seems to occur in two dimensions of uncertainty—the day from which a decision is chosen, as well as the price ratio at which a decision is implemented. Ultimately these results are consistent with the previous empirical evidence that demonstrates that the immediacy effect is moderated or eliminated by uncertainty.

The two channels of uncertainty used in this experiment are commonly used as part of popular mechanisms of experimental methodology. Accordingly, researchers should carefully consider these empirical results when designing experiments. Use of particular incentive-compatible mechanisms may confound results, especially near certainty.

A robust interaction between the immediacy effect and certainty effect is found, regardless of restrictions made on the pure discount factor δ. Finally, my estimates of present bias under certainty are far more severe than many recent experimental estimates using randomizing mechanisms. Accordingly, individuals in such studies may be much more myopic than previously thought.

Funding statement
The Department of Economics and the Graduate Division at the University of California, Santa Barbara, provided funding.

Acknowledgments
I thank my dissertation committee of Gary Charness, Ignacio Esponda, and Erik Eyster for their dedication and encouragement. Thanks also to David Freeman, Yoram Halevy, Peter Kuhn, and participants at UCSB seminars and Economic Science Association meetings.

References

Allais, Maurice. 1953. "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine." *Econometrica* 21 (4): 503–546.

Andreoni, James, and Charles Sprenger. 2012. "Estimating time preferences from convex budgets." *American Economic Review* 102 (7): 3333–56.

Augenblick, Ned. 2018. "Short-term time discounting of unpleasant tasks." http://faculty.haas.berkeley.edu/ned/Augenblick_ShortTermDiscounting.pdf.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger. 2015. "Working over time: Dynamic inconsistency in real effort tasks." *Quarterly Journal of Economics* 130 (3): 1067–1115.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy. 2018. "Incentives in experiments: A theoretical analysis." *Journal of Political Economy* 126 (4): 1472–1503.

Baucells, Manel, and Franz H. Heukamp. 2010. "Common ratio using delay." *Theory and Decision* 68 (1–2): 149–158.

———. 2012. "Probability and time trade-off." *Management Science* 58 (4): 831–842.

Beattie, Jane, and Graham Loomes. 1997. "The impact of incentives upon risky choice experiments." *Journal of Risk and Uncertainty* 14 (2): 155–168.

Brown, Alexander L., and Paul J. Healy. 2018. "Separated decisions." *European Economic Review* 101 (January): 20–34.

Chakraborty, Anujit, Yoram Halevy, and Kota Saito. 2020. "The relation between behavior under risk and over time." *American Economic Review: Insights* 2 (1): 1–16.

Charness, Gary, Uri Gneezy, and Brianna Halladay. 2016. "Experimental methods: Pay one or pay all." *Journal of Economic Behavior & Organization* 131A (November): 141–150.

Cox, James C., Vjollca Sadiraj, and Ulrich Schmidt. 2015. "Paradoxes and mechanisms for choice under risk." *Experimental Economics* 18 (2): 215–250.

Cubitt, Robin P., Chris Starmer, and Robert Sugden. 1998. "On the validity of the random lottery incentive system." *Experimental Economics* 1 (2): 115–131.

Epper, Thomas, and Helga Fehr-Duda. 2018. "The missing link: Unifying risk taking and time discounting." Working Paper 96. University of Zurich, Department of Economics. https://www.zora.uzh.ch/id/eprint/66623/.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. "Time discounting and time preference: A critical review." *Journal of Economic Literature* 40 (2): 351–401.

Freeman, David J., Yoram Halevy, and Terri Kneeland. 2019. "Eliciting risk preferences using choice lists." *Quantitative Economics* 10 (1): 217–237.

Freeman, David J., and Guy Mayraz. 2019. "Why choice lists increase risk taking." *Experimental Economics* 22 (1): 131–154.

Halevy, Yoram. 2015. “Time consistency: Stationarity and time invariance.” *Econometrica* 83 (1): 335–352.

Harrison, Glenn W., and John A. List. 2004. “Field experiments.” *Journal of Economic Literature* 42 (4): 1009–1055.

Imai, Taisuke, Tom A. Rutter, and Colin F. Camerer. 2021. “Meta-analysis of present-bias estimation using convex time budgets.” *Economic Journal* 131 (636): 1788–1814.

Keren, Gideon, and Peter Roelofsma. 1995. “Immediacy and certainty in intertemporal choice.” *Organizational Behavior and Human Decision Processes* 63 (3): 287–297.

Laibson, David. 1997. “Golden eggs and hyperbolic discounting.” *Quarterly Journal of Economics* 112 (2): 443–478.

Laibson, David, Andrea Repetto, and Jeremy Tobacman. 2007. *Estimating discount functions with consumption choices over the lifecycle.* Technical report. National Bureau of Economic Research.

Prelec, Drazen, and George Loewenstein. 1991. “Decision making over time and under uncertainty: A common approach.” *Management Science* 37 (7): 770–786.

Rana, Preetika. 2020. “Uber retools California fares in response to gig economy law; Move is latest change by ride-hailing giant to bolster case that drivers are independent workers.” *Wall Street Journal (Online)* (January). https://www.wsj.com/articles/uber-retools-california-fares-in-response-to-gig-economy-law-11578527202.

Reddinger, J. Lucas. 2019. “Temptation: Immediacy and certainty.” AEA RCT Registry. https://doi.org/10.1257/rct.4651.

Samuelson, Paul A. 1937. “A note on measurement of utility.” *Review of Economic Studies* 4 (2): 155–161.

Starmer, Chris, and Robert Sugden. 1991. “Does the random-lottery incentive system elicit true preferences? An experimental investigation.” *American Economic Review* 81 (4): 971–978.

Weber, Bethany, and Gretchen Chapman. 2005. “The combined effects of risk and time on choice: Does uncertainty eliminate the immediacy effect? Does delay eliminate the certainty effect?” *Organizational Behavior and Human Decision Processes* 96 (2): 104–118.

Appendix A Additional results and experimental details

Subsection A.1 Robustness to a single choice datum for each subject-day

Recall that some subjects made five incentivized decisions on each decision-day, while subjects in other treatments only made one incentivized decision on each decision-day. Here I show that my main result holds using only the choice data for $R = 1.25$. Because the variation in R jointly identified δ and α, we will not be able to estimate these parameters. I instead use $\widehat{\alpha} = 1.2824$ as previously estimated. This approach is sufficient to test whether the β estimates differ between treatments.

To this end, let us define alternate non-linear transformations of β as follow.

$$\widetilde{\beta} = \exp\left((\widehat{\alpha} - 1)\theta_{\text{present}}\right) \qquad \widetilde{\beta}_{\text{cp}} = \exp\left((\widehat{\alpha} - 1)\left(\theta_{\text{present}} + \theta_{\text{pr-cp}}\right)\right)$$

$$\widetilde{\beta}_{\text{cd}} = \exp\left((\widehat{\alpha} - 1)\left(\theta_{\text{present}} + \theta_{\text{pr-cd}}\right)\right) \quad \widetilde{\beta}_{\text{cp,cd}} = \exp\left((\widehat{\alpha} - 1)\left(\theta_{\text{present}} + \theta_{\text{pr-cp}} + \theta_{\text{pr-cd}} + \theta_{\text{pr-cp,cd}}\right)\right)$$

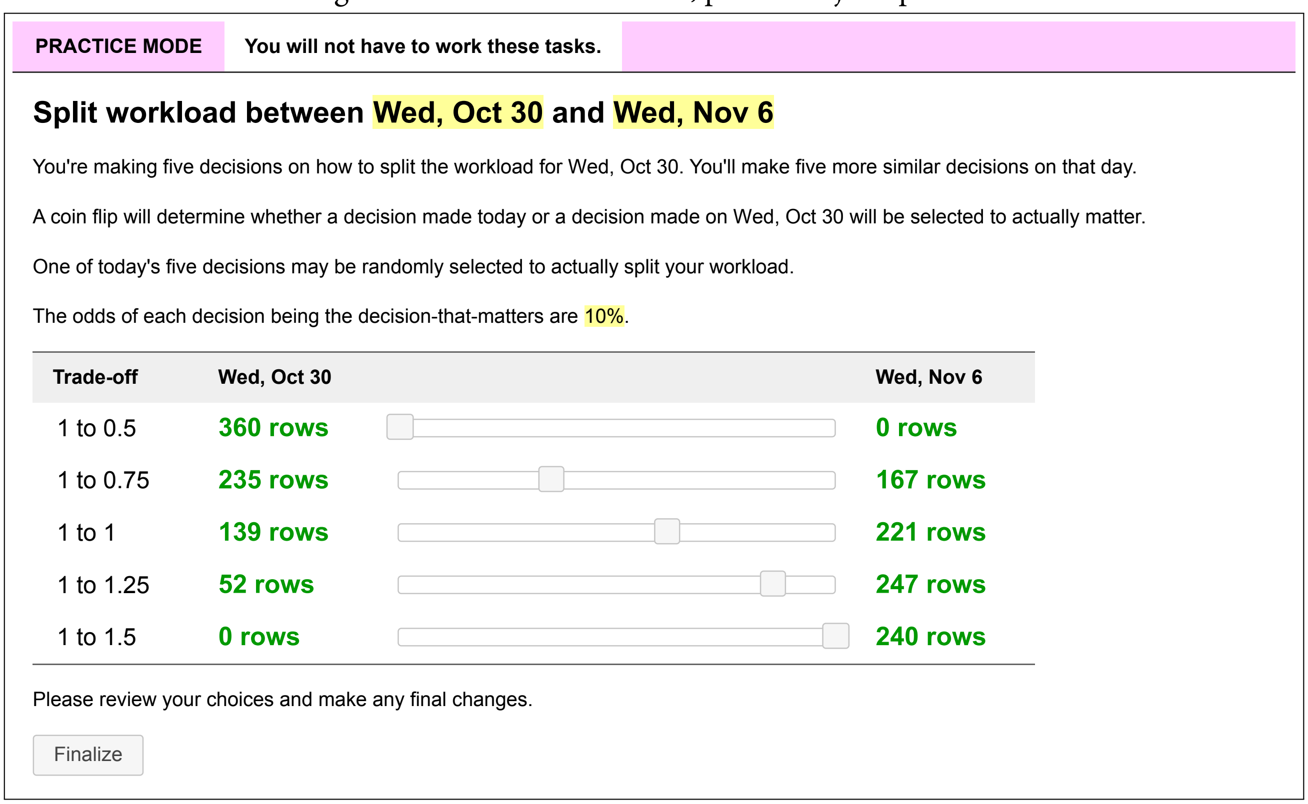
We may then simply test whether these differ, thus demonstrating a treatment effect with this data restriction. Panel B of Table 10 shows that the primary result holds—the treatment with certain price and day differs from the treatment with risky price and day. Note that some of the secondary results lose significance. The tests involving $\widetilde{\beta}$ and $\widetilde{\beta}_{\text{cd}}$ suffered the greatest loss of power, with the number of observations per subject in these treatments decreasing from ten to two. Despite this loss of power, the primary result remains highly significant.

Table 10: Result robustness to the restriction that $R = 1.25$

Parameter	Hypothesis p-value that the parameter is equal to		
PANEL A: ALL INCENTIVIZED CHOICE DATA (897 OBS., 180 SUBJECTS)			
	β_{cd}	β_{cp}	$\beta_{cp,cd}$
β	0.3404	0.0109**	0.0008***
β_{cd}		0.0445*	0.0043**
β_{cp}			0.4562
PANEL B: ONLY INCENTIVIZED CHOICES AT $R = 1.25$ (242 OBS., 121 SUBJECTS)			
	$\widetilde{\beta}_{cd}$	$\widetilde{\beta}_{cp}$	$\widetilde{\beta}_{cp,cd}$
$\widetilde{\beta}$	0.2717	0.0518+	0.0083**
$\widetilde{\beta}_{cd}$		0.3751	0.1052
$\widetilde{\beta}_{cp}$			0.4547

Notes: Excludes attrited subjects. Robust standard errors in parentheses, clustered on subject, from a two-limit Tobit model. $^{+}p<0.10$, $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.

Figure 12: Allocation interface, presented juxtaposed



Subsection A.2 Robustness to various levels of background effort

We have thus far only considered a background effort level of $\omega = 10$ tasks for the log-effort-ratio

$$E_{i,d} \coloneqq \ln \frac{e_{i,d}^{\text{Day 2}} + \omega}{e_{i,d}^{\text{Day 9}} + \omega}, \quad \text{with background effort } \omega. \tag{9}$$

Recall that total effort e costs are $c(e) = (e + \omega)^{\alpha}$ within a day; both α and ω alter the fit of the curve to the elicited data. Results for the log-effort-ratio outcome are qualitatively similar for $\omega = 10, 100, 1\,000$, and $10\,000$, as shown in Tables 11 and 12.

Table 11: Results with other levels of background effort ω

	$\omega = 10$	$\omega = 1\,000$	$\omega = 10\,000$	$\omega = 100\,000$
α	1.298	4.667	33.60	322.7
	(0.052)	(0.673)	(6.135)	(60.77)
β	0.928	0.885	0.879	0.879
	(0.033)	(0.052)	(0.055)	(0.055)
β_{cd}	0.898	0.861	0.854	0.853
	(0.053)	(0.070)	(0.074)	(0.074)
β_{cp}	0.887	0.881	0.879	0.878
	(0.094)	(0.096)	(0.098)	(0.098)
$\beta_{\text{cp,cd}}$	0.688	0.705	0.700	0.700
	(0.101)	(0.110)	(0.112)	(0.112)
δ	0.997	1.001	1.002	1.002
	(0.010)	(0.011)	(0.012)	(0.012)
δ_{cd}	0.941	0.925	0.924	0.924
	(0.062)	(0.074)	(0.077)	(0.078)
δ_{cp}	0.682	0.657	0.651	0.651
	(0.129)	(0.135)	(0.137)	(0.137)
$\delta_{\text{cp,cd}}$	0.746	0.697	0.691	0.690
	(0.105)	(0.109)	(0.111)	(0.112)
$\mathbb{P}(\beta = \beta_{\text{cd}})$	0.603	0.765	0.759	0.758
$\mathbb{P}(\beta = \beta_{\text{cp}})$	0.674	0.975	0.995	0.997
$\mathbb{P}(\beta = \beta_{\text{cp,cd}})$	0.018	0.112	0.120	0.121
$\mathbb{P}(\beta_{\text{cd}} = \beta_{\text{cp}})$	0.921	0.859	0.833	0.830
$\mathbb{P}(\beta_{\text{cp,cd}} = \beta_{\text{cd}})$	0.053	0.193	0.209	0.211
$\mathbb{P}(\beta_{\text{cp,cd}} = \beta_{\text{cp}})$	0.136	0.209	0.211	0.211
$\mathbb{P}(\beta = 1)$	0.029	0.027	0.027	0.027
$\mathbb{P}(\beta_{\text{cd}} = 1)$	0.056	0.047	0.047	0.047
$\mathbb{P}(\beta_{\text{cp}} = 1)$	0.231	0.217	0.215	0.214
$\mathbb{P}(\beta_{\text{cp,cd}} = 1)$	0.002	0.007	0.008	0.008
$\mathbb{P}(\delta = \delta_{\text{cd}})$	0.432	0.363	0.374	0.376
$\mathbb{P}(\delta = \delta_{\text{cp}})$	0.018	0.013	0.013	0.013
$\mathbb{P}(\delta = \delta_{\text{cp,cd}})$	0.023	0.008	0.008	0.008
$\mathbb{P}(\delta_{\text{cd}} = \delta_{\text{cp}})$	0.032	0.033	0.033	0.033
$\mathbb{P}(\delta_{\text{cp,cd}} = \delta_{\text{cd}})$	0.037	0.018	0.017	0.017
$\mathbb{P}(\delta_{\text{cp,cd}} = \delta_{\text{cp}})$	0.603	0.745	0.750	0.750
$\mathbb{P}(\delta = 1)$	0.760	0.901	0.873	0.871
$\mathbb{P}(\delta_{\text{cd}} = 1)$	0.342	0.310	0.325	0.327
$\mathbb{P}(\delta_{\text{cp}} = 1)$	0.014	0.011	0.011	0.011
$\mathbb{P}(\delta_{\text{cp,cd}} = 1)$	0.016	0.006	0.005	0.005

Notes: $N = 897$ observations from $G = 180$ subjects, with $N_l = 160$ left- and $N_r = 89$ right-censored observations. Excludes attrited subjects. Robust standard errors in parentheses, clustered on subject, from a two-limit Tobit model.

Table 12: Results with other levels of background effort ω, restricted model

	$\omega = 10$	$\omega = 1\,000$	$\omega = 10\,000$	$\omega = 100\,000$
α	1.282	4.447	31.59	302.9
	(0.045)	(0.579)	(5.267)	(52.15)
β	1.007	0.982	0.979	0.978
	(0.053)	(0.065)	(0.068)	(0.068)
β_{cd}	0.924	0.888	0.882	0.882
	(0.056)	(0.071)	(0.075)	(0.075)
β_{cp}	0.682	0.669	0.663	0.662
	(0.108)	(0.112)	(0.114)	(0.114)
$\beta_{cp,cd}$	0.583	0.573	0.566	0.565
	(0.108)	(0.112)	(0.114)	(0.114)
δ	0.986	0.987	0.987	0.987
	(0.004)	(0.005)	(0.005)	(0.005)
$\mathbb{P}(\beta = \beta_{cd})$	0.340	0.381	0.389	0.390
$\mathbb{P}(\beta = \beta_{cp})$	0.011	0.019	0.020	0.020
$\mathbb{P}(\beta = \beta_{cp,cd})$	< 0.001	0.002	0.002	0.002
$\mathbb{P}(\beta_{cd} = \beta_{cp})$	0.045	0.081	0.085	0.086
$\mathbb{P}(\beta_{cp,cd} = \beta_{cd})$	0.004	0.011	0.012	0.012
$\mathbb{P}(\beta_{cp,cd} = \beta_{cp})$	0.456	0.479	0.478	0.478
$\mathbb{P}(\beta = 1)$	0.896	0.781	0.751	0.748
$\mathbb{P}(\beta_{cd} = 1)$	0.177	0.116	0.115	0.115
$\mathbb{P}(\beta_{cp} = 1)$	0.003	0.003	0.003	0.003
$\mathbb{P}(\beta_{cp,cd} = 1)$	< 0.001	< 0.001	< 0.001	< 0.001

Notes: N = 897 observations from G = 180 subjects. Excludes attrited subjects. Robust standard errors in parentheses, clustered on subject, from a two-limit Tobit model.

Table 13: Subjects by treatment

	Treatment				Non-attrited subjects on		
	Day	Price	Selects[1]	Price Order[2]	Day 0	Day 2	Day 9
1	Risky	Risky	Day 2	A	8	8	8
2	Risky	Certain	Day 2	A	10	10	10
3	Risky	Risky	Day 0	A	9	8	7
4	Risky	Certain	Day 0	A	8	7	7
5	Certain	Risky	Day 2	A	17	16	15
6	Certain	Certain	Day 2	A	17	16	15
7	Certain	Risky	Day 0	A	18	16	16
8	Certain	Certain	Day 0	A	15	15	13
9	Risky	Risky	Day 2	B	9	8	7
10	Risky	Certain	Day 2	B	8	8	7
11	Risky	Risky	Day 0	B	9	8	8
12	Risky	Certain	Day 0	B	9	9	8
13	Certain	Risky	Day 2	B	17	15	14
14	Certain	Certain	Day 2	B	17	16	15
15	Certain	Risky	Day 0	B	18	16	15
16	Certain	Certain	Day 0	B	17	16	15
Total					206	192	180

Notes: Gray rows represent treatments in which subjects made no incentivized decision on Day 2, and thus are not used for analysis. [1] The day ultimately selected as the decision-day that matters. [2] Decisions are presented in one of two sequential orderings of R_i.