

# When Fair Isn't Fair: Sophisticated Time Inconsistency in Social Preferences\*

James Andreoni      Deniz Aydın      Blake Barton      B. Douglas Bernheim  
Jeffrey Naecker

March 22, 2016

## Abstract

How do people think about fairness in settings with uncertainty? One view holds that fairness requires equality of opportunity; another holds that it requires equality of outcomes. Relative to the resolution of uncertainty, the first view takes an *ex ante* perspective, while the second takes an *ex post* perspective. In this paper, we conduct a laboratory experiment designed to determine which perspective people adopt, and under what conditions. We find that most people view fairness from an *ex ante* perspective when making decisions *ex ante*, and from an *ex post* perspective when making decisions *ex post*. As a result, they exhibit the hallmark of time-inconsistency: after making an initial plan that is fully state-contingent, they revise it upon learning that certain states will not occur. These patterns are robust and persist even when people are aware of their proclivities. Indeed, subjects who switch from *ex ante* fair to *ex post* fair choices, and who are aware of this proclivity, generally avoid precommitments and intentionally retain the flexibility to manifest time inconsistency. We argue that these patterns are best explained by a theory of *nominal fairness*.

## 1 Introduction

A large and growing body of research establishes that people care to varying degrees about fairness. Classic experimental results include the tendency to divide a prize equally in the dictator game and reject lopsided offers in the ultimatum game.<sup>1</sup> Related behavioral patterns are commonly observed in the field.<sup>2</sup> And yet, despite its apparent importance, fairness is a slippery concept. In any given

---

\*We would like to thank participants at the 2015 SITE Psychology and Economics Workshop, the 2016 AEA meetings, and seminars at Texas A&M, Cornell, USC, and Claremont Graduate University for helpful comments. Andreoni: Department of Economics, UCSD, andreoni@ucsd.edu; Aydın: Department of Economics, Stanford University, daydin@stanford.edu; Barton: Department of Economics, Stanford University, blakeabarton@gmail.com; Bernheim: Department of Economics, Stanford University, bernheim@stanford.edu; Naecker: Department of Economics, Wesleyan University, jnaecker@wesleyan.edu.

<sup>1</sup>See, for example, Forsythe et al. (1994), Hoffman et al. (1996), Camerer (1997), Bohnet and Frey (1999), Andreoni and Miller (2002), Andreoni et al. (2003), Andreoni and Bernheim (2009), and Andreoni et al. (2002). Early attempts to model concerns about fairness include Rabin (1993), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Charness and Rabin (2002).

<sup>2</sup>For instance, equal sharing is common in the context of joint ventures among corporations (Veugelers and Kesteloot (1996), Dasgupta and Tao (1998), Hauswald and Hege (2003)), share tenancy in agriculture (De Weaver and Roumasset (2002), Agrawal (2002)), bequests to children (Wilhelm (1996), Menchik (1980, 1988), Bernheim and Severinov (2003)), and arbitration (Bloom (1986)).

context, people may disagree as to what constitutes a fair decision.

An important dimension of disagreement concerns the question of whether a fair society should pursue equality of opportunity or equality of outcomes. Those who favor standards based on equality of opportunity tend to view fairness from an *ex ante* perspective. They tolerate even highly unequal outcomes provided all parties had comparable shots at success. In contrast, those who favor standards based on equality of outcomes tend to think about fairness from an *ex post* perspective. They are particularly disinclined to accept disparities in circumstances that result from luck rather than choice.<sup>3</sup>

The current paper is motivated by the following simple observation: if the *ex ante* perspective on fairness is compelling *ex ante* and the *ex post* perspective is compelling *ex post*, then any particular individual may shift from the first to the second as events evolve. The resulting conflict between the perspectives adopted at different points in time would then give rise to time inconsistency. Moreover, if this type of inconsistency is common, it would have important practical implications. A society populated by such individuals would design policies to promote equality of opportunity, only to undermine those policies by consensus once winners and losers emerge.<sup>4</sup>

Several previous studies have examined whether people care about *ex ante* fairness, *ex post* fairness, or both.<sup>5</sup> However, to our knowledge, none has directly addressed the important empirical question of whether fairness preferences are time-inconsistent. We conduct a laboratory experiment in which each subject allocates chances to win cash between two impoverished Kenyan households. One of those households initially has a greater baseline chance of success, of which the subject is always aware. However, an event may occur that neutralizes the initial disparity. The subject makes decisions either before or after the neutralizing event. Time inconsistency will emerge in this context if a subject wishes to compensate for the initial disparity prior to the event, but not after it.

We structure our inquiry around five questions. First, for a newly encountered collection of allocation problems, does the framing of *initial* decisions determine the initial perspective on fairness? That is, when initial decisions are made *ex ante*, do they tend to be *ex ante* fair, and when they are made *ex post*, do they tend to be *ex post* fair? Our answer is a decisive “yes.”

Second, does the framing of initial decisions establish a *persistent* perspective on fairness? That is, if initial decisions are made *ex ante*, does the individual think through the application of their principles to this class of problems from an *ex ante* perspective, and continue to apply them in this manner even when they later face *ex post* choices? Conversely, if initial decisions are made *ex post*, does an *ex post* perspective on fairness dictate subsequent choices, even when they are made

---

<sup>3</sup>See Cappelen et al. (2013).

<sup>4</sup>Coate (1995) makes a similar point in a setting where the inconsistency arises from a different source (the Samaritan’s dilemma).

<sup>5</sup>References include Bolton et al. (2005), Karni et al. (2008), Krawczyk and Le Lec (2010), Kircher et al. (2013), Brock et al. (2013), and Cappelen et al. (2013).

ex ante? We find no evidence that perspectives on fairness are persistent. Instead, people switch readily between ex ante and ex post perspectives according to the framing of the current decision problem.

Our answers to these first two questions establish important prerequisites for time inconsistency, but do not necessarily imply it. A subject who switches between ex ante and ex post perspectives on fairness when moving from one allocation problem to another may still adopt a single perspective for a given allocation problem and adhere to it resolutely as events unfold.<sup>6</sup> Accordingly, we pose our third question: does concern for fairness generate time-inconsistency? We confirm that time inconsistency is prevalent. Indeed, the most common single pattern is for subjects to select the ex ante fair alternative ex ante, and switch to the ex post fair alternative ex post, thereby undermining their apparent ex ante objectives.

Conceivably, time inconsistency may result from lack of familiarity with the choice setting. As subjects gain experience, they may notice their inconsistency, think through their principles more carefully, and adopt a consistent perspective. This possibility motivates our fourth question: does self-awareness ameliorate time-inconsistency? We find that it does not.

As is well-understood in the literature on time-inconsistency, a sophisticated individual who anticipates undermining her own objectives should manifest a preference for commitments. Accordingly, we pose a fifth question: does self-awareness of time-inconsistency regarding fairness create such a preference? While we find that many subjects do take advantage of opportunities to make commitments, that finding is heavily qualified. Most importantly, the subjects who are prone to exhibit the characteristic time-inconsistent choice pattern (switching from ex ante to ex post fair allocations) also manifest a strong preference for retaining flexibility, rather than for making commitments.

As we explain, our findings are collectively difficult to reconcile with existing formulations of fairness preferences. They point instead to a theory of *nominal fairness*, in which people evaluate the ethics of each choice based on the consequences it most directly implies, under the assumption that nothing will subsequently happen to overturn it. To be clear, this theory does not posit a lack to sophistication or foresight. Rather, it holds that, from an ethical perspective, each action must stand on its own.

The rest of this paper is organized as follows. We provide conceptual background in Section 2, and explore the implications of several alternative theories for the decision tasks we subsequently investigate. We then describe our experiment in Section 3. Each of Sections 4 through 8 addresses one of the questions listed above in settings involving the allocation of chances to win a prize. Section 9 explores a related setting involving the allocation of dollars. Section 10 summarizes implications for various theories of fairness, and Section 11 concludes.

---

<sup>6</sup>In the pertinent literature, such an individual is said to have *resolute* preferences; see the discussion in Section 2.

## 2 Conceptual issues

### 2.1 Theoretical background

The potential tension between time consistency and preferences for fairness reflects two classic propositions. The first is that a preference for ex ante fairness is inconsistent with the independence axiom.<sup>7</sup> To understand why, imagine that you are asked to divide a handful of lottery tickets – only one of which is a winning ticket – between two equally deserving recipients. Does fairness dictate a preference for equal division of the tickets? If you believe it does, then your preferences are not linear in probabilities, which means they fall outside of the expected utility (EU) paradigm. Within the EU paradigm, the independence axiom is what delivers that linearity property.

The second classic proposition is that, subject to a qualification discussed below, EU preferences are time consistent while non-EU preferences are not.<sup>8</sup> It is important at the outset to clarify the definition of time inconsistency in contexts with uncertainty. Confusion sometimes arises because, in many settings, even a time-consistent agent will revise initial choices upon receiving additional information. However, the key point is that she will not revise choices that are contingent upon the receipt of that information. To illustrate, imagine planning out the choices you would like to make at the various nodes of a standard decision tree. For each node, you ask yourself the following question: what would I like my future self to do if it finds itself at this node? Now imagine yourself working through the decision tree in real time. If you are time consistent, you should be willing to follow through with the contingent plan you formulated before making your first choice.

Once one has that notion of time consistency in mind, it is easy to see the intuitive connection to the independence axiom. In effect, as one moves through the tree and observes some of nature’s choices, one rules out other realizations. Time consistency requires that the preferences governing choices from any given node forward are independent of the probability with which the node is reached, as well as the consequences of following any other positive-probability path.

Putting these two classic results together, one is naturally led to the inference that a preference for ex ante fairness necessarily generates time inconsistency. However, as Machina (1989) has pointed out, the purported inconsistency of non-EU preferences hinges on a consequentialist interpretation of the theory. On the one hand, if a non-EU decision maker focuses only on opportunities and outcomes, in effect “snipping off” what remains of the tree upon reaching any given node and treating the continuation problem as an entirely new decision tree, then indeed she will exhibit time inconsistency. On the other hand, if she takes past uncertainty (risks already borne) into account in a manner consistent with her original preferences, she will be time-consistent. People

---

<sup>7</sup>Classic references include Harsanyi (1955) and Diamond (1967). For a more recent perspective, see Fudenberg and Levine (2012).

<sup>8</sup>Classic references include Markowitz (1968) and Raiffa (1968). See also the excellent discussion in Machina (1989), who emphasizes that this proposition hinges on a consequentialist interpretation of non-EU theory. This is an important qualification, and we return to it below.

who behave in that manner are called *resolute*.<sup>9</sup>

Machina (1989) offers the following appealing illustration of resolute non-EU preferences. Mom has two children, Ben and Abby, as well as a single treat. She cares about outcome fairness and would ideally split the treat between them, but regrettably it is indivisible, so she must give it to one or the other. She also cares about ex ante fairness, and therefore strictly prefers a coin flip over either sure outcome. Mom flips the coin, and Abby wins. After pouting briefly, Ben has a sudden inspiration: he points out to Mom that, in light of her stated and revealed preferences, she would be better off flipping the coin again. Mom’s response: “sorry kid, you had your chance.” In this example, Mom cares about both ex ante and ex post fairness, but her preferences are resolute, so she is time-consistent.

Thus, as long as one is not wedded to consequentialism, a concern for ex ante fairness does not necessarily imply time inconsistency. Someone who has carefully thought through her principles and noticed the potential for inconsistency may choose to apply them resolutely. The issue is inherently empirical, which is why we turn to experimental evidence.

## 2.2 Applications

For the sake of clarity, it is useful to elaborate on the applicable conceptual principles within the context of the classes of decision problems we study experimentally in subsequent sections. Accordingly, we consider allocation tasks involving two equally deserving recipients,  $A$  and  $B$ . One class of tasks involves the allocation of chances to win a prize, while the other involves the allocation of dollars.

At the outset of the first type of task,  $N$  lottery tickets are allocated arbitrarily to one of the parties; assume  $B$  is the recipient. After observing this allocation, a third party (the decision maker) must distribute  $N$  additional tickets. One of the  $2N$  tickets is drawn at random and the winner receives a cash prize. (In our experiment,  $N = 10$ .)

At the outset of the second type of task, an arbitrary contingent allocation of a fixed dollar prize between the two parties is specified; assume it designates  $B$  as the recipient of the entire prize. After observing this allocation, a third party (the decision maker) must specify an alternate allocation. A coin flip determines whether the fixed allocation or the chosen alternative is implemented.

We distinguish between two versions (ex ante and ex post) of each of these tasks. In the ex post version, the decision maker selects her allocation once she learns whether her decision matters. Specifically, for the split-the-tickets task, she chooses only after learning that the winning ticket is one of hers (but not which one it is); for the split-the-prize task, she chooses only after learning that her allocation will be implemented. In the ex ante version, the decision maker selects her

---

<sup>9</sup>The phrase “resolute preferences” appears to originate with McClennen (1989), but there are earlier antecedents. See the discussion in Machina (1989)

allocation prior to receiving this additional information.

We acknowledge that the ex ante and ex post labels are incomplete. What we call the ex ante position is ex post the realization of implicit uncertainty governing the fixed allocation. There is also an *ex ante alia* position prior to that realization. Moreover, once the decision maker learns whether her choice matters, what we call the ex post position becomes an ex ante position relative to any remaining uncertainty. In the split-the-tickets task, there is also an *ex post alia* position. For the most part, these observations pertain only to labeling. However, there is one substantive exception, which we discuss below.

A time-consistent decision maker will make identical choices in the ex ante and ex post versions of our canonical tasks, regardless of whether they involve the allocation of lottery tickets or the contingent allocation of dollars. For the split-the-dollar tasks, the explanation for this proposition is immediate: the ex ante version requires the decision maker to specify precisely the same contingent allocation, but in advance of the realization of the relevant contingency. For the split-the-tickets tasks, the point is slightly more subtle.

The key to understanding the relationship between the ex ante and ex post split-the-tickets tasks is noticing that the consequences of all choices in the ex ante task are contingent to begin with. Provided people care about chances to win rather than tickets per se, there is no substantive difference between giving a particular ticket to a particular party at the outset of the ex ante task, and providing instructions that the same party should receive the prize if the ticket in question is chosen. The ticket already has the relevant contingency property built into it. In effect, in the ex ante task, the decision maker is already specifying a plan for allocating chances to win contingent on the event that one of her tickets turns out to be the winner. See the Appendix, Section A.1, for a more detailed explanation.

Next we explore how decisions makers with different types of preferences would behave in these settings. We will describe preferences within a somewhat broader class of environments, and then investigate implications for the tasks of interest. Accordingly, imagine that the state of nature is drawn from some finite set  $S$ , and that  $s \in S$  materializes with probability  $p(s) = \frac{1}{|S|}$ . For split-the-tickets tasks,  $s$  represents the selection of a particular lottery ticket; for split-the-prize tasks, it indicates whether a fixed allocation or a chosen alternative is implemented. We will use  $x_i(s) \geq 0$  to denote the payment received by party  $i$  in state  $s$ , and  $L = (p(s), x_A(s), x_B(s))_{s \in S}$  to denote the resulting lottery. States are divided into two categories,  $S_F$  and  $S_D$ , with  $|S_F| = |S_D|$ . The allocations  $(x_A(s), x_B(s))$  are fixed for  $s \in S_F$ , and set by the decision maker for  $s \in S_D$ , subject to task-specific constraints. In split-the-tickets tasks,  $(x_A(s), x_B(s))$  must equal either  $(c, 0)$  or  $(0, c)$  depending on whether  $A$  or  $B$  holds the pertinent ticket, where  $c$  is the cash prize; in split-the-prize tasks,  $x_A(s) + x_B(s) = c$ . The decision maker evaluates lotteries according to an objective function  $V$ . Ex post decisions are made after a signal reveals whether  $s$  lies in  $S_F$  or  $S_D$ ; ex ante decision

are made prior to any such revelation.

For a summary of the implications discussed below, see Table 1.

### 2.2.1 EU preferences.

First suppose the decision maker has EU preferences. In that case, she is necessarily time consistent. To understand why, observe that her objective function is

$$V(L_a) = \frac{1}{|S|} \left[ \sum_{s \in S_F} u(x_A(s), u_B(s)) + \sum_{s \in S_C} u(x_A(s), u_B(s)) \right]$$

when making decisions ex ante, and

$$V(L_p) = \frac{1}{|S_D|} \sum_{s \in S_C} u(x_A(s), x_B(s))$$

when making them ex post, where  $L_a$  and  $L_p$  are, respectively, the ex ante and ex post lotteries, and  $u$  is a Von Neumann-Morgenstern (VNM) utility function. In either case, the best choice is obviously the same: select the allocation in each state  $s \in S_C$  so as to maximize  $u(x_A(s), x_B(s))$ .

Notice that the decision of an EU maximizer is unresponsive to variations in the fixed allocations. Assuming  $u$  is asymmetric, she gives all of the tickets to the same recipient in split-the-tickets tasks, and skews the distribution in favor of one recipient in all split-the-prize tasks. Assuming  $u$  is symmetric, she is indifferent between all choices in split-the-tickets tasks; if in addition  $u$  is concave, she divides dollars equally in all split-the-prize tasks. As a general matter she is indifferent with respect to making ex ante commitments, which would have no effect on outcomes, with the following qualification. In split-the-tickets tasks, if  $u$  is symmetric, the decision maker might resolve her indifference among allocations in favor of switching, which commitments would suppress. In that case she might give the false appearance of both time inconsistency and a preference for commitment. However, this possibility involves a highly fortuitous resolution of indifference, and in any case one can rule it out by determining whether her preferences are strict.

### 2.2.2 Consequential Non-EU preferences

Next we will assume the decision maker has non-EU preferences. In that case, if she is also a consequentialist, she will manifest time inconsistency, at least in certain circumstances. Specific behavioral implications will depend on the nature of the departure from the EU paradigm. Here we consider two specific alternatives.

**EU-Distribution preferences** One natural hypothesis is that, at any given point in time, concerns about fairness pertain to the distribution of expected utility, so that preferences are governed by a utility function of the form  $W(EU_A, EU_B)$ . We will focus on symmetric cases in which the

**Table 1:** Implications of various theories.

Base	Preference	Split-the-tickets tasks				Split-the-prize tasks			
	Variant	Ex ante	Ex post	Revision	Commitment	Ex ante	Ex post	Revision	Commitment
EU		Indifferent	Indifferent	Indifferent	Indifferent, no effect	50-50, no offset	50-50, no offset	No	Indifferent, no effect
EU distribution	Consequential	Full offset	50-50, no offset	Yes	Yes, has effect	Partial offset, full if lex.	50-50, no offset	Yes	Yes, has effect
	Consequential & anticipatory	Partial offset	50-50, no offset	Yes	Yes, has effect	Partial offset	50-50, no offset	Yes	Yes, has effect
	Resolute ex ante	Full offset	Full offset	No	Indifferent, no effect	Partial offset, full if lex.	Partial offset, full if lex.	No	Indifferent, no effect
	Resolute ex post	50-50, no offset	50-50, no offset	No	Indifferent, no effect	50-50, no offset	50-50, no offset	No	Indifferent, no effect
	Resolute ex ante alia	Indifferent (sym. rules)	Indifferent (sym. rules)	Indifferent (sym. rules)	Indiff., may have effect	50-50, no offset	50-50, no offset	No	Indifferent, no effect
	Resolute anticipatory	Partial offset	Partial offset	No	Indifferent, no effect	Partial offset	Partial offset	No	Indifferent, no effect
	Nominally fair	Full offset	50-50, no offset	Yes	No	Partial offset, full if lex.	50-50, no offset	Yes	No
Probability weighting	Consequential	Full offset	50-50, no offset	Yes	Yes, has effect	50-50, no offset	50-50, no offset	No	Indifferent, no effect
	Consequential & anticipatory	Partial offset	50-50, no offset	Yes	Yes, has effect	50-50, no offset	50-50, no offset	No	Indifferent, no effect
	Resolute ex ante	Full offset	Full offset	No	Indifferent, no effect	50-50, no offset	50-50, no offset	No	Indifferent, no effect
	Resolute ex post	50-50, no offset	50-50, no offset	No	Indifferent, no effect	50-50, no offset	50-50, no offset	No	Indifferent, no effect
	Resolute ex ante alia	Indifferent (sym. rules)	Indifferent (sym. rules)	Indifferent (sym. rules)	Indiff., may have effect	50-50, no offset	50-50, no offset	No	Indifferent, no effect
	Resolute anticipatory	Partial offset	Partial offsetting	No	Indifferent, no effect	50-50, no offset	50-50, no offset	No	Indifferent, no effect
	Nominally fair	Full offset	50-50, no offset	Yes	No	50-50, no offset	50-50, no offset	No	Indifferent, no effect

Note: All implications listed in this table assume symmetric preferences between the two recipients.



decision maker intrinsically favors neither household. For the sake of simplicity, we will also assume separability. These assumptions allow us to write the utility derived from a given lottery  $L$  as follows:

$$V(L) = w \left( \sum_{s \in S} p(s) u(x_A(s)) \right) + w \left( \sum_{s \in S} p(s) u(x_B(s)) \right)$$

In the special case where the function  $w$  is linear, this formulation reduces to EU with separability over potential recipients. By assuming that  $w$  is concave, we introduce a preference for equalizing the expected utility derived by the two recipients.

To understand why this formulation gives rise to time inconsistency, suppose for simplicity that  $B$  receives all tickets associated with the fixed allocation; similar observations apply to other cases. Ex ante, the decision maker will give all of her tickets to  $A$  so as to compensate exactly for the fixed allocation. In contrast, ex post, if she discovers that the winning ticket is one of hers, she will assign the same number of tickets to  $A$  and  $B$ . Assuming she is sophisticated and knows she will have the opportunity to revisit her ex ante choice, she will anticipate the revision and draw two inferences: first, her initial choice is irrelevant; second,  $B$ 's likelihood of winning is 75%, which is suboptimal from her ex ante perspective. Consequently, she will prefer to make a commitment, giving all her tickets to  $A$  while renouncing the subsequent revision opportunity.

For split-the-prize tasks, the decision maker *partially* offsets the fixed allocation ex ante. It is straightforward to verify that the curvature of  $w$  determines the degree of the offset, and that perfect offset emerges in the limiting lexicographic case; see the Appendix, Section A.2. Equal division is optimal ex post, so the decision maker is time-inconsistent and has a preference for commitment.

**Probability weighting** For any particular class of decision problems, violations of EU are necessary but not sufficient for time inconsistency. That said, it is difficult to imagine a natural class of non-EU preferences that would robustly yield time consistency in split-the-tickets tasks, the reason being that such formulations are generally non-linear in probabilities, which change as event unfold. In contrast, it is easy to identify an arguably natural class of non-EU preferences that robustly yield time consistency in split-the-prize tasks. The general property that gives rise to this implication is separability over states of nature.

One can allow for non-linearities in probability while preserving separability over states of nature by introducing probability weighting, an assumption for which there is considerable precedent in the literature on risk and uncertainty (for example, Kahneman and Tversky (1979)). In the context of our split-the-tickets tasks, we would write the decision maker's utility function as

$$V(L) = \pi \left( \frac{|S_A|}{|S|} \right) u(c, 0) + \pi \left( \frac{|S_B|}{|S|} \right) u(0, c),$$

where  $S_A$  is the set of states in which  $A$  wins (that is, tickets allocated to  $A$ ),  $S_B$  is the set of states in which  $B$  wins (that is, tickets allocated to  $B$ ), and  $\pi$  is the probability weighting function. If  $\pi$  is linear, this formulation coincides with expected utility. But if  $\pi$  is concave, it gives rise to a preference for ex ante equality of opportunity. Indeed, if  $u$  is also symmetric, the decision maker will want to ensure that  $A$  and  $B$  have equal chances of winning the prize.

For split-the-tickets tasks, the optimum involves both potential recipients winning the prize with equal probability. Accordingly, the decision maker offsets the fixed allocation ex ante but not ex post, which creates time inconsistency. Assuming she knows she will have the opportunity to revisit her ex ante choice and that she is sophisticated (so that she anticipates the ultimate result), she will prefer to make a commitment, compensating for the fixed allocation while renouncing the subsequent revision opportunity.

Because this preference formulation is separable over states of nature, it has much different implications in the context of split-the-prize tasks. Ex ante, the decision maker chooses  $x$  to maximize<sup>10</sup>

$$V(L_a) = \pi\left(\frac{1}{2}\right)u(0, c) + \pi\left(\frac{1}{2}\right)u(c - x, x)$$

which is of course equivalent to maximizing  $u(c - x, x)$ , her ex post objective. Accordingly, she is unresponsive to the fixed allocation, and time-consistent. For example, in the special case where  $u$  is symmetric and concave, she will divide the prize equally both ex post and ex ante. As with the EU framework, she is indifferent with respect to making ex ante commitments, which would have no effect on outcomes.

### 2.2.3 Resolute non-EU preferences

So far, we have interpreted the various non-EU preference formulations discussed above through the lens of consequentialism. One can also interpret them from non-consequentialist perspectives. The most interesting possibility is that preferences are *resolute*. A resolute individual carefully thinks through the application of her principles to a particular decision problem, decides which of any competing perspectives is most compelling, and consistently evaluates her options based on a single standard.

Within our setting, a decision maker may adhere resolutely to either the ex ante or the ex post perspective on fairness. The former possibility involves some subtleties, in that she continues to take past uncertainty (risks already borne) into account even after it is resolved; see Machina (1989) for an incisive discussion.

A resolute individual may adopt different standards for fairness in different settings. For example, she may switch back and forth between ex ante fairness and ex post fairness when moving

---

<sup>10</sup>For the special case of  $x = c$ , this formula is not correct. Because this choice consolidates the two events, the correct formula is  $\pi(1)u(0, c)$ . If  $\pi$  is concave and  $u$  penalizes extreme inequality, this choice is suboptimal.

from one decision problem to another, depending on the initial decision frame in each new problem. However, for any given decision problem, she will exhibit time consistency even if she values ex ante fairness. For example, in the tasks we study, if she initially chooses to offset the fixed allocation, she will continue to adhere to that decision even after she learns that her choice will be implemented. Moreover, a resolute individual is indifferent with respect to making ex ante commitments, which would have no effect on outcomes, other than in cases involving indifference.

One special case merits further discussion: that in which the decision maker resolutely pursues fairness from the ex ante alia perspective. Assuming the initial (implicit) distribution of fixed allocations in a split-the-tickets task is equally favorable to  $A$  and  $B$ , any symmetric response to nature's choice leaves each party with a 50-50 chance of winning. To illustrate, suppose the decision maker always reinforces nature's choice, splitting her tickets to mimic the fixed allocation. Prior to nature's choice, each party will then have a 50% chance of winning, which is fair ex ante alia. Thus, the decision maker is indifferent with respect to all symmetric decision rules, whether offsetting, reinforcing, or insensitive to the fixed allocation. All of the preceding observations concerning resolute preferences continue to apply, except that migration from ex ante to ex post fairness does not necessarily rule out time consistency (because both allocation principles are maximally fair ex ante alia), nor does a commitment that deters such migration. However, to account for the signature patterns of time inconsistency, one would need to invoke a highly fortuitous resolution of indifference. In any case, one can rule out this possibility by determining whether the decision maker's distributional preferences are strict subsequent to the realization of the fixed allocation, or by examining behavior in split-the-prize tasks, where a strict ex ante preference for overall equality follows from concavity of  $u$ ; see the Appendix, Section A.2, for a demonstration.

#### 2.2.4 Anticipatory preferences

A sophisticated decision maker with non-EU preferences will anticipate any change in her perspective on fairness resulting from the realization of uncertainty. In that case, she may attach some weight to the utility or disutility that she will experience once she adopts that perspective (much as in Caplin and Leahy's (2001) model of anticipatory utility). As a result, both ex ante and ex post fairness may matter to her.

We will illustrate this possibility using the model of probability weighting described above. Let  $S_{iC}$  be the set of states in which the decision maker allocates the prize to  $i$ . If she places weight on both ex ante and ex post fairness, we can write her utility as

$$\begin{aligned}
 V(L_a) = & \lambda \left[ \pi \left( \frac{|S_{AC}|}{|S|} \right) u(c, 0) + \pi \left( \frac{|S_{BC}| + |S_F|}{|S|} \right) u(0, c) \right] \\
 & + (1 - \lambda) \left[ \pi \left( \frac{|S_{AC}|}{|S|} \right) u(c, 0) + \pi \left( \frac{|S_{BC}|}{|S|} \right) u(0, c) \right]
 \end{aligned}$$

where  $\lambda \in (0, 1)$ . Let  $x = \frac{|S_{AC}|}{|S_C|}$  represent the fraction of the decision maker's tickets that are allocated to  $A$ . Note that we can rewrite the preceding expression as

$$V(L_a) = \pi\left(\frac{x}{2}\right) u(c, 0) + \left[ \lambda\pi\left(1 - \frac{x}{2}\right) + (1 - \lambda)\pi\left(\frac{1 - x}{2}\right) \right] u(0, c)$$

Assuming  $u$  is symmetric and  $\pi$  is concave, it is easy to see that this expression is maximized at some  $x \in (\frac{1}{2}, 1)$ . In other words, the best choice lies between the ex ante and ex post fair alternatives. Thus the decision maker compensates for the initial fixed allocation partially, but not fully.

Assuming the decision maker is resolute, she adheres to the same objective and behaves consistently as events unfold. In contrast, assuming her preferences are consequential, her concern for fairness from the original ex ante perspective disappears once the ex post position is reached. At that point, the ex post perspective prevails, and she prefers to divide her tickets equally. As a result, her choices are time-inconsistent. Assuming she is sophisticated and knows she will have the opportunity to revisit her ex ante choice, she will prefer to make a commitment ex ante, giving her tickets disproportionately to  $A$  while renouncing the subsequent revision opportunity.

The lessons of the preceding example are reasonably general. Simultaneous concern for the ex ante and ex post perspectives in the ex ante position leads to choices between the ex ante and ex post fair alternatives. If those alternatives differ, and if the decision maker subscribes to consequentialism, she remains time inconsistent and prefers a commitment over flexibility.

### 2.2.5 Nominal fairness

Another non-consequentialist theory worth considering involves a concern for what we call *nominal fairness*: when making a decision, people ask themselves whether it would be fair if it were implemented. The component of their utility that involves fairness then depends on that hypothesized result, rather than the result they might rationally anticipate. Specific functional forms could, for example, involve probability weighting or non-separable utility, as above.

Why might an otherwise sophisticated individual exhibit nominally fair preferences? Some may feel that, when evaluating the *ethics* of a decision, one should proceed on the assumption that nothing will happen subsequently to overturn it. Under this view, the ends cannot ethically justify the means: one cannot rationalize an act that would, on its face, produce an unethical result by arguing that it will set in motion a chain of events that will likely achieve an ethical one instead. Rather, one must “own” each decision, and take responsibility for the consequences it most directly implies. If a choice is nominally unfair, one experiences disutility from making it, irrespective of the ultimate outcome.

Conceivably, nominal fairness could reflect a general tendency to narrowly frame evaluations of fairness. The intuition here is that, when evaluating the fairness of any decision, people do

not factor in all the world’s ills. Rather, they draw tight boundaries around the set of relevant considerations. For example, they may evaluate the fairness of a business deal only in terms of how it treats the parties to the deal, ignoring their endowments and other opportunities. Nominal fairness involves narrow framing in the temporal dimension.

A sophisticated nominally fair individual who wishes to behave ethically in all contexts, present and future, will exhibit the pattern we have associated with time inconsistency. That is, she will choose ex ante fair allocations ex ante, and ex post fair allocations ex post. Even so, her preferences are time-consistent. Choosing in advance, she will recognize that the ex post fair allocation is nominally fair in the ex post setting, and consequently she will select it. Similarly, she will affirmatively prefer flexibility to commitment so she can alter her choice to suit the ethical imperatives of the decision context as it evolves.

### 3 Design and implementation

Our experimental study involved the cooperation of well-established non-profit charity called GiveDirectly.<sup>11</sup> The organization operates a platform for donating money directly to needy households in poor African nations. We selected the households viewed by our subjects from lists of GiveDirectly’s potential recipients, and used photos they supplied.

#### 3.1 The basic allocation tasks

Each allocation task involves two impoverished Kenyan families, designated Household A and Household B. We study two types of tasks, one involving the allocation of lottery tickets, the other involving the allocation of a cash prize.

The outcome of the split-the-tickets task is that one of the two families receives a \$10 donation through GiveDirectly. A total of 20 lottery tickets are allocated between the two families. The division of 10 tickets (numbered 11-20) is fixed in advance and varies from task to task; we call this the “computer’s” ticket allocation. The subject allocates the remaining 10 tickets (numbered 1-10). Once all tickets are allocated, one is selected at random, and the family “holding” that ticket receives the \$10.

We examine multiple variants of this task, which differ according to what the subject knows and when she makes her decision. In all cases, the subject learns the computer’s allocation before dividing her own tickets.

**Ex ante decisions.** The subject makes her decision immediately after learning the computer’s allocation, without receiving any other information. Her ticket allocation is displayed on the

---

<sup>11</sup>See <http://www.givedirectly.org/>. GiveDirectly is recognized as one of the most efficient charities serving this sector. It was co-founded by a UCSD faculty member, a fact which may have enhanced its credibility with our UCSD undergraduate subjects.

screen, and she is asked to confirm or revise it. She repeats this step until she confirms her choice. A ticket is selected at random and the winner is determined.

**Ex post decisions.** The subject is told that the winning lottery ticket has been drawn. She learns whether it is one of the computer’s tickets or one of hers. In the latter case, she then allocates her own tickets without knowing which one is the winner. Her ticket allocation is displayed on the screen, and she is asked to confirm or revise it. She repeats this step until she confirms her choice. A ticket is selected at random and the winner is determined.

**Ex ante decisions with surprise ex post revisions.** The subject allocates her tickets immediately after learning the computer’s allocation, without receiving any other information. As in an ex ante decision, she is asked to confirm or revise it. Once she confirms it, she learns whether the winning ticket was one of the computer’s tickets or one of hers. In the latter case, she is not told the number of the winning ticket. Her ticket allocation is then displayed on the screen again and, as in an ex post decision, she is asked to confirm or revise it. (She is not informed in advance that she will have this opportunity to revise her choice after learning whether the winning ticket is one of hers). She repeats this step until she confirms her choice. A ticket is then selected at random and the winner is determined.

**Ex ante decisions with commitment.** The subject allocates her tickets immediately after learning the computer’s allocation, without receiving any other information. After confirming her choice (as in an ex ante task), she is informed that she will have an opportunity to reallocate her tickets after learning whether the winner is one of hers (but before learning who holds the winning ticket), unless she wishes to forego that opportunity. At that point, she must either express a preference for flexibility (“I definitely want the opportunity to revise”) or commitment (“I definitely do not want the opportunity to revise”), or express indifference (“I do not care about having an opportunity to revise”). If a subject expresses a preference for flexibility, she learns whether the winning ticket was one of the computer’s tickets or one of hers. In the later case, she is not told the number of the winning ticket, but is given an opportunity to reallocate her tickets (as in an ex post task). If a subject expresses a preference for commitment, she makes no other decisions. If a subject expresses indifference, we implement a 50-50 randomization between these two alternatives. A ticket is then selected at random and the winner is determined. Subjects are informed of all these rules in advance.

Some subjects also perform variants of these tasks designed to assess whether their preferences are strict.

**Decisions with incentivized revisions.** After setting some allocation, subjects are presented with unanticipated opportunities to switch to various alternatives. These alternatives always

involve giving all their tickets to the household they treated less favorably and none to the one they treated more favorably while increasing the prize. For example, if the subject chose to give 8 tickets to Household A and 2 tickets to Household B,<sup>12</sup> the alternative would be to give all 10 tickets to Household B while increasing the prize from \$10 to  $\$(10 + x)$ , where  $x \in \{0.10, 0.50, 1, 2, 5\}$ . Accordingly, in these cases, opportunities to switch are *incentivized*. Subjects make decisions for all five values of  $x$ ; we then select one of these choices at random and implement it.

The task presentation is structured to ensure that subjects view the two Kenyan households as equally deserving. At the outset of each task, subjects view pictures of 16 potential recipients, and are told that their Household A and Household B belong to that group.<sup>13</sup> The composition of families within each group is uniform; for instance, they might all be single women in one task, or couples with one child in another. To discourage subjects from searching for, and inflating the significance of, minor differences between families, we do not indicate which household within a group is A, and which is B.

The split-the-prize task is structured similarly. The final division of a \$10 prize between households A and B is governed by one of two allocations. The first of these is fixed in advance and varies from task to task; we call this the “computer’s” dollar allocation. The subject chooses the alternative allocation. One of these two allocations is selected at random and implemented; each is equally likely.

We examine both *ex ante* and *ex post* versions of the split-the-prize task. The subject learns the computer’s allocation at the outset of both versions. In the *ex ante* version, she chooses her allocation immediately thereafter. In the *ex post* version, she makes that choice only if she first learns that her allocation will be implemented. Details are otherwise the same as for the split-the-tickets task.

## 3.2 Treatments

In our main treatments, each subject participates in eight separate allocation tasks. These are divided into four sets of two tasks each, with sets separated by one-minute breaks. Subjects understand that they will perform at most one task involving any given household, and they view 16 new potential recipients in every round. They are also told in advance that only one of the eight resulting allocations, chosen at random at the end of the experiment, will be implemented.

We assign subjects in our main treatments to six distinct treatments, five of which involve split-the-tickets tasks, and one of which involves split-the-prize tasks. Table 2 summarizes the structure

---

<sup>12</sup>If the subject divided the tickets equally, the alternative would be to give all the tickets to a randomly selected household.

<sup>13</sup>We obtained the photographs from GiveDirectly; they are the actual recipients.

**Table 2:** Treatments

Treatment	Rounds 1&2	Rounds 3&4	Rounds 5-8	Number of subjects
4A_4A <sup>R</sup>	Ex-ante	Ex-ante	Ex-ante w/surprise revision	71
4P_4A <sup>R</sup>	Ex-post	Ex-post	Ex-ante w/surprise revision	72
2A2P_4A <sup>R</sup>	Ex-ante	Ex-post	Ex-ante w/surprise revision	48
2P2A_4A <sup>R</sup>	Ex-post	Ex-ante	Ex-ante w/surprise revision	48
4A <sup>R</sup> _4AC	Ex-ante w/surprise revision	Ex-ante w/surprise revision	Ex-ante w/commitment and anticipated revision	72

of the five split-the-tickets treatments, and indicates the number of subjects who participated in each. The first column lists treatment labels, which describes each treatment’s composition using a simple shorthand notation: “A” denotes an ex ante task, “E” denotes an ex post task, “A<sup>R</sup>” denotes an ex ante task with surprise ex post revision, and “A<sup>C</sup>” denotes an ex ante task with commitment. Consider, for example, the treatment labeled 2A2P\_4A<sup>R</sup>. “2A” means that the treatment starts with two ex ante tasks. “2P” means that rounds 3 and 4 are ex post tasks. The underscore separates the first four rounds from the second four. “4A<sup>R</sup>” means that the last four rounds are ex ante tasks with surprise ex post revision opportunities, which we present after the subject makes initial allocations in all four rounds.

The first four treatments have a common structure. In the first four rounds, subjects perform either ex ante tasks, ex post tasks, or a mixture of the two, while the last four rounds consist of ex ante decisions with surprise ex post revisions. These treatments are structured so that we can answer a variety of questions, as detailed in the next sections. For example, by comparing the first four rounds of 4A\_4A<sup>R</sup> and 4P\_4A<sup>R</sup>, we can determine whether the way in which a subject initially looks at an allocation problem (ex ante or ex post) affects her perspective on fairness. By tracking choices across those rounds, we can also determine whether perspectives drift or crystalize with further considerations. (Indeed, for treatment 4A\_4A<sup>R</sup>, we can track this for eight rounds.) By comparing ex ante choices in the last four rounds across these treatments, we can determine whether the perspective on fairness to which the subject is initially exposed (ex ante or ex post)



**Table 3:** Fixed allocation of computer’s tickets, by round

Round	1	2	3	4	5	6	7	8
Household A	7	2	10	1	8	3	9	0
Household B	3	8	0	9	2	7	1	10

affects subsequent ex ante decisions. By comparing choices in the third and fourth rounds of  $4P\_4A^R$  and  $2A2P\_4A^R$ , we can assess whether the initial perspective affects subsequent ex post decisions. By examining ex post revisions in the last four rounds of any of these treatments, we can determine whether subjects are time-inconsistent, and by comparing these patterns across treatments, we can determine whether exposure to alternative perspectives reduces the degree of inconsistency.

The fifth main treatment has a different structure: it involves four  $A^R$  tasks followed by four  $A^C$  tasks. The purpose of this treatment, unlike the first four, is to evaluate preferences for commitments, which are on display in rounds 5 through 8. However, we need to ensure that this demand is informed by appropriate experience with the pertinent decision tasks. Accordingly, the first four rounds are structured so that the subject is exposed to both the ex ante and ex post perspective, sees any conflict, and appreciates any tendency to reverse choices.

Ex ante fairness would require the subject to neutralize the computer’s distribution of lottery tickets. Accordingly, to enhance our ability to discern motives, we vary that distribution by round, as shown in Table 3. Subjects do not see this table in advance; rather, they learn the computer’s allocation at the start of each round.

The sixth main treatment involves split-the-prize tasks. The treatment has the same structure as  $4A\_4A^R$  – that is, four ex ante tasks followed by four more ex ante tasks with surprise revision. The pattern of variation across rounds for the fixed prize is also governed by Table 3, with the qualification that the numbers refer to dollars rather than tickets.

To assess whether both initial and final preferences are strict, we added two supplemental treatments involving decisions with incentivized revisions, each with four rounds. Both start with four ex ante split-the-tickets tasks. The purpose of the first supplemental treatment ( $4A^S$ ) is to evaluate whether preferences over initial allocations are strict. We accomplish this objective by giving subjects incentivized opportunities to switch after making their initial allocations. The purpose of the second supplemental treatment ( $4A^{RS}$ ) is to determine whether preferences over revised allocations are strict. Subjects are given opportunities to revise their allocations after learning whether the winning ticket is among the ones they allocated, followed by incentivized opportunities to switch their revised allocations. None of the opportunities to change allocations are anticipated.

### 3.3 Details concerning implementation

We conducted the experiment at the University of California, San Diego Economics Laboratory in November 2013, March 2014, and March 2015, within the guidelines of an IRB-approved human subjects protocol. Participants' instructions appear in Appendix B. Subjects viewed these instructions on computer screens and were directed to follow along as the study leader read them aloud. Participants made all responses using a computer interface programmed into Qualtrics survey software. Screen shots appear in Appendix C. Partitions were used to minimize the chance that one's allocations could be viewed by fellow subjects or by the experimenter. At the end of the experiment, subjects completed a short questionnaire in lieu of individual debriefing. A total of 311 subjects participated in the five main split-the-ticket treatments, distributed across the sessions as shown in Table 2, and 72 participated in the split-the-prize treatment. A total of 142 subjects participated in the supplemental treatments, divided roughly equally between them (72 in  $4A^S$  and 70 in  $4A^{RS}$ ). After completing all survey tasks, subjects filled out a short survey on demographics, including questions designed to elicit political inclinations. Surprisingly, we did not find any robust relationships between behavior and political views, but it is worth noting that our sample includes relatively few subjects who self-identified as strongly conservative. Each subject received \$15 for participating. Typically, the experiment lasted 45 minutes.

## 4 Does the framing of initial decisions affect the initial perspective on fairness?

Some of the subjects in our treatments initially encounter allocation tasks with ex ante framing, while others initially encounter them with ex post framing. In this section we investigate whether the initial decision frame colors the perspective on fairness subjects adopt before they encounter tasks with the alternative framing. Specifically, we test the hypothesis that those initially making choices with ex ante framing tend to see fairness in those tasks from an ex ante perspective, while those initially making choices with ex post framing tend to see fairness in those tasks from an ex post perspective.

For the purpose of this initial inquiry, we focus on the first four rounds of treatments  $4A\_4A^R$ , in which subjects start off by performing four allocation tasks with ex ante framing, and  $4P\_4A^R$ , in which subjects start off by performing four tasks with ex post framing.

Figure 1 shows the distributions of subjects' choices across five categories. Panels A and B pertain to subjects performing tasks with ex ante and ex post framing, respectively, during the first four rounds. The five categories are as follows.

**Excessive offsetting.** The subject allocates more tickets to the potential recipient who receives fewer from the computer, overcompensating for the disparity. For example, if the com-

puter allocates 3 tickets to A and 7 to B, the subject allocates 9 to A and 1 to B.

**Full offsetting (ex ante fair).** The subject allocates tickets so that each potential recipient ends up with the same number in total. For example, if the computer allocates 3 tickets to recipient A and 7 to B, the subject allocates 7 to A and 3 to B.

**Partial offsetting.** The subject allocates more tickets to the potential recipient who receives fewer from the computer, but does not completely compensate for the disparity. For example, if the computer allocates 3 tickets to A and 7 to B, the subject allocates 6 to A and 4 to B.

**No offsetting (ex post fair).** The subject allocates five tickets to each potential recipient.

**Reinforcing.** The subject allocates more tickets to the potential recipient who receives more from the computer. For example, if the computer allocates 3 tickets to A and 7 to B, the subject allocations none to A and 10 to B.

The differences between the distributions depicted in panels A and B of Figure 1 are striking. For panel A, which pertains to initial tasks with ex ante framing, most choices are ex ante fair (that is, fully offsetting). In contrast, for panel B, which pertains to initial tasks with ex post framing, the modal choice is ex post fair (that is, it involves no offsetting). Indeed, moving from panel A to panel B, the primary change is that the frequency of ex ante fair choices declines by 34 percentage points, while the frequency of ex post fair choices rises by 35 percentage points. Notably, ex ante fairness remains reasonably common in the ex post frame (consistent with findings in Cappelen et al. (2013)), while ex post fairness is relatively rare in the ex ante frame.

Standard tests for the equality of distributions, such as Pearson’s  $\chi^2$  test, are inapplicable here because they do not account for within-subject correlation across the four rounds. More specifically, any test that treats multiple observations of choices by the same subject as independent will tend to exaggerate the statistical significance of the differences across treatments. A resolution of this issue requires assumptions about the structure of the underlying statistical process. Accordingly, we pool the data from the two treatments, estimate a multinomial logit model with category-specific constants and category-treatment interactions, and perform a  $\chi^2$  test of the hypothesis that all the coefficients for the interaction terms are zero, clustering standard errors at the subject level. For the distributions depicted in figure 1, we reject equality decisively ( $p < 0.001$ ).

For Figure 1, we pooled observations over the first four rounds. It is natural to wonder whether the differences between initial decisions made with ex ante and ex post framing are stable, or whether they dissipate as subjects have more time to think through their attitudes toward these types of decision tasks. We investigate this possibility in the Appendix,<sup>14</sup> and find no evidence of

---

<sup>14</sup>All supplemental data analyses referenced in the text appear in Appendix Section A.3.

**Figure 1:** Distributions of initial choices conditional on initial framing



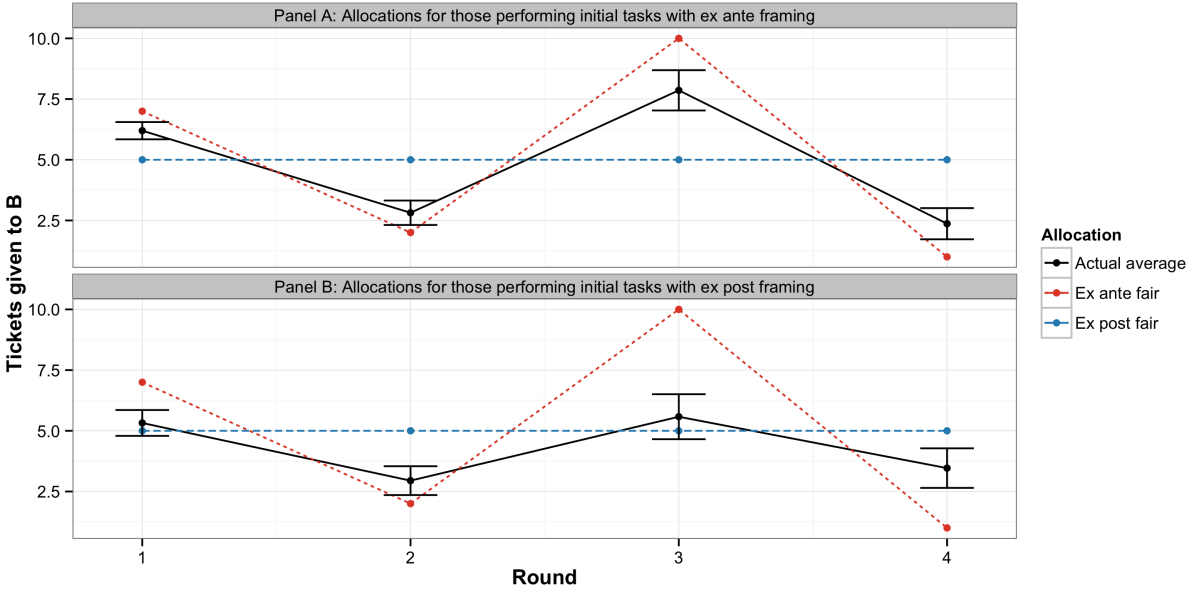
Notes: Panel A is based on the first four rounds of treatment 4A\_4A<sup>R</sup> (284 observations). Panel B is based on the first four rounds of treatment 4P\_4A<sup>R</sup> (144 observations). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

systematic changes in behavior over the course of the first four rounds of treatments 4A\_4A<sup>R</sup> and 4P\_4A<sup>R</sup>.

A closer look at the data reveals that some subjects make the same type of choice in every round, while others move around between categories. Consistency across rounds could be an indication of the seriousness and deliberateness with which subjects approached the tasks and acted on coherent decision principles. Accordingly, it is important to determine whether the differences between the distributions shown in panels A and B of Figure 1 are associated with consistent or inconsistent subjects.

In the first four rounds of 4A\_4A<sup>R</sup>, 37% of the subjects (26 of 71) made the same type of choice in every round. In every case, the choices were ex ante fair. Although we have not yet focused on rounds 5 through 8, it is notable that the degree of stability increased, perhaps because subjects arrived at coherent principles with experience. Specifically, 61% of the subjects (43 of 71) made the same type of choice in each of the last four rounds, and in all but one of those cases (42 of 43), the choices were ex ante fair. Turning next to the first four rounds of 4P-4A<sup>R</sup>, it is important to bear in mind that each subject made two decisions rather than four. Overall, 49% of subjects (35 of 72) made the same type of choice in both of those rounds. Of those, 60% (21) chose the ex post fair option, which is considerably higher than the overall frequency for this treatment (shown in

**Figure 2:** Allocations for those performing initial tasks in a single frame



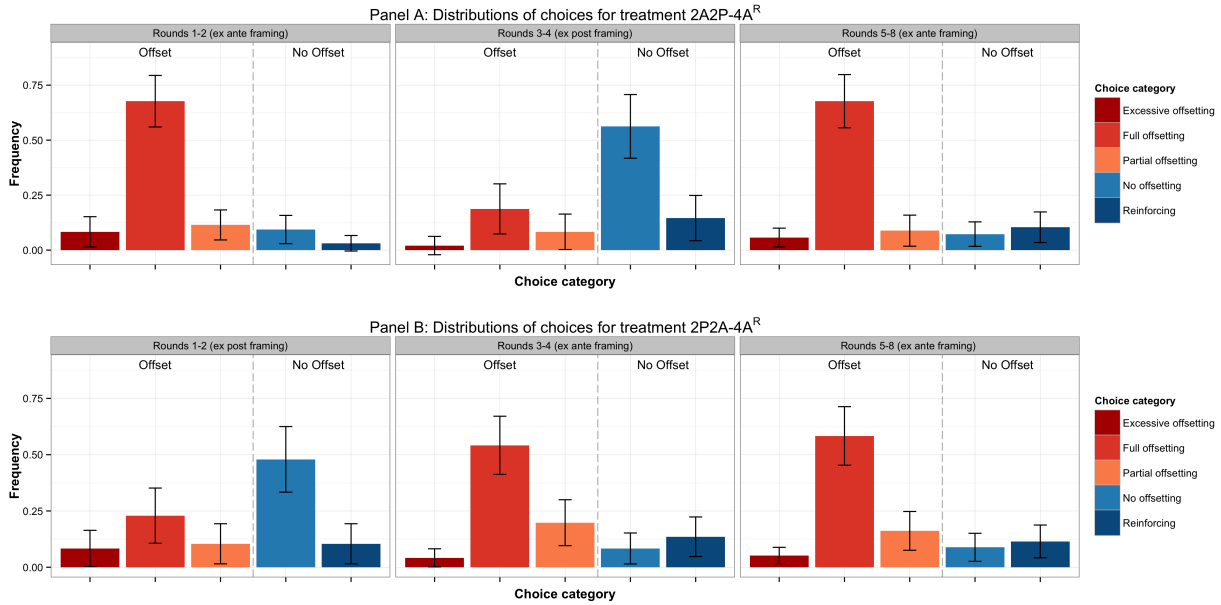
Notes: Panel A is based on the first four rounds of treatment 4A\_4A<sup>R</sup> (71 observations per round). Panel B is based on the first four rounds of treatment 4P\_4A<sup>R</sup> (36 observations per round). Error bands indicate 95% confidence intervals.

panel B of Figure 1), and only 24% (17) chose the ex ante fair option, which is noticeably lower than the overall frequency. Accordingly, we conclude that the differences between the distributions exhibited in Figure 1 are primarily attributable to consistent choosers.

An important feature of our experimental design is that the allocation of the computer’s tickets varies from one round to the next. Accordingly, the choices of an ex ante fair subject should vary in a recognizable and distinctive manner across rounds, while the choices of an ex post fair subject should remain fixed. We exhibit these patterns in the two panels of Figure 2, which plot the number of the subject’s tickets given to recipient B, by round. The red and green lines correspond, respectively, to the “fingerprints” of an ex ante fair subject, and of an ex post fair subject. Panel A superimposes a blue line representing the average choices made with ex ante framing in the first four rounds of treatment 4A\_4A<sup>R</sup>; panel B does the same for choices made with ex post framing in the first four rounds of treatment 4P\_4A<sup>R</sup>.

Notice that the actual choices resemble the ex ante fingerprint much more closely when the initial tasks involve ex ante rather than ex post framing. In the latter case, the blue line is much flatter. To quantify this difference, we estimated simple regressions of the chosen split on a constant and the computer’s split, separately for the two treatments, clustering observations at the subject level. For an ex ante fair subject, the coefficient of the computer’s split would be -1; for an ex post

**Figure 3:** Distributions of choices in treatments with changing decision frames



Notes: Panel A is based on rounds 1-8 of treatment 2A2P-4A<sup>R</sup> (72 subjects). Panel B is based on rounds 1-8 of treatment 4P-4A<sup>R</sup> (72 subjects). Results for rounds 5-8 reflect original choices, not revisions. Error bands indicate 95% confidence intervals.

fair subject, it would be 0. In fact, we find that it is -0.62 (s.e. = 0.06) for choices made with ex ante framing, and -0.29 (s.e. = 0.05) for choices made with ex post framing. We decisively reject the hypothesis that these coefficients are the same ( $p < 0.001$ ).

## 5 Does the framing of initial decisions establish a persistent perspective on fairness?

So far, we have seen that the framing of initial decisions strongly influences the initial perspective on fairness. That finding does not necessarily imply time inconsistency. After all, our experiment involves decision tasks that few if any subjects have previously encountered. Perhaps someone who initially performs such a task with one type of framing (either ex ante or ex post) thinks the fairness issues through from that perspective, and then adheres to the resulting decision principles through subsequent tasks, even if the framing changes. In that case, choices would be time-consistent. A precondition for time inconsistency is that, subject by subject, choices change as the decision frame changes.

To investigate these issues, we will focus on treatments with changing decision frames, beginning with 2A2P-4A<sup>R</sup>, in which subjects performed two tasks with ex ante framing, then two with ex post

framing, then four with ex ante framing (followed by surprise revisions), as well as 2P2A\_4A<sup>R</sup>, in which subjects performed two tasks with ex post framing, then six with ex ante framing (followed by surprise revisions for the last four). Throughout this section, when examining rounds 5-8, we will focus on the original choices, leaving the analysis of revisions to later sections.

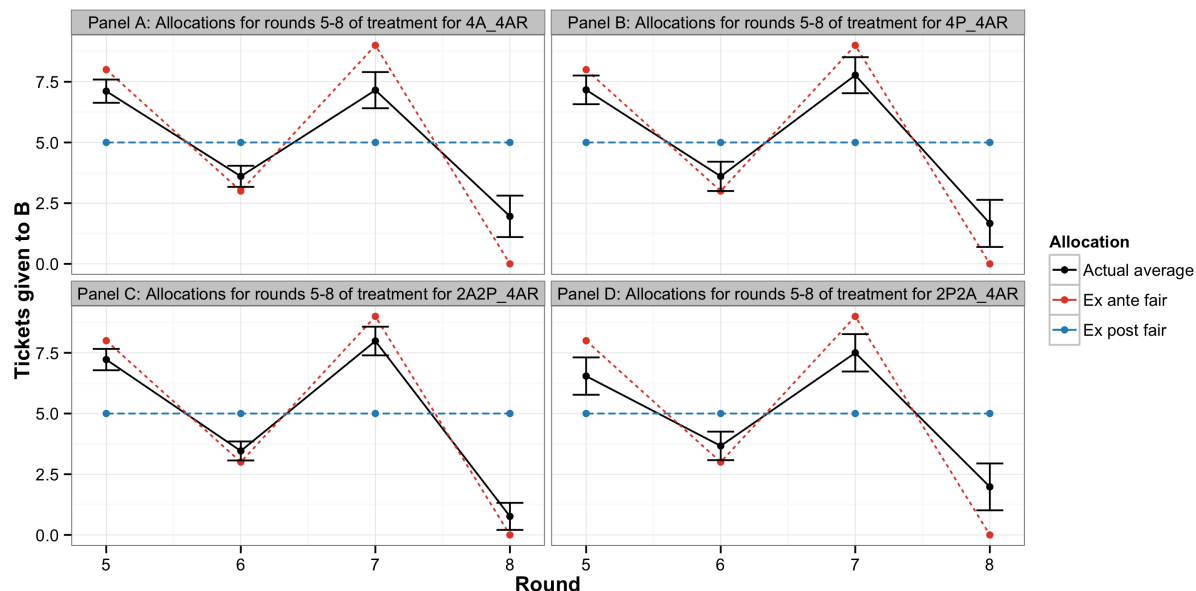
Figure 3 displays distributions of choices over the same five categories as Figure 1, except that here we report results separately for rounds 1-2, 3-4, and 5-8. Panel A pertains to treatment 2A2P\_4A<sup>R</sup>, while panel B pertains to treatment 2P2A\_4A<sup>R</sup>.

Looking at this figure, one sees a striking similarity between four of the distributions – the first and third in panel A, and the second and third in panel B. Indeed, all of these strongly resemble the distribution in panel A of Figure 1, in that ex ante fair choices are predominant. Significantly, all four involve tasks with ex ante framing. One also sees a striking similarity between the second distribution in panel A and the first in panel B. Both strongly resemble the distribution in panel B of Figure 1, in that ex post fairness is the most common decision mode. Significantly, both involve tasks with ex post framing. Thus, the main lesson from this figure is that framing effects exhibit little if any persistence. Choices depend on the framing of the current task, but not to any significant degree on the framing of initial or previous tasks. As a result, subjects shift their perspectives on fairness back and forth along with the decision frame.

Formal statistical tests confirm these visual impressions. First, the samples are comparable to those used for figure 1: we do not reject equality of the round 1-2 distributions of treatments 2A2P\_4A<sup>R</sup> and 4A\_4A<sup>R</sup> ( $p = 0.21$ ); likewise, we do not reject equality of the round 1-2 distributions of treatments 2P2A\_4A<sup>R</sup> and 4P\_4A<sup>R</sup> ( $p = 0.64$ ). Second, subjects do not simply adopt an initial perspective and adhere to it in all subsequent rounds, even when the decision frame changes: we reject the equality of the round 1-2 and round 3-4 distributions of treatment 2A2P\_4A<sup>R</sup> ( $p < 0.001$ ), and similarly for treatment 2P2A\_4A<sup>R</sup> ( $p < 0.001$ ). Third, initial exposure to the ex ante perspective does not systematically affect the subsequent proclivity to adopt the ex post perspective when the task involves ex post framing: we do not reject equality of the round 3-4 distributions for 2A2P\_4A<sup>R</sup> and 4P\_4A<sup>R</sup> ( $p = 0.38$ ). Fourth, initial exposure to the ex post perspective does not systematically affect the subsequent proclivity to adopt the ex ante perspective when the task involves ex ante framing: we do not reject equality of the round 3-4 distributions for 2P2A\_4A<sup>R</sup> and 4A\_4A<sup>R</sup> ( $p = 0.92$ ), nor do we reject equality of the round 5-8 distributions ( $p = 0.42$ ). Finally, moving back and forth between multiple perspectives does not systematically affect the subsequent proclivity to adopt the ex ante perspective when the task involves ex ante framing: we do not reject equality of the round 5-8 distributions for 2A2P\_4A<sup>R</sup> and 4A\_4A<sup>R</sup> ( $p = 0.79$ ).

A potential concern regarding the preceding results is that exposure to ex post framing is relatively brief in the treatments considered – just two rounds. Perhaps longer exposure would have a more lasting effect on the perspective adopted subsequently once the framing changes. We

**Figure 4:** Allocations for those performing tasks in the ex ante frame after varying degrees of exposure to the ex post frame



Notes: Panel A is based on 71 subjects, panel B on 72 subjects, panel C on 48 subjects, and panel D on 48 subjects. Error bands indicate 95% confidence intervals.

evaluate that possibility in the Appendix by examining choices made in the  $4P_4A^R$  treatment. We find no meaningful differences in behavior in ex ante tasks even after subjects have performed four consecutive ex post tasks to start the experiment.

Significantly, the tendency for people to make ex ante fair choices even after being exposed to the ex post perspective is even more evident when one restricts attention to consistent choosers. See the Appendix for detailed analysis.

The absence of a persistent perspective on fairness that survives changes in the decision frame is also evident from comparisons between the pattern of average allocations across rounds and the “fingerprints” associated with ex ante and ex post fairness. The various panels of Figure 4 display these fingerprints, along with average allocations in each of the last four rounds of the following treatments:  $4A_4A^R$  (panel A),  $4P_4A^R$  (panel B),  $2A2P_4A^R$  (panel C), and  $2P2A_4A^R$  (panel D). In every instance, actual choices resemble the ex ante fair fingerprint much more closely than the ex post fair fingerprint. As in section 4, we quantify this similarity by estimating simple regressions of the chosen split on a constant and the computer’s split, clustering observations at the subject level. The coefficient of the computer’s split is  $-0.61$  (s.e. =  $0.07$ ) for treatment  $4A_4A^R$ ,  $-0.79$  (s.e. =  $0.05$ ) for treatment  $4P_4A^R$ ,  $-0.68$  (s.e. =  $0.07$ ) for treatment  $2A2P_4A^R$ , and  $-0.59$  (s.e. =



0.09) for treatment 2P2A\_4A<sup>R</sup>. We do not reject equality of these coefficients ( $p = 0.32$ ), and there is certainly no indication that previous exposure to the ex post perspective pushes the coefficient away from -1 (the ex ante fair benchmark) and toward 0 (the ex post fair benchmark).

## 6 Does concern for fairness generate time-inconsistency?

The last two sections document a pronounced tendency for subjects to adopt an ex ante perspective on fairness when making decisions with ex ante framing, and an ex post perspective when making decisions with ex post framing. Those findings point to a potential source of time inconsistency, but do not actually establish that subjects are time-inconsistent. It is one thing to invoke different decision criteria in completely separate tasks, and potentially quite another to revise the choice made in a given task after arriving at a set of applicable principles for that task. Conceivably, people could apply their principles resolutely within each task while failing to do so across different tasks.

To determine whether subjects are in fact time-inconsistent, we examine the choices they make when they are unexpectedly allowed to revise decisions ex post after allocating tickets ex ante. In this section, we focus on revisions made for rounds 5-8 of treatment 4A\_4A<sup>R</sup>. The subjects in this treatment only encounter tasks with ex ante framing prior to learning that they can revise the last four choices ex post.

Revisions were the rule rather than the exception. Subjects revised 68.3% of the original round 5-8 choices, and 78.9% of subjects revised at least one choice. The revision frequency started out at 75.6% in round 5, dropped to 64.7% in round 6, and then rebounded a bit in rounds 7 and 8 (65.8% and 67.6%). Overall, there is no indication that the tendency to revise dissipates once subjects become aware of their behavior.

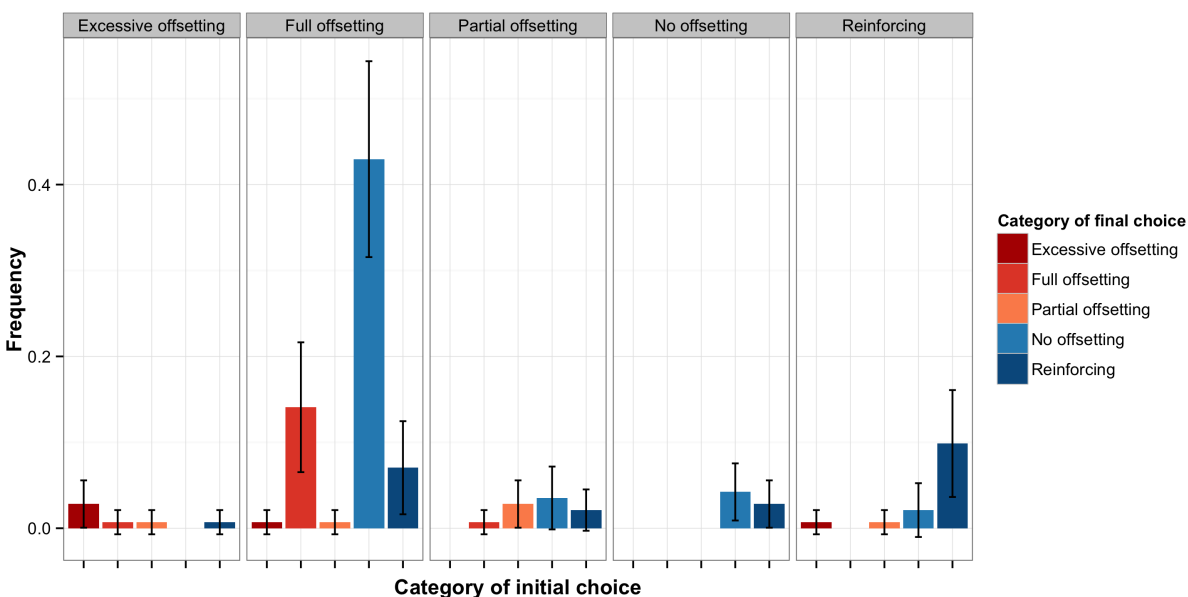
Consistent with the notion that the ex post perspective on fairness becomes compelling once the ex post position is reached, switches to 50-50 were by far the most common type of revision (71%). We provide additional details concerning the distribution of revision types in the Appendix.

Figure 5 depicts the joint distribution of original and final (revised) choices for rounds 5-8, pooling across rounds. The figure shows five groups of five bars. There is one group for each possible type of the original choice, as indicated along the top of the figure. Within each group, there is one bar for each possible type of the final choice, as indicated by the legend. Types of choices are displayed in the same order as in Figure 1, both within and across groups. Frequencies are expressed as percentages of the total number of round 5-8 original-final choice pairs, so it is easier to see which patterns are most prevalent.

Our hypothesis is that many subjects will manifest time-inconsistent behavior by making choices that are ex ante fair before the resolution of uncertainty, and then revise to ex post fair choices after the (partial) resolution of uncertainty. This is the most common pattern by far, encompassing

43.0% of all allocation tasks. The next most common pattern is time consistent: in 14.1% of tasks, subjects exhibit resolute non-EU preferences, selecting and sticking with the ex ante fair alternative. The third most common pattern (9.9% of tasks) is a reinforcing allocation followed by another reinforcing allocation (possibly the same one). These subjects may have taken the computer's allocation as a signal of relative need. Interestingly, the fourth most common pattern (7% of tasks) is an ex ante fair allocation followed by a reinforcing allocation; possibly these subjects seek to compensate the computer-favored subject ex post for bad luck. All other patterns occur with frequencies below 5%.

**Figure 5:** Joint distribution of original and final (revised) choices during the final four rounds of treatment 4A\_4A<sup>R</sup>

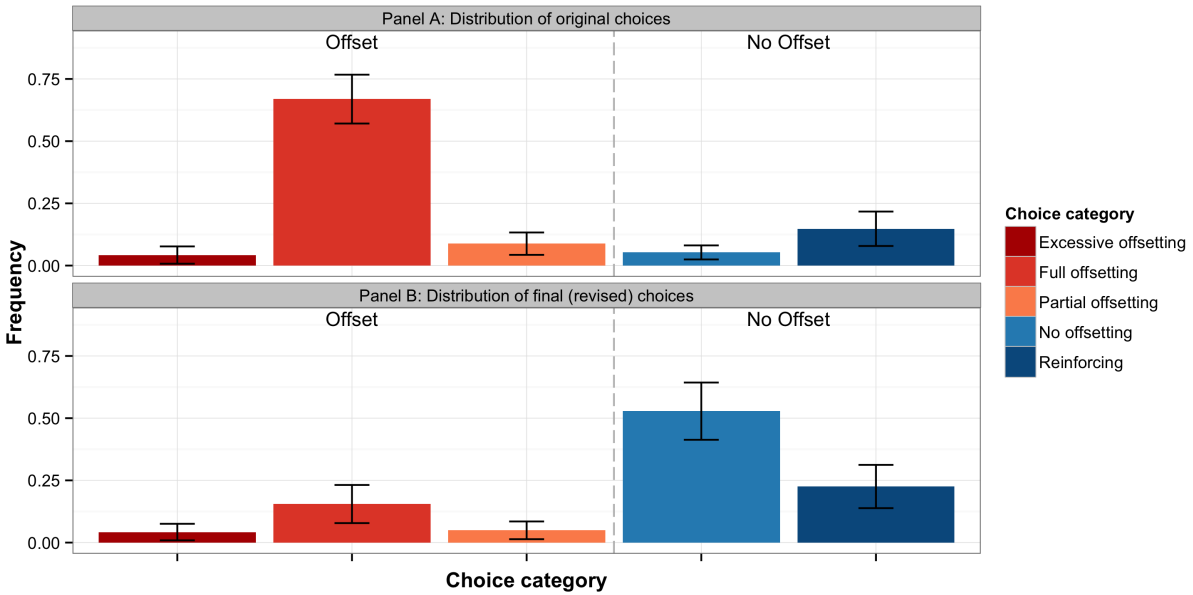


Notes: This figure is based on the final four rounds of treatment 4A\_4A<sup>R</sup> (284 observations). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

Figure 6 displays the marginal distributions for original and final choices (panels A and B, respectively) implied by the joint distribution depicted in Figure 5. A comparison of the two panels reveals the effect of unexpected revision opportunities on the distribution of allocations, which is difficult to discern from the previous figure. The overall distribution shifts dramatically from one in which ex ante fair choices predominate to one in which ex post fair choices predominate. Indeed, there is a striking resemblance between Figure 6 and Figure 1.

Formally, we reject the hypothesis that the distribution of final choices in rounds 5-8 of treatment 4A\_4A<sup>R</sup>, shown in panel B of Figure 6, is the same as the distribution of choices in rounds 1-4 of

**Figure 6:** Marginal distributions of original and final (revised) choices during the final four rounds of treatment  $4A\_4A^R$

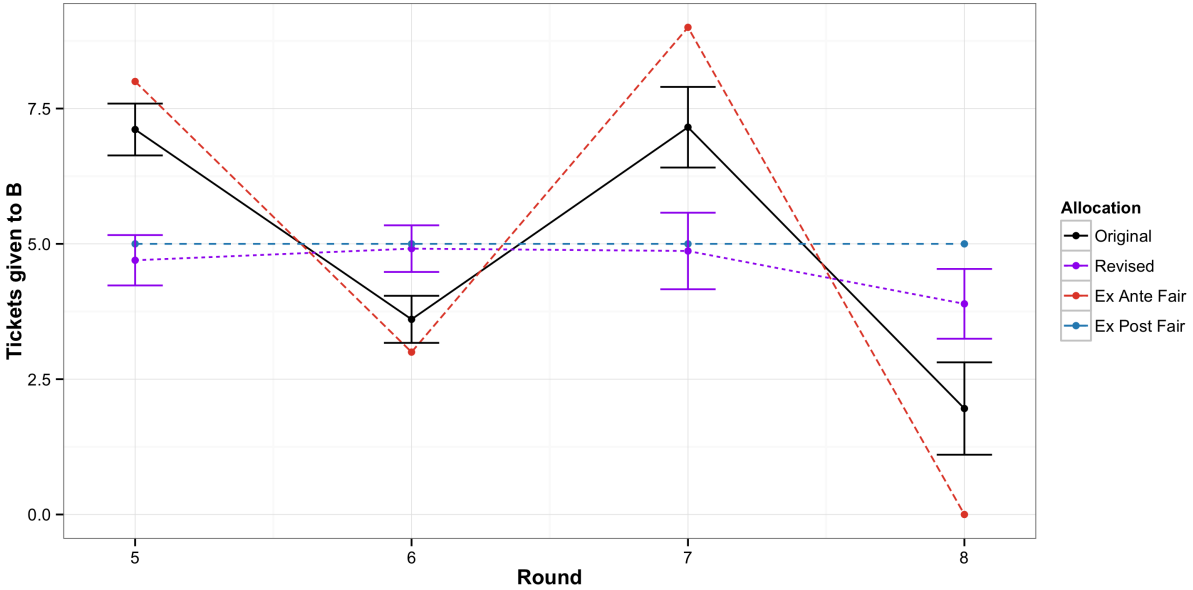


Notes: This figure is based on the final four rounds of treatment  $4A\_4A^R$  (284 observations). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

treatment  $4P\_4A^R$ , shown in panel B of figure 1 ( $p < 0.001$ ). However, this finding does not reflect a tendency to stick with the original choice made or perspective adopted when first contemplating a given allocation task. Notice in particular that the frequency of ex post fair choices is actually higher, not lower, among final choices in rounds 5-8 of treatment  $4A\_4A^R$  than in rounds 1-4 of treatment  $4P\_4A^R$ , and conversely for ex ante fair choices. Thus, the tendency to choose ex post fair allocations when making decisions ex post is undiminished when subjects revise ex ante decisions, compared to when they make ex post decisions in new allocation tasks.

The dramatic effect of revisions is also evident from comparisons between the pattern of average allocations across rounds (both before and after revisions) and the “fingerprints” associated with ex ante and ex post fairness. Figure 7 replicates panel A of Figure 4, except that we have added a line for the revised choices. The average revised choices closely resemble the benchmark for ex post fairness in rounds 5-7, and are nearly insensitive to the computer’s initial distribution. In round 8, the final choice moves a bit in the direction of the ex ante fair allocation, but to a much smaller extent than the original (unrevised) choice. As in earlier sections, we quantify the similarity to the benchmarks by estimating simple regressions of the chosen split on a constant and the computer’s split, clustering observations at the subject level. The coefficient of the computer’s split,  $-0.08$

**Figure 7:** Original and final allocations in rounds 5-8 of treatment 4A\_4A<sup>R</sup>



Notes: This figure is based on the final four rounds of treatment 4A\_4A<sup>R</sup> (71 subjects). Error bands indicate 95% confidence intervals.

(s.e. = 0.07), is not significantly different from zero, again a reflection of the fact that the ex post perspective predominantly governs revisions.

In previous sections, we investigated whether the documented patterns are attributable to subjects who choose consistently, or to those whose categorical choices vary across rounds. It is particularly important to ask this question with respect to our findings concerning revisions, because consistent choosers may be devoted to particular perspectives, and consequently less likely to change their minds as a result of changes in framing.

To address this issue, we divided subjects from the 4A\_4A<sup>R</sup> treatment into two groups: consistent choosers (those whose original decision fell into the same category in at least 7 of the 8 rounds), and inconsistent choosers (all others). Notably, most of these subjects (52%) were consistent choosers. Several patterns merit emphasis. First, nearly all (96%) of the original choices made by consistent choosers in rounds 5-8 were ex ante fair. Second, the frequency of revisions was actually higher for consistent choosers (77% of their choices) than for inconsistent choosers (58% of their choices). Thus, consistency across rounds does not translate into consistency across decision frames. Third, for this group, roughly two-thirds of choice pairs (64%) involved an original ex ante fair allocation, followed by a revision to an ex post fair allocation. Thus, consistent choosers manifest the pattern of interest to an even greater extent than the general subject population. Interestingly, nearly a

quarter of choice pairs (23%) made by consistent choosers were time consistent: these subjects exhibited resolute non-EU preferences by making and sticking to ex ante fair allocations. Roughly one in ten choice pairs entailed revisions that compensated for bad luck, in that the subject switched from an ex ante fair allocation to a reinforcing one. In the remaining choice pair, the subject switched from an ex ante fair allocation to excessive offsetting. We conclude that time inconsistency is especially prevalent for the 52% of our subjects who are consistent choosers.

In Section 2.2, we observed that several preference specifications could in principle produce apparent time inconsistency through the fortuitous resolution of uncertainty. We address this possibility by examining our two supplemental treatments, in which subjects were unexpectedly given the opportunity to redistribute all of their tickets to their less-favored household in return for increases in the prize of various magnitudes. Focusing on treatment  $4A^S$ , we see that subjects were unwilling to switch from their initial allocations in 75% of tasks for a 10 cent prize bonus. Not surprisingly, that fraction declined as the bonus increased. However, subjects were still unwilling to switch in nearly two-thirds of the tasks (65%) in return for a \$1 bonus, and in nearly half of the tasks (47%) for a \$5 bonus. For those who exhibited the choice pattern of primary interest (ex ante fairness in ex ante tasks), the fractions unwilling to switch were essentially the same. Turning to treatment  $4A^{RS}$ , we see that subjects were unwilling to switch from their revised allocations in two-thirds (67%) of tasks for a 10 cent prize bonus, in 61% of tasks for a \$1 bonus, and in 40% of tasks for a \$5 prize bonus. Results were again similar for those who exhibited the choice patterns of primary interest (revisions to ex post fairness both overall and from initial ex ante fair allocations). Thus we conclude that subjects' distributional preferences were in most cases strict.

**Table 4:** Fraction unwilling to switch by decision type, group, and magnitude of incentive

Decision	Group	Prize bonus from switching (x)				
		\$0.10	\$0.50	\$1	\$2	\$5
Initial	All	0.75	0.72	0.65	0.61	0.47
	Ex ante fair	0.76	0.73	0.66	0.63	0.52
Revised	All	0.67	0.64	0.61	0.57	0.40
	Ex post fair	0.71	0.69	0.65	0.61	0.40
	Ex ante → ex post fair	0.63	0.65	0.63	0.61	0.43

Note: Results for initial decisions are based on treatment  $4A^S$  (72 subjects, 288 observations). Results for revised decisions are based on treatment  $4A^{RS}$  (70 subjects, 280 observations).

## 7 Does awareness of framing effects ameliorate time-inconsistency?

In principle, the time-inconsistent choices documented in the previous section could be the result of subjects not considering the ex post perspective until they find themselves with ex post opportunities to revise allocations. Conceivably, those who consider both the ex ante and ex post perspectives might reconcile the conflict internally and display greater time consistency as a result. The results from section 5 suggest not: subjects continue to adopt ex ante perspectives on fairness in tasks with ex ante framing, and ex post perspectives in separate tasks with ex post framing, even after exposure to both frames. However, that evidence stops short of demonstrating that subjects continue to reverse ex ante decisions when provided with opportunities to make ex post revisions.

To address this set of issues, we examine patterns of revisions in the three treatments that expose subjects to the ex post perspective in rounds 1-4:  $4P\_4A^R$ ,  $2A2P\_4A^R$ , and  $2P2A\_4A^R$ . We have already compared the original choices made in rounds 5-8 for these treatments, as well as  $4A\_4A^R$ , and have found that they are generally similar (see section 5). Here we focus on revisions.

The frequency of revisions is almost identical in treatments  $4A\_4A^R$ ,  $4P\_4A^R$ , and  $2A2P\_4A^R$  (68.3%, 69.4%, and 69.8%, respectively). It is noticeably lower in treatment  $2P2A\_4A^R$  (53.1%), and the difference from the corresponding frequency in treatment  $4A\_4A^R$  is nearly statistically significant at the 5% level ( $p = 0.057$ ). Because treatment  $2P2A\_4A^R$  is effectively an intermediate case between treatments  $4A\_4A^R$  and  $4P\_4A^R$  (two rounds of exposure to the ex post perspective in rounds 1-4 rather than zero and four, respectively), the latter finding is surprising. As reported below, results from that treatment are somewhat anomalous. Even so, they qualitatively corroborate our central findings.

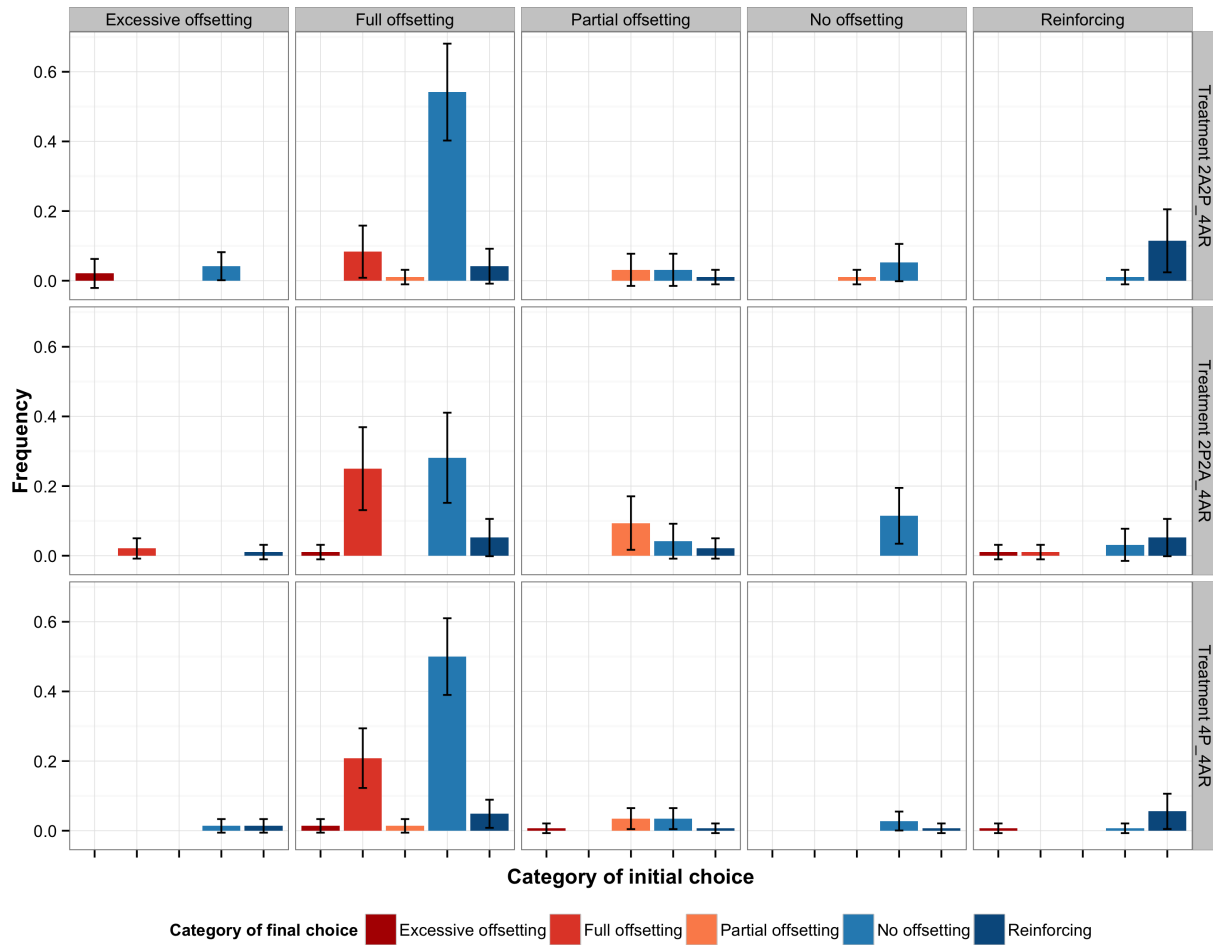
As in treatment  $4A\_4A^R$ , subjects who made revisions primarily switched to ex post fair allocations.<sup>15</sup> Moreover, differences in the distributions of revision types (whether the subject moved away from, toward, to, or past ex post fairness) between treatment  $4A\_4A^R$  on the one hand and treatments  $2A2P\_4A^R$ ,  $2P2A\_4A^R$ , and  $4P\_4A^R$  on the other were not statistically significant.

Figure 8 displays the joint distributions of original and final choices in rounds 5-8 of the same three treatments. The panels of these Figures are analogous to Figure 5, which pertains to treatment  $4A\_4A^R$ . We see that, in each case, the most common original-final choice pair is an ex ante fair allocation followed by one that is ex post fair. Pooling over the three treatments, the second most common choice pair is time consistent: the subject exhibits resolute non-EU preferences by selecting and sticking with an ex ante fair allocation. These are the same patterns we saw in our analysis of treatment  $4A\_4A^R$ . Significantly, we cannot reject the hypotheses that the distributions of final choices for any of these treatments,  $2A2P\_4A^R$ ,  $2P2A\_4A^R$ , or  $4P\_4A^R$ , are the same as for  $4A\_4A^R$  ( $p = 0.50, 0.15, \text{ and } 0.46$ , respectively).<sup>16</sup>

<sup>15</sup>For details concerning the distributions of revision types, see the Appendix.

<sup>16</sup>We provide a more detailed comparison of those distributions in the Appendix

**Figure 8:** Joint distribution of original and final (revised) choices during the final four rounds of various treatments



Notes: This figure is based on the final four rounds of the indicated treatments, in which there were 48, 48, and 72 subjects, respectively. Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

In Section 6, we compared the pattern of average allocations across rounds 5-8 of treatment  $4A\_4A^R$  both before and after revisions with the “fingerprints” associated with ex ante and ex post fairness (recall Figure 7). While the initial choices tracked the ex ante fair fingerprint reasonably well, the pattern for final choices more closely resembled the ex post fair fingerprint. In the Appendix, we perform the same analysis for treatments  $2A2P\_4A^R$ ,  $2P2A\_4A^R$ , and  $4P\_4A^R$ , and find similar results.

In summary, the predominance of the main pattern documented in the previous section – initial ex ante fair choices followed by ex post fair revisions – is undiminished when subjects are exposed to the ex post perspective prior to making any of the decisions that are subsequently subject to revision. Thus the conflict between ex ante and ex post perspectives gives rise to time inconsistency even when subjects are aware of it.

## 8 Does self-awareness of time inconsistency regarding fairness create a preference for commitment?

In Section 2, we saw that various formulations of fairness preferences give rise not only to time inconsistency, but also to a demand for commitment, provided decision makers are sophisticated. In most formulations, sophisticated time-inconsistent subjects who make allocation decisions ex ante strictly prefer to avoid revision opportunities, lest they subsequently subvert their own ex ante objectives. However, if ethical judgments are governed by the principle that each choice must stand on its own (what we have called nominal fairness), the decision maker may have no interest in precluding anticipated revisions, even ones that negate the force of earlier ethical decisions. In this section, we ask whether a demand for commitment arises in the current context.

To address these issues, we examine the choices made by subjects who participated in treatment  $4A^R\_4A^C$ . During the first four rounds, they have opportunities to experience decision making from both the ex ante and ex post frames, as well as to notice their own tendency to make revisions.<sup>17</sup> During the final four rounds, they start by making ex ante decisions, but are given options to forego subsequent revision opportunities; they are allowed to revisit each decision only if they decline the corresponding option.

In subsection 8.1, we evaluate the overall effects of these commitment opportunities. Many subjects apparently value commitment and use it to mitigate the tendency to shift from ex ante fair to ex post fair allocations. However, we show in subsection 8.2 that this perspective is not

---

<sup>17</sup>In rounds 1-4 of treatment  $4A^R\_4A^C$ , subjects generally exhibited the same patterns observed in rounds 5-8 of treatment  $4A\_4A^R$ , documented in sections 4 and 6. For instance, 60.4% of the original choices were ex ante fair, while only 10.4% were ex post fair. 54.9% of choices were revised ex post when subjects were given the opportunity. As a result, 54.9% of the final choices were ex post fair, while only 17.4% were ex ante fair. Generally, we do not reject hypotheses concerning the equivalence of behavior in rounds 1-4 of treatment  $4A^R\_4A^C$  and rounds 5-8 of treatment  $4A\_4A^R$ .



universal. As a group, those who are inclined to switch from ex ante to ex post fairness actually avoid making commitments to a greater extent than other subjects. Apparently, many of them prefer to have and to exercise the flexibility to switch. That preference is consistent with the theory of nominal fairness discussed in Section 2.

## 8.1 Overall effects of commitment opportunities

In this subsection, we address three questions. First, do subjects choose to forgo future flexibility when given the opportunity? Second, does the availability of these commitment opportunities reduce the frequency of revisions, and third, does it change the distribution of final choices?

Our first finding is that subjects take advantage of opportunities to make commitments with reasonably high frequency, but also retain flexibility with comparable frequency. Recall that we offer subjects three alternatives: express a definite preference for a commitment to the ex ante choice, express a definite preference for continued flexibility, and express indifference. Subjects expressed a strict preference for commitment 40.6% of the time, a strict preference for flexibility 30.2% of the time, and indifference 29.2% of the time. These frequencies do not vary systematically across rounds.<sup>18</sup>

The fact that many subjects chose to make commitments does not necessarily mean that the availability of commitment opportunities changed outcomes. Because commitments are costless in our experiment, subjects with time-consistent preferences – that is, those who would have resolutely stuck with their original choices in any event – may use commitments simply to emphasize strong preferences for their chosen allocations.

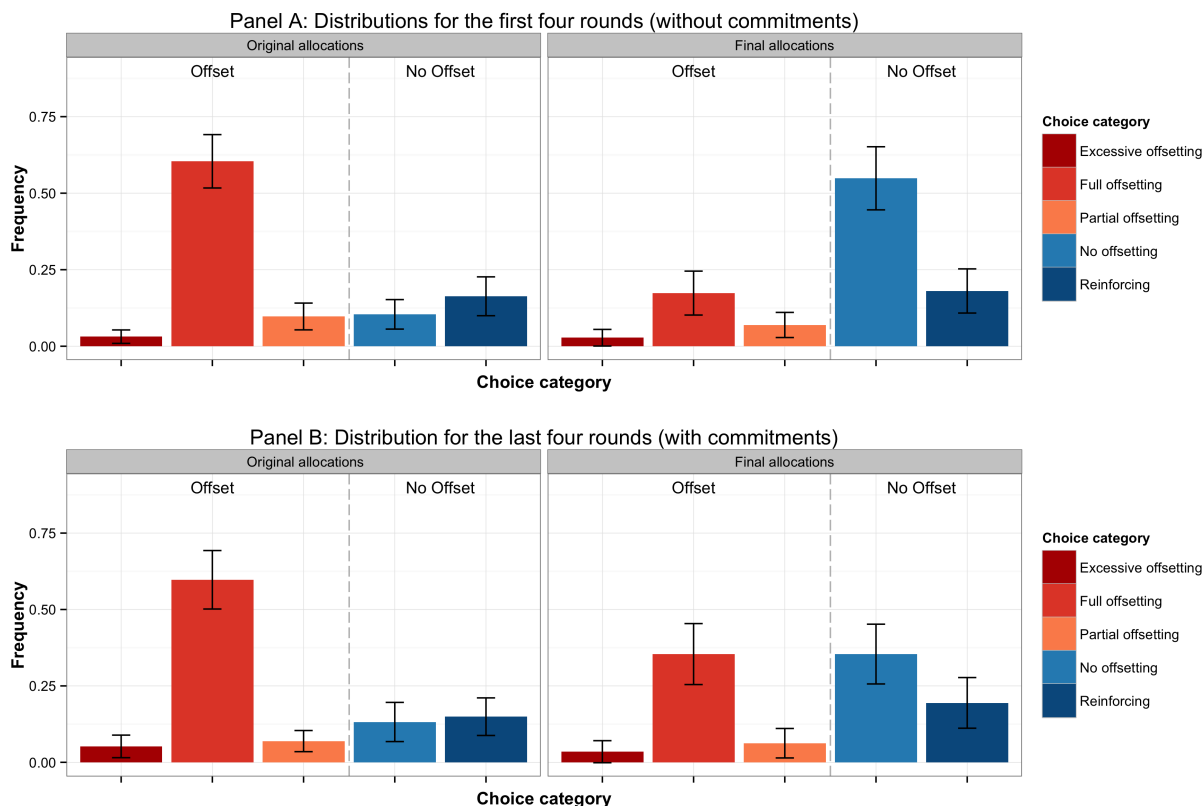
Despite the possibility mentioned in the previous paragraph, our second finding is that the frequency of revisions falls dramatically when commitment opportunities are introduced. Specifically, only 36.8% of decisions were revised in the last four rounds of  $4A^R_4A^C$ , which is a little more than half of the comparable frequencies from the first four rounds of the same treatment (65.3%) and the last four rounds of treatment  $4A_4A^R$  (68.3%); moreover, these differences are statistically significant ( $p < 0.001$  in both cases). Similarly, 51.4% of subjects revised at least one decision in the last four rounds of  $4A^R_4A^C$ , compared with 80.6% in the first four rounds of the same treatment and 78.9% in the last four rounds of treatment  $4A_4A^R$ ; these differences are also statistically significant ( $p = 0.001$  and  $p < 0.001$ , respectively). Thus, opportunities for commitment are plainly consequential.

Our third finding is that commitment opportunities significantly change the distribution of final choices. Panels A and B of Figure 9 show the distributions of allocation types for, respectively, the first and last four rounds of treatment  $4A^R_4A^C$ . In each case, the distribution of original allocations is on the left, and the distribution of final allocations is on the right. Comparing the

---

<sup>18</sup>See the Appendix for details.

**Figure 9:** Marginal distributions of original and final (revised) choices for treatment  $4A^R\_4A^C$



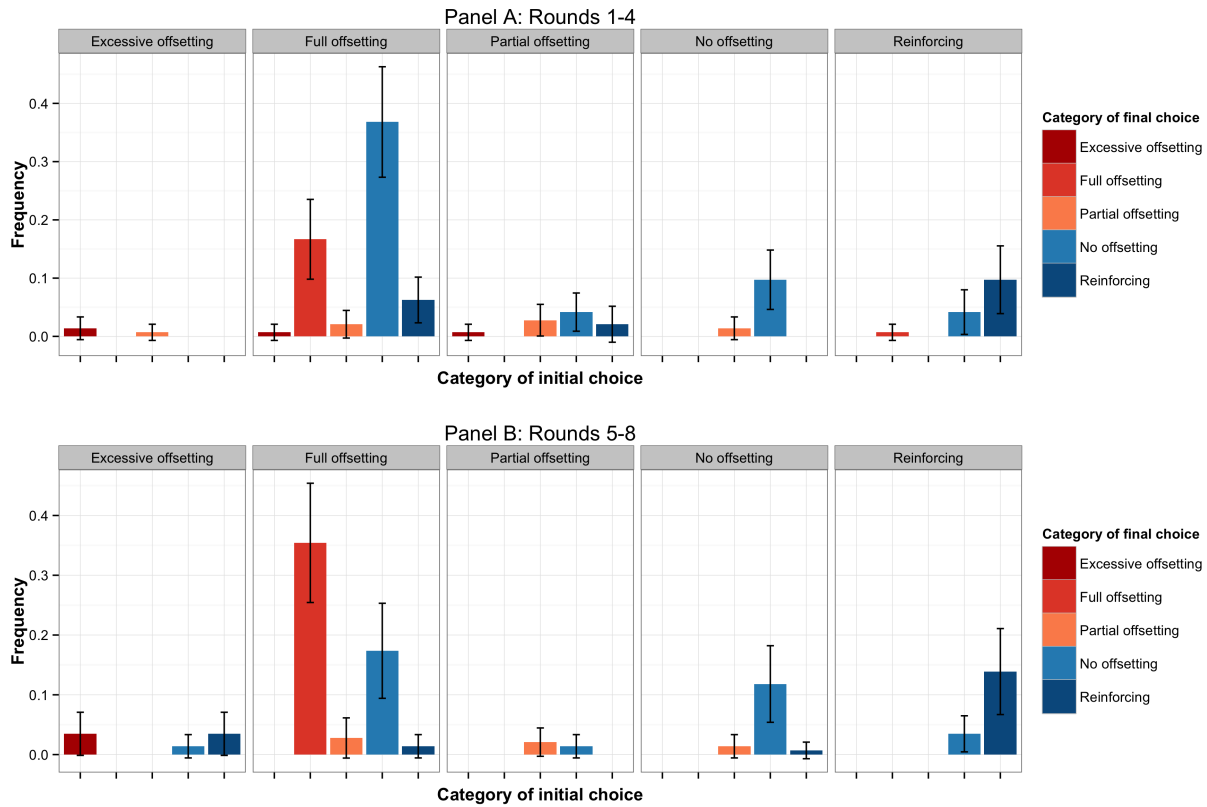
Note: This figure is based on treatment  $4A^R\_4A^C$  (72 subjects). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

distributions of original allocations, we see very little difference between the first four rounds and the last four. In fact, we do not reject the hypothesis that these two distributions are identical ( $p = 0.43$ ). In contrast, there are striking and statistically significant differences between the distributions of final outcomes ( $p < 0.001$ ).

If the availability of commitment opportunities generally works as we have hypothesized, we would expect the frequency of ex ante fair allocations to be higher, and that of ex post fair allocations to be lower, in the lower-right-hand panel of Figure 9 than in the upper-right-hand panel. That is indeed what we find: the frequency of ex ante fair allocations is about 18 percentage points higher (35.4% vs. 17.4%) in the last four rounds (with commitment) than in the first four (without commitment), and the frequency of ex post fair allocations is about 20 percentage points lower (35.4% versus 54.9%).

A look at that joint distributions of initial and final choices confirms that commitment opportunities mostly suppress migration from ex ante to ex post fair choices. Figure 10 shows these

**Figure 10:** Joint distribution of original and final (revised) choices for treatment  $4A^R_{4A^C}$



This figure is based on the treatment  $4A^R_{4A^C}$  (72 subjects). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

distributions separately for rounds 1-4 (panel A) and rounds 5-8 (panel B) of treatment  $4A^R_{4A^C}$ . The most striking aspect of the comparison between the two panels is that the frequency with which subjects choose ex ante fair allocations and stick with them increases dramatically when commitments are possible, from 16.7% to 35.4% (18.7 percentage points), while the frequency with which subjects choose ex ante fair allocations and then switch to ex post fair allocations falls by a nearly identical amount, from 36.8% to 17.4% (19.2 percentage points). The fractions of individuals choosing and sticking with three of the other four options also decline, but the changes are modest by comparison.

For the same reasons as in previous sections, it is also important to examine the behavior of subjects who exhibited a reasonable degree of consistency across tasks. For 36 of the 72 subjects in this treatment, original choices fell into the same category throughout rounds 5-8; in 30 of these

cases, the initial allocations were ex ante fair.<sup>19</sup> We will call these the “consistent” subjects, and we will call the remaining 36 subjects “inconsistent.” The preference for commitment appears to be somewhat stronger for consistent subjects, who committed themselves in 48.7% of tasks and retained flexibility in 26.8%, while the inconsistent subjects committed themselves in 29.8% of tasks and retained flexibility in 34.7%.<sup>20</sup>

Figure 11 exhibits distributions of final choices for consistent subjects who started out by choosing the ex ante fair allocation. (We do not display the rest of the joint distribution because consistent subjects started out by making other types of choices so infrequently.) Panel A pertains to rounds 1-4 of treatment  $4A^R$ – $4A^C$ , and panel B to rounds 5-8. In each case, we define a subject as consistent or inconsistent based on their behavior within the indicated rounds.<sup>21</sup> Here we see a 30 percentage point increase in the frequency of final ex ante fair allocations, from 32% in the first four rounds (without commitment), to 62% in the last four rounds (with commitment), and a 27 percentage point decline in the frequency of final ex post fair allocations decreases (62% versus 35%). Thus, among consistent subjects, offering commitment suppresses migration from ex ante fair allocations to ex post fair allocations.

## 8.2 Is there a preference for flexibility and revision?

We have seen that subjects make commitments with high frequency, and that these commitments reduce the frequency of revisions, primarily from ex ante to ex post fair allocations. As is clear from figure 10, it is also the case that many subjects opt for flexibility and then revise their allocations. How can we account for both findings?

One possibility is that the theories discussed in section 2 are correct but the population is heterogeneous. Under this view, one attributes the preference for, and effects of, commitment to time inconsistency among sophisticated subjects with non-EU preferences that are non-separable over states of nature, and the preference for flexibility and switching to nominal fairness. However, there are other possibilities. In principle, naiveté (lack of self-awareness) among time-inconsistent subjects could explain why some subjects maintain flexibility and then revise their allocations, and experimenter demand effects could account for all of these observations.

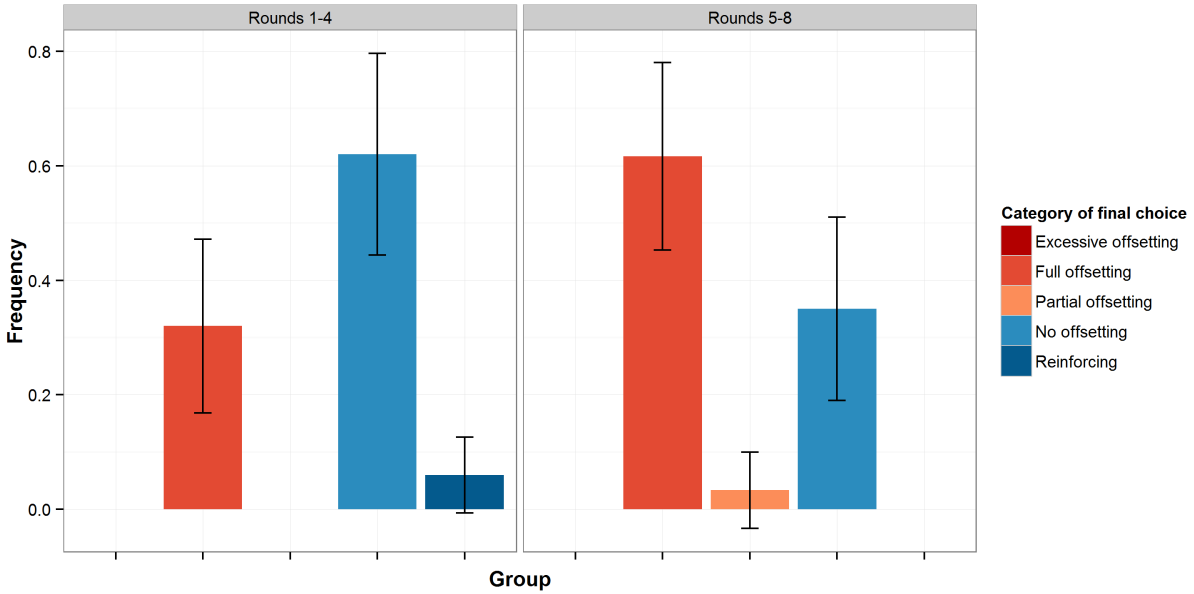
To illustrate the potential role of experimenter demand effects, imagine that, when faced with two consequential alternatives and an option to express indifference, subjects feel they are expected to choose one of the former. Suppose this causes them to make commitments in a significant fraction of allocation tasks – say 40% of them, selected at random. How would this affect the overall distribution of final outcomes? Recall two important facts: first, ex ante choices are predominantly

<sup>19</sup>Two of these subjects consistently selected reinforcing allocations, and four consistently opted for ex post fairness.

<sup>20</sup>See the Appendix for details.

<sup>21</sup>There were 29 consistent subjects in rounds 1-4, and 36 in rounds 5-8. 22 of these were the same subjects. 25 consistent subjects always chose the ex ante fair allocation in rounds 1-4, and 30 did so in rounds 5-8. 20 of these were the same subjects.

**Figure 11:** Distribution of final choices of consistent subjects for treatment  $4A^R_{-4A^C}$



Note: This figure is based on consistently initially fully offsetting subjects in treatment  $4A^R_{-4A^C}$  (25 in rounds 1-4 and 30 in rounds 5-8). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

ex ante fair; second, revisions predominantly lead to ex post fairness. Consequently, the most visible impact of the hypothesized demand effect would be an increase in the fraction of ex ante fair allocations, and a decrease in the fraction of ex post fair allocations, among final outcomes. That is of course precisely what we documented in the previous section. A similar experimenter demand effect could likewise explain why other subjects retain flexibility, but this would not account for subsequent switching unless one posits a second demand effect (specifically, that offering people the opportunity to revise induces them to do so). We designed the revision protocol to minimize that possibility, but it still merits consideration.

In this section, we present a series of findings that cast additional light on subjects' reasons for making or not making commitments. These findings speak to three questions. First, which subgroups exhibit the greatest demand for commitment? Second, what do subjects do with flexibility when they intentionally retain it? Third, do the subjects who retain and exercise flexibility appreciate ex ante that they are likely to revise their allocations ex post?

Consider the first question: which subgroups have the greatest demand for commitment? It is important to acknowledge that experimenter demand effects may establish baseline frequencies with which subjects opt for commitment and flexibility. However, even if that is the case, our theories of fairness remain testable because they imply different patterns of deviations from the baseline. If

the primary purpose of commitments is to impede undesired revisions from ex ante fair to ex post fair allocations, then the demand for commitment should be greater among subjects who choose initial allocations that entail a degree of fairness, and especially among those who then tend to switch to ex post fair allocations when no commitments are allowed. In contrast, if migration from ex ante fair to ex post fair allocations reflects nominal fairness, those same groups should exhibit a greater demand for flexibility. As we explain next, the evidence points to nominal fairness.

First, we find that the demand for commitment is lower, and the demand for flexibility higher, when subjects choose allocations they are more likely to revise (specifically, ones that entail a degree of ex ante fairness). When subjects started out by selecting the ex post fair allocation, the frequency with which they chose commitment was roughly three times as high as that with which they chose flexibility (52.6% vs. 15.8%). When they started out by selecting reinforcement (the only other non-offsetting category), the relative prevalence of commitment choices was nearly as large (41.8% vs. 18.6%). In contrast, when subjects started out by selecting the ex ante fair option, the frequency with which they chose commitment was only slightly larger than that with which they chose flexibility (42.4% vs. 33.7%). When they started out by choosing either an excessively offsetting or partially offsetting allocation, the relative frequency of a preference for commitment was even lower (20.0% vs. 40.0%, and 15.0% vs 45.0%, respectively).<sup>22</sup>

Second, we find that the demand for commitment is lower, and the demand for flexibility higher, among subjects who exhibit a greater tendency to migrate from ex ante fair to ex post fair allocations when no commitments are allowed. Recall that every subject had two opportunities to revise initial allocations during the first four rounds, and no opportunities to make commitments. Twenty-six subjects always switched from ex ante to ex post fair allocations when given the opportunity to do so during the first four rounds, and thirty-one subjects never switched from one category to another. Over the course of the final four rounds, the first group (consistent switchers) expressed a preference for commitment 28% of the time and preference for flexibility 46% of the time, while the second group (consistent non-switchers) expressed a preference for commitment 53% of the time and a preference flexibility 22% of the time. These differences are statistically significant ( $p = 0.0257$ ). Thus, those inclined to switch during the first four rounds were much more likely to prefer flexibility, and much less likely to prefer commitment, during the last four rounds than those lacking that inclination.

Now we consider the second question: how do subjects exercise flexibility when they intentionally retain it? Here we acknowledge that experimenter demand effects may establish a baseline frequency for revisions. However, that possibility does not explain the specific observation that revisions tend to yield ex post fair outcomes. Closer examination of revisions allows us to differentiate between the hypotheses of interest. If the primary purpose of commitments is to impede

---

<sup>22</sup>See the Appendix for additional details.

undesired revisions from ex ante fair to ex post fair allocations, then we would expect to find that the subjects who opt for flexibility are disproportionately time-consistent, in which case they should exhibit relatively low rates of revision and migration from ex ante fair to ex post fair allocations. In contrast, if migration from ex ante fair to ex post fair allocations reflects nominal fairness, that pattern should be particularly prevalent among those who affirmatively choose flexibility. As we explain next, the evidence again points to nominal fairness.

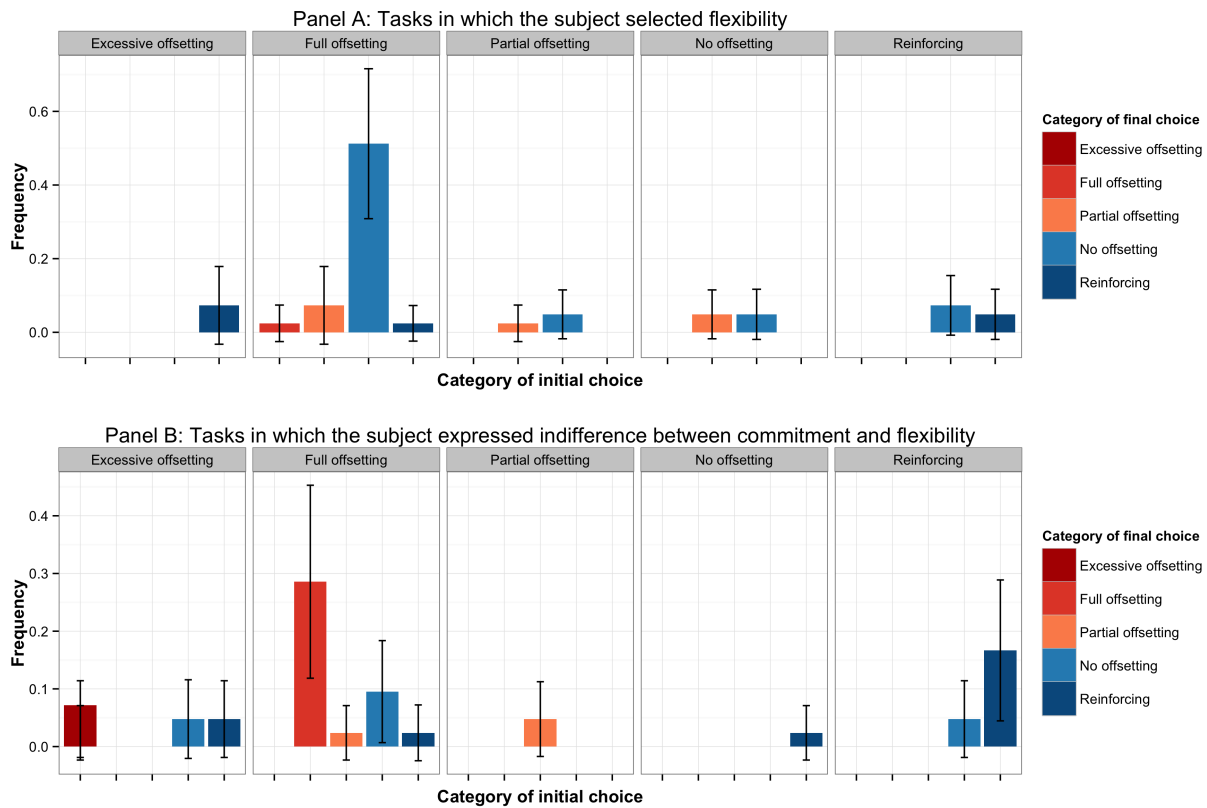
First, we find that, in tasks with commitment options, the revision rate is exceptionally high among those who opt for flexibility. Overall, subjects revised 85.4% of decisions in tasks where they chose flexibility over commitment. Significantly, that figure is higher, not lower, than the comparable figures for the first four rounds (65.3%), and for the last four rounds of treatment 4A\_4A<sup>R</sup> (68.3%).

Second, we find that those who opt for flexibility are disproportionately inclined to migrate from ex ante fair to ex post fair allocations. Figure 12 shows the joint distributions of original and final choices for those who affirmatively retained the flexibility to revise (panel A), as well those who said they were indifferent between commitment and flexibility, half of whom were allowed to revise (panel B). Focusing on panel A, we see that 66.7% of the original choices were ex ante fair. Of those, 80.8% were revised to ex post fair choices. Thus, migration from ex ante to ex post fairness predominates among uncommitted choices to a greater extent than in any of the previous figures – it accounts for 51.2% of the choice pairs (versus 36.8% in panel A of Figure 10); we reject equality of these fractions ( $p < 0.001$ ). This pattern suggests that those who migrate from ex ante fair to ex post fair choices actually prefer the flexibility to migrate.

Now consider panel B, the joint distribution for those who expressed indifference. While the sample here is small, it is worth noting that only 9.5% of the choice pairs exhibited migration from ex ante to ex post fairness. Instead, the most common pattern was to select the ex ante fair allocation and stick with it. Despite the small sample, we can reject the hypothesis that migration from ex ante to ex post fairness is as common in panel B as in panel A ( $p = 0.03$ ). This contrast between panels A and B again suggests that those who intentionally avoid commitments affirmatively value the ability to switch from an ex ante fair choice to an ex post fair one, and have no desire to preclude this migration.

Finally, consider the third question: do the subjects who retain and exercise flexibility appreciate ex ante that they are likely to revise their allocations ex post? Although we cannot observe what people think, we can track their experience during the experiment. If naive time inconsistency accounts for the retention of flexibility and the subsequent switching from ex ante fair to ex post fair allocations, we would expect the same individuals to have made few if any prior decisions that would have revealed their inconsistency. In contrast, if this pattern reflects nominal fairness, we would expect the same individuals to have behaved similarly throughout the experiment, also

**Figure 12:** Joint distribution of original and final (revised) choices in uncommitted allocation tasks



Notes: This figure is based on the final four rounds of treatments  $4A^R$  and  $4A^C$ . Panels A and B are based on 87 and 84 uncommitted allocation tasks, respectively. Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

migrating from ex ante fair to ex post fair allocations during the first four rounds. Once again, the evidence points to nominal fairness.

Altogether, during the last four rounds, we observed the “initial ex ante fair & no commitment & revised ex post fair” pattern in 21 tasks involving 14 subjects, and the “initial ex ante fair & commitment” pattern in 38 tasks involving 25 subjects. Focusing on the first group of tasks (in which the subject opted for flexibility and then switched), in 61.9% of those cases the same subject always migrated from ex ante fair to ex post fair allocations when given the opportunity during the first four rounds, and in 95.2% of cases did so at least once.<sup>23</sup> In contrast, focusing on the second group of tasks (in which the subject opted for commitment), the corresponding frequencies

<sup>23</sup>Of the 7 cases where subjects did not migrate from the ex ante fair to the ex post fair allocation in the first four round, two involved migration from partial offsetting to the ex post fair allocation, two involved migration from reinforcement to the ex post fair allocation, and three involved ex post allocations with no switching.



are considerably lower: in 21.1% of those cases the same subject always migrated from ex ante fair to ex post fair allocations when given the opportunity during the first four rounds, and in 60.5% of those cases did so at least once. Thus, those who both chose and exploited the flexibility to switch from ex ante to ex post fair allocations during the last four rounds were disproportionately likely to have made that switch during the first four rounds. We conclude that those preserving the flexibility to migrate from ex ante fair to ex post fair allocations likely understood and anticipated their inclination to do so.

## 9 Analysis of split-the-prize tasks

Throughout the previous sections, we have focused on divide-the-tickets tasks. Here we ask whether similar results emerge in the context of divide-the-prize tasks. As explained in section 2, the conditions giving rise to time inconsistency are more restrictive in this setting. Specifically, for divide-the-prize tasks, time inconsistency requires non-separability across states of nature (as in the EU-distribution formulation); it does not arise in separable formulations with probability weighting.

The following analysis is based on a  $4A\_4A^R$  treatment with divide-the-prize tasks. It is analogous to the corresponding portions of Sections 4 and 6, which are otherwise structured similarly.

Figure 13 shows the distribution of choices for rounds 1-4 (panel A), as well as the marginal distributions of the original and final choices (panels B and C, respectively) for rounds 5-8; it is analogous to Figures 1 and 6. For the moment, we will focus on the ex ante choices (panels A and B), and return to the revisions (panel C) below. Notice that, when subjects choose ex ante, the most common type of allocation is ex ante fair. The tendency to make ex ante fair choices with ex ante framing is not quite as pronounced as with divide-the-tickets tasks (Figure 1), but it is still readily evident. Significantly, the somewhat lower frequency of ex ante fair choices with ex ante framing in divide-the-prize tasks (compared with divide-the-tickets tasks) goes hand-in-hand with a somewhat higher frequency of ex post fair allocations. This is expected in light of the theoretical considerations discussed in Section 2: subjects with separable but probability-weighted preferences will prefer ex ante fair allocations in ex ante divide-the-tickets tasks, and ex post fair allocations in ex ante divide-the-prize tasks.

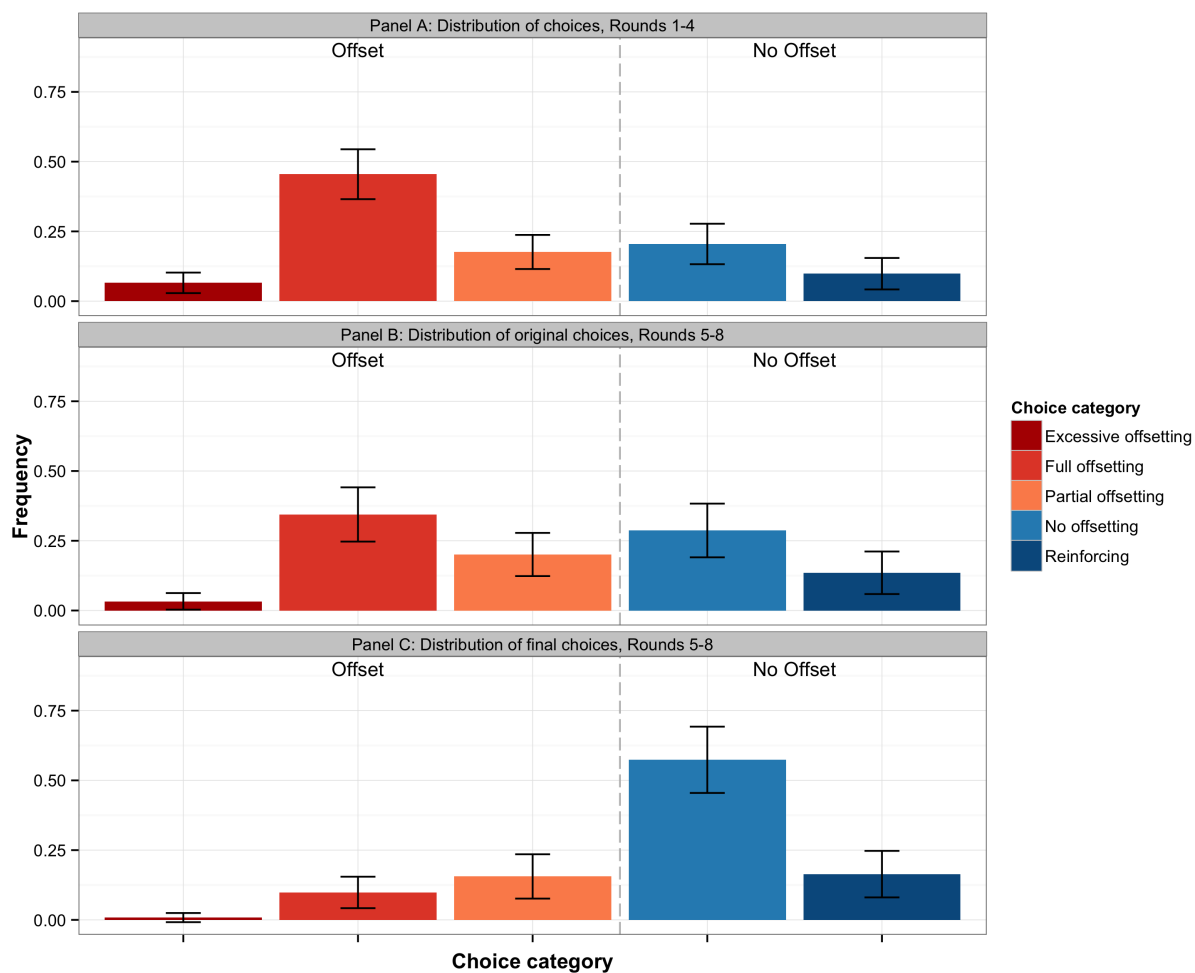
Significantly, revisions were common in rounds 5-8 of this treatment. Overall, 42.6% of choices were revised, and 55.7% of subjects revised at least one choice. Furthermore, the vast majority of revisions (73.1%) involved migration to ex post fair allocations, just as with divide-the-tickets tasks.<sup>24</sup>

Figure 14 displays the joint distribution of the original and final choices for rounds 5-8; it is analogous to Figure 5. Although migration from ex ante fair to ex post fair choices is not quite as common as for divide-the-tickets tasks, it remains the most common pattern (26.2% of tasks).

---

<sup>24</sup>See the Appendix for additional details.

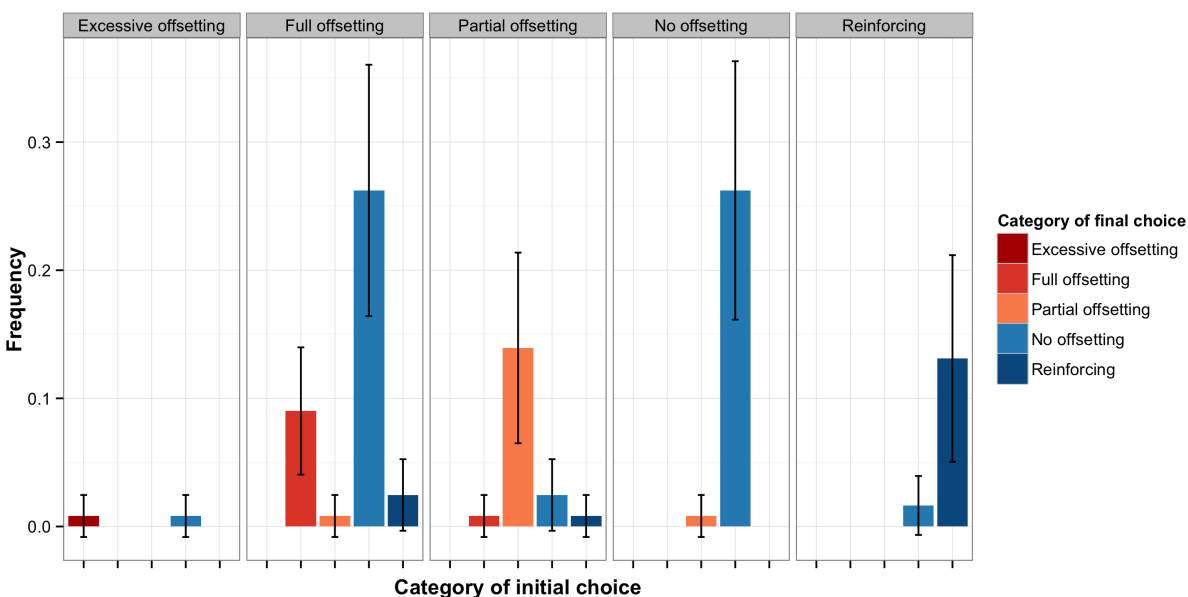
**Figure 13:** Distributions of choices in divide-the-prize tasks



This figure is based on treatment 4A\_4A<sup>R</sup> with divide-the-prize tasks (61 participants). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

Significantly, in this case it is tied with a time-consistent pattern: selecting and sticking with the ex post fair allocation. The prevalence of time-consistent ex post fair choices is expected in light of our observations concerning the implications of preferences with separability across states of nature. The next three most common patterns are also time-consistent. In 13.93% of tasks, subjects made and resolutely stuck to partially offsetting allocations. This pattern was relatively rare in divide-the-tickets tasks; apparently, the divide-the-prize setting is more conducive to reconciling the conflict between ex ante and ex post fairness by adopting and resolutely sticking to a compromise standard. In 13.11% of tasks, subjects made and stuck to choices that reinforced the computer’s allocation, and in 26.23% of tasks, they selected and stuck to the ex ante fair allocation. The latter two frequencies are comparable to those observed in the context of divide-the-tickets tasks.

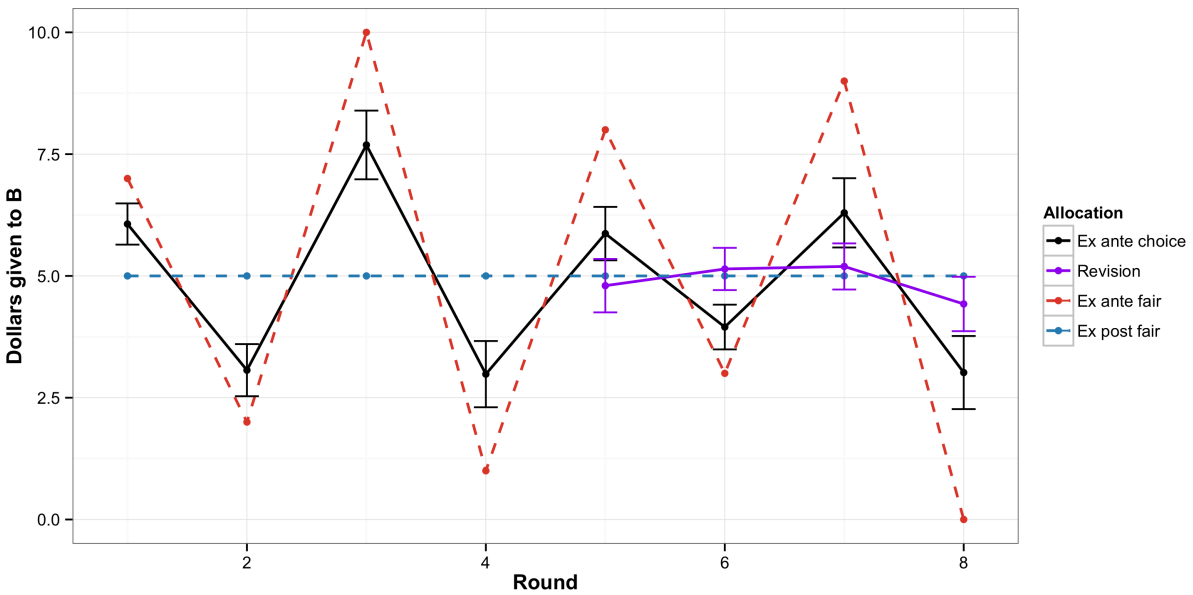
**Figure 14:** Joint distribution of original and final (revised) choices during the final four rounds of treatment 4A.4A<sup>R</sup>



Notes: This figure is based on the final four rounds of treatment 4A.4A<sup>R</sup> with divide-the-prize tasks (244 observations). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

Figure 15 compares the patterns of the average original and final allocations for rounds 5-8 with the “fingerprints” associated with ex ante and ex post fairness; it is analogous to figures 2 and 7. The average ex ante choices resemble the ex ante fair benchmark, except that responses to the computer’s allocation are dampened. Revisions in rounds 5-8 flatten the line further, moving it toward the ex post fair benchmark. As in earlier sections, we quantify the similarity to the benchmarks by estimating simple regressions of the chosen split on a constant and the computer’s

**Figure 15:** Original and final allocations for all rounds of treatment 4A\_4A<sup>R</sup> with divide-the-prize tasks



Notes: This figure is based on treatment 4A\_4A<sup>R</sup> with divide-the-prize tasks (61 participants). Error bands indicate 95% confidence intervals.

split, clustering observations at the subject level. The coefficient of the computer's split is  $-0.54$  (s.e. = 0.06) for ex ante decisions in the first four rounds,  $-0.37$  (s.e. = 0.07) for ex ante decisions in the last four rounds, and  $-0.06$  (s.e. = 0.05) for revised decisions in the last four rounds.

All of the results reported in this section are therefore qualitatively similar to their counterparts in previous sections. The patterns of interest are somewhat less striking, but this is expected given that some subjects' preferences may be separable across states of nature.

## 10 Implications for theories of fairness

Here we summarize the implications of our results for the various theories of fairness preferences discussed in section 2. When evaluating these theories, it is important to bear in mind that the population is likely heterogeneous, with different principles governing the choices of different subjects. Our object is therefore not to determine whether or not each theory is correct, but rather to shed some light on the size of the population segment to which it likely applies.

**EU preferences:** Given the demonstrated strictness of preferences, this formulation cannot account for systematic equalization of overall chances to win in both ex ante and ex post split-the-tickets tasks, full or partial offsetting of fixed allocations in ex ante split-the-prize tasks, revisions

in both contexts, or the voluntary adoption of consequential commitments. In split-the-prize tasks, the EU formulation has the strong implication that a decision maker with symmetric preferences will divide the prize equally both ex ante and ex post. Yet this pattern materializes in only a little over one-quarter of tasks, and even those choices are consistent with other theories. Thus we find little support for the EU formulation.

**Consequential EU distribution preferences:** This formulation can account for systematic equalization of overall forward-looking prospects, as well as time-inconsistency, in both split-the-ticket tasks (as a limiting case) and split-the-prize tasks. However, it cannot explain the lower prevalence of these patterns in split-the-prize tasks. It can account for the voluntary adoption of consequential commitments, but not for the fact that those who migrate from ex ante to ex post fair allocations exhibit a predominant preference for flexibility.

**Non-EU preferences with probability weighting:** This formulation can account for the systematic equalization of overall forward-looking chances to win, as well as time-inconsistency, in split-the-tickets tasks. It is inconsistent with the appearance of analogous patterns in split-the-prize tasks, but consistent with their lower prevalence. Thus, it is possible to account for our results concerning frame-dependence and time-inconsistency by positing that the population includes a mix of individuals with EU-distribution and probability-weighted preferences. However, our results concerning commitments contradict that theory. Like EU-distribution preferences, the probability-weighted formulation can account for the voluntary adoption of consequential commitments, but not for the fact that those who migrate from ex ante to ex post fair allocations exhibit a predominant preference for flexibility. The latter finding limits the explanatory power of both formulations.

**Resolute non-EU preferences:** Given the demonstrated strictness of preferences, this formulation cannot account for the prevalence of revisions or for consequential commitments. That said, we do observe some apparently resolute decision makers. A small but significant number of subjects appear to have resolute preferences for ex ante fairness. A similar number exhibit resolute anticipatory EU-distribution preferences (partial offsetting) in split-the-prize tasks, but not in split-the-tickets tasks. In both settings, a similar number resolutely reinforce the fixed allocation, but we suspect that this finding reflects a tendency among those subjects to draw false inferences about relative need from the “computer’s” choice. Thus we find a small degree of support for this class of preference formulations.

**Anticipatory preferences:** Regardless of whether one assumes EU-distribution or probability-weighted utility functions, this class of formulations implies partial offsetting of fixed allocations in ex ante split-the-tickets tasks. In fact, partial offsetting is relatively rare in those settings. Thus we see no indication that subjects place weight on future fairness evaluations that are expected to differ from current evaluations.

**Nominal fairness:** This formulation receives the greatest support because it rationalizes the

most common modes of behavior. Assuming the subject pool consists of a blend of nominally fair individuals with either EU-distribution or probability-weighted preferences, we would observe systematic equalization of overall forward-looking prospects as well as time-inconsistency in all tasks, but a lower prevalence of those patterns in split-the-prize tasks, where selecting the ex post fair allocation ex ante and sticking with it ex post would be more common. In addition, we would observe a preference for flexibility among those who migrate from ex ante fair to ex post fair allocations. All of these patterns occur with high frequency, and they are in fact the most common patterns by wide margins. Thus we conclude that nominal fairness likely accounts for a substantial fraction of observed behavior.

## 11 Conclusions

In this paper, we have explored experimentally how people think about fairness in settings with uncertainty, where there is a tension between pursuing equality of opportunity and equality of outcomes. Our inquiry is motivated by the observation that each individual may experience this tension as a conflict between the perspectives she takes at different points in time. As a result, concerns for fairness could in principle give rise to time inconsistency.

Our analysis supports the following conclusions concerning behavior in tasks involving allocations of chances to win a cash prize. First, when someone confronts a new collection of allocation problems, the framing of the initial decisions – whether they are made ex ante or ex post the resolution of uncertainty – strongly influences the perspective on fairness expressed through choices. Second, people do not carry over that initial perspective into subsequent allocation tasks. Rather, they move back and forth between ex ante and ex post perspectives on fairness, according to whether the decision frame for the task in question is ex ante or ex post. Third, this tendency to switch between perspectives generates time inconsistency. Some people resolutely adhere to the ex ante perspective, but they are a small minority. Fourth, experience does not ameliorate time inconsistency. Even when people are confronted with their inconsistencies, many of them continue to switch from one perspective to another and revise their contingent plans as events unfold. Finally, self-awareness of time-inconsistency regarding fairness may create some demand for commitments, but many of those who migrate from ex ante to ex post perspectives apparently prefer having the flexibility to switch. We have also verified that the first and third patterns are also present when subjects perform tasks involving the allocation of dollars, although they are a bit less pronounced, whereas the consistent application of ex post standards is more common.

We have discussed the implications of our findings for theories of fairness preferences. While the population is plainly heterogeneous, it is possible to account for the most common behavioral patterns by assuming that the subject pool consists of nominally fair individuals with a mix of EU-distribution and probability-weighted preferences, along with smaller numbers of individuals

with resolute non-EU preferences. Expected utility formulations are rejected, as are theories in which subjects evaluate the fairness of a choice based on where it will actually lead, rather than where it would lead if no further action were taken.

These findings have potentially important implications for public policy. Analogous issues arise in the context of the Samaritan's dilemma, wherein ex post altruism undermines ex ante altruistic objectives through incentive mechanisms. Our analysis demonstrates that concerns for fairness potentially create similar issues even when ex post actions do not compromise the ex ante incentives of the affected parties.

On a more conceptual level, one can think of this paper as a positive investigation of normative ethics. We do not attempt to derive criteria for judging whether a choice is ethical. Instead, our research sheds light on the criteria people actually use. It points toward a deontological perspective, wherein people judge the morality of an action non-consequentially, according to its consistency with ethical rules. Whether the judgments we identify resonate with a particular flavor of deontology (such as Kantianism) is an interesting question, but one that ventures beyond the more pragmatic objectives of the current study.

Indeed, our findings raise other important questions that are well worth consideration in future research. Why, for instance, do many *sophisticated* subjects cling to time inconsistency? Are they habitually applying a familiar ethical rule, or are they proceeding more thoughtfully from a value system or moral paradigm? Would they continue to adopt a time-inconsistent posture toward fairness if they had a direct stake in the outcome, or would they rationalize a self-serving ethical perspective? Would one perspective become more compelling if one of the recipient households were viewed as more deserving? Additionally, what factors (such as political beliefs or other demographic variables) predict the mix of different preference types? Investigating these and other central questions raised by this study will, we hope, contribute to a deeper and more complete understanding of social preferences.

## References

- Arun Agrawal. Indigenous knowledge and the politics of classification. *International Social Science Journal*, 54(173):287–297, 2002.
- James Andreoni and B. Douglas Bernheim. Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5):1607–1636, September 2009.
- James Andreoni and John Miller. Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2):pp. 737–753, 2002.
- James Andreoni, Paul M. Brown, and Lise Vesterlund. What makes an allocation fair? some experimental evidence. *Games and Economic Behavior*, 40(1):1–24, 2002.

- James Andreoni, Marco Castillo, and Ragan Petrie. What do bargainers' preferences look like? experiments with a convex ultimatum game. *American Economic Review*, 93(3):672–685, 2003.
- B. Douglas Bernheim and Sergei Severinov. Bequests as signals: An explanation for the equal division puzzle. *Journal of Political Economy*, 111(4):pp. 733–764, 2003.
- David E. Bloom. Empirical models of arbitrator behavior under conventional arbitration. *The Review of Economics and Statistics*, 68(4):pp. 578–585, 1986.
- Iris Bohnet and Bruno S. Frey. Social distance and other-regarding behavior in dictator games: Comment. *American Economic Review*, 89(1):335–339, 1999.
- Gary E Bolton and Axel Ockenfels. Erc: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90(1):pp. 166–193, 2000.
- Gary E Bolton, Jordi Brandts, and Axel Ockenfels. Fair procedures: Evidence from games involving lotteries. *Economic Journal*, 115(506):1054–1076, October 2005.
- J Michelle Brock, Andreas Lange, and Erkut Y Ozbay. Dictating the risk: Experimental evidence on giving in risky environments. *The American Economic Review*, 103(1):415–437, 2013.
- Colin F. Camerer. Progress in behavioral game theory. *Journal of Economic Perspectives*, 11(4): 167–188, 1997.
- Alexander W. Cappelen, James Konow, Erik O. Sorensen, and Bertil Tungodden. Just luck: An experimental study of risk-taking and fairness. *American Economic Review*, 103(4):1398–1413, September 2013.
- Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3):817–869, August 2002.
- Stephen Coate. Altruism, the samaritan's dilemma, and government transfer policy. *The American Economic Review*, pages 46–57, 1995.
- Sudipto Dasgupta and Zhigang Tao. Incomplete contracts, ownership rights, and the optimality of equity joint ventures. *Journal of Economic Behavior and Organization*, 37(4):391–413, 1998.
- M. De Weaver and J. Roumasset. Risk aversion as effort incentive: A correction and prima facie test of the moral hazard theory of share tenancy. *Economics Bulletin*, 15:1–16, 2002.
- Peter A Diamond. The role of a stock market in a general equilibrium model with technological uncertainty. *The American Economic Review*, pages 759–776, 1967.



- Ernst Fehr and Klaus M. Schmidt. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868, August 1999.
- Robert Forsythe, Joel L. Horowitz, N.E. Savin, and Martin Sefton. Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3):347 – 369, 1994.
- Drew Fudenberg and David K Levine. Fairness, risk preferences and independence: Impossibility theorems. *Journal of Economic Behavior & Organization*, 81(2):606–612, 2012.
- J Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of welfare. *J. Polit. Econ*, 63, 1955.
- Robert Hauswald and Ulrich Hege. Ownership and control in joint ventures: Theory and evidence. In *Centre for Economic Policy Research*, 2003.
- Elizabeth Hoffman, Kevin A. McCabe, and Vernon L. Smith. On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory*, 25(3):289–301, 1996.
- Edi Karni, Tim Salmon, and Barry Sopher. Individual sense of fairness: an experimental study. *Experimental Economics*, 11(2):174–189, 2008.
- Philipp Kircher, Sandra Ludwig, and Alvaro Sandroni. On the difference between social and private goods. *The BE Journal of Theoretical Economics*, 13(1), 2013.
- Michal Krawczyk and Fabrice Le Lec. ‘give me a chance!’ an experiment in social decision under risk. *Experimental economics*, 13(4):500–511, 2010.
- Mark J Machina. Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, pages 1622–1668, 1989.
- Harry M Markowitz. *Portfolio selection: efficient diversification of investments*, volume 16. Yale university press, 1968.
- Edward F McClennen. Justice and the problem of stability. *Philosophy & public affairs*, pages 3–30, 1989.
- Paul L. Menchik. Primogeniture, equal sharing, and the u.s. distribution of wealth. *The Quarterly Journal of Economics*, 94(2):pp. 299–316, 1980.
- Paul L. Menchik. *Modelling the Accumulation and Distribution of Wealth*, chapter Unequal Estate Division: Is It Altruism, Reverse Bequests, or Simply Noise? Oxford University Press, 1988.
- Matthew Rabin. Incorporating fairness into game theory and economics. *The American Economic Review*, 83(5):pp. 1281–1302, 1993.

Howard Raiffa. *Decision Analysis Introductory Lectures On Choices Under Uncertainty*. Addison-Wesley, 1968.

Reinhilde Veugelers and Katrien Kesteloot. Bargained shares in joint ventures among asymmetric partners: Is the matthew effect catalyzing? *Journal of Economics*, 64(1):23–51, 1996.

Mark O. Wilhelm. Bequest behavior and the effect of heirs' earnings: Testing the altruistic model of bequests. *The American Economic Review*, 86(4):pp. 874–892, 1996.

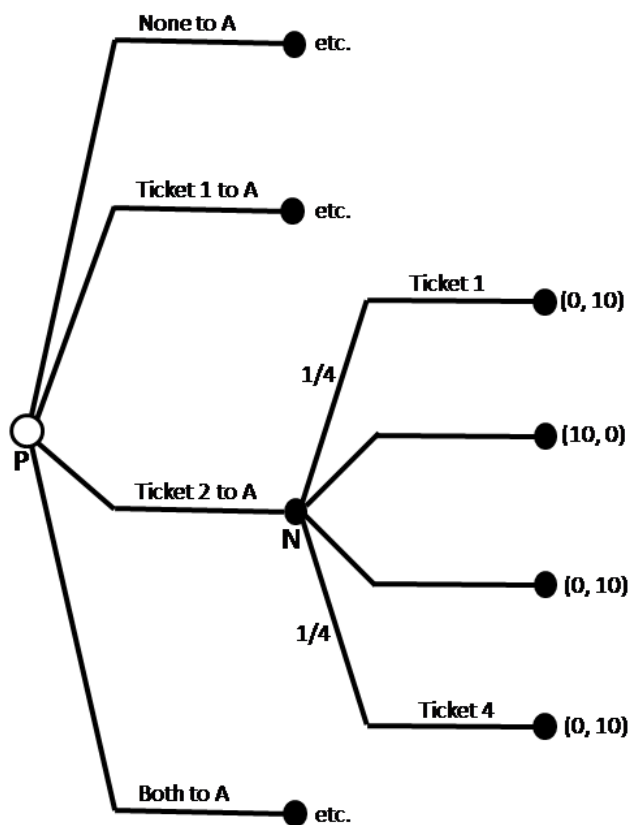
# Appendices

## A Additional Analyses

### A.1 Recognizing time inconsistency

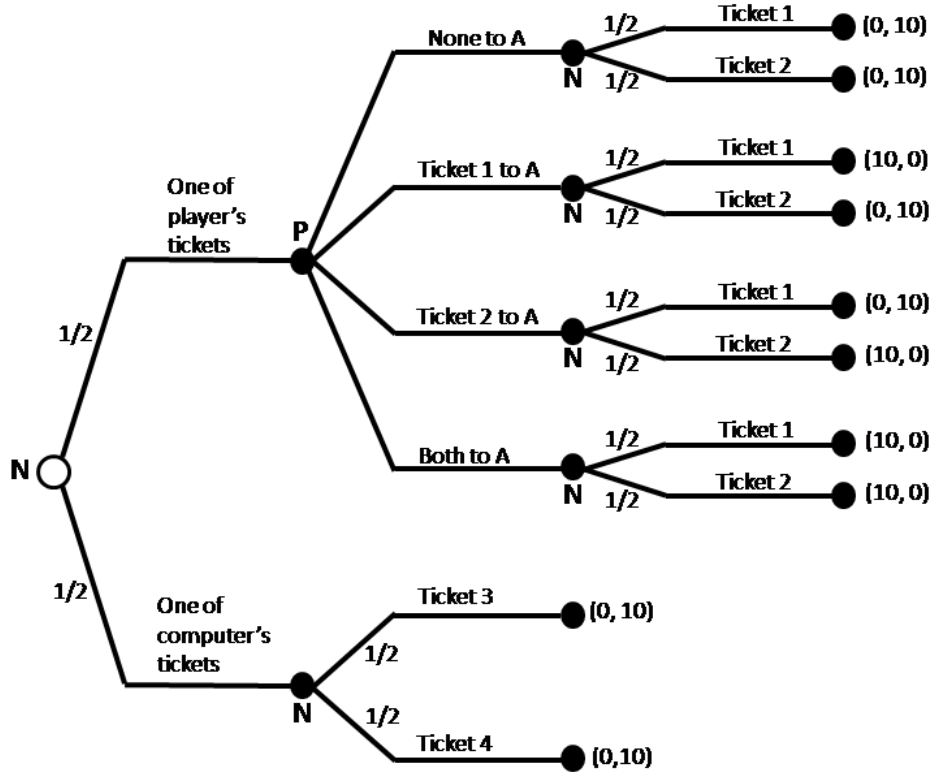
The decision trees for the ex ante and ex post split-the-tickets tasks appear in Figures 1 and 2, respectively. In Figure 1, “etc.” signifies that the continuation tree is analogous to the one shown.

Figure A.1: Game tree for the ex-ante version.



We claim that a time-consistent agent should allocate her tickets the same way in the ex ante and ex post tasks. To understand why, imagine that the decision maker is presented with the decision tree shown in Figure 2, but is told that she must plan her allocation in advance and present the plan to an agent, who will then execute it. If she is time-consistent, this procedural change will not affect the outcome. Advance decision making transforms the tree in Figure 2 into the one in Figure 3 (where “etc.” once again signifies that the continuation tree is analogous to the

Figure A.2: Game tree for the ex-post version.



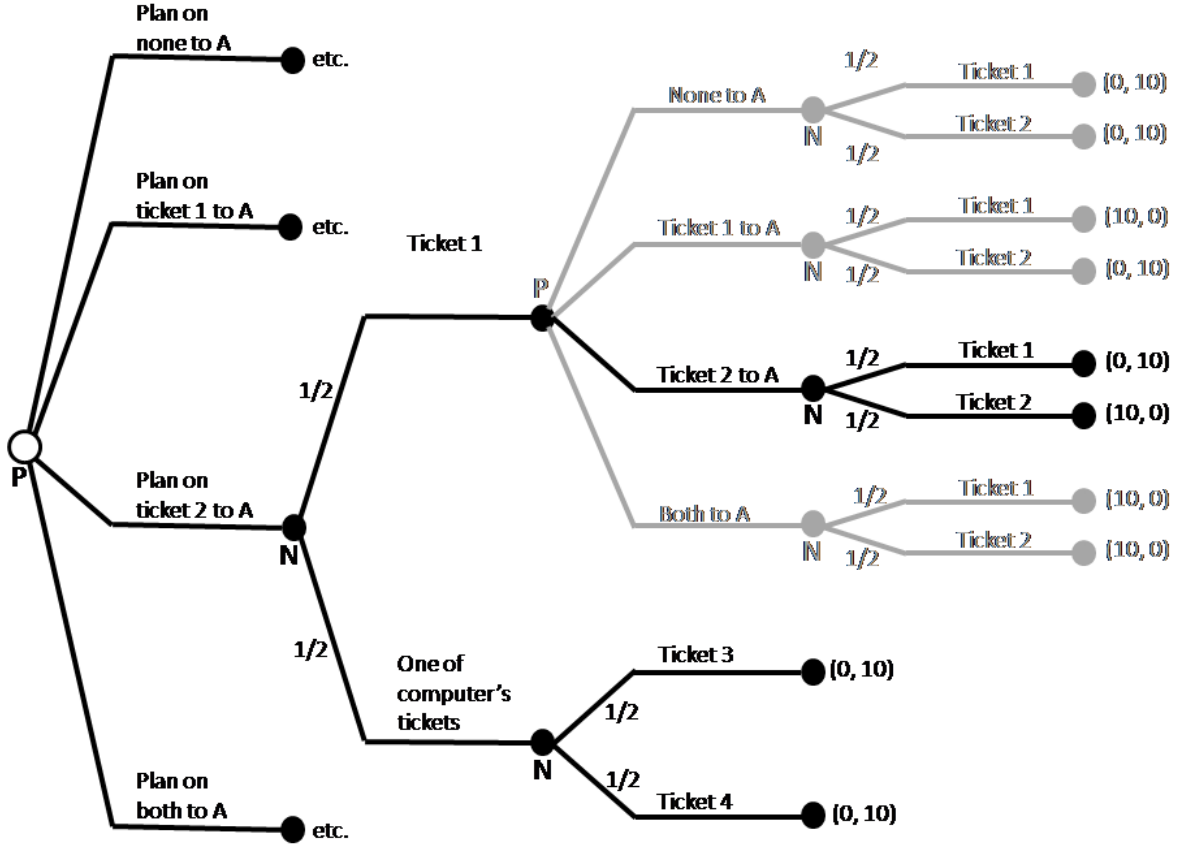
one shown). For expositional purposes, we have “greyed out” the paths the agent will not take given the decision maker’s choice. Now notice that, in Figure 3, the task ends with two consecutive choices by nature. Reducing the two-stage resolution of uncertainty into a single step, we obtain the same tree as in Figure 1. Thus, a time-consistent decision maker will make the same choice not only in Figures 2 and 3, but also in Figure 1.

## A.2 Theoretical implications

Most of the theoretical implications in Table 1 are either immediate or explained in the text. For the exceptions, the explanations appear in this section.

**Split-the-prize tasks with consequential EU-distribution preferences** In the text, we made the following claim: consumers with consequential EU-distribution preferences will partially

Figure A.3: Reduced game tree for the ex-ante version.



offset the fixed allocation in ex ante split-the-prize tasks, with the degree of offset depending on the curvature of  $w$ , and with full offset in the limit for lexicographic preferences.

Let  $x$  and  $y$  denote the fractions of the prize given to household  $A$  in the consumer's allocation and the fixed allocation, respectively. The decision maker's utility is:

$$V(L) = w\left(\frac{1}{2}u(y) + \frac{1}{2}u(x)\right) + w\left(\frac{1}{2}u(1-y) + \frac{1}{2}u(1-x)\right)$$

The first-order condition is:

$$w'\left(\frac{1}{2}u(y) + \frac{1}{2}u(x)\right)u'(x) = w'\left(\frac{1}{2}u(1-y) + \frac{1}{2}u(1-x)\right)u'(1-x)$$

Assuming concavity of  $w$  and  $u$ , this expression characterizes the optimum,  $x^*(y)$ , subject to corner constraints.

First consider the case of  $y = 0.5$ . It is immediate from the first-order condition that  $x^*(0.5) = 0.5$ .

Now suppose the  $y > 0.5$ . Evaluating the derivative of the objective function at  $x = 1 - y$ , we have

$$\left. \frac{dV}{dx} \right|_{x=1-y} = \frac{1}{2} w' \left( \frac{1}{2} u(y) + \frac{1}{2} u(1-y) \right) [u'(1-y) - u'(y)] > 0$$

Thus,  $x^*(y) > 1 - y$ . Evaluating the derivate of the objective function at  $x = 0.5$ , we have

$$\left. \frac{dV}{dx} \right|_{x=0.5} = \left[ w' \left( \frac{1}{2} u(y) + \frac{1}{2} u(0.5) \right) - w' \left( \frac{1}{2} u(1-y) + \frac{1}{2} u(0.5) \right) \right] u'(0.5) < 0$$

Thus,  $x^*(y) < 0.5$ . Plainly,  $x^*(y) \in (1 - y, 0.5)$  implies partial offset.

To understand the role of curvature of  $w$  in determining the degree of offset, consider the isoelastic specification,  $w(z) = \frac{z^{1-\alpha}}{1-\alpha}$ . To ensure that the decision maker's objective is well-defined, assume also that  $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ . For any given value of  $\alpha$ , we will write the optimum as  $x^*(y, \alpha)$ . Consider two values of  $\alpha$ ,  $\alpha' < \alpha''$ . For  $\alpha'$ , we can write the derivative of utility, evaluated at  $x^*(y, \alpha')$ , as

$$\begin{aligned} \left. \frac{dV}{dx} \right|_{x=x^*(y, \alpha'), \alpha=\alpha'} &= \left[ \left( \frac{\frac{1}{2}u(y) + \frac{1}{2}u(x^*(y, \alpha'))}{\frac{1}{2}u(1-y) + \frac{1}{2}u(1-x^*(y, \alpha'))} \right)^{1-\alpha'} - \left( \frac{u'(1-x^*(y, \alpha'))}{u'(x^*(y, \alpha'))} \right) \right] \\ &\quad \times \left( \frac{1}{2}u(1-y) + \frac{1}{2}u(1-x^*(y, \alpha')) \right)^{1-\alpha'} u'(x^*(y, \alpha')) = 0 \end{aligned}$$

For the first-order condition to hold, the first term must be zero. Now consider the same derivative evaluated at  $x = x^*(y, \alpha')$ , but for  $\alpha''$  rather than  $\alpha'$ :

$$\begin{aligned} \left. \frac{dV}{dx} \right|_{x=x^*(y, \alpha'), \alpha=\alpha''} &= \left[ \left( \frac{\frac{1}{2}u(y) + \frac{1}{2}u(x^*(y, \alpha'))}{\frac{1}{2}u(1-y) + \frac{1}{2}u(1-x^*(y, \alpha'))} \right)^{1-\alpha''} - \left( \frac{u'(1-x^*(y, \alpha'))}{u'(x^*(y, \alpha'))} \right) \right] \\ &\quad \times \left( \frac{1}{2}u(1-y) + \frac{1}{2}u(1-x^*(y, \alpha')) \right)^{1-\alpha''} u'(x^*(y, \alpha')) \end{aligned}$$

Notice that the expression in the second line is strictly positive. Thus the sign of this derivative depends entirely on the first line. Because we have already established that the decision maker partially offsets the fixed allocation, we know that

$$\frac{\frac{1}{2}u(y) + \frac{1}{2}u(x^*(y, \alpha'))}{\frac{1}{2}u(1-y) + \frac{1}{2}u(1-x^*(y, \alpha'))} > 1$$

Furthermore, with  $K > 1$ , we have

$$\frac{d}{d\alpha} K^{1-\alpha} = -K^{1-\alpha} \ln K < 0$$

Therefore,

$$\left( \frac{\frac{1}{2}u(y) + \frac{1}{2}u(x^*(y, \alpha'))}{\frac{1}{2}u(1-y) + \frac{1}{2}u(1-x^*(y, \alpha'))} \right)^{1-\alpha''} < \left( \frac{\frac{1}{2}u(y) + \frac{1}{2}u(x^*(y, \alpha'))}{\frac{1}{2}u(1-y) + \frac{1}{2}u(1-x^*(y, \alpha'))} \right)^{1-\alpha'}$$

which means that  $\frac{dV}{dx}|_{x=x^*(y,\alpha'),\alpha=\alpha''} < 0$ . From the concavity of the objective function, we then know that  $x^*(y,\alpha'') < x^*(y,\alpha')$ . It follows that the optimum involves a greater degree of offset with  $\alpha''$  than with  $\alpha'$ .

In the case of lexicographic preferences, the decision maker's utility becomes

$$V(L) = \min \left\{ \frac{1}{2}u(y) + \frac{1}{2}u(x), \frac{1}{2}u(1-y) + \frac{1}{2}u(1-x) \right\}$$

Trivially,  $x^*(y) = 1 - y$  is then the best choice because it equates the two arguments; hence we obtain full offset. For the isoelastic specification  $w(z) = \frac{z^{1-\alpha}}{1-\alpha}$ , we obtain the lexicographic case in the limit as  $\alpha \rightarrow \infty$ .

**Resolute ex ante alia preferences** In the text, we make the following claim. Assume nature chooses the fixed allocation from a distribution that is symmetric with respect to the two potential recipients. Then individuals with resolute ex ante alia preferences, whether of the EU-distribution or probability-weighted variety, would have a strict preference for equal division of the funds they control in all split-the-prize tasks. The claim is trivial for probability-weighted preferences given the assumed separability over states of nature. Here we prove it for EU-distribution preferences.

Nature's choice of the fixed allocation is governed by a lottery  $(p_1, y_1; p_2, y_2; \dots; p_N, y_N)$ , where  $p_N$  is the probability that nature will allocate the fraction  $y_1$  of the prize to household  $A$  and  $1 - y_2$  to household  $B$ . (Extending the argument to a continuous distribution is straightforward.) Let  $x_n$  be the fraction the decision maker allocates to household  $A$  in state  $n$ .

The decision maker's utility is:

$$V(L) = w \left( \sum_{n=1}^N p_n \left( \frac{1}{2}u(y_n) + \frac{1}{2}u(x_n) \right) \right) + w \left( \sum_{n=1}^N p_n \left( \frac{1}{2}u(1-y_n) + \frac{1}{2}u(1-x_n) \right) \right)$$

In light of our symmetry assumption, we can rewrite this expression as

$$V(L) = w \left( K + \frac{1}{2} \sum_{n=1}^N p_n u(x_n) \right) + w \left( K + \frac{1}{2} \sum_{n=1}^N p_n u(1-x_n) \right)$$

where

$$K = \frac{1}{2} \sum_{n=1}^N p_n u(y_n)$$

Thus we have a collection of first-order conditions (for  $n = 1, \dots, N$ ):

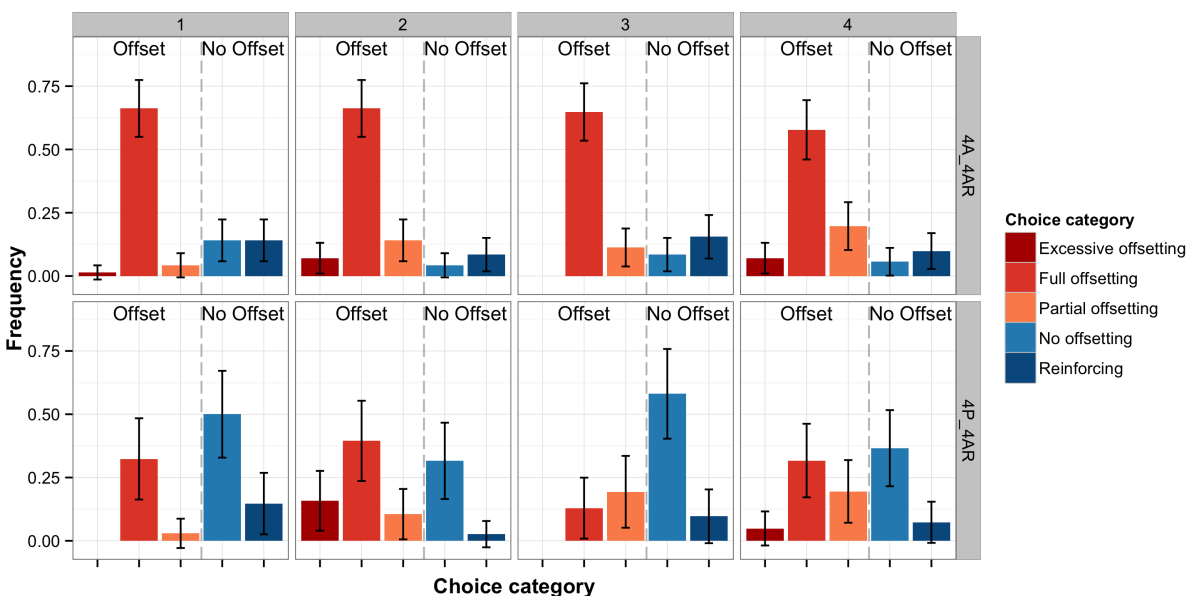
$$w' \left( K + \frac{1}{2} \sum_{n=1}^N p_n u(x_n) \right) \frac{p_n}{2} u'(x_n) = w' \left( K + \frac{1}{2} \sum_{n=1}^N p_n u(1-x_n) \right) \frac{p_n}{2} u'(1-x_n)$$

These conditions are satisfied when  $x_n = 0.5$  for all  $n$ . With concavity of  $w$  and  $u$ , this configuration is globally optimal.

### A.3 Additional data analysis

**Stability of choices across rounds** In general we found no evidence of systematic changes in behavior across rounds in which subjects encountered similar tasks. The figure below shows the distributions over choice categories for the first four rounds of treatments 4A\_4A<sup>R</sup> and 4P\_4A<sup>R</sup>.

**Figure A.4:** Choice category frequencies in each of the first four rounds of treatments 4A\_4A<sup>R</sup> and 4P\_4A<sup>R</sup>.

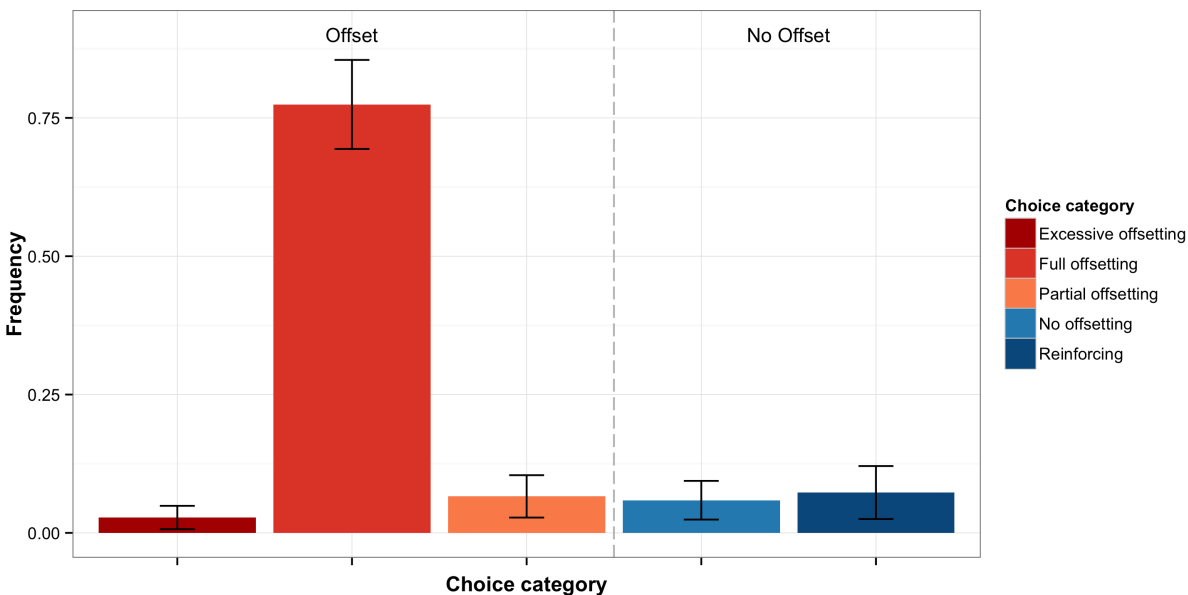


Notes: This figure is based on treatment Treatment 4A\_4A<sup>R</sup> (71 participants) and 4P\_4A<sup>R</sup> (72 participants). Error bands indicate 95% confidence intervals.

**The effect of extended exposure to ex post framing on ex ante choices** To evaluate the effect of extended exposure to ex post framing on ex ante choices, we examine choices made in the the 4P\_4A<sup>R</sup> treatment. Figure A.5 displays the unrevised choice distributions for rounds 5-8, during which subjects perform tasks with ex ante framing after experiencing four rounds with ex post framing. (Recall that Figure 1, panel B, exhibits the round 1-4 choice distribution for this treatment.) As in Figure 1, panel A, choices are predominantly ex ante fair. We reject equality of the round 1-4 and round 5-8 distributions ( $p < 0.001$ ), which tells us that subjects do not simply adhere to their initial perspective once the decision frame changes, even after four rounds of reinforcement. We also fail to reject equality of the round 5-8 distributions for the 4P\_4A<sup>R</sup> and 4A\_4A<sup>R</sup> treatments ( $p = 0.32$ ). This is precisely opposite what one would expect if initial perspectives on fairness were persistent. Thus, we find no support for the persistence hypothesis.



**Figure A.5:** Distributions of choices for tasks with ex ante framing after extended exposure to ex post framing



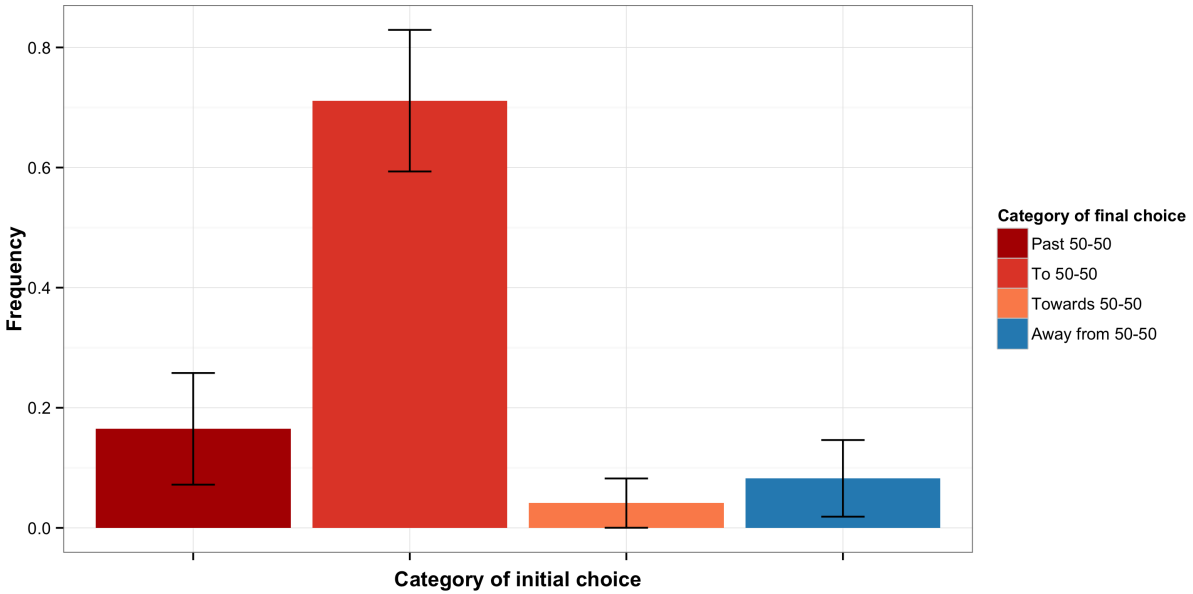
Notes: This figure is based on rounds 5-8 of treatment 4P\_4A<sup>R</sup> (72 subjects). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

The tendency for people to make ex ante fair choices even after being exposed to the ex post perspective is even more evident if one restricts attention to consistent decision makers. Two-thirds of subjects participating in the 4P\_4A<sup>R</sup> treatment displayed consistency in rounds 5-8, in the sense that they made the same type of choice in every round. We cannot reject the hypothesis that this fraction is the same as for rounds 5-8 of treatment 4A\_4A<sup>R</sup> ( $p = 0.45$ ). Of the consistent choosers, all but two chose the ex ante fair alternative in every round. Analyses of consistent choosers in rounds 5-8 of treatments 2A2P\_4A<sup>R</sup> and 2P2A\_4A<sup>R</sup> yield similar conclusions.

**The distribution of revision types for split-the-tickets tasks** Focusing just on decisions that were revised, we can usefully classify them according to whether the subject switched to a 50-50 division of his or her own tickets (ex post fairness), moved part of the way toward 50-50, moved past 50-50, or moved away from 50-50. Figure A.6 shows the distribution of revisions across these categories in the last four rounds of treatment 4A\_4A<sup>R</sup>.

The panels of Figure A.7 are analogous to Figure A.6, except they pertain to treatments 2A2P\_4A<sup>R</sup>, 2P2A\_4A<sup>R</sup>, and 4P\_4A<sup>R</sup>. All are qualitatively similar, in that revisions predominantly lead to ex post fair allocations. Notably, we cannot reject the hypotheses that each of these distributions is the same as for 4A\_4A<sup>R</sup> ( $p = 0.15, 0.21, \text{ and } 0.58$ , respectively).

**Figure A.6:** Distribution of revision types during the final four rounds of treatment 4A\_4A<sup>R</sup>

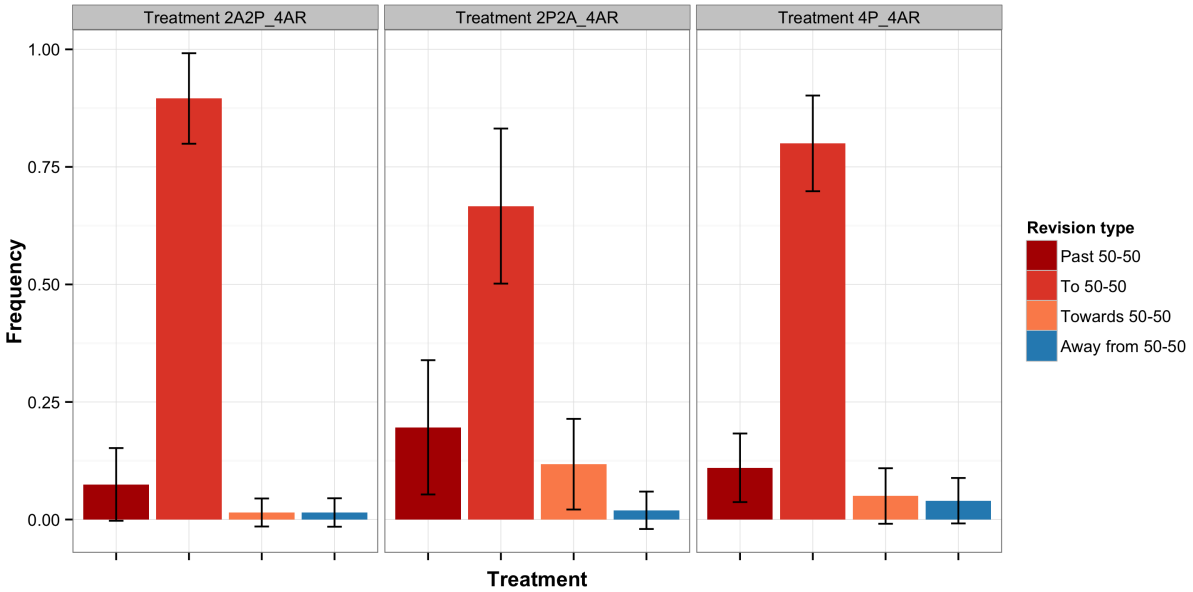


Notes: This figure is based on the final four rounds of treatment 4A\_4A<sup>R</sup>, in which there were a total of 97 revisions. Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

**Marginal distributions of final (revised) choices for various treatments** Figure A.8 exhibits the marginal distributions of final (revised) choices during the last four rounds of treatments 2A2P\_4A<sup>R</sup>, 2P2A\_4A<sup>R</sup>, or 4P\_4A<sup>R</sup>. The panels of these Figures are analogous to panel B of Figure 6, which pertains to treatment 4A\_4A<sup>R</sup>. We see that final allocations are predominantly ex post fair in all three treatments. Moreover, we cannot reject the hypotheses that the distributions for any of these treatments, 2A2P\_4A<sup>R</sup>, 2P2A\_4A<sup>R</sup>, or 4P\_4A<sup>R</sup>, are the same as for 4A\_4A<sup>R</sup> ( $p = 0.50$ , 0.15, and 0.46, respectively). Recall from Figures 3 and A.5 that the original (unrevised) choices for these same rounds were predominantly ex ante fair. Thus we see striking time-inconsistent shifts from the ex ante to the ex post perspective in all of these settings, just as in treatment 4A\_4A<sup>R</sup>.

**Comparisons of average choice patterns with “fingerprints” for ex ante and ex post fairness, various treatments** Figure A.9 compares the patterns of average allocations across rounds 5-8 (both before and after revisions) for treatments 2A2P\_4A<sup>R</sup>, 2P2A\_4A<sup>R</sup>, and 4P\_4A<sup>R</sup> with the “fingerprints” associated with ex ante and ex post fairness. The panels of this figures are analogous to Figure 7, which pertain to treatment 4A\_4A<sup>R</sup>. We see qualitatively similar patterns: the initial choices track the ex ante fair fingerprint fairly closely (as reported in Section 5), while the lines for the final (revised) allocations are flatter, more closely resembling the ex post fingerprint.

**Figure A.7:** Distribution of revision types during the final four rounds of various treatments



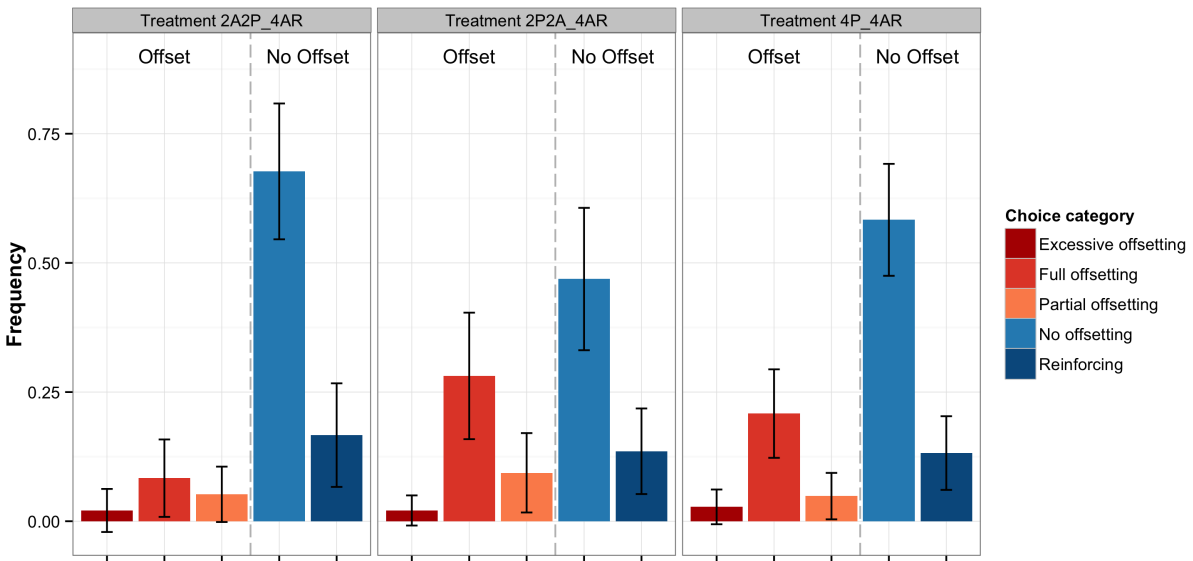
Notes: This figure is based on the final four rounds of the indicated treatments, in which there were (respectively) 67, 51, and 100 revisions. Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

As in the main text, we quantify the similarity to the benchmarks by estimating simple regressions of the chosen split on a constant and the computer’s split, clustering observations at the subject level. Focusing on final choices, the coefficient of the computer’s split is 0.02 (s.e. = 0.07) for treatment  $2A2P\_4A^R$ , -0.27 (s.e. = 0.07) for treatment  $2P2A\_4A^R$ , and -0.17 (s.e. = 0.06) for treatment  $4P\_4A^R$ . All of these coefficients are much further from the ex ante benchmark (-1) and closer to the ex post benchmark (0) than the corresponding coefficients for the original choices, which we reported in Section 5.

**Commitment choices by round** As seen in Figure A.10, the frequencies with which subjects express preferences for commitment or flexibility in rounds 5-8 of treatment  $4A^R\_4A^C$  do not vary systematically across rounds.

**Commitment choices for consistent and inconsistent subjects** As shown in Figure A.11, consistent subjects exhibited stronger preferences for commitment in rounds 5-8 of treatment  $4A\_4A^C$  than inconsistent ones.

**Figure A.8:** Marginal distributions of final (revised) choices during the final four rounds of various treatments

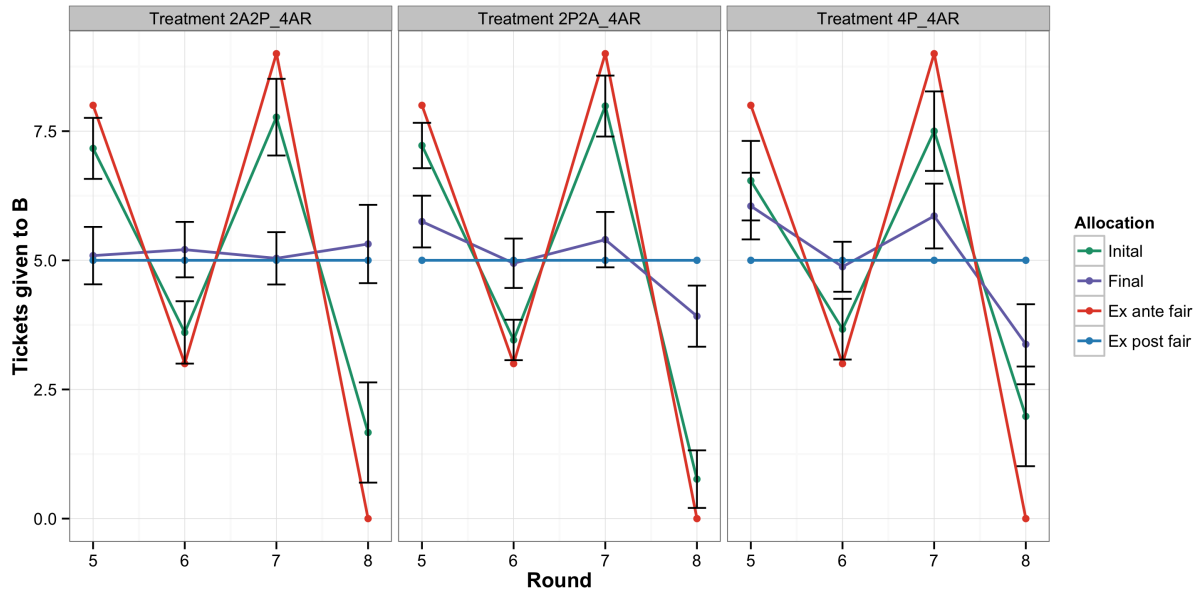


Notes: This figure is based on the final four rounds of the indicated treatments, in which there were 48, 48, and 72 subjects, respectively. Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

**Commitment choices by category of initial allocation** Figure A.12 divides the allocation tasks performed in rounds 5-8 of treatment  $4A^R_4A^C$  into five categories based on the type of the subject's original selection, and plots the distribution of commitment choices for each. As noted in the text, the propensity to commit is lower relative to the propensity to retain flexibility when subjects select initial allocations that are more vulnerable to revision.

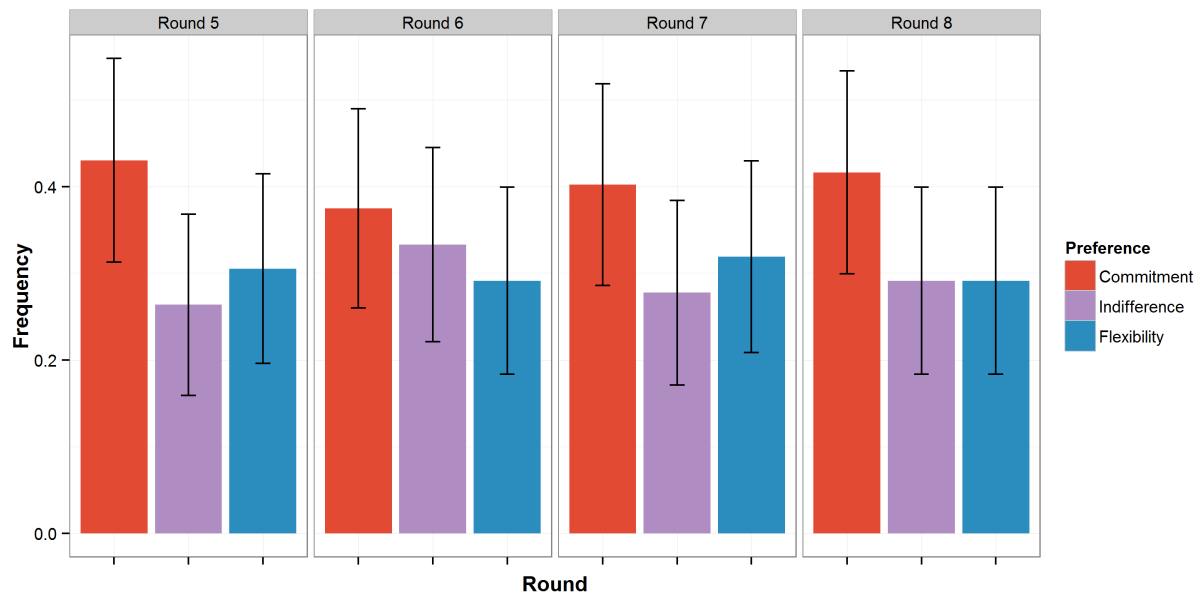
**The distribution of revision types for split-the-prize tasks** Figure A.13 focuses on the decisions that were revised in the last four rounds of the split-the-prize session, and groups them into the same four categories used for this purpose in section 6; it is analogous to figure A.6. Notice that the vast majority of those who revised (73.1%) migrated to ex post fair allocations, just as with divide-the-tickets tasks.

**Figure A.9:** Original and final allocations in rounds 5-8 of the indicated treatments.



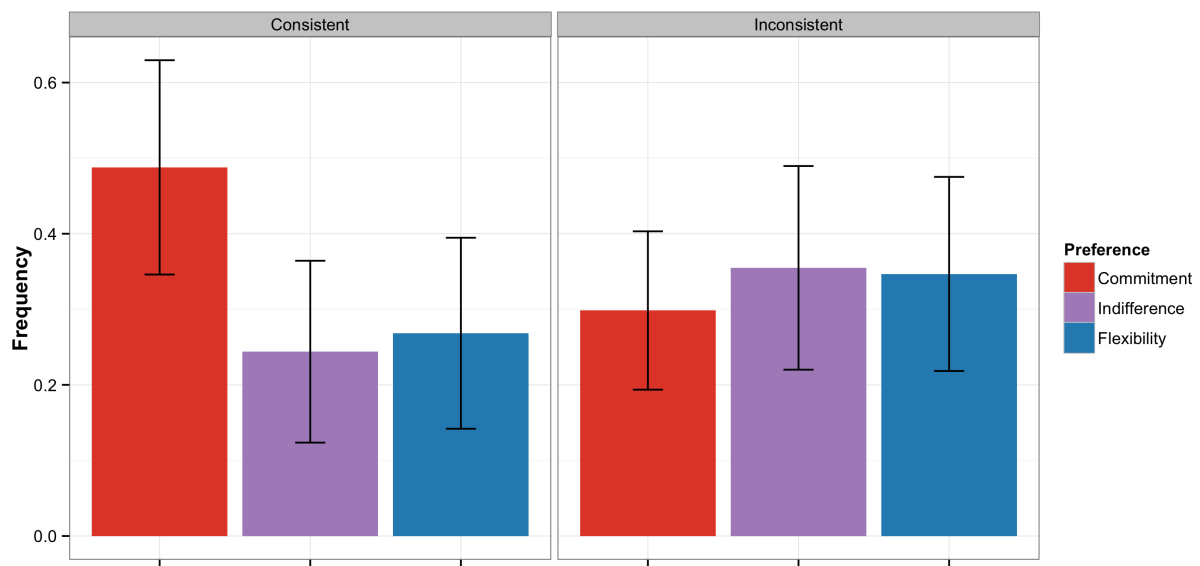
Notes: This figure is based on the final four rounds of the indicated treatments, in which there were 48, 48, and 72 subjects, respectively. Error bands indicate 95% confidence intervals.

**Figure A.10:** Commitment choices



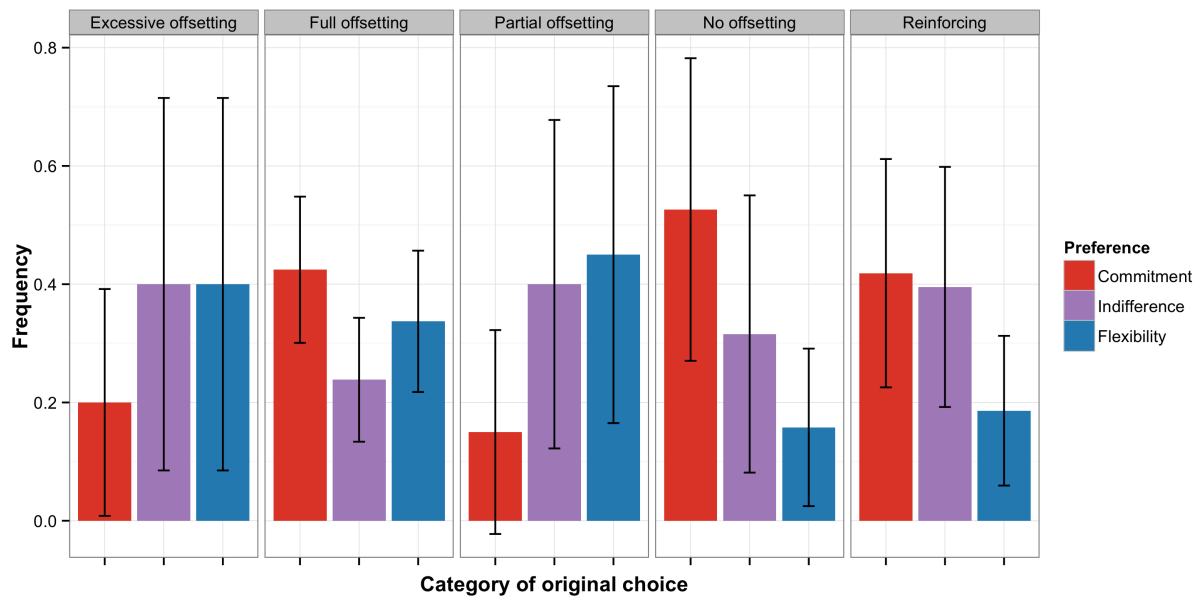
Note: This figure is based on the final four rounds of treatment  $4A^R_4A^C$  (72 subjects). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

**Figure A.11:** Commitment choices for consistent and inconsistent subjects



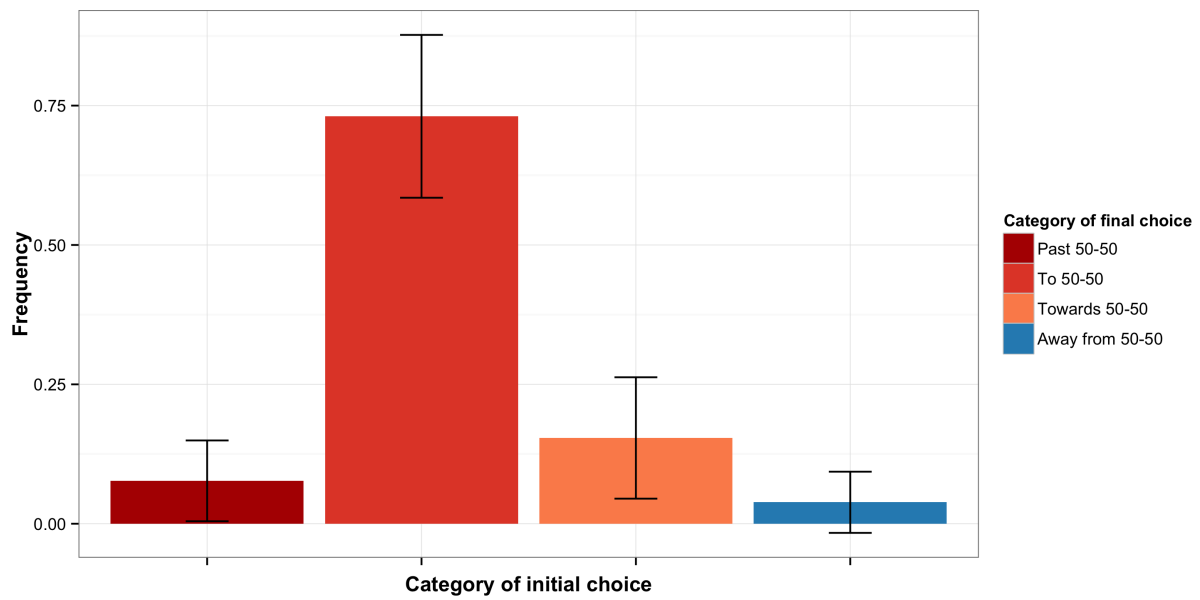
Note: This figure is based on the final four rounds of treatment  $4A^R_4A^C$  (72 subjects, 36 of whom were consistent, and 36 of whom were inconsistent). Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

**Figure A.12:** Commitment choices by category of original choice



Note: This figure is based on the final four rounds of treatment  $4A^R_4A^C$  (72 subjects). The distributions are based on 172 ex ante fair choices, 20 partially offsetting choices, 15 excessively offsetting choices, 38 ex post fair choices, and 43 reinforcing choices. Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.

**Figure A.13:** Distribution of revision types during the final four rounds of treatment 4A\_4A<sup>R</sup> with divide-the-prize tasks



Notes: This figure is based on the final four rounds of treatment 4A\_4AR (dollars), in which there were a total of 52 revisions. Error bands indicate 95% confidence intervals, calculated to allow for within-subject correlation.